

北航学报  
赠阅

ISSN 1001-5965  
CODEN BHHDE8



# 北京航空航天大学 学报

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS



**2015-10**

Vol.41 No.10

## 目 次

## 电磁理论与应用

- 随机粗糙面的双站散射模型分析 ..... 陈银山, 李召良, 刘玉 (1765)
- 带状注分布作用速调管多间隙谐振腔高频特性 ..... 阮存军, 戴军, 陈姝媛, 李仁杰, 占敏 (1777)
- 周期结构电磁特性在高频真空器件中的应用 ..... 冯进军, 蔡军, 胡银富, 邬显平 (1785)
- 接收模式下加罩天线系统特性的快速电磁仿真 .....  
..... 何芒, 刘金波, 王斌斌, 徐晓文, 章传芳, 孙厚军 (1792)
- 双反射面信道中 6 极化 MIMO 信道特性 ..... 朴大志, 陈鹏 (1801)
- 基于有限元方法的场磨式电场传感器标定装置优化设计 ..... 崔勇, 袁海文, 赵录兴, 胡多 (1807)
- 热效应布局下的缓冲器插入时序优化方法 ..... 王新胜, 韩良, 喻明艳 (1813)
- 电磁兼容测量天线系数温度误差修正方法 ..... 何洋, 苏东林, 李艳, 赵子华, 刘洪颐 (1821)
- 改进的基于本征滤波的时域宽带波束形成 ..... 王剑书, 樊养余, 杜瑞, 吕国云 (1830)
- 谐振系统集成光学陀螺中电流调谐致强度噪声研究 ..... 邝银周, 冯丽爽, 王俊杰, 唐依创 (1836)
- 协同体制被动毫米波成像系统天线阵布局优化 ..... 尚晓舟, 孙鹏, 胡岸勇, 苗俊刚 (1842)
- 电磁多尺度和非线性效应分析 ..... 毛煜茹, 谢拥军 (1848)
- 有源可调微波吸收体分析与优化 ..... 戚开南, 汪勇峰, 侯新宇, 陈军文 (1853)
- 受限空间中天线产生的电波覆盖研究 ..... 李大伟, 王均宏, 何显时 (1859)
- 等效原理算法在不同等效面下的误差分析 ..... 党训旺, 李懋坤, 杨帆, 许慎恒 (1867)
- 飞行器雷达隐身测试评估技术及发展 ..... 肖志河, 高超, 白杨, 袁晓峰 (1873)
- 毫米波-太赫兹电子回旋器件的几个基础问题 ..... 杜朝海, 罗里, 刘濮鲲 (1880)
- W 波段多注交错双栅行波放大器的高频特性 ..... 张慕武, 阮存军 (1887)

## 电路与系统

- 航空电子设备 NSEE 试验评价方法 ..... 薛海红, 王群勇, 陈冬梅, 陈宇, 阳辉, 李红军 (1894)
- 融合粗糙集与 D-S 证据理论的航空装备故障诊断 ..... 孙伟超, 李文海, 李文峰 (1902)

## 算法与应用

- 基于子目标进化的高维多目标优化算法 ..... 雷宇曜, 姜文志, 刘立佳, 马向玲 (1910)
- 基于带宽优化的载波跟踪算法 ..... 牟伟清, 刘荣科, 宋青平, 段瑞枫 (1918)
- 汉语双模情感语音数据库标注及一致性检测 ..... 景少玲, 毛峡, 陈立江, 张娜娜 (1925)
- 混合点状和非规则军标的在线手绘识别 ..... 邓维, 吴玲达, 张友根, 赵志鹏 (1935)
- 基于块平均预处理的 GNSS-R 延迟映射处理方法 ..... 王峰, 杨东凯, 李伟强, 张彦仲 (1943)
- 基于均方根容积粒子的 SMC-PHD 算法 ..... 刘哲, 王祖林, 徐迈, 刘景贤, 杨蓝 (1950)

## 网络体系与应用

- 基于可编程控制网关的一体化网络体系结构 ..... 刘锋, 王渝斐 (1959)
- 构建广义立方体感知网络安全态势 ..... 文志诚, 陈志刚 (1966)

期刊基本参数: CN11-2625/V \* 1956 \* m \* A4 \* 210 \* zh \* P \* ¥20.00 \* 900 \* 28 \* 2015-10

CONTENTS

**Electromagnetic Theory and Application**

- Model analysis of bistatic scattering from randomly rough surfaces  
..... CHEN Kunshan, LI Zhaoliang, LIU Yu (1765)
- RF characteristics of multiple-gap resonant cavity for sheet beam extended interaction klystron  
..... RUAN Cunjun, DAI Jun, CHEN Shuyuan, LI Renjie, ZHAN Min (1777)
- Application of dispersion characteristics of periodic structures for high frequency vacuum devices  
..... FENG Jinjun, CAI Jun, HU Yinfu, WU Xianping (1785)
- Fast electromagnetic simulation of characteristics of radome-antenna system in receiving mode  
..... HE Mang, LIU Jinbo, WANG Binbin, XU Xiaowen, ZHANG Chuanfang, SUN Houjun (1792)
- Characteristics of hexapolarized MIMO channel in two-mirror channel  
..... PIAO Dazhi, CHEN Peng (1801)
- Optimum design of calibration device for field mill type electric field sensor based on finite element method  
..... CUI Yong, YUAN Haiwen, ZHAO Luxing, HU Duo (1807)
- Thermal aware floor planning timing optimal method for buffer insertion  
..... WANG Xinsheng, HAN Liang, YU Mingyan (1813)
- Method for EMC antenna factor temperature error correction  
..... HE Yang, SU Donglin, LI Yan, ZHAO Zihua, LIU Hongyi (1821)
- Improved eigenfilter-based time-domain wideband beamforming  
..... WANG Jianshu, FAN Yangyu, DU Rui, LYU Guoyun (1830)
- Study on current-tuning-induced intensity noise in resonator integrated optic gyro  
..... ZHI Yinzhou, FENG Lishuang, WANG Junjie, TANG Yichuang (1836)
- Antenna array layout optimization of collaborative passive millimeter-wave imaging system  
..... SHANG Xiaozhou, SUN Peng, HU Anyong, MIAO Jungang (1842)
- Analysis of electromagnetic multi-scale structure and non-linear effects  
..... MAO Yuru, XIE Yongjun (1848)
- Analysis and optimization of active tunable microwave absorber  
..... QI Kainan, WANG Yongfeng, HOU Xinyu, CHEN Junwen (1853)
- Research on radio coverage generated by antennas in confined space  
..... LI Dawei, WANG Junhong, HE Xianshi (1859)
- Error analysis of equivalence principle algorithm on different equivalence surfaces  
..... DANG Xunwang, LI Maokun, YANG Fan, XU Shenheng (1867)
- Aircraft radar stealth test and evaluation technology and progress  
..... XIAO Zhihe, GAO Chao, BAI Yang, YUAN Xiaofeng (1873)
- Several fundamental challenges of millimeter-to-terahertz electron cyclotron devices  
..... DU Chaohai, LUO Li, LIU Pukun (1880)
- High frequency characteristics for W-band multiple beam staggered double-vane traveling wave tube amplifier  
..... ZHANG Muwu, RUAN Cunjun (1887)

**Circuit and System**

- Neutron single event effects testing and evaluation method for avionics  
..... XUE Haihong, WANG Qunyong, CHEN Dongmei, CHEN Yu, YANG Hui, LI Hongjun (1894)
- Avionic devices fault diagnosis based on fusion method of rough set and D-S theory  
..... SUN Weichao, LI Wenhai, LI Wenfeng (1902)

**Algorithm and Application**

- Many-objective optimization based on sub-objective evolutionary algorithm  
..... LEI Yuyao, JIANG Wenzhi, LIU Lijia, MA Xiangling (1910)
- Carrier tracking algorithm based on bandwidth optimization  
..... MU Weiqing, LIU Rongke, SONG Qingping, DUAN Ruifeng (1918)
- Annotations and consistency detection for Chinese dual-mode emotional speech database  
..... JING Shaoling, MAO Xia, CHEN Lijiang, ZHANG Nana (1925)
- Online sketch recognition for mixed point and irregular military symbols  
..... DENG Wei, WU Lingda, ZHANG Yougen, ZHAO Zhipeng (1935)
- GNSS-R delay mapping processing method based on block averaging pre-processing  
..... WANG Feng, YANG Dongkai, LI Weiqiang, ZHANG Yanzhong (1943)
- SMC-PHD algorithm based on squared cubature particles  
..... LIU Zhe, WANG Zulin, XU Mai, LIU Jingxian, YANG Lan (1950)

**Network System and Application**

- Architecture of unified network based on programmable control gateway  
..... LIU Feng, WANG Yufei (1959)
- Constructing general cube to be aware of network security situation  
..... WEN Zhicheng, CHEN Zhigang (1966)

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0205

# 随机粗糙面的双站散射模型分析

陈崑山<sup>\*1,2</sup>, 李召良<sup>2</sup>, 刘玉<sup>1</sup>

(1. 中国科学院遥感与数字地球研究所 遥感科学国家重点实验室, 北京 100101;

2. 中国农业科学院农业资源与区域规划研究所 农业部农业信息技术重点实验室, 北京 100081)

**摘 要:** 以先进积分方程模型(AIEM)为工作模型,从理论上系统分析了随机粗糙面的双站散射特征,探讨了不同粗糙程度下散射系数在方位平面上的谷点位置,并利用数值仿真模拟、实测值对理论预测的散射系数进行了说明与验证.结果显示,散射系数在方位向上的谷点位置与粗糙程度和极化特征密切相关.利用数值仿真模拟,可从物理的角度更系统深入地分析统计粗糙表面的散射特性,从而有利于设计以后的更有效的双站探测模式.

**关键词:** 双站散射; 散射系数; 随机粗糙面; 积分方程模型(IEM); 先进积分方程模型(AIEM); 数值仿真模拟

**中图分类号:** O571.41<sup>+8</sup>

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1765-12

## 1 Introduction

Electromagnetic wave scattering from a randomly rough surface requires complex mathematical treatment, but is of palpable importance in many fields of disciplines and bears itself in various applications spanned from surface treatment to remote sensing of terrain and sea<sup>[1-6]</sup>. For example, it has been a common practice to retrieve, by analyzing the sensitivity of the scattering behavior and mechanisms, the geophysical parameters of interest from the scattering and/or emission measurements. Another example is that by knowing the scattering patterns, one may be able to detect the presence of the undesired random roughness of a reflective surface such as antenna reflector, and thus accordingly devise a means to correct or compensate the

phase errors. Therefore, it has been both theoretically and practically motivated to study the electromagnetic wave scattering from the randomly rough surfaces. Research and progress, being a long historical track, of this topic have been well documented and still been keeping updates.

In order to tackle the complex and sometimes intricate mathematical derivations, while to remain a high level of accuracy beyond conventional models as well, notably, Kirchhoff and small perturbation method (SPM), the integral equation model (IEM) has been developed by Fung, et al.<sup>[3, 6-8]</sup> under few physical-justified assumptions. Driven by the need of predicting bistatic scattering and microwave emissivity, considerable efforts have been devoted to further improving the IEM accuracy<sup>[9-19]</sup> by removing some of the assumptions that imposed for the

收稿日期: 2015-04-09; 录用日期: 2015-05-05; 网络出版时间: 2015-07-02 10:25

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150702.1025.001.html

基金项目: 中国科学院遥感与数字地球研究所所长基金

\* 通讯作者: 陈崑山(1959—),男,中国台湾人,教授,chenks@radi.ac.cn,主要研究方向为微波遥感.

引用格式: 陈崑山, 李召良, 刘玉. 随机粗糙面的双站散射模型分析[J]. 北京航空航天大学学报, 2015, 41(10): 1765-1776.

Chen K S, Li Z L, Liu Y. Model analysis of bistatic scattering from randomly rough surfaces[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1765-1776 (in Chinese).

purpose of mathematical simplicity during the course of derivation. Another step leaping forward was the introduction of a transition function into the Fresnel reflection coefficients to take spatial dependences into account, removing the restrictions on the limits of surface roughness and permittivity<sup>[3,11]</sup>. Though the approach is heuristic but self-consistent, it proves working perfectly for a broad range of surface dielectric and geometric parameters<sup>[6,11,14]</sup>.

## 2 Formulation of wave scattering from a rough surface

Referring to Fig. 1 with the direction of incidence  $(\theta_i, \phi_i)$  and scattering  $(\theta_s, \phi_s)$ , suppose that a plane wave impinges onto a dielectric rough surface that scatters waves up into the incident plane and down into the lower medium, with the electric field  $\mathbf{E}_i$  and magnetic field  $\mathbf{H}_i$  being written as

$$\mathbf{E}_i = \hat{\mathbf{p}} E_0 \exp[-j(\mathbf{k}_i \cdot \mathbf{r})] \quad (1)$$

$$\mathbf{H}_i = \frac{1}{\eta} \hat{\mathbf{k}}_i \times \mathbf{E}_i \quad (2)$$

where  $j = \sqrt{-1}$ ;  $\hat{\mathbf{p}}$  denotes the unit polarization vector,  $E_0$  the amplitude of the incident electric field  $\mathbf{E}_i$ ,  $\eta$  and  $\eta_i$  denote the intrinsic impedance of the upper and lower transmitted media, respectively. The position vector  $\mathbf{r}$ ,  $k$ -plane in incident direction  $\mathbf{k}_i$  and scattering direction  $\mathbf{k}_s$  are defined as follows, respectively

$$\left\{ \begin{aligned} \mathbf{r} &= x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} \\ \mathbf{k}_i &= k\hat{\mathbf{k}}_i = \hat{\mathbf{x}}k_{ix} + \hat{\mathbf{y}}k_{iy} + \hat{\mathbf{z}}k_{iz} \\ k_{ix} &= k \sin \theta_i \cos \phi_i \\ k_{iy} &= k \sin \theta_i \sin \phi_i, \\ k_{iz} &= k \cos \theta_i \\ \mathbf{k}_s &= k\hat{\mathbf{k}}_s = \hat{\mathbf{x}}k_{sx} + \hat{\mathbf{y}}k_{sy} + \hat{\mathbf{z}}k_{sz} \\ k_{sx} &= k \sin \theta_s \cos \phi_s \\ k_{sy} &= k \sin \theta_s \sin \phi_s \\ k_{sz} &= k \cos \theta_s \end{aligned} \right. \quad (3)$$

where  $x, y$  and  $z$  denote a coordinate system,  $\hat{\mathbf{x}}, \hat{\mathbf{y}}$  and  $\hat{\mathbf{z}}$  unit vector of  $x, y$  and  $z$ , respectively,  $\hat{\mathbf{k}}_i$  and  $\hat{\mathbf{k}}_s$  incident and scattering directions, respectively,  $k_{ix}, k_{iy}$  and  $k_{iz}$   $x, y$  and  $z$  components of the incident wave number, respectively,  $k_{sx}, k_{sy}$  and  $k_{sz}$   $x, y$  and  $z$  components of the scattering wave number, respectively and  $k$  denotes wave number.

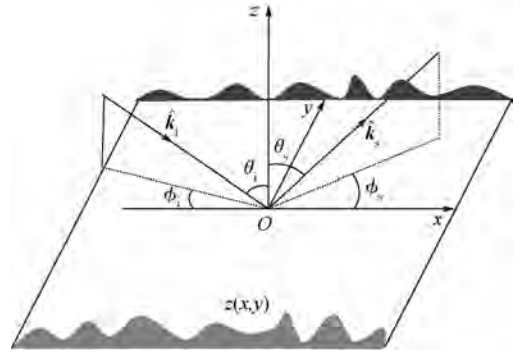


Fig. 1 Geometry of bistatic scattering.

The polarization vector  $\hat{\mathbf{p}}$  is defined as  $\hat{\mathbf{h}}_i$  and  $\hat{\mathbf{v}}_i$  for incident linearly horizontal-polarized wave and vertical-polarized waves, respectively; and as  $\hat{\mathbf{h}}_s$  and  $\hat{\mathbf{v}}_s$  for scattering linearly horizontal-polarized wave and vertical-polarized waves, respectively.

$$\left\{ \begin{aligned} \hat{\mathbf{h}}_i &= -\hat{\mathbf{x}} \sin \phi_i + \hat{\mathbf{y}} \cos \phi_i \\ \mathbf{v}_i &= \hat{\mathbf{h}}_i \times \hat{\mathbf{k}}_i = \\ &\quad -(\hat{\mathbf{x}} \cos \theta_i \cos \phi_i + \hat{\mathbf{y}} \cos \theta_i \sin \phi_i + \hat{\mathbf{z}} \sin \theta_i) \\ \hat{\mathbf{h}}_s &= \hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}} \sin \phi_s + \hat{\mathbf{y}} \cos \phi_s \\ \mathbf{v}_s &= \hat{\boldsymbol{\theta}} = \hat{\mathbf{h}}_s \times \hat{\mathbf{k}}_s = \\ &\quad \hat{\mathbf{x}} \cos \theta_s \cos \phi_s + \hat{\mathbf{y}} \cos \theta_s \sin \phi_s - \hat{\mathbf{z}} \sin \theta_s \end{aligned} \right. \quad (4)$$

where  $\hat{\boldsymbol{\phi}}$  and  $\hat{\boldsymbol{\theta}}$  denote azimuthal and polar unit vectors, respectively.

The total electric field  $\mathbf{E}(\mathbf{r})$  and magnetic field  $\mathbf{H}(\mathbf{r})$  according to the Stratton-Chu formula could be expressed as<sup>[2,5]</sup>

$$\mathbf{E}(\mathbf{r}) = \iint_{S'} dS' [ i\omega\mu(\hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}')) G(\mathbf{r}, \mathbf{r}') + (\hat{\mathbf{n}} \cdot \mathbf{E}(\mathbf{r}')) \nabla' G(\mathbf{r}, \mathbf{r}') + (\hat{\mathbf{n}} \times \mathbf{E}(\mathbf{r}')) \times \nabla' G(\mathbf{r}, \mathbf{r}') ] \quad (5a)$$

$$\mathbf{H}(\mathbf{r}) = \iint_{S'} dS' [ -i\omega\varepsilon(\hat{\mathbf{n}} \times \mathbf{E}(\mathbf{r}')) G(\mathbf{r}, \mathbf{r}') + (\hat{\mathbf{n}} \cdot \mathbf{H}(\mathbf{r}')) \nabla' G(\mathbf{r}, \mathbf{r}') + (\hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}')) \times \nabla' G(\mathbf{r}, \mathbf{r}') ] \quad (5b)$$

where  $S'$  denotes source surface,  $i$  imaginary unit;  $\omega$  angular frequency,  $\mu$  permeability;  $\hat{\mathbf{n}}$  normal vector,  $\mathbf{H}(\mathbf{r}')$  magnetic field,  $G(\mathbf{r}, \mathbf{r}')$  Green function and  $\mathbf{r}'$  source point,  $\varepsilon$  permittivity. The total field is the sum of the incident field, which is known, and the scattered field, which is unknown and is to be solved. mathematically, the scattered field, for a source free region as in our case, is expressed as<sup>[20]</sup>

$$\mathbf{E}(\mathbf{r}) = \gamma \mathbf{E}_i(\mathbf{r}) - \frac{\gamma}{4\pi} \cdot$$

$$\iint dS' [i\omega\mu(\hat{\mathbf{n}} \times \mathbf{H})G - (\hat{\mathbf{n}} \times \mathbf{E}) \times \nabla'G - (\hat{\mathbf{n}} \cdot \mathbf{E}) \nabla'G] \quad (6a)$$

$$\mathbf{H}(\mathbf{r}) = \gamma \mathbf{H}_i(\mathbf{r}) + \frac{\gamma}{4\pi} \cdot$$

$$\iint dS' [i\omega\varepsilon(\hat{\mathbf{n}} \times \mathbf{E})G + (\hat{\mathbf{n}} \times \mathbf{H}) \times \nabla'G + (\hat{\mathbf{n}} \cdot \mathbf{H}) \nabla'G] \quad (6b)$$

$$\text{where } \gamma = (1 - \Omega/4\pi)^{-1}, \Omega = \begin{cases} 0 & \mathbf{r} \notin S \\ 2\pi & \mathbf{r} \in S \end{cases}$$

Once the scattered field is solved, the scattering coefficient with  $q$  polarization is calculated as<sup>[2]</sup>

$$\sigma_{qp}^0 = \frac{4\pi R^2 \text{Re} \{ \langle |\mathbf{E}_{qp}^s|^2 \rangle \}}{A_0 \cos \theta_i \text{Re} \{ |\mathbf{E}_i|^2 \}} \quad (7)$$

where  $R$  denotes the range from the surface to radar observation point,  $\sigma_{qp}^0$  scattering coefficient,  $\text{Re}\{\cdot\}$  real part,  $\mathbf{E}_{qp}^s$  the scattered field of  $q$  polarization with the incident field of  $p$  polarization and  $A_0$  the illuminated overlapping area, confined by transmitting and receiving antenna beam patterns, over the rough surface.

## 2.1 Surface tangential fields

In view of Eq. (5), Ref. [5] states the Huygens' principle. The field solution in a given volume  $V'$  is completely determined by the tangential fields specified over the surface  $S'$  enclosing  $V'$ . To find the surface fields, one has to solve the pair of Fredholm integral equations of 2nd kind. In our case, exact analytic solution is not possible. Approximate surface fields are estimated by iterative approach. The rough surface with irregular boundary is presented in Fig. 2, where  $\mathcal{F}^\pm$ ,  $\mathcal{G}^\pm$  denote the upward re-radiation and downward re-radiation, going through upper medium and lower medium, respectively,  $\theta_t$  denotes transmission angle,  $\eta_0$  wave impedance in air,  $\varepsilon_0$  permittivity in air,  $\eta_1$  wave impedance in lower medium,  $\varepsilon_1$  permittivity in lower medium,  $\hat{\mathbf{v}}$  vertical unit vector and  $\hat{\mathbf{d}}$  tangential unit vector normal to surface contour.

The completely analytic solution of rough surface is almost prohibitive. Instead, we seek the approximate estimate of the surface tangential fields by taking vector product with the unit surface normal on both sides of Eq. (6a), Eq. (6b) and, after reformulations<sup>[21]</sup>, allowing an iterative scheme to find the estimates.

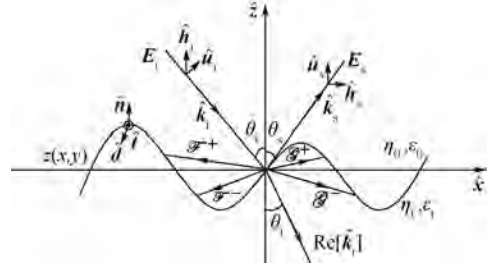


Fig. 2 Geometry of wave scattering and propagation from a rough surface

In IEM modelling<sup>[3,7]</sup>, the tangential surface fields is the sum of the Kirchhoff field and the complementary field, which are

$$(\hat{\mathbf{n}} \times \mathbf{E}_p) = (\hat{\mathbf{n}} \times \mathbf{E}_p)_k + (\hat{\mathbf{n}} \times \mathbf{E}_p)_c \quad (8a)$$

$$(\hat{\mathbf{n}} \times \mathbf{H}_p) = (\hat{\mathbf{n}} \times \mathbf{H}_p)_k + (\hat{\mathbf{n}} \times \mathbf{H}_p)_c \quad (8b)$$

where the Kirchhoff fields can be expressed as

$$(\hat{\mathbf{n}} \times \mathbf{E}_p)_k = \hat{\mathbf{n}} \times [(1 - R_v)\hat{\mathbf{p}} + (R_v + R_h)(\hat{\mathbf{p}} \cdot \hat{\mathbf{t}})\hat{\mathbf{t}}] \mathbf{E}_i \quad (9a)$$

$$\eta_1 (\hat{\mathbf{n}} \times \mathbf{H}_p)_k = \hat{\mathbf{n}} \times \hat{\mathbf{k}}_i \times [(1 + R_v)\hat{\mathbf{p}} + (R_v + R_h)(\hat{\mathbf{p}} \cdot \hat{\mathbf{t}})\hat{\mathbf{t}}] \mathbf{E}_i \quad (9b)$$

where  $\mathbf{E}_p$  denotes electric field for  $p$  polarization,  $\mathbf{H}_p$  magnetic field for  $p$  polarization,  $\hat{\mathbf{p}}$  polarization unit vector,  $\hat{\mathbf{t}}$  tangential unit vector and is tangential to the edge of surface contour;  $\eta_1$  wave impedance in medium 1,  $R_h$  and  $R_v$  denote Fresnel reflection coefficients for horizontally and vertically polarized waves, respectively.

The complementary surface field, which corrects the Kirchhoff estimates, is given as

$$(\hat{\mathbf{n}} \times \mathbf{E}_v)_c = -\frac{1}{4\pi} [\hat{\mathbf{n}} \times \int (1 - R_v) \mathcal{F}_v ds' + \hat{\mathbf{n}} \times \int (1 + R_v) \mathcal{F}_v ds' - (R_v + R_h)(\hat{\mathbf{n}} \times \hat{\mathbf{t}})(\hat{\mathbf{n}} \times \hat{\mathbf{t}}) \cdot \hat{\mathbf{n}} \times \int (\mathcal{F}_v - \mathcal{F}_{v_t}) ds'] \quad (10a)$$

$$(\hat{\mathbf{n}} \times \mathbf{E}_h)_c = -\frac{1}{4\pi} [\hat{\mathbf{n}} \times \int (1 + R_h) \mathcal{F}_h ds' + \hat{\mathbf{n}} \times \int (1 - R_h) \mathcal{F}_h ds' + (R_v + R_h)\hat{\mathbf{t}} \cdot \hat{\mathbf{n}} \times \int (\mathcal{F}_h - \mathcal{F}_{h_t}) ds'] \quad (10b)$$

$$(\hat{\mathbf{n}} \times \mathbf{H}_v)_c = \frac{1}{4\pi} [\hat{\mathbf{n}} \times \int (1 + R_v) \mathcal{G}_v ds' + \hat{\mathbf{n}} \times \int (1 - R_v) \mathcal{G}_v ds' - (R_v + R_h)\hat{\mathbf{t}} \cdot \hat{\mathbf{n}} \times \int (\mathcal{G}_v - \mathcal{G}_{v_t}) ds'] \quad (11a)$$

$$\begin{aligned}
 (\hat{n} \times \mathbf{H}_h)_c &= \frac{1}{4\pi} [\hat{n} \times \int (1 - R_h) \mathcal{G}_h ds' + \hat{n} \times \\
 &\int (1 + R_h) \mathcal{G}_{ht} ds' + (R_v + R_h)(\hat{n} \times \hat{t})(\hat{n} \times \hat{t}) \cdot \\
 &\hat{n} \times \int (\mathcal{G}_h - \mathcal{G}_{ht}) ds'] \quad (11b)
 \end{aligned}$$

where  $\mathbf{E}_h$  and  $\mathbf{E}_v$  denote horizontal and vertical polarized electric fields, respectively,  $\mathbf{H}_h$  and  $\mathbf{H}_v$  horizontal and vertical polarized magnetic fields, respectively and  $s'$  denotes source surface.

The electric and magnetic fields that appear inside the integrals above,  $\mathcal{E}_v$ ,  $\mathcal{E}_{vt}$ ,  $\mathcal{E}_h$ ,  $\mathcal{E}_{ht}$ ,  $\mathcal{H}_v$ ,  $\mathcal{H}_{vt}$ ,  $\mathcal{H}_h$  and  $\mathcal{H}_{ht}$  are expressed as

$$\mathcal{E}_p = jk_i \eta (\hat{n} \times \mathbf{H}_p) G_i - (\hat{n} \times \mathbf{E}_p) \times \nabla' G_i - (\hat{n} \cdot \mathbf{E}_p) \nabla' G_i \quad (12a)$$

$$\mathcal{H}_p = j \frac{k_i}{\eta} (\hat{n} \times \mathbf{E}_p) G_i - (\hat{n} \times \mathbf{H}_p) \times \nabla' G_i - (\hat{n} \cdot \mathbf{H}_p) \nabla' G_i \quad (12b)$$

$$\mathcal{E}_{pt} = - \left[ jk_i \eta_t (\hat{n} \times \mathbf{H}_p) G_t - (\hat{n} \times \mathbf{E}_p) \times \nabla' G_t - \frac{1}{\varepsilon_r} (\hat{n} \cdot \mathbf{E}_p) \nabla' G_t \right] \quad (13a)$$

$$\mathcal{H}_{pt} = - \left[ j \frac{k_t}{\eta_t} (\hat{n} \times \mathbf{E}_p) G_t - (\hat{n} \times \mathbf{H}_p) \times \nabla' G_t - \frac{1}{\mu_r} (\hat{n} \cdot \mathbf{H}_p) \nabla' G_t \right] \quad (13b)$$

where  $p$  denotes h or v, h and v denote horizontal and vertical polarization, respectively,  $G_t$  and  $G_i$  denote Green functions in upper and lower media as shown in Eqs. (12) and (13), respectively,  $\mu_r$  and  $\varepsilon_r$  denote the relative permittivity and permeability, respectively.

In Eq. (12) and Eq. (13), there involve Greens functions and theirs gradients in the upper and lower medium. To seek solutions by iterative scheme we make use of the spectral form, instead of spatial form, of the Green's function [3, 7]

$$G_i = \frac{1}{2\pi} \iint \frac{j}{q_i} \exp(j\Phi) dudv \quad (14a)$$

$$\nabla' G_i = \frac{1}{2\pi} \iint \frac{\mathbf{g}_i}{q_i} \exp(j\Phi) dudv \quad (14b)$$

$$G_t = \frac{1}{2\pi} \iint \frac{j}{q_t} \exp(j\Phi_t) dudv \quad (15a)$$

$$\nabla' G_t = \frac{1}{2\pi} \iint \frac{\mathbf{g}_t}{q_t} \exp(j\Phi_t) dudv \quad (15b)$$

where  $u$  denotes coordinate values after change,  $v$  denotes coordinate values after change, the phase terms  $\Phi$  associated with the upper and lower medium

are, respectively,

$$\Phi_i = j [u(x - x') + v(y - y') - q_i |z - z'|] \quad (16a)$$

$$\Phi_t = j [u(x - x') + v(y - y') - q_t |z - z'|] \quad (16b)$$

where  $x'$ ,  $y'$  and  $z'$  denote coordinate values of source point, respectively,  $q_i = \sqrt{k_i^2 - u^2 - v^2}$ ,  $q_t = \sqrt{k_t^2 - u^2 - v^2}$ ;  $\mathbf{g}_i = \hat{x}u + \hat{y}v \mp \hat{z}q_i$ ,  $\mathbf{g}_t = \hat{x}u + \hat{y}v \mp \hat{z}q_t$ . By substituting the Kirchhoff surface fields in Eq. (9a), Eq. (9b) into Eq. (12) and Eq. (13), we obtain the estimates of the complementary fields of Eq. (10) and Eq. (11). This may be seen as a 2nd iteration of seeking the solution of the integral equations governing the surface fields using Kirchhoff fields as the initial guess which is indeed a very good choice for fast convergence.

### 2.2 Far-zone scattered field

Now, with the surface tangential field estimates available,  $\mathbf{E}_{qp}^s$  at far-zone distance  $R$  is readily calculated by making use of the Stratton-Chu formula.

$$\begin{aligned}
 \mathbf{E}_{qp}^s &= KE_0 \int [\hat{q} \times \hat{k}_s \cdot (\hat{n} \times \mathbf{H}_p) + \eta \hat{q} \cdot \\
 &(\hat{n} \times \mathbf{H}_p)] \exp[j(k\hat{k}_s \cdot \mathbf{r})] dS \quad (17)
 \end{aligned}$$

where  $\hat{q}$  denotes unit polarization vector,  $K = -\frac{jk_i}{4\pi R} \exp(-jk_i R)$ ,  $E_0$  amplitude of the incident electric field.

Corresponding to the Kirchhoff and the complementary surface fields in Eq. (8a), Eq. (8b), the far-zone scattered field may also be expressed as a sum of the Kirchhoff  $\mathbf{E}_{qp}^k$  and the complementary scattered fields  $\mathbf{E}_{qp}^c$  [3, 7]:

$$\mathbf{E}_{qp}^s = \mathbf{E}_{qp}^k + \mathbf{E}_{qp}^c \quad (18)$$

where the Kirchhoff field is given by

$$\mathbf{E}_{qp}^k = KE_0 \iint f_{qp} \exp(j\Phi) dx dy \quad (19)$$

with the phase term

$$\Phi = k [(\hat{k}_s - \hat{k}_i)x + (\hat{k}_s - \hat{k}_i)y + (\hat{k}_s - \hat{k}_i)z]$$

The complementary scattered field, propagating upward and downward, may be written as

$$\begin{aligned}
 \mathbf{E}_{qp}^c &= \frac{KE_0}{8\pi^2} \int (\mathcal{F}_{qp} e^{j(\Phi_i + \hat{k}_s \cdot \mathbf{r} - k_i \cdot \mathbf{r}')} + \mathcal{G}_{qp} e^{j(\Phi_t + \hat{k}_s \cdot \mathbf{r} - k_t \cdot \mathbf{r}')} ) \\
 &dudvdxdydx'dy' \quad (20)
 \end{aligned}$$

where  $\mathcal{F}_{qp}$  and  $\mathcal{G}_{qp}$  denote propagators. The Kirchhoff field coefficient  $f_{qp}$  (both  $q$  and  $p$  denote h or v)

appearing in Eq. (19) may be more explicitly written into the following form:

$$f_{vv} = - [(1 - R_v) \hat{\mathbf{h}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{v}}) + (1 + R_v) \hat{\mathbf{v}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{h}})]_{s_1} - (R_h + R_v) (\hat{\mathbf{v}} \cdot \hat{\mathbf{t}}) [(\hat{\mathbf{h}}_s \cdot \hat{\mathbf{d}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i) - (\hat{\mathbf{n}} \cdot \hat{\mathbf{d}}) (\hat{\mathbf{h}}_s \cdot \hat{\mathbf{k}}_i) - (\hat{\mathbf{v}}_s \cdot \hat{\mathbf{t}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i)]_{s_1} \quad (21a)$$

$$f_{vh} = [(1 - R_h) \hat{\mathbf{v}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{v}}) - (1 + R_h) \hat{\mathbf{h}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{h}})]_{s_1} - (R_h + R_v) (\hat{\mathbf{h}} \cdot \hat{\mathbf{d}}) [(\hat{\mathbf{h}}_s \cdot \hat{\mathbf{t}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i) - (\hat{\mathbf{n}} \cdot \hat{\mathbf{d}}) (\hat{\mathbf{v}}_s \cdot \hat{\mathbf{k}}_i) + (\hat{\mathbf{v}}_s \cdot \hat{\mathbf{d}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i)]_{s_1} \quad (21b)$$

$$f_{hv} = [(1 - R_v) \hat{\mathbf{v}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{v}}) - (1 + R_v) \hat{\mathbf{h}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{h}})]_{s_1} - (R_h + R_v) (\hat{\mathbf{v}} \cdot \hat{\mathbf{t}}) [(\hat{\mathbf{h}}_s \cdot \hat{\mathbf{t}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i) - (\hat{\mathbf{n}} \cdot \hat{\mathbf{d}}) (\hat{\mathbf{v}}_s \cdot \hat{\mathbf{k}}_i) + (\hat{\mathbf{v}}_s \cdot \hat{\mathbf{d}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i)]_{s_1} \quad (21c)$$

$$f_{hh} = [(1 + R_h) \hat{\mathbf{v}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{h}}) + (1 - R_h) \hat{\mathbf{h}}_s \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{v}})]_{s_1} - (R_h + R_v) (\hat{\mathbf{h}} \cdot \hat{\mathbf{d}}) [(\hat{\mathbf{h}}_s \cdot \hat{\mathbf{d}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i) - (\hat{\mathbf{n}} \cdot \hat{\mathbf{d}}) (\hat{\mathbf{h}}_s \cdot \hat{\mathbf{k}}_i) - (\hat{\mathbf{v}}_s \cdot \hat{\mathbf{t}}) (\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}_i)]_{s_1} \quad (21d)$$

where  $s_1 = \sqrt{1 + z_x^2 + z_y^2}$ ,  $z_x$  and  $z_y$  denote partial differential on  $x$  and  $y$ , respectively.

Noted that in IEM model, the terms involved  $(R_h + R_v)$  in Eq. (21) are all dropped off. Keeping in mind that the Kirchhoff field coefficient  $f_{qp}$  is spatially dependent via the Fresnel reflection coefficients  $R_p$  ( $p$  denotes h or v) and the surface slope term  $s_1$ . To make the integrals in Eq. (19) and Eq. (20) mathematically manageable in calculation of the average scattered power, we apply a stationary phase approximation while ignoring the edge diffraction, to obtain an estimate of the surface slopes. That is, the surface slopes are presumably independent of spatial variable, and are approximately determined by the directions of incident and scattering waves. To further tackle the mathematical manipulations, removal of the spatial dependence of the reflection coefficient will be made in Section 2.3. Keeping in mind that to what extent such estimate is valid or at least sufficiently accurate remains to be solved in further investigations.

Now go back to the complementary scattered

field, which is much more complicated to deal with. Recalled that the complementary scattered field is contributed from the reradiated fields that may propagate through medium 1 and medium 2, in way of both upwardly and downwardly. The physical mechanism may be graphically represented by the field coefficients or the propagators,  $\mathcal{F}_{qp}$  and  $\mathcal{G}_{qp}$ , as illustrated in Fig. 2. Further processing of the phase term involving the surface height, the propagators may be decomposed into the upward components designated by  $\mathcal{F}_{qp}^+$  and  $\mathcal{G}_{qp}^+$  and the downward components by  $\mathcal{F}_{qp}^-$  and  $\mathcal{G}_{qp}^-$ , mathematically appearing as the absolute terms in Eq. (16), physically denoting the change of propagation velocity in different media.

In what follows, the complete phase terms are kept and all possible propagation waves are included. After straightforward but tedious mathematical manipulations, the complementary field coefficients can be obtained and put into compact forms for both the upward and downward propagations. Explicit expressions that are easy for numerical computation are given<sup>[14,18]</sup>. For cross polarizations, we use an approximate reflection coefficient by taking the average of  $R_h$  and  $R_v$ :  $R \approx \frac{1}{2}(R_h + R_v)$ <sup>[3]</sup>.

### 2.3 Bistatic scattering coefficients

Now that with the scattered fields calculated, we perform ensemble averaging to compute the scattered power and the scattering coefficient. To gain more physical insights into the field interactions that produce the average power, the following expression for the incoherent average power  $P_{qp}^s$  is written as a sum of three terms: the Kirchhoff power  $P_{qp}^k$ , the cross power  $P_{qp}^{kc}$  due by the Kirchhoff field and the complementary, and the complementary power  $P_{qp}^c$ .

$$P_{qp}^s = \langle E_{qp}^s E_{qp}^{s*} \rangle - \langle E_{qp}^s \rangle \langle E_{qp}^{s*} \rangle = \langle E_{qp}^k E_{qp}^{k*} \rangle - \langle E_{qp}^k \rangle \langle E_{qp}^{k*} \rangle + 2\text{Re}[\langle E_{qp}^c E_{qp}^{rk*} \rangle - \langle E_{qp}^c \rangle \langle E_{qp}^{rk*} \rangle] + \langle E_{qp}^c E_{qp}^{c*} \rangle - \langle E_{qp}^c \rangle \langle E_{qp}^{c*} \rangle \triangleq P_{qp}^k + P_{qp}^{kc} + P_{qp}^c \quad (22)$$

where  $E_{qp}^s$  and  $E_{qp}^{s*}$  denote scattered  $q$ -polarized electric scattering incident  $p$ -polarized wave and its conjugate field, respectively;  $E_{qp}^k$  and  $E_{qp}^{k*}$  similar due to Kirchhoff field, respectively,  $E_{qp}^c$  denotes so due to complementary power.

Geometrically, it is readily recognized that the cross power is the result of the interactions between the Kirchhoff field and the complementary field, being involved by four terms—two accounting for the upper medium propagation and the other two for the lower medium:

$$P_{qp}^{kc} = P_{qp}^{kc} | \mathcal{F}_{qp}^+ + P_{qp}^{kc} | \mathcal{F}_{qp}^- + P_{qp}^{kc} | \mathcal{G}_{qp}^+ + P_{qp}^{kc} | \mathcal{G}_{qp}^- \quad (23)$$

Similarly, the radiated power by the complementary field itself is mutually generated by re-radiation fields from every point on the surface and thus consists of sixteen terms resulting from interactions of two upward fields and two downward fields propagating through medium 1 and medium 2:

$$\begin{aligned} P_{qp}^c &= P_{qp}^c | \mathcal{F}_{qp}^+ \mathcal{F}_{qp}^+ + P_{qp}^c | \mathcal{F}_{qp}^+ \mathcal{F}_{qp}^- + P_{qp}^c | \mathcal{F}_{qp}^- \mathcal{F}_{qp}^+ + P_{qp}^c | \mathcal{F}_{qp}^- \mathcal{F}_{qp}^- + \\ &P_{qp}^c | \mathcal{G}_{qp}^+ \mathcal{G}_{qp}^- + P_{qp}^c | \mathcal{G}_{qp}^- \mathcal{G}_{qp}^+ + P_{qp}^c | \mathcal{G}_{qp}^- \mathcal{G}_{qp}^- + \\ &P_{qp}^c | \mathcal{F}_{qp}^- \mathcal{G}_{qp}^+ + P_{qp}^c | \mathcal{F}_{qp}^- \mathcal{G}_{qp}^- + P_{qp}^c | \mathcal{G}_{qp}^+ \mathcal{F}_{qp}^+ + \\ &P_{qp}^c | \mathcal{G}_{qp}^+ \mathcal{F}_{qp}^- + P_{qp}^c | \mathcal{G}_{qp}^+ \mathcal{F}_{qp}^+ + P_{qp}^c | \mathcal{G}_{qp}^- \mathcal{F}_{qp}^+ + \\ &P_{qp}^c | \mathcal{G}_{qp}^- \mathcal{F}_{qp}^- + P_{qp}^c | \mathcal{G}_{qp}^- \mathcal{F}_{qp}^- \end{aligned} \quad (24)$$

Now by substituting the Kirchhoff field and the complementary field into Eq. (24) and carrying out the ensemble averages, we may obtain an explicit expression of the incoherent average power. When the medium is very lossy, contributions from the lower medium propagation is expected to be small. In general, all modes must be included to get a more complete scattered power, as claimed in Ref. [19]. Whether it really so needs further investigation, however. After some algebraic manipulations and arrangements by Eq. (7) we can reach the final expression of a relatively compact form as

$$\begin{aligned} \sigma_{qp}^s &= \frac{k_1^2}{2} \exp[-\sigma^2(k_{iz}^2 + k_{sz}^2)] \cdot \\ &\sum_{n=1}^{\infty} \frac{\sigma^{2n}}{n!} | \mathcal{F}_{qp}^n |^2 W^{(n)}(k_{sx} - k_{ix}, k_{sy} - k_{iy}) \end{aligned} \quad (25)$$

where  $\sigma$  denotes bistatic scattering coefficient;  $k_1$  denotes wave number in medium 1;  $\sigma_{qp}^s$  denotes cross-polarized scattering coefficient determined in for incident  $p$ -polarized and scattering  $q$ -polarized wave;  $\sigma^{2n}$  denotes the  $n$ th order scattering coefficient;  $n$  denotes  $n$ th order;  $\mathcal{F}_{qp}^n$  denotes scattering amplitude;  $W^{(n)}(k_{sx} - k_{ix}, k_{sy} - k_{iy})$  denotes the

surface roughness spectrum of the surface related to the  $n$ th power of the surface correlation function by the two-dimensional Fourier transform, assuming that the surface height follows a Gaussian distribution.

$R_h$  and  $R_v$ , are defined as follows, respectively:

$$R_h = \frac{\mu_1 k_i \cos \theta_i - \mu_0 k_{iz}}{\mu_1 k_i \cos \theta_i + \mu_0 k_{iz}} \quad (26a)$$

$$R_v = \frac{\varepsilon_0 k_{iz} - \varepsilon_1 k_i \cos \theta_i}{\varepsilon_1 k_i \cos \theta_i + \varepsilon_0 k_{iz}} \quad (26b)$$

Where

$$k_{iz} = \text{Re}\{k_{iz}\} + j\text{Im}\{k_{iz}\} \quad (26c)$$

With

$$\begin{aligned} \text{Re}\{k_{iz}\} &= \left[ \frac{1}{\sqrt{2}} [\text{Re}\{k_i^2\} - k_i^2 \sin^2 \theta_i]^{1/2} + \right. \\ &\left. \sqrt{(\text{Re}\{k_i^2\} - k_i^2 \sin^2 \theta_i)^2 + (\text{Im}\{k_i^2\})^2} \right]^{1/2} \end{aligned} \quad (27a)$$

$$\begin{aligned} \text{Im}\{k_{iz}\} &= -\frac{1}{\sqrt{2}} \left[ -(\text{Re}\{k_i^2\} - k_i^2 \sin^2 \theta_i) + \right. \\ &\left. \sqrt{(\text{Re}\{k_i^2\} - k_i^2 \sin^2 \theta_i)^2 + (\text{Im}\{k_i^2\})^2} \right]^{1/2} \end{aligned} \quad (27b)$$

where  $k_i$  denotes wave number in transmitted lower layer.

After the above extensions, we shall name the IEM-based model as the AIEM model.

### 3 Comparison theoretical predictions with numerical simulations and experimental data

#### 3.1 Comparison theoretical predictions with numerical simulations

The prediction of scattering coefficient in scattering plane between the present model and numerical results of SSA (Small Slope Approximation) and MoM (Method of Moment) are shown in Fig. 3(a). The simulation data is adopted from Ref. [22] for a Gaussian correlated surface with  $\varepsilon_r = 4 - j1$ ,  $k\sigma = 0.5$ ,  $kl = 3.0$  at incident angle of  $30^\circ$  and scattering angle between  $-60^\circ$  and  $60^\circ$ , where  $l$  denotes correlation length. Obviously, all the three predictions are quite close to each other except at larger back-scattering angle. There is dip in specular direction shown by MoM and SSA, but not by AIEM. Doubling the surface roughness, results are given in Fig. 3(b) from which we can see that the angular trends by three predictions are similar to those in Fig. 3(a).

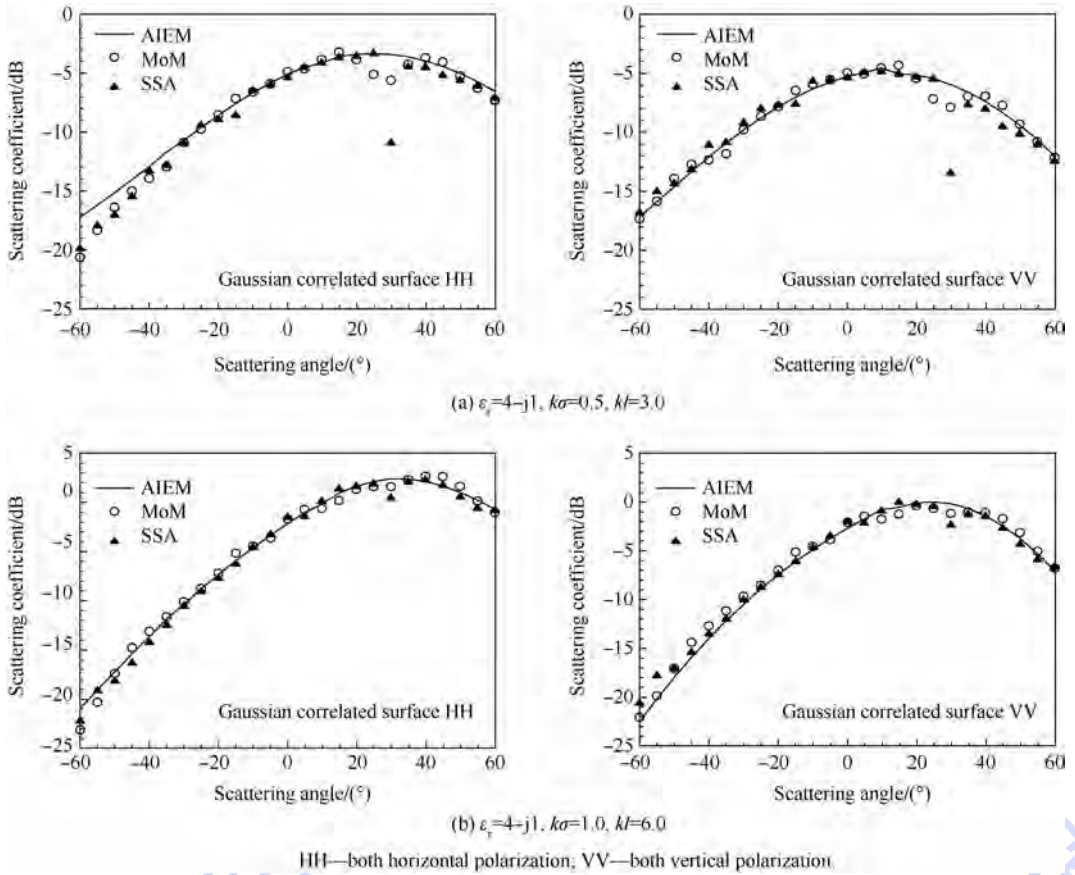


Fig. 3 Comparison of scattering coefficient between AIEM model and numerical results of MoM and SSA for both horizontal and vertical polarization for a Gaussian correlated surface with incident angle of  $30^\circ$

### 3.2 Comparison theoretical predictions with measurement data

An excellent experimental data set for the purpose of comparison is adopted from Ref. [23] for a Gaussian correlated surface with  $\sigma = 0.4$  cm,  $l = 6$  cm; the scattering coefficient was measured at two incident angles,  $20^\circ$  and  $40^\circ$ ; incident frequencies of 2, 5, and 10 GHz, resulting in three different roughness scales but keeping at same surface slope as shown in Fig. 4. At this surface slope of Gaussian correlation, it is believed that the multiple scattering is very small. Fig. 4(a) shows the bistatic scattering behavior from which it is seen that both the AIEM and SPM agree well with the measured data except at scattering angles near  $20^\circ$  and beyond  $20^\circ$ . This might be due to a strong coherent contribution to the measurements, but still remains to be further confirmed. Measurements at frequency of 5 GHz are presented in Fig. 4 (b). The surface is rougher for this case, so we plotted the GOM (Geometrical Optics Model) prediction as a reference. At this roughness scale with incident angle of  $20^\circ$ , neither

GOM nor SPM matches with the experimental data, while the AIEM is in excellent agreement with the measured data. The scattering coefficient of HH and VV are closer compared to the measurement at 2 GHz for higher frequency. Further increasing the incident angle to  $40^\circ$ , similar observations can be drawn, as is evident from Fig. 4 (c). Also it is clear that along forward scattering near specular direction and beyond, the experimental data presented some degree of fluctuations—may be due to the remaining of coherent scattering which is stronger for smoother surface and tends to reduce as surface becomes rougher. At incident frequency of 10 GHz, by closely inspecting Fig. 4(d) and Fig. 4 (e), both the AIEM and GOM present similar angular trends in forward scattering region and both agree well with the measured data. However, at backscattering with incident angle larger than  $20^\circ$ , the measurement reveals excessively higher than the model predictions of the AIEM and GOM. The difference seems not possibly be explained from the exclusion of multiple scattering, for the surface slope remains the same and small.

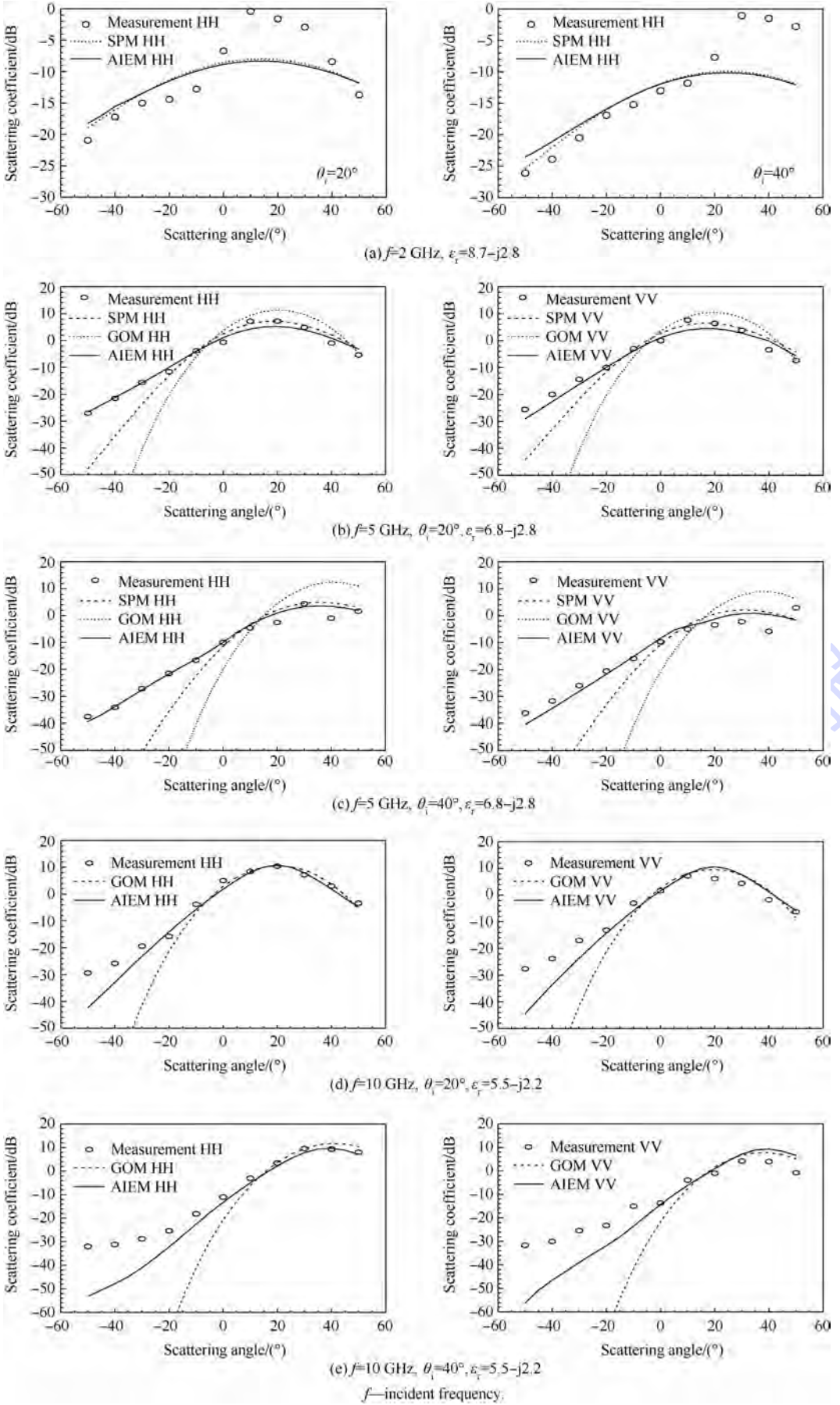


Fig. 4 Comparison of bistatic scattering coefficient between model predictions and measurement as function of scattering angle for a Gaussian correlated surface ( $\sigma = 0.4\text{ cm}$ ,  $l = 6\text{ cm}$ )

## 4 Bistatic scattering features in azimuthal plane

In bistatic scattering, it is of interest to observe the azimuthal pattern. From Fig. 1, if the transmitter is located at  $\phi_i = 0^\circ$ , the null or dip of the scattering return presumably will occur near  $\phi_s = 90^\circ$  for plane (flat) surface. Things are much more complicated, however. In what follows, we present two numerical cases to exam the roughness effect on the bistatic scattering pattern. The incident and scattering angles were chosen as  $\theta_i = 20^\circ, \theta_s = 40^\circ$ , respectively, and the surface relative permittivity of  $\epsilon_r = 10 + j0.05$ . Fig. 5 displays the scattering coefficients for both HH and VV polarizations, where Fig. 5(a) is the case of smooth surface with  $\sigma = \frac{\lambda}{20}, l = \frac{\lambda}{2}$ , while in Fig. 5(b) the roughness is doubled ( $\sigma = \frac{\lambda}{10}, l = \lambda$ ),  $\lambda$  being the radar wavelength. For reference, results of SPM and GOM models are also plotted.

In small roughness cases, we see that both AIEM and SPM are in agreement, as expected. For

VV polarization, the dip occurs ahead of that of HH polarization. As surface gets rougher, AIEM follows closely with GOM, also as expected. The dip now moves forward for HH polarization, but moves backward for VV polarization. Together the preceding features suggest that the dip location, which is highly polarization dependent, is an effective indicator of surface roughness. It is profoundly found that unlike SPM and GOM, the AIEM model consistently traces the dip location in the sense of roughness and polarization dependence.

Fig. 6 displays the bistatic scattering coefficients of the SSA2 (Fig. 6(a)) and AIEM (Fig. 6(b)) in all four polarizations, VV, HH, HV, and VH<sup>[24]</sup> from a rough surface having rms height  $\sigma = \lambda/20$ , a Gaussian correlation function with correlation length  $l = \lambda/2$  and a surface relative permittivity of  $\epsilon_r = 10 + j0.05$ . Generally, similar behavior between SSA2 (second order small slope approximation) and AIEM, which are hemispherical patterns, is observed, including the locations of the dips that are dependent on the polarization<sup>[25]</sup>.

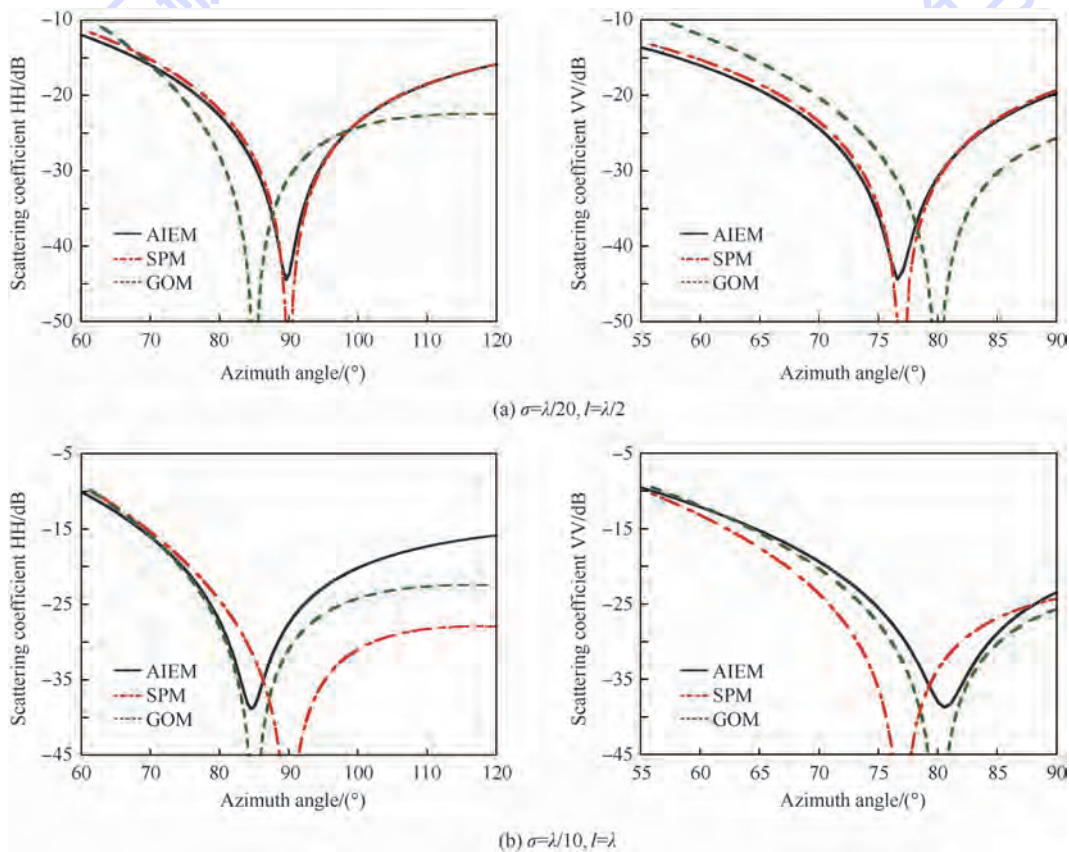


Fig. 5 Effect of surface roughness on dip location in azimuthal plane ( $\theta_i = 20^\circ, \theta_s = 40^\circ, \epsilon_r = 10 + j0.05$ )

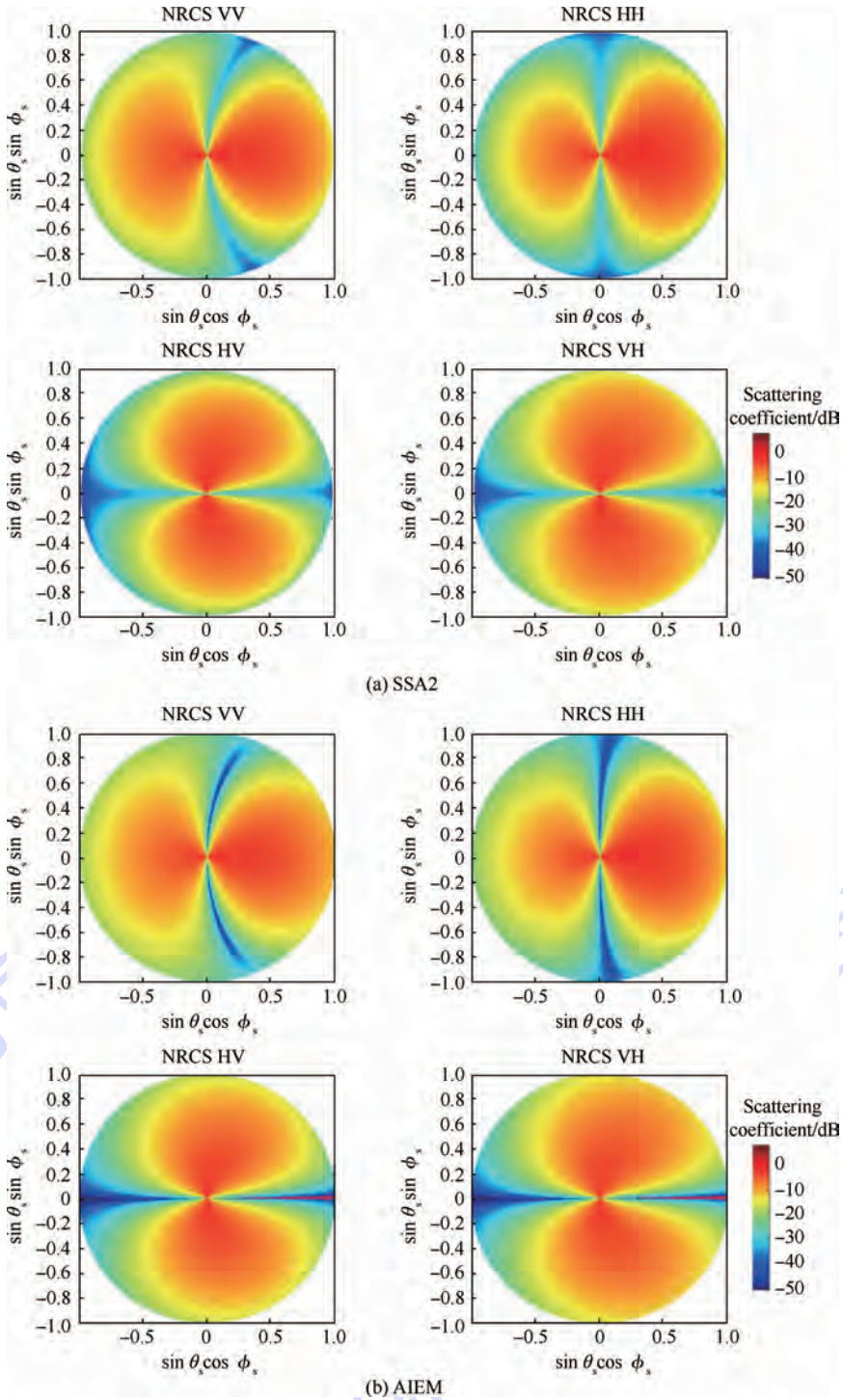


Fig. 6 Bistatic scattering coefficients of VV, HH, HV, and VH for SSA2, and AIEM<sup>[24]</sup>

## 5 Conclusions

Analysis of bistatic scattering from randomly rough surface was presented. The results can be summarized as follows.

1) Validation of the AIEM for bistatic scattering is made by comparison with numerical simulations and experimental measurements.

2) The AIEM model consistently traces the dip location in the sense of roughness and polarization

dependence.

3) The hemispherical plots of AIEM, SSA2, MoM further confirm that the locations of the dips are dependent on the polarization.

4) Scattering in azimuthal plane is an effective indicator of surface roughness. The dependence of surface relative permittivity should be of importance and is under investigation.

5) The study of bistatic scattering through model simulations is useful for designing future radar

measurement configuration as far as surface parameters retrieval is mattered.

## References

- [1] Beckman P, Spizzichino A. The scattering of electromagnetic waves from rough surfaces[M]. Oxford: Pergamon Press, 1963: 215-285.
- [2] Ulaby F T, Moore R K, Fung A K. Microwave remote sensing: Active and passive, Vol. II-radar remote sensing and surface scattering and emission theory[M]. Norwood, MA: Artech House, 1982: 922-1025.
- [3] Fung A K. Microwave scattering and emission models and their applications[M]. Norwood, MA: Artech House, 1994: 1-10.
- [4] Voronovich A G. Wave scattering from rough surfaces[M]. Berlin: Springer-Verlag, 1994: 1-5.
- [5] Tsang L, Kong J A. Scattering of electromagnetic waves: Advanced topics[M]. New York: John Wiley & Sons, 2001: 65-118.
- [6] Fung A K, Chen K S. Microwave scattering and emission models for users[M]. Norwood, MA: Artech House, 2009: 1-6.
- [7] Fung A K, Li Q, Chen K S. Backscattering from a randomly rough dielectric surface[J]. IEEE Transactions on Geoscience and Remote Sensing, 1992, 30(2): 356-369.
- [8] Chen K S, Fung A K, Weissman D E. A backscattering model for sea surfaces[J]. IEEE Transactions on Geoscience and Remote Sensing, 1992, 30(4): 811-817.
- [9] Hsieh C Y, Fung A K, Nesti G, et al. A further study of the IEM surface scattering model[J]. IEEE Transactions on Geoscience and Remote Sensing, 1997, 35: 901-909.
- [10] Chen K S, Wu T D, Tsay M K, et al. A note on the multiple scattering in an IEM model[J]. IEEE Transactions on Geoscience and Remote Sensing, 2000, 38(1): 249-256.
- [11] Wu T D, Chen K S, Shi J C, et al. A transition model for the reflection coefficient in surface scattering[J]. IEEE Transactions on Geoscience and Remote Sensing, 2001, 39(9): 2040-2050.
- [12] Álvarez-Pérez J L. An extension of the IEM/IEMM surface scattering model[J]. Waves in Random Media, 2001, 11(3): 307-329.
- [13] Fung A K, Liu W Y, Chen K S, et al. An improved IEM model for bistatic scattering[J]. Journal of Electromagnetic Wave and Applications, 2002, 16(5): 689-702.
- [14] Chen K S, Wu T D, Tsang L, et al. Emission of rough surfaces calculated by the integral eq. method with comparison to three-dimensional moment method simulations[J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(1): 90-101.
- [15] Wu T D, Chen K S. A reappraisal of the validity of the IEM model for backscattering from rough surfaces[J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(8): 743-753.
- [16] Fung A K, Chen K S. An update on IEM surface backscattering model[J]. IEEE Geoscience and Remote Sensing Letters, 2004, 1(2): 75-77.
- [17] Du Y. A new bistatic model for electromagnetic scattering from randomly rough surfaces[J]. Waves Random Complex Media, 2008, 18(1): 109-128.
- [18] Wu T D, Chen K S, Shi J C, et al. A study of AIEM model for bistatic scattering from randomly surfaces[J]. IEEE Transactions on Geoscience and Remote Sensing, 2008, 46(9): 2584-2598.
- [19] Álvarez-Pérez J L. The IEM2M rough-surface scattering model for complex-permittivity scattering media[J]. Waves in Random and Complex Media, 2012, 22(2): 207-233.
- [20] Poggio A J, Miller E K. Computer techniques for electromagnetics[M]. Oxford: Pergamon Press, 1973: 159-260.
- [21] Li Z, Fung A K. A Reformulation of the surface field integral equation[J]. Journal of Electromagnetic Waves and Application, 1991, 5: 195-203.
- [22] Ewe H T, Johnson J T, Chen K S. A comparison study of the surface scattering models and numerical model[C] // IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ: IEEE Press, 2001: 2692-2694.
- [23] Macelloni G M, Nesti G, Pampaloni P, et al. Experimental validation of surface scattering and emission models[J]. IEEE Transactions on Geoscience and Remote Sensing, 2000, 38(1): 459-469.
- [24] Yardim C, Johnson J T, Burkholder R J, et al. An intercomparison of models for predicting bistatic scattering from rough surfaces[C] // Proceedings of IGARSS. Piscataway, NJ: IEEE Press, 2015, in press.
- [25] Johnson J T, Ouellette J D. Polarization features in bistatic scattering from rough surfaces[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(3): 1616-1626.

# Model analysis of bistatic scattering from randomly rough surfaces

CHEN Kunshan<sup>\*1,2</sup>, LI Zhaoliang<sup>2</sup>, LIU Yu<sup>1</sup>

(1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth,  
Chinese Academy of Science, Beijing 100101, China;

2. Key Laboratory of Agri-informatics, Ministry of Agriculture, Institute of Agricultural Resources and Regional Planning,  
Chinese Academy of Agricultural Sciences, Beijing 100081, China)

**Abstract:** Model analysis was presented of bistatic scattering features from randomly rough surfaces theoretically. In particular, a statistical advanced integral equation model (AIEM) was selected as the working model. The scattering coefficients in the location of the dip in azimuthal plane were investigated for different roughness scale. Theoretical predictions of scattering coefficients were illustrated, demonstrated, and validated by comparisons with numerical simulations, and measurements. The results show that the dip location or region is both rough and polarization dependent. Simulation results offer deeper physical insights into the scattering features of statistically rough surface and thus are useful or even vital for designing future and yet more effective bistatic measurement configurations.

**Key words:** bistatic scattering; scattering coefficients; randomly rough surfaces; integral equation model (IEM); advanced integral equation model (AIEM); numerical simulation

**Received:** 2015-04-09; **Accepted:** 2015-05-05; **Published online:** 2015-07-02 10:25

**UPL:** [www.cnki.net/kcms/detail/11.2625.V.20150702.1025.001.html](http://www.cnki.net/kcms/detail/11.2625.V.20150702.1025.001.html)

**Foundation item:** The Director General's Innovative Funding-2015 of the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences.

\* **Corresponding author:** CHEN Kunshan (1959—), Male, Taiwan, China, Professor, chenks@radi.ac.cn, main research interest is microwave remote sensing.

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0207

# 带状注分布作用速调管多间隙谐振腔高频特性

阮存军<sup>\*1</sup>, 戴军<sup>1</sup>, 陈姝媛<sup>2</sup>, 李仁杰<sup>1</sup>, 占敏<sup>1</sup>

(1. 北京航空航天大学 电子信息工程学院, 北京 100191; 2. 试验物理与计算数学国家重点实验室, 北京 100076)

**摘 要:** 带状注分布作用速调管(SBEIK)的典型特征是平面多间隙谐振腔及其分布式注波相互作用系统. 对应于W波段微型化SBEIK的一种强耦合式五间隙分布作用谐振腔进行了研究,并结合传统的弱耦合式谐振腔与输出腔对其高频特性进行了深入分析,结果表明强耦合式腔体具备谐振模式隔离、各间隙高频场(RF)的耦合及其能量输出的良好技术优势.此外,对五间隙谐振腔周期结构引入的轴向筒并模式竞争问题进行了研究,得到其工作模式与竞争模式之间的频率间隔在600 MHz以上,完全可以满足SBEIK整体方案设计中100 MHz带宽的要求.最后,利用三维粒子模拟(PIC)软件对优化后的强耦合式五间隙输出腔注波相互作用性能进行了初步的模拟仿真,验证了其在相应工作模式上具备了高功率产生与输出的技术特征.本研究对毫米波和太赫兹频段高功率SBEIK的物理设计与工程研究具有重要的价值.

**关键词:** W波段; 带状电子注; 强耦合式; 五间隙谐振腔; 分布作用速调管(EIK)

**中图分类号:** O441.3

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1777-08

带状注速调管(Sheet Beam Klystron, SBK)采用了可平面扩展的带状电子注,因而具有了低空间电荷力、易于传输强流电子注、大平面空间所带来的高功率容量等技术优势<sup>[1-2]</sup>. 分布作用速调管(Extended Interaction Klystron, EIK)则采用周期性多间隙耦合谐振腔高频(RF)结构,因而具有了大特性阻抗和高注波相互作用效率等技术优势<sup>[3]</sup>. 带状注分布作用速调管(Sheet Beam Extended Interaction Klystron, SBEIK)则很好地兼顾了SBK和EIK的技术优点,不仅克服了传统速调管向高频段发展时所带来的工作频率与几何尺寸共度性效应物理限制,而且可以有效地避免由于高频段小尺寸结构带来的高频击穿问题,具有了在高频段产生高效率、高增益和高功率输出的优良技术特性. 近几年, SBEIK已成为国内外在毫米波和太赫兹频段产生高性能相干辐射输出重要的研究方向

和研究热点,并已在W波段和G波段取得了一定的研究进展<sup>[4-5]</sup>.

与传统SBK采用的平面周期性结构中多个相互作用间隙及其耦合腔相互分离的弱耦合式谐振高频系统不同, SBEIK通过采用连通的大平面结构将周期结构的多个间隙及其谐振腔进行有效耦合,以构成具有分布式特性的强耦合式多间隙谐振腔高频及注波相互作用系统. 但是,无论采用弱耦合式还是强耦合式的平面高频结构,它不仅要求高频系统与平面带状电子注之间形成良好的匹配,而且应采用多个具有周期结构的多间隙谐振腔以进一步增加腔体的功率容量,并在毫米波和太赫兹频段极大地提高注波相互作用效率和功率增益. 但是,随着谐振腔体间隙周期数的增加,其大平面过模结构所带来的工作模式产生及其互作用场型建立等问题的复杂度增加,模式密集所导致

收稿日期: 2015-04-10; 录用日期: 2015-05-14; 网络出版时间: 2015-06-25 15:49

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150625.1549.001.html

基金项目: 国家自然科学基金(61222110); 北京航空航天大学基本科研基金(30418501)

\* 通讯作者: 阮存军(1974—),男,甘肃天水人,研究员, ruancunjun@buaa.edu.cn,主要研究方向为微波、毫米波与太赫兹电子学.

**引用格式:** 阮存军, 戴军, 陈姝媛, 等. 带状注分布作用速调管多间隙谐振腔高频特性[J]. 北京航空航天大学学报, 2015, 41(10): 1777-1784. Ruan C J, Dai J, Chen S Y, et al. RF characteristics of multiple-gap resonant cavity for sheet beam extended interaction klystron [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1777-1784 (in Chinese).

的模式竞争问题愈加严重,以及高频功率产生与输出特性的稳定性等问题更加突出.这些难点问题都给这种复杂多间隙腔体的研究与应用带来很大技术挑战.

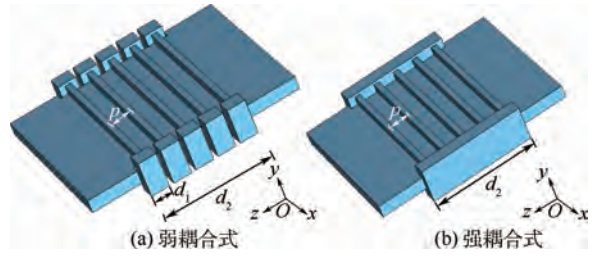
近几年,基于平面带状注技术发展毫米波和太赫兹频段的新颖高功率真空电子器件已成为了本领域非常重要的研究热点,并在新型平面高频系统构建方面提出了两种技术途径:一种是早期提出并广泛研究的传统弱耦合式多间隙谐振腔及其 SBK 方案<sup>[6-8]</sup>,其主要研究工作大都集中于低频微波波段(如 X 波段)和毫米波段(如 W 波段)<sup>[9-14]</sup>;另一种则是近几年开始提出并逐步发展的强耦合式多间隙谐振腔及其 SBEIK 研究工作,主要集中在毫米波段(如 W 波段)和太赫兹波段(如 G 波段)<sup>[1,15,16]</sup>.美国海军实验室(NRL)于2009年设计的 G 波段 SBEIK,模拟仿真获得了 453 W 的放大功率输出,增益为 41.6 dB,效率为 5.28%<sup>[3]</sup>.美国 UC-Davis 也于 2009 年开展了 W 波段高功率 SBEIK 的研究工作,对其注波相互作用系统进行了三维仿真设计,获得了 50 kW 的功率输出,增益为 40 dB,效率为 18.8%,瞬时带宽为 150 MHz<sup>[15]</sup>.2013 年,NRL 对 W 波段 SBEIK 进行了详细研究,在国际上首次研制出了输出功率为 7.5 kW、增益为 40 dB、效率为 10.7%、带宽为 150 MHz 的大功率高频输出实验样管<sup>[16]</sup>.

本文结合正在开展的 W 波段 SBEIK 的研究工作,并与传统的弱耦合式多间隙谐振腔和输出腔进行详细比较与分析,对基于强耦合式的五间隙分布作用谐振腔高频及其注波相互作用系统构建并对其特性进行了深入的研究.获得的结果,为下一步毫米波和太赫兹频段基于平面化和微型化 SBEIK 的大功率相干辐射源研制与应用奠定了良好基础.

### 1 强耦合式五间隙腔高频特性

图 1 为弱耦合式和强耦合式的五间隙谐振腔示意图.传统 SBK 采用了弱耦合式的平面结构哑铃型多间隙谐振腔以进行电子注与高频系统之间的高效率注波相互作用和能量交换,如图 1(a)所示.这种谐振腔每个间隙两端均有各自独立的耦合谐振结构,各个间隙之间只能通过电子注传输的漂移通道进行连接和耦合.而 SBEIK 则对这种弱耦合式谐振腔进行改进,它将每个间隙两侧的谐振耦合结构合并连通为一个大的谐振腔,如图 1(b)所示.可以看出,这种强耦合式多间隙谐

振腔各间隙之间的耦合包含两部分:一部分是由两端各自连通的大耦合腔产生;另一部分则由相邻间隙之间的电子注漂移通道建立相应的耦合.



$d_1$ —弱耦合式腔中每个间隙的耦合腔沿 z 轴向上的尺寸;  
 $d_2$ —强耦合式腔体整体耦合腔沿 z 轴向上的尺寸;  
 $p$ —多间隙腔间隙周期长度.

图 1 弱耦合式和强耦合式五间隙谐振腔示意图  
Fig. 1 Configuration of weak-couple and strong-couple five-gap resonant cavity

#### 1.1 五间隙谐振腔

为了对弱耦合式和强耦合式多间隙谐振腔高频特性进行研究,利用三维仿真软件 CST MICROWAVE STUDIO 对应用于 W 波段 SBK 的五间隙谐振腔体进行建模和仿真,如图 1 所示.两种腔体所建模型除了各间隙两边的谐振耦合结构及其尺寸有所差别外(图 1(a)中 5 个分离的耦合腔沿 z 轴向上的总长度则与图 1(b)一致,均为  $d_2$ ),其他尺寸(包括间隙尺寸、各间隙之间的距离、电子注漂移管尺寸等)完全一致.同时,两种模型中电子注漂移管对所研究的工作模式(如  $2\pi$  模式)不截止,可以在相邻间隙之间建立起一定的耦合.图 2 给出了经过仿真设计与参数优化得到的两种五间隙谐振腔  $2\pi$  工作模式的电场形分布图,可以看出无论两边的耦合腔是否连通,对于谐振电场分布没有明显影响,均集中在各个间隙中,且最中间间隙的谐振场强最大.图 3 进一步比较了弱耦合式与强耦合式谐振腔轴向电场强度沿 z 轴向( $E_z$ )的分布,可以看出两条曲线几乎完全重合.

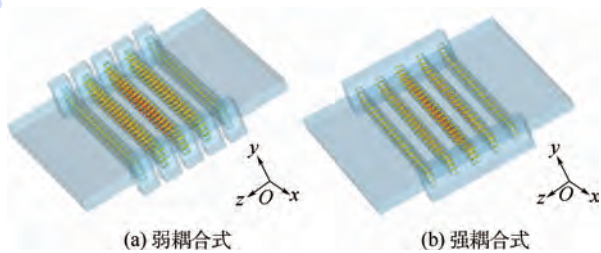


图 2 工作在  $2\pi$  模式的弱耦合式和强耦合式五间隙谐振腔沿 z 轴向电场分布图  
Fig. 2  $2\pi$  mode electric fields distribution diagram along z-axis for weak-couple and strong-couple five-gap resonant cavity

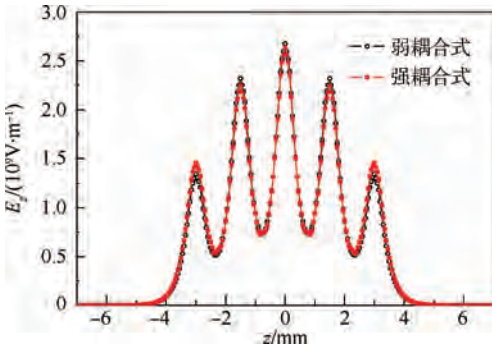


图 3 工作在  $2\pi$  模式的弱耦合式和强耦合式五间隙谐振腔  $z$  轴向电场对比图

Fig. 3 Comparison diagram of  $2\pi$  mode electric fields distribution along  $z$ -axis for weak-couple and strong-couple five-gap resonant cavity

表 1 给出了弱耦合式与强耦合式五间隙谐振腔的高频特性参数计算结果的对比,可以看出,其谐振频率  $f$ 、总特性阻抗  $R/Q$  及本征品质因数  $Q_0$  都几乎相等。

表 1 弱耦合式和强耦合式五间隙谐振腔高频参数对比

Table 1 Comparison of RF parameters for weak-couple and strong-couple five-gap resonant cavity

谐振腔种类	高频参数		
	$f/\text{GHz}$	$\frac{R}{Q}/\Omega$	$Q_0$
弱耦合式	94.65	70.43	1347.3
强耦合式	94.54	71.68	1348.3

其中:  $R/Q$  的计算公式为

$$\frac{R}{Q} = \frac{\left( \int_{-\infty}^{\infty} |E_z| dz \right)^2}{2wW_s} \quad (1)$$

式中:  $w$  为谐振角频率;  $W_s$  为谐振腔的储能。因此,当将五间隙腔作为 SBK 谐振腔使用时,在漂移管对工作模式不截止时,弱耦合式和强耦合式腔的高频特性几乎保持不变,因此二者将具有基本相同的注波相互作用过程。但是,若缩小电子注漂移通道尺寸至对工作模式截止时,强耦合式谐振腔将依然保持其良好的整体耦合和谐振特性,而弱耦合式腔体各间隙只能转化为互相分离的单个间隙谐振腔。此外,将 SBK 应用于毫米波和太赫兹频段时,对其微型化结构的复杂腔体工程实现及其腔体特性的动态调谐而言,充分利用强耦合式多间隙谐振腔连通的大尺寸耦合结构将具有更重要的价值。

更进一步,由上述五间隙谐振腔分析及 SBK 注波相互作用要求可知,所采用的多间隙谐振腔其电场应尽可能集中于各间隙中间并保持较高的电

场强度,同时也应保持各间隙电场及其谐振模式之间形成良好的隔离度以避免各自之间的干扰和互相影响。为此,采用强耦合式五间隙谐振腔将是一种更优的技术方案,可以通过将各间隙之间电子注漂移通道设计在截止状态,以确保各间隙之间的高频隔离并避免干扰。同时,通过其两侧连通的耦合腔可以实现各间隙之间很好的模式耦合并产生所需的电场强度幅值。相比较而言,采用弱耦合式多间隙谐振腔将很难同时对这两个方面给予兼顾并同时达到良好的工作状态。

### 1.2 五间隙输出腔

在 SBK 研究中,采用更多间隙的输出腔可以获得更高的注波相互作用效率和功率容量,并产生高功率的相干辐射输出。为此,仍然以五间隙腔为研究对象,在 CST(Computer Simulation Technology)微波工作室中分别建立弱耦合式和强耦合式的五间隙输出腔三维模型,如图 4 所示,其各个间隙及耦合腔结构尺寸与图 1 中相应谐振腔的几何参数基本一致,同时两者采用的能量耦合口与输出波导结构也完全一致。图 5 给出了通过仿真计算和优化得到的  $2\pi$  工作模式的输出腔电场分布,图 6 给出了其工作模式和轴向电场强度沿  $z$  轴向的分布图,可以看出两者存在很大的区别。强耦合式五间隙腔在不开外接耦合口和具有外接耦合口两种情况下,其场形几乎没有太大变化,这是因为其外接耦合口与五间隙整体的耦合腔体相连,故而

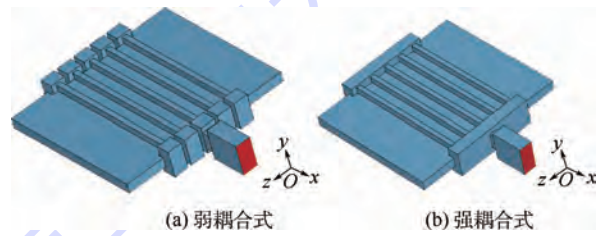


图 4 弱耦合式和强耦合式五间隙输出腔示意图  
Fig. 4 Configuration of weak-couple and strong-couple five-gap output cavity

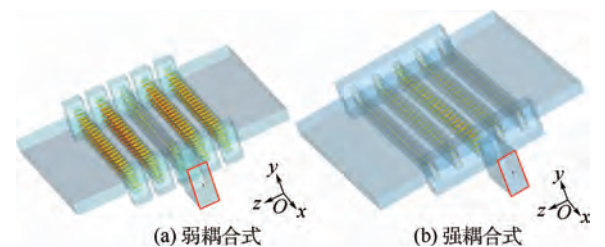


图 5 工作在  $2\pi$  模式的弱耦合式和强耦合式五间隙输出腔  $z$  轴向电场分布图  
Fig. 5  $2\pi$  mode electric fields distribution diagram along  $z$ -axis for weak-couple and strong-couple five-gap output cavity

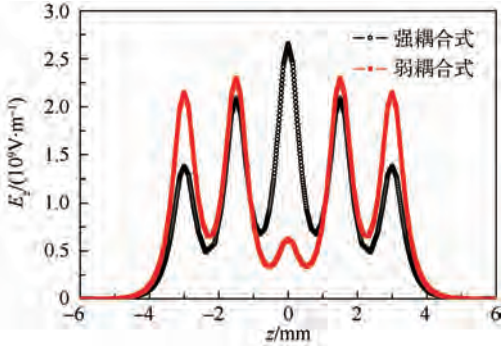


图6 工作在 $2\pi$ 模式的弱耦合式和强耦合式五间隙输出腔沿 $z$ 轴向电场对比图

Fig. 6 Comparison diagram of  $2\pi$  mode electric fields distribution along  $z$ -axis for weak-couple and strong-couple five-gap output cavity

对各个间隙的场强分布影响不大,作为输出腔时能够将各个间隙存储的谐振能量有效提取出去.而弱耦合式五间隙腔在外接耦合口后场形发生了很大改变,其主要原因在于每个间隙只能通过电子注漂移通道进行耦合,输出波导也只能与最中间间隙及谐振腔相连而存在较强的耦合,与不连接的其他几个间隙和谐振腔耦合较弱,从而导致中间间隙处的场强大大降低,而两侧间隙的场强稍稍升高.这也表明,当将弱耦合式五间隙结构作为输出腔时,只有中间间隙的能量能够有效提取,而与之不连通的间隙及其谐振腔能量难以输出.此外,在相同电子注漂移通道下,与弱耦合式输出腔相比,强耦合式结构对各个间隙模式和谐振电场的隔离稍微明显一些,这也能在一定程度上避免各间隙之间注波相互作用过程的互相干扰与影响.

为了比较,表2给出了用CST微波工作室后处理模块进一步分析与计算得到的两种耦合方式输出腔的关键高频特性参数.可以看出,弱耦合式五间隙输出腔与强耦合式五间隙输出腔在外接耦合口和输出波导后,其 $f$ 、 $Q_0$ 和 $R/Q$ 并没有明显的变化和区别,但其外品质因数 $Q_e$ 及有载 $Q$ 值( $Q_L$ )相差甚远.弱耦合式五间隙输出腔的 $Q_e$ 及

表2 弱耦合和强耦合式五间隙输出腔高频参数对比

Table 2 Comparison of RF parameters for weak-couple and strong-couple five-gap output cavity

谐振腔种类	高频参数				
	$f/\text{GHz}$	$\frac{R}{Q}/\Omega$	$Q_0$	$Q_e$	$Q_L$
弱耦合式	94.60	70.43	1358.2	4530.8	1045.0
强耦合式	94.50	71.68	1343.5	694.9	460.4

$Q_L$ 均远远大于强耦合式五间隙输出腔,其 $Q_e$ 甚至高达强耦合式输出腔的6~7倍.研究表明,高的 $Q_e$ 将导致输出腔的能量难以有效提取出去,另一方面 $Q_L$ 值太高也容易在输出腔中激起振荡,使得其稳定输出性能大大下降.

由此可以看出,选择将强耦合式五间隙结构作为SBEIK输出腔时有着良好的优势.首先,它允许电子注漂移管被设计为对工作频率的截止模式,这样将大大增加各个间隙的特性阻抗及其与电子注之间的耦合系数,并进一步提高注波相互作用效率.其次,强耦合式结构通过多个间隙之间的整体耦合谐振结构来实现将整个输出腔能量有效提取到外接波导中去;而对于弱耦合式输出腔,当电子注漂移管对工作模式截止时,其各个间隙之间将无法建立起足够的耦合,从而导致其与输出波导不直接相连的各间隙谐振腔几乎无法将其能量有效耦合出去.此外,采用强耦合式五间隙输出腔结构,将能更好地抑制腔体的非工作模式振荡,更有利于SBEIK注波相互作用过程的稳定性及其高功率提取的有效性.

## 2 强耦合式五间隙腔体的模式竞争

在SBEIK高频系统设计中,除了第1节对 $2\pi$ 工作模式的场形、频率、特征阻抗和 $Q$ 值等关键高频参数重点研究外,也存在其他频率靠近 $2\pi$ 模的非工作竞争模式,从而对工作模式、高频特性的稳定性及其与电子注之间的注波相互作用过程产生影响.对于研究的W波段SBEIK五间隙谐振腔,不仅其每个间隙可能存在着高次谐振模式的竞争,还存在由于其轴向周期性结构所引起的轴向简并模式竞争.相比较而言,每个间隙的高次谐振模式引起的竞争,由于其频率间隔比较大(在设计中,工作模式 $\text{TM}_{110}$ 频率为94.50 GHz,距离它最近的低频段高次模式 $\text{TM}_{520}$ 频率为88.25 GHz;而距离它最近的高频段高次模式 $\text{TM}_{310}$ 频率为97.59 GHz),可以看出其与工作模式 $\text{TM}_{110}$ 之间的频率间隔均在3 GHz以上,如图7给出了通过仿真得到的五间隙腔存在的相邻高次模式所示,因此这些高次模式可以完全控制在工作频带范围外.所以,在此更为关注的问题是由多间隙腔周期性结构引起的轴向简并模式竞争.

对于SBEIK多间隙谐振腔,其周期数越大,引入的轴向简并谐振模式越多,模式竞争问题也更加突出.由第1节可知,作为五间隙谐振腔,弱耦合和强耦合谐振腔高频特性基本一致,在此以强耦合式五间隙谐振腔为重点进行研究与讨论.

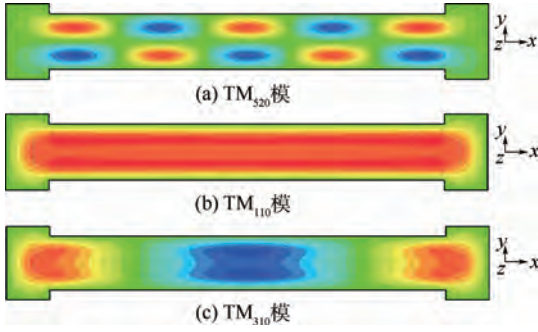


图 7 五间隙腔中各间隙存在的相邻高次模式  
Fig. 7 Existed nearby high modes for each gap of strong-coupling five-gap cavity

对于强耦合式五间隙腔的  $TM_{110}$  模,其周期数为 5,因而存在 5 种轴向简并模式,分别为  $2\pi$  模、 $1/4\pi$  模、 $2/4\pi$  模、 $3/4\pi$  模和  $4/4\pi$  模. 该命名规则是根据其相邻间隙的相位变化来确定的,如  $1/4\pi$  模,每两个相邻间隙的相位相差为  $1/4\pi$  (如第 1 个间隙到第 2 个间隙之间). 而根据 SBEIK 工作特性,需要的工作模式要求每个间隙沿  $z$  方向的电场方向要保持一致并同步变化,由此可知  $2\pi$  模式完全符合此要求. 相比较而言, $1/4\pi$  模第 1、2 间隙和第 3、4 间隙的电场方向相反,并且第 3 个间隙的场强为 0; $2/4\pi$  模第 1、3、5 间隙的电场方向相反,第 2、4 间隙电场为 0; $3/4\pi$  模第 1、2 间隙的电场方向相反,第 3 间隙电场为 0,第 4、5 间隙的电场方向相反; $4/4\pi$  模第 1、3、5 间隙与第 2、4 间隙的电场方向相反,相邻间隙相位相差  $\pi$ . 因此这几种模式均不能满足与电子注与谐振电场同步与互作用的要求,称之为非工作模式.

图 8 给出了利用 CST 微波工作室仿真设计得到的 W 波段 SBEIK 强耦合式五间隙谐振腔中,由周期结构引起的 5 个轴向简并模式在  $xz$  截面的电场场形及其强度分布图. 图 9 给出了各轴向简并模式其轴向电场沿  $z$  轴的电场强度变化. 表 3 给出了相应 5 个模式的  $f$ 、 $Q_0$ 、 $R/Q$  和耦合系数  $M$ ,其中

$$M = \frac{\left| \int_{-\infty}^{\infty} E_z e^{i\beta_z z} dz \right|}{\int_{-\infty}^{\infty} |E_z| dz} \quad (2)$$

式中:  $i = \sqrt{-1}$ ;  $\beta_z$  为电磁波沿  $z$  轴向的传播常数. 沿  $z$  轴向的积分遍历所有间隙及漂移通道区域. 可以看出,通过优化高频结构尺寸,设计的强耦合式五间隙谐振腔其  $2\pi$  工作模式与之最近的  $1/4\pi$  模式之间的频率间隔超过 600 MHz,可以完全满足对 SBEIK 整管设计提出的带宽为 100 MHz 的要求.

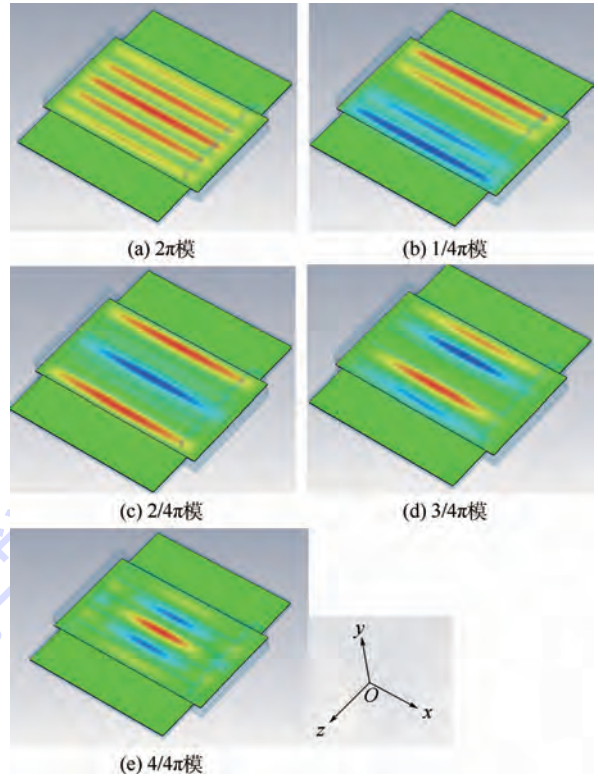


图 8 强耦合式五间隙腔中存在的 5 种轴向简并谐振模式及其沿  $z$  轴向电场强度分布图  
Fig. 8 Existed five types of axial degeneration resonant modes with their electric field distribution for strong-couple five-gap cavity along  $z$ -axis

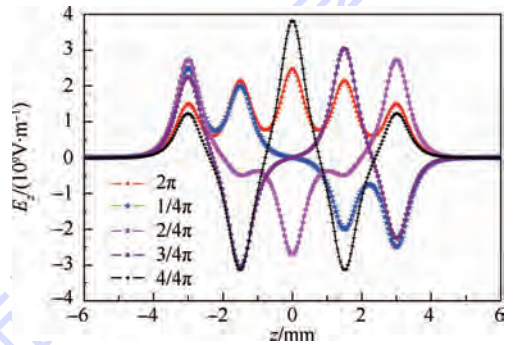


图 9 强耦合式五间隙腔中存在的 5 种轴向简并模式电场沿  $z$  轴分布图  
Fig. 9 Electric field distribution diagram along  $z$ -axis for existed five types of axial degeneration resonant modes in strong-couple five-gap cavity

根据表 3 得到的数据可以做进一步的分析. 对于给定的高频谐振腔系统,其与电子注之间注波相互作用过程中高频电场对电子注的调制幅度正比于  $M^2(R/Q)$ ,这也从一方面反映了电子注与高频场之间的注波相互作用强度. 对于强耦合式五间隙谐振腔,由于其轴向周期结构产生了 5 个轴向简并模式. 相比较而言, $2\pi$  工作模式的  $R/Q$  值远高于其他 4 个简并模式,而其耦合系数也大于其

他几个模式. 由此可知, 选择  $2\pi$  模式与电子注进行相互作用, 将使得谐振腔电场对于电子注的调制幅度远高于其他 4 个简并模式, 这也将更有利于 SBEIK 注波相互作用效率的提高并产生高功率输出.

表 3 五间隙谐振腔中存在的 5 个轴向简并模式高频参数比较

**Table 3 RF parameters comparison for five types of axial degeneration resonant modes in strong-coupling five-gap output cavity**

简并模式	高频参数			
	$f/\text{GHz}$	$Q_0$	$\frac{R}{Q}/\Omega$	$M$
$2\pi$ 模	94.50	1343.5	71.68	0.2715
$1/4\pi$ 模	95.14	1345.1	$1.8 \times 10^{-31}$	0.1217
$2/4\pi$ 模	95.59	1362.3	2.60	0.0527
$3/4\pi$ 模	96.07	1379.1	$3.1 \times 10^{-31}$	0.0205
$4/4\pi$ 模	97.12	1450.5	0.0028	0.0014

### 3 强耦合式五间隙输出腔的粒子模拟与验证

根据对 W 波段 SBEIK 的整管设计, 要求其中中心工作频率 94.50 GHz, 工作带宽 100 MHz 左右. 利用第 1~2 节的强耦合式五间隙输出腔几何参数及其高频特性优化结果, 通过在三维 PIC (Particle In Cell) 粒子模拟软件中建立物理模型并设置输入参数, 采用电压 75 kV、电流 4 A 的理想预群聚带状电子注通过强耦合式五间隙输出腔中, 对这种复杂输出腔的注波相互作用特性进行热模拟仿真, 得到的计算结果如图 10 所示. 图 10(a) 给出了五间隙输出腔电场在  $xy$  截面的场强分布, 可以看出, 这完全是预期的  $\text{TM}_{110}$  模式, 通过输出端口以磁耦合的方式将高频场能量耦合到输出波导; 图 10(b) 给出了电场在  $xz$  截面的分布, 可以看出, 五间隙输出腔沿  $z$  轴向工作在  $2\pi$  模, 完全满足了 SBEIK 模式要求; 由图 10(c) 可以看出粒子经过输出腔时发生了强烈的群聚; 图 10(d) 给出了电子注轴向能量的变化过程, 当电子注通过输出腔时, 将其相当一部分的直流能量转化成了高频能量, 从而导致电子能量降低. 此外, 通过详细的仿真设计与分析, 验证了该强耦合式五间隙输出腔工作性能的稳定性, 频谱纯净, 没有出现振荡或非工作模式的竞争, 完全符合在第 1~2 节中对该输出腔仿真设计所获得的高频特性及其稳定性分析结论.

进一步而言, 为对该强耦合式五间隙输出腔的工作带宽特性进行分析, 利用三维 PIC 粒子模拟软件对不同工作频率时的功率输出特性进行了详细计算, 在前述电子注参数和高频结构下, 得到

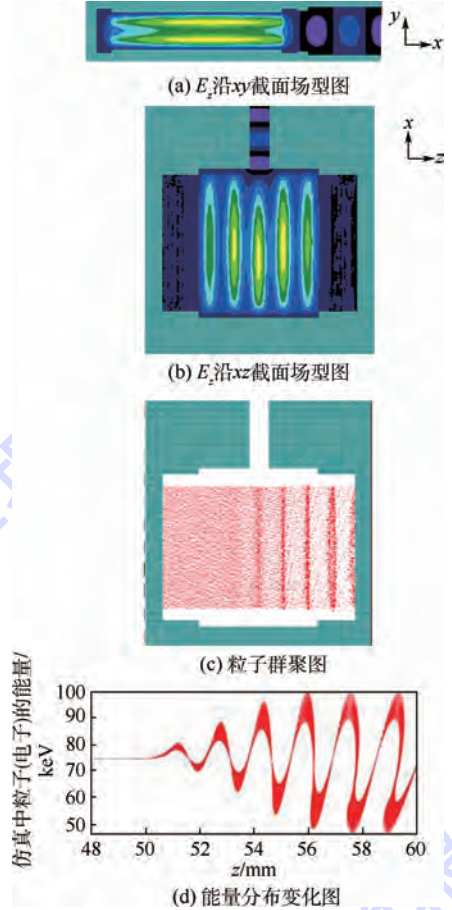
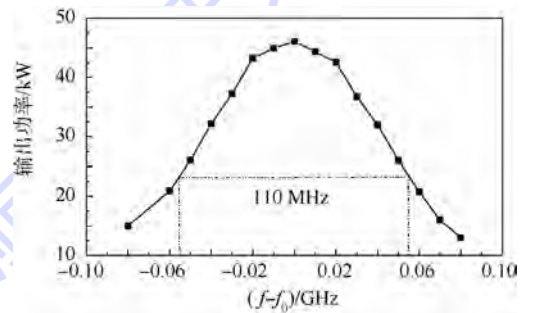


图 10 强耦合式五间隙输出腔注波相互作用过程的三维 PIC 模拟仿真结果

Fig. 10 Simulation result of beam-wave interaction process for strong-couple five-gap output cavity using three dimensional PIC



$f_0$ —仿真设计时的中心频率 ( $f_0 = 94.5 \text{ GHz}$ ).

图 11 三维 PIC 仿真计算得到的输出功率相对于频率带宽图

Fig. 11 Bandwidth diagram calculated by three dimensional PIC simulation and its output power corresponding to frequency

的输出腔频率功率特性曲线如图 11 所示. 可以看出, 计算出的最大输出功率为 46 kW 左右, 由曲线半高宽得到的 3 dB 带宽约为 110 MHz. 同时, 在该工作带宽下, 不会出现高次模式竞争, 以及由于五间隙输出腔周期结构引起的  $1/4\pi$  模、 $2/4\pi$  模、

$3/4\pi$ 模和 $4/4\pi$ 模等非工作模式的竞争,以便更好地保证了该输出腔功率产生与输出稳定性,并为后续工程化设计与研制奠定坚实的研究基础。

## 4 结 论

本文针对 W 波段 SBEIK 设计,对基于平面结构哑铃形谐振腔的弱耦合式和强耦合式五间隙谐振腔及输出腔的高频特性进行了详细研究。结果表明:

1) 当作为谐振腔使用且电子注漂移管对工作模式不截止时,弱耦合式和强耦合式五间隙腔体高频特性基本等效。

2) 当作为输出腔时,相比较而言强耦合式五间隙腔可以完全工作在漂移管对工作模式截止情况下,以进一步提高腔体的耦合阻抗和注波相互作用效率,而且可以很好地将各个间隙的谐振能量有效耦合到输出波导中。因此,在 SBEIK 中采用强耦合式多间隙输出腔是一种更为优化的技术方案。

3) 针对强耦合式五间隙谐振腔,对由于周期结构引起的模式竞争问题进行了详细研究,验证了其相邻非工作模式的频率间隔足以保证 SBEIK 整管的工作带宽。

4) 利用三维 PIC 粒子模拟软件对强耦合式五间隙输出腔的注波相互作用特性进行了进一步设计与验证,在理想预群聚电子注入输出腔时,可以获得 46 kW 左右的功率输出,且工作带宽可以达到 110 MHz 左右,完全能够满足 SBEIK 整管设计要求。

本文分布式五间隙腔体高频特性的研究工作,为亚毫米波和太赫兹频段 SBEIK 更为复杂多间隙腔高频系统的设计、注波相互作用性能的改进与优化,以及输出功率技术优势的进一步挖掘具有非常重要的意义。

### 参考文献 (References)

- [1] Shin Y M, Wang J X, Barnett L R, et al. Particle-in-cell simulation analysis of a multicavity W-band sheet beam klystron [J]. IEEE Transactions on Electron Devices, 2011, 58(1): 251-258.
- [2] Pasour J, Nguyen K T, Wright E L, et al. Demonstration of a 100-kW solenoidally focused sheet electron beam for millimeter-wave amplifiers [J]. IEEE Transactions on Electron Devices, 2011, 58(6): 1792-1797.
- [3] Roitman A, Berry D, Steer B. State-of-the-art W-band extended interaction klystron for the CloudSat program [J]. IEEE Transactions on Electron Devices, 2005, 52(5): 895-898.
- [4] Nguyen K T, Pershing D, Wright E L, et al. Sheet-beam 90 GHz and 220 GHz extend-interaction-klystron designs [C] // 2007 IEEE International Vacuum Electronics Conference (IVEC). Piscataway, NJ: IEEE Press, 2007: 1-2.
- [5] Nguyen K T, Pasour J, Wright E L, et al. Design of a G-band sheet-beam extended-interaction klystron [C] // 2009 IEEE International Vacuum Electronics Conference (IVEC). Piscataway, NJ: IEEE Press, 2009: 298-299.
- [6] Eppley K R, Herrmannsfeldt W B, Miller R H. Design of a Wiggler-focused, sheet beam klystron [C] // Proceedings of the IEEE 1987 Particle Accelerator Conference. Piscataway, NJ: IEEE Press, 1987: 1809-1811.
- [7] Yu D, Wilson P. Sheet-beam klystron RF cavity [C] // Proceedings of the IEEE 1993 Particle Accelerator Conference. Piscataway, NJ: IEEE Press, 1993: 2681-2683.
- [8] Solyga S, Schmolke M, Henke H. Mm-wave sheet beam klystron: Performance at different voltages [C] // Proceedings of the IEEE 1999 Particle Accelerator Conference. Piscataway, NJ: IEEE Press, 1999: 1034-1036.
- [9] Scheitrum G, Caryotakis G, Burke A, et al. W-band sheet beam klystron research at SLAC [C] // IEEE 2006 International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2006, 481-482.
- [10] Garyotakis G. Design of a 11.4 GHz, 150 MW, sheet beam, PPM-focused klystron [C] // AIP Conference Proceedings. Melville, NY: AIP Publishing, 2003, 691: 22-33.
- [11] Scheitrum G, Caryotakis G, Burke A, et al. W-band sheet beam klystron design [C] // Conference Digest of the 2004 Joint 29th International Conference. Piscataway, NJ: IEEE Press, 2004: 525-526.
- [12] Burke A, Besong V, Granlund K, et al. W-band sheet beam klystron PCM focusing design [C] // Proceedings of IVEC/IVESC 2006. Piscataway, NJ: IEEE Press, 2006: 485-486.
- [13] Cusick M, Atkinson J, Balkcum A. X-Band sheet beam klystron (XSBK) [C] // IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2009: 296-297.
- [14] Zhao D, Lu X, Liang Y, et al. Researches on an X-Band sheet beam klystron [J]. IEEE Transactions on Electron Devices, 2014, 61(1): 151-158.
- [15] Shin Y M, Barnett L R, Luhmann N C. Quasi-optical output-cavity design for a 50-kW multicavity W-band sheet-beam klystron [J]. IEEE Transactions on Electron Devices, 2009, 56(12): 3196-3202.
- [16] Pasour J, Wright E, Nguyen K, et al. Sheet beam extended interaction klystron (EIK) in W band [C] // IEEE 2013 International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2013: 1-2.

## RF characteristics of multiple-gap resonant cavity for sheet beam extended interaction klystron

RUAN Cunjun<sup>\*1</sup>, DAI Jun<sup>1</sup>, CHEN Shuyuan<sup>2</sup>, LI Renjie<sup>1</sup>, ZHAN Min<sup>1</sup>

(1. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China;

2. National Key Laboratory of Science and Technology on Test Physics & Numerical mathematical, Beijing 100076, China)

**Abstract:** The sheet beam extended interaction klystron (SBEIK) has the typical and attractive characteristics with the plan structural multiple-gap cavity and extended beam-wave interaction system. A type of novel miniaturization strong-coupling five-gap extended resonant cavity was investigated combined with the design of W-band SBEIK, which was compared with the traditional weak-coupling five-gap resonant and output cavity. The obtained characteristics shows that the strong-coupling five-gap cavity may have many advantages of resonant mode separation, radio frequency (RF) fields coupling in the gap and RF energy exporting, etc. Then, the degeneracy mode competition for the strong-coupling five-gap cavity was studied due to its axial period structures. The result shows that the frequency interval between the work mode and vicinal non-working mode is above 600 MHz, which is higher enough than the required 100 MHz bandwidth for the SBEIK. Finally, the 3D particle in cell (PIC) simulation software was used to calculate the beam-wave interaction for the strong-coupling five-gap extended output cavity, the expected high output power was observed with the corresponding performance and favorable stability. Thus, the research is very important for the physical design and engineering development of the SBEIK in the millimeter-wave and terahertz band with high power output radiation later.

**Key words:** W-band; sheet beam electron; strong-coupling; five-gap resonant cavity; extended interaction klystron (EIK)

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0209

# 周期结构电磁特性在高频真空器件中的应用

冯进军\*, 蔡军, 胡银富, 邬显平

(北京真空电子技术研究所 微波电真空器件国家重点实验室, 北京 100015)



**摘 要:** 作为行波类真空电子器件的核心组件,慢波结构是一种周期结构,其场可以有无限多个模式,每个模式由无穷多个空间谐波构成.每个空间谐波有相应的色散曲线且曲线各段有不同的特性.提出了周期结构色散特性的全维度开发的概念,并以一种可用微电机系统(MEMS)技术加工的折叠波导(FWG)慢波结构为例,对其色散特性进行了分析,利用这些色散特性开展了行波管(TWT)、返波管(BWO)等传统器件的研究工作,同时提出了过模器件、带边振荡器(BO)和谐波放大器(THAT)等新型器件,这些器件的实验研究则以W波段及其以上频率为主,最后给出了突破的关键技术以及测试得到的器件的主要性能.

**关键词:** 周期结构;色散特性;太赫兹;行波管(TWT);返波管(BWO);多维度;真空器件

中图分类号: V221+.3

文献标识码: A      文章编号: 1001-5965(2015)10-1785-07

在行波类真空电子器件中,电子注和电磁波之间产生持续相互作用的一个条件就是电磁波的相速小于光速,通常有两种方法来降低相速,一种是采用介质材料加载的波导,另一种就是采用周期性的金属或者介质结构.微波真空器件通常采用金属周期结构,其传输损耗比较小,可以维持较高的射频电场,还可以接收高频散焦等杂散电子注电流.介质结构(含介质周期结构)的慢电磁波系统使用较少,只用在切伦科夫辐射器件中.介质结构(含介质周期结构)的慢电磁波系统一般适用于光学频段,因为金属结构对光学频段的电磁波传输损耗比较大,非线性效应成为主要矛盾.

真空射频器件采用了周期结构来降低电磁波的相速,如前向波放大管、行波管(Traveling Wave Tube, TWT)、返波管(Backward Wave Oscillator, BWO)、回旋行波管等,其主要的慢波结构有各种

叶片加载结构、螺旋线型、耦合腔型、梳状线、梯形线、叉指线、盘荷波导、螺旋波纹波导等.磁控管是利用周期结构激励出来的电磁波在谐振腔中形成振荡.在高频器件中,如Smith-Purcell器件、斜注管<sup>[1]</sup>等,近年来研究比较多的周期结构有折叠波导(Folded Waveguide, FWG)、双/多排齿结构等一系列与微机电系统(Micro-Electromechanical System, MEMS)技术相兼容的二维结构,利用UV-LIGA或深反应离子刻蚀技术来加工<sup>[2-3]</sup>.

行波类真空电子器件的工作原理可以概括为电子枪组件中的阴极产生一束电子注,由聚焦系统维持一定的电子注形状,通过慢波结构中的电子注通道;高频电磁场通过输入系统进入慢波结构中传输,其沿电子运动方向的相速度与电子速度同步,电磁场通过注波相互作用从电子注中获得能量,被放大的电磁场通过输出系统耦合出来,收

收稿日期: 2015-04-09; 录用日期: 2015-05-08; 网络出版时间: 2015-06-25 16:09

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150625.1609.003.html

基金项目: 国家重点基础研究发展计划(2013CB933602)

\* 通讯作者: 冯进军(1966—),男,山西运城人,研究员, fengjinjun@tsinghua.org.cn, 主要研究方向为真空电子学、微电子技术及毫米波太赫兹电子学等.

**引用格式:** 冯进军, 蔡军, 胡银富, 等. 周期结构电磁特性在高频真空器件中的应用[J]. 北京航空航天大学学报, 2015, 41(10): 1785-1791. Feng J J, Cai J, Hu Y F, et al. Application of dispersion characteristics of periodic structures for high frequency vacuum devices[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1785-1791 (in Chinese).

集极则用来收集通过慢波结构后已经交出能量的电子。

作为真空电子器件的核心组件,慢波结构对器件的性能起到决定性作用,也在线性同步辐射光源等大设备中发挥了重要作用,因此被称为20世纪的英雄<sup>[4]</sup>。慢波结构和目前基于光学以及材料学的“光子带隙<sup>[5]</sup>”和“超材料<sup>[6]</sup>”概念一脉相承,对其进一步开发和拓展将可以在W波段及太赫兹等高频产生特点鲜明的新型真空电子器件。本文首先分析了周期结构的电磁特性,然后以一种适合于高频真空功率器件的全金属折叠波导慢波结构为例,分析了其作为周期结构的电磁特性,提出了色散曲线全维度开发的概念,即根据色散特性的不同开发不同的真空功率器件,如利用其中的前向波特性和返波特性和带边特性等研制行波管、返波管和带边振荡器等,最后给出了利用这些特性在W波段及以上频率研制的器件及其性能。

## 1 周期结构的电磁特性

慢波结构是一种周期结构,可以是一组周期结构加载的波导,也可以是一系列相互耦合的谐振腔。周期结构通过多点反射延缓了波的传输,最终形成了无限多个幅度比例一定的空间谐波。根据Floquet定理,无耗传输线上的电场可以表示为一系列空间谐波和的形式:

$$E_z = e^{-j\beta z} E_p(z) = \sum_{m=-\infty}^{\infty} E_{zm}(x, y) e^{-j\beta_m z} \quad (1)$$

$$\beta_m = \beta_0 + \frac{2\pi m}{p}$$

式中: $z$ 为纵向距离; $E_z$ 为 $z$ 轴方向的电场强度分量; $E_p(z)$ 为周期性函数,在不同的周期,函数只存在相位变化; $m$ 为空间谐波次数; $p$ 为周期长度; $E_m(x, y)$ 为 $m$ 次空间谐波的电场强度, $E_{zm}(x, y)$ 为其在 $z$ 轴的分量; $e$ 为自然对数的底数; $j$ 为虚数的单位; $\beta$ 为电磁波的传播常数; $\beta_m$ 和 $\beta_0$ 分别为 $m$ 次和0次空间谐波的传播常数。电磁波的相速度与频率的关系称为色散特性,空间谐波中只有一小部分的相速大于或等于光速,大部分都具有低于光速的相速度,因而可能与电子注处于同步,从而发生注波相互作用,放大电磁波。为了描述电子注和电磁波的相互作用强弱程度,引入耦合阻抗 $K$ 的概念<sup>[7]</sup>:

$$K = \frac{E_{zm}}{2\beta_m^2 P_{rf}} \quad (2)$$

式中: $E_{zm}$ 为 $m$ 次空间谐波电场纵向分量的幅值;

$P_{rf}$ 为系统中所有空间谐波的总功率流。周期结构中存在许多通带,通带之间是止带。每一个通带对应一种模式;每个模式都是由各个空间谐波叠加而成的沿传播方向呈非正弦分布的行波。也就是说,周期结构中 $z$ 方向非正弦分布的场由一系列 $z$ 方向正弦分布的行波场,即空间谐波场叠加而成。真空电子器件通常采用闭合周期结构而不是开敞结构。

随着工作频率的升高,最常用的螺旋线和耦合腔慢波结构的加工变得非常困难,结构尺寸已达到几十 $\mu\text{m}$ 至几百 $\mu\text{m}$ 的量级<sup>[8]</sup>,而且高频工作时螺旋线慢波结构的功率容量较低,散热也存在很大的困难。在太赫兹频段,微加工技术成为一种理想的工艺技术,而传统的螺旋线及其变态结构都不适合采用微加工技术。此时,折叠波导慢波结构展现出了它的价值。折叠波导是一种周期性波导,它是将矩形波导沿电场面( $E$ 面)像曲折线一样弯曲形成的一种慢波结构,如图1所示。

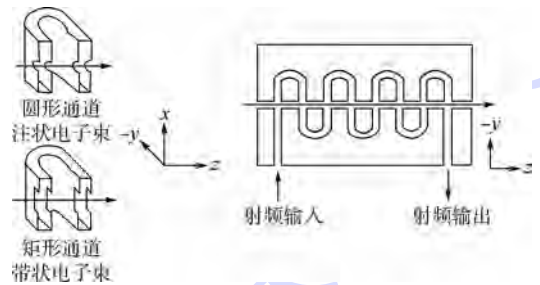


图1 FWG结构示意图

Fig. 1 Structure diagram of FWG

折叠波导慢波结构主要具有机械强度高、散热好、功率容量大、易于加工以及输入/输出耦合结构简单等优点。通过引入最先进的微加工技术,以折叠波导慢波结构为核心的微型真空电子器件很有潜力成为工作在毫米波甚至太赫兹频段的一种大功率、小型辐射源。折叠波导的色散特性可以通过3种方法进行分析:理论公式计算、等效电路法和软件模拟,虽然前2种方法计算速度更快,但是在分析折叠波导的止带和过模等方面时适用性差,因此利用CST微波工作室软件进行冷特性的精确仿真真是得到折叠波导慢波结构色散曲线的一种最重要的分析方法。图2为利用CST微波工作室计算的某折叠波导慢波结构的色散曲线示意图,其中字母F为前向波,此时相速和群速的方向一致。反之,字母B为返波。字母前的数字为空间谐波的次数,如式(1)所示, $m=0$ 时为基波, $m$ 为其他整数时为 $m$ 次空间谐波。虚线框内的标

识 1 为折叠波导行波管利用的  $-2$  次空间谐波(前向波);点划线区域中的标识 2 为折叠波导返波振荡器利用的  $+2$  次空间谐波(返波);实线区域中的标识 3 为折叠波导慢波结构的第一止带,标识 4 为非线性的二次时间谐波,标识 5 为折叠波导慢波结构的高次模式,即  $TE_{20}$  模式. 为了实现周期结构电磁特性在真空射频器件中的全维度应用,一方面是利用常规色散区域进行高频率行波管和返波振荡器的研制,另一方面是开拓色散的新区域,开展新器件的研究及实验验证.

22 kV 和 180 mA 时,瞬时带宽达到 10 GHz,脉冲输出功率大于 100 W,全频带小信号增益大于 40 dB;工作比为 1%,流通率大于 96%<sup>[17]</sup>.

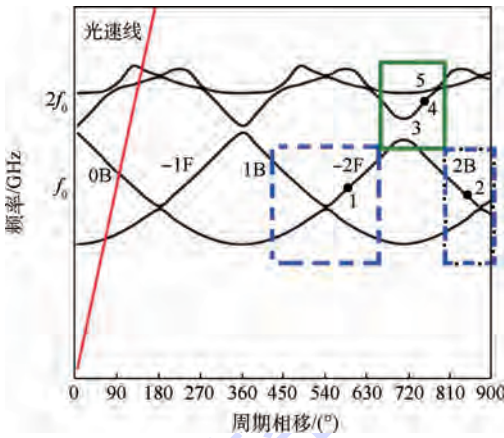


图 2 利用 CST 微波工作室软件计算得到的 W 波段 FWG 慢波结构的色散曲线  
Fig. 2 Dispersion curve of W-band FWG structure using CST microwave studio

## 2 高频率的开发和应用

微波电真空器件国家级重点实验室对 W 波段行波管开展了 10 年的研究,通过自主研发突破了多项关键技术<sup>[9-11]</sup>,目前,W 波段连续波和脉冲行波管的峰值/平均功率、带宽、电子效率等技术指标均达到国际先进水平<sup>[12-15]</sup>. 同时,实验室也开展了 W 波段返波振荡器以及太赫兹返波振荡器的研究,以研制实用化的高频率真空辐射源.

### 2.1 W 波段连续波行波管

W 波段连续波行波管功率及饱和增益测试曲线如图 3 所示. 结果显示,研制的两支行波管在 10 GHz 带宽内功率大于 30 W,最大功率大于 40 W,饱和增益大于 30 dB,工作电压和工作电流分别为 16 kV 和 80 mA,流通率大于 98.5%<sup>[16]</sup>.

### 2.2 W 波段脉冲行波管

W 波段脉冲行波管脉冲功率测试曲线如图 4 所示. 结果显示,在工作电压和工作电流分别为

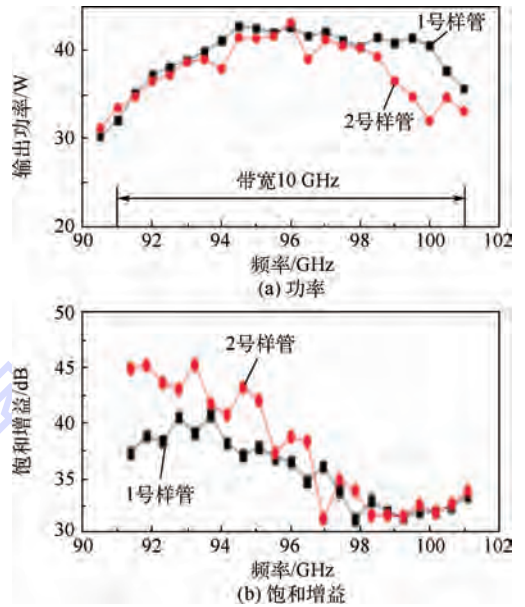


图 3 W 波段连续波 TWT 功率及饱和增益测试曲线<sup>[16]</sup>  
Fig. 3 W-band continuous wave TWT power and saturation gain tested curves<sup>[16]</sup>

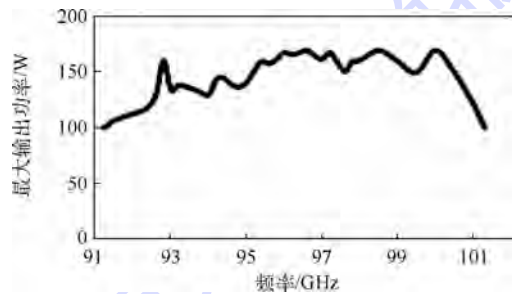


图 4 W 波段脉冲 TWT 脉冲功率测试曲线<sup>[17]</sup>  
Fig. 4 Power tested curve of W-band pulsed TWT<sup>[17]</sup>

### 2.3 G 波段行波管

实验室正在研制 G 波段折叠波导行波管,目前已经完成了电子光学系统及注波相互作用模拟和设计,图 5 为利用 Opera 和 CST 粒子工作室模拟 G 波段折叠波导行波管的电子轨迹和输出功率,结果显示 G 波段折叠波导行波管可以实现 99% 以上的流通率,10 W 输出功率的带宽大于 10 GHz<sup>[18]</sup>.

### 2.4 W 波段返波振荡器

利用折叠波导慢波结构中的返波特性,研制了 W 波段返波振荡器,完成了平面阴极发射实验,最大电流密度可达 44.23 A/cm<sup>2</sup>. 利用有效长度为 80 mm 的均匀永磁聚焦系统,流通管的流通率大于 50%. 图 6 为研制的 W 波段折叠波导返波管样管.

大功率、小型化的慢波结构带边振荡器,利用止带附近高耦合阻抗的特点实现强注波互作用,获得更高的功率.

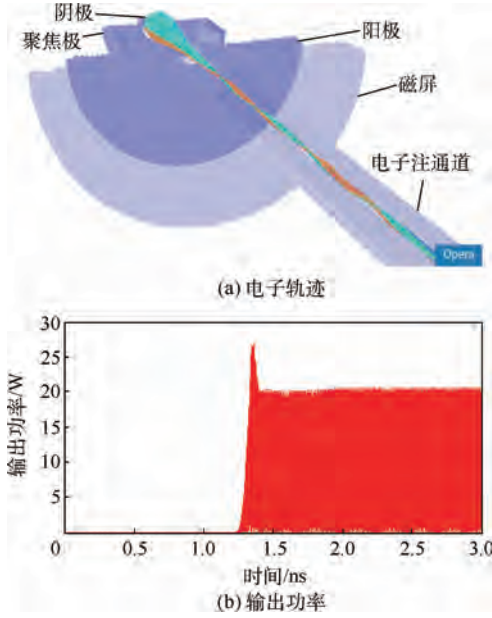


图5 G波段FWG TWT的电子轨迹和输出功率模拟结果

Fig.5 Electron optics and output power simulation results of G-band FWG TWT



图6 W波段FWG BWO样管

Fig.6 Prototype of W-Band FWG BWO

### 3 色散新区域的开拓研究及实验验证

#### 3.1 太赫兹折叠波导带边振荡器

为了实现太赫兹新频段的开拓,满足太赫兹应用对实用化功率源的需求,研发了太赫兹折叠波导慢波结构带边振荡器.器件工作在慢波结构的止带附近,利用高耦合阻抗的特点实现强注波互作用,可以作为大功率、小尺寸的太赫兹源.

折叠波导慢波结构中的电磁波在波导弯曲和电子注通道的位置由于存在不匹配会产生反射,这些反射虽然很小,但是不能忽略,多个反射相互叠加,导致折叠波导慢波结构产生了止带.图7为带边振荡器的色散和耦合阻抗.对于同样的折叠波导慢波结构几何尺寸,以行波管的方式如果工作在W波段,对应第一止带的振荡器可以工作在更高的F波段,如图7(a)和图7(b)所示,因此,在器件尺寸相同的条件下可以获得更高频率的电磁波.图7(b)中为区分不同周期相移分为A、B、C和D 4段曲线,图7(c)为图7(b)中C段与D段曲线对应的耦合阻抗.如图7(c)所示,止带附近耦合阻抗值远大于行波管工作频率附近的耦合阻抗值.因此,基于带边振荡,提出了这种高频率、

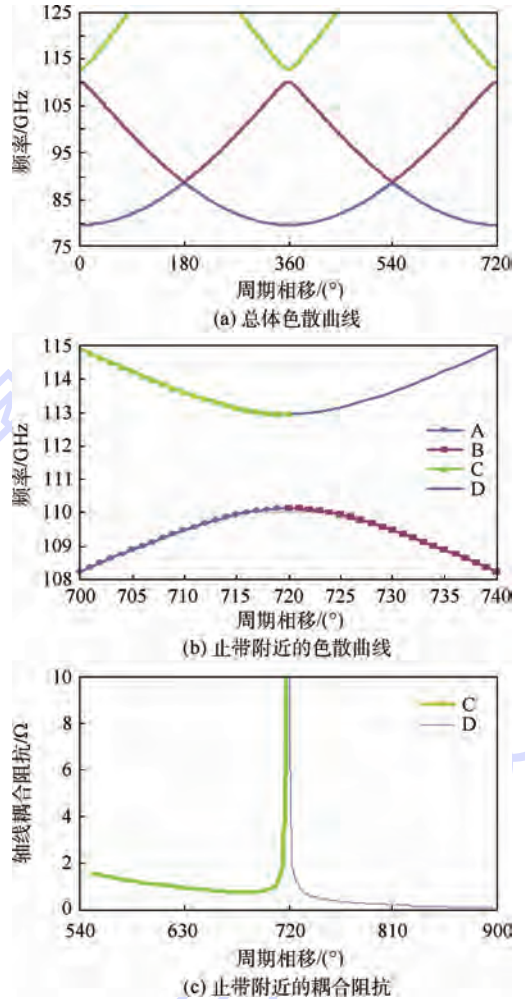


图7 带边振荡器的色散和耦合阻抗

Fig.7 Dispersion and impedance of band-edge oscillator

研制了D波段折叠波导带边振荡器作为概念验证管,其工作电压为23.1 kV,工作电流为150 mA,测试得到其振荡频率为124.45 GHz,最大脉冲输出功率达到32 W.图8为太赫兹折叠波导慢波结构带边振荡器的实物照片,图9为折叠波导慢波结构带边振荡器的功率和频率测试曲线.

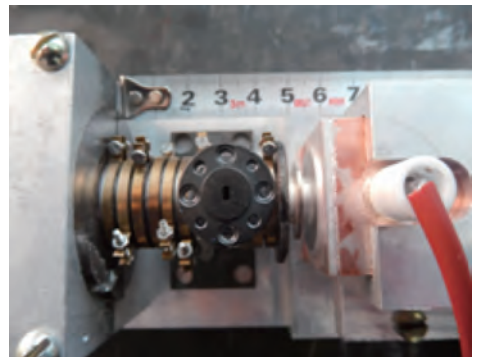


图8 太赫兹带边振荡器照片

Fig.8 Photo of Terahertz band-edge oscillator

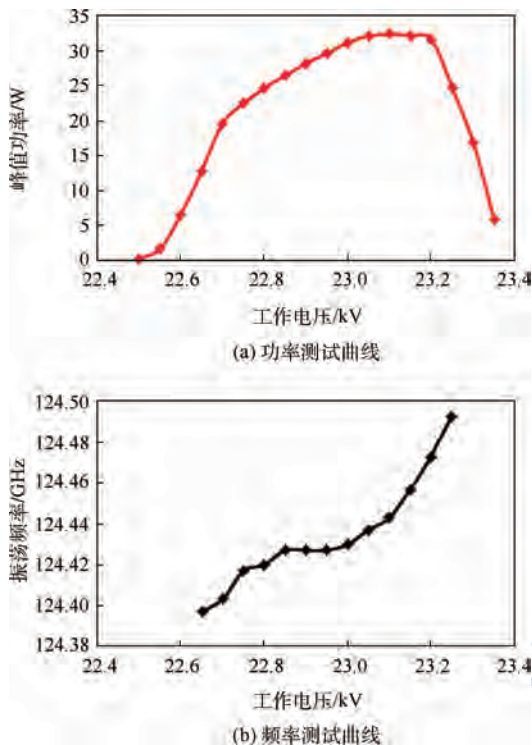


图 9 FWG 慢波结构带边振荡器功率和频率测试曲线  
Fig. 9 Tested power and frequency-tuning curve of FWG band-edge oscillator

### 3.2 太赫兹行波管谐波放大器

为了满足高性能、实用化的太赫兹源需求,提出了太赫兹行波管谐波放大器 (TWT Harmonic Amplifier in THz, THAT) 的概念. 放大器立足于对传统行波管相互作用后电子注中谐波信息的再利用,获得谐波产生、选择、放大及耦合. 为了验证这一概念,通过 W 波段折叠波导脉冲行波管以及级联的二次谐波提取系统,研制了 G 波段 (170 ~ 260 GHz) 行波管二次谐波放大器作为概念验证管. 实验测试显示,在 11.4 GHz 的频带范围内,二次谐波的输出功率大于 100 mW,最大功率达到 500 mW. 本项工作证实了器件概念及可行性,根据需求可以进一步开展更高频率、更大功率的太赫兹源的研发.

图 10 为太赫兹折叠波导行波管谐波放大器的照片,图 11 为太赫兹行波管谐波放大器的功率和增益测试曲线.



图 10 太赫兹 FWG 结构的 THAT 照片  
Fig. 10 Photo of THAT using FWG structure

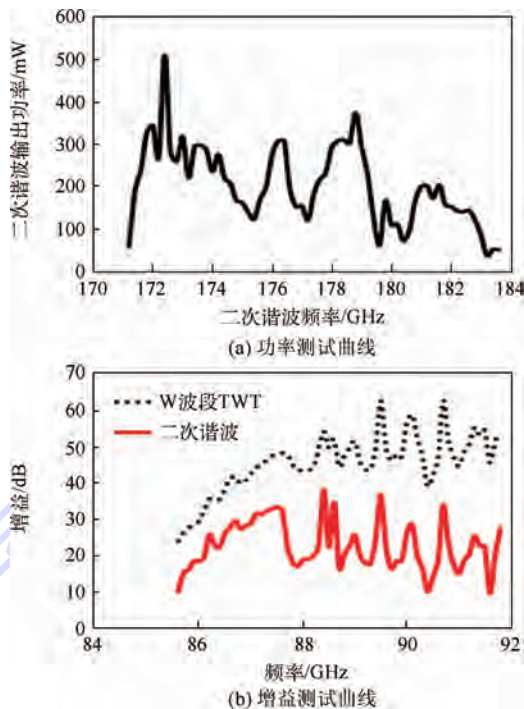


图 11 THAT 功率和增益测试曲线  
Fig. 11 Tested power and gain curve of THAT

### 3.3 过模行波管

真空器件随着频率的提高能够达到的平均输出功率呈指数下降,采用过模相互作用电路,是解决传统的工作在基模的行波管到更高频率后遇到的尺寸限制的一种有效方法. 所谓过模相互作用电路是指高频电路在工作频率下有不只一个传输模式,电子注与多个模式同时进行相互作用. 针对过模相互作用电路中工作模式的选择和非工作模式的抑制两项关键技术,开展了设计和模拟研究,结果显示过模行波管在 W 波段可实现 kW 级的功率输出. 图 12 为利用 CST 粒子工作室建立的 W 波段过模行波管模型,以及其中单个几何周期的结构.

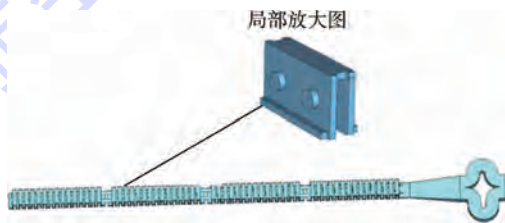


图 12 过模 FWG TWT 模型  
Fig. 12 Solid model of overmode FWG TWT

图 13 为过模折叠波导慢波结构完整的布里渊曲线图,对曲线中所有的分支进行了模式甄别,图 13 中主要列出了  $TE_{10}$ 、 $TE_{20}$  以及  $TE_{30}$  3 个模式. 常规的基模折叠波导行波管利用的是  $TE_{10}$  模式,该过模行波管利用的是  $TE_{20}$  模式. 25 kV 为 25 kV 电子注所在位移.

为了能够使所设计的  $TE_{20}$  模式折叠波导行

波管正常工作,设计了一个  $TE_{10}$  到  $TE_{20}$  的模式变换器,模拟结果表明  $TE_{10}$  到  $TE_{20}$  的模式转换效率能够达到 99% 以上。

整管粒子仿真得到的放大输出信号如图 14(a) 所示,可以看出输出功率达到 800 W 以上。对输出信号进行幅度傅里叶分析,信号频谱如图 14(b) 所示,输出频谱中只有放大的 94 GHz 信号<sup>[19]</sup>。

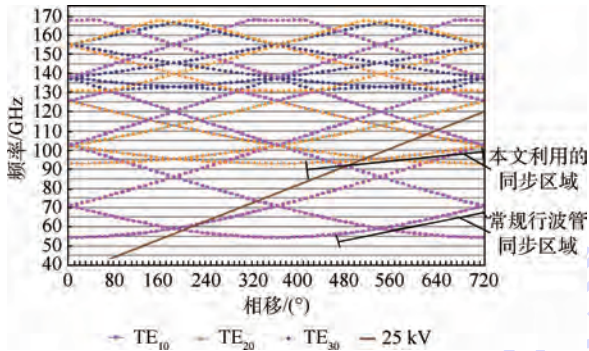


图 13 过模 FWG 慢波结构完整的布里渊曲线  
Fig. 13 Complete Brillouin curve of overmode FWG slow wave structure

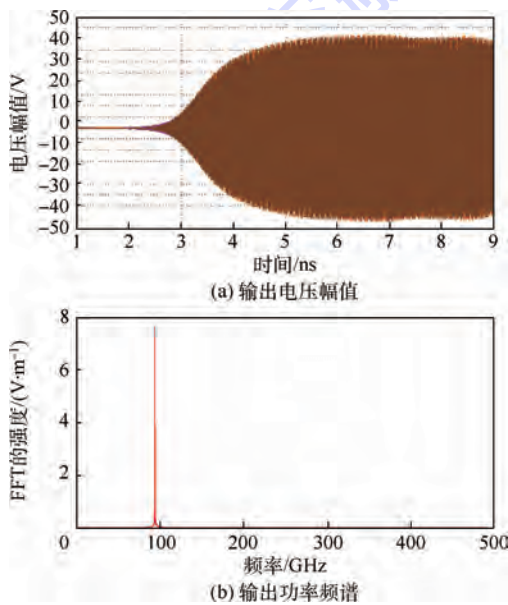


图 14 过模 FWG TWT 输出电压幅值和输出功率频谱  
Fig. 14 Output voltage amplitude and output power spectrum of overmode FWG TWT

## 4 结 论

1) 周期结构慢波系统的色散特性比较复杂,其中具有多个通带,也存在止带;各次空间谐波中既有返波,也有前向波。

2) 利用周期结构的全维度电磁特性,以及与 MEMS 技术相兼容的折叠波导慢波结构,在 W 波段及太赫兹领域开发了多种器件,如行波管、返波

管等传统器件,以及带边振荡器、谐波放大器、过模行波管等新型器件。

3) 下一步将要探索新型场发射阴极在高频率器件中的应用,并进一步提高器件的功率、效率、寿命和可靠性等性能,实现工程应用。

## 参考文献 (References)

- [1] 任大鹏,冯进军. 太赫兹真空源斜注管高频结构研究[J]. 微波学报,2010(S1):551-553.  
Ren D P, Feng J J. Study on slow wave structure of the THz clinotron[J]. Journal of Microwaves, 2010(S1):551-553 (in Chinese).
- [2] 冯进军. 集成真空电子学[J]. 真空电子技术,2010(2):1-7.  
Feng J J. Integrated vacuum electronics[J]. Vacuum Electronics, 2010(2):1-7 (in Chinese).
- [3] 冯进军,廖复疆,朱敏,等. 微型真空电子器件技术研究[J]. 真空电子技术,2005(6):8-16.  
Feng J J, Liao F J, Zhu M, et al. Study on micro vacuum electron devices technology[J]. Vacuum Electronics, 2005(6):8-16 (in Chinese).
- [4] Shvets G. Applications of electromagnetic metamaterials to vacuum electronics devices and advanced accelerators[C]//Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2014:14466902.
- [5] 杨乐,廖复疆. 高阻抗电磁表面及其在真空电子器件中的应用[J]. 真空电子技术,2008(5):29-31.  
Yang L, Liao F J. High impedance electromagnetic surface and application in vacuum electron devices[J]. Vacuum Electronics, 2008(5):29-31 (in Chinese).
- [6] Caloz C, Itoh T. Metamaterials for high-frequency electronics[J]. Proceedings of the IEEE, 2005, 93(10):1744-1752.
- [7] 张克潜,李德杰. 微波与光电子学中的电磁理论[M]. 北京:电子工业出版社,1994:403.  
Zhang K Q, Li D J. Electromagnetic theory for microwaves and optoelectronics[M]. Beijing: Publishing House of Electronics Industry, 1994:403 (in Chinese).
- [8] 冯进军,唐焯,李含雁,等. 短毫米波和太赫兹线性注真空器件研究[J]. 真空电子技术,2013(1):1-9.  
Feng J J, Tang Y, Li H Y, et al. Study on linear beam vacuum electron devices in sub-millimeter wave and THz[J]. Vacuum Electronics, 2013(1):1-9 (in Chinese).
- [9] Cai J, Feng J J, Wu X P, et al. Analysis and test preparation of attenuator for W-band FWG TWT[C]//Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2007:9855417.
- [10] Hu Y F, Feng J J. 3D simulation of electron gun and beam transport for W-band TWT[C]//Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2007:9855438.
- [11] Hu Y F, Feng J J, Cai J, et al. A broadband microwave window for W-band TWT[C]//Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2008:376-377.
- [12] Feng J J, Cai J, Wu X P, et al. Design investigation of 10 W W-

- band FWG TWT[C] // Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2007: 9855428.
- [13] Hu Y F, Feng J J, Cai J, et al. Performance enhancement of W-band CW TWT[C] // Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2011: 21-22.
- [14] Hu Y F, Feng J J, Cai J, et al. Development of W-band CW TWT amplifier[C] // Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2012: 295-296.
- [15] Feng J J, Cai J, Hu Y F. Development of W-band FWG pulsed TWTs[J]. IEEE Transactions on Electron Devices, 2014, 61(6): 1721-1725.
- [16] Hu Y F, Feng J J, Cai J, et al. Design and experimental study of wide bandwidth W-band FWG continuous-wave TWT[J]. IEEE Transactions on Plasma Science, 2014, 42(10): 3380-3386.
- [17] Cai J, Feng J J, Hu Y F, et al. 10 GHz bandwidth 100 watt W-band FWG pulsed TWTs[J]. IEEE Microwave and Wireless Components Letters, 2014, 24(9): 620-621.
- [18] Pan P, Hu Y F, Liu J K, et al. Preliminary design of a 220 GHz FWG TWT[C] // Proceedings of IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2014: 249-250.
- [19] 胡银富,冯进军,蔡军,等. W 波段大功率梯形线行波管的设计与仿真[J]. 真空科学与技术学报, 2014, 34(3): 294-299.
- Hu Y F, Feng J J, Cai J, et al. Design and simulation of W-band high power ladder TWT[J]. Chinese Journal of Vacuum Science and Technology, 2014, 34(3): 294-299 (in Chinese).

## Application of dispersion characteristics of periodic structures for high frequency vacuum devices

FENG Jinjun\*, CAI Jun, HU Yinfu, WU Xianping

(National Key Laboratory of Science and Technology on Vacuum Electronics, Beijing Vacuum Electronics Research Institute, Beijing 100015, China)

**Abstract:** As the key component of vacuum electron device (VED), slow wave structure is a kind of periodic structure, whose dispersion properties can be presented as the sum of an infinite number of space harmonic wave. Each space harmonic wave exits as a dispersion curve with its characteristic. We proposed the idea of the full use of dispersion properties of periodic structures for high frequency vacuum power devices, and carried out the research on folded waveguide (FWG) slow wave structure, which was comparable with MEMS technology. The dispersion curve was analyzed, and was used for different kind of traditional devices including travelling wave tubes (TWTs), backward wave oscillator (BWO) and band-edge oscillators (BO). In addition, over-mode TWTs and TWT harmonic amplifier in THz (THAT) were proposed as novel devices. All the experimental investigation was focused on W-band and frequencies above, and the key technologies and the specifications obtained from these devices were given.

**Key words:** periodic structure; dispersion properties; terahertz; travelling wave tube (TWT); backward wave oscillator (BWO); full use; vacuum devices

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0212

# 接收模式下加罩天线系统特性的快速电磁仿真



何芒\*, 刘金波, 王斌斌, 徐晓文, 章传芳, 孙厚军

(北京理工大学 信息与电子学院, 北京 100081)

**摘 要:** 针对接收模式下电大尺寸的天线-天线罩系统电磁特性分析问题, 提出了一种全波法和高频法结合的混合算法: 积分方程/修正型表面积分 + 多层快速多极子 (IE/MSI + MLFMA) 方法. 该方法将由高频方法 MSI 得到的天线罩内场分布作为天线阵列的激励场, 再利用基于体-面混合积分方程 (VSIE) 和 MLFMA 的快速全波分析方法精确计算天线阵; 同时应用球谐函数展开、预处理技术、混合并行等技术进一步改进算法的计算效率. 同传统的全波数值法相比, 该方法在保证计算精度、满足工程需要的前提下, 解决了求解计算时间长、计算效率低的问题, 实现了对电大尺寸天线-天线罩系统的快速、高效仿真.

**关键词:** 加罩天线系统; 体-面混合积分方程 (VSIE); 表面积分; 多层快速多极子算法 (MLFMA); 球谐函数展开; 并行计算

**中图分类号:** V219; TN958.92

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1792-09

现代天线系统中, 天线罩在保护天线免受外界环境干扰的同时, 也会影响到天线本身的电性能, 如主瓣展宽、波束指向偏移、副瓣升高等<sup>[1]</sup>. 因此, 对带罩天线系统电磁特性分析的研究, 不仅具有较高的理论价值, 更具有重要的现实意义.

对于电中小尺寸的天线-天线罩系统, 全波数值法可以精确高效地得到仿真结果. 但是, 工作于微波尤其是毫米波段的天线罩通常具有很大的电尺寸, 随着系统规模的增加, 待求未知量也急剧增多, 现有的全波数值分析方法对这种结构电磁特性的分析往往需要消耗大量的内存和计算时间, 即便使用加速算法如多层快速多极子 (Multilevel Fast Multipole Algorithm, MLFMA)<sup>[2]</sup> 等方法降低了计算复杂度, 但计算效率仍然不高. 而且, 已有文献大多研究发射模式下天线罩对天线特性的影响, 而极少有文献讨论接收模式下电大尺寸天线-天线罩系统的电性能分析, 而后者在工程中应用极多, 如带罩

相控阵的校准测量一般都是在接收模式下进行. 因此, 接收模式下带罩天线系统电磁特性的有效分析是计算电磁学领域一个颇具挑战性的研究课题.

针对这一问题, 本文结合全波快速分析法与高频射线算法各自的优势, 提出了一种基于多层快速多极子加速的 IE/MSI (Integral Equation/Modified Surface Integration) + MLFMA 混合算法, 在保证计算精度满足工程需要的情况下, 消除了传统方法求解此类问题时存在的计算时间长、效率低下的弊端. 本文详细地论述了该方法的基本理论和技术细节, 给出的数值算例证明了其高效性与精确性.

## 1 理论基础

被动接收模式下的天线-天线罩系统模型如图 1 所示, 模型可等效为入射平面波照射下  $N$  层介质天线罩加载的天线阵列中等效电流分布的求

收稿日期: 2015-04-10; 录用日期: 2015-04-24; 网络出版时间: 2015-06-24 16:46

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150624.1646.003.html

基金项目: 国家自然科学基金 (61471040)

\* 通讯作者: 何芒 (1976—), 男, 安徽安庆人, 副教授, hemang@bit.edu.cn, 主要研究方向为天线理论与技术、计算电磁学.

**引用格式:** 何芒, 刘金波, 王斌斌, 等. 接收模式下加罩天线系统特性的快速电磁仿真[J]. 北京航空航天大学学报, 2015, 41(10): 1792-1800. He M, Liu J B, Wang B B, et al. Fast electromagnetic simulation of characteristics of radome-antenna system in receiving mode [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1792-1800 (in Chinese).

解.如引言所述,本文提出了一种新的全波分析与高频近似方法相结合的混合算法.整个算法可分为 2 个步骤:①由基于 MSI 的高频方法得到天线罩内的场分布  $\mathbf{E}$  和  $\mathbf{H}$ ;②再将此场分布作为天线阵列的激励场,利用基于体-面混合积分方程 (Volume Surface Integral Equation, VSIE) 和 MLFMA 的快速全波分析方法精确计算天线阵列的电流分布,便可得到所有感兴趣的电磁参数.本文称此算法为 IE/MSI 方法<sup>[3]</sup>.与纯粹的全波数值法不同,在此方法中只有天线阵列部分采用全波求解,而天线罩部分则只用高频射线法计算,因此该方法极大地减少了内存用量.

上述过程中的第②步可以按照已有的全波方法精确计算出阵列中的电流分布,问题在于第①步中天线罩内场分布  $\mathbf{E}, \mathbf{H}$  的计算.传统的口径积分-表面积分法 (Aperture-Integration Surface-Integration, AI-SI)<sup>[4]</sup>,只能得到处于发射工作模式下天线加罩后的辐射方向图等信息,但由于等效原理的失效,这种方法不能计算天线工作于接收状态下的场分布.为此,本文将天线罩的底面  $S_0$  向外做任意无限延伸,形成一个带有空洞结构的无限大虚拟面  $S_0'$  (图 1 中虚线部分).此虚拟面  $S_0'$  和天线罩的内表面  $S_1$ 、天线罩的底面  $S_0$  一起形成了一个完整的无限大曲面,此曲面将整个三维空间分成了 2 个无限大半空间,将其中的下半空间视为天线罩内场区域.那么,根据面等效原理,天线罩内场区域的电磁场  $\mathbf{E}$  和  $\mathbf{H}$  可以由 3 个不同曲面上的等效电磁流作表面积分得到.同时,假定由非实际结构虚拟面  $S_0'$  上的等效电磁流产生的部分场,不随天线罩实际结构的改变而发生变化.这样,2 次使用高频近似方法 SI 就可得到实际天线罩结构下,天线罩内场区域中的电磁场分布  $\mathbf{E}$  和  $\mathbf{H}$ .然后,再以此场分布作为天线阵列的激励源,计算出阵列中等效电流分布.

在天线阵列中,设  $S$  为其中任意形状的理想导体表面,  $V$  为任意形状且介电常数为  $\epsilon$  的非磁性介质体,那么若入射电场为  $\mathbf{E}^i$ ,则电场在天线阵列的导体表面上和介质体中应分别满足积分方程:

$$\begin{cases} \mathbf{n} \times (\mathbf{E}^i(\mathbf{r}) + \mathbf{E}^s(\mathbf{r})) = \mathbf{0} & \mathbf{r} \in S \\ \frac{\mathbf{D}(\mathbf{r})}{\epsilon(\mathbf{r})} = \mathbf{E}^i(\mathbf{r}) + \mathbf{E}^s(\mathbf{r}) & \mathbf{r} \in V \end{cases} \quad (1)$$

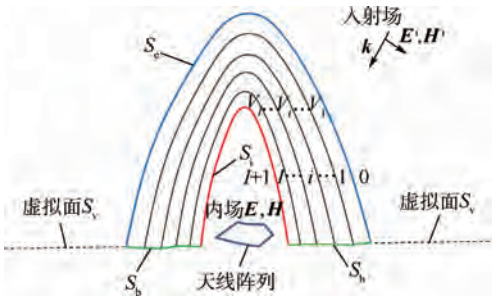
式中:  $\mathbf{E}^s$  为散射电场;  $\mathbf{D}$  为介质体内的电位移矢量;  $\mathbf{r}$  为坐标矢量;  $\mathbf{n}$  为表面单位法向矢量.式(1)两式联立,即构成了基于电场的 VSIE.其中,  $\mathbf{E}^s$  由导体表面的等效面电流与介质体内的等效体电流所产生的部分场叠加而成:

$$\begin{aligned} \mathbf{E}^s(\mathbf{r}) = & -j\omega\mathbf{A}_s(\mathbf{r}) - \nabla\Phi_s(\mathbf{r}) - j\omega\mathbf{A}_v(\mathbf{r}) - \nabla\Phi_v(\mathbf{r}) \end{aligned} \quad (2)$$

式中:  $\mathbf{A}_s$  和  $\Phi_s$  分别为面电流产生的磁矢位和电标位;  $\mathbf{A}_v$  和  $\Phi_v$  分别为体电流产生的相应位函数;  $j = \sqrt{-1}$ ;  $\omega$  为角频率.

利用 MLFMA,可将式(2)的 VSIE 计算模型的计算复杂度降至  $O(N \log N)$  量级,其中  $N$  为未知量个数.但是同时也注意到,MSI 计算天线罩内区域透射场过程的计算复杂度为  $O(MN)$  量级,其中  $M$  和  $N$  分别为天线罩内壁三角形和天线罩内场区域采样点的个数.这就意味着 MSI 将消耗混合算法的大部分时间,尤其是在电大尺寸罩复杂天线阵列的情况下,罩内场计算将占据整个算法 90% 以上的计算时间.这就导致在某些情况下,IE/MSI 的计算时间甚至比全波算法的计算时间更长.

为了加速 MSI 的计算,进一步提高计算效率,将整个天线-天线罩系统用 2 套 MLFMA 树形结构覆盖,如图 2 所示.其中,树形结构 1 覆盖整



$\mathbf{E}^i, \mathbf{H}^i$ —入射电磁场;  $S_0$ —天线罩外表面;  $\mathbf{k}$ —波矢量;  
 $i = 0, 1, \dots, J+1$ —介质天线罩的层数;  
 $V_i = V_1, V_2, \dots, V_J$ —各层天线罩实体.

图 1 天线-天线罩系统模型

Fig. 1 Model of antenna-radome system structure

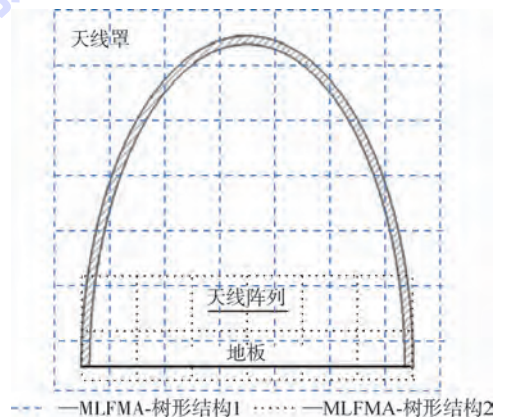


图 2 IE/MSI + MLFMA 方法中的 MLFMA 树形结构

Fig. 2 MLFMA-tree structure applied for

MSI + MLFMA method

个带罩天线系统,首先用于 MSI 的 MLFMA 加速计算<sup>[5]</sup>.第 2 套树形结构 2 仅覆盖天线阵列及附属结构(如反射地板),用于天线阵列所对应的 VSIE 模型的 MLFMA 快速求解.本文称此改进后的方法为 IE/MSI+MLFMA 法.可以看出,改进型算法中 MSI 部分的计算复杂度也降到了  $O(N \log N)$  量级,这就使得该算法不仅占用内存少,而且计算时间也大为缩短,从而实现了对外大尺寸带罩天线系统快速高效仿真的目标.

## 2 相关数值计算技术

### 2.1 基于改进型球谐函数展开的 MLFMA 技术求解 VSIE

在传统 MLFMA 求解积分方程的过程中,远区盒子之间的相互作用是通过单位球面上平面波的积分来实现的.为了提高计算效率,基函数在 Ewald 球面上的过采样信息,即基函数的方向图,需要事先计算并存储下来.采样率由底层对角转移矩阵的带宽决定,但过采样将浪费大量的核心内存.为解决这一问题,国外学者针对面积分方程提出了一种基于球谐函数展开的多层快速多极子方法(Spherical harmonics Expansion-MLFMA, SE-MLFMA)<sup>[6-7]</sup>,这种方法不再直接存储方向图信息,而是转化为对方向图球谐函数展开系数的存储,由于展开系数的个数远小于球面上的过采样点数,因此 SE-MLFMA 方法可将传统 MLFMA 的内存用量由  $2N(L+1)^2 c_1 c_2 B$  降至  $(3N/2)(P+1)(P+2)c_1 c_2 B$ .其中,  $L$  为底层多极子模式数,  $P$  为球谐函数的截断参数,一般取  $P=L/2-1$ ,  $N$  为未知量个数,  $c_1$  和  $c_2$  为两个常数:  $c_1$  与积分方程类型有关;电场积分方程对应  $c_1=1$ ,而混合场积分方程对应  $c_1=2$ ;  $c_2$  是与数值计算精度相关的常数,单精度时  $c_2=8$ ,双精度时  $c_2=16$ .而且,底层平面波发散过程的计算可以利用球谐函数的正交性质来加速,这就使得在保证精度与减少内存用量的前提下能够提高计算速度.

与 SIE 不同,在 VSIE 的 MLFMA 求解中,由于体基函数在底层盒子中密集存在而导致方向图的存储需要大量内存,传统的 SE-MLFMA 方法在内存用量和计算效率方面仍存在不足.为此,本文传统方法做了两方面的改进<sup>[8]</sup>:①有别于传统 MLFMA 中积分方程的并矢形式,将 VSIE 写成了混合位形式;②不同于传统 MLFMA 中 RWG 与 SWG 基函数<sup>[9-10]</sup>基于公共边/公共面的分组方式,采用了基于三角形/四面体的分组形式.改进后的算法与传统方法相比,计算性能上有 3 点

提高:

1) 方向图存储量进一步下降至  $2N_i(P+1)(P+2)c_2 B$ .此时,方向图的内存消耗仅与三角形与四面体的总数  $N_i$  成正比,而与公共边和公共面的总数  $N$  无关.而对于通常的网格剖分,  $N_i$  的值要远小于  $N$ ,因此方向图的内存用量大为降低.

2) 方向图存储量不再与 SIE 类型有关.对于传统的 SE-MLFMA,混合场积分方程中方向图存储量需求是电场积分方程的 2 倍,而对改进型算法而言,以上两种不同类型积分方程的方向图内存用量完全相同.

3) 更快的计算速度.为了避免 Gibbs 现象以及获得更小的  $P$  值,传统并矢形式的 SE-MLFMA 算法的聚集、转移及发散过程在极坐标系下进行,而底层方向图的聚合与发散需要在直角坐标系下完成.因此,传统算法中方向图函数需要在直角坐标系与极坐标系下进行反复转换;而对于改进算法,MLFMA 的所有过程均可在直角坐标系下完成,因此省去了坐标系间多次转换步骤,提高了计算速度.

### 2.2 预处理技术

天线阵列的 VSIE 模型经 MLFMA 离散化后通常得到形如  $\mathbf{Z}\mathbf{I}=\mathbf{V}$  的大型线性方程组,其中,  $\mathbf{Z}$  为非奇异广义阻抗矩阵,  $\mathbf{V}$  为激励矢量,  $\mathbf{I}$  为待求的展开系数矢量. MLFMA 中采用迭代法求解该线性方程组,其收敛速度由矩阵  $\mathbf{Z}$  的条件数所决定.通常 VSIE 对应的  $\mathbf{Z}$  矩阵条件数较差,导致迭代求解的收敛速度一般较慢,甚至可能出现不收敛的情况.因此,采用预处理技术改善问题的数值性态是提高计算效率的有效途径之一.

预处理技术构造出新的矩阵方程

$$\mathbf{M}\mathbf{Z}\mathbf{I}=\mathbf{M}\mathbf{V} \quad (3)$$

或

$$\begin{cases} \mathbf{Z}\mathbf{M}\mathbf{y}=\mathbf{V} \\ \mathbf{I}=\mathbf{M}\mathbf{y} \end{cases} \quad (4)$$

式中:  $\mathbf{M}$  为预条件矩阵;  $\mathbf{y}$  为预条件中间向量.新线性系统的条件数由  $\mathbf{M}\mathbf{Z}$  或  $\mathbf{Z}\mathbf{M}$  来决定.显然,若  $\mathbf{M}$  为  $\mathbf{Z}^{-1}$  的一个很好的近似,则新系统的数值性态将得到大大改善.但是,对于 MLFMA 而言,由于仅有近区场矩阵  $\mathbf{Z}_{\text{near}}$  被显式存储,因此常用  $\mathbf{Z}_{\text{near}}$  来构造  $\mathbf{M}$ .

在对 VSIE 的求解中,本文采用了 2 种不同的预处理技术构造矩阵  $\mathbf{M}$ ,即不完全 LU 分解(Incomplete Low-Up, ILU)方法和稀疏近似逆(Sparse Approximate Inverse, SAI)方法.

### 2.2.1 ILU 预处理技术

ILU 预处理将近场作用矩阵  $\mathbf{Z}_{\text{near}}$  做不完全 LU 分解,然后将预条件矩阵  $\mathbf{M}$  表示为

$$\mathbf{M} \approx (\mathbf{L}\mathbf{U})^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1} \quad (5)$$

式中: $\mathbf{L}$  和  $\mathbf{U}$  分别为下三角和上三角矩阵. 由于  $\mathbf{L}^{-1}$  与  $\mathbf{U}^{-1}$  并不显式存储,因此 ILU 实质上是隐式预处理. 为了降低内存用量、加快分解速度,在进行 ILU 分解时必须舍弃一些元素. 根据舍弃策略的不同,常用的 ILU 预处理方式有无填充不完全分解预处理 ILU(0)、多层填充不完全分解预处理 ILU( $k$ ) ( $k$  为填充度)、双门槛不完全分解预处理 ILUT(Threshold ILU)、带选主元形式的 ILUTP (ILUT with Pivot) 等. 其中,由于 ILUT 与 ILUTP 具有鲁棒性好、精度可调节和内存用量可控的优势,本文选择了这 2 种预处理技术加速 VSIE 的 MLFMA 求解. 大量的计算结果表明,对于中等规模的天线阵列电磁仿真问题,ILUT 及 ILUTP 可以取得良好的预条件效果<sup>[11-12]</sup>.

然而,ILU 预条件也有自身的缺陷. 由于其自身结构所限,ILU 预条件的分解及迭代求解过程均难以并行. 近年来,虽有学者对可并行化的多水平不完全 LU 分解法 (ILUM) 进行了探讨,但未见深入研究. 这限制了计算程序的可扩展性,使得在求解大规模电磁散射问题时,ILU 预处理成为了技术瓶颈.

### 2.2.2 SAI 预处理技术

SAI 预处理<sup>[13]</sup> 的基本思想是通过求解稀疏矩阵  $\mathbf{Z}$  的 Frobenius 范数最小化问题构造预条件矩阵  $\mathbf{M}$ ,以期所得的  $\mathbf{M}$  与稀疏矩阵  $\mathbf{Z}$  的乘积能尽可能近似等于单位矩阵  $\mathbf{I}$ . 由于矩阵  $\mathbf{M}$  为显式存储形式,因此 SAI 为显式预处理技术. MLFMA 中一般用  $\mathbf{Z}_{\text{near}}$  来构造  $\mathbf{M}$  矩阵,其 Frobenius 范数最小化问题可以分解为该矩阵单独的各列向量的最小二乘问题:

$$\min \|\mathbf{I} - \mathbf{Z}_{\text{near}}\mathbf{M}\|_F^2 = \sum_{k=1}^N \|\mathbf{e}_k - \mathbf{Z}_{\text{near}}\mathbf{m}_k\|_2^2 \quad (6)$$

式中: $\mathbf{e}_k$  和  $\mathbf{m}_k$  分别为单位矩阵  $\mathbf{I}$  与预条件矩阵  $\mathbf{M}$  的第  $k$  列元素组成的列向量. 显然,相互独立的  $\mathbf{m}_k$  可以被单独求解,因此 SAI 预条件具有天然并行性. 实际操作过程中,对于特定的第  $k$  个最小二乘问题,常采用  $\mathbf{Z}_{\text{near}}$  的缩减矩阵  $\mathbf{Z}_k$  并通过 QR 分解来计算. 如设  $J_k$  为矩阵  $\mathbf{Z}_{\text{near}}$  的第  $k$  行非零元素的列号组成的集合:

$$J_k = \{j' | z(k,j) \neq 0, j=1, 2, \dots, N\} \quad (7)$$

式中: $j'$  为列编号. 同样用  $I_k$  表示  $J_k$  所对应的各列非零元素的行号的集合:

$$I_k = \{i' | z(i,j) \neq 0, j \in J_k, i=1, 2, \dots, N\} \quad (8)$$

式中: $i'$  为行编号. 这样对于任意的  $\mathbf{m}_k, \mathbf{Z}_{\text{near}}$  可被缩减为相应的子矩阵  $\mathbf{Z}_k = \mathbf{Z}(I_k, J_k)$ . 显然,当  $\mathbf{Z}_{\text{near}}$  足够稀疏时,对  $\mathbf{Z}_k$  进行 QR 分解的计算量为  $O(1)$  量级.

然而,通过式(6)可知,构造  $\mathbf{M}$  时需要进行  $N$  次最小二乘计算,这导致了即使用缩减矩阵来构造  $\mathbf{M}$  的耗时依然很“漫长”. 因此,本文的改进方法是将 SAI 与 MLFMA 的分组方式结合起来. 从 MLFMA 的特点可知,未知量之间的相互作用不再是一一对应,而是通过包含未知量的盒子来进行近、远相互作用. 这时,每个特定盒子所包含的所有未知量相对应的近场元素排列方式(即非零元素的位置)是一致的. 例如,假设  $G$  为某一底层非空盒子所包含的所有未知数编号的集合,则对于所有  $k \in G, \mathbf{Z}_{\text{near}}$  的缩减矩阵  $\mathbf{Z}_k$  均相同. 因此,对于所有  $k \in G$ ,仅需进行一次 QR 分解即可求得所有  $\mathbf{m}_k$ . 这使得构造  $\mathbf{M}$  时所需求解的最小二乘问题降为  $N_{\text{group}}$  ( $N_{\text{group}}$  为底层非空盒子数,个). 由于  $N_{\text{group}} \ll N$ ,因此  $\mathbf{M}$  的构造时间得以成倍下降,效率大幅提升.

需要指出的是,SAI 对于矩阵的稀疏程度非常敏感. 对于某些包含复杂精细结构的仿真目标,其网格剖分往往过于密集,这就导致了  $\mathbf{Z}_{\text{near}}$  的稀疏程度较低. 此时,即使采用了与 MLFMA 相结合的 SAI 来构造预条件矩阵  $\mathbf{M}$ ,其耗时也将难以忍受. 若采用同 ILUT 预条件类似的舍弃策略,则预处理效果往往大打折扣. 另外,随着仿真目标电尺寸的增大,近场矩阵  $\mathbf{Z}_{\text{near}}$  对全局矩阵  $\mathbf{Z}$  的近似性将变得越来越差,这是所有预条件技术都会遇到的问题,这需要通过人为地扩大用于构造  $\mathbf{M}$  的矩阵规模来缓解.

### 2.3 并行计算技术

VSIE 的求解过程中,数值仿真时间将随着目标电尺寸的增加而成平方量级增长,对于电尺寸较大的问题,这样的计算速度显然是不可接受的. 现代并行计算技术是提高仿真速度的重要手段之一,在计算电磁学领域,这一技术已经得到国内外学者的广泛研究<sup>[14-18]</sup>. 针对大型带罩天线阵列的电磁仿真问题和现有的多核计算平台,本文提出了一种新的结合单程序多数据以及单指令多数据并行计算模型的混合并行计算模式,很好地结合了基于共享内存式架构的 OpenMP 并行模型与基于矢量逻辑运算单元 (Vector Arithmetic Logic Unit, VALU) 的硬件矢量加速技术两者的优点. 与

基于 GPU (Graphic Processing Unit) 的加速技术<sup>[14]</sup>相比, VALU 加速技术<sup>[15]</sup>具有 2 个明显优势:①由于 VALU 仅是 CPU 的一个基本单元,因此不需要添加任何其他硬件设备而仅需做程序上的改动;②应用了 VALU 加速技术的程序其计算速度一定比不使用该技术时快,而同样的结论并不适用于 GPU 加速.

以下将着重讨论在共享内存式的计算机架构下,应用该混合并行技术加速 VSIE 的 MLFMA 求解的几个关键技术<sup>[17-18]</sup>.

### 2.3.1 OpenMP 并行程序设计中的关键问题

MLFMA 算法的实现中,最耗时的两个部分为近场矩阵填充与矩阵向量相乘 (Matrix-Vector Product, MVP). 前一部分的 OpenMP 并行实现十分简单,直接将使用了 Guided 策略的 OpenMP 命令应用到程序的最外重循环即可获得良好的并行效果;而 MVP 部分包含两个过程:近场相互作用 (Near-Field Interaction, NFI) 与远场相互作用 (Far-Field Interaction, FFI).

其中, FFI 通过 3 个步骤实现:聚集、转移和发散. 聚集和发散过程的程序代码主要包含三重循环:父盒子循环 (Father Box, FB)、子盒子循环 (Son Box, SB) 以及平面波循环 (Plane Wave, PW). 传统串行程序中,按照 FB-SB-PW 循环顺序设计的代码将会获得最高的执行效率,然而并行程序中若将 OpenMP 的指令直接应用到最外重的 FB 循环上,随着聚集过程的进行,父盒子数目迅速下降,甚至有可能出现父盒子数目少于线程数的情况,导致各线程间的负载越来越难以达到平衡. 因此,为了取得更高的并行效率,需要调整程序的循环次序. 数值实验结果表明,在高层循环中,虽然 PW 循环作为最外重循环可以获得最高的并行效率,但计算时间会明显增加. 同时,由于每一层中子盒子数目都要多于父盒子数目,故而在不增加计算时间的前提下,将子盒子循环设置为最外重循环更易于实现负载平衡. 因此,对于聚集与发散过程的并行设计而言,最佳的循环结构应为 SB-FB-PW. 转移过程的算法同样包含三重循环:盒子循环、每个盒子的次相邻盒子循环以及平面波循环. 同样,从低层到高层每层包含非空盒子的数目迅速下降,在高层若直接将 OpenMP 指令应用到盒子循环将不能获得良好的并行效率. 不仅如此,由于盒子与其次相邻盒子都处于同一层中,应用到聚集和发散过程中的循环策略也不适用于转移过程. 因此,需要重新考虑循环策略. 实验数据表明,在高层中,平面波循环作为最外重

循环可以获得更高的并行效率以及更短的并行时间,而在低层将盒子循环作为最外重循环更为合适. 因此必然存在一个过渡层,过渡层的两侧应采用不同的循环策略以获得最优的并行效率. 然而,过渡层的选择取决于总层数以及高层盒子内的平面波数量,若选择不当则并行效率将会大打折扣. 大量的实验数据表明,当在某一层分配到每个线程中的循环数大于 60 时,选择这一层作为过渡层是较为合适的.

近场相互作用 NFI 的实质是一个 CSR 格式存储的稀疏矩阵与一个向量相乘,简单的代码结构造成了内存数据的存取速率滞后于 CPU 的计算速率. OpenMP 中,随着参与 NFI 计算的线程数增加,这一滞后现象将变得越来越严重,最终使得 NFI 的并行计算会遭遇“内存墙”问题. 这一问题导致 NFI 的计算时间不会再随参与计算的线程数的增加而减少,严重影响了程序的并行效率与可扩展性. 为了解决这一问题,本文的策略是将 NFI 的代码与聚集过程中底层盒子方向图聚合的代码合并. 由于 NFI 的计算与底层方向图的聚合都是针对基函数的操作,因此可以使这两部分的计算共享同一内存寻址时间,从而将 NFI“隐藏”到方向图的聚合过程中. 这时,由于方向图的聚合过程较为复杂,与内存的寻址时间相比, CPU 的计算时间将起主导作用,这样就减弱了“内存墙”现象带来的影响. 实验数据表明,采用合并策略以后,并行效率大幅提升.

### 2.3.2 VALU 硬件加速技术的关键问题

VALU 为 CPU 的一个基本单元,与浮点运算单元 (FPU) 每次只能对一个数据进行操作不同, VALU 可以对 4 个数据同时进行操作. 其运算规则依赖于 SSE 指令集,而就 Windows 系统而言, Win98 之后的所有系统均支持此指令集,并且从 SSE2 起, VALU 开始支持双精度的浮点数运算. 因此,理论上应用了 VALU 加速的程序比普通程序最多可以快 4 倍. 然而,应用了 VALU 加速技术的代码,并非仅是对原有程序进行简单调整,而是要从算法结构上做出根本改变<sup>[17-18]</sup>. 同 OpenMP 相比, VALU 加速技术粒度更细,对程序细节的要求也更高. 一般而言,只有纯粹的赋值语句与多重循环的最内重循环才可利用此项技术. 不仅如此,应用 VALU 的程序代码还需满足以下 3 个条件:①循环体内的每一个语句能够独立执行;②除编译器的标准内部函数外,循环体内不能调用任何自定义函数;③对于需要退出条件的循环,只能有一个出入口.

影响 VALU 加速性能的一个重要因素是数据对齐. 由于只有当 CPU 每次从缓存中读取的数据个数为 4 时才可应用 VALU 进行加速, 因此, 若参与计算的数组第一维度较小并且不是 4 的整数倍时, VALU 的加速性能将会大受影响. 这就需要在进行代码的编写工作时, 仔细考虑算法流程, 尽量将能够应用 VALU 加速技术进行计算的数组的第一维度存储成 4 的整数倍.

另外, VALU 单元只能针对整数与实数进行运算, 而对于 MLFMA 运算中最常用的复数类型是不适用的. 因此, 首先需要将代码中所有的复数运算全部转化为实数形式. 实际操作过程中, 4 个同种类型的数据(实数或整数)被收集成数据流的形式同时输入到 VALU 中; 经过同一个计算周期后, 4 个计算结果又以数据流的形式同时输出. 需要指出的是, 这 4 个数据必须进行的是同类型的运算才可应用 VALU 进行加速.

### 3 数值算例

本节采用两个数值算例来验证上述算法的正确性与计算效率. 其中, 计算代码基于 Intel Visual FORTRAN 平台编写, 并且采用普通 PC 机进行计算, 其配置水平如表 1 所示.

表 1 计算平台配置水平

Table 1 Configuration level of computing platform

CPU 类型	核心数	内存大小	操作系统
Intel Core™ i5-4570 @ 3.2 GHz	4	DDR3 16 GB	Windows7 Ultimate x64

另外, 为分析程序的并行效率, 定义并行加速比为

$$S = T / T_p \quad (9)$$

式中:  $T$  为串行程序(serial)执行时间;  $T_p$  为并行程序(OpenMP)执行时间. 所有数值算例均采用重启型 GMRES 求解器进行求解, 其中重启次数为 100, 收敛精度为 0.001.

#### 3.1 微带天线阵列

带 A 夹层罩的微带天线阵列系统工作频率为 10.3 GHz, 模型如图 3(a) 所示. A 夹层天线罩关于  $yOz$  面对称, 表 2 中给出了天线罩各层的具体参数. 天线阵列由 4 个如图 3(b) 所示的方形切角微带贴片单元组成, 4 个贴片单元分别放在坐标系  $x$  轴和  $y$  轴上并距离坐标原点 50 mm, 基板介电常数为 2.55. 阵列剖分所得的三角形个数为 23 300, 四面体个数为 20 656, 总的未知量为 82 352; 天线罩剖分所得三角形个数为 45 824.

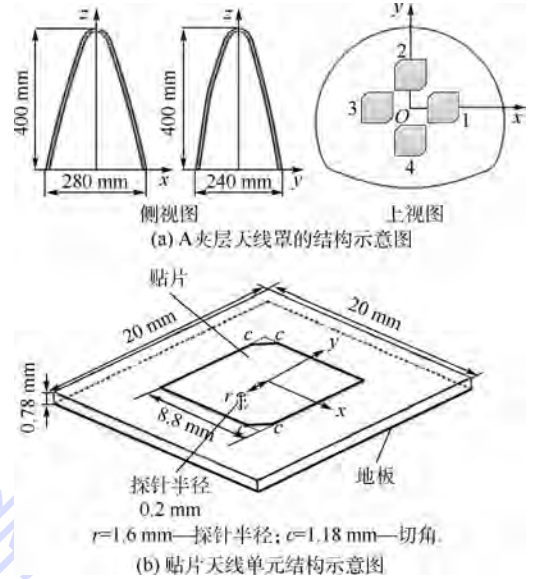


图 3 带罩微带天线阵列

Fig. 3 Radome-enclosed microstrip patch antenna array

表 2 A 夹层天线罩的参数

Table 2 Parameters of A-type sandwiched radome

层数	相对介电常数	损耗角正切	厚度/mm
蒙皮 1	4.0	0.02	0.6
夹芯 2	1.07	0.002	8.8
蒙皮 3	4.0	0.02	0.6

图 4 给出了分别应用传统 MLFMA(conv.) 与 SE-MLFMA(se) 算法并结合 ILU 预处理技术得到的不同入射角度下端口 1 的接收电压值, 入射角度分辨率为  $2^\circ$ . 其中, 对于 SE-MLFMA,  $P$  的取值为 2. 从图 4 中可以看出, SE-MLFMA 与传统 MLFMA 相比精度虽稍有下降, 但完全可以满足工程需求. 为了对比, 图 4 中也给出了当  $P=3$  时 SE-MLFMA 的仿真结果, 可以看出, 此时 SE-MLFMA 与传统 MLFMA 的仿真结果高度吻合. 总的计算时间、加速比与方向图的内存需求如表 3 中 ex1 所示, 其中 se2 和 se3 分别对应  $P=2$  和  $P=3$ .

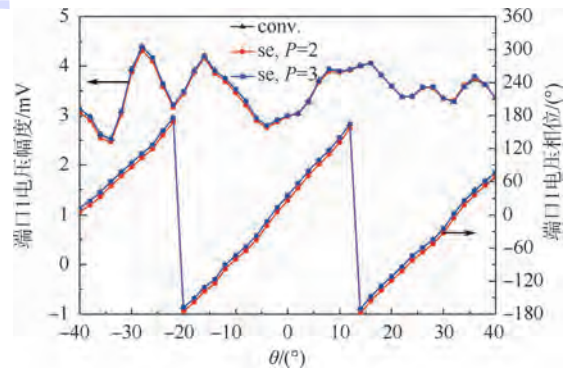


图 4 不同入射角度下端口 1 的接收电压变化  
Fig. 4 Variation of receiving voltage at port 1 changing with different incidence angles

表3 仿真所耗的总时间、并行加速比及方向图内存用量

Table 3 Total solving time, parallel ratio and pattern memory cost of simulations

算例	类型	总时间/s	加速比	方向图内存/MB	峰值内存/MB
ex1	Serial&conv.	2995			
	OpenMP&conv.	1743	1.72	21.55	6049
	Serial&se2	2860			
	OpenMP&se2	1672	1.71	8.07	6036
ex21	Serial&se3	2951			
	OpenMP&se3	1724	1.71	13.40	6041
	Serial&se	66647			
	OpenMP&se	18389	3.62	60.61	4978
ex22	Serial&se	81968			
	OpenMP&se	22860	3.59		

从表3中可以看出,SE-MLFMA的方向图内存用量仅为传统MLFMA的37.4%,而总的计算时间大体持平.这就使得SE-MLFMA在保证精度与计算速度的前提下,大幅降低了仿真时的核心内存用量.另外,此例中程序的并行加速比仅为1.7,这是由ILU预处理不可并行所导致.由此可见,ILU预处理技术对并程序的可扩展性有较大限制.

### 3.2 带罩天线阵的接收方向图

1604个半波阵子单元组成天线阵,工作频率为16.5GHz.天线罩为半波长壁罩,带罩天线阵列的模型如图5所示.

图5中阵列俯仰间距为9.375mm,方位间距为21.25mm,反射板为直径750mm的理想导体圆盘.阵列剖分所得的三角形个数为330383,总的未知量个数为483724;天线罩剖分所得三角形个数为803556.其中,超球体方程天线罩外形为

$$x^2 + y^2 = \left(\frac{D}{2L}\right)^2 (L^v - z^v)^{2/v} \quad (10)$$



图5 1604个单元的带罩天线阵列模型结构

Fig.5 Structure of a radome enclosed antenna array model with 1604 elements

式中:超球体的直径 $D$ 与高度 $L$ 均为750mm,系数 $v=1.8$ ; $x$ 、 $y$ 和 $z$ 分别为3个方向的尺度.

应用SE-MLFMA并结合SAI预条件分别对不加罩的孤立阵与加罩后的天线阵进行仿真,得到的接收方向图如图6所示,其中分辨率为 $0.25^\circ$ .从图6中可以看出,孤立阵与加罩阵的接收方向图相差不大,表明所设计的天线罩具有较好的电磁特性.仿真的计算时间、加速比与方向图的内存需求分别如表3中的ex21(孤立阵)、ex22(加罩阵)所示.从表3可以看出,加罩前后天线阵列的总体仿真时间差距并不大,这说明IE/MSI+MLFMA算法极大降低了天线罩对天线阵列仿真效率的影响.此外,由于采用了可并行的SAI预处理技术,与第3.1节中的算例相比并行程序的加速比有了大幅提升,达到了3.6.这说明SAI预处理技术具有良好的并行特性,尤其适用于电大尺寸目标的仿真计算.

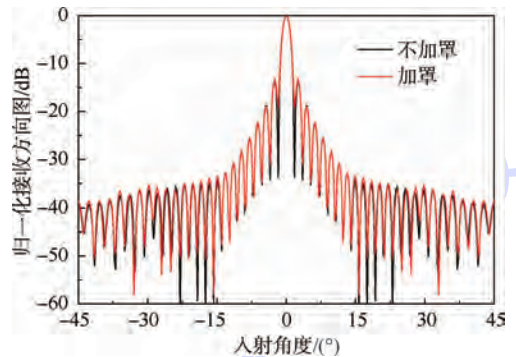


图6 偶极子阵列接收方向图

Fig.6 Receiving pattern of dipole array

## 4 结论

本文针对接收模式下的天线-天线罩系统,提出了一种全波法结合高频法的混合算法,结合仿真结果,得到:

1) 将IE/MSI+MLFMA混合算法成功应用于加罩天线阵系统的电磁仿真计算.在具体的算法实现过程中,球谐函数展开、ILU与SAI预处理技术、OpenMP+VALU的混合并行等技术的应用使得该算法得到了相当大的改进.

2) 通过具体算例可知,加罩前后天线阵列的总体仿真时间差距并不大,这说明IE/MSI+MLFMA算法极大降低了天线罩对天线阵列仿真效率的影响,实现了对电大尺寸天线-天线罩系统的快速、高效仿真,并且具有良好的可扩展性.

3) 具体的程序细节方面,对于SE-MLFMA的应用,当 $P=2$ 时已经完全可以满足工程需求;相较于ILU,SAI预处理技术具有良好的并行特

性,尤其适用于电大尺寸目标的仿真计算,使得并行程序的加速比得到大幅提升。

### 参考文献 (References)

- [1] Kozakoff D J. Analysis of radome-enclosed antennas [M]. Norwood, MA: Artech House, 2010: 5-8.
- [2] Chew W C, Jin J M, Michielssen E, et al. Fast and efficient algorithms in computational electromagnetics [M]. London: Artech House, 2001: 77-85.
- [3] Wang B B, He M, Liu J B, et al. An efficient integral equation/modified surface integration method for analysis of antenna-radome structures in receiving mode [J]. IEEE Transactions on Antennas and Propagation, 2014, 62(9): 4884-4889.
- [4] Paris D T. Computer-aided radome analysis [J]. IEEE Transactions on Antennas and Propagation, 1970, 18(1): 7-15.
- [5] Wang B B, He M, Liu J B, et al. Efficient method for analysis of radome in receiving mode [C] // Proceedings of Asia-Pacific Microwave Conference. Piscataway, NJ: IEEE Press, 2014: 1408-1410.
- [6] Eibert T F. A diagonalized multilevel fast multipole method with spherical harmonics expansion of the  $k$ -space integrals [J]. IEEE Transactions on Antennas and Propagation, 2005, 53(2): 814-817.
- [7] Ismatullah I, Eibert T F. Surface integral equation solutions by hierarchical vector basis functions and spherical harmonics based multilevel fast multipole method [J]. IEEE Transactions on Antennas and Propagation, 2009, 57(7): 2084-2093.
- [8] He M, Liu J B, Zhang K. Improving the spherical harmonics expansion based multilevel fast multipole algorithm [J]. IEEE Antennas and Wireless Propagation Letters, 2013, 12: 551-554.
- [9] Rao S M, Wilton D R, Glisson A W. Electromagnetic scattering by surfaces of arbitrary shape [J]. IEEE Transactions on Antennas and Propagation, 1982, 30(3): 409-418.
- [10] Schaubert D H, Wilton D R, Glisson A W. A tetrahedral modeling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies [J]. IEEE Transactions on Antennas and Propagation, 1984, 32(1): 77-85.
- [11] Sertel K, Volakis J L. Incomplete LU preconditioner for FMM implementation [J]. Microwave and Optical Technology Letters, 2000, 26(4): 255-257.
- [12] Lee J, Zhang J, Lu C C. Incomplete LU preconditioning for large scale dense complex linear systems from electromagnetic wave scattering problems [J]. Journal of Computational Physics, 2003, 185: 158-175.
- [13] Malas T, Gurel L. Accelerating the multilevel fast multipole algorithm with the sparse-approximate-inverse preconditioning [J]. Society for Industrial and Applied Mathematics, 2009, 31(3): 1968-1984.
- [14] Peng S X, Nie Z P. Acceleration of the method of moments calculations by using graphics processing units [J]. IEEE Transactions on Antennas and Propagation, 2008, 56(7): 2130-2133.
- [15] 余文华, 李文兴. 高等时域有限差分法 [M]. 哈尔滨: 哈尔滨工程大学出版社, 2011: 23-39.
- Yu W H, Li W X. The advanced FDTD method [M]. Harbin: Harbin Engineering University Press, 2011: 23-39 (in Chinese).
- [16] Velamparambil S, Chew W C. Analysis and performance of a distributed memory multilevel fast multipole algorithm [J]. IEEE Antennas and Propagation Magazine, 2005, 53(8): 2719-2727.
- [17] Liu J B, He M, Zhang K, et al. Parallelization of the multilevel fast multipole algorithm by combined use of OpenMP and VALU hardware acceleration [J]. IEEE Transactions on Antennas and Propagation, 2014, 62(7): 3884-3889.
- [18] 刘金波, 何芒, 基于 OpenMP 与 VALU 硬件加速的表面积分方程矩量法混合并行求解技术 [J]. 北京理工大学学报, 2014, 34(1): 50-55.
- Liu J B, He M. A hybrid parallelization technique based on OpenMP and VALU acceleration for the method of moments solution of the surface integral equations [J]. Journal of Beijing Institute of Technology, 2014, 34(1): 50-55 (in Chinese).

## Fast electromagnetic simulation of characteristics of radome-antenna system in receiving mode

HE Mang<sup>\*</sup>, LIU Jinbo, WANG Binbin, XU Xiaowen, ZHANG Chuanfang, SUN Houjun

(School of Electronics and Information, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** An efficient hybrid approach, which combines the full-wave and high-frequency method, i. e., the multilevel fast multipole algorithm (MLFMA) based volume-surface integral equation (VSIE) and the modified surface integration (MSI) method, was proposed to analyze the electromagnetic characteristics of the electrically large radome-antenna system in receiving mode. The MSI method determined the transmitted field distributions through the radome firstly and treated the fields as the excited source of the antennas, and then the MLFMA accelerated VSIE solver was used to analyze the antennas accurately. The method employed spherical harmonic expansion, preconditioner and hybrid parallel techniques to further improve the calculation efficiency of the algorithm. This new IE/MSI + MLFMA method has reasonable accuracy and dramatically reduces the computational time and improves the efficiency when compared to the conventional full-wave numerical methods, and produces fast and efficient simulation of electrically large antenna-radome system.

**Key words:** radome-antenna system; volume-surface integral equation (VSIE); surface integration; multilevel fast multipole algorithm (MLFMA); spherical harmonic expansion; parallel computing

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0215

# 双反射面信道中6极化MIMO信道特性



朴大志\*, 陈鹏

(中国传媒大学 信息工程学院, 北京 100024)

**摘 要:** 为研究双反射面信道中多径丰富程度对6极化多输入多输出系统(HPMIMO)信道容量和特征值的影响,设计了一个工作于28 GHz的共点正交的双极化天线,此天线由一个电偶极子和一个环天线组成,并以此天线为基本单元,得到了一个虚拟的共点正交6极化MIMO天线.从计算结果中发现,随着两个反射面间夹角的减小,信道中的多径数目逐渐增加,对应的信道容量也逐渐增加,可达到单极化系统信道容量的2.47~5.10倍.夹角为0°时多径数目趋于无穷的情况比自由空间(FS)和夹角为180°等多径较少的情况下的6极化MIMO信道容量会有30%~45%的增加.此外,在自由空间信道和双反射面信道中,6极化MIMO信道矩阵都会得到6个非零特征值,但在多径数目较少时,有一些特征值会接近于0.因此,为了获得足够高的信道复用和分集增益,多极化MIMO天线需要工作在多径较为丰富的信道中.

**关键词:** 极化; MIMO; 信道容量; 特征值; 双反射面

**中图分类号:** TN929.5

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1801-06

无线频谱是一种非常稀缺而昂贵的资源,为了应对无线互联网、无线视频等高流量业务的爆发性增长,人们不得不寻找一些能够充分利用电磁场的矢量特性、具有更高频谱效率的无线通信技术.多输入多输出系统(Multiple-Input Multiple-Output, MIMO)技术在提高频谱效率方面的独特优势已经引起了广泛关注,但在传统的利用空间分集来实现的MIMO系统中,为了获得足够大的自由度,收发阵元间距在手持设备端至少要达到半个波长,在基站端则可能要达到10个波长以上,从而限制了MIMO技术的应用.而采用极化分集构成的MIMO系统可以通过空间共点的天线来实现,因此多极化MIMO技术在密集天线阵MIMO通信系统和高分辨率到达角(Direction of Arrival, DOA)估计MIMO雷达系统中都具有广阔的应用前景.

在传统的极化分集系统中,一般只考虑用电

场的垂直和水平2个分量来得到2个独立的信道.文献[1]中通过理论分析指出如果在发射和接收端都使用3个正交的电偶极子天线和3个正交的磁偶极子天线,通过对电磁场中 $E_x$ 、 $E_y$ 、 $E_z$ 和 $H_x$ 、 $H_y$ 、 $H_z$ 这6种极化状态的利用,可以获得与单极化系统相比更多的信道容量增益.接下来,人们分别通过基于统计分析的模型<sup>[2-5]</sup>或基于确定性的信道模型<sup>[6-7]</sup>对一个6极化MIMO系统的自由度进行了研究.在这些基于统计分析的建模方法,虽然这种方法可以从宏观上对多极化MIMO信道的统计特性有所认识,但从微观角度这种信道模型无法从物理上体现多极化MIMO信道特性的细节.在这些分析中,一般考虑理想的天线在一个散射丰富的环境中,通过多径的俯仰角和水平角的不同分布来描述不同的信道散射环境,用随机相关矩阵模型来进行信道建模.然而,这种建模方法

收稿日期: 2015-04-09; 录用日期: 2015-06-01; 网络出版时间: 2015-06-09 18:00

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150609.1800.001.html

基金项目: 国家自然科学基金(61201235); 北京高等学校青年英才计划(YETP0595)

\* 通讯作者: 朴大志(1977—),女,辽宁营口人,教授, piaodazhi@hotmail.com, 主要研究方向为多极化MIMO信道建模理论和天线设计.

**引用格式:** 朴大志, 陈鹏. 双反射面信道中6极化MIMO信道特性[J]. 北京航空航天大学学报, 2015, 41(10): 1801-1806.

Piao D Z, Chen P. Characteristics of hexapolarized MIMO channel in two-mirror channel[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1801-1806 (in Chinese).

中没有体现信道散射过程对极化耦合的作用,也无法直接体现 MIMO 信道容量与其传播环境间的关系以及收发阵的结构和位置等参数对 MIMO 信道特性的影响.此外,在基于模式理论的模型<sup>[8]</sup>中,多考虑无源的环境中远场平面波的叠加,而天线辐射特性的影响则无法得到清楚的体现.也有一些实验测量工作,利用一些共点的 $2 \times 2$ 、 $3 \times 3$ 和 $4 \times 4$ 的多极化天线原型<sup>[9-12]</sup>,此外,文献[13]提出了一种6极化 MIMO,HPMIMO(Hexpolarized MIMO)天线,但并非真正共点的结构.由于尺寸和耦合的限制,完全共点、正交的 HPMIMO 天线很难实现,因此对于发射端和接收端都采用6极化天线的 MIMO 信道特性测量试验也很少看到.

考虑到上述问题,本文利用电磁场仿真软件(HFSS),采用数值计算的方法,对多极化 MIMO 信道特性进行研究.通过 HFSS,近场和远场的精确解都可得到,比采用简化的统计模型更加可靠.文献[5]采用虚源理论,针对无限小的理想、无耦合的电偶极子和磁偶极子天线,对自由空间和几种多径环境中的 HPMIMO 信道特性进行了研究.

本文在前期的研究中发现,具有均匀电流分布的环天线构成的多极化 MIMO 天线阵所获得的信道容量较大且鲁棒性更高<sup>[14-15]</sup>.因此在本文中,采用半波电偶极子天线和一种具有均匀电流分布的环天线,研究多径数目逐渐增加时,HPMIMO信道容量和信道矩阵特征值的变化情况.为了实现上述研究目标,本文选择了具有不同夹角的双反射面信道,因为在这种环境下,通过改变两个反射面之间的夹角,很容易实现对信道中多径个数和分布的控制.虽然这种双反射面信道在实际的通信环境中不常遇到,但在这种信道中可以很方面地研究多径的丰富程度对多极化 MIMO信道特性的影响,也可以为更复杂信道环境下的建模理论提供参考.

## 1 天线设计

### 1.1 共点正交双极化天线

由于共点、正交、低耦合的6极化天线难以实现,本文首先得到了由一个电偶极子和一个环天线组成的双极化天线,如图1所示.然后利用这个双极化天线得到一个虚拟的6极化天线.

考虑到28 GHz的毫米波频段是第5代移动通信系统的一个重要目标频段<sup>[16]</sup>,因此将双极化天线的工作频段设为28 GHz.工作于其中的电偶极子天线为普通的半波偶极子天线,环天线采用了一种具有均匀电流分布的天线结构<sup>[17]</sup>,这样可

以得到近似磁偶极子天线的辐射特性.为了使它们的工作频率位于28 GHz,对文献[17]中的天线参数进行了调整和优化.在设计中本文发现,如果两个单个的天线都按照28 GHz的工作频带设计,则把它们放在一起以后,由于相互耦合,会使每个天线的谐振频率都发生较大的变化.因此本文对这种耦合作用首先进行了仿真,使单个天线的工作频带稍微偏离28 GHz.单个天线与合成天线回波损耗 $S_{11}$ 的比较见图2,这样可以使合成以后的电偶极子天线和环天线能够共同工作于28 GHz.双极化天线的S参数( $S_{11}$ ,  $S_{12}$ ,  $S_{21}$ ,  $S_{22}$ )见图3,其中1端口为环,2端口为电偶极子.

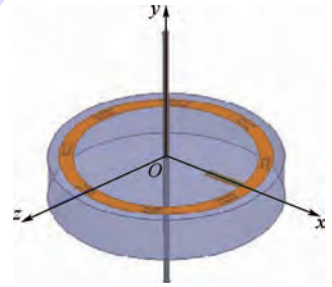


图1 共点正交电偶极子与环天线

Fig. 1 Co-located orthogonal electric dipole and loop antenna

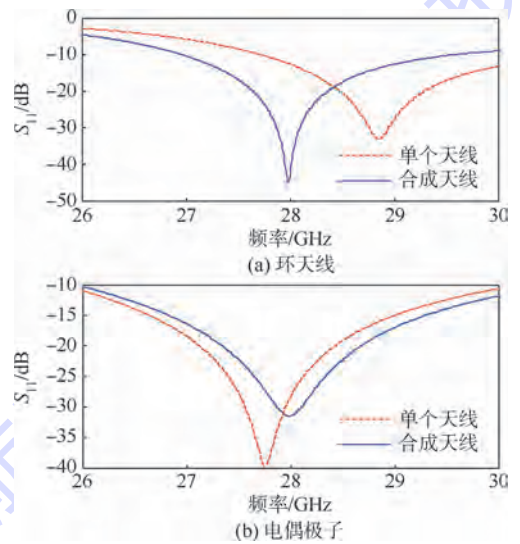


图2 单个天线与共点合成天线  $S_{11}$  的比较

Fig. 2  $S_{11}$  comparison between single antenna and co-located combined antenna

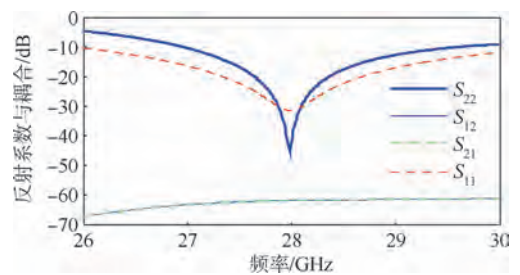


图3 双极化天线的S参数

Fig. 3 S-parameters of dual-polarized antenna

半波偶极子天线的两个金属圆柱长度均为 2.487 8 mm, 半径为 0.052 1 mm, 馈电点间隙为 0.026 mm. 在环天线中, 通过周期性负载来实现均匀的电流分布. 此天线印制在 Teflon 基板上, 基板的性能参数为介电常数  $\epsilon_r = 2.6$ ,  $\tan \delta = 0.002$ , 厚度  $h = 0.8$  mm. 整个环被均匀地分割成 8 个部分, 其中第 1 部分上连接着阻抗匹配转换器. 该环的具体参数为: 外半径  $R_1 = 1.9$  mm, 内半径  $R_2 = 1.65$  mm, 每一段弧形带状线的角度为  $\alpha_1 = 44^\circ$ , 相邻弧形带状线之间的间隔为  $\alpha_2 = 1^\circ$ . 为了实现周期性负载, 在每一弧形带状线的一端加上互相交织的耦合线. 这一部分的参数为: 内部的弧形带状线宽  $w_s = 0.09$  mm, 角度  $\alpha_3 = 11^\circ$ , 且与镂空带状线的内边界相隔  $g_s = 0.036$  mm, 与镂空带状线底部相隔角度  $\alpha_4 = 1^\circ$ . 并行的两根带状线阻抗匹配器的长度为  $l_p = 1$  mm, 宽  $w_p = 0.0425$  mm, 它们间隔为 0.068 mm. 这些参数的具体含义可以参考文献[17].

图 3 中可见, 环天线和半波偶极子天线的耦合很低,  $S_{12}$  和  $S_{21}$  参数低于  $-60$  dB. 考虑  $S_{11} < -10$  dB, 环天线的带宽约为 2.5 GHz, 偶极子天线的带宽高于 5 GHz.

## 1.2 虚拟 6 极化 MIMO 天线

对于共点正交的 3 个电偶极子和 3 个磁偶极子构成的收发阵, 其  $6 \times 6$  的信道矩阵通过对接收点处电磁场分量的计算来获得, 对于  $m$  或  $n = \{1, 2, 3\}$ , 信道矩阵的元素  $h_{mn}$  分别对应于沿着  $x$ 、 $y$  和  $z$  轴发射或接收的电偶极子; 类似的, 对于  $m$  或  $n = \{4, 5, 6\}$ , 信道矩阵的元素  $h_{mn}$  分别对应于沿着  $x$ 、 $y$  和  $z$  轴发射或接收的磁偶极子<sup>[1-3]</sup>.

完成图 3 所示的双极化天线设计以后, 为了得到一个虚拟的 6 极化天线, 本文首先让电偶极子天线和环天线的轴都位于  $x$  轴, 把它们放到仿真的环境中, 对它们分别进行馈电, 得到相应的信道响应. 然后让它们的中心保持不变, 对它们整体进行旋转, 使电偶极子天线和环天线的轴分别位于  $y$  轴和  $z$  轴, 对相应的信道响应进行重新计算, 然后把把这些结果按照信道矩阵的定义组合起来, 就可以得到  $6 \times 6$  的 MIMO 信道矩阵.

## 2 仿真实验及结果

### 2.1 双反射面信道

如图 4 所示的双反射面信道由两个夹角为  $\theta$

的无限大理想导体平面组成, 这里以  $\theta = 60^\circ$  为例, 对双反射面中多径的分布进行了说明. 其中  $h_t$  为发射天线 Tx 的高度,  $h_r$  为接收天线 Rx 的高度,  $d$  为 Tx 到两平面交线的距离,  $R$  为 Tx 和 Rx 的  $x$  轴坐标间的距离.

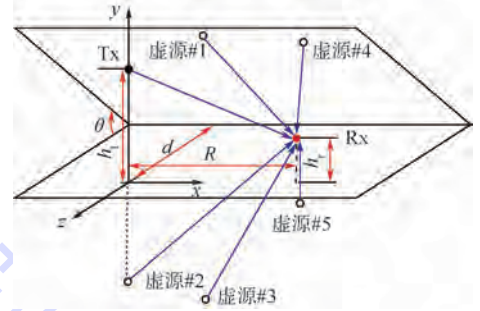


图 4 双反射面多径结构示意图

Fig. 4 Multipath structure schematic of two-mirror

根据镜像理论, 当  $\theta = 60^\circ$  时, 在两个反射面中产生的虚源个数为 5, 加上直射信号, 则信道中的多径个数为 6. 在仿真实验中, 本文将水平面固定, 另一个面绕着两平面的交线进行旋转. 根据多径个数  $N$  与两反射面间夹角  $\theta$  的关系:  $N = 360^\circ/\theta$ , 则当  $\theta$  按照  $180^\circ$ 、 $90^\circ$ 、 $60^\circ$ 、 $45^\circ$ 、 $30^\circ$  和  $0^\circ$  变化时, 信道中对应的多径的个数  $N$  为 2、4、6、8、12 和无穷多. 其中  $\theta = 180^\circ$  时两个反射面合成为一个单反射面,  $\theta = 0^\circ$  时两个反射面变成两个平行板.

在 HFSS 仿真中, 所用的计算机为 12 核 Intel 处理器, 内存为 64 GB. 考虑到计算机的计算能力和计算量的大小, 本文选择了两个反射面的尺寸相等, 均为  $153$  mm  $\times$   $38$  mm.

### 2.2 信道容量

信道容量一直被广泛用于评估一个 MIMO 系统的性能, 它被定义为一个信道中能够以任意低的误比特率所实现的最高信息传输速率. 如果信道的状态信息仅在接收端可以获得而发射端无法得到, 一般采用发射功率均匀分配的方式, 其信道容量<sup>[18]</sup>可以写为

$$C = \text{lb det}[\mathbf{I} + (\text{SNR}/n_T)\mathbf{H}\mathbf{H}^\dagger] \quad (1)$$

式中:  $\mathbf{I}$  为  $6 \times 6$  的单位阵;  $n_T = 6$  为发射天线数; SNR (Signal-to-Noise Ratio) 为信噪比, 本文设为 20 dB;  $\mathbf{H}^\dagger$  是信道矩阵  $\mathbf{H}$  的共轭转置. 当信噪比较高时, 采用发射功率均匀分配与最优的注水算法分配方式所获得的信道容量非常接近. 为了消除计算中路径损耗对总接收信号功率的影响, 一般采用归一化的信道矩阵:

$$\|\mathbf{H}\|_F^2 = n_T n_R \quad (2)$$

式中:  $\|\cdot\|_F$  为 Frobenius 范数;  $n_R$  为接收天线的个数. 采用这种归一化方式可以较好地反映信道中多径的丰富程度对 MIMO 系统信道容量的影响.

图 5 为 HPMIMO 信道容量与收发阵距离的关系. 其中: Tx 固定, Rx 沿着  $x$  轴移动, 在图 5(a) 和图 5(b) 中分别考虑了发射和接收天线高度相等与不相等的情况. 其中  $\theta = 0^\circ$  时两平行板之间的距离为 22 mm,  $\lambda_0$  是工作频率为 28 GHz 时自由空间 (Free Space, FS) 中的波长. 由图 5 可以看到,  $\theta$  不同时对应的 HPMIMO 信道容量有明显的区别. 总的来看,  $\theta = 0^\circ$  时的信道容量最大, 随着  $\theta$  的增加, 多径数目会逐渐减少, 相应的信道容量也会逐渐减小,  $\theta = 180^\circ$  时最小, 与自由空间信道的结果接近. 其中  $\theta = 0^\circ$  时的信道容量约为  $\theta = 180^\circ$  时的 1.3 ~ 1.45 倍.  $R$  为发射和接收天线间的距离,  $\lambda_0$  为自由空间中的波长. 当  $R < 7\lambda_0$  时, 自由空间和  $\theta = 180^\circ$  的单反射面信道中的信道容量要明显小于其他情况. 当  $R > 7\lambda_0$  时,  $\theta = 0^\circ$  和  $30^\circ$  时的信道容量结果要明显高于其他情况.

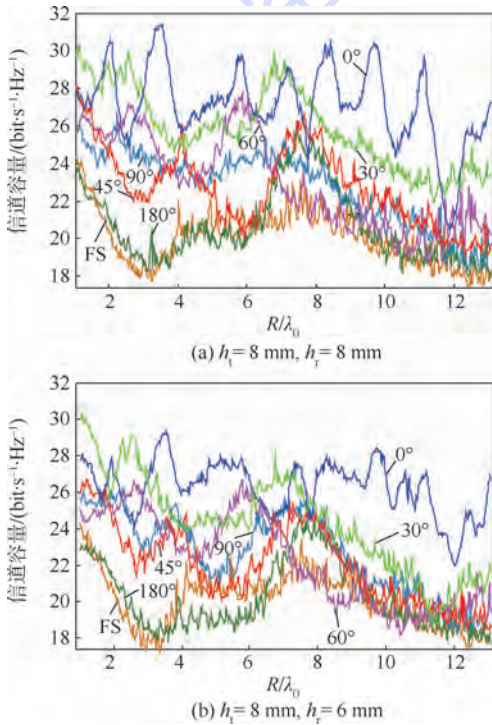


图 5 HPMIMO 信道容量与  $R$  的关系

Fig. 5 Relation between channel capacity and  $R$  of HPMIMO

根据式 (1) 和式 (2), 可得到信噪比为 20 dB 时, 单极化收发天线所对应的信道容量为  $6.66 \text{ bit} \cdot \text{s}^{-1} \cdot \text{Hz}^{-1}$ . 在图 6 中以  $\theta = 0^\circ$  和  $180^\circ$  为例, 给出了当 Tx 位于坐标原点, Rx 的  $x$  坐标固定

时, 6 极化相对于单极化系统信道容量增益随 Rx 的  $y, z$  坐标变化的情况. 从图 6 中可以看到, 信道

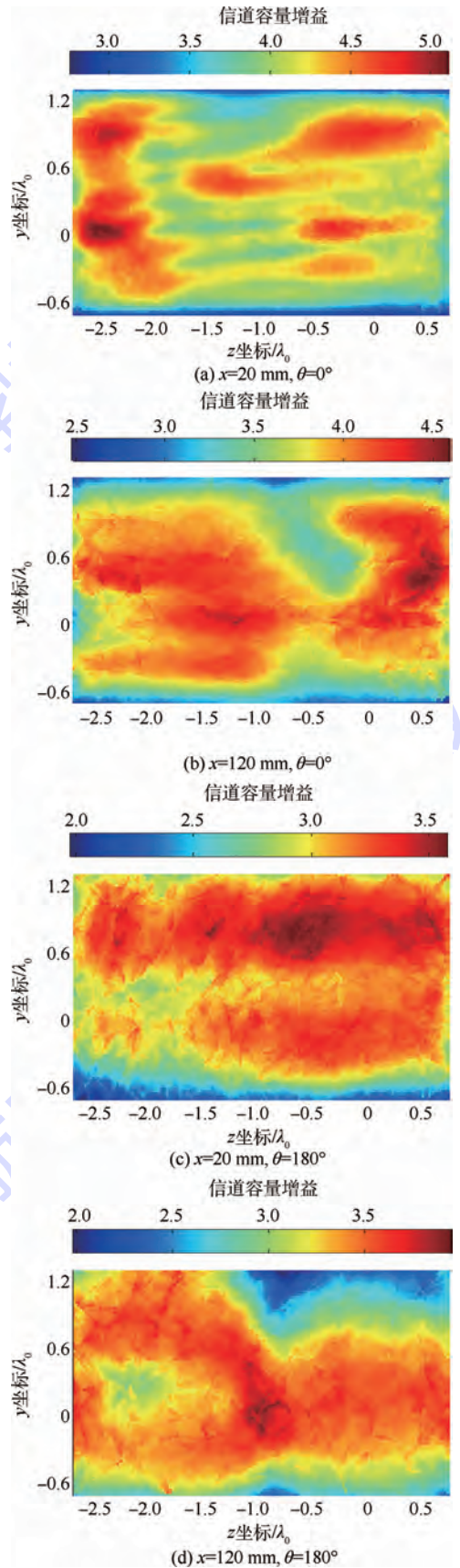


图 6 信道容量增益空间分布图

Fig. 6 Spatial distribution diagram of channel capacity gain

容量在空间中呈现出较强的不均匀性,而且对于不同的水平间距,其分布情况也有明显的区别.在双平行板信道中,HPMIMO 信道容量的增益在 2.47 ~ 5.10 之间.此外,双平行板信道中的增益要明显大于单反射面信道.为了更清楚地描述信道容量增益在图 6 中这 4 种不同环境下的空间统计特性,在图 7 中对这些情况下信道容量增益的累积分布函数 (Cumulative Distribution Function, CDF) 进行了比较.

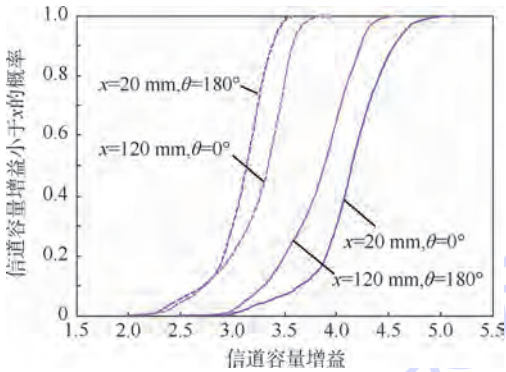


图 7 HPMIMO 系统相对于单极化系统信道容量的 CDF  
Fig. 7 Capacity gain CDF of HPMIMO system over single-polarized system

从图 7 中可以看到,双平板信道中的 HPMIMO 信道容量增益约为单反射面信道中的 1.3 倍.考虑 10% 中断概率,以上这 4 种情况下对应的增益分别为 3.65、3.30、2.72 和 2.73.

### 2.3 特征值

通过奇异值分解 (SVD), MIMO 信道矩阵  $\mathbf{H}$  可以被分解为一组特征信道<sup>[19]</sup>. 如果信道矩阵  $\mathbf{H}$  的奇异值被表示成  $\lambda_1, \lambda_2, \dots, \lambda_M$ , 其中  $M$  是  $\mathbf{H}$  的秩, 则信道容量表达式 (1) 也可以写为

$$C = \sum_{i=1}^M \text{lb}[1 + (\text{SNR}/n_T)\lambda_i^2] \quad (3)$$

其中:每个非零的奇异值  $\lambda_i$  对应于一个特征信道,每个特征信道可以支持一路数据流.每个奇异值的平方被称为矩阵  $\mathbf{H}\mathbf{H}^\dagger$  的特征值.通过式 (3) 可以清楚地体现每个特征值对信道容量的贡献.在图 8 中,以  $\theta = 0^\circ$  和  $\theta = 180^\circ$  为例,描述了特征值随距离的变化,其中  $h_t = 8 \text{ mm}$ ,  $h_r = 8 \text{ mm}$ .从图 8 中可见,在这两种情况下,HPMIMO 信道中都存在 6 个非零的特征值,对应于图 8 (a) 和图 8 (b) 中的 6 条曲线,但在自由空间信道中只有 2 个特征值较大,其他 4 个都只是略大于 0,而在双平行板信道中,这 4 个特征值都明显增加了,从而获得的信道容量增益更大一些.

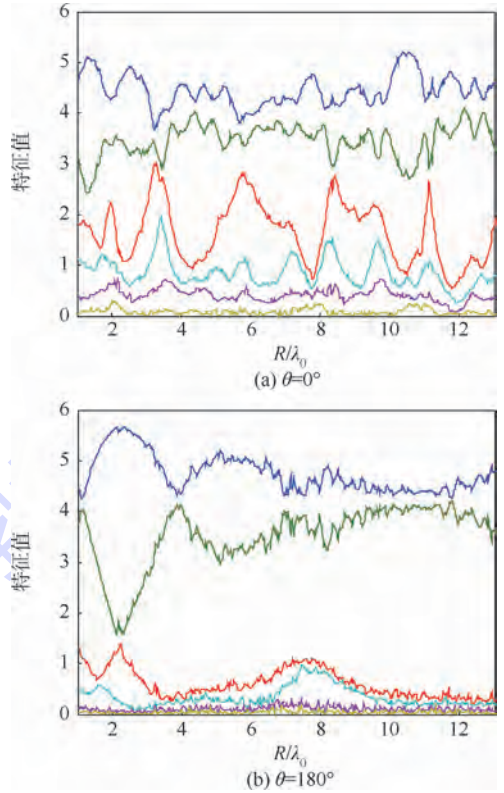


图 8 特征值随 R 的变化  
Fig. 8 Eigenvalues changing with R

### 3 结论

本文从信道容量和特征值的数值计算结果中可以得到:①在夹角不同的双反射面信道和自由空间信道中,HPMIMO 系统都可以获得 6 个非零的特征值,但在多径较少的环境中,有些特征值会接近于 0,因此对总信道容量的贡献很小;②随着双反射面间夹角度数的增加,信道中多径的数目逐渐减小,对应的 HPMIMO 信道容量也逐渐减小.夹角为  $180^\circ$  的单反射面信道中的信道容量最小,接近于自由空间信道,夹角为  $0^\circ$  的双平行板信道中的信道容量最大,可以达到单反射面信道中的 1.3 ~ 1.45 倍;③与单极化信道情况相比,HPMIMO 系统可以获得的信道容量增益最大值为 5.1,考虑 10% 中断概率,在双平行板信道中可获得的增益接近 4 倍,而在单反射面信道和自由空间信道中接近 3 倍.

### 参考文献 (References)

[1] Andrews M R, Mitra P, de Carvalho R. Tripling the capacity of wireless communications using electromagnetic polarization[J]. Nature, 2001, 409: 316-318.  
[2] Svantesson T, Jensen M A, Wallace J W. Analysis of electromagnetic field polarizations in multi-antenna systems[J]. IEEE Transactions on Wireless Communications, 2004, 3(2): 641-646.  
[3] Poon S Y, Tse D N C. Degree-of-freedom gain from using polari-

- metric antenna elements [J]. *IEEE Transactions on Information Theory*, 2011, 57(9):5695-5709.
- [4] 黄丘林, 郭万有, 史小卫. 极化分集时 MIMO 系统的信道容量 [J]. *电子与信息学报*, 2006, 28(8):1443-1446.  
Huang Q L, Guo W Y, Shi X W. Channel capacity of MIMO wireless systems in the presence of polarization diversity [J]. *Journal of Electronics & Information Technology*, 2006, 28(8):1443-1446 (in Chinese).
- [5] 伍裕江, 聂在平. 多天线系统中的多维极化分集性能分析 [J]. *电波科学学报*, 2007, 22(3):365-369.  
Wu Y J, Nie Z P. Performance analysis of multi-polarization diversity in MIMO wireless communications [J]. *Chinese Journal of Radio Science*, 2007, 22(3):365-369 (in Chinese).
- [6] Elnaggar M S, Chaudhuri S K, Safavi-Naeini S. Multi-polarization dimensionality of multi-antenna systems [J]. *Progress in Electromagnetics Research B*, 2009, 14:45-63.
- [7] Piao D. Characteristics of the hexapolarized MIMO channel over free-space and three non-free-space scenarios [J]. *IEEE Transactions on Wireless Communications*, 2013, 12(8):4174-4182.
- [8] Xu J. Field dimensions and polarization gains in electromagnetic systems [J]. *IEEE Transactions on Antennas Propagation*, 2012, 60(11):5336-5347.
- [9] Stancil D, Berson A, Hof J P V, et al. Doubling wireless channel capacity using co-polarised, co-located electric and magnetic dipoles [J]. *Electronics Letters*, 2002, 38(5):746-747.
- [10] Gupta G, Hughes B, Lazzi G. On the degrees of freedom in linear array systems with tri-polarized antennas [J]. *IEEE Transactions on Wireless Communications*, 2008, 7(7):2458-2462.
- [11] Konanur S, Gosalia K, Krishnamurthy S H, et al. Increasing wireless channel capacity through MIMO systems employing co-located antennas [J]. *IEEE Transactions on Microwave Theory and Techniques*, 2005, 53(6):1837-1844.
- [12] 王鹏, 郑剑峰, 高旭, 等. 室内信道中多极化 MIMO 系统的容量 [J]. *清华大学学报:自然科学版*, 2011, 51(7):993-998.  
Wang P, Zheng J F, Gao X, et al. Capacity of a multi-polarized MIMO system in an indoor channel [J]. *Journal of Tsinghua University: Science & Technology*, 2011, 51(7):993-998.
- [13] Tian R, Lau B K. Experimental verification of degrees of freedom for colocated antennas in wireless channels [J]. *IEEE Transactions on Antennas Propagation*, 2012, 60(7):3416-3423.
- [14] Piao D, Mao Y, Liu L, et al. Effect of antenna circuit distribution on the characteristics of hexapolarized MIMO system over free space and ground reflected channel [C] // *Proceedings of the 2nd International Conference on Connected Vehicles & Expo (ICCVE 2013)*. Piscataway, NJ: IEEE Press, 2013:526-530.
- [15] Piao D, Yang L. MIMO capacity comparisons of three types of colocated dual-polarized loop antennas [J]. *Progress In Electromagnetics Research C*, 2014, 54:41-48.
- [16] Rappaport T S, Shu S, Mayzus R, et al. Millimeter wave mobile communications for 5G cellular: It will work! [J]. *IEEE Access*, 2013, 1:335-349.
- [17] Wei K, Zhang Z, Feng Z. Design of a wideband horizontally polarized omnidirectional printed loop antenna [J]. *IEEE Antennas Wireless Propagation Letters*, 2012, 11(1):49-52.
- [18] Foschini G J, Gans M J. On the limits of wireless communications in a fading environment when using multiple antennas [J]. *Wireless Personal Communications*, 1998, 6(3):311-335.
- [19] Telatar I E. Capacity of multi-antenna Gaussian channels [J]. *European Transactions on Telecommunications*, 1999, 10(6):585-595.

## Characteristics of hexapolarized MIMO channel in two-mirror channel

PIAO Dazhi\*, CHEN Peng

(College of Information Engineering, Communication University of China, Beijing 100024, China)

**Abstract:** To study the effect of multipath richness in the two-mirror channel on the characteristics of the hexapolarized multiple-input multiple-output (HPMIMO) channel, a co-located and orthogonal dual-polarized antenna was designed, composed of an electric dipole and an electric loop working at 28 GHz, based on which a virtual HPMIMO antenna was constructed. From the computation results, it is seen that with the reduction of the angle between the two reflection planes, the richness of the multipath increases gradually, and the capacity gain of the HPMIMO channel over the single-polarized channel gets large correspondingly, which can achieve 2.47–5.10. If the angle between the two planes is  $0^\circ$ , the capacity of the HPMIMO in the two-mirror channel is larger than that in the free space (FS) and other channel with little multipath by 30%–45%. Furthermore, in both of the FS channel and the two-mirror channel, the HPMIMO channel matrix can obtain 6 non-zero eigenvalues, but some of them are very small and may approach zero if the multipath are not sufficiently rich. Thus, in order to get a high multiplexing and diversity gain, the multi-polarized MIMO antenna should be used in a rich multipath environment.

**Key words:** polarized; multiple-input multiple-output (MIMO); channel capacity; eigenvalue; two-mirror channel

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0783

# 基于有限元方法的场磨式电场传感器标定装置优化设计



崔勇<sup>\*1</sup>, 袁海文<sup>1</sup>, 赵录兴<sup>2</sup>, 胡多<sup>3</sup>

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100191; 2. 中国电力科学研究院 高电压所, 北京 100192;  
3. 国网山西省电力公司 检修分公司, 太原 030000)

**摘 要:** 场磨式电场传感器常被用来测量特高压直流(HVDC)输电线路下地面合成电场强度. 传统对其标定的方法存在标定装置体积较大、户外标定移动不方便的缺点. 针对此问题研究了一种应用于特高压直流输电线路下地面合成电场测量的电场传感器便携式缩尺标定装置. 基于有限元方法建立了传感器的三维电场模型, 基于该模型对便携式标定装置的尺寸、结构等关键参数做出了模拟分析, 在此基础上进行了便携式标定装置的结构参数的优化设计. 将便携式缩尺标定装置与标准的标定装置进行了实验比较. 仿真及实验结果表明该标定装置能够方便、准确地对场磨式电场传感器进行标定.

**关键词:** 特高压直流(HVDC)输电; 场磨式电场传感器; 标定装置; 有限元分析; 电场测量

**中图分类号:** TM726.1; V248.1; TM937.1

**文献标识码:** A      **文章编号:** 1001-5965(2015)10-1807-06

发展建设特高压直流输电线路作为国家重要战略政策, 得到了快速的发展<sup>[1-2]</sup>. 而特高压直流输电线路周边的电磁问题日益得到人们的重视, 在线路的建造、调试和运行阶段, 都需要对线路的电磁环境进行检测. 可听噪声、无线电干扰、合成电场等物理量被用来衡量电磁水平<sup>[3-5]</sup>.

场磨式电场传感器大量地被用于测量输电线路下方的地面合成电场<sup>[6-10]</sup>. 除此之外, 场磨式电场传感器在航空航天、大气电场测量与监测<sup>[11-13]</sup>、高电压非接触式测量<sup>[14]</sup>等其他领域也得到了广泛应用. 根据 IEEE Std 1227TM—1990 (R2010)<sup>[15]</sup> 标准的规定, 场磨式电场传感器放置在由两个平行电极产生的均匀场中, 由于需要产

生均匀电场, 同时考虑到传感器具有一定的体积, 因而需要较大的平行板面积和足够尺寸的平行板间距. 两极板间距的增加, 一方面使标定装置体积变大, 另一方面要求产生电场用的直流电源能输出较高电压, 这都会使标定装置整体变得庞大, 不适合移动标定的使用<sup>[16]</sup>. 直流输电线路下地面合成电场的测量, 经常需要电场传感器在偏远地区野外长时间地工作. 传感器经过长时间运输或长时间连续工作后, 再次使用前, 需要对其进行重新标定. 为此, 传统的大尺寸标定装置存在着诸多不便, 给合成电场的准确测量会带来不利影响. 因此, 非常需要有一种能够便于携带、适用于在野外对相应电场传感器进行准确标定的装置. 基于此,

收稿日期: 2014-12-12; 录用日期: 2015-02-13; 网络出版时间: 2015-04-02 09:50

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150402.0950.005.html

基金项目: 中央高校基本科研业务费专项资金(YWF-14-ZDHXY-13); 国家“973”计划(2011CB209402); 国家自然科学基金(61273165)

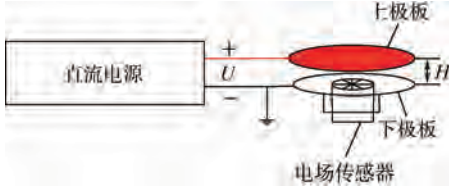
\* 通讯作者: 崔勇(1982—), 男, 河南漯河人, 讲师, cuiyong@buaa.edu.cn, 主要研究方向为电磁场测量、无线传感器网络.

**引用格式:** 崔勇, 袁海文, 赵录兴, 等. 基于有限元方法的场磨式电场传感器标定装置优化设计[J]. 北京航空航天大学学报, 2015, 41(10): 1807-1812. Cui Y, Yuan H W, Zhao L X, et al. Optimum design of calibration device for field mill type electric field sensor based on finite element method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1807-1812 (in Chinese).

本文研发出一种适合于便携使用的小型电场传感器标定装置.

### 1 设计原理

场磨式电场传感器标定装置由两块上下相互平行的金属极板组成,如图 1 所示.



U, H—上、下极板间的电压和距离.

图 1 场磨式电场传感器的标定装置

Fig. 1 Calibration device for field mill type electric field sensor

在上、下极板间施加一定的电压 U(由外接的直流电源提供),两极板间的电场强度可表征为

$$E_0 = \frac{U}{H} \quad (1)$$

根据式(1),可以计算出两极板间电场强度 E<sub>0</sub> 的理论值. 将被检的电场传感器放入该标定装置中,施加不同大小的直流电压 U,可以获得不同的电场强度的理论值,如此,就建立了输出电压与电场强度理论值 E<sub>0</sub> 的关系,从而可计算出电场传感器的灵敏度. 尽管式(1)中,电场强度的理论值 E<sub>0</sub> 仅与极板间距 H 有关,但由于电场会在上、下极板边缘处发生畸变,使得电场传感器在标定装置中感应到的电场强度的实际值与电场值之间有一定的差距. 只有当极板长度与极板间距满足一定条件时<sup>[15]</sup>,放置在两极板中心处的电场传感器测得的电场强度才最接近电场强度的理论值.

本文工作中设计的便携式电场传感器标定装置,其基本结构由两个平行极板构成,标定装置的下极板开有圆孔,并采用特殊夹具固定被检电场传感器. 被检电场传感器的动片与标定装置的下极板平齐,使得被检电场传感器无需进入标定装置的上、下极板之间的空间,即可感应到其电场.

### 2 结构参数的有限元分析

基于有限元的相关理论,首先对标定装置的机械结构建立模型,如图 2 所示. 然后,对几何模型进行单元剖分、加载,可求解出标定装置两极板间的电场分布情况,如图 3 和图 4 所示. 根据求得的电场分布情况,可进行标定装置结构参数的设计. 在计算求解过程中,改变加载在两极板间的电

压,使两极板间形成的电场强度的理论值始终为 20 kV/m. 被标定的场磨式电场传感器外壳直径 8 cm,感应片直径 6 cm,传感器外壳与标定装置的下极板接触.

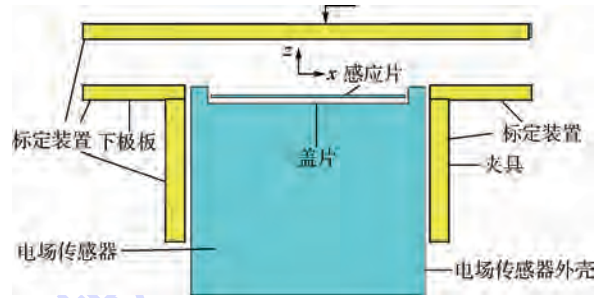


图 2 标定装置和被检电场传感器的几何模型

Fig. 2 Geometric model of calibration device and tested electric field sensor

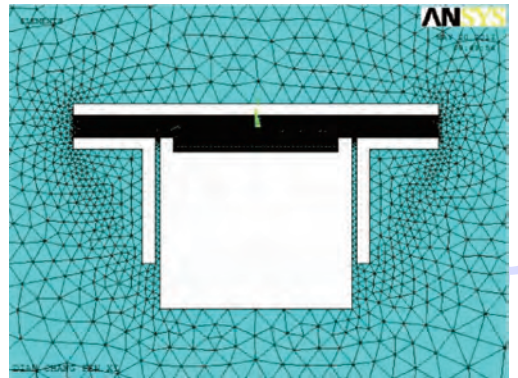


图 3 标定装置的剖分图

Fig. 3 Mesh diagram of calibration device

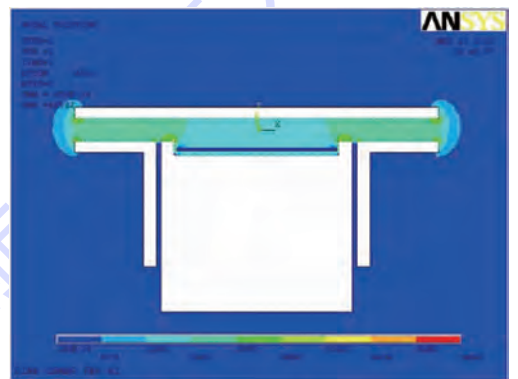


图 4 标定装置极板间电场分布图

Fig. 4 Distribution diagram of electric field between two plates of calibration device

#### 2.1 极板间距和极板直径对电场的影响

在标定装置的设计上,受限于被检电场传感器的尺寸,以及要考虑标定装置的便携性,把标定装置的极板直径 L 固定为 16 cm. 在 L 固定的条件下,分析 H 对 E 的影响,并以此确定极板间距 H. 依照图 2 所建立的模型,取 H 值分别为 1、2、3、4 和 5 cm,可得到如图 5 所示的结果. 由图 5 可以看

出,在极板尺寸( $L = 16\text{ cm}$ )、以及电场强度理论值  $20\text{ kV/m}$  不变的条件下,随着极板间距  $H$  的增加,极板间电场强度的实际值逐渐接近电场强度的理论值  $20\text{ kV/m}$ . 同时,在感应片的敏感范围( $x < 0.03\text{ m}$ )内,电场强度并非恒定值,而是随着与标定装置中心距离的增加发生了畸变. 图6为极板间电场强度实际值的畸变情况. 理想情况下,在感应片的敏感范围内,电场强度应保持不变,但由于标定装置中极板边缘效应的存在,使得感应片敏感区域内的电场不是一个恒定值,距离电场传感器的外壳越近,畸变程度越大. 定义在感应片敏感范围( $x < 0.03\text{ m}$ )内各个位置处电场强度的平均值与理论值之比为电场强度的畸变率,并用该值来衡量电场强度的变化程度. 畸变率越小,说明所产生的电场越接近均匀分布. 从图6可以看出,在两极板间距离  $H = 5\text{ cm}$  时,电场的畸变最小. 综合图5和图6可以看出, $H = 5\text{ cm}$  时,电场强度的实际值更接近理论计算值,同时电场的畸变率也较低.

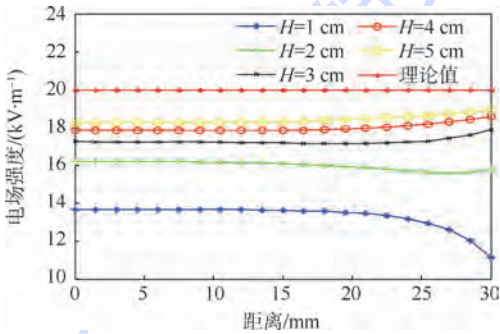


图5 不同  $H$  下的标定装置极板间的电场分布

Fig. 5 Distribution of electric field between two plates of calibration device with different  $H$

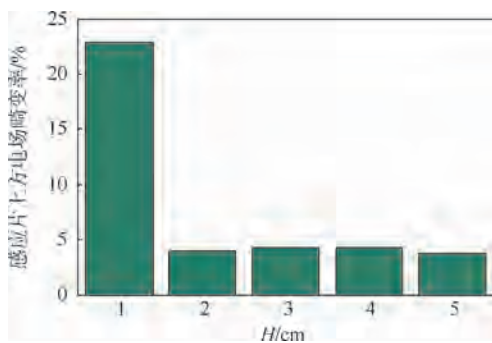


图6 不同  $H$  下电场传感器感应片上方电场的畸变

Fig. 6 Distortion of electric field over sensing plate between two plates of calibration device with different  $H$

如果固定极板间距  $H = 5\text{ cm}$  不变,使极板直径  $L$  分别等于  $16\text{ cm}$  和  $32\text{ cm}$  时,感应片上方的电场分布情况如图7所示,可以看出  $L$  的增加对降低极板间电场畸变的作用是微弱的. 这与通常认

为的  $H/L$  的比值越小,极板间电场的畸变越小似乎有所矛盾. 这是因为,如果在  $H/L$  小于  $0.5$  或更小的范围内,电场强度的畸变率都非常小,可认为  $H/L$  在此范围内的变化不会对电场造成影响. 在上述仿真中, $H/L$  的范围为  $0.031 \sim 0.156$  ( $5/16 \sim 5/32$ ),远小于  $0.5$ . 因而  $L$  变化造成的  $H/L$  的比值变化对电场畸变的影响基本可以忽略. 这就解释了图7中改变  $H/L$  比值后的两条曲线基本重合的原因.

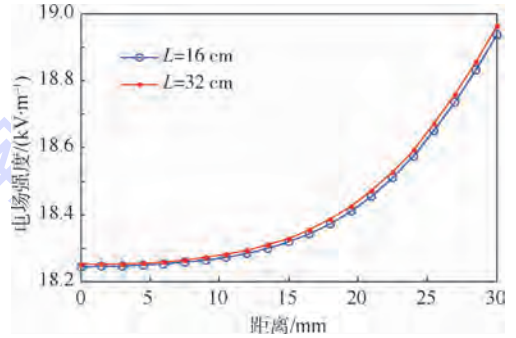


图7  $H = 5\text{ cm}$  时不同直径极板的电场分布

Fig. 7 Distribution of electric field with different  $L$  when  $H = 5\text{ cm}$

综上,在极板直径固定为  $16\text{ cm}$ 、极板间距为  $5\text{ cm}$  时,电场强度的实际值与理论值最为接近,且在电场传感器感应片感应区域内电场的畸变最小. 同时,在保证  $H/L$  小于  $0.5$  的条件下,极板直径  $L$  对实际电场的实际影响非常小.

## 2.2 传感器外壳与标定装置的相对位置

当标定装置与被检电场传感器配合不好时,容易使被检电场传感器相对于标定装置发生倾斜,标定装置相对传感器倾斜时的几何模型如图8所示. 模型中,极板直径为  $16\text{ cm}$ ,极板间距为  $1\text{ cm}$ ,倾斜角度为  $1.5^\circ$ . 标定装置的倾斜,会对被检电场传感器感应片上方的电场分布造成较大影响. 图9是基于图8的倾斜模型计算得到的感应片上方的电场强度的横向分布. 由于相对倾斜后,模型不再对称,因此分析了整个感应片上方 ( $-3 \sim 3\text{ cm}$ ) 的电场强度的横向分布,并将结果与没有相对倾斜时的感应片上方电场分布作了比较. 由图9可以看出,被检电场传感器与标定装置在相对倾斜角为  $1.5^\circ$  时的电场的畸变情况,比没有相对倾斜时严重. 有相对倾斜时,感应片上方电场分布更加不均匀,因而被检电场传感器与标定装置间的相对倾斜会对标定结果产生较大影响. 在标定装置设计中,应使标定装置与被检电场传感器的外壳的直径尽可能接近(极限情况是外径与孔径的差值为零),以使得两者紧密结触,从而

保证被检电场传感器与标定装置之间不会发生相对倾斜。

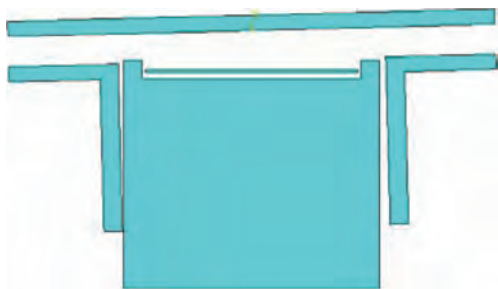


图8 标定装置相对传感器倾斜时的几何模型

Fig. 8 Geometric model when calibration device inclined relative to sensor

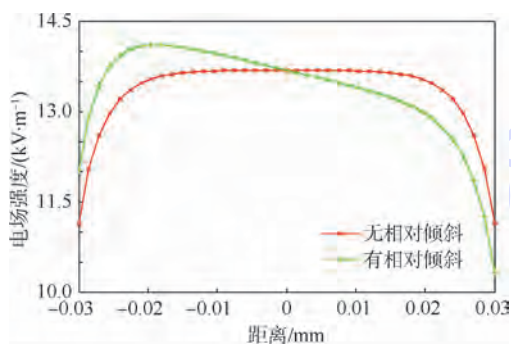


图9 标定装置与传感器间有相对倾斜时与没有相对倾斜时的电场分布对比

Fig. 9 Electric field distribution when calibration device and sensor with and without relative incline

此外,被检电场传感器的感应片与标定装置下极板的相对位置必须固定.为方便设计,可使被检电场传感器旋转的动片与标定装置下极板处于同一平面.如果两者位置不固定,则感应片所感应的电场也会发生变化,这对标定结果也会产生影响.对被检电场传感器动片(屏蔽片)与下极板平齐,被检电场传感器动片向上偏离下极板1 mm(高度误差1 mm),以及被检电场传感器动片向下偏离下极板1 mm(高度误差-1 mm)这3种情况下,被检电场传感器感应片上方的电场分布进行仿真,3种模型中,极板直径均为16 cm,极板间距均为1 cm.电场传感器与标定装置相对位置变化时感应片上方电场分布的变化仿真结果如图10所示.由图10可以看出,传感器动片与下极板位置的改变,会对电场的分布造成影响,位置误差1 mm和位置误差-1 mm所引起的电场强度的最大变化分别为7.3%和5.9%.仿真结果表明,标定装置对位置误差比较敏感.这是因为被检电场传感器的动片是与下极板一起接地的,动片与开有孔的下极板共同构成了标定装置的下极板.被检电场传感与下极板相对位置的变化相当于标定

装置极板的间隔发生了变化.因此,在设计标定装置时,必须保证被检电场传感器动片与下极板的相对位置保持不变.

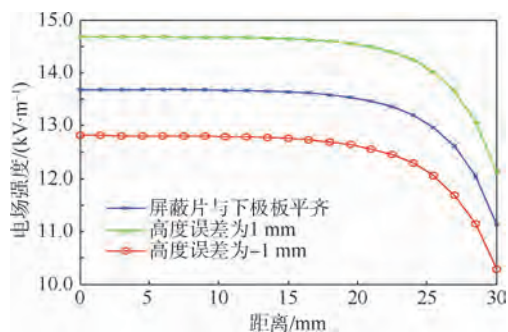


图10 电场传感器与标定装置相对位置变化时感应片上方电场分布的变化

Fig. 10 Change of electric field distribution over sensing plate with different relative location between electric field sensor and calibration device

### 3 实验结果与分析

基于前述仿真分析,本文工作中设计的标定装置的参数如下: $L = 16 \text{ cm}$ , $H = 5 \text{ cm}$ ,极板材料为钢,便携式电场传感器标定装置实物如图11所示.



图11 便携式电场传感器标定装置

Fig. 11 Portable calibration device for electric field sensor

当输出为 $-3 \sim +3 \text{ kV}$ 的可调直流电源加在两极板上时,两极板间的电场强度理论值的范围为 $-60 \sim +60 \text{ kV/m}$ .使用在标准电场标定装置如图12所示,图中标定好的电场传感器测量本文所设计的便携式标定装置中的实际电场,便携式标定装置产生的实际电场值结果如图13所示.从图13中可以看出,实测电场强度与所加电源电压之间有良好的线性关系,同时,实测电场小于理论电场,两者的比值约为0.92,这与图5给出的仿真结果吻合.在野外的实际标定过程中,保持被检电场传感器与标定装置的位置不变,使得电场强

度理论值与实际值的比值保持不变,在此基础上,可以通过加在两极板间的电压计算出电场强度的理论值,进而可根据图13所示关系,计算出电场强度的实际值.然后,通过电场强度实际值与被检电场传感器输出值两者间的关系,计算出被检电场传感器的灵敏度,实现对被检电场传感器的标定.



图12 标准的电场传感器标定装置

Fig.12 Standard calibration device for electric field sensor

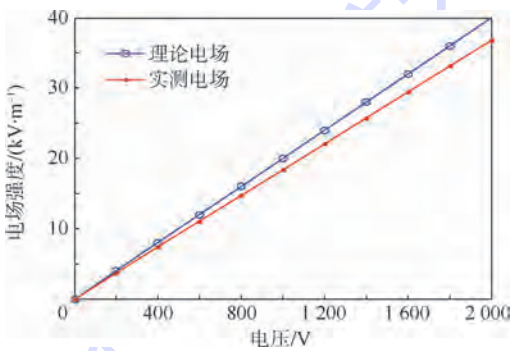


图13 便携式标定装置产生的实际电场强度

Fig.13 Real electric field strength of portable calibration device

经过较长时间的现场使用,所研发的便携式标定装置能够方便、快捷地对场磨式电场传感器进行校准.目前,该校准装置已经应用于中国电力科学研究院特高压直流实验基地中特高压直流输电线路地面合成电场测量系统中,并已取得了良好的效果.

## 4 结论

1) 本文给出了标定装置及电场传感器的有限元模型,基于该模型,采用有限元方法对电场传感器标定装置进行了结构参数的分析和设计,所设计的标定装置克服了传统标定装置体积大、不便于移动的缺点.

2) 标定装置两极板间距和极板直径对电场的影响较大,当两极板间距离为5 cm时,产生的

电场的畸变率较小.

3) 被检电场传感器与标定装置的相对位置对标定结果影响较大,在实际使用中,两者的相对位置必须固定.

本文工作的研究成果,可为野外长时间地面合成电场测量和监控工作提供良好的基础条件和保障.

## 参考文献 (References)

- [1] 刘振亚. 特高压直流输电工程电磁环境[M]. 北京:中国电力出版社,2009:14-20.  
Liu Z Y. UHVDC transmission engineering electromagnetic environment[M]. Beijing: China Electric Power Press, 2009: 14-20 (in Chinese).
- [2] 陆家榆,鞠勇.  $\pm 800$  kV 直流输电线路电磁环境限值研究[J]. 中国电力,2006,39(10):37-42.  
Lu J Y, Ju Y.  $\pm 800$  kV DC transmission line electromagnetic environment research[J]. China Power, 2006, 39(10): 37-42 (in Chinese).
- [3] 田丰,余占清,曾嵘,等. 高海拔地区 UHVDC 输电线路电晕可听噪声影响因素分析[J]. 高电压技术,2012,38(12):3125-3132.  
Tian F, Yu Z Q, Zeng R, et al. Influence factors of audible noise caused by corona discharge on UHVDC transmission lines under high altitude condition[J]. High Voltage Engineering, 2012, 38(12): 3125-3132 (in Chinese).
- [4] 邓军,肖瑶,楚金伟,等. 云南-广东  $\pm 800$  kV 特高压直流线路无线电干扰仿真计算与测试分析[J]. 高电压技术,2013,39(3):597-604.  
Deng J, Xiao Y, Chu J W, et al. Analysis of simulative calculation and measurement of radio interference from yunnan-guangdong  $\pm 800$  kV UHVDC transmission lines[J]. High Voltage Engineering, 2013, 39(3): 597-604 (in Chinese).
- [5] 崔翔,周象贤,卢铁兵. 高压直流输电线路离子流场计算方法研究进展[J]. 中国电机工程学报,2012,32(36):130-141.  
Cui X, Zhou X X, Lu T B. Recent progress in the calculation methods of ion flow field of HVDC transmission lines[J]. Proceedings of the CSEE, 2012, 32(36): 130-141 (in Chinese).
- [6] 张建功,马士新,张广洲,等. 直流电场测量装置研制[J]. 高电压技术,2009,35(12):3027-3030.  
Zhang J G, Ma S X, Zhang G Z, et al. Development of measurement device for DC electric-field strength[J]. High Voltage Engineering, 2009, 35(12): 3027-3030 (in Chinese).
- [7] 杨勇. 高压直流模拟试验短线段下地面3维合成电场的计算分析[J]. 高电压技术,2013,39(3):2843-2847.  
Yang Y. Calculation analysis on three-dimensional total electric field at ground level under short HVDC simulation test line section[J]. High Voltage Engineering, 2013, 39(3): 2843-2847 (in Chinese).
- [8] 张洪钊,袁海文,陆家榆,等. 交直流混合电场旋转式一体化测试仪的研制[J]. 电网技术,2012,36(12):182-188.  
Zhang H C, Yuan H W, Lu J Y, et al. Development of an integrative rotary test instrument for AC/DC hybrid electric field[J].

- Power System Technology, 2012, 36 ( 12 ): 182-188 ( in Chinese ).
- [ 9 ] Cui Y, Lv J, Yuan H, et al. Development of a wireless sensor network for distributed measurement of total electric field under HVDC transmission lines [ J ]. International Journal of Distributed Sensor Networks, 2014 : 850842.
- [ 10 ] Cui Y, Wang Q, Yuan H, et al. Relative localization in wireless sensor networks for measurement of electric fields under HVDC transmission lines [ J ]. Sensors, 2015, 15 ( 2 ): 3540-3564.
- [ 11 ] Fort A, Mugnaini M, Vignoli V, et al. Design, modeling, and test of a system for atmospheric electric field measurement [ J ]. IEEE Transactions on Instrumentation and Measurement, 2011, 60 ( 8 ): 2778-2785.
- [ 12 ] Bateman M G, Stewart M F, Podgorny S J, et al. A low-noise, microprocessor-controlled, internally digitizing rotating-vane electric field mill for airborne platforms [ J ]. Journal of Atmospheric and Oceanic Technology, 2007, 24 ( 7 ): 1245-1255.
- [ 13 ] 郑凤杰, 夏善红, 陈贤祥. 三维电场传感器仿真优化及性能测试 [ J ]. 传感技术学报, 2008, 21 ( 6 ): 946-950.
- Zheng F J, Xia S H, Chen X X. Simulation optimization and performance test of three dimensional electric field sensor [ J ]. Chinese Journal of Sensors and Actuators, 2008, 21 ( 6 ): 946-950 ( in Chinese ).
- [ 14 ] Tant P, Bolsens B, Sels T, et al. Design and application of a field mill as a high-voltage DC meter [ J ]. IEEE Transactions on Instrumentation and Measurement, 2007, 56 ( 4 ): 1459-1464.
- [ 15 ] IEEE Std 1227TM—1990 ( R2010 ). IEEE guide for the measurement of DC electric-field strength and ion related quantities [ S ]. Piscataway, NJ: IEEE Press, 1990: 1-50.
- [ 16 ] Selezneva A N. Instrument for calibrating means of measuring electric field strength [ J ]. Measurement Techniques, 1983, 26 ( 10 ): 850-852.

## Optimum design of calibration device for field mill type electric field sensor based on finite element method

CUI Yong<sup>\*1</sup>, YUAN Haiwen<sup>1</sup>, ZHAO Luxing<sup>2</sup>, HU Duo<sup>3</sup>

- ( 1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China;  
2. High Voltage Department, China Electric Power Research Institute, Beijing 100192, China;  
3. Maintenance Branch, State Grid Shanxi Electric Power Company, Taiyuan 030000, China )

**Abstract:** Field mill type electric field sensors are widely used to measure the total electric field at ground level under the high voltage direct current (HVDC) transmission lines. The size of the traditional calibration device is extremely large and it is not very convenient to move for outdoor measurement usage. To overcome this disadvantage, a new kind of reduced scale calibration device with preferable portability was elaborated applied in measurement of the total electric field at ground level under the high voltage direct current transmission lines. The finite element analysis method was employed to establish the three dimensional model of the electric field sensor calibration device. The numerical simulation analysis based on the finite element method was made with the respect to the dimension and structure and other key parameters for the portable calibration device. The structure parameter of the calibration device was optimally designed based on the simulation result. In addition, the reduced-scale portable calibration device was compared with the traditional calibration device by experiments. Based on the experimental and the simulation results, the proposed portable device demonstrates that it can accomplish the accurate and flexible demands of the electric field calibration.

**Key words:** HVDC power transmission; field mill type electric field sensors; calibration device; finite element analysis; electric field measurement

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0809

# 热效应布局下的缓冲器插入时序优化方法

王新胜<sup>1</sup>, 韩良<sup>\*1</sup>, 喻明艳<sup>2</sup>

(1. 哈尔滨工业大学 信息与电气工程学院, 威海 264209; 2. 浙江大学 宁波理工学院, 宁波 315100)



**摘 要:** 随着集成电路的集成度越来越高, 芯片的发热量越来越大且其内部温度呈不均匀分布, 这会影响关键路径的传播延时, 进而影响基于缓冲器插入的关键路径性能. 提出了一种考虑芯片热效应布局优化的缓冲器插入时序优化方法, 在版图设计的早期估计芯片的热分布和温度分布并且将其应用到版图布局优化和 RC 延时模型中. 同时利用模拟退火算法基于热分布调整并优化布局, 最后在最优布局下利用提出的缓冲器插入模型和快速插入算法进行时序优化. 仿真结果表明相对于不考虑温度效应布局优化的缓冲器插入方法, 缓冲器插入延时优化方法能有效降低最坏延时和缓冲器插入数目, 最坏延时比传统方法降低 9% ~ 18%, 比文献已经提出的最好方法降低 5% ~ 7%, 缓冲器插入数比其少 10 ~ 20 个.

**关键词:** 大规模集成电路; 布局规划; 缓冲器插入; 互连线; 模拟退火算法

**中图分类号:** TN405.97; TN402

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1813-08

随着互补金属氧化物半导体 (Complementary Metal Oxide Semiconductor, CMOS) 集成电路工艺的持续进步, 已经能够实现高度集成也就是在一个芯片内集成多个 IP 核实现系统级芯片 (System on Chip, SoC) 系统. 尽管 SoC 设计有很多优点如强大的功能、比较低的系统成本和比较快的上市时间; 但是也带来很多问题, 如功耗增加导致的热问题<sup>[1]</sup>, 热问题会使互连和单元电阻增加, 已经成为恶化芯片性能的主要因素. 因此, 解决芯片热问题及其对性能的影响在物理设计阶段越早越好.

热效应对缓冲器插入的影响已有文献对其研究<sup>[2-3]</sup>, 但主要说明热效应对缓冲器插入位置的影响, 即不考虑温度效应的缓冲器插入位于线的中间, 而考虑驱动单元的热效应缓冲器插入位置向负载端移动, 如果考虑线上的热效应会使插入缓冲器的位置向驱动端移动, 如果同时考虑驱动

端和线上的热效应插入缓冲器的位置也是向驱动端移动, 这是由于线上的热效应比驱动单元的热效应明显, 如果驱动一个长线可能增加额外的缓冲器. 为了估计由于热效应额外插入的缓冲器必须在布局布线之前知道热条件下的时序预算. 文献[4]已经开发一个在布局阶段估计片内温度分布的工具, 文献[5]提出通过 Hotspot 工具进行热条件下布局. 为了满足时序收敛, 基于温度分布的缓冲器插入时序优化技术也必须被开发. 根据最近研究结果, 43.6% 的缓冲器被插在块与块间的互连线上<sup>[6-7]</sup>. 基于以上分析, 存在的考虑温度效应的缓冲器插入技术很少并且不充分.

## 1 热条件下的布局规划

### 1.1 热条件下布局规划定义

热条件下布局规划问题是存在温度分布的条件下放置  $n$  个矩形块使其满足:

收稿日期: 2014-12-23; 录用日期: 2015-02-12; 网络出版时间: 2015-04-14 16:00

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150414.1600.002.html

基金项目: 国家自然科学基金(61201307)

作者简介: 王新胜(1978—), 男, 山东威海人, 讲师, xswang@hit.edu.cn

\* 通讯作者: 韩良(1969—), 男, 吉林榆树人, 副教授, hanliang@hit.edu.cn, 主要研究方向为大规模集成电路设计.

**引用格式:** 王新胜, 韩良, 喻明艳. 热效应布局下的缓冲器插入时序优化方法[J]. 北京航空航天大学学报, 2015, 41(10): 1813-1820. Wang X S, Han L, Yu M Y. Thermal aware floor planning timing optimal method for buffer insertion[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1813-1820 (in Chinese).

- 1) 每个模块  $i$  有一个固定的面积  $A_i$  并且其纵横比满足一定约束,也就是纵横比是有界的.
- 2) 芯片面积给定  $A = W \times H$ ,  $W$  为芯片的宽度,  $H$  为芯片的高度,其纵横比也是有界的.
- 3) 给定每个模块的能量消耗  $P_i$ ,其用于计算模块温度.
- 4) 布局规划的目标是低芯片面积、低总线长度  $L$  和低最高温度.

### 1.2 模拟退火算法

模拟退火是一种通用概率算法,用来在固定时间内在一个大的搜索空间内找到最优解<sup>[8]</sup>,其流程如图1所示.模拟退火算法与初值和算法迭代的起点无关,具有渐进收敛性,已在理论上证明是一种以概率为1收敛于全局最优的优化算法,并且具有并行性.

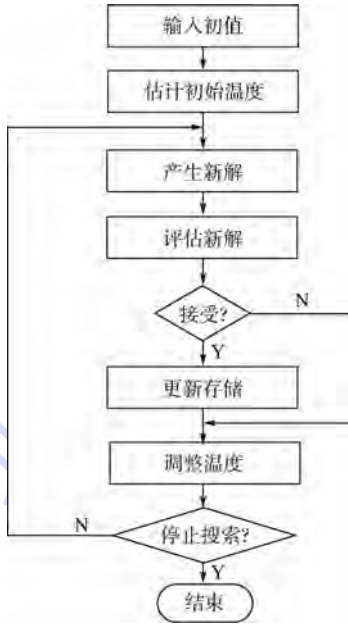


图1 模拟退火算法流程

Fig.1 Flow of simulated annealing algorithm

实验表明,初温越大,获得高质量解的几率越大,本文初始温度选择为  $T_{init} = 30\ 000$ . 假定一个设计有  $N$  个块,温度以一个因子  $\omega < 1$  速率降低可表示为

$$T_{current} = \omega \cdot T_{old} \quad (1)$$

式中:  $T_{current}$  为当前温度;  $T_{old}$  为之前温度.

本文降温方式为在初始阶段快速降温结束阶段降温很慢,也就是初始阶段  $\omega = 0.85$ ,结束阶段  $\omega = 0.98$ . 因此大部分退火时间花费在低温部分.

考虑热条件布局的目标函数可表示为

$$\Delta = \frac{C_A + C_{AR}}{2} \times A + \frac{C_W + C_{AR}}{2} \times W + C_T \times T \quad (2)$$

式中:  $C_A$ 、 $C_W$  和  $C_T$  分别为面积、线长和最高温度的权重;  $C_{AR} = 0.2$  为纵横比权重,并且  $C_A$ 、 $C_W$ 、 $C_{AR}$  和  $C_T$  之和总为1. 由于式(2)在不同模块单元都要被测量,因此会取得一个线性关系的目标函数. 为此,归一化此目标函数的每一项,具体为

$$\Delta A = \frac{(M_{area} - C_{area}) \cdot 1.2 \cdot T_{init}}{C_{area}} \quad (3)$$

$$\Delta W = \frac{(M_{wl} - C_{wl}) \cdot 1.2 \cdot T_{init}}{C_{wl}} \quad (4)$$

$$\Delta T = \frac{(M_{mt} - C_{mt}) \cdot 1.2 \cdot T_{init}}{C_{mt}} \quad (5)$$

式中:  $\Delta A$ 、 $\Delta W$  和  $\Delta T$  分别为归一化的面积、线长和最高温度;  $C_{area}$ 、 $C_{wl}$  和  $C_{mt}$  分别为当前解的面积、线长和最高温度;  $M_{area}$ 、 $M_{wl}$  和  $M_{mt}$  分别为修改后解的面积、线长和最高温度. 因此经过归一化后的目标函数可表示为

$$\Delta = \frac{C_A + C_{AR}}{2} \times \Delta A + \frac{C_W + C_{AR}}{2} \times \Delta W + C_T \times \Delta T \quad (6)$$

当  $\Delta$  为负值时表示相对于以前解的改进; 如果  $R_{0-1} < \exp\left(\frac{-\Delta \cdot M_{init}}{M_{curr}}\right)$  ( $M_{init}$  为初始的面积、线长和最高温度参数值;  $M_{curr}$  为当前上述相应参数值),  $\Delta$  可以为正值 ( $R_{0-1}$  为0到1之间的随机数), 这说明接受一个错误移动的概率在降低,也就是随着  $M_{curr}$  降低而降低.

### 1.3 HotSpot 热建模工具

HotSpot 能精确而快速进行热建模,主要用于处理器结构研究<sup>[9]</sup>. 针对输入布局图 HotSpot 产生等效热阻热容电路来计算温度分布. 本文主要利用其计算布局中所有块中的最高温度,此最高温度被用于模拟退火算法的优化布局,同时给出最优布局的温度分布用于缓冲器插入. 基于热条件布局的流程如图2所示.

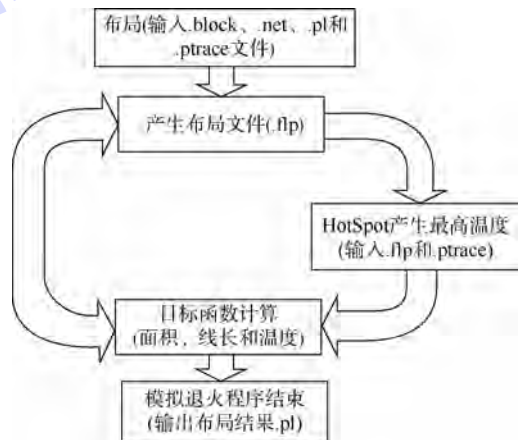


图2 基于热条件布局流程

Fig.2 Placement flow based on thermal condition

HotSpot 中有很多模型可用于温度计算<sup>[10]</sup>, 本文主要利用块模型即基本模型来计算温度. 该模型利用输入的功耗迹文件和布局文件, 产生相应块的温度迹分布. 本文进行修改使其根据温度迹文件和布局文件生成此布局的最高温度.

### 1.4 缓冲器插入

在一个驱动单元驱动一根长线或多个负载条件下, 缓冲器插入是有效减少延时的方法. 为了在布局阶段进行时序估计, 假定互连网络为块与块之间通过引脚互连, 连接最近的驱动端和负载端, 具体如图 3 所示<sup>[7]</sup>. 全局布线是通过互连所有内部块网络, 其中网络拓扑基于最小生成树被提取. 经典的时序优化是在给定布线树下, 从负载端到驱动端通过 van Ginneken<sup>[11]</sup> 算法找到优化的缓冲器插入位置和数目, 使其满足整个网络弛豫时间最大和负载电容最小. 但传统的缓冲器插入算法

没有考虑互连模块之间的热效应. 为了找到优化的缓冲器插入位置和数目, 延时模型必须包括线延时和驱动端延时部分. van Ginneken 算法用的线模型是  $\pi$  模型, 延时模型用的是一阶近似 Elmore 延时模型<sup>[12-13]</sup>. 为求节点  $i$  和节点  $j$  之间的延时, 下面分两种情况说明.

1) 插入线时的延时为

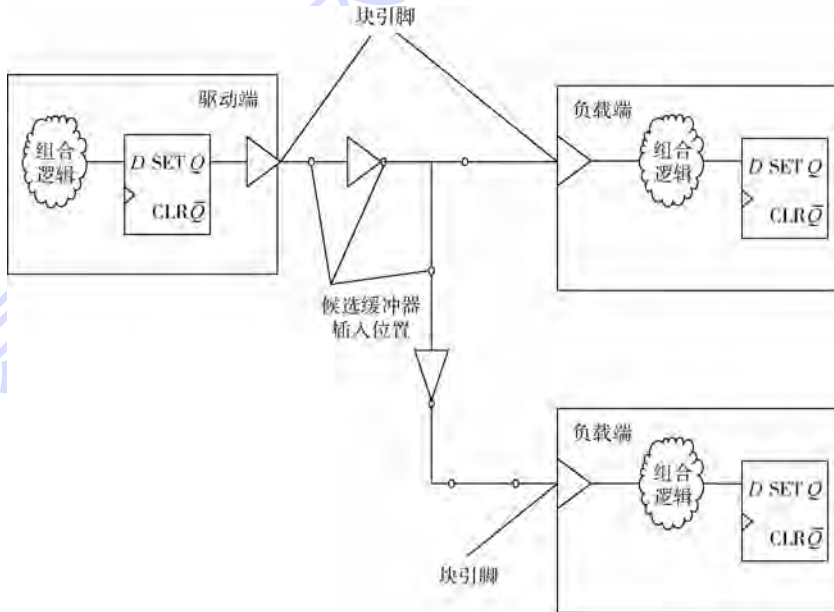
$$D_{\text{line}} = R \cdot \left( \frac{C}{2} + C(v_j) \right) \quad (7)$$

式中:  $R$  和  $C$  分别为节点  $i$  与节点  $j$  之间的线电阻和线电容;  $C(v_j)$  为节点  $j$  处的向下负载电容.

2) 如果一个缓冲器插入在  $i$  节点, 则插入缓冲器后的延时为

$$D_{\text{buf}} = R_B \cdot C(v_i) + D_B \quad (8)$$

式中:  $R_B$  和  $D_B$  分别为缓冲器的驱动电阻和内在延时;  $C(v_i)$  为  $i$  节点处的向下负载电容.



$D$ —输入端;  $Q$ —输出端;  $\bar{Q}$ —输出取反端; CLR—复位端; SET—置位端.

图 3 布局阶段缓冲器插入拓扑<sup>[7]</sup>

Fig. 3 Buffer insertion topology in placement stage<sup>[7]</sup>

## 2 热条件最优布局延时优化方法

大规模集成电路特别是 SoC 系统包含许多功能模块, 由于这些功能模块功耗不一致, 导致芯片内温度梯度分布. 在基于缓冲器插入优化延时过程中, 如果不考虑温度对互连线和驱动端的影响, 可能会错误地估计路径延时和所插入的缓冲器数目. 缓冲器插入数目不足可能引起关键路径时序违反, 因此在布局阶段就考虑温度对布局和缓冲器插入的影响显得十分必要. 下面分别从考虑温度延时建模和考虑温度最优布局两方面说明基于

热条件最优布局的延时模型.

首先给出插入互连线时考虑温度的延时模型. 互连电阻特性受温度影响的程度远大于电容, 文献[3-4]对此进行了说明, 因此本文忽略电容随温度变化. 互连线电阻随温度变化的模型可表示为

$$R_T = R_0 [1 + \alpha_1 (T - T_0)] \quad (9)$$

式中:  $R_0$  为温度为  $T_0 = 293 \text{ K}$  下的电阻;  $R_T$  为温度为  $T$  下的电阻;  $\alpha_1$  为电阻温度系数. 首先把布局平面划分成面积相等的小网格叫做热网格, 每个小网格内温度相同, 不同网格间温度不同, 具体

网格大小由待布局电路的最小模块大小和精度决定,具体如图4<sup>[7]</sup>所示.假定节点*i*和节点*j*之间连线经过*n*个热网格单元,则*i*和*j*两节点间的延时可表示为

$$D = \sum_{k=1}^n r_k l_k \left[ \frac{c}{2} \left( L - \sum_{k=1}^n l_{k-1} \right) + C(v_j) \right] \quad (10)$$

式中: $r_k$ 和 $l_k$ 分别是在第*k*个热网格单元的单位电阻和线长; $L$ 为节点*i*和节点*j*间的总线长; $c$ 为单位长度的线电容,其不随温度变化.

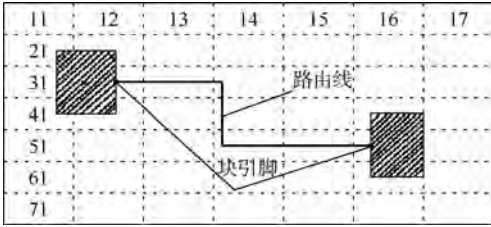


图4 互连线穿过热网格示意图<sup>[7]</sup>

Fig. 4 Sketch of interconnect cross thermal grid<sup>[7]</sup>

接下来给出插入缓冲器时考虑温度的延时模型.缓冲器受温度的影响主要改变其驱动电阻和内在延时<sup>[14]</sup>,具体为

$$R_{BT} = R_{B0} \cdot [1 + \delta(T - T_0)] \quad (11)$$

$$D_{BT} = D_{B0} \cdot [1 + \beta_{in}(T - T_0)] \quad (12)$$

式中: $R_{BT}$ 和 $R_{B0}$ 分别为温度*T*和0时缓冲器驱动电阻; $D_{BT}$ 和 $D_{B0}$ 分别为温度*T*和0时缓冲器内在延时; $\beta_{in} = S \cdot \tau$ 为缓冲器内在延时随温度变化的温度系数, $S$ 为缓冲器的大小;当*S*等于10倍到70倍最小尺寸时, $\delta$ 和 $\tau$ 为常数,其值分别为0.005和0.000 013.因此插入缓冲器时考虑温度的延时模型为

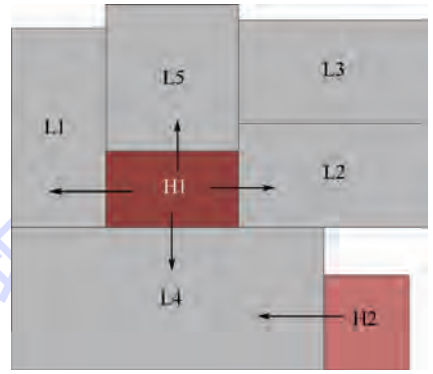
$$D_B = R_{BT} \cdot C(v_i) + D_{BT} \quad (13)$$

该模型充分考虑了热分布的局部特性,互连线的电阻和驱动器的输出电阻在不同位置上其大小随温度变化,基于热网格的方法能有效地加快考虑温度条件下缓冲器插入延时优化计算.

热条件下的最优布局是找到基于面积、线长和温度成本函数最优的布局.同时热条件下的布局也是一种降低芯片温度的一种有效方式,热量通过横向的临近衬底进行扩散<sup>[15]</sup>,也就是一个温度高的模块旁边紧邻一个温度低的模块,整体温度就较低,如图5所示,而图6布局温度就比较高,由于两个高温模块放在一起.具体布局理论已经在第1节叙述,在此不再累述.下面主要给出基于热条件最优布局的延时优化的实现流程.

基于热条件最优布局实现的延时优化如图7所示.输入网表文件来源于MCNC(Microelectronics Center of North Carolina)和GSRC(Gigascale Silicon

Research Center)标准测试电路<sup>[16]</sup>,首先基于热条件下的最优布局生成布局文件和热分布文件;接着调用考虑温度条件的缓冲器插入时序优化程序,输出优化结果.下面主要说明一下考虑温度分布的FBI(Fast Buffer Insertion)快速时序优化方法.



L1~L5—低温模块;H1,H2—高温模块.

图5 热块离得比较远布局

Fig. 5 Placement based on far from thermal block

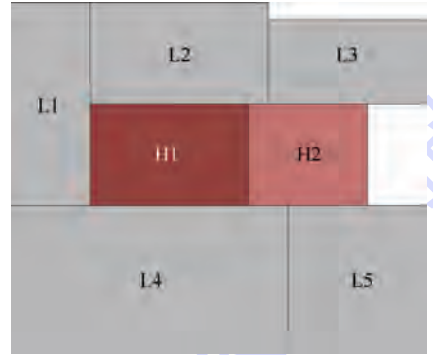


图6 热块离得比较近布局

Fig. 6 Placement based on near thermal block

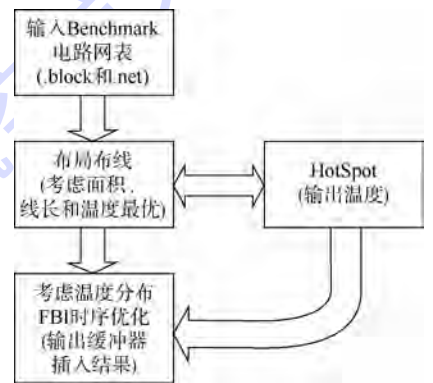


图7 热条件最优布局的延时优化结构

Fig. 7 Delay optimization structure in thermal optimal layout condition

首先假定一个给定的路由树  $T = (V, E)$ ,其中: $V$ 为所有节点, $E$ 为节点之间的连接弧.每个终端负载节点包括负载电容 $C_s$ 和要求到达时间(Required Arrival Time, RAT).整个路径的延时采用第1.4节的延时模型计算,在此不再累述.给定

任意节点,假定  $T(v)$  是节点  $v$  以下的子树,也就是节点  $v$  被看做根节点. 如果已经知道在  $T(v)$  中哪里插入缓冲器,用一个参数  $\alpha$  来标识,从节点  $v$  到终端负载节点  $s$  的延时表示为

$$D(v, s, \alpha) = \sum_{e=(v_i, v_j)} (D(v_i) + D(e)) \quad (14)$$

式中:求和为从  $v$  到  $s$  的所有连接弧. 如果节点  $v_i$  插入缓冲器,则  $D(v_i)$  是插入缓冲器的延时;如果  $v_i$  节点没有缓冲器插入则  $D(v_i) = 0$ . 因此从节点  $v$  到节点  $s$  的弛豫时间  $Q$  为

$$Q(v, \alpha) = \min_{s \in T(v)} \{ \text{RAT}(s) - D(v, s, \alpha) \} \quad (15)$$

缓冲器插入算法本质上可能从两方面改进:快速计算解方案和冗余解裁剪. FBI 算法在快速计算方面已有很大改进,本文主要从冗余裁剪方面进行改进,下面基于定理 1 给出详细说明.

**定理 1** 假设在节点  $v$  处,经过 van Ginneken 算法裁剪后剩下的解方案  $\alpha_i, \alpha_j$  和  $\alpha_k$ ,将其按  $Q$  值从小到大进行排列,如果  $\alpha_j$  位于  $\alpha_i$  和  $\alpha_k$  连线下与  $Q$  轴形成的区域内,则  $\alpha_j$  冗余.

**证明**  $\alpha_i, \alpha_j, \alpha_k$  和  $\alpha_m$  解方案分布如图 8 所示,  $\alpha_j$  位于  $\alpha_i$  和  $\alpha_k$  连线之下与  $Q$  轴形成的区域内,则其满足:

$$\frac{C_j - C_i}{Q_j - Q_i} > \frac{C_k - C_i}{Q_k - Q_i} \quad (16)$$

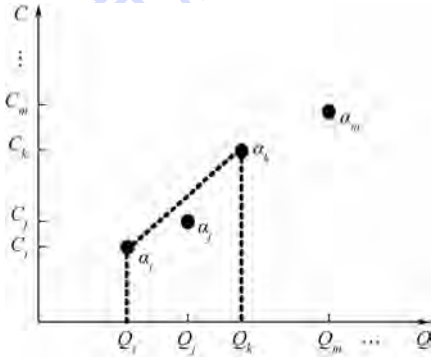


图 8 按  $Q$  值大小排列的解方案集合  
Fig. 8 Solution set based on  $Q$  value by size

假设  $v^*$  是离  $v$  最近插入缓冲器的节点并且靠近驱动节点,同时令  $D = D_{\text{wire}} + D_{\text{buf}}, R = R_{\text{wire}} + R_{\text{buf}}$  和  $C(v, \alpha)$  为节点  $v$  处的负载电容. 根据式(7)、式(8)、式(14)和式(15)可以得到  $v^*$  处的弛豫时间  $Q$  为

$$Q(v^*, \alpha_i) = Q(v, \alpha_i) - \frac{r \times c \times L^2}{2} - r \times L \times C(v, \alpha_i) = D_{\text{buf}} - R_{\text{buf}} \times (C(v, \alpha_i) + c \times L) = Q(v, \alpha_i) - C(v, \alpha_i) - D \quad (17)$$

式中:  $r$  和  $c$  分别为单位长度的线电阻和线电容;

$L$  为  $v^*$  节点和  $v$  节点之间的总线长;  $D_{\text{wire}}$  和  $D_{\text{buf}}$  分别为线延时和缓冲器延时;  $R_{\text{wire}}$  和  $R_{\text{buf}}$  分别为线电阻和缓冲器输出电阻. 因此在节点  $v^*$  从节点  $v$  继承的解方案的弛豫时间之差可表示为

$$Q(v^*, \alpha_k) - Q(v^*, \alpha_j) = Q_k - Q_j - R \times (C_k - C_j) \quad (18)$$

$$Q(v^*, \alpha_j) - Q(v^*, \alpha_i) = Q_j - Q_i - R \times (C_j - C_i) \quad (19)$$

同时节点  $v^*$  从节点  $v$  继承的负载电容为

$$C(v^*, \alpha_k) = C(v^*, \alpha_j) = C(v^*, \alpha_i) = C_{\text{buf}} \quad (20)$$

基于式(18)和式(19)大于 0、小于 0 和等于 0 所有 6 种情况分别讨论:

1) 如果  $Q(v^*, \alpha_k) - Q(v^*, \alpha_j) > 0$ , 且  $Q(v^*, \alpha_j) - Q(v^*, \alpha_i) > 0$ , 由式(18)、式(19)和式(16)可得

$$\frac{1}{R} > \frac{C_j - C_i}{Q_j - Q_i} > \frac{C_k - C_j}{Q_k - Q_j} \quad (21)$$

基于式(20)和式(21),由 van Ginneken 冗余裁剪规则<sup>[11]</sup>,  $\alpha_j$  冗余.

2) 如果  $Q(v^*, \alpha_k) - Q(v^*, \alpha_j) < 0$ , 且  $Q(v^*, \alpha_j) - Q(v^*, \alpha_i) > 0$  和  $Q(v^*, \alpha_k) - Q(v^*, \alpha_i) > 0, Q(v^*, \alpha_j) - Q(v^*, \alpha_i) > 0$  情况类似,可以得到

$$\frac{C_j - C_i}{Q_j - Q_i} > \frac{C_k - C_j}{Q_k - Q_j} > \frac{1}{R} \quad (22)$$

基于式(20)和式(22),由 van Ginneken 冗余裁剪规则,  $\alpha_j$  冗余.

3) 如果  $Q(v^*, \alpha_k) - Q(v^*, \alpha_j) > 0$ , 且  $Q(v^*, \alpha_j) - Q(v^*, \alpha_i) < 0$ , 由式(16)、式(18)和式(19)可得

$$\frac{C_j - C_i}{Q_j - Q_i} > \frac{1}{R} > \frac{C_k - C_j}{Q_k - Q_j} \quad (23)$$

基于式(20)和式(23),由 van Ginneken 冗余裁剪规则,  $\alpha_j$  冗余.

4) 如果  $Q(v^*, \alpha_k) - Q(v^*, \alpha_j) < 0$ , 且  $Q(v^*, \alpha_j) - Q(v^*, \alpha_i) > 0$ , 则

$$\frac{C_j - C_i}{Q_j - Q_i} < \frac{1}{R} < \frac{C_k - C_j}{Q_k - Q_j} \quad (24)$$

这与式(16)相矛盾,因此不满足图 8 所示分布.

5) 如果  $Q(v^*, \alpha_k) - Q(v^*, \alpha_j) = 0, Q(v^*, \alpha_j) - Q(v^*, \alpha_i)$  任意,同时基于式(20),由 van Ginneken 冗余裁剪规则明显  $\alpha_j$  冗余.

6) 如果  $Q(v^*, \alpha_k) - Q(v^*, \alpha_j)$  任意,  $Q(v^*, \alpha_j) - Q(v^*, \alpha_i) = 0$ , 同时基于式(20),由 van Ginneken 冗余裁剪规则明显  $\alpha_j$  冗余.

### 3 仿真结果与分析

输入网表文件来源于 MCNC 和 GSRC 标准测试电路,软件实现主要用到 C++ 和 Perl 脚本语言,在 Linux 操作系统下完成,Gcc 版本为 3.46. CPU 主频 2.4 GHz,内存 12 GB. 目标 CMOS 工艺技术为 65 nm,并且互连线和缓冲器所有参数参考预测技术模型 PTM (Predictive Technology Model)<sup>[17]</sup>,详细参数如表 1 所示.为了和参考文献对比,也只用一种固定类型缓冲器在 293 K 温度下输出电阻为 363  $\Omega$ ,输入负载电容为 23.4 fF,固有延时为 36.4 ps.

表 1 互连线和缓冲器电气参数  
Table 1 Electrical parameters on interconnect and buffer

参数	参数值
单位长度电容/(fF· $\mu\text{m}^{-1}$ )	0.118
单位长度电阻/( $\Omega$ · $\mu\text{m}^{-1}$ )	0.074
缓冲器输出电阻/ $\Omega$	363
缓冲器输入电容/fF	23.4
缓冲器固有延时/ps	36.4

在给出考虑温度分布的布局测试结果之前,首先给出如何产生功耗迹文件.功耗迹文件是利用 Perl 脚本语言自动生成,其中块名来自预布局生成的文件.每一个在功耗迹文件中的块有 100 个功耗数,其由一个随机数发生器产生,并且高功耗数、中等功耗数和低功耗数分布是随机分布.考虑温度分布的布局与不考虑温度分布布局的结果对比如图 9 所示.图 9 中 hp 和 apte 为 MCNC 标准测试电路

测试结果,n10、n30 和 n200 为 GSRC 标准测试电路测试结果.图 10 给出考虑热效应布局温度降低情况,从中可以看出,考虑温度分布布局与不考虑温度分布布局,芯片最高温度能降低 5%~18%.图 11 给出由于考虑温度效应,造成布线长度变化情况,从中可以看出线长增加最多不到 16%.

表 2 给出温度分布范围及相应温度下所有基准电路的线电阻变化情况.在优化布局后,热条件下的 RC 延时模型和快速缓冲器插入算法被加到 FBI 程序中,优化整个电路的时序,给出优化结果.为了本文的简洁性,本文只考虑 2 种情况的测试:①关键路径插入缓冲器个数;②最坏路径延时.

关键路径插入缓冲器插入个数情况,为了和参考文献比较,本文也采用每个测试电路保留 10% 路径做为关键路径.表 3 列出和不考虑温度效应及文献[7]的比较结果,从中可以看出相比参考文献,本文提出的考虑温度效应布局下缓冲器插入方法需要缓冲器数比文献[7]的少,这主要由于考虑温度布局降低芯片温度.

对于最坏路径延时情况,每个标准测试电路延时降低情况如表 4 所示.这些电路中包括单驱动多负载和单驱动单负载情况,基于考虑温度效应布局优化缓冲器插入时序优化方法,最坏延时降低 9%~18% 相比传统方法,相对于文献[7]的方法最坏延时降低 5%~7%,这是由于基于本文方法虽然线长增加,但温度降低了.以上结果表明,温度效应对于集成电路时序优化有重要影响,必须进行优化布局.

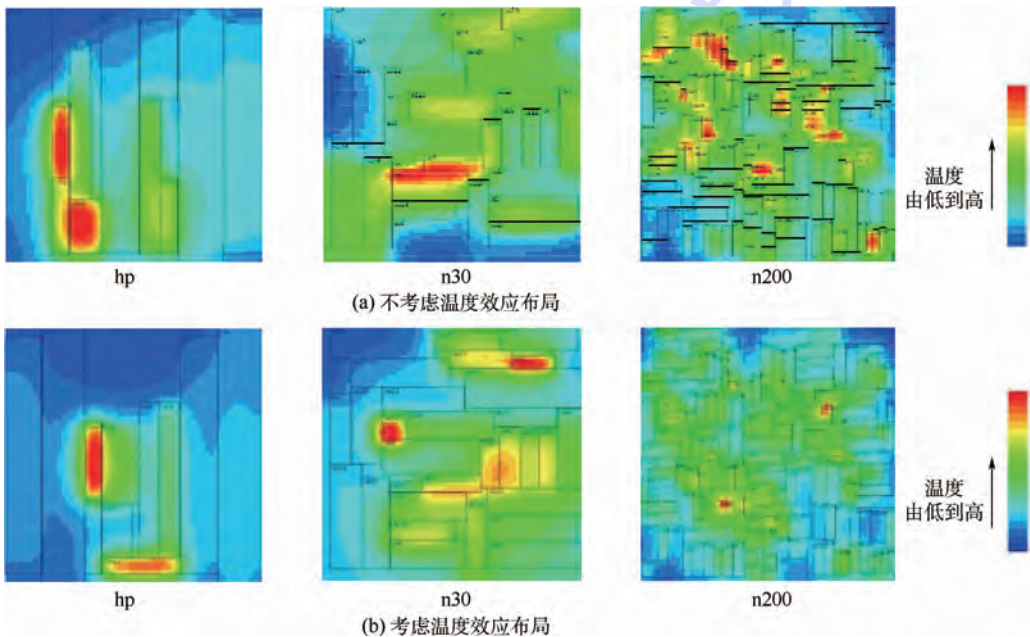


图 9 不同布局的比较

Fig. 9 Comparison of different placements

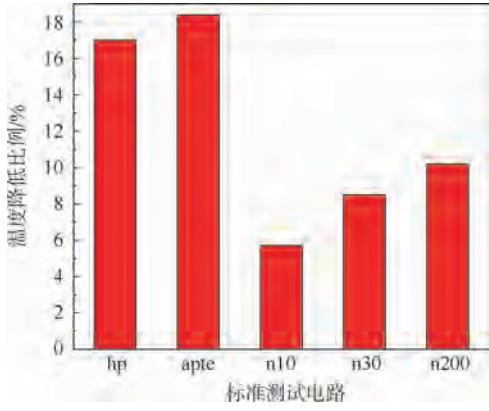


图 10 考虑热效应布局温度降低情况

Fig. 10 Temperature reduction based on thermal effect placement

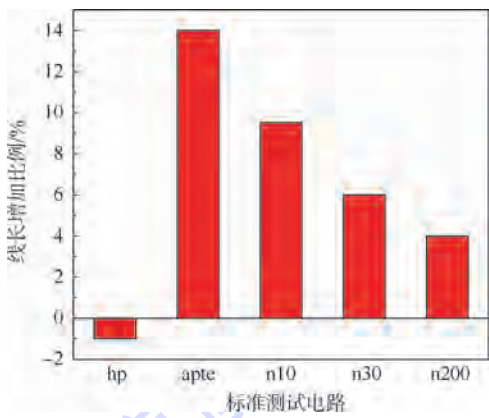


图 11 考虑热效应布局线长增加情况

Fig. 11 Wire length increasing based on thermal effect placement

表 2 不同温度下的互连线电阻参数

Table 2 Interconnect resistor parameter in different temperature

T/K	线电阻/( $\Omega \cdot \mu\text{m}^{-1}$ )
396.17 ~ 871.01	0.104 ~ 0.24
310.15 ~ 375.63	0.079 ~ 0.098
303.61 ~ 335.13	0.077 ~ 0.086
306.63 ~ 479.14	0.078 ~ 0.128
305.56 ~ 396.15	0.078 ~ 0.104

注:  $T$  初值为 293.15 K; 线电阻初值为  $0.074 \Omega/\mu\text{m}$ .

表 3 关键路径插入缓冲器数对比

Table 3 Buffer insertion number comparison in critical path

缓冲器插入方法	插入缓冲器数				
	apte	hp	n10	n30	n200
本文方法	145	122	18	39	370
不考虑温度效应	80	73	10	24	210
文献[7]方法	165	132	-	-	-

注:“-”一文献中未有此数据.

表 4 支持向量机与 RBF 网络分类性能比较

Table 4 Comparison between SVM and RBF network

基准电路	延时/ps		延时降低/%	
	不考虑温度效应	文献[7]方法	本文方法	文献[7]方法
apte	1078.67	1014.78	941.54	12.71
hp	605.16	569.83	541.43	10.53
n10	842.35	-	760.21	9.75
n30	790.45	-	650.78	17.67
n200	1405.79	-	1168.36	16.89

## 4 结 论

1) 根据集成电路时序要求,提出基于热条件布局的缓冲器插入方法.首先在热条件下基于模拟退火算法进行优化布局(相对传统布局温度能降低 5% ~ 20%);接着进行热条件下缓冲器插入.

2) 构建基于热效应的温度模型,并且采用快速缓冲器插入算法进行互连时序优化.采用 C++ 语言实现对 MCNC 和 GSRC 标准测试电路的仿真模拟,模拟结果表明算法的有效性.

3) 根据电路规模不同,选择不同的测试电路进行仿真对比测试,仿真结果表明本文提出的方法的有效性包括:①需要缓冲器数少;②最坏延时的更低,比传统方法降低 9% ~ 18%,比 Kim 等<sup>[7]</sup>提出的方法降低 5% ~ 7%.

## 参考文献 (References)

- [1] Tsai C H, Kang S M. Cell-level placement for improving substrate thermal distribution[J]. IEEE Transactions on Computer-Aided Design for Integrated Circuits and Systems, 2000, 19(2): 253-266.
- [2] Ajami A H, Banerjee K, Pedram M. Modeling and analysis of non-uniform substrate temperature effects on global ULSI interconnects[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2005, 24(6): 849-861.
- [3] Ajami A H, Banerjee K, Pedram M. Analysis of substrate thermal gradient effects on optimal buffer insertion[C]// IEEE ACM International Conference on ICCAD. Piscataway, NJ: IEEE Press, 2001: 44-48.
- [4] Huang W, Ghosh S, Velusamy S, et al. HotSpot: A compact thermal modeling methodology for early-stage VLSI design[J]. IEEE Transactions on Very Large Scale Integration Systems, 2006, 14(5): 501-513.
- [5] Sankaranarayanan K, Velusamy S, Stan M R, et al. A case for thermal-aware floorplanning at the microarchitectural level[J]. Journal of Instruction-Level Parallelism, 2005(8): 1-16.
- [6] Saxena P, Menezes N, Cocchini P, et al. Repeater scaling and its impact on CAD[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2004, 23(4): 451-463.
- [7] Kim M, Ahn B G, Kim J, et al. Thermal aware timing budget for

- buffer insertion in early stage of physical design[C] // IEEE International Symposium on Circuit and Systems. Piscataway, NJ: IEEE Press, 2012:357-360.
- [8] Kirkpatrick S. Optimization by simulated annealing: Quantitative studies[J]. Journal of Statistical Physics, 1983, 34 (5-6): 975-986.
- [9] Kevin S. HotSpot[EB/OL]. 2014 [2015-03-22]. <http://lava.cs.virginia.edu/hotspot>.
- [10] Kevin S. HotSpot[EB/OL]. 2014 [2015-03-22]. <http://lava.cs.virginia.edu/HotSpot/HotSpotHowTo.htm>.
- [11] van Ginneken L P P P. Buffer placement in distributed RC-tree networks for minimal elmore delay [C] // IEEE International Symposium on Circuits and Systems. Piscataway, NJ: IEEE Press, 1990:865-868.
- [12] Gupta R, Tutuianu B, Pileggi L T. The Elmore delay as a bound for RC trees with generalized input signals[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 1997, 16(1):95-104.
- [13] Abou A I, Nowak B, Chu C. Fitted Elmore delay: A simple and accurate interconnect delay model[J]. IEEE Transactions on Very Large Scale Integration System. Piscataway, NJ: IEEE Press, 2004, 12(7):691-696.
- [14] Athikulwongse K, Zhao X, Lim S K. Buffered clock tree sizing for skew minimization under power and thermal budgets [C] // Proceedings of the 2010 Asia and South Pacific Design Automation Conference. Piscataway, NJ: IEEE Press, 2010:474-479.
- [15] Skadron K, Stan M, Velusamy S, et al. A case for thermal aware floor planning at the microarchitectural level[J]. Journal of Instruction-level Parallelism, 2005, 7:1-16.
- [16] Patrick H M, Ameya R A. GSRC bookshelf benchmarks[EB/OL]. 2001 (2007-06-01) [2015-03-22]. <http://vlsicad.cs.binghamton.edu/benchgsrc.html>.
- [17] Yu C. Predictive technology model[EB/OL]. Arizona: [s. n.], 2005 (2012-06-01) [2015-03-22]. <http://ptm.asu.edu/>.

## Thermal aware floor planning timing optimal method for buffer insertion

WANG Xinsheng<sup>1</sup>, HAN Liang<sup>\*1</sup>, YU Mingyan<sup>2</sup>

(1. School of Information and Electrical Engineering, Harbin Institute of Technology, Weihai 264209, China;

2. Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China)

**Abstract:** With the integration degree of integrated circuit (IC) is increasingly high, the heat on a chip is also growing, which leads to an uneven temperature distribution intra-die and affects the propagation delay of the critical path thereby affecting the performance of buffer insertion path. A buffer insertion timing optimization method which considered the heat distribution condition optimization floor-planning was proposed. It estimates the temperature and heat distribution of chip in the early stages of layout design and is applied to layout optimization floor-planning. The thermal aware floor planning based on simulated annealing algorithm was used to adjust and optimize planning, and then we made an optimization for timing by proposed buffer insertion model and fast buffer insertion algorithm. Simulation results show that the use of the proposed buffer insertion delay optimization method can effectively reduce the worst delay and the number of buffer insertion, worst delay is 9% - 18% lower than traditional methods, 5% - 7% lower than the best method shown in reference, and the insertion buffer numbers are 10 to 20 less than its.

**Key words:** very large scale integrated circuit; floor planning; buffer insertion; interconnect; simulated annealing algorithm

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0746

# 电磁兼容测量天线系数温度误差修正方法

何洋<sup>1</sup>, 苏东林<sup>\*1</sup>, 李艳<sup>2</sup>, 赵子华<sup>1</sup>, 刘洪颀<sup>1</sup>

(1. 北京航空航天大学 电子与信息工程学院, 北京 100191; 2. 63863 部队 二室, 白城 137001)



**摘 要:** 为减小外场环境温度的影响,提高电磁兼容性(EMC)测试设备的测量精度,提出了一种电磁兼容测量天线系数温度误差修正方法.首先,根据1 m距离两天线校准原理,理论推导了近场情况下采用透波材料制作的温控装置对接收天线接收特性的影响,结合仿真分析与试验结果,验证了使用温控装置带来的测试误差可控制在允许的范围.然后,在此基础上,通过实际测试获得了典型天线的天线系数随温度的变化规律及误差修正曲面.最后,采用对比标准条件下和开阔试验场条件下的测试结果的方法来验证该方法的有效性.结果表明,在保证测量精度的前提下,通过天线系数温度误差修正,可将天线的使用环境温度扩展到-40~+50℃,进而减小外场试验测试误差.适用于双锥、对数周期和双脊喇叭等天线的天线系数温度误差修正.

**关键词:** 电磁兼容性(EMC); 天线系数; 温度; 修正; 校准

**中图分类号:** TM930.1

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1821-09

为确保试验结果的准确性和可对比性,目前电磁兼容性(Electromagnetic Compatibility, EMC)试验大都在温湿度条件可控的微波暗室中进行.但大型设备系统级电磁兼容性试验<sup>[1]</sup>、现场快速干扰诊断排查、电磁环境监测等一般在外场进行,外场环境温度无法像内场(如屏蔽室)那样做到可控,如果不做修正,往往会对测量结果产生影响,主要表现在两个方面:①对测试设备的测量精度产生影响;②可导致被试大型设备的电磁兼容性产生变化.被试大型设备在外场实际使用环境下的电磁兼容性,恰恰反映了其使用状态的真实电磁兼容性能;而测试设备的测量精度受到的影响必须得到控制和修正,方可确保测量结果的有效性.以开阔试验场(Open Area Test Site, OATS)电磁兼容性辐射发射测试为例<sup>[2]</sup>,目前主要的控制方法为将接收机、计算机、衰减器等测试设备放置于温度可控

的方舱内,使其不会受到外界自然温度环境影响.但测量天线这一核心设备必须置于开阔试验场场坪的测试点位上,暴露于自然环境下,其测试精度不可避免地会受到外界环境温度的影响.

天线系数是用来表征天线接收特性的重要参数,直接关系到电磁兼容性测量结果的准确度.国内外许多学者对于电磁兼容测量天线系数的校准方法进行了研究,文献[3-9]提出了标准条件下天线系数校准的改进方法,文献[10-15]研究了开阔试验场、混响室等其他条件下天线系数的校准方法,但都只给出一种温度条件下的校准方法和结果,即在实验室标准温度条件下(20℃左右)的天线系数,由温度因素引起的天线系数变化未见文献报道.因此,需要剔除由环境温度引起的测量天线系数误差,提高测试设备的精度,确保测试结果的有效性和准确性.

收稿日期: 2014-12-01; 录用日期: 2015-01-23; 网络出版时间: 2015-03-18 14:37

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150318.1437.003.html

基金项目: 国家自然科学基金(61427803, 61221061)

作者简介: 何洋(1974—),男,吉林松原人,博士研究生, nudt\_hy@163.com

\* 通讯作者: 苏东林(1960—),女,山东莱芜人,教授, sdl@buaa.edu.cn, 主要研究方向为电磁兼容、计算电磁学、抗干扰理论与应用、射频微波电路与系统、新型飞行器机载共形/共用/小型化天线等.

**引用格式:** 何洋, 苏东林, 李艳, 等. 电磁兼容测量天线系数温度误差修正方法[J]. 北京航空航天大学学报, 2015, 41(10): 1821-1829. He Y, Su D L, Li Y, et al. Method for EMC antenna factor temperature error correction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1821-1829 (in Chinese).

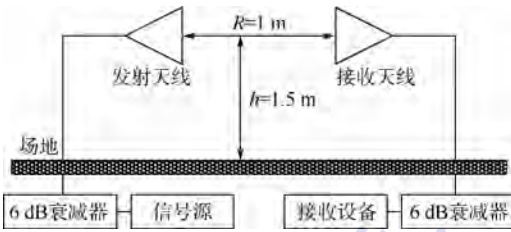
# 1 天线系数校准原理

## 1.1 标准 1 m 法天线系数校准原理

GJB J5410—2005 中给出了距被测件 1 m 的两天线法天线系数校准方法<sup>[16]</sup>,如图 1 所示.校准时,需要两个相同的天线,两天线间的插入损耗与天线系数之间的关系为

$$A_{FT} + A_{FR} = I_L + 20 \lg \frac{\eta_0}{Z_0} - 20 \lg \lambda \quad (1)$$

式中: $A_{FT}$ 和 $A_{FR}$ 分别为发射天线和接收天线的天线系数,dB/m; $I_L$ 为发射天线和接收天线端口之间的插入损耗,dB; $\eta_0 = 120 \pi \Omega$ 为自由空间波阻抗; $Z_0$ 为系统阻抗,通常为 $50 \Omega$ ; $\lambda$ 为电磁波波长,m.



R—发射天线与接收天线之间的距离;h—测试距离地面高度.

图 1 两天线法天线系数校准

Fig. 1 Antenna factor calibration using two antennas

如果已知其中一个天线的天线系数,则可计算待测天线的天线系数:

$$A_{F1} = I_L + 20 \lg f_M - 32 - A_{F2} \quad (2)$$

式中: $A_{F1}$ 和 $A_{F2}$ 分别为待测天线和已知天线的天线系数,dB/m; $f_M$ 为测量频率,MHz.

## 1.2 天线系数温度误差补偿原理

若要得到温度对天线性能的影响变化规律,需将被试天线置于温度可控的环境内,见图 2.

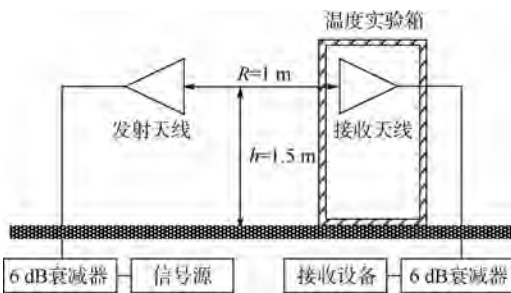


图 2 不同温度条件下天线系数校准

Fig. 2 Antenna factor calibration under different temperature conditions

设信号源输出为 $P$ ,使用发射天线产生一定场强,在标准场地常温条件下,接收天线接收到的场强为 $E_0$ ,则

$$P - A_{FR} - I_L - A_{FT} = E_0 \quad (3)$$

保持测试条件不变,将接收天线置于温度实

验箱中,此时:

$$P - A_{FR} - I_L - S - A_{FT} = E'_0 \quad (4)$$

式中: $S$ 为场地引入的误差; $E'_0$ 为接收天线置于温度实验箱后接收到的场强.

保持测试条件不变,改变温度实验箱内的温度,此时:

$$P - A_{FR} - I'_L - S' - A'_{FT} = E''_0 \quad (5)$$

式中: $I'_L$ 、 $S'$ 、 $A'_{FT}$ 和 $E''_0$ 分别为不同温度条件下的线缆损耗、场地引起的误差、接收天线的天线系数和接收天线接收到的场强.假设

$$\begin{cases} I'_L \approx I_L \\ S' \approx S \end{cases} \quad (6)$$

则由式(4)和式(5)可得

$$S = E_0 - E'_0 \quad (7)$$

$$A'_{FT} = E'_0 - E''_0 + A_{FT} \quad (8)$$

式中: $A'_{FT}$ 为天线在不同温度下的天线系数.

## 1.3 测试误差分析

由天线系数误差补偿原理可知,产生测试误差的来源主要为温度对线缆和温度实验箱的影响,以及温度实验箱对接收天线的影

响.温度对线缆的影响主要通过形变引起其阻抗等电参数变化.金属的热膨胀系数一般为 $10^{-5} \sim 10^{-6} \text{ m}/^\circ\text{C}$ ,自然环境温度变化一般为 $-40 \sim 50^\circ\text{C}$ ,实验用线缆最大尺寸一般不超过 3 m,由温度变化引起的电参数变化非常小,可以忽略.

但温度实验箱体对被试天线接收性能产生影响,为尽量减小箱体内部对电磁波的反射及由密闭腔体导致的谐振效应,本文采用以二氧化硅体系为主的透波材料作为温度实验箱体,该种材料具有极小的线膨胀系数 $\alpha$ (约 $0.5 \times 10^{-6} \text{ K}^{-1}$ )、较好的抗热冲击性能、较低的介电常数 $\epsilon$ ( $2.8 \sim 3.5$ )和损耗角正切 $\tan \sigma$ ( $< 0.004$ ),且随温度变化小的优点,同时导热系数小,热防护能力好,制造工艺相对来说较简单,成本较低.

选取辐射发射测试中常用的 LPDA-9531 型对数周期天线为例,见图 3,研究箱体对被试天线接收性能的影响规律.由于军用标准规定天线与被试品的距离为 1 m,此时,对数周期天线处于近



图 3 LPDA-9531 型对数周期天线

Fig. 3 LPDA-9531 type log periodic antenna

场使用条件,很难进行理论推导和计算,因此,本文采用仿真分析与实测相结合的方法进行分析。

### 1.3.1 典型天线建模仿真分析

使用 CST 电磁场仿真软件对 LPDA-9531 型对数周期天线进行精细建模<sup>[17]</sup>,LPDA-9531 型对数周期天线仿真模型如图 4 所示。

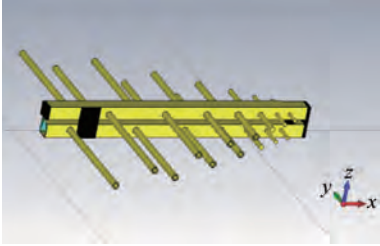


图 4 LPDA-9531 型对数周期天线仿真模型

Fig.4 A simulation model of log periodic LPDA-9531 type

模型尺寸及材料性质均采用实际数据.以此为基础,分别讨论温度实验箱体对被试天线性能的影响。

#### 1) 实验箱谐振效应仿真分析.

设置平面波作为辐射源,对温度实验箱进行照射,如图 5 所示.箱体内部尺寸 1 600 mm × 1 200 mm × 1 800 mm,材料介电常数  $\epsilon = 3.0$ ,  $\tan \sigma = 0.005$ .仿真频率  $f$  范围为 0.3 ~ 1 GHz,按

照表 1 所示实验箱谐振效应仿真参数设置进行仿真。

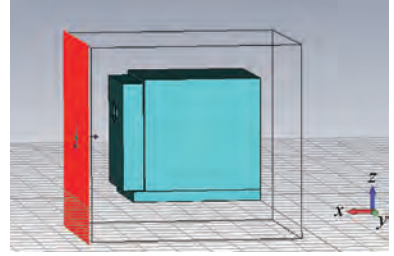


图 5 实验箱谐振效应仿真模型

Fig.5 Experiment box resonance effect simulation model

表 1 实验箱谐振效应仿真参数设置

Table 1 Resonance effect simulation parameter settings of experiment box

$f/\text{MHz}$	仿真位置/mm	
	$x$ 平面	$z$ 平面
400,600,800	400	600
	1 100	1 000
	1 400	1 400

通过对不同频率、不同位置的仿真结果可知,当  $x = 1 400 \text{ mm}$  且  $z = 1 400 \text{ mm}$  时的附近区域电磁场分布较均匀,且谐振效应不明显,场强在  $0.3 \text{ V/m}$  以下. $x = z = 1 400 \text{ mm}$  时 3 种频率下的仿真结果如图 6 所示。

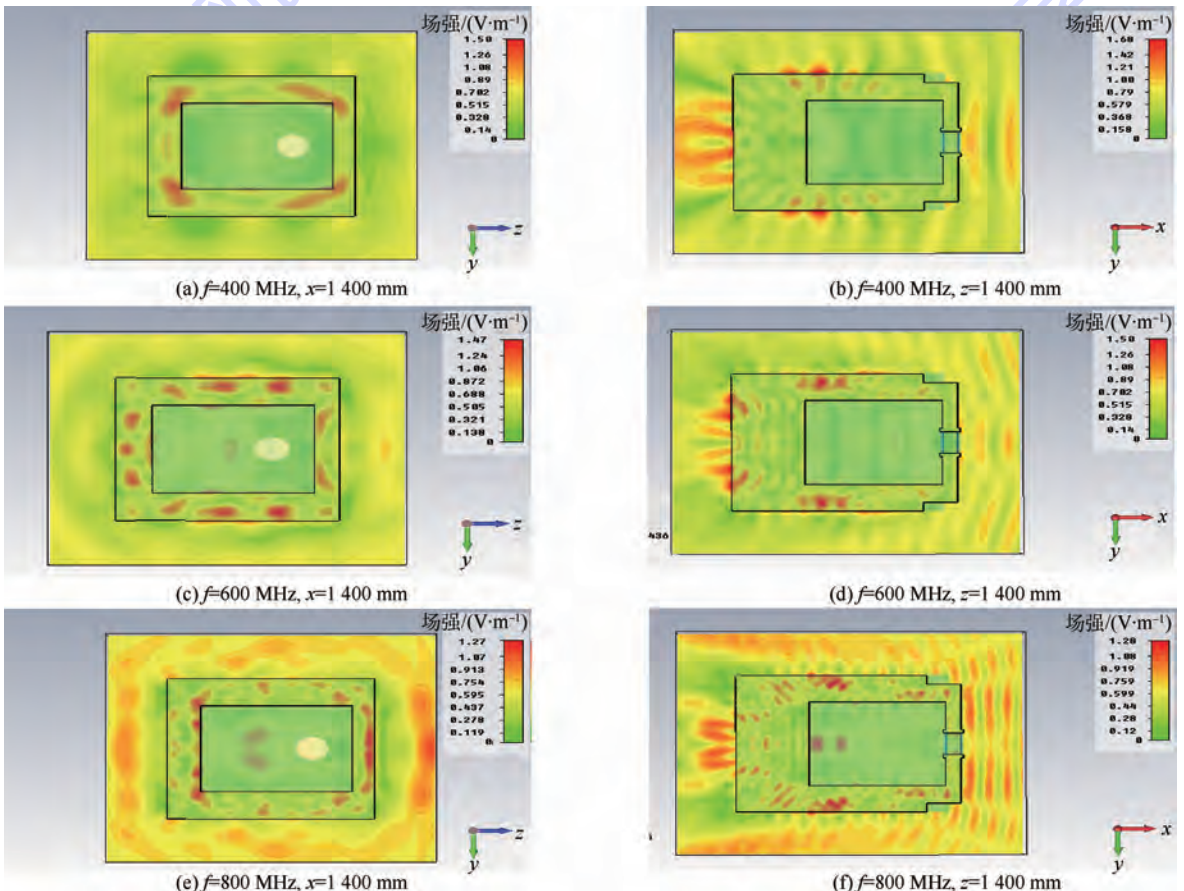


图 6  $x = z = 1 400 \text{ mm}$  时 3 种频率下的仿真结果

Fig.6 Simulation results in three kinds of frequency with  $x = z = 1 400 \text{ mm}$

### 2) 自由空间天线接收性能仿真.

建立自由空间天线系数测试仿真模型,如图7所示.收发天线距地面1.5 m,两天线之间距离为1 m.图8为自由空间接收天线的频率-端口电压( $V$ )变化曲线.



图7 自由空间天线系数测试仿真模型

Fig.7 Simulation model of antenna factor test in free space

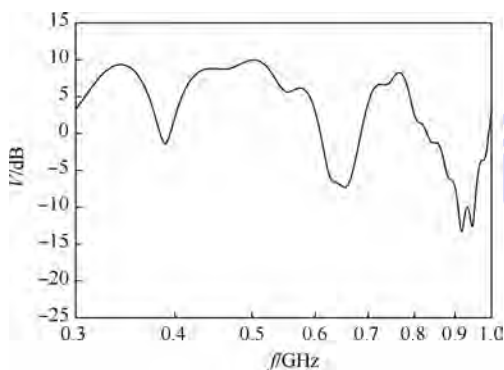


图8 自由空间接收天线的频率-端口电压变化曲线

Fig.8 Frequency-port voltage changing curve of receiving antenna in free space

### 3) 温度实验箱对天线性能影响仿真.

使用CST仿真软件,对实验中使用的温度实验箱进行建模仿真,接收天线置于实验箱内时的仿真模型和实验箱内接收天线的频率-端口电压变化曲线分别如图9和图10所示.

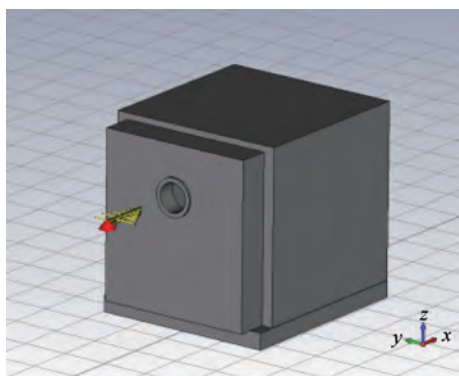


图9 接收天线置于实验箱内时的仿真模型

Fig.9 Simulation model when receiving antenna put into experimental box

发射功率为13 dBm,  $f$ 为0.3~1 GHz.首先,在标准实验室温度20℃条件下,按照校准频点逐一测试,使用EMI接收机测得天线接收的幅值数据;其次,保持测试配置不变,将接收天线放置于温度实验箱内,设置箱内温度为20℃,测得此时天线在各个校准频点的接收场强幅值  $A$ .20℃时两种条件下的接收天线实测数据对比如图11所示.

仿真结果与实测结果的相对误差对比,如图12所示.可见,仿真结果与实测结果的相对误差

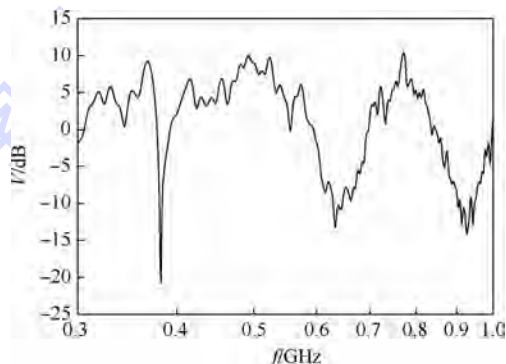


图10 实验箱内接收天线的频率-端口电压变化曲线

Fig.10 Frequency-port voltage changing curve of receiving antenna in experimental box

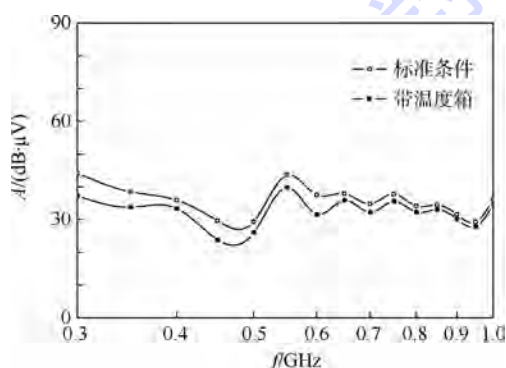


图11 20℃时两种条件下接收天线实测数据对比

Fig.11 Measured data of receiving antenna under two conditions when temperature is 20℃

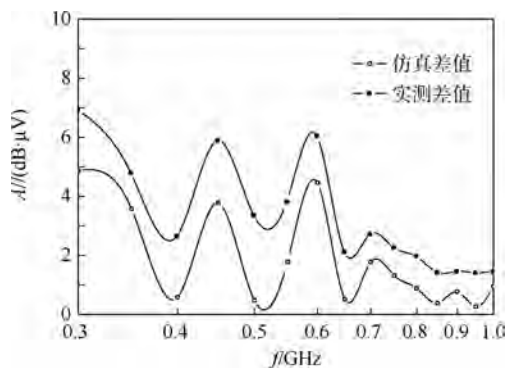


图12 仿真结果与实测结果的相对误差对比

Fig.12 Comparison of relative errors obtained from simulation results and measured results

### 1.3.2 实验测试分析

选取两副相同型号(LPDA-9531)的对数周期天线,测试配置与仿真布局相同,设置信号源

曲线变化趋势基本一致,两条曲线之间的平均误差为 1.46 dB,最大误差为 2.88 dB,说明当其他测试条件相同时,引入温度实验箱后的系统测试误差与仿真值(近似为理论值)相比可满足 3 dB 的测试要求.

## 2 天线系数温度误差修正

### 2.1 测试方法

天线系数温度误差校准流程如图 13 所示,主要测试步骤为:

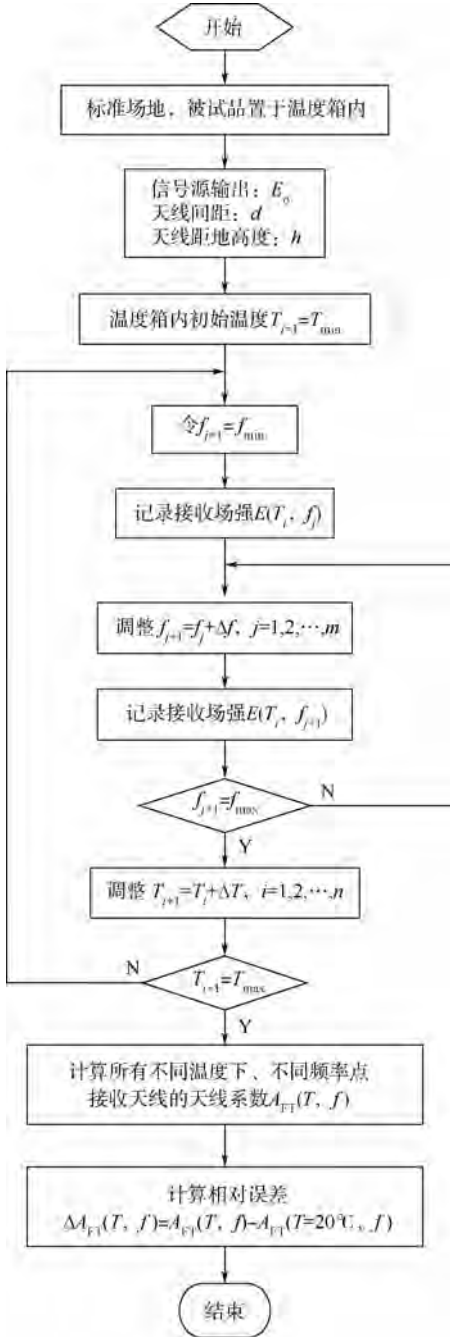


图 13 天线系数温度误差校准流程

Fig. 13 Calibration process of antenna factor temperature error

1) 如图 2 所示,将两个型号相同的天线互相对准,距离 1 m,保证两天线轴线在同一直线上,并且极化匹配.将接收天线置于温度实验箱内并保持  $T = 20^\circ\text{C}$ ,在此情况下,按照天线校准证书中给出的校准频率  $f_j (j = 1, 2, \dots, m)$ ,在校准频率最小值  $f_{\min}$  和最大值  $f_{\max}$  之间,按照频率变化量  $\Delta f$  步进,分别测量每一个校准频点的  $A_{FT}$ .

2) 保持信号源输出和测试配置不变,假设  $T_{\min} \leq T \leq T_{\max}$ ,首先设置实验箱内温度为最低温度  $T = T_{\min}$ ,对接收天线进行保温.之后,按照步骤 1) 中的配置进行测量,得到此时接收天线的天线系数  $A_{FT_{\min}}$ .

3) 保持信号源输出和测试配置不变,令温度实验箱内温度  $T_{i+1} = T_i + \Delta T (\Delta T$  为温度变化量,  $i = 1, 2, \dots, n)$ ,重复步骤 2),获取不同温度条件下被试天线的天线系数  $A'_{FT}$ .

4) 将测得的不同温度条件下对应的天线系数  $A'_{FT}$  分别与标准温度条件下得到的天线系数  $A_{FT}$  进行比较,获得不同温度对天线系数的影响变化量  $\Delta A_{FT}$ .

5) 使用插值拟合算法对步骤 4) 中获得的测试数据进行处理,获得天线系数随温度变化的温度-频率-天线系数相对变化量曲面.

### 2.2 典型天线系数温度误差修正曲线

应用第 2.1 节中的方法,在测试频率 0.3 ~ 1 GHz 的范围内,实验温度  $-40 \sim 50^\circ\text{C}$  之间,每隔  $10^\circ\text{C}$  对 LPDA-9531 型对数周期天线进行一次测试,获得其在典型温度下 LPDA-9531 型对数周期天线接收场强幅值变化曲线,如图 14 所示.采用 3 次样条插值算法对上述典型温度下天线接收场强幅值的相对误差进行拟合处理,得到对数周期天线的频率-温度-天线接收场强幅值相对误差 ( $\delta$ ) 修正曲面,如图 15 所示.

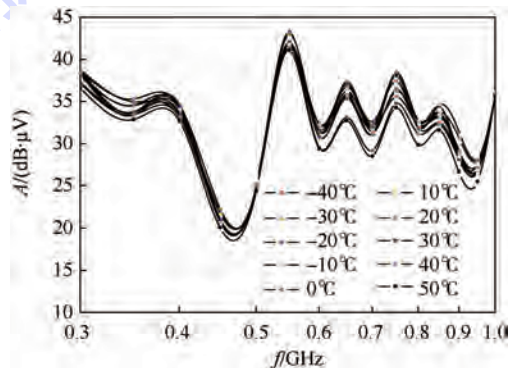


图 14 典型温度下 LPDA-9531 型对数周期天线系数幅值变化曲线

Fig. 14 Antenna factor amplitude curve of LPDA under typical temperature

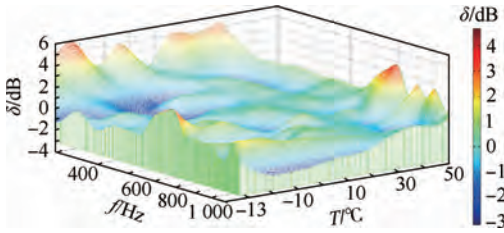


图 15 LPDA-9531 型对周期天线频率-温度-天线系数相对误差修正曲面

Fig. 15 Correction surface of periodic antenna frequency-temperature-antenna factor relative error of LPDA-9531 type

### 3 实例验证

为进一步验证该方法的正确性和有效性,采取如下方法进行验证.

#### 3.1 验证方法

分别在标准实验室 20℃ 条件下和开阔试验场自然环境温度条件下进行辐射发射测试,采用本文的天线系数温度误差修正方法对开阔试验场条件下的测试数据进行修正,逆推预测常温下的结果,再与标准实验室的测试结果进行比对分析,检验两种结果之间的误差,验证该方法的有效性.

为确保两种场地情况下测试结果的可比性,需要选取标准被试品并使其状态保持一致,为此,本文选取 CGC-515 型标准辐射源并将其置于非金属透波材料制作的保温箱内,保温温度设为 20℃,测试频率为 0.3 ~ 1 GHz.

1) 标准实验室 20℃ 条件下标准被试品的辐射发射测试.

在标准实验室 20℃ 条件下,将辐射参考源置于保温箱内并使其温度保持在 20℃,将被试天线放置在距离辐射参考源被试面  $d = 1\text{ m}$  处,天线与参考源高度  $h = 1.5\text{ m}$ ,进行辐射发射测试,获取辐射参考源的辐射发射数据,标准 EMC 实验室辐射发射测试布置如图 16 所示.

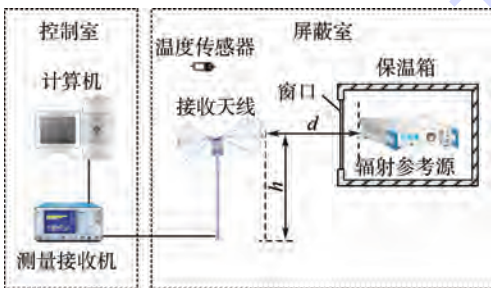


图 16 标准 EMC 实验室辐射发射测试布置  
Fig. 16 Radiation emission test layout in standard EMC laboratory

2) 开阔试验场高低温环境条件下标准被试品辐射发射测试.

将辐射参考源置于保温箱内并使其温度保持在 20℃,测试配置与标准实验室测试布置一致.利用温度传感器实时监测开阔试验场环境温度的变化情况,选取合适的环境温度点,使用被试天线测量辐射参考源的辐射发射情况,记录此时辐射参考源的辐射发射数据及对应的环境温度值,开阔试验场辐射发射测试布置如图 17 所示.

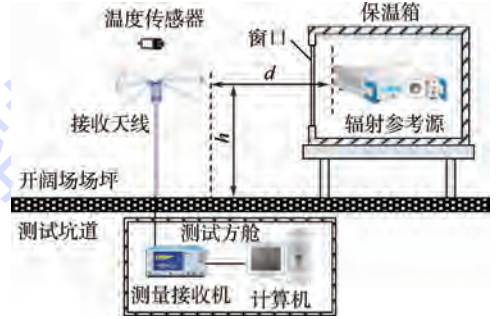


图 17 开阔试验场辐射发射测试布置  
Fig. 17 Radiation emission test layout in OATS

#### 3.2 实验结果比对分析

1) 标准实验室 20℃ 条件下标准被试品的辐射发射测试数据.

在北京无线电计量测试研究所完成 20℃ 条件下标准被试品的辐射发射测试,将测试结果作为基准比对数据,标准 EMC 实验室 20℃ 时辐射发射测试数据,如图 18 所示.

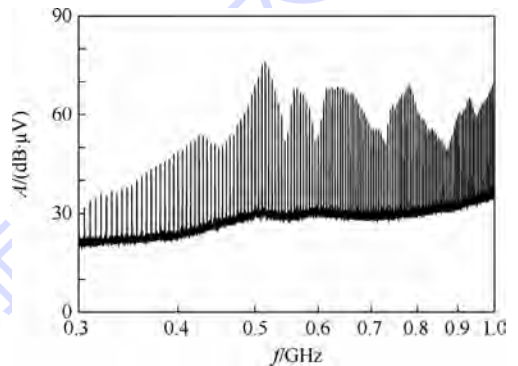


图 18 标准 EMC 实验室 20℃ 时辐射发射测试数据  
Fig. 18 Radiation emission test data at 20℃ in standard EMC laboratory

2) 开阔试验场自然环境温度条件下标准被试品辐射发射测试数据.

在开阔试验场实际环境温度 6℃ 条件下进行标准被试品的辐射发射测试,开阔试验场辐射发射测试布置如图 19 所示,开阔试验场 6℃ 去除噪声后的开阔试验场背景环境电平及辐射发射测试数据如图 20(a) 和图 20(b) 所示.

为达到对比的有效性和准确性,首先将开阔



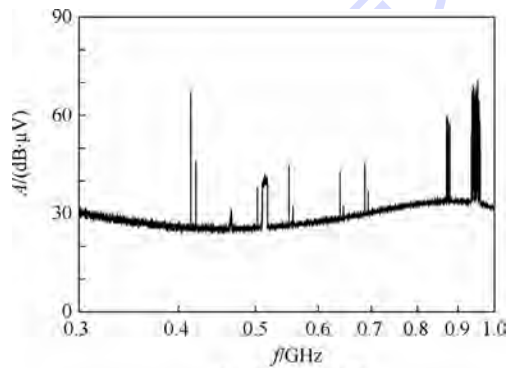
(a) 0.2~1 GHz辐射参考源



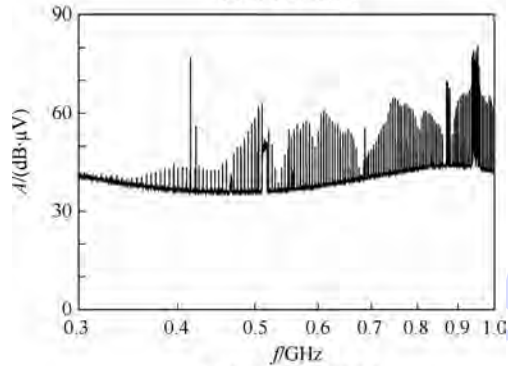
(b) 开阔试验场辐射发射测试布置实物图

图 19 开阔试验场辐射发射测试布置

Fig. 19 Radiation emission test layout in OATS



(a) 背景环境电平



(b) 辐射发射测试数据

图 20 开阔试验场 6℃ 时背景环境电平及辐射发射测试数据

Fig. 20 Background environmental level and radiation emission test data at 6℃ in OATS

试验场典型背景环境干扰剔除,得到可用于对比的开阔试验场辐射发射测试数据,剔除背景干扰的开阔试验场辐射发射测试数据,如图 21 所示。

3) 测试数据比对。

应用图 15 所示频率-温度-天线系数相对误差修正曲面将开阔试验场测试数据修正到 20℃,剔除数据异常值后将开阔试验场实测数据、修正后测试数据与标准实验室 20℃ 时的测试数据进行对比分析,对比频段 0.3 ~ 1 GHz,检验 3 种情况下测试结果之间的误差.实测数据、修正数据与标准数据的比对,如图 22 所示。

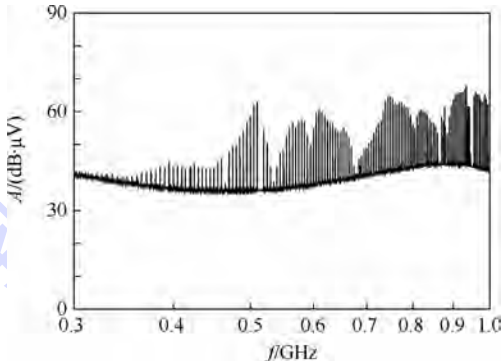


图 21 剔除背景干扰的开阔试验场辐射发射测试数据

Fig. 21 Radiation emission test data in OATS which has eliminated background interference

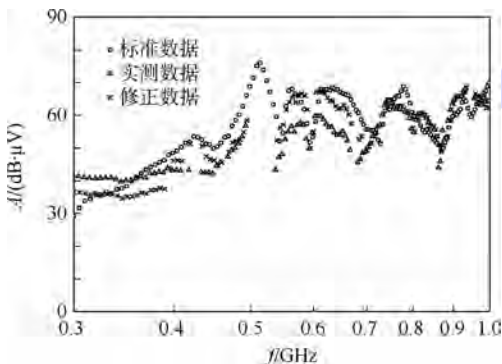


图 22 实测数据、修正数据与标准数据的比对

Fig. 22 Comparison between measured data and revised data and standard data

经计算,实测数据与标准数据之间的平均误差为 2.51 dB,均方根误差为 6.34 dB,修正后的实测数据与标准数据之间的平均误差为 1.16 dB,均方根误差为 4.10 dB.结果表明,通过天线系数温度误差修正,减小了测试误差。

### 4 结 论

本文针对外场电磁兼容性试验,提出了一种电磁兼容测量天线系数温度误差修正方法,通过实验验证表明:

1) 该方法的测量最大误差为 2.88 dB,能够有效保证测量精度。

2) 该方法能够对典型对数周期天线的天线系数进行温度误差修正,实例测试表明,可有效减

小环境温度对测量结果的影响.

3) 通过天线系数温度误差修正,可将典型对数周期天线的使用环境温度扩展到  $-40 \sim +50^{\circ}\text{C}$ .

4) 适用于军标中规定的基于 1 m 法校准的电磁兼容测量天线,如双锥、对数周期和双脊喇叭等天线的天线系数温度误差修正.

为使本文提出的方法能够推广应用到其他类型的电磁兼容性测量天线,并使其能在更大的温度范围内使用,还需进一步优化修正参数和控制测量误差.

### 参考文献 (References)

- [1] GJB/1389A—2005. 系统电磁兼容性要求[S]. 北京:中国人民解放军总装备部,2005:2-13.  
GJB/1389A—2005. Electromagnetic compatibility requirements for systems[S]. Beijing: PLA General Armament Department, 2005:2-13 (in Chinese).
- [2] GJB/151B—2013. 军用设备和分系统电磁发射和敏感度要求与测量[S]. 北京:国防科学技术工业委员会,2013:2-8.  
GJB/151B—2013. Requirements and measurement of electromagnetic emission and susceptibility for military equipment and subsystems[S]. Beijing: Commission of Science, Technology and Industry for National Defense of the PRC, 2013:2-8 (in Chinese).
- [3] 苏东林,戴飞,谢树果,等. 天线系数的测试误差与 NSA 测试的改进[J]. 北京航空航天大学学报,2007,33(11):1291-1294.  
Su D L, Dai F, Xie S G, et al. Errors of antenna factor and improved method for NSA test[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(11):1291-1294 (in Chinese).
- [4] 沈国连. EMC 测试中天线系数的误差分析[J]. 航空电子技术,2001,32(2):1-5.  
Shen G L. Error analysis of antenna factor for EMC test[J]. Chinese Avionics Technology, 2001, 32(2):1-5 (in Chinese).
- [5] CISPR16-1-6—2008. Standard for the calibration of antennas used for radiated emission measurements[S]. Geneva: International Electrotechnical Commission, 2008:2-5.
- [6] ANSI C63.5—2004. American national standard for electromagnetic compatibility-radiated emission measurements in electromagnetic interference (EMI) control-calibration of antennas (9 kHz to 40 GHz)[S]. New York: Institute of Electrical and Electronic Engineers, Inc, 2004:3-16.
- [7] Garn H, Buchmayr M, Mullner W, et al. Primary standards for antenna factor calibration in the frequency range of (30 to 1000) MHz[J]. IEEE Transactions on Instrumentation & Measurement, 1997, 46(2):544-558.
- [8] Meng D L, Alexander M J. Calibration of biconical antennas by vertically stacking method[J]. IEEE Transactions on Electromagnetic Compatibility, 2014, 56(6):1262-1270.
- [9] Das S K, Rao P H, Kanda M. Calibration of antenna factor of a tuned dipole using two antenna method, NIST method and a proposed reference antenna[C]//Proceedings of International Conference on Electromagnetic Interference & Compatibility. Piscataway, NJ: IEEE Press, 1995:263-265.
- [10] Jeong H K, Jeong I P, Ung T K, et al. Antenna factor calibration in 30-1 000 MHz frequency range[C]//Proceedings of Conference on Precision Electromagnetic Measurements Digest. Piscataway, NJ: IEEE Press, 1996:423-424.
- [11] de marinis J. Quality control of antenna calibration and site-attenuation data from open area test sites[C]//Proceedings of IEEE International Symposium on Electromagnetic Compatibility, Symposium Record. Piscataway, NJ: IEEE Press, 1988:290-294.
- [12] Morioka T, Hirasawa K. Proper antenna factors for the normalized site attenuation above a ground plane[J]. IEEE Transactions on Electromagnetic Compatibility, 2013, 56(2):246-258.
- [13] Chen Y M, Chen H Y, Ting Y T, et al. Calibration of antenna factor in nonstandard environments[C]//Proceedings of Asia-Pacific Microwave Conference. Piscataway, NJ: IEEE Press, 2001, 3:1318-1321.
- [14] Liu X X, Ma W Y, Zhang T. The research about the double ridge horn calibration in local EMC testing chambers measurements and uncertainty evaluations[C]//Proceedings of IEEE 5th International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications. Piscataway, NJ: IEEE Press, 2013:311-314.
- [15] Hong T, Chou C L, Kuo A. Improving calibration of broadband antenna factors in a GTEM cell[C]//Proceedings of International Symposium on Electromagnetic Compatibility. Piscataway, NJ: IEEE Press, 1999:592-595.
- [16] GJB/J5410—2005. 电磁兼容测量天线的天线系数校准规范[S]. 北京:国防科学技术工业委员会,2005:5-6.  
GJB/J5410—2005. Antenna factor calibration standard of EMC measurement antenna[S]. Beijing: Commission of Science, Technology and Industry for National Defense of the PRC, 2005:5-6 (in Chinese).
- [17] 张敏. CST 微波工作室®用户全书[M]. 成都:电子科技大学出版社,2014:50-210.  
Zhang M. CST Microwave studio® user encyclopedia[M]. Chengdou: University of Electronic Science and Technology of China Press, 2014:50-210 (in Chinese).

## Method for EMC antenna factor temperature error correction

HE Yang<sup>1</sup>, SU Donglin<sup>\*1</sup>, LI Yan<sup>2</sup>, ZHAO Zihua<sup>1</sup>, LIU Hongyi<sup>1</sup>

(1. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China;

2. The 2nd Department, 63863 Army, Baicheng 137001, China)

**Abstract:** In order to reduce the influence of the outfield environmental temperature and improve the measuring precision of electromagnetic compatibility (EMC) test equipment, a method for antenna factor temperature error correction was presented. Firstly, according to the principle of two antenna calibration with 1 m distance, the antenna receiving characteristic influenced by temperature control device made by radome material under the condition of near field was theoretically deduced. Combined with the simulation analysis and test results proved that the test error can be controlled in the permitted range by using the temperature control device. Secondly, on this basis the antenna factor variation with temperature and temperature error correction surface for typical antenna was obtained by the actual test. In the end, the method was verified by comparing the test result under standard conditions with the test result under the open area test site (OATS) condition. The results prove that this method can expand the environmental temperature range of use from  $-40^{\circ}\text{C}$  to  $50^{\circ}\text{C}$  by antenna factor temperature error correction on the premise of not reducing measurement precision, and suitable for the antenna factor temperature error correction in EMC double cone antenna and log periodic antenna and the double ridged horn antenna, etc.

**Key words:** electromagnetic compatibility (EMC); antenna factor; temperature; correction; calibration

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0778

# 改进的基于本征滤波的时域宽带波束形成

王剑书, 樊养余\*, 杜瑞, 吕国云

(西北工业大学 电子信息学院, 西安 710129)



**摘 要:** 为实现基于本征滤波宽带波束形成的频率不变性和零陷设计, 提出一种改进的方法. 基于最大能量阵列代价函数和主极大方向阵列响应线性约束构造的优化模型, 将阵列空间响应偏差整合入代价函数, 从而实现频率不变性的设计; 提出了两种零陷设计的方法, 将扩展旁瓣能量整合入代价函数和添加干扰方向阵列响应线性约束, 构造两个新的约束优化模型, 最后均由本征滤波的方法求解. 仿真实验表明, 该方法能在通带范围内实现指定角度范围内阵列响应的频率不变性和零陷效果, 同时具有低旁瓣增益, 验证了本文算法的有效性和良好性能.

**关键词:** 宽带波束形成; 本征滤波; 频率不变性; 零陷; 旁瓣

**中图分类号:** TN911.7

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1830-06

波束形成是一种空间滤波技术, 主要任务是通过传感器阵列接收特定方向的信号, 同时抑制其他方向的信号(视为干扰信号或噪声)<sup>[1]</sup>. 波束形成技术广泛应用于雷达<sup>[2]</sup>、声呐<sup>[3,4]</sup>、通信<sup>[5]</sup>和地震学<sup>[6]</sup>等领域. 文献[1]提供了3种时域宽带固定波束形成的设计方法: 凸优化法、最小二乘法和本征滤波(eigenfilter)的方法. 基于本征滤波的波束形成算法, 不需要逐步迭代优化, 也可以避免最小二乘法中的大矩阵求逆运算, 是一种高效的波束形成设计方法.

本征滤波器的系数由一组特征向量的元素组成, 该方法最早为设计数字滤波器提出<sup>[7]</sup>, 后来扩展到波束形成器的设计. Korompis等提出一种最大能量阵列代价函数<sup>[8]</sup>, 即通带能量与阻带能量之比, 可以通过特征滤波法求其最大解, 从而设计固定时域宽带波束形成器. 2002年, Doclo和Moonen提出一种总体最小均方本征滤波方法<sup>[9]</sup>, 优化宽带阵列响应与参考响应的均方误差, 构造

波束形成器, 2003年又提出远场、近场和两者混合的情况下基于本征滤波的宽带波束形成方法<sup>[10]</sup>. 文献[8-10]提出的本征滤波的波束形成方法的缺点是, 未设计阵列响应的频率不变性和零陷. 由于传统宽带波束形成的主瓣宽度随频率变化, 当目标信号从主瓣方向但非主极大方向入射时, 信号波束输出频谱就会失真, 设计频率不变性可以解决这个问题, 主瓣内保持频率不变性可以得到固定束宽波束形成器. 近年来, 国内外学者对波束形成器的频率不变性进行了深入的研究<sup>[11-15]</sup>. Liu和Weiss将阵列响应构造为与频率无关的二维傅里叶变换形式<sup>[11]</sup>, 实现了频率不变性. Duan等<sup>[12]</sup>提出了阵列空间响应偏差(Spatial Response Variation, SRV), Zhao等<sup>[13]</sup>、范展和梁国龙<sup>[14]</sup>、陈明建和罗景青<sup>[15]</sup>分别使用SRV来获取波束形成的频率不变性. 另外, 随着应用环境越来越复杂, 波束形成的零陷设计也尤为重要. 范展和梁国龙在基于凸优化的宽带波束形成方法中,

收稿日期: 2014-12-11; 录用日期: 2015-01-12; 网络出版时间: 2015-02-25 15:32

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150225.1532.006.html

作者简介: 王剑书(1989—), 男, 四川南充人, 博士研究生, 565563085@163.com

\* 通讯作者: 樊养余(1960—), 男, 陕西西安人, fan\_yangyu@sina.com, 主要研究方向为图像处理、信号处理、虚拟现实技术等.

**引用格式:** 王剑书, 樊养余, 杜瑞, 等. 改进的基于本征滤波的时域宽带波束形成[J]. 北京航空航天大学学报, 2015, 41(10): 1830-1835. Wang J S, Fan Y Y, Du R, et al. Improved eigenfilter-based time-domain wideband beamforming[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1830-1835 (in Chinese).

对不等式约束的干扰方向进行展宽,获得了展宽零陷的效果<sup>[14]</sup>.

为实现基于本征滤波波束形成的频率不变性和零陷设计,本文提出了基于最大能量阵列代价函数和本征滤波的波束形成的改进方法. 首先将 SRV 整合入最大能量阵列代价函数,以获取频率不变性;然后使用提供的两种零陷设计方法构造约束优化问题;最后由本征滤波法求解. 两种零陷设计方法分别为:扩展旁瓣能量法(对旁瓣能量中干扰方向的阵列响应赋予一个相对较大的权值,形成扩展旁瓣能量)和线性约束法(对干扰方向阵列响应增加线性约束). 仿真结果验证了本文两种方法在频率不变性和零陷设计上的有效性,同时均能获得低旁瓣增益,证明了本文方法的有效性和良好性能.

### 1 时域宽带波束形成模型

考虑  $M$  元各向同性的阵元组成的任意结构阵列,图 1 为时域宽带波束形成器的结构图. 图中  $\Delta$  为时延一个采样间隔,每个阵元接收到的信号经过一个  $J$  阶的有限长单位冲激响应(FIR)滤波器,然后在输出端进行叠加,形成波束输出,写成向量形式为

$$y(t) = \mathbf{w}^H \mathbf{x}(t) \quad (1)$$

式中:  $\mathbf{w} = [w_{0,0}, w_{1,0}, \dots, w_{M-1,0}, w_{0,1}, w_{1,1}, \dots, w_{M-1,J-1}]^T$ ;  $\mathbf{x} = [x_0(t), x_1(t), \dots, x_{M-1}(t), x_0(t - T_s), x_1(t - T_s), \dots, x_{M-1}(t - (J-1)T_s)]^T$ ,  $T_s$  为一个采样间隔时间;  $t$  为时间;  $H$  为共轭转置运算;  $T$  为转置运算. 根据傅里叶反变换理论,假设角频率范围为  $[\omega_1, \omega_h]$  的有限带宽信号为  $x_0(t) = \frac{1}{2\pi} \int_{\omega_1}^{\omega_h} X_0(\omega) e^{j\omega t} d\omega$ , 其中  $j = \sqrt{-1}$ ,  $\omega$  为频率,  $X_0(\omega)$  为  $x_0(t)$  的频谱密度函数,则

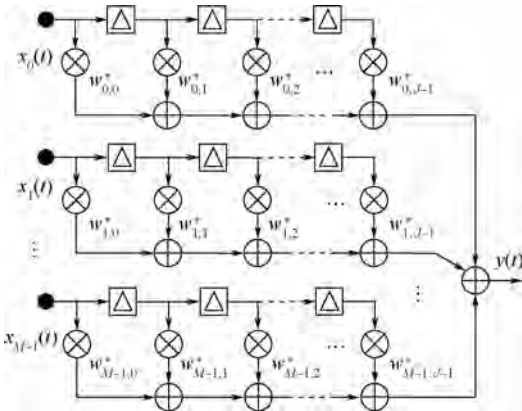


图 1 时域宽带波束形成的结构图

Fig. 1 Structure diagram of time-domain wideband beamforming

$$x_m(t) = \frac{1}{2\pi} \int_{\omega_1}^{\omega_h} X_0(\omega) e^{j\omega[t - (\tau_m + iT_s)]} d\omega \quad (2)$$

式中:  $\tau_m$  为第  $m+1$  个传感器接收信号相对参考传感器的时延,  $m=0, 1, \dots, M-1$ ;  $i=0, 1, \dots, J-1$ . 阵列输出为

$$y(t) = \frac{1}{2\pi} \int_{\omega_1}^{\omega_h} X_0(\omega) e^{j\omega t} \sum_{m=0}^{M-1} \sum_{i=0}^{J-1} e^{-j\omega(\tau_m + iT_s)} w_{m,i}^* d\omega = \frac{1}{2\pi} \int_{\omega_1}^{\omega_h} X_0(\omega) e^{j\omega t} P(\omega, \theta) d\omega \quad (3)$$

式中:  $P(\omega, \theta)$  为波束形成器关于角度  $\theta$  和频率  $\omega$  独立的响应函数,称为阵列响应,表示  $x_0(t)$  中角频率为  $\omega$  的窄带分量的加权系数;  $w_{m,i}^*$  为第  $m$  阵元对应的 FIR 滤波器的第  $i$  个系数.  $P(\omega, \theta)$  可以表示为向量形式:

$$P(\omega, \theta) = \mathbf{w}^H \mathbf{a}(\omega, \theta) \quad (4)$$

式中:  $\mathbf{a}(\omega, \theta)$  为时域宽带波束形成器的阵列流型向量,其元素为

$$\mathbf{a}(\omega, \theta) = [e^{-j\omega\tau_0}, e^{-j\omega\tau_1}, \dots, e^{-j\omega\tau_{M-1}}, e^{-j\omega(\tau_0 + T_s)}, e^{-j\omega(\tau_1 + T_s)}, \dots, e^{-j\omega(\tau_{M-1} + (J-1)T_s)}]^T \quad (5)$$

### 2 基于本征滤波的宽带波束形成

#### 2.1 最大能量阵列代价函数

最大能量阵列代价函数定义为通带能量与阻带能量的比率,对于宽带波束形成,即为主瓣能量与旁瓣能量的比率:

$$G = \frac{\int_{\Omega_{PB}} \int_{\Theta_{ML}} |\mathbf{w}^H \mathbf{a}(\omega, \theta)|^2 d\omega d\theta}{\int_{\Omega_{PB}} \int_{\Theta_{SL}} |\mathbf{w}^H \mathbf{a}(\omega, \theta)|^2 d\omega d\theta} \quad (6)$$

式中:  $\Omega_{PB}$ ,  $\Theta_{ML}$  和  $\Theta_{SL}$  分别为频率通带范围、主瓣角度范围和旁瓣角度范围. 对式(6)运算得

$$G = \frac{\int_{\Omega_{PB}} \int_{\Theta_{ML}} \mathbf{w}^H \mathbf{A}(\omega, \theta) \mathbf{w} d\omega d\theta}{\int_{\Omega_{PB}} \int_{\Theta_{SL}} \mathbf{w}^H \mathbf{A}(\omega, \theta) \mathbf{w} d\omega d\theta} = \frac{\mathbf{w}^H \mathbf{A}_{ML} \mathbf{w}}{\mathbf{w}^H \mathbf{A}_{SL} \mathbf{w}} \quad (7)$$

式中:  $\mathbf{A}(\omega, \theta) = \mathbf{a}(\omega, \theta) \mathbf{a}(\omega, \theta)^H$

$$\mathbf{A}_{ML} = \int_{\Omega_{PB}} \int_{\Theta_{ML}} \mathbf{A}(\omega, \theta) d\omega d\theta$$

$$\mathbf{A}_{SL} = \int_{\Omega_{PB}} \int_{\Theta_{SL}} \mathbf{A}(\omega, \theta) d\omega d\theta$$

#### 2.2 主极大方向阵列响应线性约束

为保证从主极大方向  $\theta_0$  入射的信号,通带内频谱和相位不发生畸变,对该方向阵列响应进行约束:

$$P(\omega, \theta_0) = \mathbf{a}(\omega, \theta_0)^H \mathbf{w} = 1 \quad (8)$$

式中:  $\omega \in \Omega_{PB}$ .

将式(8)写成向量形式:

$$C^H w = g \tag{9}$$

式中:  $C = [a(\omega_0, \theta_0), a(\omega_1, \theta_0), \dots, a(\omega_k, \theta_0)]$ ,  $\omega_0, \omega_1, \dots, \omega_k$  为  $\Omega_{PB}$  内频率的离散采样;  $g$  为元素均为 1 的  $K$  维列向量. 对目标函数(式(7))增加该线性约束, 可得优化问题:

$$\max_w \frac{w^H A_{ML} w}{w^H A_{SL} w} \quad \text{s. t. } C^H w = g \tag{10}$$

### 2.3 优化问题求解

将式(10)优化问题变换进行变换:

$$\max_{\hat{w}} \frac{\hat{w}^H B_1 \hat{w}}{\hat{w}^H B_2 \hat{w}} \quad \text{s. t. } \hat{C}^H \hat{w} = 0 \tag{11}$$

式中:  $\hat{w} = [w^H, -1]^H$

$$B_1 = \begin{bmatrix} A_{ML} & 0 \\ 0 & 0 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} A_{SL} & 0 \\ 0 & 0 \end{bmatrix}$$

$$\hat{C} = [C^H, g]^H$$

由等式约束条件易知  $\hat{w}$  位于  $\hat{C}^H$  的零空间  $Z$ , 则存在非零向量  $h$  使  $\hat{w} = Zh$  成立, 式(11)约束优化问题转化为无约束优化问题<sup>[9]</sup>:

$$\max_h \frac{h^H Z^H B_1 Z h}{h^H Z^H B_2 Z h} \tag{12}$$

由广义本征滤波问题可知,  $h$  的最优解为  $Z^H B_1 Z$  相对于  $Z^H B_2 Z$  的最大广义特征值对应的广义特征向量.

## 3 改进的方法

对常规的使用最大能量阵列代价函数的本征滤波宽带波束形成算法, 本文对其目标函数和线性约束条件进行改进, 从而实现波束形成器的频率不变性与零陷设计.

### 3.1 频率不变性设计

SRV 的值可以用来衡量波束形成器的频率不变性, 其定义为

$$Q_{SRV} = \int_{\Omega_{PB}} \int_{\Theta_{FI}} |w^H a(\omega, \theta) - w^H a_r(\omega, \theta)|^2 d\omega d\theta = w^H A_{FI} w \tag{13}$$

式中:  $A_{FI} = \int_{\Omega_{PB}} \int_{\Theta_{FI}} (a(\omega, \theta) - a_r(\omega, \theta))(a(\omega, \theta) - a_r(\omega, \theta))^H d\omega d\theta$ ;  $\Theta_{FI}$  为需要设计频率不变性的角度范围;  $a_r(\omega, \theta)$  为参考频率处的阵列流型.  $Q_{SRV}$  为衡量通带内各频率阵列响应与参考频率阵列响应之间的距离, 其值越小, 波束形成器在  $\Omega_{PB}$  和  $\Theta_{FI}$  内具有越好的频率不变性. 将 SRV 整合入最大能量目标函数, 形成优化模型:

$$\max_w \frac{w^H A_{ML} w}{w^H (A_{SL} + \alpha A_{FI}) w} \quad \text{s. t. } C^H w = g \tag{14}$$

式中:  $\alpha \geq 0$  为控制频率不变性强弱的权值.

### 3.2 零陷设计

针对具有频率不变性的优化模型(式(14)), 提出两种零陷设计方法: 扩展旁瓣能量法和线性约束法.

#### 3.2.1 扩展旁瓣能量法

最大能量阵列代价函数中, 旁瓣能量定义为  $w^H A_{SL} w$ , 对其适当变形可得

$$w^H A_{SL} w = \int_{\Omega_{PB}} \int_{\Theta_{SL}} |P(\omega, \theta) - 0|^2 d\omega d\theta \tag{15}$$

式(15)可以看做通带内旁瓣阵列响应与 0 的差值加权平方和(权值相等), 对零陷角度范围  $\Theta_{NL}$  ( $\Theta_{NL} \subset \Theta_{SL}$ ) 内的阵列响应分配更大的权值可以使旁瓣能量突出零陷处阵列响应对其的贡献. 将式(15)改写为  $\int_{\Omega_{PB}} \int_{\Theta_{SL}} |\beta_\theta P(\omega, \theta) - 0|^2 d\omega d\theta$ , 称为扩展旁瓣能量, 其中  $\beta_\theta$  为权系数,  $\theta \in \Theta_{SL}$ , 当  $\theta \in \Theta_{NL}$  时  $\beta_\theta$  为一个大于 1 的参数, 其他角度范围  $\beta_\theta = 1$ . 化为更简洁的形式:

$$\int_{\Omega_{PB}} \int_{\Theta_{SL}} |\beta_\theta P(\omega, \theta) - 0|^2 d\omega d\theta = w^H (A_{SL} + \lambda A_{NL}) w \tag{16}$$

式中:  $A_{NL} = \int_{\Omega_{PB}} \int_{\Theta_{NL}} A(\omega, \theta) d\omega d\theta$ ;  $\lambda \geq 0$  为控制零陷深度的参数. 使用扩展旁瓣能量代替目标函数中的旁瓣能量, 可得优化模型:

$$\max_w \frac{w^H A_{ML} w}{w^H (A_{SL} + \lambda A_{NL} + \alpha A_{FI}) w} \quad \text{s. t. } C^H w = g \tag{17}$$

#### 3.2.2 线性约束法

原优化模型对主瓣主极大方向的阵列响应进行了线性约束, 同样可以对干扰方向阵列响应进行约束, 从而实现零陷效果. 设干扰方向阵列响应为

$$P(\omega_k, \theta_u) = a(\omega_k, \theta_u)^H w = \xi(\omega_k, \theta_u) \exp(j\varphi(\omega_k, \theta_u)) \tag{18}$$

式中:  $\omega_k \in \Omega_{PB}$ ;  $\theta_u \in \Theta_{NL}$ ;  $\xi(\omega_k, \theta_u)$  和  $\varphi(\omega_k, \theta_u)$  分别为  $P(\omega_k, \theta_u)$  的幅度和相位; 由于主瓣主极大方向阵列响应的幅度约束为 1, 可得零陷  $(\omega_k, \theta_u)$  处阵列响应幅度增益为  $-20 \lg(1/\xi(\omega_k, \theta_u))$  dB. 式(18)可改写为

$$C_{NL}^H w = g_{NL} \tag{19}$$

式中:  $C_{NL} = [C_1, C_2, \dots, C_U]$

$$C_u = [a(\omega_0, \theta_u), a(\omega_1, \theta_u), \dots, a(\omega_K, \theta_u)]$$

$$g_{NL} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_U \end{bmatrix}$$

$$l_u = \begin{bmatrix} \xi(\omega_0, \theta_u) \exp(j\varphi(\omega_0, \theta_u)) \\ \xi(\omega_1, \theta_u) \exp(j\varphi(\omega_1, \theta_u)) \\ \vdots \\ \xi(\omega_K, \theta_u) \exp(j\varphi(\omega_K, \theta_u)) \end{bmatrix}$$

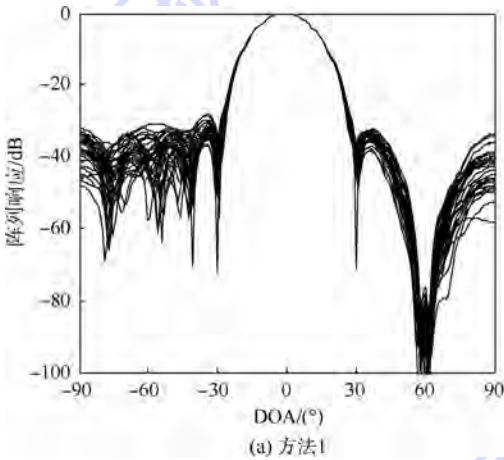
其中:  $u = 1, 2, \dots, U$ ;  $\theta_1, \theta_2, \dots, \theta_U$  为  $\Theta_{NL}$  内角度的离散采样. 将式(19)与式(9)合并, 代入优化模型式(14)得

$$\max_w \frac{w^H A_{ML} w}{w^H (A_{SL} + \alpha A_{FI}) w} \quad \text{s. t.} \quad \begin{bmatrix} C^H \\ C_{NL}^H \end{bmatrix} w = \begin{bmatrix} g \\ g_{NL} \end{bmatrix} \quad (20)$$

优化模型式(20)可以获得指定大小的零陷.

## 4 仿真分析

考虑 12 元的均匀线阵, 阵元各向同性, 阵元间距为通带最大频率对应的半波长. 定义采样频率的一半为归一化频率 1, 频率通带范围为  $[f_l, f_h] = [0.4, 1]$ , 参考频率选为 0.6. 将通带范围划



分为 36 个等宽子带, FIR 滤波器阶数  $J = 15$ . 设定主极大方向为  $0^\circ$  (线阵法线方向), 主瓣范围为  $\Theta_{ML} = [-15^\circ, 15^\circ]$ , 旁瓣范围为  $\Theta_{SL} = [-90^\circ, -25^\circ] \cup [25^\circ, 90^\circ]$ , 以  $2^\circ$  为间隔进行角度离散化.

### 4.1 频率不变性和零陷验证

设置频率不变性范围  $\Theta_{FI} = [-15^\circ, 15^\circ]$ , 干扰角度范围  $\Theta_{NL} = [-58^\circ, 62^\circ]$ , 扩展旁瓣能量法 (记为“方法 1”) 优化模型式(17)中, 设置参数  $\alpha = 200, \lambda = 10^5$ , 得到二维波束, 图如图 2(a) 所示. 图 2 中横坐标为信号的来向 (Direction of Arrival, DOA), 波束图绘制的角度间隔为  $1^\circ$ . 线性约束法 (记为“方法 2”) 优化模型式(20)中, 设置  $\alpha = 1, \xi(\omega_k, \theta_u) = 10^{-4}$  (对应零陷深度  $-80$  dB),  $\varphi(\omega_k, \theta_u) = 0$ , 其中  $k = 1, 2, \dots, K, u = 1, 2, \dots, U$ , 得到二维波束响应图, 如图 2(b) 所示.

由图 2 可以看出, 两种方法在  $\Theta_{FI}$  范围内各频率的阵列响应均具有很好的 consistency, 验证了本文方法频率不变性设计的有效性; 两种方法各频率的阵列响应均在  $[-58^\circ, 62^\circ]$  范围形成了较深的零陷, 并且线性约束法准确地形成了  $-80$  dB 的深度, 验证了本文两种零陷设计方法的有效性.

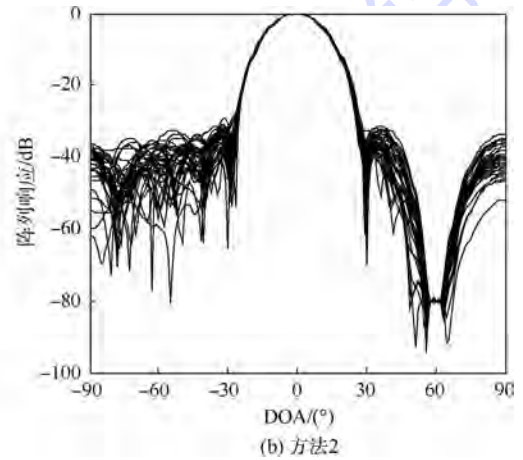


图 2 方法 1 和方法 2 的二维波束响应图

Fig. 2 Two-dimensional array response of method 1 and method 2

### 4.2 性能分析

为证明本文方法的良好性能, 下面将对主瓣 SRV 值、波束宽度、旁瓣增益 (定义为旁瓣最大阵列响应幅度) 进行综合分析, 并与文献[14]提供的基于凸优化的波束形成法进行对比. 本文方法参数设置与第 4.1 节相同, 文献[14]的方法使用与本文仿真一样的假设, 相关参数设置为, 恒定束宽加权因子  $F(\omega, \theta) = 1$ , 约束值  $\delta = 1 \times 10^{-6}$ , 零陷深度  $\xi = 10^{-4}$  (对应  $-80$  dB), 不作白噪声输出

功率约束, 得到二维波束响应图如图 3 所示. 对 3 种方法得到的波束图计算其评价指标: 主瓣 SRV 值、波束宽度和旁瓣增益, 结果如表 1 所示.

由图 2、图 3 可以看出, 3 种方法均在  $[-58^\circ, 62^\circ]$  范围内获得了约为  $-80$  dB 的零陷, 同时主瓣具有良好不变性, 即实现了固定束宽. 由表 1 可知, 3 种方法均获得非常小的主瓣 SRV 值, 保证了波束图的主瓣  $\Theta_{ML}$  范围内的频率不变性和固定波束宽度, 本文两种方法获得了更低的主瓣 SRV

值;3种方法均获得的固定束宽,均为 $[-10^\circ, 9^\circ]$ ;本文两种方法获得的旁瓣增益在 $-25\text{ dB}$ 以下,略低于文献[14]的方法,能更好地抑制旁瓣入射的干扰或噪声.另外,本文方法不需要迭代优化,具有更高的运算效率.

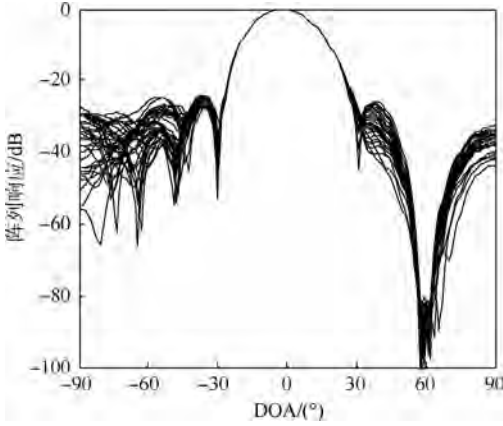


图3 文献[14]二维波束响应图

Fig.3 Two-dimensional array response of method in Ref. [14]

表1 3种方法评价指标

Table 1 Evaluation of three methods

评价指标	主瓣 SRV 值	波束宽度	旁瓣增益/dB
文献[14]方法	$5.88 \times 10^{-7}$	$[-10^\circ, 9^\circ]$	-24.03
方法1	$1.47 \times 10^{-10}$	$[-10^\circ, 9^\circ]$	-28.4
方法2	$4.97 \times 10^{-10}$	$[-10^\circ, 9^\circ]$	-29.54

## 5 结论

本文基于最大能量阵列代价函数和主极大方向线性约束的优化模型,提出了改进的基于本征滤波的时域宽带波束形成方法,以获得阵列响应的频率不变性和零陷效果,经实验验证表明:

- 1) 两种方法均能实现阵列响应在指定角度范围内的频率不变性或零陷.
- 2) 在获得足够零陷深度和相同波束宽度条件下,两种方法均获得比文献[14]中方法更小的主瓣 SRV 值和旁瓣增益.

## 参考文献 (References)

[1] Liu W, Weiss S. Wideband beamforming: Concepts and techniques[M]. Chichester: Wiley-Blackwell, 2010:4-5.  
 [2] Ström M, Viberg M, Falk K. Robust transceiver design for wideband MIMO radar utilizing a subarray antenna structure[J]. Sig-

nal Processing, 2013, 93(12):3541-3552.  
 [3] Somasundaram S D. Wideband robust capon beamforming for passive sonar[J]. IEEE Journal of Oceanic Engineering, 2013, 38(2):308-322.  
 [4] Feng Z G, Yiu K F C, Nordholm S E. Placement design of microphone arrays in near-field broadband beamformers [J]. IEEE Transactions on Signal Processing, 2012, 60(3):1195-1204.  
 [5] Tanaka S, Harada A, Sawahashi M, et al. Experiments on coherent adaptive antenna array diversity for wideband DS-CDMA mobile radio[J]. IEEE Journal on Selected Areas in Communications, 2000, 18(8):1495-1504.  
 [6] Huang B S, Huang Y L, et al. Estimation of the rupture velocity and fault length of the 2004 Sumatra-Andaman earthquake using a dense broadband seismic array in Taiwan[J]. Journal of Asian Earth Sciences, 2011, 40(3):762-769.  
 [7] Makhoul J. On the eigenvectors of symmetric Toeplitz matrices [J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1981, 29(4):868-872.  
 [8] Korompis D, Yao K, Lorenzelli F. Broadband maximum energy array with user imposed spatial and frequency constraints[C]// Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway, NJ: IEEE Press, 1994, 4:529-532.  
 [9] Doclo S, Moonen M. Comparison of least-squares and eigenfilter techniques for broadband beamforming[C]// Proceedings of 3rd IEEE SPS. Piscataway, NJ: IEEE Press, 2002:73-76.  
 [10] Doclo S, Moonen M. Design of far-field and near-field broadband beamformers using eigenfilters [J]. Signal Processing, 2003, 83(12):2641-2673.  
 [11] Liu W, Weiss S. Design of frequency invariant beamformers for broadband arrays[J]. IEEE Transactions on Signal Processing, 2008, 56(2):855-860.  
 [12] Duan H, Ng B P, See C M S, et al. Applications of the SRV constraint in broadband pattern synthesis [J]. Signal Processing, 2008, 88(4):1035-1045.  
 [13] Zhao Y, Liu W, Langley R J. Adaptive wideband beamforming with frequency invariance constraints[J]. IEEE Transactions on Antennas and Propagation, 2011, 59(4):1175-1184.  
 [14] 范展, 梁国龙. 基于凸优化的最小旁瓣恒定束宽时域宽带波束形成[J]. 电子学报, 2013, 41(5):943-948.  
 Fan Z, Liang G L. Broadband beamforming with minimum side-lobe and constant beamwidth based on convex optimization[J]. Acta Electronica Sinica, 2013, 41(5):943-948 (in Chinese).  
 [15] 陈明建, 罗景青. 基于空间响应偏差约束的最小二乘宽带波束形成[J]. 电路与系统学报, 2013, 18(1):84-89.  
 Chen M J, Luo J Q. Least squares broadband beamforming based on the spatial response variation constraints[J]. Journal of Circuits and System, 2013, 18(1):84-89 (in Chinese).

## Improved eigenfilter-based time-domain wideband beamforming

WANG Jianshu, FAN Yangyu<sup>\*</sup>, DU Rui, LYU Guoyun

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China)

**Abstract:** An improved eigenfilter-based wideband beamforming method with frequency invariance and nulls was proposed. Based on the optimization model formed by maximum energy array cost function and linear constraints of array response in main great direction, array spatial response variation was integrated into the cost function to achieve frequency invariance. Two methods, integrating the extended sidelobe energy into cost function and using linear constraints of array response in the direction of inference, were designed to achieve nulls, so that two new constraint optimization models were formed, which could be solved by eigenfilter. Simulation results show that frequency invariance and nulls of beam response are achieved at the specified angle region in the frequency passband, and the gain of the sidelobe is small, the effectiveness and satisfactory performance of the proposed method are proved.

**Key words:** wideband beamforming; eigenfilter; frequency invariance; nulls; sidelobe

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0780

# 谐振式集成光学陀螺中电流调谐致强度噪声研究

鄧銀周, 馮麗爽\*, 王俊杰, 唐依創

(北京航空航天大学 仪器科学与光电工程学院, 北京 100191)



**摘 要:** 采用半导体激光器(LD)是谐振式集成光学陀螺(RIOG)小型化发展的重要趋势,但是LD的电流调谐特性会引入强度噪声.理论分析了电流调谐致强度噪声对开环RIOG性能的影响,得到频率锁定位置与LD输出光强之间为线性关系.随后,提出了基于频率锁定位置的强度噪声抑制方案,通过频率锁定位置实时监测LD输出光强的变化,根据监测值得到补偿系数,并补偿陀螺输出.建立了抑制电流调谐致强度噪声的理论模型,并进行了实验验证,RIOG的零偏稳定性由 $0.768(^{\circ})/s$ 提高到 $0.103(^{\circ})/s$ .

**关键词:** 谐振式集成光学陀螺(RIOG); 半导体激光器(LD); 电流调谐; 强度噪声; 补偿

中图分类号: TN815; TN911.74

文献标识码: A 文章编号: 1001-5965(2015)10-1836-06

谐振式集成光学陀螺(Resonator Integrated Optic Gyro, RIOG)是基于Sagnac效应的惯性角速度传感器,它具有全固态结构、精度高、体积小、动态范围大、可批量生产等优点<sup>[1-3]</sup>.RIOG利用谐振原理,要求光源具有窄线宽、可调谐特性<sup>[4]</sup>.YAG固体激光器<sup>[5-6]</sup>、He-Ne气体激光器<sup>[7-8]</sup>和分布式反馈光纤激光器<sup>[9-10]</sup>均可满足RIOG的需求,并且已经被用于RIOG系统中.虽然这些激光器的线宽窄、频率性能优越,但是它们的体积大、功耗高,难以实现工程化样机.随着激光器工艺技术的进步,目前半导体激光器(Laser Diode, LD)不仅具备窄线宽、可调谐、频率稳定等特点,还具有体积小优势,采用LD是RIOG小型化发展的重要趋势<sup>[11-12]</sup>.

基于LD的RIOG中,为了检测Sagnac效应产生的谐振频差,通过电流调谐将LD的中心频率锁定在谐振腔的谐振频率处<sup>[13]</sup>.由于驱动电流

的调节同时改变LD的输出频率和光强,必然引入强度噪声.文献[14]已经分析了光源强度噪声对开环RIOG性能的影响,得到光源强度噪声对陀螺性能的影响与固有零偏值相关.为了消除开环RIOG中的这种电流调谐致强度噪声,文献[15]提出了基于模拟减法电路的强度噪声抑制方案.但是该方案存在两个不足之处:①需要增加新的电路器件,不利于系统的小型化;②模拟减法电路控制精度低,容易受环境因素的影响.文献[16]基于声光移频器实现了RIOG的闭环检测,理论上可以消除光源强度噪声对陀螺性能的影响,但是声光移频器不仅增加了光路器件,还引入了新的噪声源,不利于系统集成和陀螺精度的提高.本文在研究LD电流调谐致强度噪声的特性基础上,提出了基于频率锁定位置的强度噪声抑制方案,该方案采用数字补偿方法,不需要增加任何光路和电路器件,利于系统的小型化.本文还搭

收稿日期: 2014-12-12; 录用日期: 2015-02-25; 网络出版时间: 2015-04-14 16:00

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150414.1600.001.html

基金项目: 国家自然科学基金(61171004); 北京航空航天大学博士研究生创新基金(YWF-14-YJSY-039)

作者简介: 鄧銀周(1986—),男,河南开封人,博士研究生, zhiyinzhou@163.com

\* 通讯作者: 馮麗爽(1968—),女,河北辛集人,研究员, fenglishuang@buaa.edu.cn, 主要研究方向为微光机电系统、先进光纤传感技术.

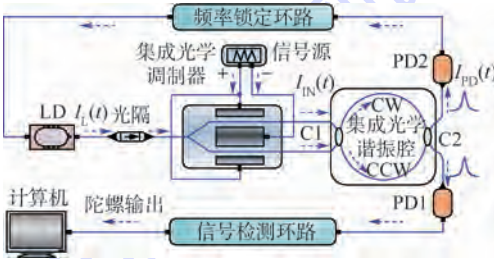
**引用格式:** 鄧銀周, 馮麗爽, 王俊杰, 等. 谐振式集成光学陀螺中电流调谐致强度噪声研究[J]. 北京航空航天大学学报, 2015, 41(10): 1836-1841. Zhi Y Z, Feng L S, Wang J J, et al. Study on current-tuning-induced intensity noise in resonator integrated optic gyro[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1836-1841 (in Chinese).

建了实验系统,验证了方案的可行性,有效抑制了开环 RIOG 的电流调谐致强度噪声.

## 1 电流调谐致强度噪声分析与抑制

### 1.1 RIOG 基本原理

图 1 给出了基于 LD 的开环 RIOG 结构框图, LD 发出的光经光隔和集成光学调制器后,被分为两束相等的光,并受到集成光学调制器的三角波相位调制作用. 两束光经耦合器 C1 耦合进入集成光学谐振腔,在谐振腔内分别沿顺时针和逆时针方向传输. 传输不同圈数的顺时针光和逆时针光经耦合器 C2 耦合出谐振腔时分别发生多光束干涉,干涉光强由光电探测器 PD1 和 PD2 探测. 当陀螺旋转时,顺时针光与逆时针光的谐振频差与旋转角速度成正比. 为了检测该谐振频差信号,频率锁定环路根据 PD2 的解调输出反馈控制 LD 的驱动电流,将 LD 的中心频率锁定在逆时针光的谐振频率处,而信号检测环路根据 PD1 的解调输出得到正比于角速度的陀螺输出信号.



$I_L(t)$ —LD 的输出光强; $I_{IN}(t)$ —集成光学谐振腔的输入光强;  
 $I_{PD}(t)$ —集成光学谐振腔的输出光强;CW—顺时针方向;  
CCW—逆时针方向

图 1 基于 LD 的开环 RIOG 结构框图

Fig. 1 Structure block diagram of open-loop RIOG based on LD

### 1.2 电流调谐致强度噪声分析

根据图 1 所示的开环 RIOG 检测原理,LD 的中心频率与逆时针光的谐振频率相等,但随着应力、温度、旋转角速度等因素的变化,逆时针光的谐振频率在不断变化. 当 LD 的中心频率与逆时针光的谐振频率不相等时,频率锁定环路根据 PD2 解调得到的误差信号反馈控制 LD 的驱动电流,调整其中心频率,实现 LD 中心频率对逆时针光谐振频率的实时锁定. 此时,LD 的驱动电流  $i(t)$  与逆时针光的谐振频率  $f_{CCW}$  之间的关系为

$$K_f i(t) + f_0 = f_{CCW} = \frac{qc}{n(t)L_{CCW}(t)} \quad (1)$$

式中: $K_f$  为 LD 的电流调谐系数; $f_0$  为 LD 的初始频率; $q$  为纵模阶数; $c$  为真空中的光速; $n(t)$  为集

成光学谐振腔的折射率; $L_{CCW}(t)$  为逆时针光路的等效腔长.

LD 的输出光强  $I_L(t)$  和驱动电流  $i(t)$  的关系如图 2 所示,其中  $i_{th}$  为 LD 的阈值电流. 当 LD 的驱动电流小于阈值电流时,如图 2 中区域 I 所示,LD 的输出只有很微弱的光,并且输出光强与驱动电流无关. 当 LD 的驱动电流大于阈值电流时,输出光强随着驱动电流的增大而增加,并且输出光强和驱动电流之间近似为线性关系,如图 2 中区域 II 所示. RIOG 中 LD 工作在区域 II,在该区域内光强  $I_L(t)$  和驱动电流  $i(t)$  的关系可以近似表示为<sup>[15]</sup>

$$I_L(t) = K_1(i(t) - i_{th}) \quad (2)$$

式中: $K_1$  为 LD 的强度调节系数.

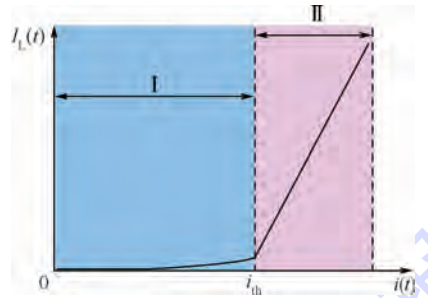


图 2 LD 的输出光强与驱动电流之间的关系示意图

Fig. 2 Illustration of relationship between intensity of LD and its driven current

将式(1)代入式(2),得到

$$I_L(t) = \frac{K_1 qc}{K_f n(t) L_{CCW}(t)} - \frac{K_1}{K_f} f_0 - K_1 i_{th} \quad (3)$$

因此,通过电流调谐实现频率锁定的同时,必然改变 LD 的输出光强,引入光源强度噪声. 这种电流调谐致强度噪声同时受环境因素和旋转角速度的影响,频率分量比较复杂,无法通过传统数字滤波滤除,严重影响 RIOG 的性能<sup>[13]</sup>.

RIOG 中的电流调谐致强度噪声可以等效为光源强度波动噪声,它通过改变谐振曲线影响陀螺的性能. 对于图 1 中 LD 的输出光强  $I_L(t)$ ,由于光隔和集成光学调制器的作用,逆时针光入腔前的光强可以表示为  $I_{IN}(t) = K_{loss} I_L(t)$ ,  $K_{loss}$  为光强入腔前的损耗系数. 当 LD 的中心频率与逆时针光的谐振频率之差为  $\Delta f_{CCW}(t)$  时,集成光学谐振腔的输出光强  $I_{PD}(t)$  可以表示为<sup>[17]</sup>

$$I_{PD}(t) = K_{loss} I_L(t) \frac{1}{1 + \left( \frac{\Delta f_{CCW}(t)}{\Gamma/2} \right)^2} \quad (4)$$

式中: $\Gamma$  为谐振曲线的半高全宽.

根据式(4),对于一定的频差  $\Delta f_{CCW}(t)$ ,谐振

腔的输出光强  $I_{PD}(t)$  与 LD 的输出光强  $I_L(t)$  之间为线性关系. 图 3 给出了当谐振腔的输入光强  $I_{IN}$  分别为  $0.9I_0$ 、 $1.0I_0$ 、 $1.1I_0$ 、 $1.2I_0$  时的仿真谐振曲线, 从图 3 中可以看出, 当 LD 的输出光强波动引起入腔光强变化时, 会线性改变谐振曲线的形状. 根据三角波相位调制解调原理, 对于相同的调制参数, 谐振曲线的变化会改变陀螺输出环路的增益, 造成陀螺标度因数的变化, 最终影响 RIOG 的性能<sup>[18-19]</sup>.

中的电流调谐致强度噪声提供一种方法.

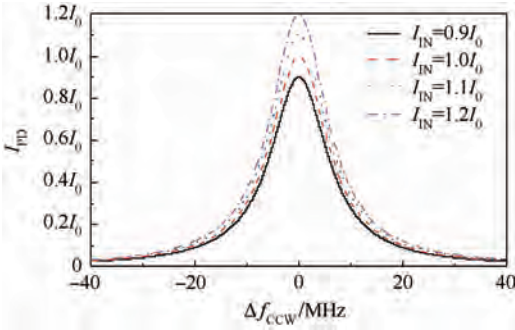


图3 谐振腔输入光强波动对谐振曲线的影响  
Fig. 3 Fluctuation influence of resonator's input intensity on resonance curve

文献[14]已经分析了光源强度波动对开环 RIOG 输出的影响, 陀螺输出  $G_{out}(t)$  与光源强度  $I_L(t)$  之间为线性关系, 并且与陀螺的固有零偏值相关. 开环陀螺输出  $G_{out}(t)$  可以表示为

$$G_{out}(t) = I_L(t) K_C \Delta f(t) \quad (5)$$

式中:  $\Delta f(t)$  为顺逆时针光之间的谐振频差, 其中包含陀螺的固有非互易频差;  $K_C$  为陀螺输出与  $\Delta f(t)$  之间的归一化比例系数. 依据式(5)所示的线性关系, 为了消除电流调谐致强度噪声对陀螺性能的影响, 可监测 LD 输出光强的变化, 然后对 RIOG 的输出进行实时补偿.

三角波相位调制等效于方波频率调制<sup>[20]</sup>, 当 LD 的中心频率被锁定在逆时针光的谐振频率时, LD 的中心频率与逆时针光的谐振频率之差为零. 此时, 如图 4 所示, 逆时针光路的输出为一恒定光强值  $I_{lock}$ , 结合式(4), 可以得到频率锁定时的输出光强  $I_{lock}$  可以表示为

$$I_{lock}(t) = K_{loss} I_L(t) \frac{1}{1 + \left(\frac{f_{FM}}{f/2}\right)^2} \quad (6)$$

式中:  $f_{FM}$  为三角波相位调制等效的方波频率调制幅度.

根据式(6),  $I_{lock}(t)$  与  $I_L(t)$  之间为线性关系. 因此, 可以建立  $I_{lock}(t)$  与陀螺输出之间的理论模型, 然后通过  $I_{lock}(t)$  实时地监测 LD 输出光强的变化, 并对光源强度波动进行补偿, 为消除开环 RIOG

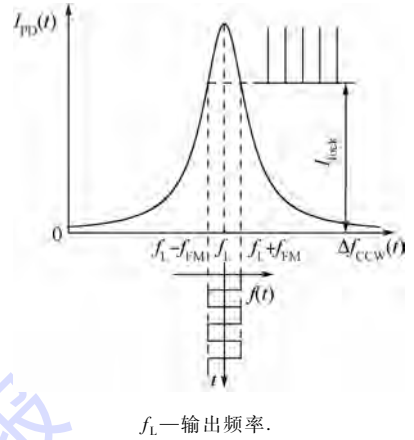


图4 频率锁定条件下的谐振腔输出光强  
Fig. 4 Output intensity of resonator under condition of frequency locking

### 1.3 抑制方案

基于第 1.2 节分析, 提出了基于频率锁定位置的强度噪声抑制方案. 基本原理方案如图 5 所示, PD2 的输出电压给出了频率锁定位置, 它与频率锁定时逆时针光路的输出光强成正比, 并且实时反应了 LD 输出光强的变化趋势. 信号检测环路一方面采集 PD2 输出的电压直流信号, 用于监测 LD 的光强波动情况, 得到补偿系数; 另一方面根据 PD1 得到解调输出信号, 将解调输出信号与补偿系数相乘, 消除光强波动对陀螺输出的影响, 达到抑制电流调谐致强度噪声的目的.

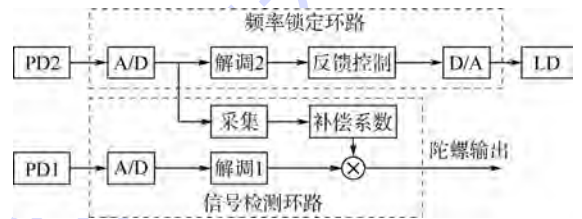


图5 基于频率锁定位置的强度噪声抑制方案示意图  
Fig. 5 Scheme illustration of intensity noise suppression based on frequency locking position

在频率锁定条件下, 假定图 5 中 LD 的初始光强  $I_0$  与 PD2 输出的初始频率锁定位置  $D_{lock}$  之间有如下的比例关系:

$$D_{lock} = K_{lock} I_0 \quad (7)$$

式中:  $K_{lock}$  为两者之间的比例系数. 式(7)对光强  $I_0$  求导, 可以得到 PD 输出频率锁定位置的变化量为

$$\Delta D_{lock} = K_{lock} \Delta I_0 \quad (8)$$

根据式(8), 当 LD 的初始光强  $I_0$  受强度变化量  $\Delta I_0(t)$  影响时, RIOG 的输出可以表示为

$$G_{out}(t) = (\Delta I_0(t) + I_0) K_C \Delta f(t) \quad (9)$$

结合式(8)和式(9),得到反映光源强度变化趋势的补偿系数为

$$K_p(t) = \frac{D_{\text{lock}}}{\Delta D_{\text{lock}}(t) + D_{\text{lock}}} = \frac{I_0}{\Delta I_0(t) + I_0} \quad (10)$$

将 RIOG 的输出  $G_{\text{out}}(t)$  乘以补偿系数  $K_p(t)$ , 得到

$$G_{\text{out}}(t) = (\Delta I_0(t) + I_0) K_p(t) K_C \Delta f(t) = I_0 K_C \Delta f(t) \quad (11)$$

因此,补偿之后 RIOG 的输出  $G_{\text{out}}(t)$  中不再包含强度变化因子  $\Delta I_0(t)$ , 消除了强度噪声对陀螺性能的影响。

## 2 实验结果与讨论

根据图 1 建立了 RIOG 实验系统,其中 LD 的中心波长为 1550 nm,洛伦兹线宽为 1.17 kHz,最大输出光功率为 20 mW;集成光学谐振腔的长度为 18.8 cm,半高全宽为 13.4 MHz;调制三角波的频率为 1 MHz,调制电压峰峰值为 20 V;光电探测器 PD1 和 PD2 均采用 PIN-FET 组件,响应度为 0.92 A/W. 光电探测器 PD1 和 PD2 的输出电压经 A/D 转换后到达可编程逻辑器件的现场可编程门阵列(Field-Programmable Gate Array, FPGA),频率锁定、信号检测和实时补偿都基于 FPGA 实现。

根据 RIOG 的谐振特性,当集成光学谐振腔的入射光频率与谐振腔的谐振频率之差线性变化时,谐振腔的输出可探测到谐振曲线. 实验中默认短期内谐振腔的谐振频率保持不变,通过控制 FPGA 线性调节 LD 的驱动电流对 LD 扫频,测试得到集成光学谐振腔输出的谐振曲线如图 6 所示,验证了驱动电流对 LD 的线性调谐特性. 当扫频出现谐振峰时,启动 FPGA 内部的频率反馈控制模块,完成 LD 中心频率对逆时针光谐振频率的锁定. 频率锁定之后,PD2 探测到的谐振腔输出如图 7 所示,由于三角波相位调制等效的方波频率调制作用,在每个方波频率调制周期内,调制频率切换时会产生一个尖峰脉冲信号,造成谐振腔的输出并非理想的直流电压信号. 为了提高频率锁定位置的采集精度,信号采样时应避开尖峰脉冲区域,一个解调周期内的理想信号采样区域如图 7 中阴影所示. 根据三角波相位调制解调原理<sup>[17]</sup>,图 7 中信号采样区域①和②的共模分量反应了频率锁定位置,并且该共模分量不受频率锁定误差的影响,可用于实时监测 LD 输出光强的变化。

为了验证频率锁定位置与 LD 输出光强之间的线性关系,在频率锁定条件下,同时采集了光电探测器 PD2 的输出和 LD 的光强输出. 图 8 和图 9

分别给出了 PD2 输出的频率锁定位置数字量和 LD 的光强输出数字量,从图 9 中可以看出,受电流调谐和环境因素的影响,LD 输出的光强中叠加了一个低频噪声,该噪声无变化规律,但它与频率锁定位置的变化趋势一致. 将频率锁定位置数字量与光强数字量相除,图 10 给出了数据处理结果,两者的比值与光强波动无关,表明频率锁定位置可以实时地反应 LD 输出光强的变化趋势。

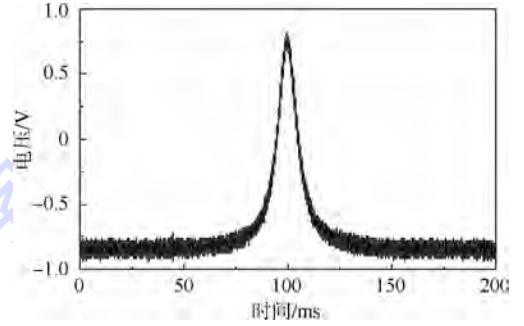


图 6 集成光学谐振腔的谐振曲线

Fig. 6 Resonance curve of integrated optic resonator

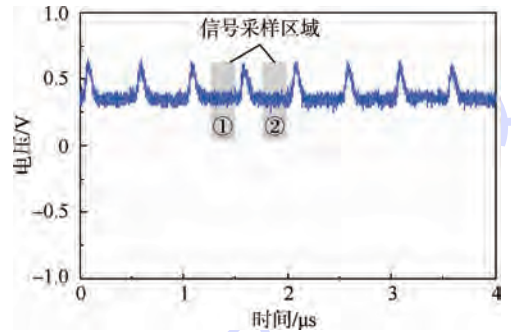


图 7 频率锁定时的 PD2 输出

Fig. 7 Output of PD2 under condition of frequency locking

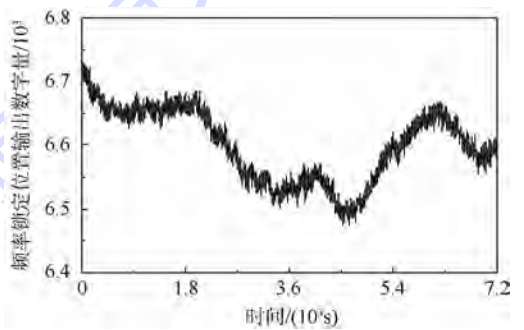


图 8 频率锁定位置输出

Fig. 8 Output of frequency locking position

理论模型的精确度严重制约了实时补偿算法的效果,而频率锁定位置数字量与光强数字量的比值在一定程度上反映了模型的精确度. 从图 10 中可以看出,两者之间的比值并非严格的直线,它围绕均值存在  $\pm 0.5$  的误差,分析主要有 3 个方面的原因:①频率锁定位置与输出光强的信号采集环路非互易,存在时间延时,影响两者的相关

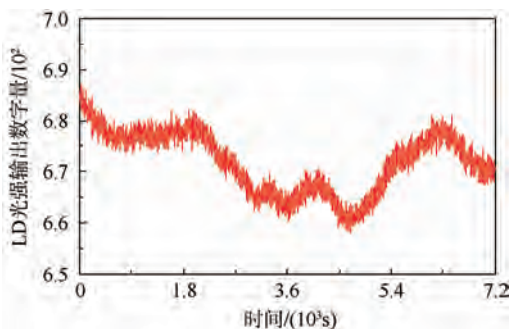


图9 LD光强输出

Fig. 9 Intensity output of LD

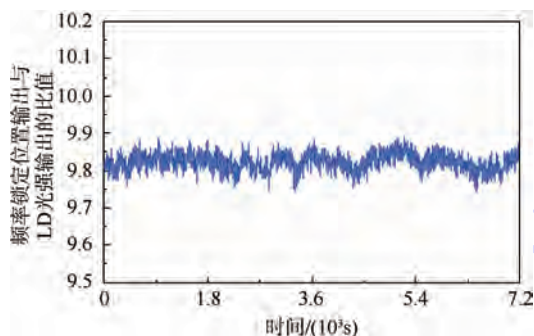


图10 频率锁定位置与LD光强之比

Fig. 10 Ratio between frequency locking position and LD intensity

性;②信号采集环路中的器件受环境噪声的影响,使两路信号分别产生增益误差,并将误差引入比值当中;③RIOG系统中光学器件受环境因素产生了光学误差,并随着光电转换引入采集结果当中.因此,为了提高实时补偿效果,实验中不但选择了环境性能较好的光路和电路器件,还保持频率锁定环路和信号检测环路的器件互易性,尽可能提高理论模型的精确度.

采集得到频率锁定位置的基础上,根据式(10)计算得到补偿系数,然后将补偿系数与PD1的解调输出相乘,实现对LD强度噪声的实时补偿,消除了强度噪声产生的陀螺标度因数变化.为了对比补偿算法对电流调谐致强度噪声的抑制效果,图11给出了实时补偿前后2h的陀螺静态测试数据,从图中可以看出,补偿之前陀螺输出与光强变化保持一致,显然受到电流调谐致强度噪声的影响.而补偿之后陀螺输出与光强波动不相关,证明光源强度噪声得到了有效抑制.对比补偿前后的陀螺性能,RIOG的零偏稳定性由 $0.768(^{\circ})/s$ 提高到了 $0.103(^{\circ})/s$ ,有效改善了陀螺输出精度.

基于频率锁定位置的强度噪声抑制方案不仅适用于陀螺静态测试,还同样适用于陀螺转动实验.当陀螺发生转动时,由于集成光学谐振腔中逆时针光的谐振频率会发生变化,频率锁定环路会

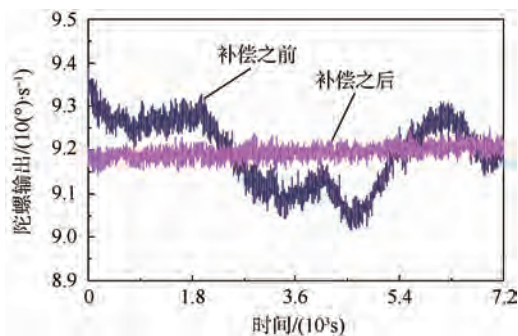


图11 补偿前后陀螺输出结果

Fig. 11 Gyro output before and after compensation

调节LD的中心频率实时跟踪并锁定逆时针光的谐振频率.虽然在频率跟踪与锁定过程中会存在频率锁定误差,但频率锁定位置为图7中采样区域①和采样区域②的共模分量,不受频率锁定误差的影响,仍可实时监测LD输出光强的变化,对光源强度波动进行补偿.此外,该方案基于FPGA对开环RIOG的输出进行实时数字补偿,既不会影响陀螺的动态特性,又不需要增加任何光路和电路器件,利于系统的小型化.

### 3 结论

本文在分析RIOG中电流调谐致强度噪声特性的基础上提出了一种基于频率锁定位置的强度噪声抑制方案,经理论分析和实验验证表明:

1) RIOG中的电流调谐致强度噪声可等效为光源强度波动噪声,它通过改变谐振曲线影响陀螺的性能.

2) 频率锁定位置与LD输出光强之间为线性关系,利用频率锁定位置可实时监测LD输出光强的波动.

3) 该方案基于FPGA对开环RIOG的输出进行实时数字补偿,既不会影响陀螺的动态特性,又不需要增加任何光路和电路器件.

4) RIOG的零偏稳定性由 $0.768(^{\circ})/s$ 提高到 $0.103(^{\circ})/s$ ,验证了方案的可行性.

为进一步提高该方案对电流调谐致强度噪声的抑制的效果,仍需改善理论模型的精确度.

### 参考文献 (References)

- [1] Hervé C L. 光纤陀螺仪[M]. 张桂才,王巍,译.北京:国防工业出版社,2002:157-167.  
Hervé C L. The fiber-optic gyroscope[M]. Translated by Zhang G C, Wang W. Beijing: National Defense Industry Press, 2002: 157-167 (in Chinese).
- [2] Mario N A, Caterina C, Francesco D O, et al. Advances in gyroscope technologies[M]. Berlin: Springer-Verlag, 2010: 37-46.

- [3] Feng L S, Wang J J, Zhi Y Z, et al. Transmissive resonator optic gyro based on silica waveguide ring resonator [J]. *Optics Express*, 2014, 22(22): 27565-27575.
- [4] Suzuki K, Takiguchi K, Hotate K. Monolithically integrated resonator microoptic gyro on silica planar lightwave circuit [J]. *Journal of Lightwave Technology*, 2000, 18(1): 66-72.
- [5] Sanders G A. Critical review of resonator fiber optic gyroscope technology [J]. *SPIE Fiber Optic Sensors*, 1992, CR44: 115-133.
- [6] Takiguchi K, Hotate K. Partially digital-feedback scheme and evaluation of optical Kerr-effect induced bias in optical passive ring-resonator gyro [J]. *IEEE Photonics Technology Letters*, 1991, 3(7): 679-681.
- [7] Meyer R E, Ezekiel S, Stowe D W, et al. Passive fiber-optic ring resonator for rotation sensing [J]. *Optics Letters*, 1983, 8(12): 644-646.
- [8] Zarinetchi F, Ezekiel S. Observation of lock-in behavior in a passive resonator gyroscope [J]. *Optics Letters*, 1986, 11(6): 401-403.
- [9] Smith P, Zarimetchi F, Ezekiel S. Narrow-linewidth stimulated Brillouin fiber laser and applications [J]. *Optics Letters*, 1991, 16(6): 393-395.
- [10] Wang X J, He Z Y, Hotate K. Reduction of polarization-fluctuation induced drift in resonator fiber optic with twin 90° polarization-axis rotated splices gyro by a resonator [J]. *Optics Express*, 2010, 18(2): 1677-1683.
- [11] 肖树妹, 梅海平, 黄宏华, 等. 半导体激光器电流频率调制率的测量 [J]. *红外与激光工程*, 2010, 39(增刊): 74-77.  
Xiao S M, Mei H P, Huang H H, et al. Measuring current-induced frequency modulation ratio in semiconductor diode laser [J]. *Infrared and Laser Engineering*, 2010, 39(Sup.): 74-77 (in Chinese).
- [12] Jin Z H, Yu X H, Ma H L. Resonator fiber optic gyro employing a semiconductor laser [J]. *Applied Optics*, 2012, 51(15): 2856-2864.
- [13] Lei M, Feng L S, Zhi Y Z, et al. Current modulation technique used in resonator micro-optic gyro [J]. *Applied Optics*, 2013, 52(2): 307-313.
- [14] Lei M, Feng L S, Zhi Y Z, et al. Effect of intensity variation of laser in resonator integrated optic gyro [J]. *Applied Optics*, 2013, 52(19): 4576-4581.
- [15] Lei M, Feng L S, Zhi Y Z, et al. Experiments on resonator micro-optic gyro using external cavity laser diode [J]. *Optical Engineering*, 2012, 51(10): 104602.
- [16] 王文怡, 王慧泉, 陈妍, 等. 双路闭环谐振式微光学院螺 [J]. *中国激光*, 2012, 39(12): 12080003.  
Wang W Y, Wang H Q, Chen Y, et al. Double closed-loop resonator micro optic gyro [J]. *Chinese Journal of Lasers*, 2012, 39(12): 12080003 (in Chinese).
- [17] Feng L S, Zhi Y Z, Lei M, et al. Suppression of frequency locking noise in resonator fiber optic gyro by differential detection method [J]. *Optics & Laser Technology*, 2014, 62: 109-114.
- [18] Zhi Y Z, Feng L S, Lei M, et al. Low-delay, high-bandwidth frequency-locking loop of resonator integrated optic gyro with triangular phase modulation [J]. *Applied Optics*, 2013, 52(33): 8024-8031.
- [19] 洪灵菲, 张春熹, 冯丽爽, 等. 谐振式微光学院螺标度因数及其非线性度分析 [J]. *北京航空航天大学学报*, 2012, 38(8): 1046-1050.  
Hong L F, Zhang C X, Feng L S, et al. Investigation on scale factor and output nonlinearity in resonator micro-optic gyro [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2012, 38(8): 1046-1050 (in Chinese).
- [20] 洪灵菲, 张春熹, 冯丽爽, 等. 谐振式微光学院螺偏置频率调制技术研究 [J]. *光学技术*, 2011, 37(5): 627-631.  
Hong L F, Zhang C X, Feng L S, et al. Bias-frequency-modulation used in resonator-optic gyro [J]. *Optical Technology*, 2011, 37(5): 627-631 (in Chinese).

## Study on current-tuning-induced intensity noise in resonator integrated optic gyro

ZHI Yinzhou, FENG Lishuang\*, WANG Junjie, TANG Yichuang

(School of Instrumentation Science and Opto-to-electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** Resonator integrated optic gyro (RIOG) employing laser diode (LD) is beneficial to miniaturization developing. However, when the frequency of the LD is tuned by current, accompany intensity noise is produced. Influence of current-tuning-induced intensity noise on the open-loop RIOG was theoretically analyzed; linear relationship between frequency locking position and output intensity of the LD was obtained. Subsequently, method of the intensity noise elimination was proposed based on the frequency locking position. Intensity fluctuation of the LD was monitored in real time by the frequency locking position. Compensation coefficients were calculated according to the monitor values, and they were used to compensate the gyro output. Finally, theoretical model that suppressing the current-tuning-induced intensity noise was set up, and the experiments were demonstrated, in which the bias stability of the RIOG was improved from 0.768(°)/s to 0.103(°)/s.

**Key words:** resonator integrated optic gyro (RIOG); laser diode (LD); current-tuning; intensity noise; compensation

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0811

# 协同体制被动毫米波成像系统天线阵布局优化

尚晓舟, 孙鹏, 胡岸勇\*, 苗俊刚

(北京航空航天大学 电子信息工程学院, 北京 100191)



**摘 要:** 针对人体安检应用中的高分辨率实时成像需求,分析了被动毫米波成像系统中相控阵体制与综合孔径体制的关系,结合两者优势提出了一种协同体制被动毫米波成像系统.针对水平方向相控阵体制布局,利用模拟退火算法进行优化,提出采用不同孔径喇叭天线作为阵列单元的方式提升阵列性能,并对主波束效率最大化及副瓣电平最平坦两种准则下的优化结果进行比较;针对垂直方向综合孔径体制布局,设计了满足基线不缺失条件的冗余度最平均稀疏阵列布局.仿真分析结果表明优化后的阵列天线结构可满足被动成像中提升功率测量信噪比(SNR)的需求.

**关键词:** 综合孔径;相控阵;阵列优化;模拟退火算法;被动毫米波成像系统

**中图分类号:** V221<sup>+</sup>.3

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1842-06

近年来,随着安全形势日趋严峻,毫米波成像技术在安检领域的应用得到了世界范围内的重视和研究<sup>[1]</sup>.由于被动成像技术中成像仪不主动辐射电磁波,仅检测目标自身辐射,而毫米波可以穿透衣物,且金属、粉末和液体等不同种类目标在毫米波段具有不同的辐射特性,因此被动毫米波成像技术是一种针对人体安检应用的有效技术<sup>[2-4]</sup>.被动毫米波成像技术最早采用实孔径扫描方式,但其成像速度慢,无法满足人体安检应用要求,且成像分辨率受限于实天线孔径.因此,为了实现人体安检所需的高分辨率实时成像要求,人们提出了基于阵列天线成像方式的焦平面<sup>[5-6]</sup>、相控阵<sup>[7-8]</sup>与综合孔径体制<sup>[9-10]</sup>.

焦平面体制是通过微波透镜或聚焦天线将目标辐射能量进行聚焦,将接收天线阵列布置于透镜或聚焦天线的焦平面处,以检测目标的辐射特性.焦平面体制不需进行扫描,因此能够实现实时成像需求.但由于聚焦仅对目标中心位置而言是

准确的,随着位置偏移目标中心距离加大,散焦将极速加剧,导致成像结果严重失真.因此焦平面体制只适用于小范围目标成像,限制了人体安检的应用场景.

相控阵体制与相控阵雷达原理类似,通过对每路信号进行移相实现波束的空间扫描,获得目标辐射特性.相控阵体制信号处理简单,采用电扫描方式可以不受限于机械扫描装置速度.但进行辐射测量时波束需要在图像的每一个像素点驻留一定时间以获得足够的测量灵敏度,因此图像的分辨率与灵敏度制约了实时成像的应用场景.

综合孔径体制采用干涉测量原理,通过相关运算获得目标可视度,进而反演得到目标辐射特性.由于采用干涉测量方式,因此综合孔径体制中只需要稀疏度很高的接收天线阵列,且后端的相关运算可以并行处理,易于实现人体安检领域的实时成像需求.但是,综合孔径体制的后端信号处理复杂,相关运算需要大量的相关器与大规模的

收稿日期: 2014-12-23; 录用日期: 2015-04-03; 网络出版时间: 2015-08-03 14:58

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150803.1458.011.html

基金项目: 中央高校基本科研业务费专项资金(30344001)

作者简介: 尚晓舟(1986—),男,江苏南京人,博士研究生,whisperwind.shang@gmail.com

\* 通讯作者: 胡岸勇(1980—),男,湖南娄底人,讲师,hu\_anyong@buaa.edu.cn,主要研究方向为被动毫米波成像系统设计与成像算法.

**引用格式:** 尚晓舟, 孙鹏, 胡岸勇, 等. 协同体制被动毫米波成像系统天线阵布局优化[J]. 北京航空航天大学学报, 2015, 41(10): 1842-1847. Shang X Z, Sun P, Hu A Y, et al. Antenna array layout optimization of collaborative passive millimeter-wave imaging system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1842-1847 (in Chinese).

数字处理系统,限制了综合孔径成像系统的规模.

针对不同被动毫米波成像体制的特性及人体安检应用的需求,本文提出了一种相控阵体制与综合孔径体制相结合的被动毫米波成像方式,通过两种体制相结合,在保证系统实时成像要求的前提下降低了系统复杂度,减少了硬件需求,适应于实际人体安检应用.针对相控阵与综合孔径结合的方式,本文重点对其天线阵布局方式进行研究,给出了一种一维相控阵一维综合孔径的矩形阵布阵方式.针对相控阵维,结合仿真分析结果,引入了两种新的评价标准以适用于被动成像体制,并与常规的副瓣电平最小化标准的优化结果进行了对比.针对综合孔径维,给出了适合实际应用的稀疏布阵方式.

### 1 相控阵体制与综合孔径体制

被动毫米波成像方式中,成像系统接收目标自身辐射的电磁波.由于物体自身辐射的电磁波具有随机信号的特征,其功率谱由物体温度决定<sup>[11]</sup>,因此被动成像方式测量目标是物体的功率谱分布,或称辐射亮温.相控阵体制对接收信号进行移相叠加,将天线波束指向测量位置,通过对总接收信号的功率测量获得波束指向处的辐射亮温,继而扫描整个场景获得场景亮温分布.综合孔径体制的基本原理与其不同,综合孔径体制采用干涉测量方式<sup>[12-13]</sup>,将接收信号两两相关,获得目标场景的可视度函数,可视度函数在远场条件下或经近场修正后与场景亮温呈快速傅里叶变换(Fast Fourier Transformation, FFT)关系,因此完成相关运算后进行傅里叶逆变换可直接获得整个场景的亮温分布.由于相控阵体制与综合孔径体制测量方式不同,首先对两者的关系进行分析.

以图 1 与图 2 所示的 4 单元阵列进行说明.为了方便对比,此时综合孔径体制采用满阵中的所有相关结果进行计算,称其为相关阵.

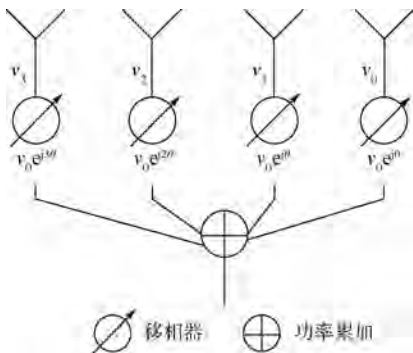


图 1 相控阵体制

Fig. 1 Phased array system

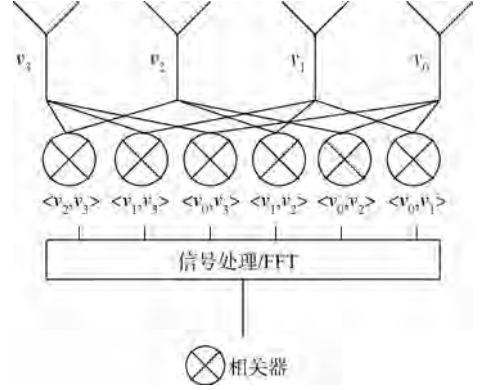


图 2 综合孔径体制

Fig. 2 Synthetic aperture system

设天线单元的输出信号分别为  $v_n (n = 0, 1, 2, 3)$ . 在相控阵列中,为了使波束指向  $\theta$  方向,各接收通道中的移相值分别为  $e^{jnk d_\lambda \sin \theta} (n = 0, 1, 2, 3)$ ,其中  $j$  为虚数单位, $k$  为波矢量, $d_\lambda$  为天线单元间隔.此相控阵输出为

$$P(\theta) = \left( \sum_{n=0}^3 v_n e^{jnk d_\lambda \sin \theta} \right)^2 = \left( \sum_{n=0}^3 v_n e^{jnk d_\lambda \sin \theta} \right) \left( \sum_{n=0}^3 v_n^* e^{-jnk d_\lambda \sin \theta} \right) = \sum_{n=0}^3 \langle v_n, v_n \rangle + \langle v_0, v_1 \rangle e^{-jkd_\lambda \sin \theta} + \langle v_0, v_2 \rangle e^{-j2kd_\lambda \sin \theta} + \langle v_0, v_3 \rangle e^{-j3kd_\lambda \sin \theta} + \langle v_1, v_2 \rangle e^{-jkd_\lambda \sin \theta} + \langle v_1, v_3 \rangle e^{-j2kd_\lambda \sin \theta} + \langle v_2, v_3 \rangle e^{-jkd_\lambda \sin \theta} = \sum_{n=0}^3 \langle v_n, v_n \rangle e^{-j0d_\lambda \sin \theta} + (\langle v_0, v_1 \rangle + \langle v_1, v_2 \rangle + \langle v_2, v_3 \rangle) e^{-jkd_\lambda \sin \theta} + (\langle v_0, v_2 \rangle + \langle v_1, v_3 \rangle) e^{-j2kd_\lambda \sin \theta} + \langle v_0, v_3 \rangle e^{-j3kd_\lambda \sin \theta}$$

式中: $v_n^*$  为  $v_n$  的共轭.

在相关阵中,获得的可视度函数是接收信号的两两相关值  $\langle v_m, v_n \rangle$ . 进行傅里叶变换可视为对所有相关值  $\langle v_m, v_n \rangle$  进行相位加权并叠加.若希望获得  $\theta$  方向的亮温,需要附加的相位项为两天线间基线  $d_{mn}$  与方向正弦  $\sin \theta$  的乘积,即与相控阵中的输出结果一致.

因此,如果不考虑可视度函数的冗余性,将所有可能可视度函数都利用上,则相关阵列的输出将与相控阵列的输出相同.

### 2 天线阵列整体设计

对人体安检应用而言,需要在近距离获得较高的分辨率,即需要大等效孔径阵列天线.同时为了保障足够范围的不混叠视场范围,阵列天线中的单元天线间隔不能过大,这造成了阵列中需要大量的单元天线,特别是对二维成像而言,所需的

单元天线数目非常庞大. 对被动成像而言, 由于需要对随机信号的功率进行测量, 为了保证毫米波图像有足够的温度灵敏度, 需要足够的积分时间. 若采用相控阵体制, 则需要数目众多的移相单元并对其进行控制, 且总体测量时间是单像素积分时间乘以像素数目, 无法满足实时成像要求. 若采用综合孔径体制, 虽然可以满足实时成像要求, 且前端可以使用稀疏阵列结构进行简化, 但为了保证可视度函数不缺失, 后端的复相关器数目仍然十分庞大, 使得后端硬件及信号处理复杂度高, 对系统实现造成很大困难.

为了满足人体安检的实时性需求, 同时保证被动毫米波图像有足够的温度灵敏度, 又不会使得综合孔径复相关运算复杂度过高, 本文采用相控阵与综合孔径结合的成像方式, 并对此设计了布阵方法. 下面分析新型被动毫米波成像系统对天线布阵的需求.

对人体安检应用而言, 希望成像的视场范围包括整个人体, 且水平和垂直方向有相似的分辨率. 新型被动毫米波成像系统设计的系统指标为: 水平方向视场范围为  $20^\circ$ , 竖直方向视场范围为  $40^\circ$ , 在  $2\text{ m}$  处的空间分辨率约为  $5\text{ cm}$ . 由此可估计出, 垂直方向上的单元天线个数约为水平方向上的 2 倍. 为了减少前端结构复杂度, 利用综合孔径体制的稀疏阵列优势, 宜在垂直方向采用综合孔径体制, 水平方向采用相控阵体制.

天线阵列设计简图如图 3 所示. 水平方向为相控阵体制, 垂直方向为稀疏布阵的综合孔径体制, 其中每一行的相控阵经移相后叠加的输出将作为综合孔径的一个通道.

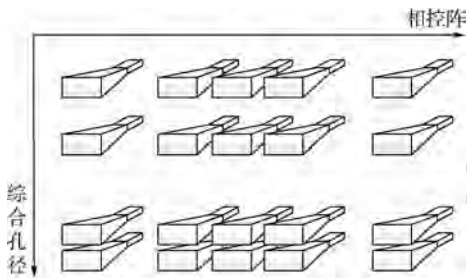


图 3 天线阵列设计简图

Fig. 3 Antenna array design schema

### 3 天线阵列详细仿真结果

#### 3.1 相控阵维度仿真结果

对于水平方向的相控阵, 本文采用模拟退火算法<sup>[14]</sup>对其阵列构成进行优化. 一般而言, 对相控阵的优化准则往往是以副瓣电平最小化为准,

期望副瓣信号电平对主瓣的影响小. 但被动成像体制中, 由于进行的是功率测量, 期望副瓣信号功率对主瓣影响较小, 且人体安检成像中只需要获得功率的相对值而非绝对值. 因此, 本文在优化计算中引入了两种新的准则: 主波束效率最大化 (即主瓣  $-3\text{ dB}$  宽度范围) 和副瓣电平最平坦. 前一种准则主要是用来反映实际中希望接收来自波束指向目标位置的能量尽可能多, 后一种准则则是希望干扰信号变化程度较小, 保证相对功率变化较小. 此外, 为进一步提升主波束效率, 本文引入一种新的方法: 采用口径大小不一致的喇叭天线作为阵列单元, 通过不同单元天线的方向图相干叠加以进一步降低阵列副瓣电平. 表 1 为仿真参数设置.

表 1 仿真参数设置

Table 1 Simulation parameter setting

仿真参数类型	仿真参数数值
波长/mm	8
喇叭天线口径/mm	15, 20, 23
视场范围/ $(^\circ)$	$-10 \sim 10$

对相控阵成像方式而言, 当波束扫描至视场边缘时, 天线阵的方向性最差. 因此, 在仿真过程中, 主要关注两个位置: 扫描角位于  $0^\circ$  与扫描角位于  $10^\circ$  时的仿真效果.

首先, 为了验证采取不同口径天线方案的可行性, 此处仿真统一采取副瓣电平最小化为优化标准, 副瓣电平最小化准则下, 不同扫描角时仿真结果如图 4 和图 5 所示, 并汇总于表 2.

从表 2 中可以看到, 采用口径大小不一的喇叭天线作为阵列单元可以使天线阵的性能提高, 主要体现在以下两方面: 一是主波束效率提高了

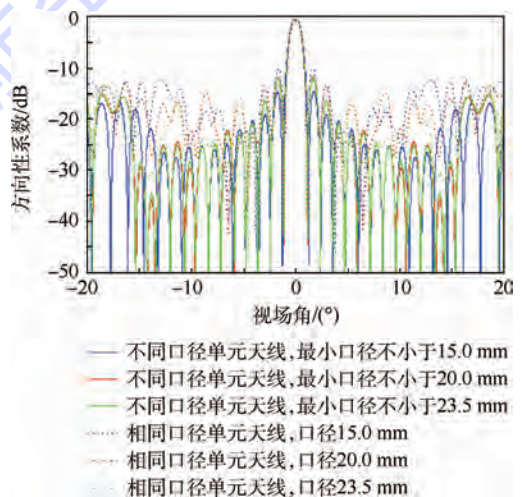


图 4 副瓣电平最小化准则下,  $0^\circ$  扫描角仿真结果

Fig. 4 Simulation result at sweep angle of  $0^\circ$  under minimum sidelobe principle

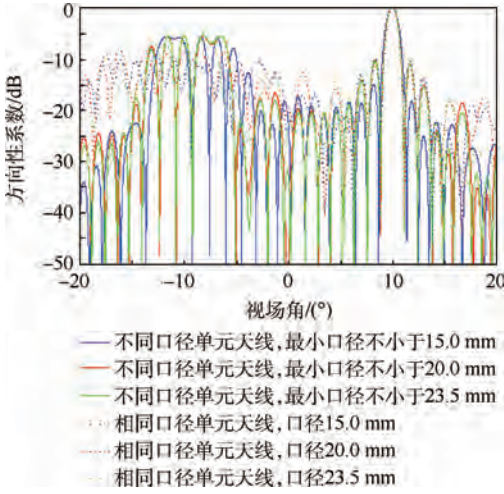


图 5 副瓣电平最小化准则下,10°扫描角仿真结果

Fig. 5 Simulation result at sweep angle of 10° under minimum sidelobe principle

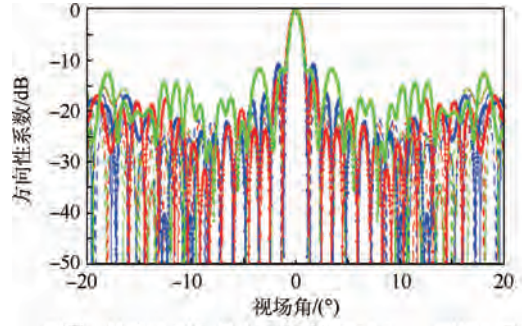


图 6 不同优化准则下,0°扫描角仿真结果

Fig. 6 Simulation result at sweep angle of 0° under different optimization criterion

表 2 副瓣电平最小化准则优化结果

Table 2 Optimization result of minimum sidelobe principle

天线长度标准	初始扫描角 0°			初始扫描角 10°		
	-3 dB 带宽/(°)	归一化最大旁瓣电平/dB	主波束效率/%	-3 dB 带宽/(°)	归一化最大旁瓣电平/dB	主波束效率/%
均为 15.0 mm	0.96	-10.22	56.6	0.98	-9.71	41.29
均为 20.0 mm	1.00	-11.73	57.2	1.02	-8.59	41.15
均为 23.5 mm	1.00	-11.42	57.9	1.02	-6.47	41.46
最小为 15.0 mm	1.10	-14.62	75.3	1.15	-5.77	45.10
最小为 20.0 mm	1.20	-13.37	72.7	1.20	-8.38	42.10
最小为 23.5 mm	1.20	-13.32	72.8	1.20	-6.30	42.20

近 30% ;二是在视场边缘,副瓣电平并没有显著变化.

接下来,本文将验证比较 3 种优化原则的结果.此处仿真将统一采用多口径单元天线方案.

不同优化准则下,不同扫描角时仿真结果如图 6 和图 7 所示,并汇总于表 3.从表 3 中可以看出,主波束效率最大化原则既可以保证主波束能量最大,也可以兼顾视场边缘副瓣电平不至过高,基本符合设计初衷.

综上所述,多口径单元天线布局和主波束效率最大化准则都可以在一定程度上提高接收天线阵的性能,是一种提升被动成像系统性能的有效方式.

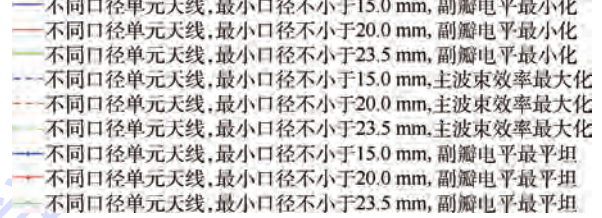


图 7 不同优化准则下,10°扫描角仿真结果

Fig. 7 Simulation result at sweep angle of 10° under different optimization criteria

### 3.2 综合孔径维度仿真结果

根据相控阵与综合孔径体制的理论分析,为保证相控阵体制与综合孔径体制的一致性,需要保证综合孔径维的可视度不缺失<sup>[15]</sup>,即阵列稀疏后仍保证二元干涉仪单元的基线长度种类不缺失,同时为了充分利用冗余基线,希望不同长度基线的冗余度接近,以提升可视度信噪比.此外还需要保证实际工程装配简便性.对此,本文提出了一种方法来解决这些问题.

为保证相控阵维与综合孔径维的分辨率基本相同,通过计算可知,满阵情况下,阵列中有 30 个单元天线位置.考虑到系统实现时的方便性,在 30 个单元天线位置中选取 16 个位置进行稀疏

化,稀疏后的阵列应保证基线不缺失且各长度基线冗余度接近.采用不同长度基线冗余度的标准差来衡量冗余度接近程度.基线不缺失且冗余度标准差最小时稀疏阵单元天线位置如下:11010 11011 00001 01010 00110 11011(其中1代表此处放置天线,0代表不放置天线).

考虑到实际工装中,此种布阵安排无法使用一种弯波导完成与后端接收机的装配.因此将综合孔径维满阵单元天线数目增加到32个,在其中选取16个位置进行稀疏化.基线不缺失且冗余度标准差最小时稀疏阵单元天线位置如下:10101 00110 10100 10100 11001 01100 11,此时不同方案基线对比如图8所示,冗余度标准差为2.0889.

表3 不同优化准则优化结果

Table 3 Optimization results of different optimization principles

优化原则	天线长度标准	初始扫描角 0°			初始扫描角 10°		
		归一化			归一化		
		-3 dB 带宽/ (°)	最大旁瓣电平/ dB	主波束效率/ %	-3 dB 带宽/ (°)	最大旁瓣电平/ dB	主波束效率/ %
旁瓣电平最小化	最小为 15.0 mm	1.1	-14.62	75.30	1.15	-5.770	45.1
	最小为 20.0 mm	1.2	-13.37	72.70	1.20	-8.380	42.1
	最小为 23.5 mm	1.2	-13.32	72.80	1.20	-6.300	42.2
主波束效率最大化	最小为 15.0 mm	1.1	-18.23	78.60	1.45	-8.260	47.3
	最小为 20.0 mm	1.1	-16.36	77.70	1.35	-8.492	47.1
	最小为 23.5 mm	1.1	-14.44	76.00	1.20	-8.426	45.9
副瓣电平最平坦	最小为 15.0 mm	1.2	-13.17	71.20	1.20	-9.470	44.3
	最小为 20.0 mm	1.2	-13.44	73.08	1.20	-8.380	42.3
	最小为 23.5 mm	1.2	-13.23	72.30	1.20	-6.200	41.2

### 4 结论

本文在分析相控阵体制和综合孔径体制成像原理的基础上提出了相控阵成像体制与综合孔径成像体制相结合的协同体制被动毫米波成像系统,并针对其特性进行优化设计.

1) 对水平方向相控阵的优化结果表明:采用口径大小不一的喇叭天线作为阵列单元使主波束效率提高了近30%,且在视场边缘副瓣电平并没有显著变化;采用主波束效率最大为优化原则可保证主波束能量最大且视场边缘处副瓣电平不至过高.

2) 对垂直方向综合孔径维的优化结果表明:满足综合孔径成像基线不缺失的条件;基线冗余度接近,有利于提升可视度信噪比.

### 参考文献 (References)

[1] Bhartia P, Bahl I J. Millimeter wave engineering and applications [M]. New York: Wiley, 1984:660-671.

[2] Appleby R, Anderton R N, Price S, et al. Compact real-time (video rate) passive millimeter-wave imager [J]. Passive Millimeter-Wave Imaging Technology III, 1999, 3703:13-19.

[3] Williams T D, Vaidya N M. A compact, lowcost, passive MMW security scanner [C] // Proceedings of International Society for Optics and Photonics. Bellingham; SPIE, 2005:109-116.

[4] 钱嵩松. 被动毫米波阵列探测成像的关键技术研究 [D]. 南京:南京理工大学, 2006.

Qian S S. Key technology of passive millimeter-wave array detecting and imaging [D]. Nanjing: Nanjing University of Science and Technology, 2006 (in Chinese).

[5] Goldsmith P F, Hsieh C T, Huguenin G R, et al. Focal plane imaging systems for millimeter wavelength [J]. IEEE Transactions on Microwave Theory and Techniques, 1993, 41 (10): 1664-1675.

[6] 章勇, 李兴国, 王华力. 毫米波焦平面阵列成像技术 [J]. 制导与引信, 1998, 12(3): 38-41.

Zhang Y, Li X G, Wang H L. Millimeter-wave focal plane array imaging [J]. Guidance & Fuze, 1998, 12 (3): 38-41 (in Chinese).

[7] Martin C A, Clark S E, Lovberg J A, et al. Passive millimeter-wave imaging technology for phased array systems [C] // Proceedings of the Passive Millimeter-Wave Imaging Technology VI and Radar Sensor Technology VII. Orlando; SPIE, 2003:33-41.

[8] Lovberg J A, Martin C, Kolinko V. Video-rate passive millimeter-wave imaging using phased arrays [C] // Proceedings of the Microwave Symposium. Piscataway, NJ: IEEE Press, 2007: 1689-1692.

[9] Le Vine D M L, Jackson T J, Swift C T, et al. Passive microwave remote sensing with the synthetic aperture radiometer, ESTAR, during the southern great plains experiment [C] // Proceedings of

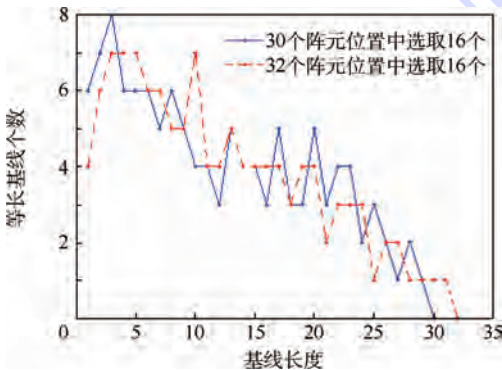


图8 不同方案基线对比

Fig. 8 Baseline comparison for different projects

- International Geoscience and Remote Sensing Symposium. Piscataway, NJ: IEEE Press, 1998, 5: 2606-2608.
- [10] Zheng C, Yao X, Hu A, et al. Initial results of a passive millimeter-wave imager used for concealed weapon detection BHU-2D-U[J]. Progress in Electromagnetics Research C, 2013, 43: 151-163.
- [11] Ulaby F T, Moore R K, Fung A K. Microwave remote sensing, active and passive, Vol. I, microwave remote sensing fundamentals and radiometry[M]. Massachusetts: Addison-Wesley Publishing Company, 1981: 22-25.
- [12] Swift C T, Levine D M, Ruf C S. Aperture synthesis concepts in microwave remote sensing of the earth[J]. IEEE Transactions on Microwave Theory and Techniques, 1991, 39(12): 1931-1935.
- [13] Vine D M L. Synthetic aperture radiometer systems[J]. IEEE Transactions on Microwave Theory and Techniques, 1999, 47(12): 407-410.
- [14] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing[J]. Science, 1983, 220(4598): 606-615.
- [15] Thompson A R, Moran J M, Swenson G W. Interferometry and synthesis in radio astronomy[M]. Malabar: Krieger Publishing Company, 1994: 126-128.

## Antenna array layout optimization of collaborative passive millimeter-wave imaging system

SHANG Xiaozhou, SUN Peng, HU Anyong\*, MIAO Jungang

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** Based on the advantages of phased array and synthetic aperture imaging method in passive millimeter-wave imaging, a collaborative passive millimeter-wave imaging system was developed for real-time and high resolution imaging in human security application. In the horizontal orientation, by applying annealing algorithm, different aperture horn antennas were determined for the improvement of the phased array performance, which was generally compared with the maximum main beam effective and most smoothable sidelobe level. In vertical orientation, the synthetic aperture, namely, the sparse antenna array was designed by the criterion of the most uniform redundant baseline, while baseline absence should be avoided. Finally, simulation results reveal that with the optimized antenna array the requirement of increasing signal-to-noise ratio (SNR) in power measurement in the collaborative passive imaging system can be fulfilled.

**Key words:** synthetic aperture; phased array; array optimization; simulated anneal algorithm; passive millimeter-wave imaging system

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0210

# 电磁多尺度和非线性效应分析

毛煜茹, 谢拥军\*

(北京航空航天大学 电子信息工程学院, 北京 100191)



**摘 要:** 微观尺度上的结构和电接触特性对于宏观尺度上的电磁特性具有重要的影响. 针对多尺度结构的非线性无源互调(PIM)无法用频域方法直接计算的问题, 提出了基于Floquet理论的全波方法, 对具有周期性微观结构的多尺度电大尺寸问题进行计算, 从而准确地计算其电特性. 此外, 将产生非线性的金属结的等效电路模型引入到全波分析方法中, 通过全波方法与等效电路方法相结合的方式来分析多尺度结构的无源互调问题. 数值计算结果表明该方法可以有效地评估微观结构对于宏观结构电性能的影响, 并解决了时域方法求解电大尺寸微波部件产生的计算累积误差较大这一难题.

**关键词:** 多尺度结构; 电大尺寸; 周期性微观结构; 非线性效应; Floquet理论

**中图分类号:** TN823

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1848-05

多尺度电大尺寸工程问题是指电磁工程问题的宏观规模达到几十乃至几百个电波长, 但其中存在几十分之一波长尺度上的微观结构, 这些微观结构虽然尺度很小, 却由于数量规模巨大、材料或者结构特殊, 对于电磁问题的宏观特性的影响不可忽略. 多尺度问题因其重要性在包括电磁学的很多领域都进行了大量的研究. 文献[1]综述了材料科学中多尺度问题研究的发展趋势和面临的挑战. 文献[2]给出了一种不依赖于方程的计算机辅助分析的方法. 在电磁场领域, 对于含有多尺度的电大散射和辐射问题多采用空域分解技术<sup>[3-4]</sup>或者混合有限元-边界积(FEBI)-多层快速多极子(MLFMA)-一致性几何绕射(UTD)高低频混合技术<sup>[5]</sup>, 文献[6]使用了时域积分方程方法(TDIE)-改进节点分析法(MNA)-基于TDIE的传输线求解器混合的方法来分析电大尺寸平台上的电缆耦合问题. 文献[7]给出了结合高阶场点采样技术的等效方法来解决多尺度电磁分析. 为解

决多尺度问题产生的稠密不均匀网格导致的病态阻抗矩阵, 国内胡俊教授课题组跟踪研究了区域分解技术<sup>[8-9]</sup>, 通过将问题尺度划分为若干小区域, 独立求解每个子区域, 再利用区域边界之间的连续性求解出整个问题. 此外, 利用预条件技术加速矩量法(MoM)稠密矩阵方程的迭代收敛也较常见, 盛新庆教授利用插值分解构造预条件的方法加速多层快速多极子算法的收敛<sup>[10]</sup>. 陈如山教授课题组在多尺度问题的MoM矩阵方程迭代收敛性及其预条件和大规模MoM矩阵方程快速求解算法方面也进行了深入研究<sup>[11-14]</sup>, 提出了3种基于快速傅里叶变换(FFT)和低秩压缩的快速积分方法, 降低了矩阵方程存储和迭代求解矩阵矢量乘的计算复杂度, 同时还研究了两种用于解决多尺度问题在低频和高频时矩阵性态差、收敛速度慢的预条件技术.

以上多尺度问题中的“微观”部分, 其尺度在波长的分数量级, 而很多的实际工程问题中, 其微

收稿日期: 2015-04-10; 录用日期: 2015-06-05; 网络出版时间: 2015-06-25 16:36

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150625.1636.004.html

作者简介: 毛煜茹(1984—), 女, 山西晋城人, 博士研究生, maoyuru218@126.com

\* 通讯作者: 谢拥军(1968—), 男, 陕西西安人, 教授, yjxie@buaa.edu.cn, 主要研究方向为太赫兹、天线工程.

**引用格式:** 毛煜茹, 谢拥军. 电磁多尺度和非线性效应分析[J]. 北京航空航天大学学报, 2015, 41(10): 1848-1852.

Mao Y R, Xie Y J. Analysis of electromagnetic multi-scale structure and non-linear effects [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1848-1852 (in Chinese).

观尺度在波长的几十分之一,无法进行网格剖分及求解。

对于典型的具有松散连接结构如网状天线,其网格尺度均在几十分之一波长以下,而电性能则需要几十个乃至上百个波长尺度的范围内进行求解,由于多尺度电磁结构中的微观结构的剖分密度远大于其他部分,致使矩阵的条件数恶化,在求解中收敛很慢,甚至不收敛。丝网结构会造成电磁波因为绕射效应的泄露,而丝网节点处的金属结的非理想接触会引起网状天线的无源互调。该结构属于典型的多尺度、非线性无源互调(PIM)的电大尺寸电磁问题。

金属接触非线性引起的无源互调的散射机理及其数值计算最早由意大利 Pelosi 课题组于 1999 年提出<sup>[15-17]</sup>,该课题组将时域物理光学方法(TD-PO)的应用推广于非理想导体和非线性环境,给出由于两种不同的金属相互接触时所产生的非线性无源互调的计算模型。在 TD-PO 方法中,假设入射波为平面波,通过引入模拟两种不同金属相互搭接而产生的金属结的非线性因子,求解相互搭接的金属在电磁波照射下所产生的空间散射场,并对该散射场进行傅里叶变换求得各阶无源互调产物,实现对不同金属相互接触而产生的无源互调的预测。

但是,TD-PO 方法中的入射波被近似为平面波,仅适用于辐射源位于散射目标的散射远区时的情形,只能对一些典型简单的微波结构进行理论验证,无法计算实际复杂微波部件工作时产生的无源互调,并且在电大环境下采用 TD-PO 时域算法所带来的误差积累是不可忽视的。

含有多尺度、非线性的电大尺寸电磁问题的分析具有重要的实际工程背景,对于计算电磁学的发展带来了新的挑战。本文提出的多尺度等效方法、场路结合数值分析方法是解决多尺度、非线性电大尺寸电磁工程问题分析的关键技术。

## 1 多尺度问题的等效分析方法

多尺度电磁工程问题无法利用现有的数值计算方法如有限元、矩量法及其快速算法等进行几何剖分运算,会导致矩阵求解无法收敛,甚至连几何剖分都无法完成。但是,微观结构由于数量规模和材料结构的原因,其宏观影响却必须通过数值计算来分析评估。等效方法是解决这一矛盾的有效手段,等效方法是指将微观结构对于宏观结构中电磁行为的影响以边界条件的形式等效,并加

入宏观结构中进行求解。这种等效方法将问题分为两个阶段:第 1 个阶段是研究微观结构的电磁响应,并将其等效为边界条件;第 2 个阶段计算不含有具体微观结构、但含有其等效边界条件的宏观电磁问题。多尺度问题的等效方法化难为易,既解决了直接求解方法中的数值离散化带来的计算难题,也充分考虑了微观结构的电磁效应。具有松散接触的周期性微观结构如丝网反射面、频率选择表面、天线罩高频段特性分析等的多尺度电大尺寸问题均可采用等效分析方法进行数值分析。文献[18]推导了多层极薄层边界条件,解决了天线罩高频段特性分析多尺度电大尺寸的计算,这一技术已集成入 ANSYS HFSS 软件,并为国内多个重要军工单位解决了新型产品的论证和设计。

松散连接是指金属相互接触面之间存在间隙的连接情况,并形成 MIM (Metal Insulator Metal) 结构,如图 1 所示。图 1 中有 5 种接触状态:①金属接触;②接触面之间夹有金属膜氧化物;③接触面之间夹有绝缘介质;④微小空气间隙;⑤较大空气间隙。①和②形成电流的主要通道,形成收缩电阻和接触电阻;②中的氧化物依靠隧道效应和穿透薄膜的金属桥进行导电,属于半导体接触导电,是非线性的;③则不导电,电流绕到金属接触处通过;在空气间隙④和⑤处电流环绕间隙流过。当电流遇到阻抗  $Z$ ,产生间隙电压,间隙电压是潜在的,可能激活任何一个半导体而产生无源互调效应<sup>[19]</sup>。接触面的等效电路模型如图 2 所示,其  $V-I$  特性是非线性的。

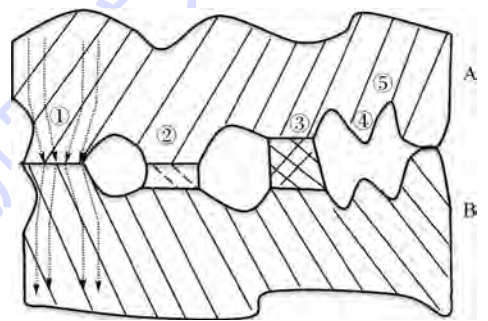
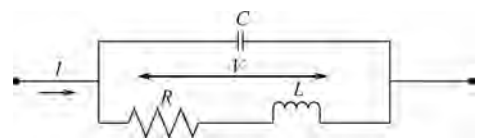


图 1 松散连接接触面

Fig. 1 Contact surface of metallic connection



$I$ —电流; $V$ —电压; $C$ —电容; $R$ —电阻; $L$ —电感。

图 2 松散连接接触面的等效电路

Fig. 2 Equivalent circuit of metallic connection contact surface

美国 Rahmat-Samii 教授等应用 Floquet 方法研究了不考虑电接触模型(即理想搭接)情况下周期性反射面天线(mesh antenna)的特性,主要研究周期结构单元的尺寸和形状等因素对于反射特性的影响<sup>[20]</sup>. 对于搭接的影响,以理想连接和完全断开两种极端情况作为极限来评估. 而本文的多尺度等效分析方法首先要对松散连接的周期性单元结构进行电磁建模,并在金属结处引入金属搭接的电接触模型,即图 2 所示的等效 RLC 边界条件,采用 Floquet 方法求解周期平面结构的电性能参数,多尺度等效方法的计算流程如图 3 所示.

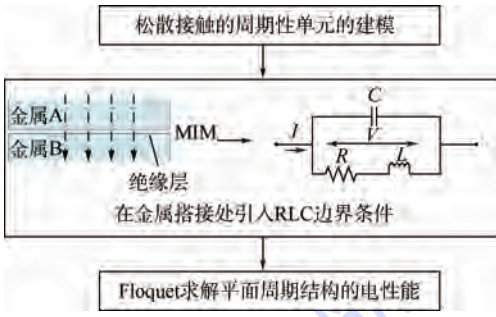


图 3 多尺度结构等效方法流程图

Fig. 3 Flow diagram of equivalent method for multi-scale structure

## 2 非线性问题的场路结合分析方法

实际微波结构中金属部件上的各种不均匀性(镀层、法兰连接面、搭接、焊点、螺钉等)在电磁场的激励下,其金属结中的自由电子运动会产生非线性响应. 如果有两个或者以上频率的电磁波同时激励,则会产生高阶交互调频谱上电磁波,这种效应称为无源互调效应. 在收发系统中,3 阶互调波等会恰好落在工作频带,会对微波接收机产生干扰,对于宽带的收发系统而言,7 阶互调波甚至 9 阶互调波都会对接收机的正常工作产生影响.

对于实际工程中电大尺寸环境中的非线性问题,由于非线性效应非常弱,一般均在 -170 dB 以下,很容易和远大于波长级电尺寸问题中的数值误差积累混淆. 所以,时域方法很难应用于实际电大尺寸问题中.

本文提出了场路结合的数值分析方法,在计算电大尺寸范围内的激励电磁场分布时采用频域全波方法,获得精确的非线性结处的激励场分布,然后结合非线性结的伏安特性求取其各阶互调分量,再采用频域全波方法评估其在电大尺寸范围内的影响,如图 4 所示.



图 4 场路结合方法分析无源互调的流程图

Fig. 4 Flow diagram of coupled field-circuit method for PIM analysis

## 3 数值实例

### 3.1 多尺度问题的等效分析方法的应用

建立图 5 所示的印刷平面网状周期结构,其表面由铜条带构成,介质层的相对介电常数为 4.7,厚度为 1.78 mm. 采用频域全波计算方法结合 Floquet 理论对该结构的传输特性进行数值计算,计算频率为 2 ~ 15 GHz<sup>[20]</sup>.



图 5 周期结构的平面网状反射面

Fig. 5 Planar mesh reflector with periodic structure

使用图 3 所示的等效计算方法对上述所示的平面网状周期结构进行计算,计算结果如图 6 所示. 与文献[20]中的实验数据进行比较,在不考虑电接触特性时本文方法计算与其预测结果和测量数据符合,验证了本文等效分析方法的准确性.

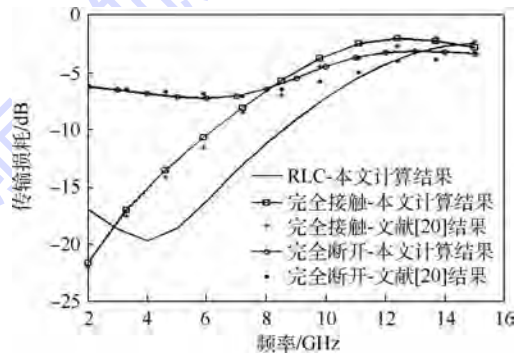


图 6 不同接触状态下的传输损耗

Fig. 6 Transmission loss for different connections

从图 6 可以看出,平面丝网结构在金属结搭接处的开路、短路和存在电阻、电容 3 种状态下引起的传输损耗是有很大的区别的.

### 3.2 场路结合数值分析方法的应用

考虑双载波频率情况,互调的频率  $f_{IM}$  可以表示为:  $f_{IM} = mf_1 + nf_2$ , 其中  $f_1, f_2$  为输入载波频率,

$m, n$  为非零整数. 建立如图 3 所示的平板验证模型, 使用喇叭天线作为馈源, 由金属平板形成的单个金属结距离喇叭天线 0.3 m, 金属结部分的等效电路模型采用式 (1). 假设频率分别为  $f_1 = 11 \text{ GHz}$ ,  $f_2 = 12.6 \text{ GHz}$  的两个电磁波垂直照射到平板上, 计算分析该金属结的 3 阶、5 阶互调干扰.

天线和微波接头处的金属结的非线性效应可用式 (1) 所示的电路模型来等效<sup>[16]</sup>:

$$h_x = h_0 \arctan(e_y/e_T) \quad (1)$$

式中:  $h_0$  和  $e_T$  为待定的系数;  $h_x$  为沿非线性金属结的磁场强度;  $e_y$  为垂直于非线性金属结的电场强度; 而  $h_0, e_T$  的值决定了非线性金属结的材料与非线性程度.

使用图 4 所示的场路结合数值分析方法对图 7 所示的验证模型进行计算, 首先对图 3 的单个金属结以及喇叭馈源进行电磁建模, 采用频域计算方法求得金属结的表面电场, 并在非线性区域对表面电场进行积分求得非线性区域的电压降, 求得 3 阶和 5 阶互调产物的功率电平. 再次利用频域计算方法求解 3 阶和 5 阶互调产物作为激励所激发的空间二次辐射场, 在接收端口处对电场进行积分得到的无源互调功率结果如表 1 所示, 与测量数据<sup>[16]</sup>比较可以看出, 虽然存在一定的误差, 但仍可认为预测结果与测量数据基本符合, 初步验证了本文场路结合数值分析方法的准确性.

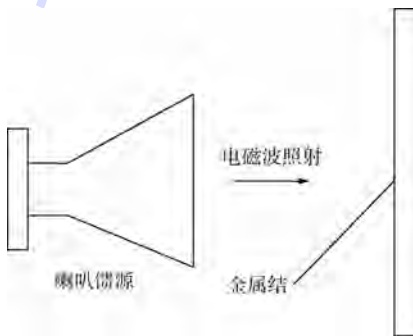


图 7 金属结无源互调计算模型

Fig. 7 PIM computed model of metallic junction

表 1 3 阶和 5 阶无源互调测试结果和计算结果  
Table 1 Measured and computed PIM levels of third order and fifth order

互调阶数	每个频点处 输入功率/W	测试结果/dBm	本文计算 结果/dBm
3 阶	100	-104	-102
	170	-100	-97
5 阶	100	-125	-123
	170	-116	-114

## 4 结 论

本文介绍了含有微观结构的多尺度电大尺寸的松散连接的周期结构的等效计算方法, 以及计算电大尺寸微波部件无源互调的场路结合数值分析方法, 可以得出以下结论:

1) 由丝网反射面电性能的计算结果可知, 网面上金属结的电接触特性对整个丝网的电性能会产生较大的影响, 必须建立正确合理的丝网单元的电磁模型, 才能得到准确的丝网反射面的电性能. 而多尺度等效分析方法可以有效地对含有微观结构的电大尺寸微波部件的电性能以及该结构对宏观结构电性能的影响进行评估.

2) 提出的场路结合数值分析方法解决了频域计算方法无法求解非线性物质结构的电磁响应问题.

3) 场路结合数值分析方法可以分析具有松散连接结构的电大尺寸微波部件的无源互调, 相比时域计算方法, 产生的误差较小.

## 参考文献 (References)

- [1] Geers M G D, Kouznetsova V G, Brekelmans W A M. Multi-scale computational homogenization: Trends and challenges [J]. Journal of Computational and Applied mathematics, 2010, 234 (7): 2175-2182.
- [2] Kevrekidis I G, Gear C W, Hummer G. Equation-free: The computer-aided analysis of complex multiscale systems [J]. AIChE Journal, 2004, 50 (7): 1346-1355.
- [3] Umashankar K R, Nimmagadda S, Taflove A. Numerical analysis of electromagnetic scattering by electrically large objects using spatial decomposition technique [J]. IEEE Transactions on Antennas and Propagation, 1992, 40 (8): 867-877.
- [4] Notaros B M, Popovic B D, Weem J P, et al. Efficient large-domain MoM solutions to electrically large practical EM problems [J]. IEEE Transactions on Microwave Theory and Techniques, 2001, 49 (1): 151-159.
- [5] Tzoulis A, Eibert T F. A hybrid FE/BI-MLFMM-UTD method for numerical solutions of electromagnetic problems including arbitrarily shaped and electrically large objects [J]. IEEE Transactions on Antennas and Propagation, 2005, 53 (10): 3358-3366.
- [6] Bagci H, Yilmaz A E, Jin J M, et al. Fast and rigorous analysis of EMC/EMI phenomena on electrically large and complex cable-loaded structures [J]. IEEE Transactions on Electromagnetic Compatibility, 2007, 49 (2): 361-381.
- [7] Li M K, Chew W C. Multiscale simulation of complex structures using equivalence principle algorithm with high-order field point sampling scheme [J]. IEEE Transactions on Antennas and Propagation, 2008, 56 (8): 2389-2397.
- [8] Jiang M, Hu J, Zhao R, et al. Hybrid IE-DDM-MLFMA with forward-backward method for conducting body of translation [C] // Proceedings of the 2012 International Conference on Electromag-

- netics in Advanced Applications. Piscataway: IEEE Press, 2012: 263-265.
- [9] Shao H, Hu J, Lu W, et al. Analyzing large-scale arrays using tangential equivalence principle algorithm with characteristic basis functions [J]. Proceedings of the IEEE, 2013, 101 (2): 414-422.
- [10] Pan X M, Sheng X Q. Preconditioning technique in the interpolative decomposition multilevel fast multipole algorithm [J]. IEEE Transactions on Antennas and Propagation, 2013, 61 (6): 3373-3377.
- [11] Li M M, Chen R S, Wang H X, et al. A multilevel FFT method for the 3-D capacitance extraction [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2013, 32 (2): 318-322.
- [12] Li M M, Li C Y, Ong C J, et al. A novel multilevel matrix compression method for analysis of electromagnetic scattering from PEC targets [J]. IEEE Transactions on Antennas and Propagation, 2012, 60 (3): 1390-1399.
- [13] Li M M, Francavilla M A, Vipiana F, et al. A doubly hierarchical MoM for high-fidelity modeling of multiscale structures [J]. IEEE Transactions on Electromagnetic Capability, 2014, 56 (5): 1103-1111.
- [14] Li M M, Francavilla M A, Vipiana F, et al. Nested equivalent source approximation for the modeling of multiscale structures [J]. IEEE Transactions on Antennas and Propagation, 2014, 62 (7): 3664-3678.
- [15] Bolli P, Naldini A, Pelosi G, et al. Passive intermodulation scattering in a time domain physical optics framework [C] // Proceedings of International Symposium on Nonlinear Electromagnetic Systems (ISEM). Pavia: ISEM, 1999: 121-124.
- [16] Selli S, Bolli P, Pelosi G. A time domain physical optics heuristic approach to passive intermodulation scattering [J]. IEEE Transactions on Electromagnetic Compatibility, 2001, 43 (2): 203-209.
- [17] Bolli P, Pelacchi P, Pelosi G, et al. Passive intermodulation on large reflector antennas [J]. IEEE Transactions on Antennas and Propagation Magazine, 2002, 44 (5): 13-20.
- [18] 刘莹, 谢拥军. 多层介质天线罩的数值分析 [J]. 电波科学学报, 2011, 26 (3): 550-555.
- Liu Y, Xie Y J. Numerical analysis of multi-layer dielectric radome [J]. Chinese Journal of Radio Science, 2011, 26 (3): 550-555 (in Chinese).
- [19] 罗一锋. 舰船通信系统的无源互调研究 [J]. 现代电子技术, 2010 (23): 39-44.
- Luo Y F. Approach to passive intermodulation in ship communication system [J]. Modern Electronics Technique, 2010 (23): 39-44 (in Chinese).
- [20] Imbriale W A, Galindo-Israel V, Rahmat-Samii Y. On the reflectivity of complex mesh surfaces [J]. IEEE Transactions on Antennas and Propagation, 1991, 39 (9): 1352-1365.

## Analysis of electromagnetic multi-scale structure and non-linear effects

MAO Yuru, XIE Yongjun\*

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** The microstructures and the electrical contact property may affect the performance of the macro scale objects, and the passive intermodulation (PIM) problems cannot be directly calculated by frequency-domain methods. Firstly, the Floquet's theorem combined with the full-wave method was proposed to analyze the electrical performance of the multi-scale structure with periodic microstructures. Secondly, an equivalent circuit model of non-linear metallic junctions coupled with full-wave frequency domain method was adopted to analyze the PIM problems of multi-scale structure. The numerical results show that the method proposed can evaluate the influence of microstructure on the performance of macrostructure, and eliminate the accumulation of computational error appeared in the time domain method which is unacceptable in electrically large objects.

**Key words:** multi-scale structure; electrically large objects; periodic microstructures; non-linear effects; Floquet's theorem

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0211

# 有源可调微波吸收体分析与优化

戚开南<sup>1,2</sup>, 汪勇峰<sup>1,2</sup>, 侯新宇<sup>2</sup>, 陈军文<sup>\*2</sup>

(1. 中国传媒大学 信息工程学院, 北京 100024; 2. 电磁散射重点实验室, 北京 100854)



**摘 要:** 为解决无源雷达吸波材料带宽有限的问题, 基于传输线理论设计了一种有源可调微波吸收体. 该吸收体是在 Salisbury 屏的拓扑结构基础上, 用基于 PIN 二极管可控的有源频率选择表面(FSS)代替传统的电阻层. 反射率测量结果表明, 通过改变二极管的偏置电流可以动态调节吸收体的最佳吸波频率; 当偏置电流在 0~0.5 mA 之间变化时, 吸收体的最佳吸波频率在 7~12.5 GHz 频率范围内动态调整. 基于遗传算法, 优化设计了一种具有双极化吸波特性的十字型可调微波吸收体, 仿真结果表明该吸收体在 5.6~17.6 GHz 范围内实现了最佳吸波频率的可控迁移.

**关键词:** 传输线理论; 微波吸收体; 频率选择表面(FSS); 偏置电流; 优化

**中图分类号:** V221<sup>+</sup>.3; TB553

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1853-06

隐身技术是当今世界各国普遍关注的重要军事技术之一. 其实质是降低目标的雷达散射截面(RCS)<sup>[1]</sup>, 加载雷达吸波材料(RAM)是实现 RCS 减缩的一种主要技术途径.

20 世纪 30 年代就出现了雷达吸波材料, 并应用于天线领域. 随着雷达技术的发展, 20 世纪 50 年代以后, 出现了众所周知的 Salisbury 屏<sup>[2]</sup>、Dallenbach 和 Jaumann 吸收体<sup>[3]</sup>, 它们的吸波原理主要有两种: 相位抵消或者直接吸收入射电磁波的能量. 后来, 为了提高 RAM 的性能, 开发了一些新型材料, 诸如铁氧体、频率选择表面(FSS)<sup>[4]</sup>、纳米材料吸收剂等. 但是, 常规的 RAM 一旦制造出来其特性便是固定的, 吸收带宽是有限的<sup>[5]</sup>, 如果战场环境和所面临的威胁发生改变, 它的隐身效能便会大打折扣.

智能隐身概念的提出, 为 RAM 的发展提供了更为广阔的空间. 智能隐身系统是一种能够从自身的表层或内部获取关于环境条件及其变化的信息, 随后进行判断、处理、反应, 以改变自身的结

构与功能, 使之很好地与外界相协调地具有自适应性的材料系统. 智能隐身系统主要由电子侦察设备、控制器、控制电压或调制波形产生器和有源可调微波吸收体组成. 智能隐身系统的工作流程是: 首先由电子侦察设备捕捉敌方雷达的威胁频率, 将它反馈给控制器; 然后控制器根据威胁频率作用于控制电压或调制波形产生器, 使其产生相应的控制电压或调制波形; 最后所产生的控制电压或调制波形作用于可调微波吸收体, 使其在威胁频率处获得最佳的吸波性能.

有源可调微波吸收体是智能隐身系统的关键部件, 它可以根据武器装备所面临的威胁频率, 采用电控方式动态调整吸收体阻抗的实部或虚部, 使在威胁频率处获得最佳吸波性能. 因此, 研究有源可调微波吸收体具有重大意义.

20 世纪 90 年代, 英国的 Parker 等提出在频率选择表面中加入 PIN 二极管构成有源 FSS 结构, 通过调节二极管的偏置电压或偏置电流, 可以改变 FSS 的谐振特性<sup>[6]</sup>. 2004 年英国谢菲尔德大

收稿日期: 2015-04-10; 录用日期: 2015-05-08; 网络出版时间: 2015-06-25 15:55

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150625.1555.002.html

作者简介: 戚开南(1986—), 男, 山东德州人, 博士研究生, qikanan@163.com

\* 通讯作者: 陈军文(1959—), 男, 河南新乡人, 研究员, lixiaoshen@chinasarft.gov.cn, 主要研究方向为雷达目标特性.

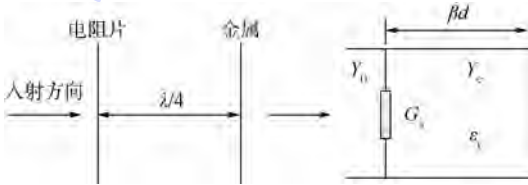
**引用格式:** 戚开南, 汪勇峰, 侯新宇, 等. 有源可调微波吸收体分析与优化[J]. 北京航空航天大学学报, 2015, 41(10): 1853-1858.  
Qi K N, Wang Y F, Hou X Y, et al. Analysis and optimization of active tunable microwave absorber[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1853-1858 (in Chinese).

学教授 Tennant 和 Chambers 设计了领结型 FSS 有源雷达吸收体<sup>[7]</sup>. 该结构选择 PIN 管作为可控元件, 经测试该有源 FSS 吸波屏对垂直极化的平面波在 9 ~ 13 GHz 内实现了吸波特性的可调; 该成果主要实现了反射率深度的电流控制, 而谐振点或工作频段的可控迁移效果并不明显. 东南大学的徐欧和朱敏设计了双方环 FSS 单元可调微波吸收体, 并开展了进一步的实验研究<sup>[8]</sup>.

本文设计并研制了一种方环可调微波吸收体的实验样板, 并进行了反射率测量. 在此基础上, 从工程应用角度出发, 研究了十字型单元可调微波吸收体的优化设计技术, 为宽频段、高吸收、双极化雷达吸波材料的研究和开发提供了一种新途径.

### 1 可调微波吸收体

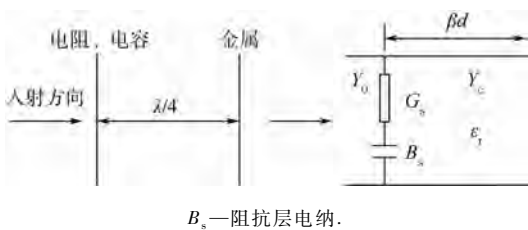
无源结构型吸波材料包括 Salisbury 屏或 Jaumann 型及渐变介质型两大类. 有源吸波结构是在上述两种类型基础上发展起来的. 综合考虑隐身效果、设计和加工难度及响应速度等多种因素, 本文采用以 Salisbury 屏为基础的可调微波吸收体. 可调微波吸收体是在 Salisbury 屏(见图 1)的基础上发展起来的. 在 Salisbury 屏的阻抗层中引入电抗成分可形成电路模拟吸收体<sup>[9-10]</sup>(见图 2); 对阻抗层的电阻或电抗进行电控, 就形成了可调微波吸收体. 可调微波吸收体由金属背衬、介质间隔层、可变阻抗层组成, 如图 3 所示(电阻与电导、电抗与电纳、阻抗与导纳互为倒数关系).



$\lambda$ —波长;  $Y_0$ —自由空间导纳;  $Y_c$ —间隔层本征导纳;  
 $d$ —间隔层材料厚度;  $\beta$ —相移常数;  
 $\epsilon_r$ —间隔层介电常数;  $G_s$ —阻抗层电导.

图 1 Salisbury 屏及传输线等效电路

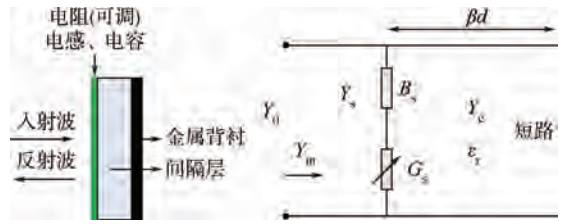
Fig. 1 A Salisbury screen and transmission line equivalent circuit



$B_s$ —阻抗层电纳.

图 2 电路模拟吸收体及传输线等效电路

Fig. 2 A circuit analog absorber and transmission line equivalent circuit



$Y_{in}$ —输入导纳;  $Y_s$ —可变阻抗层导纳.

图 3 可调微波吸收体及其传输线等效电路

Fig. 3 A tunable microwave absorber and transmission line equivalent circuit

根据传输线理论, 法向入射时的反射系数为

$$\rho = \frac{Y_0 - Y_{in}}{Y_0 + Y_{in}} = \frac{Y_0 - Y_s + jY_c \cot(\beta d)}{Y_0 + Y_s - jY_c \cot(\beta d)} \quad (1)$$

1) Salisbury 屏, 频带窄, 不可调.

当  $Y_s = Y_0, d = \lambda/4, \cot(\beta d) = 0, \rho \rightarrow 0$ , 其中反射系数  $\rho$  用传输线理论对不同隔离层介质的反射率进行计算, 结果见图 4(取电阻  $R = 377 \Omega$ , 隔离层厚度为 7.5 mm).

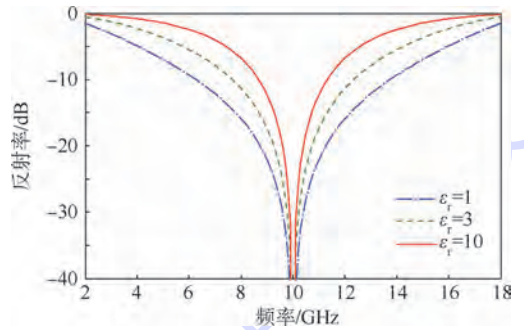


图 4 加载不同材料后 Salisbury 屏的反射特性

Fig. 4 Reflectivity characteristics of Salisbury screens designed with different materials

从图 4 看出, Salisbury 屏频带窄, 且随着间隔层介电常数增大, 吸波性能变差.

2) 电路模拟吸收体, 频带宽, 不可调.

用等效电路法<sup>[11-13]</sup>和 HFSS 对单层电路模拟吸收体(间隔层、蒙皮材料介电常数为 4, 厚度分别为 1.87 mm 和 3.16 mm, 阻抗层采用方形贴片: FSS 单元周期  $p = 7.5 \text{ mm}$ , FSS 单元间距  $g = 1.3 \text{ mm}$ , 电阻  $R = 60 \Omega$ )的反射率进行计算, 仿真曲线如图 5 所示.

3) 可调微波吸收体, 高吸收, 频带宽, 且可调.

图 3 中, 当  $B_s$  为容性时

$$Y_{in} = \frac{w^2 C^2 R}{1 + w^2 C^2 R^2} + j \left( \frac{wC}{1 + w^2 C^2 R^2} - Y_c \cot(\beta d) \right) \quad (2)$$

式中:  $w$  为 FSS 铜箔宽度;  $C$  为电容.

有源吸波结构的谐振频率由式(2)中虚部确

定. 调节  $R(G_s)$  或  $C$  可以调谐频率, 同时改变反射率的“0”深.

图 6 所示为固定  $B_s$ , 改变  $G_s(R)$  时的可调微波吸收体(厚度为 5.3 mm) 反射率仿真结果, 当  $R$  变化时, 最低反射率“0”深和反射率位置相应改变.

根据表 1, 比较 Salisbury 屏与电路模拟吸收体看出, 引入电纳可以扩展频带; 但两者的不足是一旦设计加工完成, 其吸波性能也固定. 可调微波

吸收体的阻抗可调, 因而具有更大的吸波带宽.

表 1 3 种吸收体吸波性能对照

Table 1 Reflectivity comparison of three absorbers

吸波结构类型	电阻	电纳	厚度/波长	-20 dB 相对带宽/%	带宽/厚度
Salisbury 屏	固定	无	0.25	25.3	1.01
电路模拟吸收体	固定	固定	0.187	34.7	1.85
可调微波吸收体	可调	固定	0.185	66.7	3.61

从工程实际角度出发, 实现可变阻抗层的阻抗可调有 3 种途径:

- 1) 固定  $X$ , 改变  $R$ . 例如 FSS 串联 PIN 二极管, 通过改变 PIN 偏置电流改变  $R$ .
- 2) 固定  $R$ , 改变  $X$ . 例如有耗 FSS 串联变容管, 通过改变变容管偏置电压, 改变电容.
- 3)  $R, X$  均可调.

本文采用由 FSS 结合 PIN 二极管组成的阻抗表面方案. 通过电流控制 PIN 二极管的电阻, 从而改变整个可变阻抗层的阻抗. PIN 二极管响应速度快, 响应时间在  $\mu\text{s}$  级, 功耗约为  $25 \mu\text{W}/\text{cm}^2$ .

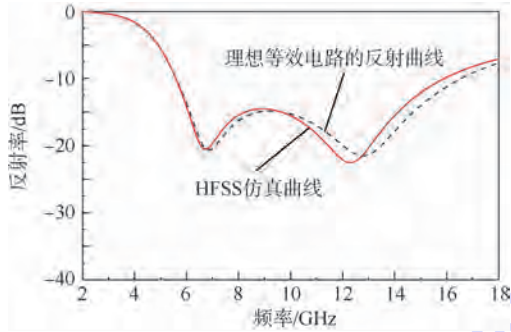


图 5 单层电路模拟吸收体的反射特性  
Fig. 5 Reflectivity characteristics of a single-layer circuit analog absorber

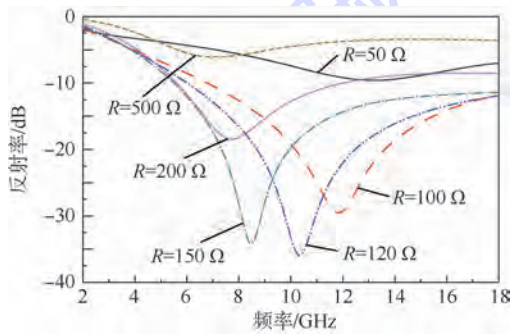


图 6  $R$  取值不同时可调微波吸收体的反射特性  
Fig. 6 Reflectivity characteristics of a tunable microwave absorber with different  $R$

## 2 反射率预估

可调微波吸收体如图 7 所示: 蒙皮层为 FR4 材料, 间隔层为工程泡沫(介电常数为 1.1). 可变阻抗层由 FSS 加载 PIN 二极管组成; 本文选取方环 FSS 单元, 并采用整体并联馈电方式加载电流.

方环单元周期  $p = 9 \text{ mm}$ , 间隔  $g = 2 \text{ mm}$ , 线宽  $w = 0.5 \text{ mm}$ . 利用有限元分析软件仿真, 得到方环可调微波吸收体在垂直入射时的反射率预估曲线(见图 8). 二极管电阻在  $10 \sim 1\,000 \Omega$  之间变化时, 反射率谐振频率相应地变化, 低于  $-15 \text{ dB}$  的反射率曲线覆盖了  $7 \sim 13 \text{ GHz}$ .

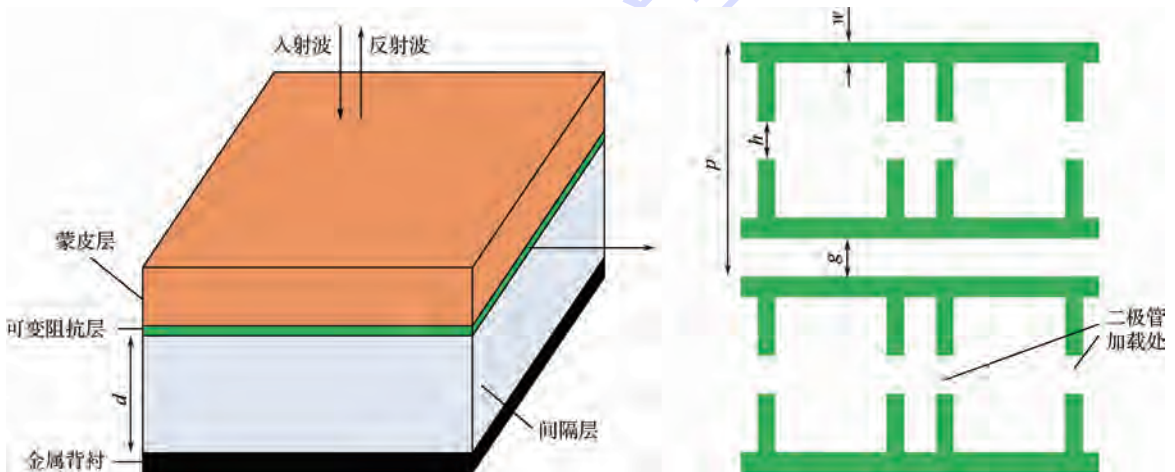


图 7 方环 FSS 单元可调微波吸收体示意图

Fig. 7 A tunable microwave absorber schematic with square FSS unit

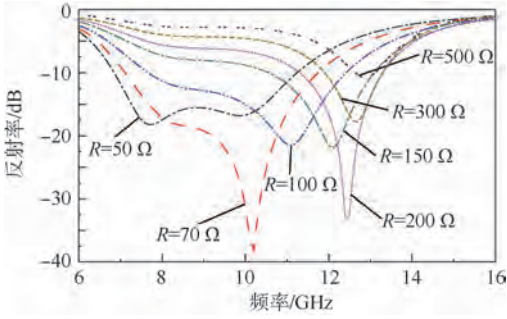


图8 可调微波吸收体反射率预估曲线

Fig.8 Predicted reflectivity characteristics curves of tunable microwave absorber

### 3 样件研制与实验测量

根据上述结构参数,制作方环可调微波吸收体的实验样板(见图9). FSS 样板长 200 mm, 宽 200 mm.

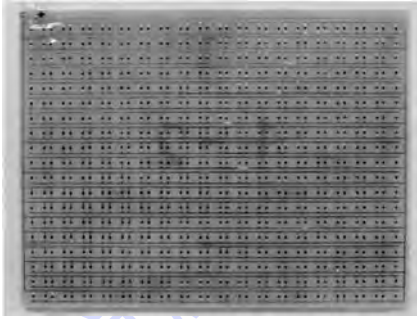


图9 可调微波吸收体实验样品板

Fig.9 Experimental sample plate of tunable microwave absorber

实验测试在微波暗室中进行,测量方法采用远场 RCS 测试法,适合 2 ~ 18 GHz 的反射率测量. 图 10 为测量装置示意图. 其中用到的主要仪器设备有: 矢量网络分析仪、发射接收天线、定向耦合器、微波放大器、扫频信号源、样品支架与转台、电流源等. 反射率测量过程按照 GJB 2038—1994 标准执行,得到的测试曲线如图 11 所示. 当二极管偏置电流  $I$  为 0 mA 时,吸收体几乎没有吸波效果;偏置电流为 0.05 mA 时,吸收体的最佳吸波频

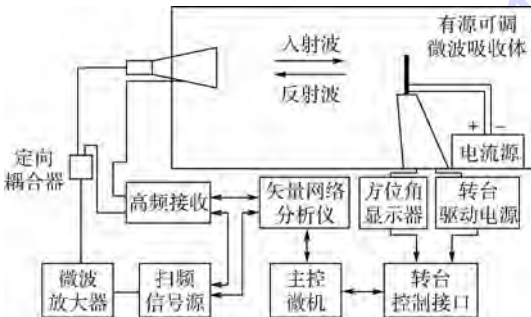


图10 远场 RCS 测试系统组成

Fig.10 Composition of far field RCS measured system

率为12 GHz,最低反射率为 - 20 dB. 随着偏置电流越来越大,吸收体的最佳吸波频率越来越小(见表2);偏置电流为 0.5 mA 时,吸收体的最佳吸波频率达到 7.5 GHz.

从图 11 可以看出,当二极管偏置电流在 0 ~ 0.5 mA 变化时,该实验板对垂直极化的平面波在 7 ~ 12.5 GHz 内实现了吸波特性的可调. 与图 8 比较可以看出,实验与仿真结果基本一致.

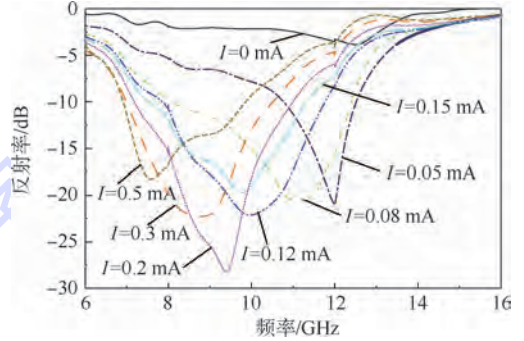


图11 可调微波吸收体反射率特性测量曲线

Fig.11 Measured reflectivity characteristics curves of tunable microwave absorber

表2 偏置电流与最佳吸波频率对照

Table 2 Correspondence between bias current and optimized absorbing frequency

偏置电流/mA	最佳吸波频率/GHz
0.05	12
0.08	11
0.12	10
0.15	9.7
0.2	9.4
0.3	9
0.5	7.5

### 4 优化设计

可调微波吸收体的设计是一个多变量的优化问题. 因为反射系数是涉及材料参数、间隔层厚度、FSS 的形状尺寸等的函数,这些变量有合理的取值范围. 优化的目的就是在这些取值范围内,找出一组最优解,使得设计出的可调微波吸收体具有最优的吸波性能. 应用遗传算法<sup>[14-16]</sup>使这个问题变得切实可行,只要变量的搜索空间不是特别庞大的话,很快就可以找到期望的答案.

针对可调微波吸收体,本文编制了基于遗传算法的优化程序. 为了实现超宽带以及双极化的目的,本文设计了一种中心对称的十字型 FSS 单元结构(见图 12).

利用反射率优化程序,设定参数后,得到最终优化结果:  $p = 5.3 \text{ mm}$ ,  $w = 0.5 \text{ mm}$ , 二极管长度

$h = 1 \text{ mm}$ , 间隔层厚度为  $5.6 \text{ mm}$ , 蒙皮层厚度为  $1.8 \text{ mm}$ . 仿真结果如图 13 所示, 吸收体的反射率随二极管两端的偏置电阻大小变化而变化, 并且垂直极化和水平极化的仿真结果完全一致; 在  $5.6 \sim 17.6 \text{ GHz}$  范围内, 整体反射率包络低于  $-15 \text{ dB}$ , 带宽达到  $12 \text{ GHz}$ . 从而得到了一种宽频段、高吸收、双极化的可调微波吸收体.

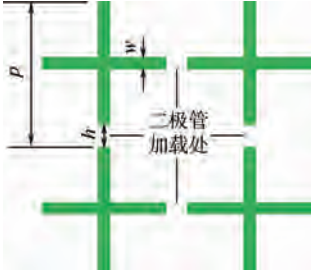


图 12 十字型有源 FSS 示意图

Fig. 12 Schematic diagram of cross active FSS

以有效控制有源可调微波吸收体的最佳吸波频率; 当电流大小在  $0 \sim 0.5 \text{ mA}$  变化时, 该吸收体对垂直极化的入射波在  $7 \sim 12.5 \text{ GHz}$  范围内实现了吸波特性的可控迁移.

3) 有源可调微波吸收体可以根据敌方雷达的威胁频率动态调整自身的吸波频率, 具有智能化特点.

4) 有源可调微波吸收体在同等厚度基础上吸波带宽更大.

优化设计了具有双极化吸波特性的有源可调微波吸收体, 为宽频段、高吸收、双极化的高性能吸波材料研究和开发提供了一种新途径.

## 参考文献 (References)

- [1] Knott E F, Shaeffler J F, Tuley M T. Radar cross section [M]. 2nd ed. Norwood: Artech House, 1993: 156-158.
- [2] Fante R L. Reflection properties of the salisbury screen [J]. IEEE Transactions on Antennas and Propagation, 1988, 36(10): 1443-1454.
- [3] du Toit L J. The design of Jaumann absorbers [J]. IEEE Antennas and Propagation Magazine, 1994, 36(6): 17-25.
- [4] Munk B A. Frequency selective surfaces: Theory and design [M]. New York: Wiley, 2000: 15-50.
- [5] Rozanov K N. Ultimate thickness to bandwidth ratio of radar absorbers [J]. IEEE Transaction Antennas Propagation, 2000, 48(8): 1230-1234.
- [6] Chang T K, Langley R J, Parker E A. An active square loop frequency selective surface [J]. IEEE Transaction Microwave and Guided Wave Letters, 1993, 3(10): 387-388.
- [7] Tennant A, Chambers B. A single-layer tunable microwave absorber using an active FSS [J]. IEEE Microwave and Wireless Components Letters, 2004, 14(1): 46-47.
- [8] 徐欧, 朱敏. 双层方环可电控 FSS 吸波屏设计和实验研究 [J]. 电波科学学报, 2009, 24(5): 837-843.  
Xu O, Zhu M. Design and measurement of active absorber using the double layer square loop FSS [J]. Chinese Journal of Radio Science, 2009, 24(5): 837-843 (in Chinese).
- [9] 高正平, 饶力. 电路模拟技术在吸波结构中的应用 [J]. 电子科技大学学报, 1998, 27(2): 576-579.  
Gao Z P, Rao L. Circuit analog technique applied to structural absorbers [J]. Journal of University of Electronic Science and Technology of China, 1998, 27(2): 576-579 (in Chinese).
- [10] 饶克谨, 赵伯琳, 高正平. 电路模拟吸收材料—原理、特性及设计方法 [J]. 电子科技大学学报, 1995, 24(2): 164-170.  
Rao K J, Zhao B L, Gao Z P. Circuit analogue absorbing material principles, character and design [J]. Journal of University of Electronic Science and Technology of China, 1995, 24(2): 164-170 (in Chinese).
- [11] Langley R J. Equivalent circuit model for arrays of square loops [J]. IEEE Transaction Electronics Letters, 1982, 18(7): 294-296.
- [12] 王焕青. 等效电路法分析频率选择表面的双频特性 [J]. 系

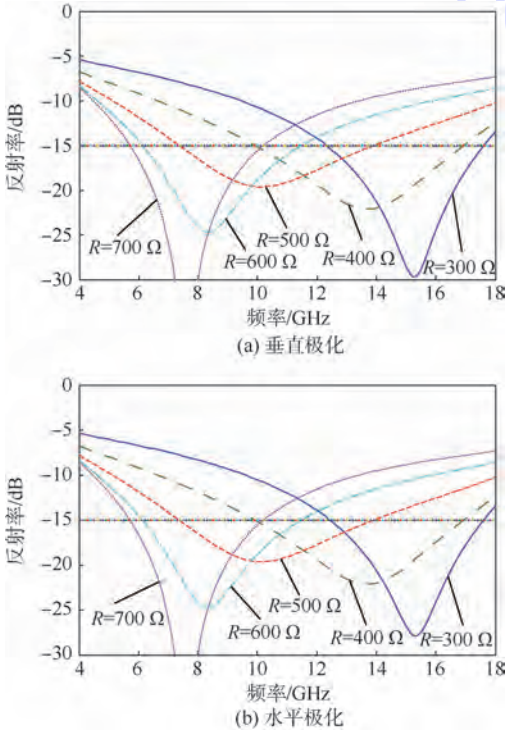


图 13 十字型可调微波吸收体优化结果

Fig. 13 Optimized results of cross tunable microwave absorber

## 5 结论

针对传统雷达吸波材料吸波特性固定的缺点, 本文设计了一种有源可调微波吸收体. 实验研究表明:

1) 可调阻抗层由 FSS 阵列和 PIN 二极管组成, 在 FSS 阵列上设计了并联馈电网络, 便于使用普通的电压/电流源和电表进行调整和校正.

2) 通过改变流过 PIN 二极管的电流大小, 可

统工程与电子技术,2008,30(11):2054-2057.

Wang H Q. Analysis of double properties of frequency selective surfaces by using equivalent circuit method[J]. Systems Engineering and Electronics, 2008, 30 ( 11 ) : 2054-2057 ( in Chinese ).

- [13] Lee C K, Langley R J. Equivalent-circuit models for frequency-selective surfaces at oblique angles of incidence [ J ]. IEEE Transaction Antennas, Microwave Circuits and Networks, 1985, 132(6):395-399.

- [14] Michalewicz Z, Krawczyk J B, Kazemi M, et al. Genetic algorithms and optimal control problem[ C ] // Proceedings of 29th IEEE Conference on Decision and Control. Piscataway, NJ: IEEE Press, 1990:1664-1666.

- [15] Weibe D S, Michiels E, Goldberg D H. Genetic algorithm design of Pareto optimal broad band microwave absorbers[ J ]. IEEE Transaction Electronics Compat, 1996, 38(3):518-524.

- [16] Chambers B. Optimum design of a Salisbury screen radar absorber[ J ]. Electronics Letters, 1994, 30(16):1353-1354.

## Analysis and optimization of active tunable microwave absorber

QI Kainan<sup>1,2</sup>, WANG Yongfeng<sup>1,2</sup>, HOU Xinyu<sup>2</sup>, CHEN Junwen<sup>\*2</sup>

(1. School of Information Engineering, Communication University of China, Beijing 100024, China;

2. Science and Technology on Electromagnetic Scattering Laboratory, Beijing 100854, China)

**Abstract:** In order to improve the limited band of passive radar absorbing materials, an active tunable microwave absorber was designed based on transmission line theory. The absorber was a planar structure based upon the topology of a Salisbury screen, but in which the conventional resistive layer was replaced by an active frequency selective surface (FSS) controlled by PIN diodes. Measured data of reflectivity are presented and show that the reflectivity response of the absorber can be controlled by changing bias current. When bias current changes from 0 to 0.5 mA, the optimal absorbing frequency can be controlled over the frequency band from 7 GHz to 12.5 GHz. A double polarization tunable microwave absorber is optimized based on genetic algorithm, simulation results show that the reflectivity response of the absorber can be controlled over the frequency band from 5.6 GHz to 17.6 GHz.

**Key words:** transmission line theory; microwave absorber; frequency selective surface (FSS); bias current; optimization

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0213



# 受限空间中天线产生的电波覆盖研究

李大伟<sup>1,2</sup>, 王均宏<sup>\*1,2</sup>, 何显时<sup>1,2</sup>

(1. 北京交通大学 电子信息工程学院, 北京 100044; 2. 全光网络与现代通信网教育部重点实验室, 北京 100044)

**摘 要:** 受限空间中的有效电波覆盖直接影响通信质量和可靠性,对天线系统的布局优化尤为重要。然而现有的研究很少考虑本来就处于隧道中的天线参数对电波覆盖的影响。针对铁路隧道这种特殊的受限空间情况首先介绍了电波覆盖的研究方法和现状,然后利用射线追踪方法、波导模式方法和矢量抛物方程方法对电波覆盖进行了预估,针对笔形方向图的特征采用高斯方向图重点研究了不同的天线主瓣宽度、波束指向以及在隧道截面的位置对电波覆盖的影响,分析了沿隧道轴向不同区域内电波覆盖的不同特征,最后研究了受限空间中电波覆盖的三段式模型。本文的方法和模型可以为隧道内的天线选型及优化布局提供理论依据和参考。

**关键词:** 天线; 电波覆盖; 受限空间; 射线追踪; 模式法; 抛物方程

**中图分类号:** TN929.5

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1859-08

受限空间包括有明确物理边界的地铁、隧道、矿井、飞机、轮船等有形通道,以及没有明确物理边界、但希望将电磁波限制在一定范围内传播的高速铁路、高速公路、地上城铁等无形通道。有形边界通道称为硬边界受限空间,无形电波通道称为软边界受限空间,但不管是硬边界还是软边界受限空间,无线通信是保障其中移动载体安全运行、人员安全工作以及对外信息交流的重要手段。而天线类型的选取和天线的优化布局以及电波的高效覆盖是受限空间中无线通信的重要基础,从根本上决定着受限空间中无线通信的质量和可靠性。

早在20世纪20年代,在隧道这类受限空间中的无线通信的尝试就已经开始,然而直到20世纪70年代,无线通信在隧道中才开始大量应用<sup>[1-2]</sup>,隧道环境的电波传播问题也成为热点被广泛研究。到目前为止,隧道环境中的电波覆盖方案主要分为基于漏缆和漏泄波导的漏波结

构<sup>[3-4]</sup>和分布式天线系统<sup>[5-6]</sup>两大类。相比漏波结构,分布式天线具有加工方便、安装灵活、成本低等优势,已经在实际中得到广泛应用。但是受隧道环境的波导效应影响,天线在隧道中的场强分布并不均匀,甚至出现盲区,严重地制约着通信速率和可靠性。所以研究天线在隧道这类受限空间中的电波覆盖规律,预测场强分布情况,一直是分布式天线系统设计无法回避的问题。

近几十年来,国内外诸多学者都对隧道、矿井这类受限空间中的电波传播机制进行了研究, Mahmoud等学者研究了隧道内的传播模式,推导了不同模式的衰减系数<sup>[7-10]</sup>; Dudley等在隧道矿井中做了大量测试和实验,分析了具体环境中的电波覆盖特点<sup>[11-14]</sup>; Zhang等应用数值仿真技术模拟了隧道环境中的场强分布,分析了窄带信号和宽带信号的时延特性,并分析和解释了隧道中的电波传播机制<sup>[15-16]</sup>。然而这些研究工作重点讨

收稿日期: 2015-04-10; 录用日期: 2015-04-24; 网络出版时间: 2015-06-17 15:17

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150617.1517.006.html

基金项目: 国家自然科学基金(61331002)

作者简介: 李大伟(1988—),男,内蒙古赤峰人,博士研究生,11111001@bjtu.edu.cn

\* 通讯作者: 王均宏(1965—),男,江苏东台人,教授, wangjunh@bjtu.edu.cn, 主要研究方向为电磁场与微波技术。

**引用格式:** 李大伟, 王均宏, 何显时. 受限空间中天线产生的电波覆盖研究[J]. 北京航空航天大学学报, 2015, 41(10): 1859-1866.  
Li D W, Wang J H, He X S. Research on radio coverage generated by antennas in confined space[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1859-1866 (in Chinese).

论的是隧道这类传播环境对电波传播的影响,对于本来就处于隧道中的天线对电波覆盖的影响讨论较少.根据文献[17-19]的测试结果可以发现,在隧道环境中天线与无线传播环境相互融合,其方向性、极化以及安装位置都对电波覆盖有显著的影响,因此深入分析天线对隧道内电波覆盖的影响对于实现高效电波覆盖显得尤为重要.

基于以上考虑,本文分别讨论了基于几何光学、波导模式方法以及矢量抛物方程方法的传播模型,重点研究了不同波瓣宽度、波束指向以及安装位置的天线发射时在隧道这类受限空间中的电波覆盖规律.本文研究可以为受限空间中的天线设计及分布式天线系统的优化布局提供重要的理论基础.

## 1 理论和方法

### 1.1 基于射线追踪的信道模型

射线追踪模型基于几何光学理论,在波长远小于反射面的尺寸时,将电磁波近似地看作射线,所有射线的能量在场点叠加,从而得到接收点处的场强.射线追踪模型广泛地应用于复杂环境中的路径损耗计算.在宽  $2a$ 、高  $2b$  的矩形隧道中,在垂直极化天线作为发射端时,基于镜像法的接收场强表达式为

$$E_r(x, y, z) = E_i \frac{1}{r_{m,n}} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(\theta, \varphi) e^{-jk_0 r_{m,n}} R_{\perp}^{|m|} R_{\parallel}^{|n|} \quad (1)$$

式中:  $E_i$  为射线波源处的电场系数,与发射功率相关;  $k_0$  为真空中的波数;  $f$  表示发射天线的方向图函数;  $\theta$  为俯仰角;  $\varphi$  为方位角;  $m$  和  $n$  分别为垂直和水平反射的次数;  $r_{m,n}$  为镜像点  $I_{m,n}$  到场点的距离:

$$r_{m,n} = \sqrt{(x_m - x)^2 + (y_n - y)^2 + z^2} \quad (2)$$

$$x_m = 2ma + (-1)^m x_0 \quad (3)$$

$$y_n = 2nb + (-1)^n y_0 \quad (4)$$

$x_0$  和  $y_0$  分别为场点横、纵坐标;  $x_m$  和  $y_n$  分别为镜像点横、纵坐标;  $R_{\parallel}$  和  $R_{\perp}$  分别为相对于水平墙壁和垂直墙壁的水平反射系数和垂直反射系数(入射波为垂直极化),其表达式为

$$R_{\parallel} = \frac{(\varepsilon_r - j\varepsilon_r') \sin \theta_{\parallel} - \sqrt{(\varepsilon_r - j\varepsilon_r') - \cos^2 \theta_{\parallel}}}{(\varepsilon_r - j\varepsilon_r') \sin \theta_{\parallel} + \sqrt{(\varepsilon_r - j\varepsilon_r') - \cos^2 \theta_{\parallel}}} \quad (5)$$

$$R_{\perp} = \frac{\sin \theta_{\perp} - \sqrt{(\varepsilon_r - j\varepsilon_r') - \cos^2 \theta_{\perp}}}{\sin \theta_{\perp} + \sqrt{(\varepsilon_r - j\varepsilon_r') - \cos^2 \theta_{\perp}}} \quad (6)$$

式中:  $\theta_{\parallel}$  和  $\theta_{\perp}$  分别为射线在水平隧道壁和垂直隧道壁上的反射角度;  $\varepsilon_r$  和  $\varepsilon_r'$  分别为隧道壁的相对介电常数的实部和虚部.该模型中涉及到发射天线的方向图  $f(\theta, \varphi)$ ,本文中假设发射天线是具有笔形方向图特征的定向天线,其方向图可以用高斯函数近似表征:

$$f(\theta, \varphi) = \exp\left[-\frac{(\theta - \theta_0)^2}{\sigma_0 \theta_{\text{BW}}^2} - \frac{(\varphi - \varphi_0)^2}{\sigma_0 \varphi_{\text{BW}}^2}\right] \quad (7)$$

式中:  $\theta_0$  和  $\varphi_0$  为波束的俯仰角和方位角;  $\theta_{\text{BW}}$  和  $\varphi_{\text{BW}}$  表示 E 面和 H 面波束的宽度;  $\sigma_0$  为 0.7213.

通过上述模型,可以建立隧道中任意点的接收场强与发射天线的指向、波束宽度以及位置的关系.

### 1.2 基于波导模式的信道模型

隧道这类受限空间可以看成是一个特大的介质边界的空心波导,隧道中的路径损耗相比自由空间中的路径损耗要小很多,场强分布呈现有规律的加强和减弱,波导效应明显.特别是在超高频(Ultra High Frequency, UHF)和甚高频(Very High Frequency, VHF)波段上,可以用波导模式模型对信道进行建模.与金属波导不同,隧道中传播的模式是横电波与横磁波的混合模式.尽管实际中的隧道多是拱形的或近似矩形的隧道截面,但根据文献[20-21]的研究发现,这些复杂截面的隧道可以等效为矩形或圆形这类规则截面的隧道,因此研究规则截面隧道对实际的电波传播预测有重要的指导意义.在矩形隧道中,不同点的场强可以表示成多种模式的叠加:

$$E(x, y, z) = \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} E_{0pq} e_{pq}(x, y) e^{-(\alpha_{pq} + j\beta_{pq})z} \quad (8)$$

式中:  $\alpha_{pq}$  和  $\beta_{pq}$  分别为衰减常数和相位常数:

$$\alpha_{pq} = \frac{1}{b} \left(\frac{q\lambda}{4b}\right)^2 \operatorname{Re}\left(\frac{\bar{\varepsilon}_b}{\sqrt{\bar{\varepsilon}_b} - 1}\right) + \frac{1}{a} \left(\frac{p\lambda}{4a}\right)^2 \operatorname{Re}\left(\frac{1}{\sqrt{\bar{\varepsilon}_a} - 1}\right) \quad (9)$$

$$\beta_{pq} = \sqrt{k_0^2 - \left(\frac{p\pi}{2a}\right)^2 - \left(\frac{q\pi}{2b}\right)^2} \quad (10)$$

$\bar{\varepsilon}_a$  和  $\bar{\varepsilon}_b$  分别为隧道垂直墙壁和水平墙壁的相对复介电常数;  $e_{pq}$  为模式的本征函数;  $E_{0pq}$  为模式的复幅度系数,与激励源紧密相关.文献[22-23]给出了点源天线辐射时  $E_{0pq}$  的表达式为

$$E_{0pq} = \frac{\pi}{ab \sqrt{1 - \left(\frac{p\pi}{2ak}\right)^2 - \left(\frac{q\pi}{2bk}\right)^2}} \cdot \sin\left(\frac{p\pi}{2a}x_0 + \varphi_x\right) \cos\left(\frac{q\pi}{2b}y_0 + \varphi_y\right) \quad (11)$$

式中:

$$\begin{cases} \varphi_x = \begin{cases} 0 & p \text{ 为偶数} \\ \frac{\pi}{2} & p \text{ 为奇数} \end{cases} \\ \varphi_y = \begin{cases} \frac{\pi}{2} & q \text{ 为偶数} \\ 0 & q \text{ 为奇数} \end{cases} \end{cases} \quad (12)$$

然而实际应用中,隧道中的天线大都是定向天线,点源天线的模型并不能用来分析这类天线的覆盖特性.文献[24]提出了一种基于射线追踪于波导模式的混合方法,这种方法首先应用射线追踪方法计算距离天线一定距离上的隧道横截面上的场强分布,然后对横截面上的场强进行模式匹配,从而得到不同模式的激励幅度:

$$E_{0pq} = \frac{\int_S (\mathbf{E}_{\text{REF}} \times \mathbf{h}_{pq}^*) \cdot \mathbf{i}_z dS}{\int_S (\mathbf{e}_{pq} \times \mathbf{h}_{pq}^*) \cdot \mathbf{i}_z dS} \quad (13)$$

式中: $\mathbf{E}_{\text{REF}}$ 为隧道参考截面上的电场分布; $\mathbf{h}_{pq}$ 为本征函数; $\mathbf{i}_z$ 为 $z$ 方向的单位矢量.这种方法一方面规避了射线追踪计算复杂的缺点,另一方面将天线辐射特性与隧道传播模式联系起来,可以从理论上对隧道内的电波传播特点进行分析和解释.

### 1.3 基于矢量抛物方程的信道模型

矢量抛物方程<sup>[25-27]</sup>是一种求解亥姆霍兹方程的近似方法,当电磁波近轴传播时,近似具有很高的精确性,非常适合隧道这种近轴传播问题,近年来广泛地应用于隧道中的电波传播预测.标准的抛物方程可以表示为

$$\frac{\partial u}{\partial z} = \frac{1}{2jk_0} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u \quad (14)$$

抛物方程的差分求解方法是基于CN(Crank-Nicolson)格式离散的交替方向隐式(alternating-direction implicit)方法,其中对二次偏微分采用PR(Peaceman-Rachford)方法进行离散<sup>[26]</sup>,这种离散具有二阶精度.

$$[1 - (r_x/4jk_0)\delta_x] \tilde{u}^{n+1/2} = [1 - (r_y/4jk_0)\delta_y] u^n \quad (15)$$

$$[1 - (r_y/4jk_0)\delta_y] u^{n+1} = [1 - (r_x/4jk_0)\delta_x] \tilde{u}^{n+1/2} \quad (16)$$

式中:

$$\delta_x u_{i,l} = u_{i+1,l} - 2u_{i,l} + u_{i-1,l} \quad (17)$$

$$\delta_y u_{i,l} = u_{i,l+1} - 2u_{i,l} + u_{i,l-1} \quad (18)$$

$r_x$ 和 $r_y$ 为与网格划分尺寸相关的常数,分别为 $\Delta z/\Delta x^2$ 和 $\Delta z/\Delta y^2$ ;  $\delta_x$ 和 $\delta_y$ 为 $x$ 和 $y$ 方向的差分; $u$ 为场量.通过插入 $z = n + 1/2$ 的中间平面,可以

得到步进平面上任意点的场强.为了提高抛物方程方法的计算精度,文献[28]给出了采用MF(Mitchell-Fairweather)离散方法的抛物方程格式:

$$[1 - (r_x/4jk_0 - 1/12)\delta_x] \tilde{u}^{n+1/2} = [1 + (r_y/4jk_0 + 1/12)\delta_y] u^n \quad (19)$$

$$[1 - (r_y/4jk_0 - 1/12)\delta_y] u^{n+1} = [1 + (r_x/4jk_0 + 1/12)\delta_x] \tilde{u}^{n+1/2} \quad (20)$$

这种离散具有四阶精度.

针对隧道这种受限空间,边界采用阻抗边界条件进行截断,Leontovich阻抗边界条件可以表示为<sup>[25]</sup>

$$\begin{cases} \left[ u \pm \frac{Z_{\parallel}}{jk_0} \cdot \frac{\partial u}{\partial x} \right]_{x=\pm a} = 0 \\ \left[ u \mp \frac{1}{jk_0 Z_{\perp}} \cdot \frac{\partial u}{\partial y} \right]_{y=\pm b} = 0 \end{cases} \quad (21)$$

式中: $Z_{\parallel}$ 和 $Z_{\perp}$ 分别为隧道侧壁和上下壁的表面阻抗.

### 1.4 模型验证

为了验证方法的正确性,本文对文献[13]给出的隧道结构进行仿真,隧道宽7.8 m,高5.3 m,发射天线与接收天线均垂直极化,分别放置在偏离隧道中心水平1.325 m、高2 m处,天线增益7 dBi,工作频率900 MHz.隧道壁的相对介电常数 $\epsilon_r$ 为5,电导率 $\sigma$ 为0.01 S/m.图1给出了文献[13]中法国Massif Central隧道的测试结果与不同

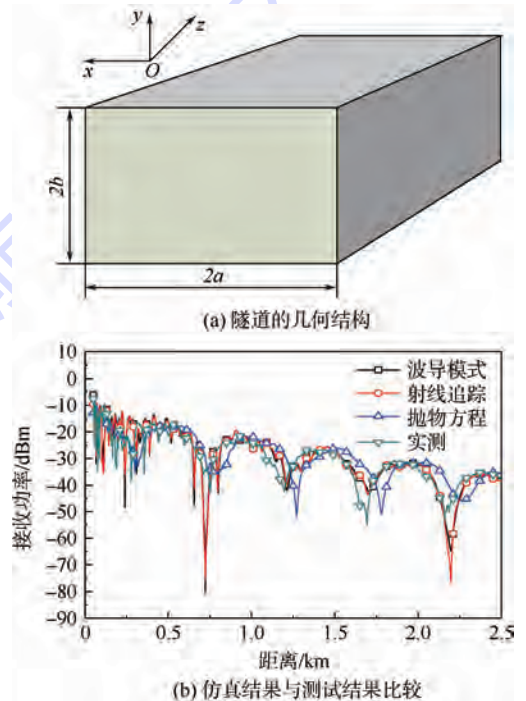


图1 不同方法所得接收功率与测试结果的比较

Fig.1 Comparison of received power among different methods with measurement result

方法仿真时接收功率随距离变化曲线的比较. 由图可以看出, 上述3种方法与测试数据吻合很好, 可以清晰地看出近场区的快速波动和远场区的波导效应.

## 2 天线产生的电波覆盖分析

本文采用射线追踪与波导模式结合的方法对第1节中的隧道模型进行了仿真, 分析了不同天线参数对隧道中电波覆盖的影响.

### 2.1 天线参数对电波覆盖的影响

本文中的天线参数主要指的是发射天线的波瓣宽度、波束指向和天线的安装位置, 隧道参数如第1.4节所述, 接收天线位于隧道截面中心, 发射天线和接收天线均垂直极化.

图2给出了天线在隧道横截面中心且波束指向隧道中心线方向时, 波瓣宽度变化对隧道中心线上电磁波功率分布的影响.

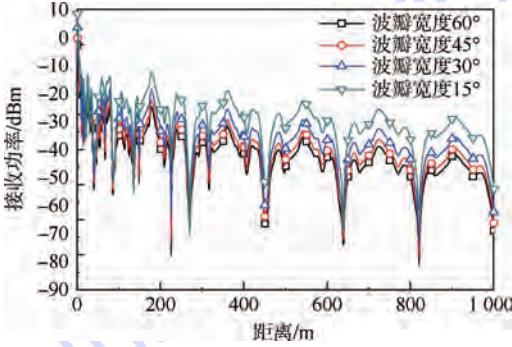


图2 发射天线波瓣宽度不同时接收功率分布沿隧道的变化

Fig. 2 Received power distribution variations in tunnel for different beam widths generated by transmitting antenna

从图2中可以发现以下规律:

1) 功率密度分布具有相同的形状, 但是每条曲线的电平有明显差异, 这是由于在相同输入功率的情况下, 波瓣较宽的天线将更多的能量直接射向天线附近的隧道墙壁, 导致近处的损耗增加, 从而远处的电平降低.

2) 在靠近发射天线的近区, 场强波动剧烈, 在远区, 场强变化相对缓慢, 这主要是由于波导中不同模式在轴向具有不同的相位常数, 传播过程中会不断地产生同相叠加和反相相消, 从而形成场强加强点和减弱点, 又由于天线近处的高次模含量较多, 并且衰减常数不同, 所以近处的加强点和减弱点变化较快, 而远区主要是低阶模式, 衰减常数相差不大, 因而叠加和相消的位置基本固定.

3) 近区的场强衰减迅速, 远区衰减缓慢, 这主要是由于近区的高次模衰减迅速, 而远处保留下来的低次模衰减缓慢所导致的.

4) 场强的加强点与减弱点位置与隧道的横截面尺寸紧密相关, 因为隧道横截面尺寸直接影响各模式的相位常数; 对于这里讨论的垂直极化的情况, 加强点间的距离为  $2\pi/(\beta_{11} - \beta_{s1})$ , 这是由于  $\text{EH}_{11}$  模式和  $\text{EH}_{s1}$  模式具有较小的衰减系数, 能够传播较远的距离, 在远场区域成为主要模式, 因此这两种模式的叠加和相消决定着接收功率的分布形态. 隧道沿中心线剖开的两个相互垂直纵向截面内的场强分布如图3所示.

图4给出了天线的波瓣宽度固定为  $15^\circ$ , 并且位于隧道横截面中心时, 波束指向变化对电波覆盖的影响 (曲线图均是半波偶极子接收功率图),

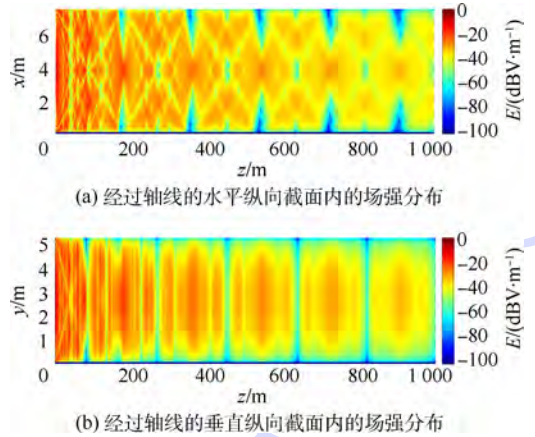


图3 波瓣宽度为  $15^\circ$  的发射天线位于隧道横截面中心时的隧道内场强分布

Fig. 3 Received E-field distribution with transmitter located in center of cross section of tunnel, and with antenna beam width of  $15^\circ$

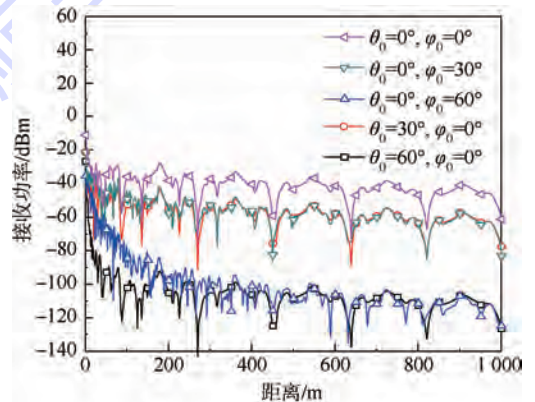


图4 发射天线波束指向不同时接收功率分布沿隧道中心线的变化

Fig. 4 Received power distribution variation along center line of the tunnel for different beam orientation angles

彩图是场强分布图). 从图 4 中可以发现当波束俯仰角变化时近区的电平要比方位角以相同角度变化时衰减快,更快地形成稳定的波导模式分布,而在远区它们的电平则基本相同,图 5 给出了天线波束方位角和俯仰角不同时隧道沿中心线剖开的两个相互垂直纵向截面内的场强分布.

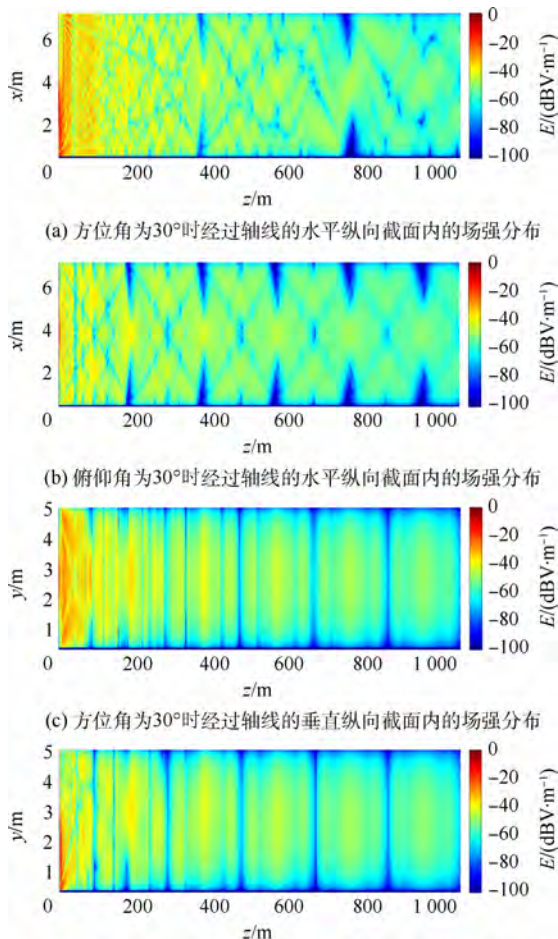


图 5 天线波束方位角和俯仰角不同时隧道内的场强分布

Fig. 5 Received E-field distribution in tunnel with different altitude and azimuth angles

图 6 给出了发射天线位于隧道横截面内不同位置时沿隧道中心线的接收功率的分布曲线(坐标原点位于隧道截面中心). 此时发射天线主波瓣指向隧道中心线方向,波瓣宽度为  $15^\circ$ . 从图 6 中可以发现:①天线沿水平方向移动时接收功率曲线的形状基本没有变化,只是在电平幅度上有所降低;②天线沿垂直方向移动时,接收功率曲线变化明显,峰值和低谷的位置发生移动.

图 7 给出了天线位置不同时隧道沿中心线剖开的两个相互垂直纵向截面内的场强分布,由图可知当天线在水平和垂直方向移动时,场强分布差异巨大,这是由于天线位置不同时,各激励模式

的复系数的幅度和相位不同,造成不同模式叠加的规律不同.

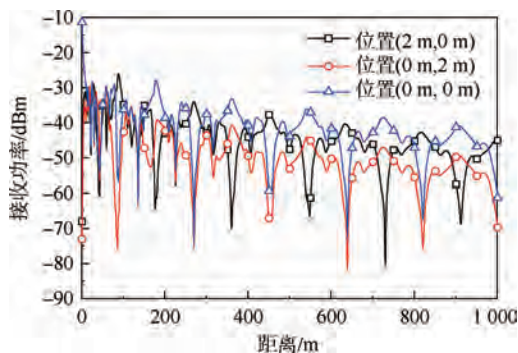


图 6 发射天线位置不同时接收功率沿隧道中心线的变化

Fig. 6 Received power distribution variations along the center line of the tunnel for different transmitting antenna positions

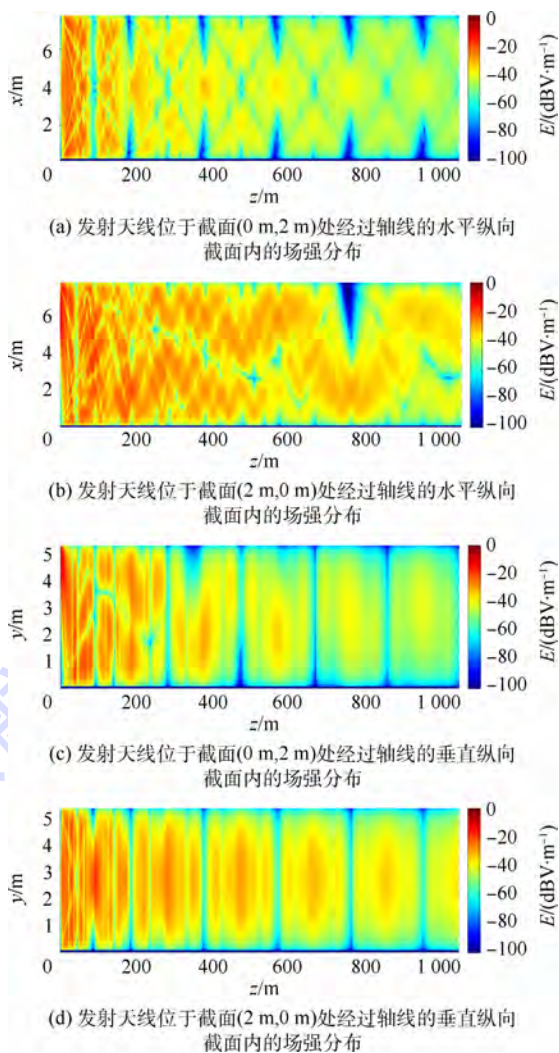


图 7 发射天线位置不同时隧道内的场强分布

Fig. 7 Received E-field distribution in tunnel with for different transmitting antenna positions

## 2.2 三段式电波传播模型及间断点

从第 2.1 节的结果和分析中可以发现,距离

发射天线不同距离处电波覆盖的特性也不同. 距离较近时接收功率衰减快, 并且伴随着快速波动; 距离远时接收功率衰减慢, 波动缓慢, 且出现固定的加强点和减弱点. 移动通信中将这两类区域称作近场区和远场区, 这两类区域的分界点叫做间断点. 间断点的确定对信道模型的建立有重要意义, 文献[5]给出了间断点的计算公式:

$$Z_{NF} = \max\left(\frac{W^2}{\lambda}, \frac{H^2}{\lambda}\right) \quad (22)$$

式中:  $W$  和  $H$  分别为隧道截面的宽度和高度;  $\lambda$  为真空中的波长. 而在上述两类区域中的近场区域内, 又可以看出, 在靠近天线的地方呈现出指数衰减, 类似于自由空间, 称为自由空间传播区域; 远离天线一点的区域高次模作用强烈, 衰落很大, 称为多模作用区域. 因此近场区也存在一个间断点. 文献[29]讨论了这个间断点的位置, 当接收天线移动到使第一菲涅耳区和隧道壁相切时的位置就是第一间断点的位置. 根据以上分析本文建立了隧道内的三段式电波传播模型, 如图8所示. 图8给出了天线波束方位角变化时的接收功率曲线及相应的三段式模型, 当方位角(偏离隧道轴线方向)增大时, 自由空间传播区域(第1段直线区域)范围会有所增大, 多模区的范围(第2段直线区域)会显著地扩大, 并且信号衰落明显, 并伴随场强的快速波动, 远场区(第3段直线区域)的开始端距发射天线的距离会明显增加. 由此可见, 隧道内的电波传播区域不仅受限于隧道的几何尺寸和波长, 与天线位置及姿态也是息息相关的.

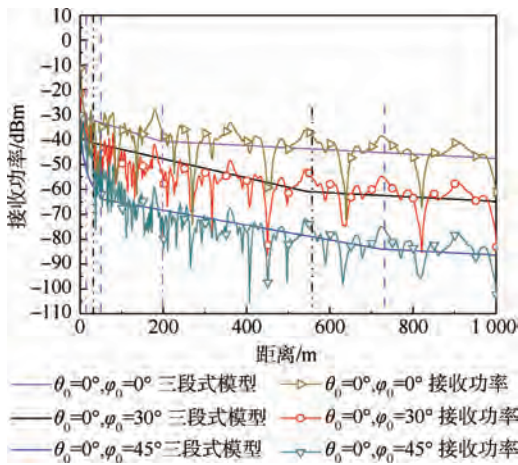


图8 发射天线波束指向不同时沿隧道中心线接收功率分布的变化及其对三段式电波传播模型的影响

Fig.8 Received power distribution variations and corresponding three slopes propagation modals along center line of tunnel for different beam orientation angles of transmitting antenna

### 3 结论

1) 分析了发射天线的波瓣宽度、波束指向以及天线位置对隧道内电波覆盖的影响. 研究表明, 天线波束宽度影响电平分布的幅度, 但不影响分布规律; 波束指向方位角的变化对场强分布的影响较大, 俯仰角的变化对场强分布的影响不甚明显; 天线位置的水平移动对场强分布影响大, 垂直移动对场强分布基本没有影响, 只是幅度上有所降低.

2) 本文还研究了三段式的电波传播模型, 给出了间断点的物理意义和位置. 本文的研究为受限空间无线通信系统天线的优化设计提供了理论依据.

### 参考文献 (References)

- [1] Large D G, Ball L, Farstad A J. Radio transmission to and from underground coal mines-theory and experiment[J]. IEEE Transactions on Communication, 1973, 21(3): 194-202.
- [2] Forooshani A E, Bashir S, Michelson D G, et al. A survey of wireless communications and propagation modeling in underground mines[J]. IEEE Communications Surveys and Tutorials, 2013, 15(4): 1524-1545.
- [3] Fishahn S, Camp M, Garbe H. Transient emission behavior of leaky section cables for wireless applications[C] // Proceedings of International Symposium on Electromagnetic Compatibility. Piscataway, NJ: IEEE Press, 2005: 477-482.
- [4] 毛万恒. 北京亦庄线商用通信无线覆盖技术[J]. 市政技术, 2011, 29(1): 89-94.  
Mao W H. Wireless coverage technology of commercial communications for Beijing Yizhuang metro line[J]. Municipal Engineering Technology, 2011, 29(1): 89-94 (in Chinese).
- [5] Briso-Rodriguez C, Cruz J M, Alonso J I. Measurements and modeling of distributed antenna systems in railway tunnels[J]. IEEE Transactions on Vehicle Technology, 2007, 56(5): 2870-2879.
- [6] Guan K, Zhong Z D, Alonso J I, et al. Measurement of distributed antenna systems at 2.4 GHz in a realistic subway tunnel environment[J]. IEEE Transactions on Vehicle Technology, 2012, 61(2): 834-837.
- [7] Mahmoud S F, Wait J R. Geometrical optical approach for electromagnetic wave propagation in rectangular mine tunnels[J]. Radio Science, 1974, 9(12): 1147-1158.
- [8] Emslie A G, Lagace R L, Strong P F. Theory of the propagation of UHF radio waves in coal mine tunnels[J]. IEEE Transactions on Antennas and Propagation, 1975, 23(2): 192-205.
- [9] Mahmoud S F. On the attenuation of monofilar and bifilar modes in mine tunnels[J]. IEEE Transactions on Microwave Theory and Technology, 1974, 22(3): 845-847.
- [10] Wait J. Propagation in rectangular tunnel with imperfectly conducting walls[J]. Electronics Letters, 1980, 16(13): 521-522.

- [11] Dudley D G, Mahmoud S F. Wireless propagation in circular tunnels[J]. *IEEE Transactions on Antennas and Propagation*, 2005, 53(1):435-441.
- [12] Lienard M, Degauque P. Propagation in wide tunnels at 2 GHz: A statistical analysis[J]. *IEEE Transactions on Antennas and Propagation*, 1988, 47(4):1322-1328.
- [13] Dudley D G, Lienard M, Mahmoud S F, et al. Wireless propagation in tunnels[J]. *IEEE Antennas and Propagation Magazine*, 2007, 49(2):11-26.
- [14] Zhang Y P, Hwang Y. Characterization of UHF radio propagation channels in tunnel environments for microcellular and personal communications[J]. *IEEE Transactions on Vehicle Technology*, 1988, 47(1):283-296.
- [15] Zhang Y P. Novel model for propagation loss prediction in tunnels[J]. *IEEE Transactions on Vehicle Technology*, 2003, 52(3):1308-1314.
- [16] Pallares F, Juan F, Juan-Llaser L. Analysis of path loss and delay spread at 900 MHz and 2.1 GHz while entering tunnels[J]. *IEEE Transactions on Vehicle and Technology*, 2001, 50(3):767-776.
- [17] Mabrouk I B, Talbi L, Nedil M, et al. Effect of antenna directivity on performance of multiple input multiple output systems in an underground gold mine[J]. *IET Microwaves, Antennas and Propagation*, 2012, 6(5):555-561.
- [18] Bashir S. Effect of antenna position and polarization on UWB propagation channel in underground mines and tunnels[J]. *IEEE Transactions on Antennas and Propagation*, 2014, 62(9):4771-4779.
- [19] Castiblanco J A, Seetharamdoo D, Ney M, et al. Determination of antenna specification and positioning for efficient railway communication in tunnels of arbitrary cross section[C]// *Proceedings on Workshop on Transceiver Design for ITS*. Piscataway, NJ: IEEE Press, 2011:678-681.
- [20] Molina-García-Pardo J M, Lienard M, Nasr A, et al. On the possibility of interpreting field variations and polarization in arched tunnels using a model for propagation in rectangular and circular tunnels[J]. *IEEE Transactions on Antennas and Propagation*, 2008, 56(4):1206-1211.
- [21] Mahmoud S F. Wireless transmission in tunnels with non-circular cross section[J]. *IEEE Transactions on Antennas and Propagation*, 2010, 58(2):613-616.
- [22] Sun Z, Akyildiz I. Channel modeling and analysis for wireless networks in underground mines and road tunnels[J]. *IEEE Transactions on Communication*, 2010, 58(6):1758-1769.
- [23] Zhou C M, Waynert J. The equivalence of the ray tracing and modal methods for modeling radio propagation in lossy rectangular tunnels[J]. *IEEE Antennas and Wireless Propagation Letters*, 2014, 13:615-618.
- [24] Fuschini F, Falciasecca G. A mixed rays-modes approach to the propagation in real road and railway tunnels[J]. *IEEE Transactions on Antennas and Propagation*, 2012, 60(2):1095-1105.
- [25] Popov A V, Zhu N Y. Modeling radio wave propagation in tunnels with a vectorial parabolic equation[J]. *IEEE Transactions on Antennas and Propagation*, 2000, 48(9):1403-1412.
- [26] Martelly R, Janaswamy R. An ADI-PE approach for modeling radio transmission loss in tunnels[J]. *IEEE Transactions on Antennas and Propagation*, 2009, 57(6):1759-1770.
- [27] Levy M. Parabolic equation methods for electromagnetic wave propagation[M]. London: Institute Electronic Engineering, 2000:35-42.
- [28] Zhang X Q, Sarris C D. A high-accuracy ADI scheme for the vector parabolic equation applied to the modeling of wave propagation in tunnels[J]. *IEEE Antennas and Wireless Propagation Letters*, 2014, 13:650-653.
- [29] Guan K, Zhong Z D, Ai B, et al. Propagation mechanism modeling in the near-region of arbitrary cross-sectional tunnels[J]. *International Journal of Antennas and Propagation*, 2012, 2012:1-11.

# Research on radio coverage generated by antennas in confined space

LI Dawei<sup>1,2</sup>, WANG Junhong<sup>\*1,2</sup>, HE Xianshi<sup>1,2</sup>

(1. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. Key Laboratory of All Optical Network & Advanced Telecommunication Network of Ministry of Education, Beijing 100044, China)

**Abstract:** Efficient radio wave coverage in confined space is one of the foundation factors to guarantee quality and reliability of communication which is important for the layout and optimization of the antenna system. However, the existing researches rarely consider the effect of the antenna parameters on field coverage in confined space. The research status and methods of radio wave coverage in confined space, aimed at railway tunnel, were reviewed. Then the deterministic models for field coverage prediction in the railway tunnel were adopted which were ray tracing method, modal method and parabolic equation method. By combining these methods, the effect of the antenna beam width, beam direction and fixed position in the cross section of the tunnel were analyzed. Then the field coverage characteristics in different regions were analyzed. At last, a three slopes propagation model was proposed for studying the field coverage characteristics in confined space. The method and model can provide the theoretical basis and a reference for antenna design and layout optimization in the tunnels.

**Key words:** antenna; radio coverage; confined space; ray tracing; modal method; parabolic equation

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0214

# 等效原理算法在不同等效面下的误差分析

党训旺<sup>1</sup>, 李懋坤<sup>\*1,2</sup>, 杨帆<sup>1,2</sup>, 许慎恒<sup>1,2</sup>

(1. 清华大学 电子工程系, 北京 100084; 2. 清华信息科学与技术国家实验室(筹), 北京 100084)



**摘 要:** 等效原理算法(EPA)借鉴了区域分解的思想,并基于惠更斯等效原理实现了对大规模散射问题的计算. 本文对EPA中的等效原理算子(EPO)进行了误差分析,选取了立方体、球体和进行了光滑处理的立方体等作为等效面寻找误差来源,并提出了误差改进的建议. 根据计算结果,等效磁流在等效面上描述的误差一般集中在等效面表面的几何不连续处,且主要受等效面上的法向量不连续性和基函数的选取的影响. 故在选取等效面时应尽量考虑使用平滑的等效面,而不是直接使用立方体等效面.

**关键词:** 惠更斯等效原理; 区域分解方法; 误差分析; 计算电磁学; RWG基函数

**中图分类号:** TN011

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1867-06

随着计算机性能的提高,计算电磁学领域在近几十年来有了很大的发展,人们越来越需要求解大规模电磁场问题. 大规模电磁场问题中未知数数量多,矩阵条件数高,难以准确求解<sup>[1]</sup>. 同时,在一个大规模电磁场计算问题中往往涉及多种尺度的结构,难以采用单一算法同时求解<sup>[2]</sup>. 由此人们发展了多种区域分解算法(Domain Decomposition Method, DDM),并在计算电磁学中进行了广泛的应用,如基于微分方程的DDM<sup>[3]</sup>、积分方程的DDM<sup>[4]</sup>等.

等效原理算法(Equivalence Principle Algorithm, EPA)是一种基于惠更斯等效原理、借鉴了区域分解思想的算法<sup>[5]</sup>. 目前EPA的研究方向较多. Li和Chew讨论了散射体穿过等效面时EPA的公式形式<sup>[6-7]</sup>. Ylä-Oijala和Taskinen报道了一种提高等效面接近散射体时算法精度的方法T-EPA(Tangential EPA)<sup>[8]</sup>, Shao等<sup>[9]</sup>对T-EPA中求逆操作做了进一步改进,提升了其计算速度. Sun等<sup>[10]</sup>对电场积分方程中的“低频崩溃”问题

进行了改进,借鉴A-EFIE(Augmented Electric Field Integral Equation)<sup>[11]</sup>的方式提出了A-EPA(Augmented EPA)方法. 针对大型阵列问题, Zhang等<sup>[12]</sup>提出了PEPA(Periodic EPA)方法,利用周期性结构的特点减少EPA等效面间相互作用的计算量. Shi等<sup>[13]</sup>将EPA推广到了时域. 此外,南京理工大学的杨晨<sup>[14]</sup>和西安电子科技大学的张博<sup>[15]</sup>针对EPA分别采用了快速多极子算法(Fast Multipole Algorithm, FMA)和自适应交叉近似(Adaptive Cross Approximation, ACA)进行加速. 可见EPA已逐渐从理论走向实际应用.

由于EPA中涉及多次矩阵乘法与基函数变换,且每一次运算都有可能引入一定误差,对等效原理算法的误差来源进行分析就变得十分必要. 本文计算了选取不同形状的等效面时其表面等效磁流的误差,并相互比较,分析误差来源.

## 1 算法简介

### 1.1 等效原理简介与符号说明

EPA的基本思路基于等效原理,并可以用电

收稿日期: 2015-04-10; 录用日期: 2015-04-24; 网络出版时间: 2015-08-03 14:23

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150803.1423.005.html

基金项目: 清华大学电子工程系传信研究基金

作者简介: 党训旺(1992—),男,河北盐山人,博士研究生, dxw13@mails.tsinghua.edu.cn

\* 通讯作者: 李懋坤(1979—),男,辽宁锦州人,副教授, maokunli@tsinghua.edu.cn, 主要研究方向为计算电磁学.

**引用格式:** 党训旺, 李懋坤, 杨帆, 等. 等效原理算法在不同等效面下的误差分析[J]. 北京航空航天大学学报, 2015, 41(10): 1867-1872. Dang X W, Li M K, Yang F, et al. Error analysis of equivalence principle algorithm on different equivalence surfaces[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1867-1872 (in Chinese).

磁场的表面积分形式表达. 自由空间中的格林函数为

$$G_0(\mathbf{r}, \mathbf{r}') = \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \quad (1)$$

式中:  $k$  为电磁波的波数;  $\mathbf{r}$  为观测点的位置;  $\mathbf{r}'$  为辐射源的位置.

本文定义了 2 个积分算子  $\bar{L}_p(\mathbf{X})$  和  $\bar{K}_p(\mathbf{X})$  表示对一个散射体  $p$  上向量场  $\mathbf{X}$  的某种表面积分运算:

$$\bar{L}_p(\mathbf{X}) = \frac{1}{jk} \nabla \times \int_{S_p} \nabla G_0(\mathbf{r}, \mathbf{r}') \times \mathbf{X}(\mathbf{r}') dS'_p \quad (2)$$

$$\bar{K}_p(\mathbf{X}) = -\text{P. V.} \int_{S_p} \nabla G_0(\mathbf{r}, \mathbf{r}') \times \mathbf{X}(\mathbf{r}') dS'_p \quad (3)$$

式中: P. V. (Principal Value) 表示取积分的主值;  $S_p$  为散射体  $p$  的表面. 根据等效原理<sup>[6]</sup>, 利用式(2)和式(3)可知, 在  $S_p$  外任意一点  $\mathbf{r}$  处的场强只与  $S_p$  表面的等效源有关, 为

$$\begin{bmatrix} \mathbf{E}(\mathbf{r}) \\ \mathbf{H}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \eta_0 \bar{L}_p & \bar{K}_p \\ -\bar{K}_p & \bar{L}_p \end{bmatrix} \begin{bmatrix} \mathbf{J}_p(\mathbf{r}') \\ \mathbf{M}_p(\mathbf{r}') \end{bmatrix} \quad (4)$$

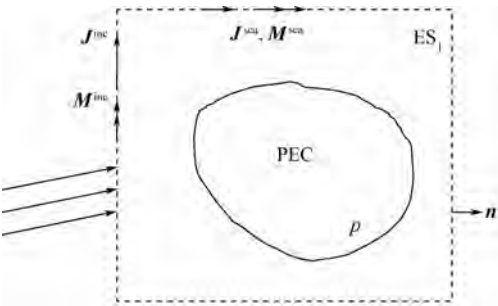
式中:  $\eta_0$  为自由空间的波阻抗;  $\mathbf{J}_p(\mathbf{r}')$  和  $\mathbf{M}_p(\mathbf{r}')$  分别为  $S_p$  上的电流和磁流, 根据等效原理, 它们与  $S_p$  表面的磁场与电场有如下关系:

$$\begin{cases} \mathbf{J}(\mathbf{r}') = \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}') \\ \mathbf{M}(\mathbf{r}') = -\hat{\mathbf{n}} \times \mathbf{E}(\mathbf{r}') \end{cases} \quad (5)$$

此时  $S_p$  即为等效面 (Equivalence Surface, ES).

### 1.2 等效原理算法简介

以单个理想电导体 (Perfect Electric Conductor, PEC) 的散射为例, 介绍等效原理算法. 图 1 为一个 PEC 散射体  $p$  和等效面  $ES_1$  的示意图.



$M^{inc}, J^{inc}$ —等效面上的等效入射磁流和电流;  
 $M^{sca}, J^{sca}$ —等效面上的等效散射磁流与电流.

图 1 EPA 示意图

Fig. 1 An illustration of EPA

$$\begin{bmatrix} \mathbf{M}^{sca} \\ \mathbf{J}^{sca} \end{bmatrix} = \begin{bmatrix} -\hat{\mathbf{n}} \times \eta_0 \bar{L} \\ -\hat{\mathbf{n}} \times \bar{K} \end{bmatrix} [\eta_0 \bar{L}_{pp}]^{-1} \cdot$$

$$[\bar{K} \quad \eta_0 \bar{L}] \begin{bmatrix} \mathbf{M}^{inc} \\ \mathbf{J}^{inc} \end{bmatrix} = \bar{S}_{11} \begin{bmatrix} \mathbf{M}^{inc} \\ \mathbf{J}^{inc} \end{bmatrix} \quad (6)$$

式(6)的物理含义是 ES 上的等效入射磁流

和电流首先通过  $[\bar{K} \quad \eta_0 \bar{L}]$  算子 (Outside In 算子)<sup>[6]</sup> 向内辐射至 PEC 上, 然后对于 PEC 采用电场积分方程 (Electric Field Integral Equation, EFIE) 得到其表面电流.  $\bar{L}_{pp}$  表示求解 PEC 散射体上电流的 EFIE 求解算子, 下标  $pp$  表示  $\bar{L}(\mathbf{X})$  算子从散射体  $p$  作用到自身. 如果散射体不是 PEC, 只需要将求解散射体表面电流的 EFIE 求解算子换成相应的混合场积分方程 (Combined Field Integral Equation, CFIE) 算子即可. 最后由

$[-\hat{\mathbf{n}} \times \eta_0 \bar{L} \quad -\hat{\mathbf{n}} \times \bar{K}]^T$  算子 (Inside Out 算子)<sup>[6]</sup> 再将 PEC 上的电流向外投影到等效面  $ES_1$  上, 得到等效散射磁流和电流. 根据等效原理, 该 PEC 对外界的辐射场与等效面上的等效磁流与电流形成的辐射场完全一致. 整个过程可以用等效面  $ES_1$  上的等效原理算子 (Equivalence Principle Operator, EPO)<sup>[6]</sup>  $\bar{S}_{11}$  表示. 当 PEC 表面的未知数较多时, 可以在等效面上使用较少的未知数等效原始 PEC 向外辐射的特性.

在多个散射体的情况下, 第  $i$  个等效面上的等效入射电流和磁流除了有原始入射场的贡献, 还有除  $i$  以外其他等效面的辐射贡献, 其中第  $j$  个等效面的辐射贡献可以用传输算子 (Transmission Operator, TO)<sup>[6]</sup>  $\bar{T}_{ij}$  表示, 如式(7)所示. 这样可以将散射体之间未知数的相互耦合转移到等效面之间的相互耦合, 可以减少计算量, 改善待求解矩阵的性态.

$$\bar{T}_{ij} = \begin{bmatrix} -\hat{\mathbf{n}} \times \bar{K}_{ij} & -\hat{\mathbf{n}} \times \eta_0 \bar{L}_{ij} \\ \hat{\mathbf{n}} \times \frac{1}{\eta_0} \bar{L}_{ij} & -\hat{\mathbf{n}} \times \bar{K}_{ij} \end{bmatrix} \quad (7)$$

### 1.3 矩量法求解

入射场的等效磁流  $\mathbf{M}^{inc}$  与等效电流  $\mathbf{J}^{inc}$  可以投影到基函数组  $\{f_i^l\}$  上, 得到二组向量  $\mathbf{m}_p^{inc}$  和  $\mathbf{j}_p^{inc}$ , 其分量如式(8)所示. 其中脚标  $p$  表示对应的分量值是入射场在基函数组上的投影值,  $\langle \cdot, \cdot \rangle$  表示向量场的内积.

$$\begin{cases} [\mathbf{m}_p^{\text{inc}}]_i = \langle \mathbf{M}^{\text{inc}}, \mathbf{f}_i^{\text{d}} \rangle \\ [\mathbf{j}_p^{\text{inc}}]_i = \langle \mathbf{J}^{\text{inc}}, \mathbf{f}_i^{\text{d}} \rangle \end{cases} \quad (8)$$

$\mathbf{M}^{\text{inc}}$  和  $\mathbf{J}^{\text{inc}}$  也可以用基函数组  $\{\mathbf{f}_i^{\text{d}}\}$  展开,得到二组向量  $\mathbf{m}_c^{\text{inc}}$  和  $\mathbf{j}_c^{\text{inc}}$  向量,如式(9)所示.其中脚标 c 表示对应的分量值是入射场用基函数组的展开系数.

$$\begin{cases} \mathbf{M}^{\text{inc}} = \sum_i [\mathbf{m}_c^{\text{inc}}]_i \mathbf{f}_i^{\text{d}} \\ \mathbf{J}^{\text{inc}} = \sum_i [\mathbf{j}_c^{\text{inc}}]_i \mathbf{f}_i^{\text{d}} \end{cases} \quad (9)$$

$\mathbf{m}_c^{\text{inc}}$  和  $\mathbf{j}_c^{\text{inc}}$  以及  $\mathbf{j}_c^{\text{inc}}$  和  $\mathbf{m}_p^{\text{inc}}$  可以通过  $\{\mathbf{f}_i^{\text{d}}\}$  的 Gram 矩阵<sup>[5]</sup>  $\mathbf{U}_1$  进行转换,以  $\mathbf{m}_c^{\text{inc}}$  和  $\mathbf{m}_p^{\text{inc}}$  为例,转换过程如式(10)、式(11)所示.

$$[\mathbf{m}_p^{\text{inc}}]_i = \sum_j \sum_l \langle \mathbf{f}_i^{\text{d}}, \mathbf{f}_j^{\text{d}} \rangle [\mathbf{m}_c^{\text{inc}}]_j \Leftrightarrow \mathbf{m}_p^{\text{inc}} = \mathbf{U}_1 \mathbf{m}_c^{\text{inc}} \quad (10)$$

$$\mathbf{m}_c^{\text{inc}} = \mathbf{U}_1^{-1} \mathbf{m}_p^{\text{inc}} \quad (11)$$

同理,可以用基函数组  $\{\mathbf{f}_i^{\text{sca}}\}$  展开散射等效源,得到相应的系数向量  $\mathbf{m}_c^{\text{sca}}$  和  $\mathbf{j}_c^{\text{sca}}$ :

$$\begin{cases} \mathbf{M}^{\text{sca}} = \sum_i [\mathbf{m}_c^{\text{sca}}]_i \mathbf{f}_i^{\text{sca}} \\ \mathbf{J}^{\text{sca}} = \sum_i [\mathbf{j}_c^{\text{sca}}]_i \mathbf{f}_i^{\text{sca}} \end{cases} \quad (12)$$

将式(6)右端的入射等效源投影到基函数组  $\{\mathbf{f}_i^{\text{sca}}\}$  上,可以得到矩阵方程:

$$\begin{bmatrix} \mathbf{m}_c^{\text{sca}} \\ \mathbf{j}_c^{\text{sca}} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_2^{-1} \\ \mathbf{U}_2^{-1} \end{bmatrix} \begin{bmatrix} -\hat{\mathbf{n}} \times \eta_0 \bar{\mathbf{L}} \\ -\hat{\mathbf{n}} \times \bar{\mathbf{K}} \end{bmatrix} [\eta_0 \bar{\mathbf{L}}_{pp}]^{-1} \cdot [\bar{\mathbf{K}} \quad \eta_0 \bar{\mathbf{L}}] \begin{bmatrix} \mathbf{U}_1^{-1} \\ \mathbf{U}_1^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{m}_p^{\text{inc}} \\ \mathbf{j}_p^{\text{inc}} \end{bmatrix} \quad (13)$$

式中:  $\mathbf{U}_1$  和  $\mathbf{U}_2$  分别为  $\{\mathbf{f}_i^{\text{d}}\}$  与  $\{\mathbf{f}_i^{\text{sca}}\}$  的 Gram 矩阵;  $\bar{\mathbf{L}}$  和  $\bar{\mathbf{K}}$  分别为  $\bar{\mathbf{L}}$  和  $\bar{\mathbf{K}}$  算子离散化后对应的矩阵.基函数可以采用 RWG(Rao-Wilton-Glisson)函数<sup>[17]</sup>、Rooftop 函数<sup>[18]</sup>等,测试基函数根据情况选择 RWG 或  $\hat{\mathbf{n}} \times \text{RWG}$  函数.

与直接使用 EFIE 算子  $[\eta_0 \bar{\mathbf{L}}_{pp}]^{-1}$  求解散射场相比,EPA 引入了两个额外的算子 Outside In 和 Inside Out,它们都有可能带来新的误差.下面针对这两个算子进行误差分析.本文的误差用均方根误差(Root Mean Square Error, RMSE)  $\varepsilon$  衡量,对于一组准确值为  $\{x_{i,0}\}$  的变量  $\{x_i\}$ ,其定义为

$$\varepsilon = \sqrt{\sum_{i=1}^N \frac{|x_i - x_{i,0}|^2}{|x_{i,0}|^2}} \quad (14)$$

## 2 计算结果

选取真空中半径  $r = 1 \text{ m}$  的 PEC 球,入射电磁波沿  $-x$  方向传播,频率 100 MHz,电场极化方向

为  $-z$  方向. PEC 上网格尺寸为  $0.2 \text{ m}$  ( $\approx 1/15\lambda$ ,  $\lambda$  为该频率电磁波在真空中的波长).求解普通的 EFIE 方程,该 PEC 球的后向双站雷达散射截面(Radar Cross Section, RCS)在  $\varphi = 0^\circ$  平面内的值如图 2 所示,与 Mie 散射的解析解<sup>[16]</sup> 的均方根误差为  $-13.8 \text{ dB}$ .

对于 Outside In 算子,用经过  $[\bar{\mathbf{K}} \quad \eta_0 \bar{\mathbf{L}}]$  算子得到的入射场在散射体表面与基函数的内积,和散射体表面的入射场在散射体表面与基函数的内积比较.等效面为球形,半径为  $1.2 \text{ m}$ ,网格尺寸为  $0.3 \text{ m}$  ( $\approx 1/10\lambda$ ),大于散射体的网格.其入射场与基函数内积的均方根误差为  $-19.1 \text{ dB}$ ,小于原始矩量法(Method of Moment, MoM)的误差,可以接受.

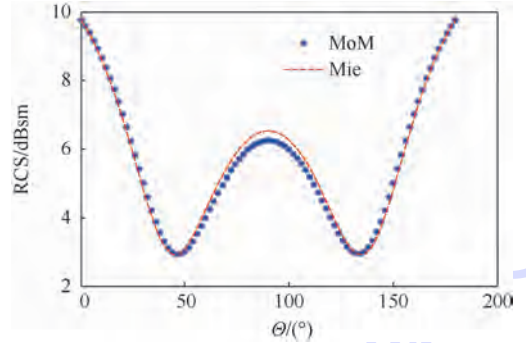


图 2 PEC 球的后向 RCS

Fig. 2 Backward RCS of PEC sphere

对于 Inside Out 算子,比较

$[-\hat{\mathbf{n}} \times \eta_0 \bar{\mathbf{L}} \quad -\hat{\mathbf{n}} \times \bar{\mathbf{K}}]^T$  中的  $\hat{\mathbf{n}} \times \bar{\mathbf{L}}$  算子在等效面上投影得到的等效散射磁流与 Mie 散射解析解的误差.其中散射体表面电流由式(15)表示,  $\{\mathbf{f}_i\}$  为 RWG 函数.分别用 RWG 函数和  $\hat{\mathbf{n}} \times \text{RWG}$  函数对 MoM 方程进行测试,如式(16)和式(17)所示.此外,从散射体向等效面辐射时也可以直接用  $\hat{\mathbf{n}} \times \bar{\mathbf{L}}$  算子对散射体表面电流作用,如式(18)所示.

$$\mathbf{J}^{\text{surf}} = \sum_i [\mathbf{j}^{\text{surf}}]_i \mathbf{f}_i \quad (15)$$

$$\mathbf{M}^{\text{es},1} \approx - \sum_i [\mathbf{U}_1^{-1} [\hat{\mathbf{n}} \times \bar{\mathbf{L}}] \mathbf{j}^{\text{surf}}]_i \mathbf{f}_i \quad (16)$$

$$\mathbf{M}^{\text{es},2} \approx - \sum_i [\mathbf{U}_2^{-1} [\hat{\mathbf{n}} \times \tilde{\mathbf{L}}] \mathbf{j}^{\text{surf}}]_i \hat{\mathbf{n}} \times \mathbf{f}_i \quad (17)$$

$$\mathbf{M}^{\text{es},3} = - \sum_i [\mathbf{j}^{\text{surf}}]_i \hat{\mathbf{n}} \times \bar{\mathbf{L}} \mathbf{f}_i \quad (18)$$

式中:

$$[\mathbf{U}_1]_{ij} = \langle \mathbf{f}_i, \mathbf{f}_j \rangle \quad (19)$$

$$[\mathbf{U}_2]_{ij} = \langle \hat{\mathbf{n}} \times \mathbf{f}_i, \hat{\mathbf{n}} \times \mathbf{f}_j \rangle \quad (20)$$

$$[\hat{\mathbf{n}} \times \tilde{\mathbf{L}}]_{ij} = \langle \hat{\mathbf{n}} \times \mathbf{f}_i, \hat{\mathbf{n}} \times \bar{\mathbf{L}} \mathbf{f}_j \rangle \quad (21)$$

$$[\mathbf{L}]_{ij} = \langle \mathbf{f}_i, \bar{\mathbf{L}} \mathbf{f}_j \rangle \quad (22)$$

此外还采用 Mie 散射近场计算公式计算了从散射体辐射至等效面的等效磁流,用作等效表面磁流的精确值.下面针对不同形状的等效面计算相应的误差情况.

首先,使用立方体形状的等效面,边长从 2.4 m 连续变化至 6 m,网格尺寸不变,均为 0.3 m ( $\approx 1/10\lambda$ ).均方根误差结果如图 3 所示,图例中的各方法说明如表 1 所示.可见随着等效面尺寸的增加,方法 1 的均方根误差减小,方法 2 和方法 3 的均方根误差基本不变,且使用 RWG 函数测试  $\hat{n} \times L$ (方法 1)效果较方法 2 好.使用方法 1 后等效磁流误差值在立方体等效面上的分布如图 4 所示,可见靠近棱边的部位误差较大.

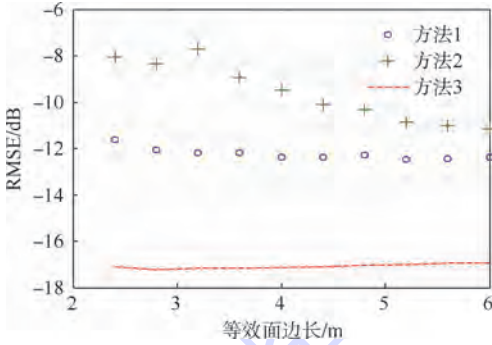


图 3 不同算法下不同尺寸立方体等效面的均方根误差  
Fig. 3 RMSE of cubic equivalence surfaces with different sizes under different methods

表 1 图例说明

Table 1 Illustration of the legend

图例	对应的计算方法
方法 1	式(16)
方法 2	式(17)
方法 3	式(18)

于磁场、电场和等效面的法向向量.散射体的散射场是连续的,但在立方体等效面的棱边附近,等效面表面的法向向量不连续,导致在棱边两侧的等效电流和等效磁流不连续,如图 6 所示.但在球形等效面上不存在这样的问题.

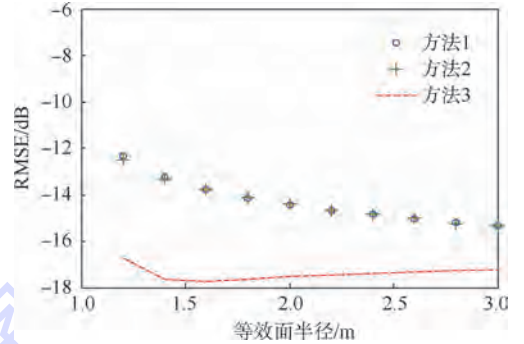


图 5 不同算法下不同尺寸球形等效面的均方根误差  
Fig. 5 RMSE of spherical equivalence surface with different sizes under different methods

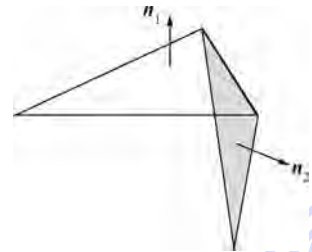


图 6 一条棱边上法向向量不连续的示意图  
Fig. 6 Illustration of discontinuity of normal vector on an edge

球形等效面不同位置的等效磁流误差值分布如图 7 所示.与图 4 相比,球形等效面上的等效磁流计算误差的分布相对均匀,主要原因是球形等效面没有棱边,不会出现立方体等效面情况下误差集中在棱边附近的情况.

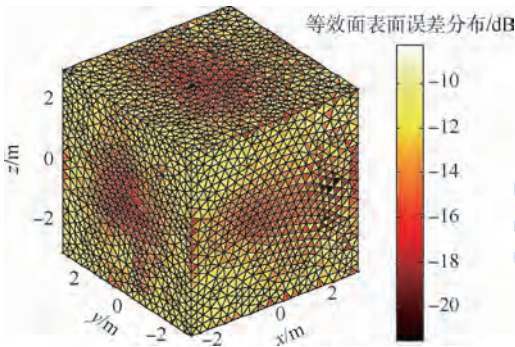


图 4 立方体等效面误差分布

Fig. 4 Distribution of error on a cubic equivalence surface

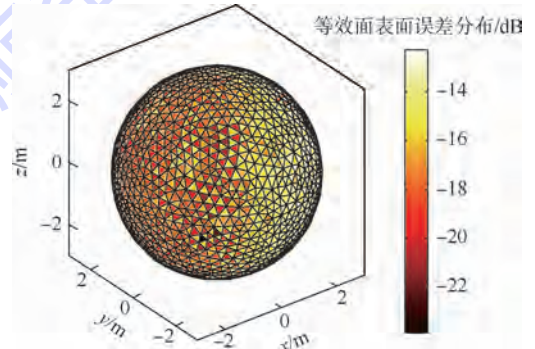


图 7 球形等效面误差分布

Fig. 7 Distribution of error on a spherical equivalence surface

将等效面换为球形,直径也从 2.4 m 扫描至 6 m,结果如图 5 所示.与图 3 相比,球形等效面上使用 RWG 和  $\hat{n} \times RWG$  函数对  $\hat{n} \times \bar{L}$  算子进行测试的效果基本相同,且随着等效面尺寸的增加而下降.主要原因是等效电流和等效磁流同时取决

将立方体等效面的棱边做平滑处理,把立方体的棱边用光滑柱面替代、顶点用球面替代,如图 8 所示,该图对应的立方体边长为 6 m,柱面和

球面的半径为 1.5 m. 随着等效面尺寸增大,均方根误差结果如图 9 所示,误差分布如图 8 所示. 与立方体等效面的结果相比,误差有所改善,方法 1 基本达到了球形等效面的误差水平,但方法 2 没有达到. 主要原因是该立方体仍然存在不完全光滑的边,而且  $\hat{n} \times \text{RWG}$  函数不是散度共形函数,不适合表示不光滑的等效面上的等效磁流.

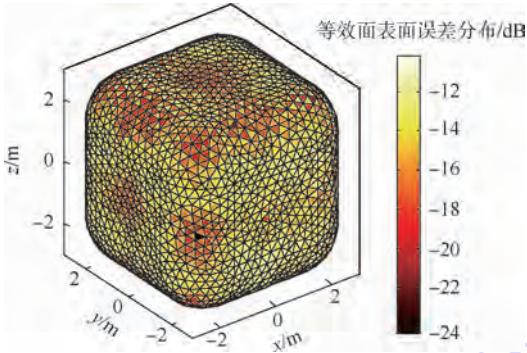


图 8 平滑立方体等效面的误差分布

Fig. 8 Distribution of error on a smooth cubic equivalence surface

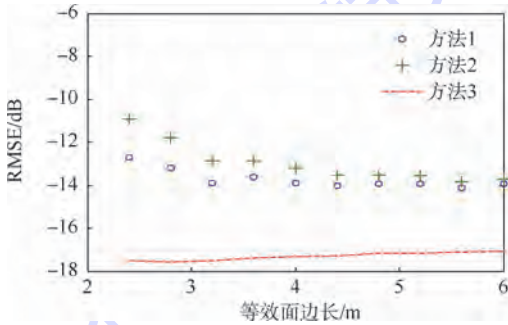


图 9 不同算法下不同尺寸平滑立方体等效面的均方根误差

Fig. 9 RMSE of a smooth cubic equivalence surface with different sizes under different methods

以上讨论了 Inside Out 算子的近场的误差,下文将根据方法 1 得到的等效面表面等效散射磁流与等效散射电流,进而计算  $\varphi = 0^\circ$  平面的 RCS. 不同形状的等效面其结果如图 10 所示. 它们与直

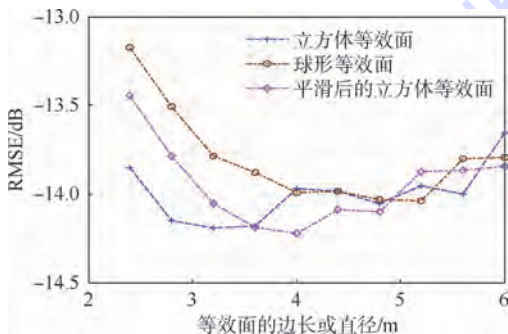


图 10 由不同等效面计算 RCS 的均方根误差

Fig. 10 RMSE of RCS calculated from different equivalence surfaces

接由 EFIE 得到的 RCS 的精度基本相同,说明等效面可以通过等效面上较为稀疏的网格得到与传统 EFIE 方法相同的精度,可以有效减小散射体之间的相互作用的计算量.

### 3 结论

通过对不同等效面的等效磁流和 RCS 的比较,可以得出如下结论:

1) 在相同的网格尺寸下,等效面尺寸越大,表面等效磁流的计算误差越小.

2) 等效面的棱边会增加表面等效磁流的误差. 对立方体等效面适当的平滑处理可以减少误差,但难以达到球形等效面的误差水平.

3)  $\hat{n} \times \text{RWG}$  函数不是散度共形函数,不适合用作对等效磁流的展开.

4) 等效面可以通过更为稀疏的网格代表散射体向外辐射作用,同时保证和原先基本相同的精度,适合用作求解多尺度大规模散射问题.

实际选取等效面时除了考虑减小其表面磁流的误差外,还需要考虑一些限制. 如对于天线阵列,选取立方体的等效面可以使两个相接触的表面的等效源相消<sup>[7]</sup>,可以减少计算量,而球形等效面不具有这样的性质.

### 参考文献 (References)

- [1] Andriulli F P, Cools K, Bagefi H, et al. A multiplicative Calderon preconditioner for the electric field integral equation [J]. IEEE Transactions on Antennas and Propagation, 2008, 56 (8): 2398-2412.
- [2] Heldring A, Rius J M, Tamayo J M, et al. Multiscale compressed block decomposition for fast direct solution of method of moments linear system [J]. IEEE Transactions on Antennas and Propagation, 2011, 59 (2): 526-536.
- [3] Li Y J, Jin J M. A new dual-primal domain decomposition approach for finite element simulation of 3-D large-scale electromagnetic problems [J]. IEEE Transactions on Antennas and Propagation, 2007, 55 (10): 2803-2810.
- [4] Peng Z, Wang X C, Lee J F. Integral equation based domain decomposition method for solving electromagnetic wave scattering from non-penetrable objects [J]. IEEE Transactions on Antennas and Propagation, 2011, 59 (9): 3328-3338.
- [5] Li M K, Chew W C, Jiang L J. A domain decomposition scheme based on equivalence theorem [J]. Microwave and Optical Technology Letters, 2006, 48 (9): 1853-1857.
- [6] Li M K, Chew W C. Wave-field interaction with complex structures using equivalence principle algorithm [J]. IEEE Transactions on Antennas and Propagation, 2007, 55 (1): 130-138.
- [7] Li M K, Chew W C. Multiscale simulation of complex structures using equivalence principle algorithm with high-order field point

- sampling scheme[J]. IEEE Transactions on Antennas and Propagation, 2008, 56(8):2389-2397.
- [8] Ylä-Oijala P, Taskinen M. Electromagnetic scattering by large and complex structures with surface equivalence principle algorithm[J]. Waves in Random and Complex Media, 2009, 19(1): 105-125.
- [9] Shao H, Hu J, Guo H, et al. Fast simulation of array structures using T-EPA with hierarchical LU decomposition[J]. IEEE Antennas and Wireless Propagation Letters, 2012, 11:1560-1563.
- [10] Sun L E, Chew W C, Jin J M. Augmented equivalence principle algorithm at low frequencies[J]. Microwave and Optical Technology Letters, 2010, 52(10):2274-2279.
- [11] Qian Z G, Chew W C. Fast full-wave surface integral equation solver for multiscale structure modeling[J]. IEEE Transactions on Antennas and Propagation, 2009, 57(11):3594-3601.
- [12] Zhang K, Ouyang J, Yang F, et al. Radiation analysis of large antenna array by using periodic equivalence principle algorithm [J]. Progress in Electromagnetics Research, 2013, 136:43-59.
- [13] Shi Y, Wang J, Liang C H. A time-domain equivalence principle and its marching-on-in-degree solution[J]. Microwave and Optical Technology Letters, 2014, 56(10):2415-2422.
- [14] 杨晨. 基于等效原理的区域分解算法[D]. 南京:南京理工大学, 2013.
- Yang C. Domain decomposition method based on equivalence principle [D]. Nanjing: Nanjing University of Science and Technology, 2013 (in Chinese).
- [15] 张博. 等效原理算法及其结合快速算法分析目标的电磁散射[D]. 西安:西安电子科技大学, 2013.
- Zhang B. Scattering analysis of target using equivalent principle algorithm and its fast algorithm[D]. Xi'an: Xidian University, 2013 (in Chinese).
- [16] Chew W C. Waves and fields in inhomogeneous media [M]. Piscataway, NJ: IEEE Press, 1995:29, 186.
- [17] Rao S M, Wilton D, Glisson A W. Electromagnetic scattering by surfaces of arbitrary shape[J]. IEEE Transactions on Antennas and Propagation, 1982, 30(3):409-418.
- [18] Jin J M. Theory and computation of electromagnetic fields[M]. Hoboken, New Jersey: John Wiley & Sons, 2011:439-440.

## Error analysis of equivalence principle algorithm on different equivalence surfaces

DANG Xunwang<sup>1</sup>, LI Maokun<sup>\*1,2</sup>, YANG Fan<sup>1,2</sup>, XU Shenheng<sup>1,2</sup>

(1. Department of Electronic Engineering, Tsinghua University, Beijing 100084, China;

2. Tsinghua National Laboratory for Information Science and Technology(TNist), Beijing 100084, China)

**Abstract:** Equivalence principle algorithm (EPA) is a domain decomposition algorithm based on the Huygens' equivalence principle to solve large scale scattering problems. The errors of the equivalence principle operator (EPO) in EPA using cubic, spherical and smooth cubic equivalence surfaces were analyzed to find the source of errors. The sources of the error were investigated and a simple scheme to improve its accuracy was developed. From the numerical examples, it shows that the error of equivalent magnetic current is mainly concentrated on geometrical discontinuity of the equivalence surface, and affected by both the discontinuity of the normal vector on the equivalence surface and the choice of basis functions. Therefore it is suggested to choose smooth equivalence surfaces if possible rather than to use cubic ones directly.

**Key words:** Huygens equivalence principle; domain decomposition method; error analysis; computational electromagnetics; RWG basis function

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0220



# 飞行器雷达隐身测试评估技术及发展

肖志河<sup>\*1</sup>, 高超<sup>1,2</sup>, 白杨<sup>1,2</sup>, 袁晓峰<sup>1</sup>

(1. 电磁散射重点实验室, 北京 100854; 2. 中国传媒大学 信息工程学院, 北京 100024)

**摘 要:** 全面隐身化, 隐身性能指标成为装备最重要的战技指标之一. 隐身测试评估是实现装备隐身性能的关键指标和不可缺少的重要环节, 涉及隐身装备从方案设计验证、研制方案筛选、部件和整机隐身效果评估、使用维护检定等全生命周期. 综述了飞行器雷达隐身性能试验测试体系和测试评估方法, 介绍了国内外典型的测试场及测量技术进展, 详细分析了低散射诊断实验测试技术的最新成果, 展望了飞行器隐身电磁特性测试技术未来发展趋势.

**关键词:** 雷达散射截面; 隐身飞行器; 测试评估; 低散射测量; 雷达成像

**中图分类号:** V221+.3; TB553

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1873-07

武器系统隐身化是当今战场电磁对抗的主流方向之一, 良好的低可探测性能成为先进隐身装备的标志性特征<sup>[1-2]</sup>. 自 20 世纪 80 年代以来, F-117A、B-2A 参加实战, F-22A 已装备军队, F-35、X-47B 即将完成研制形成作战能力. 隐身目标的核心是低 RCS、低可探测性, 目标 RCS 减缩 10 dB, 雷达威力下降约 44%; 目标 RCS 减缩 20 dB, 雷达威力下降约 68%; 目标 RCS 减缩 30 dB, 雷达威力下降约 82%. 飞行器隐身性能对雷达和整个武器系统作战效能产生致命影响. 隐身技术结合诱饵技术是提高弹道导弹突防能力的有效手段. 减小反导系统 (BMD) 探测器的作用距离, 推迟 BMD 反应时间; 增加 BMD 从诱饵中识别弹头难度, 降低拦截成功概率; 降低对有源诱饵及干扰装置的功率需求 (例如弹头 RCS 减小 10 dB, 要求的干扰功率也减小 10 倍), 提高了突防干扰的有效性. 随着隐身技术的发展, “雷达-目标”之间的探测与反探测对抗, 发展成为了新的隐身与反隐身技术的对抗.

雷达隐身技术途径主要有: 低 RCS 外形技

术<sup>[3]</sup>、吸波材料涂覆技术<sup>[4]</sup>、雷达吸波结构技术<sup>[5-6]</sup>、无源与有源阻抗加载技术、等离子隐身技术等<sup>[7-8]</sup>.

研究飞行器雷达隐身特性的方法主要包括理论建模分析和试验测试<sup>[9-10]</sup>. 当今的雷达系统面对的目标外形更加复杂, 尤其是越来越多的经过隐身性能设计的飞行器散射量级低, 并且包含大量的非金属结构, 这就加大了理论计算的研究难度, 也使得测试的方法成为获取目标电磁散射辐射特征数据的主要手段.

武器装备全面隐身化, 隐身性能指标成为装备最重要的战技指标之一<sup>[11-12]</sup>. 隐身测试评估是实现装备隐身性能不可缺少的重要环节, 涉及隐身装备从方案设计验证、研制方案筛选、部件和整机隐身效果评估、使用维护检定等全生命周期. 隐身性能评估测试已发展成为一项专业技术, 在隐身装备研制中发挥重要作用. 美、俄、欧等军事强国为促进装备隐身技术发展, 建立了强大的隐身技术研究与试验体系, 全面支撑了隐身装备研制发展<sup>[13-14]</sup>; 高度重视隐身试验测试评估技术, 建

收稿日期: 2015-04-10; 录用日期: 2015-05-14; 网络出版时间: 2015-06-29 17:08

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150629.1708.001.html

\* 通讯作者: 肖志河 (1966—), 男, 江西吉安人, 研究员, xiaoh66479@vip.sina.com, 主要研究方向为电磁散射与特征应用研究.

**引用格式:** 肖志河, 高超, 白杨, 等. 飞行器雷达隐身测试评估技术及发展[J]. 北京航空航天大学学报, 2015, 41(10): 1873-1879.  
Xiao Z H, Gao C, Bai Y, et al. Aircraft radar stealth test and evaluation technology and progress[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1873-1879 (in Chinese).

立了国立雷达目标散射特性测试场,用以飞行器隐身性能测试评估<sup>[15]</sup>. 本文介绍了飞行器雷达隐身测试评估技术,综述了飞行器雷达隐身性能试验测试体系和测试评估方法,介绍了国内外典型的测试场及测量技术进展,研究了低散射诊断试验测试技术,分析展望了电磁特性测试技术未来发展趋势.

## 1 测试评估体系

目标电磁特性试验测试研究中采用的各类手段和要求构成了目标特性测试的试验体系,其中的要素包含测试场地和平台、测试目标和对象、测试方式和参量以及测试技术规范. 在进行某种目标的具体测试时,需要从上述要素的角度,考虑目标的本体、目标姿态、背景环境和目标特征信息空间的需要,根据实际情况选择测试平台和具体实施技术.

### 1.1 测试场地

测试场地是目标电磁散射试验测试的基础条件,场地的类型决定了其适合进行的目标试验测试类型. 测试场的分类方法很多,按场地建设方式可以分为室内测试场和室外测试场,按照目标状态和设备搭载平台类型可以分为静态测试场和动态测试场,按照测试电磁波属性可以分为近场测试场和远场测试场. 所有这些测试场绝大部分都是为了获得目标区域的平面波照射,以便测得正确的 RCS 值.

室内测试场通常是以微波暗室的形式实现,诸如室内锥形暗室、球面波微波暗室、近场非消波测试场、室内紧凑场、近场扫描测试场等. 室内测试场受自然环境影响小,测试精度高,保密条件较好. 但受到场地尺寸和配套设备技术水平的影响,室内测试场对被测目标在各方面的限制相对较多<sup>[16]</sup>.

室外测试场是在室外建设的测试场地,常见的有地平测试场、倒 V 形测试场、高架测试场、近场扫描测试场、动态测试场等. 相对于室内测试场而言,室外测试场能适应更多尺寸类型目标的测试,某些测试过程中还引入了目标真实使用环境的背景条件,真实度较高. 但室外测试场受自然条件影响较大,保密条件相对不好.

### 1.2 测试设备

不论是哪种测试场,除场地之外,最重要的是有完整配套的测试设备. 为了获取目标的电磁散射特性,测试设备至少应该包含以下 5 个部分:

1) 电磁波收发系统,用于发射和接收所需的

电磁波信号.

2) 目标姿态控制系统,用于在测试过程中调整目标的姿态. 如室内测试场中常用到的转台、低散射支架等设施. 在室外动态测试场中这部分功能由目标与测试仪器间的相对运动实现.

3) 目标,包括内测目标和系统标校目标. 常见的系统标校目标有金属球、金属平板、低散射载体、有源标校器、标校卫星和自然天体等.

4) 满足噪声要求的测试环境,使得噪声、杂波对所需测试信号的干扰最小的环境.

5) 控制和数据处理系统,能够多渠道的采集和分析目标特征信息.

### 1.3 试验测试的研究对象

目标电磁散射特性试验测试技术的研究对象既包括具体装备、部件和形体这类狭义的目标,也包括雷达作战背景环境和气象环境的电磁散射特征,内容涉及“海、陆、空、天、电”5 个维度中的各类目标与环境. 主要归结为以下几大类:

第 1 类,测试研究己方军事目标特性,目的在于提升己方装备的生存能力并支持其进一步发展.

第 2 类,测试研究国外的军事目标特性,目的在于对潜在作战对象特性信息的理解掌握.

第 3 类,测试研究复杂战场环境的特性,用于解决作战环境掌握和装备环境相匹配方面的技术问题.

第 4 类,测试研究用于装备研制试验过程中所需要的靶机、靶标的电磁散射特性.

目标电磁散射特性试验测试还作为武器装备隐身突防性能测评的手段,在装备研制过程中占据着重要地位.

### 1.4 试验测试标准

目标电磁散射特性试验测试技术具有很明显的专业性特征,需要有一套专门的标准体系来规范测试技术的应用与发展. 电磁散射特性测试标准具体说明了采用各类先进测试平台、测试技术获取目标特性各类数据所需要遵循的测试规范. 其中最常用的包括:《室内缩比目标雷达散射截面测试方法》、《室外场雷达目标散射特性测试方法》、《空中目标雷达散射截面动态测试方法》等标准规范.

在目标特性测试技术规范方面,其他国家也已进行了大量相关工作. 特别是美国,由美军方、国家标准与技术研究所、IEEE 等通力合作制定了相关的标准与通则,形成了《IEEE RCS 测试程序推荐实施通则》、《美国国防部 RCS 测试场认

证标准》、《RCS 测试质量保证手册》等,这些文件建立了 RCS 测试场的质量认证体系。

### 1.5 隐身测试评估应用体系

飞行器隐身性能的测试评估,主要利用室内微波暗室(紧缩场)、室外静态测试场、室外动态

试验设施 3 种试验平台,分别完成隐身飞行器从概念设计到装备部队全寿命周期的隐身性能测评工作,不同阶段需要不同的测量手段和设备来支撑,从而形成了一个完整的研制生命周期如图 1 所示。

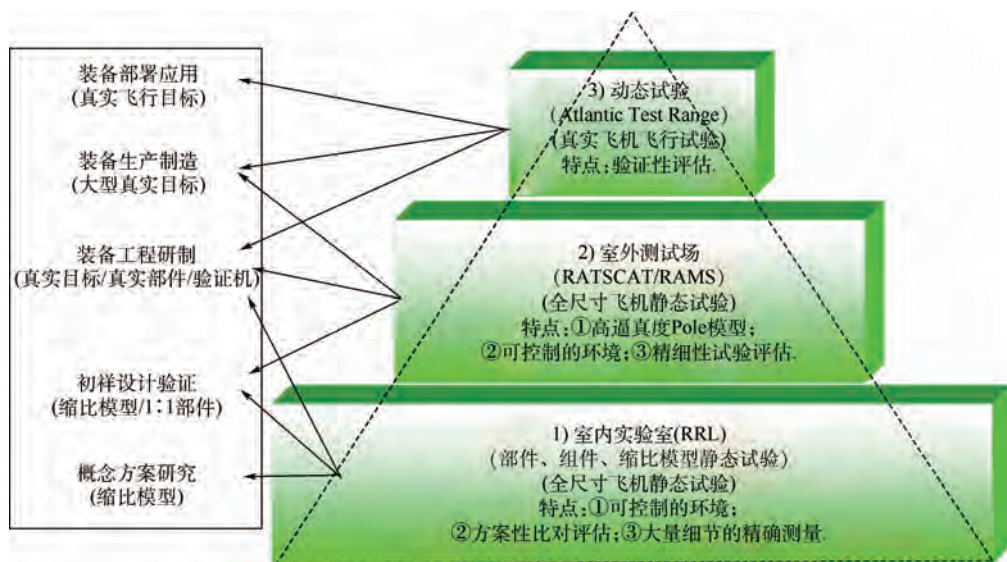


图 1 隐身性能试验评估应用体系

Fig. 1 Stealth performance test evaluation and application system

装备研制过程中,为了能更好地对装备目标特性进行设计和评估,在研制的不同阶段需要借助不同的测试手段对目标的散射性能加以测评。在装备的概念研究、设计验证和装备研制阶段,常需要验证一些设计样件、模型和零部件,室内静态测试实验室由于环境可控制度好,测量精确,比较适合此阶段的使用。当研制进入设计试生产阶段,对于全尺寸结构和模型的测试在精度相对较高的室外全尺寸静态测试场进行就会比较有效。当装备进入了生产应用阶段,室外动态试验设施将发挥对装备在真实应用环境中的评估与测试作用。如此根据各阶段的特点,测试评估体系可以全程准确支撑装备的研发。

## 2 国外测试场技术与发展

### 2.1 室外全尺寸散射特性测量

室外静态场主要用于全尺寸目标包括实际武器装备的隐身性能的检测,它通过在室外搭建测量转台和轨道,形成大距离范围的测量系统,以此来满足散射特性测试所需的远场条件。

为了进行飞行器隐身性能评估和电磁散射特性测量,美国在全国建立了几十个骨干实验室、研究中心和测试基地<sup>[17]</sup>,技术已经十分先进。最具代表性的是建立于 1963 年的美国国立雷达目标

散射特性测试场(RATSCAT),目前有主测试场地和先进测试系统(RAMS)两个外场、静态雷达特性测试场地(见图 2)。其他军工企业测试场还包括:洛克希德-马丁公司海伦达尔室外测试场、波音公司波德曼室外测试场(见图 3)。

室外静态测试场历史长、直观、技术较成熟,是衡量武器系统低可探测设计效果的尺码,由于



图 2 美国国立测试场(主测试场和先进测试系统)

Fig. 2 USA National RCS test field (main test area and advanced measurement system)



图 3 洛克希德-马丁公司与波音公司室外测试场

Fig. 3 Lockheed-Martin and Boeing outdoor test range

非金属结构材料大量用于军用飞行器等武器系统的设计制造. 图 4 所示为外军典型隐身飞行器外场测试图例.

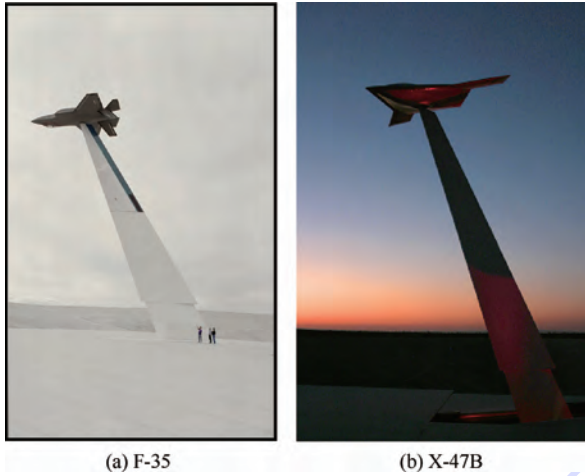


图 4 外军典型隐身飞行器外场测试图例  
Fig. 4 Schematic of typical foreign military stealth aircraft field testing

### 2.2 紧缩场散射特性测量

紧缩场通过在微波暗室内采用精密的反射面, 将点源产生的球面波在近距离内转换为平面波, 从而实现散射特性测量要求的单一平面波照射和单一平面波接收的条件(见图 5). 典型的室内紧缩场测量系统如林肯实验室的紧缩场(见图 6), 该紧缩场可以测量尺寸为 3.6 m 的目标, 波段为 0.4 ~ 100 GHz.

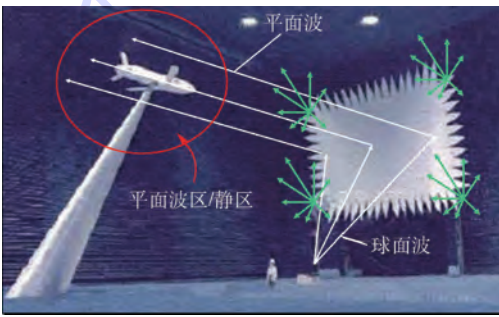


图 5 室内紧缩场测量示意图  
Fig. 5 Schematic of indoor compact range measurement

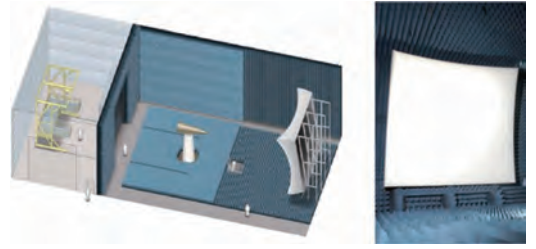


图 6 林肯实验室紧缩场及其反射面  
Fig. 6 Lincoln laboratory compact range and its reflector

紧缩场占地小、效率高, 不仅可以高精度地检测目标的 RCS, 提取其点频信号、频响特性、多维高分辨率成像, 而且可以完成材料反射率及其电参数的测量, 所有这些都是完成目标低可探特性研究必不可少的, 随着紧缩场设计和制造趋于成熟, 需求更加迫切, 超大型紧缩场将得到发展, 国外的波音公司已研制出静区达 12 m 的超大紧缩场.

### 2.3 近场扫描散射特性测量

近年来发展的室内近场测量技术, 采用近场变换测量与远场测试相比较的方法进行产品隐身性能的出场验收及使用维护后的检定测试. 该技术效率高、成本低, 逐渐成为散射测量领域关心的热点问题, 各国都在不断加强与关注其理论与技术的发展.

近年来, 近场散射特性测量在目标特性领域受到越来越多的重视. 其中, 美国、俄罗斯、欧洲各国、印度、南非等都加快了近场散射特性测量研究, 并已开发出相应的近场测量设备, 使近场测量真正走向了实用化. 典型的近场测量系统是美国洛克希德-马丁公司的 ATF (Acceptance Test Facility), 该装置通过在有限的距离内测得的近场数据外推出其远场散射特性, 在 F-22 战机和 F-35 战机的研制与出厂验收(见图 7)中, 隐身效果的验证均是利用该测量系统来完成的. 该测量系统效率高, 与相关的硬件相配合, 现在已经实现在一天内完成 F-35 隐身性能的初步验证.

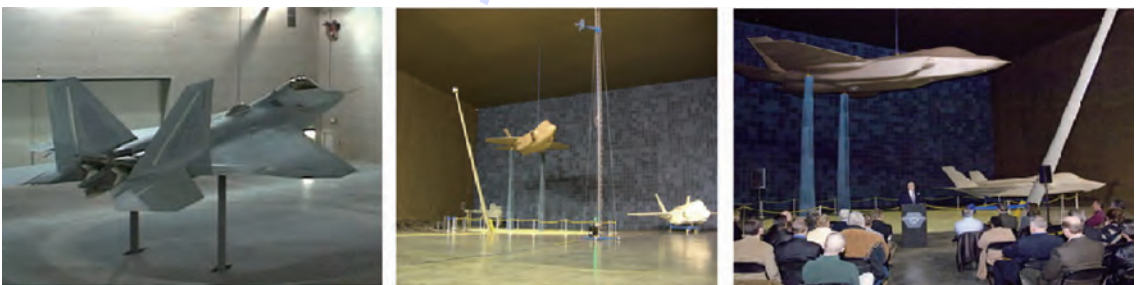


图 7 F22、F35 近场测量示意图  
Fig. 7 Schematic of F22, F35 near field measurement

### 3 低散射诊断测试技术

随着隐身技术的发展与日趋成熟,隐身化已逐步成为飞行器发展的重要需求与特征,各种飞行器都提出了隐身化设计需求,它通常要求飞行器具有很低的电磁散射特征信号以此来降低被雷达发现的概率.对于飞行器这类大型复杂复合材料目标,利用现有的手段计算其散射特性非常困难.因此面向隐身装备的低散射测试技术,特别是低散射高分辨成像测试技术是分析诊断强散射源、指导隐身飞行器研制、评估飞行器隐身性能等的最有效、快捷、准确的手段<sup>[18]</sup>.

#### 3.1 低散射测试技术

随着装备越来越注重隐身性能的设计,电磁散射测试所面对的目标散射截面越来越小,各种闪烁噪声体现的也越明显,这就对测试场和测试系统都提出了更高的要求.在 RCS 测试中,需要保证测试场背景回波水平比目标散射水平低 20 dB,才能使测试结果不确定度达到 1 dB,所以低散射测试技术主要解决的就是场地背景回波的问题.为了达到这个目的,测试中通常要使用专门设计的低散射支架来支撑目标,采用硬件脉冲距离选通、软件波门距离选通和矢量背景对消等方法.通过上述方法可以实现室内测试场有效目标 RCS 测试值达到  $10^{-4} \text{ m}^2$  量级,室外测试场达到  $10^{-3} \text{ m}^2$  量级.随着技术的发展,目标支撑结构的进步也为降低测试场地背景提供了更多的方法<sup>[19]</sup>.除了采用更好的低散射外形设计和更高水平的表面吸波处理,带有目标平动功能的新型支撑系统配合后期信号处理,可以将支撑结构本身对场地背景和测试结果的影响进一步消除.

#### 3.2 宽带幅相精密测试技术

宽带幅相精密测试技术是指采用高稳定度全相参测试系统,配合精密目标姿态控制设备对目标进行宽带电磁散射特性测试的技术<sup>[20]</sup>.测试过程中,要使用精密定标体对测试系统和环境进行精确的幅相标定.经过标定后,系统对目标散射场相位信息的测量精度可以达到优于 2 dB 的水平,进一步可以获得目标的宽带幅相精密测量数据.这些数据可以为目标高分辨处理、高精度目标识别方法研究等工作提供丰富的目标特征信息.

#### 3.3 高分辨成像及散射中心诊断技术

高分辨成像及散射中心诊断技术是建立在目

标宽带特征和目标多散射中心理论基础上的,通过对目标的大带宽频率阶跃和多孔径角合成的方法,可以得到目标的高分辨图像,对目标散射中心分布情况进行分析诊断<sup>[21-22]</sup>.

在高分辨成像时,测试结果的分辨率是由测试带宽和测试孔径张角决定的.以二维高分辨为例,成像的径向距离分辨率为  $\delta_y = \frac{c}{2B}$ ,横向距离分辨率为  $\delta_x = \frac{\lambda}{2\Delta\theta}$ ,其中  $c$  为光速, $B$  为测试带宽, $\lambda$  为波长, $\Delta\theta$  为测试孔径角,如图 8 所示,图中  $xOy$  为目标坐标系, $uOv$  为目标坐标系旋转  $\theta$  角度后的新坐标系, $(r, \varphi)$  为目标在  $uOv$  坐标系下的极坐标, $R_0$  为目标旋转中心与发射接收天线连线的距离, $2d$  为发射天线与接收天线的距离, $R_t$  与  $R_r$  分别为发射天线、接收天线与目标上散射中心  $(x, y)$  的距离, $l_e$  为散射中心  $(x, y)$  在旋转过程中的投影线.

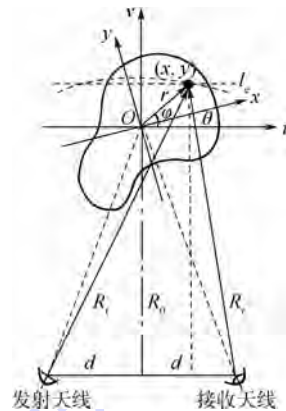


图 8 微波成像几何关系图

Fig. 8 Relation schematic of microwave imaging geometry

图 9 所示为雷达目标的二维、三维高分辨成像图例.高分辨成像测试为雷达目标散射机理分析及强散射源诊断提供了十分重要的方法,是进行隐身目标检测评估的主要手段.

#### 3.4 隐身材料性能评估测试技术

使用隐身材料是现代装备隐身设计中常使用的重要手段,由于隐身材料电磁参数复杂、材料成分种类繁多,对其性能进行精确理论计算分析是非常困难的,所以测试成为对隐身材料性能进行评估的最有效、最直接的手段.

隐身材料的性能评估测试主要包括材料电磁参数测试和材料反射率测试,其中反射率测试是电磁散射测试中比较重要的一部分.常用的测试方法有远场 RCS 测试法和弓形测试法.

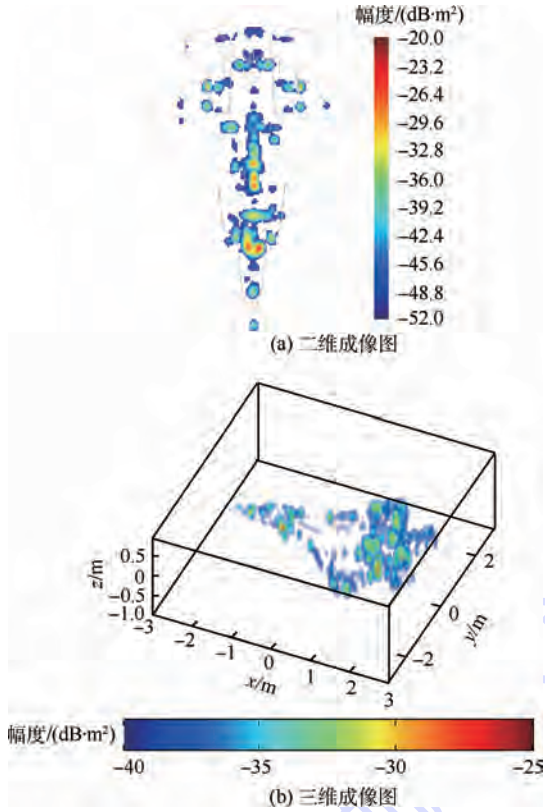


图9 目标高分辨成像及散射中心诊断  
Fig.9 Target high resolution imaging and scattering centers diagnostic

#### 4 雷达隐身测试技术发展趋势

随着装备隐身反隐身、探测反探测、识别反识别技术的对抗发展,目标电磁散射特性研究始终是一个热点,得到越来越多的关注,同时促进了目标电磁散射特性试验测试技术的发展.以美国和部分欧洲国家为代表,总结其电磁散射测试技术发展,表现出以下特点:

1) 向多波段、多极化、更宽的带宽方向发展.低频隐身特性、太赫兹散射特性需求,测试带宽覆盖0.03~2000 GHz(从VHF到THz)的频带范围.新型冲击雷达超宽带测试,全极化、厘米级超分辨率成像诊断技术.

2) 外场场地大型化,单站向双站和多站测试发展.双站/多基地反隐身需求促进了外场双站和多站测试技术发展.

3) 室内测试系统、近场扫描测试系统日益先进、完善.基于近远场传递关系的近场扫描测试极大的提高了测试的操作性,能更好地满足大型、新型装备的现场、实时评估测试需要.

4) 测试技术的规范化方向发展.统一标准的测试规范文件和测试场地的认证系统,有利于测试技术的规范化,也有利于目标特征数据的交流和应用,推动测试技术的发展.

#### 5 结论

综上所述,国外雷达隐身测试评估技术的发展主要围绕装备隐身反隐身、探测反探测、识别反识别技术与对抗发展需求,以装备隐身研制试验及装备隐身伪装突防性能评估为背景,重点突破了大宽带/多频谱测量、动平台动态跟踪测量、极低RCS测量、双/多站测量、新型隐身材料和隐身技术的试验测试等技术,拓展和完善了室内(室外)场测试系统,建立并完善了地、海、空、天的试验评估体系,形成了满足各类装备隐身突防伪装性能试验测试评估能力,支撑装备隐身技术快速发展.

#### 参考文献 (References)

- [1] 黄培康.雷达目标特征信号[M].北京:中国宇航出版社,1993:13-14.  
Huang P K. Radar target characteristic signal[M]. Beijing: China Astronautic Publishing House,1993:13-14 (in Chinese).
- [2] 黄培康,殷红成,许小剑.雷达目标特性[M].北京:电子工业出版社,2005:265-269.  
Huang P K, Yin H C, Xu X J. Radar target characteristic[M]. Beijing: Publishing House of Electronics Industry,2005:265-269 (in Chinese).
- [3] 桑建华.飞行器隐身技术[M].北京:航空工业出版社,2013:182-196.  
Sang J H. Low-observable technologies of aircraft[M]. Beijing: Aviation Industry Publishing House,2013:182-196 (in Chinese).
- [4] 陈文.隐身飞机的维修经验[J].国际航空,1990,12(3):58-60.  
Chen W. Stealth aircraft maintenance international aviation[J]. International Aviation,1990,12(3):58-60 (in Chinese).
- [5] 许占显.隐身飞机的质量探测与维修技术研究[J].航空维修与工程,2012,12(1):68-70.  
Xu Z X. Quality detection and maintenance technology research for stealth aircraft[J]. Aviation Maintenance & Engineering,2012,12(1):68-70 (in Chinese).
- [6] 杨亮.雷达吸波涂层损伤及修复研究进展[J].新技术新工艺,2009,12(1):97-99.  
Yang L. The repair technique and the damage of radar absorbing [J]. New Technology & New Process,2009,12(1):97-99 (in Chinese).
- [7] 张卫东,冯小云,孟秀兰.国外隐身材料研究进展[J].宇航材料工艺,2000,30(3):1-4.  
Zhang W D, Feng X Y, Meng X L. Status and development of foreign study on new stealthy materials[J]. Aerospace Materials and Technology,2000,30(3):1-4 (in Chinese).
- [8] 赵敬,杨雪.等离子体隐身技术浅议[J].航空科学技术,2001(4):34-35.  
Zhao J, Yang X. Plasma stealth technology[J]. Aeronautical Sci-

- ence & Technology, 2001(4):34-35 (in Chinese).
- [9] 张麟兮,李南京,胡楚风,等.雷达目标散射特性测试与成像诊断[M].北京:中国宇航出版社,2009:61-67.  
Zhang L X, Li N J, Hu C F, et al. Radar target scattering characteristics measurement and image diagnostic [M]. Beijing: China Astronautic Publishing House, 2009:61-67 (in Chinese).
- [10] 何国瑜,卢才成,洪家才,等.电磁散射的计算和测量[M].北京:北京航空航天大学出版社,2006:163-167.  
He G Y, Lu C C, Hong J C, et al. Calculation and measurement of electromagnetic scattering [M]. Beijing: Beihang University Press, 2006:163-167 (in Chinese).
- [11] Obaid Y. Lessons learned from the F/A-22 and F/A-18E/F development programs [M]. Santa Monica: the RAND Corporation, 2005:29-40.
- [12] 余宏明.雷达/电子战中的现代隐身技术[J].现代雷达, 2007, 29(6):25-27.  
Yu H M. Modern stealth technique in radar/EW [J]. Modern Radar, 2007, 29(6):25-27 (in Chinese).
- [13] Bill S. IDR technology and aerospace editor [J]. JSF-Hidden Problems of Stealth, 2002, 1(31):111-116.
- [14] 王瑞凤,杨宪江,张彦朴.解析武器装备的隐身问题[J].探测与控制学报, 2008, 30(增刊):77-79.  
Wang R F, Yang X J, Zhang Y P. Analysis on stealth technology of weapon equipment [J]. Journal of Detection & Control, 2008, 30(Supplement):77-79 (in Chinese).
- [15] Grant R. The radar game [M]. Virginia: IRIS Independent Research, 1998:25-27.
- [16] 张麟兮,郭静远,魏世京,等.内场 RCS 测试环境中误差的消减方法研究[J].现代电子技术, 2011, 34(13):13-15.  
Zhang L X, Guo J Y, Wei S J, et al. Investigation on reducing environment errors in indoor RCS measurement [J]. Modern Electronics Technique, 2011, 34(13):13-15 (in Chinese).
- [17] 麻连凤,桑建华,陈颖闻.隐身目标 RCS 外场测试概述 [C]//2013 航空试验测试技术学术交流会论文集.北京:中国航空学会, 2013:51-53.  
Ma L F, Sang J H, Chen Y W. Overview of stealth target test in outdoor range [C]//The Proceedings of Aviation Testing Technology in 2013. Beijing: Chinese Society of Aeronautics and Astronautics, 2013:51-53 (in Chinese).
- [18] 刘君,马瑶,渠立永,等.微波暗室低散射目标 RCS 测量方法[J].解放军理工大学学报:自然科学版, 2013, 14(1):300-303.  
Liu J, Ma Y, Qu L Y, et al. Research on radar cross section measurement method of low-scattering objects in microwave anechoic chamber [J]. Journal of PLA University of Science and Technology: Natural Science Edition, 2013, 14(1):300-303 (in Chinese).
- [19] 安大卫,李志平,陈五一.低散射目标支撑金属支架的外形参数优化[J].电讯技术, 2015, 55(3):333-335.  
An D W, Li Z P, Chen W Y. Shape optimization of target support low-scattering metal pylons [J]. Telecommunication Engineering, 2015, 55(3):333-335 (in Chinese).
- [20] 莫锦军,袁乃昌.一套先进的自动测量计算超宽带紧缩场测试系统[J].计算机测量与控制, 2002, 10(9):573-574.  
Mo J J, Yuan N C. An advanced automatic measuring computational UWB compact range system [J]. Computer Measurement & Control, 2002, 10(9):573-574 (in Chinese).
- [21] 肖志河.旋转目标微波成像的滤波-逆投影法[D].北京:航天科工集团 207 所, 1989.  
Xiao Z H. Research on CBP imaging technique for rotating targets [D]. Beijing: CASIC 207 Institute, 1989 (in Chinese).
- [22] 王保平,张麟兮,范九伦,等.新的滤波-逆投影成像算法[J].计算机工程与应用, 2009, 45(32):16-18.  
Wang B P, Zhang L X, Fan J L, et al. New filtered back-projection algorithm [J]. Computer Engineering and Applications, 2009, 45(32):16-18 (in Chinese).

## Aircraft radar stealth test and evaluation technology and progress

XIAO Zhihe<sup>\*1</sup>, GAO Chao<sup>1,2</sup>, BAI Yang<sup>1,2</sup>, YUAN Xiaofeng<sup>1</sup>

(1. Science and Technology on Electromagnetic Scattering Laboratory, Beijing 100854, China;

2. School of Information Engineering, Communication University of China, Beijing 100024, China)

**Abstract:** Weapon equipment stealth is comprehensive. Stealth performance becomes one of the most important combat skills. Stealth testing and evaluation is an important indicator and indispensable link to realize the equipment stealth performance, involving stealth equipment from the design verification, design scheme selection, parts and machine maintenance stealth effect evaluation, maintenance and other life-cycle test. Aircraft radar stealth performance measurement system was reviewed and the evaluation method, the typical test range and measurement technology progress at home and abroad were introduced, the latest low scattering diagnostic test technology was studied and the analysis of electromagnetic characteristics testing technology for the future development trend was also forecasting.

**Key words:** radar cross section; stealth aircraft; test and evaluation; low scattering measurement; radar image

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0231

# 毫米波-太赫兹电子回旋器件的几个基础问题

杜朝海<sup>1</sup>, 罗里<sup>2,3</sup>, 刘濮鲲<sup>\*1</sup>

(1. 北京大学 信息科学技术学院, 北京 100871; 2. 中国科学院 电子学研究所, 北京 100190;

3. 中国科学院大学 电子电气与通信工程学院, 北京 100049)



**摘 要:** 探索和发展更高频率和更高功率水平的电磁波源是电子器件长期以来的重要发展方向. 本文介绍了基于电子回旋脉塞原理发展起来的电子回旋器件, 该类器件在毫米波-太赫兹波段具有高功率的优势. 系统探讨了电子回旋器件所面临的欧姆损耗、模式竞争以及对强磁场的依赖性等几个基础问题, 指出在深入研究模式竞争机理的基础上发展高阶模式和次谐波系统将有助于推动电子回旋器件实现高功率、高效率和高稳定性, 这对促进器件向太赫兹频段发展具有参考意义.

**关键词:** 电子回旋脉塞; 回旋器件; 模式竞争; 欧姆损耗; 次谐波

**中图分类号:** O441.3; O463<sup>+</sup>.1

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1880-07

电子回旋脉塞 (Electron Cyclotron Maser, ECM) 是一种基于相对论效应的电磁辐射机制<sup>[1-3]</sup>. 电子回旋器件基于 ECM 机制, 采用回旋电子注与电磁波相互作用, 通过电磁波对电子的同步调制使得电子动能、相对论质量因子、回旋频率等参数发生变化, 进而引起电子注角向群聚, 使电子注与波之间发生持续的能量交换. 电子回旋器件属于快波器件, 能在毫米波-太赫兹频段产生高功率相干电磁辐射<sup>[1-3]</sup>. 由于电子回旋器件具有高频率、高功率和高效率等突出优点, 因而在远程毫米波成像雷达、定向能武器、电子对抗、核聚变的等离子体加热等方面具有不可替代的重要性<sup>[1-5]</sup>. 探索和发展更高频率和更高功率的电磁波源始终是电子器件发展的重要方向<sup>[6]</sup>. 但是, 在高功率电子回旋器件从毫米波向太赫兹方向发展的同时, ECM 电子注-波相互作用系统面临着严重的欧姆耗散、模式竞争和对强磁场的强依赖性等关键

问题<sup>[3]</sup>. ECM 相互作用电路通常采用金属结构, 由于金属材料的电导率有限, 当频率提高到毫米波-太赫兹波段时, 欧姆耗散引起的功率损耗与输出功率处于同一量级, 欧姆耗散成为影响输出功率的关键因素; 模式竞争容易引起 ECM 系统自激励振荡, 使得系统输出效率低, 频谱出现较多杂谱分量; 次谐波相互作用能够成倍降低系统所需磁场强度, 但系统会面临较低的相互作用效率和更加复杂的模式竞争问题. 对这些问题进行深入研究并探索相应的解决方案对毫米波-太赫兹电子回旋脉塞器件的发展具有重要意义.

## 1 原理、器件及应用

电磁频谱是一种宝贵的电磁资源. 典型的大气传输电磁波谱特性如图 1 所示<sup>[2,7]</sup>.

由于传统的真空电子器件和基于能级跃迁的量子电子器件都难以在毫米波-太赫兹波段有

收稿日期: 2015-04-16; 录用日期: 2015-05-14; 网络出版时间: 2015-06-03 15:27

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150603.1527.005.html

基金项目: 国家自然科学基金面上项目(61471007, 11275206); 北京大学医学与信息联合种子基金(2014-MI-01); 国家“985”重点建设项目

作者简介: 杜朝海(1982—), 男, 研究员, duchaochai@pku.edu.cn

\* 通讯作者: 刘濮鲲(1965—), 男, 教授, pkliu@pku.edu.cn, 主要研究方向为微波、毫米波与太赫兹电子学.

**引用格式:** 杜朝海, 罗里, 刘濮鲲. 毫米波-太赫兹电子回旋器件的几个基础问题[J]. 北京航空航天大学学报, 2015, 41(10): 1880-1886. Du C H, Luo L, Liu P K. Several fundamental challenges of millimeter-to-terahertz electron cyclotron devices[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1880-1886 (in Chinese).

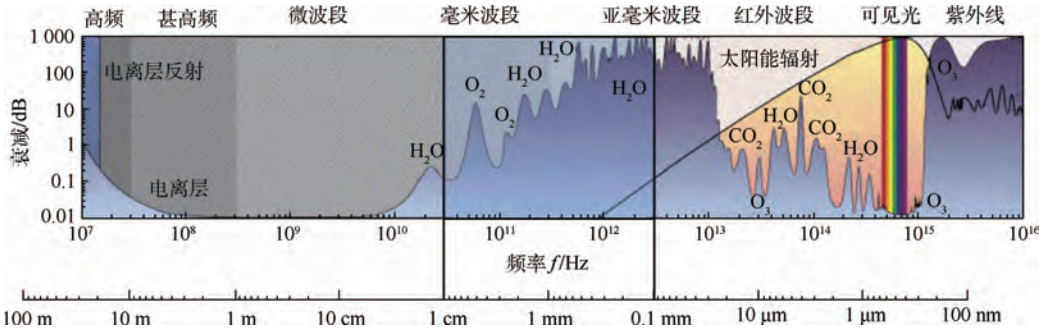


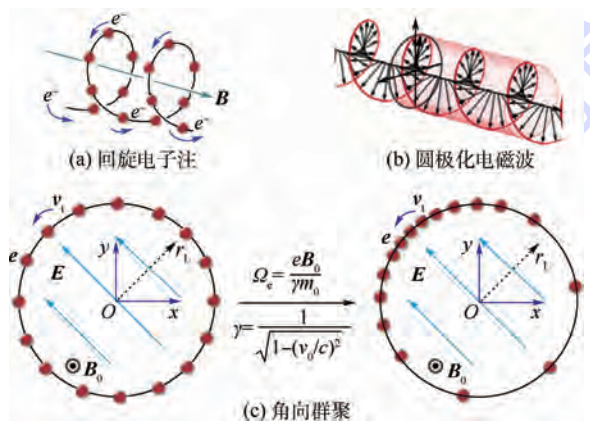
图 1 大气中不同波段电磁波谱<sup>[2,7]</sup>

Fig. 1 Electromagnetic spectrum of different frequency bands in atmosphere<sup>[2,7]</sup>

效地产生电磁辐射,导致电磁频谱上出现了所谓的太赫兹“空白”<sup>[1-5]</sup>. 由于对应的电磁波源、检测技术和应用技术都相对匮乏,使得太赫兹频段成为一段亟待充分开发的电磁频谱<sup>[1]</sup>.

ECM 能够在毫米波-太赫兹频段产生高功率相干电磁辐射<sup>[1-10]</sup>. 常规真空器件相互作用结构尺寸与工作波长之间存在“共渡效应”,而 ECM 辐射可以采用高阶电磁模式工作,其辐射频率由磁场决定,突破了“共渡效应”的限制,具有高频率和高功率的突出优势,这为解决毫米波-太赫兹“空白”提供了可行方案<sup>[1-2,11-14]</sup>. ECM 辐射基于相对论效应,采用单电子实现多光子辐射,具有高效率的能量转换能力;且由于 ECM 辐射采用多波长相互作用空间,因而具有大功率容量的优势. ECM 采用回旋电子注与波导模式相互作用,电子的轴向能量保证了电子与波导模式之间的多普勒频移回旋谐振(Doppler-shifted cyclotron resonance),保证电子与波之间的宽频带同步,电子通过角向群聚把横向动能转变为电磁波辐射.

图 2 给出了 ECM 的原理示意图<sup>[2]</sup>. 电子在磁场的引导下沿螺旋轨迹前进,具有回旋频率  $\Omega_e$ , 如图 2(a) 所示. 当回旋电子注遇上如图 2(b) 所



$E$ —场的大小;  $v_1$ —电子的横向速度;  $r_L$ —电子回旋的拉莫半径.

图 2 回旋脉塞中电子群聚原理<sup>[2]</sup>

Fig. 2 Principle of electron bunching in cyclotron maser<sup>[2]</sup>

示的极化电磁波,且二者满足相位同步关系  $\omega - k_z v_z \approx s\Omega_e$  时,能够发生持续的电子注-波交互作用,在电子的角向运动方向上发生净的能量交换,其中  $\omega$  和  $k_z$  分别为电磁波角频率和波数;  $v_z$  为  $z$  方向速度,  $s$  为谐波次数. 如图 2(c) 中所示,  $\Omega_e = eB_0 / (\gamma m_0)$ , 其中  $e$  为电子电荷量,  $B_0$  为磁场大小,  $m_0$  为电子质量,  $\gamma$  为相对论质量因子,且有

$$\gamma = 1 / \sqrt{1 - (v_0/c)^2}$$

式中:  $v_0$  为电子的速度;  $c$  为光速.

波相互作用过程中,回旋电子受到角向电场的加速或者减速,导致相对论质量因子变化. 由于引导磁场维持常数,被加速的电子相对论质量因子变大,回旋频率减小;被减速的电子相对论质量因子变小,回旋频率增加. 回旋频率的变化使得在拉莫尔圆上的电子在注-波相位空间上形成群聚,促使电子注与电磁波发生持续的能量交换<sup>[1-3]</sup>.

对 ECM 辐射机制的研究直接导致了回旋管振荡器、回旋管放大器和回旋自谐振脉塞(Cyclotron Auto-Resonance Maser, CARM)等一系列高功率毫米波-太赫兹电子器件的诞生<sup>[1-10]</sup>,并开辟出了微波电子学的一个重要分支——相对论电子学<sup>[8-10]</sup>.

图 3 给出了基于 ECM 原理的 4 大类电子回旋器件的原理性示意图<sup>[2]</sup>. 回旋管振荡器中,回旋单腔管(gyromonotron)是基于腔体的相互作用电路,具有高功率的优势;而回旋返波振荡器(Gyrotrotron Backward-Wave Oscillator, Gyro-BWO)是基于回旋谐波与反向波相互作用的可调谐 ECM 系统<sup>[1-2]</sup>. 回旋管放大器中,回旋速调管<sup>[15]</sup>采用群聚腔的调制作用,在输出腔获得高功率电磁波,其功率相对较高,稳定性较好,但是频带较窄;回旋行波管(Gyro-TWT)则是基于 ECM 对流不稳定性工作的宽频带放大器<sup>[1-3,16]</sup>. 电子回旋器件可在毫米波-太赫兹频段范围内产生瓦级至兆瓦级输出功率的脉冲或连续波辐射. 该领域的研究推动了高

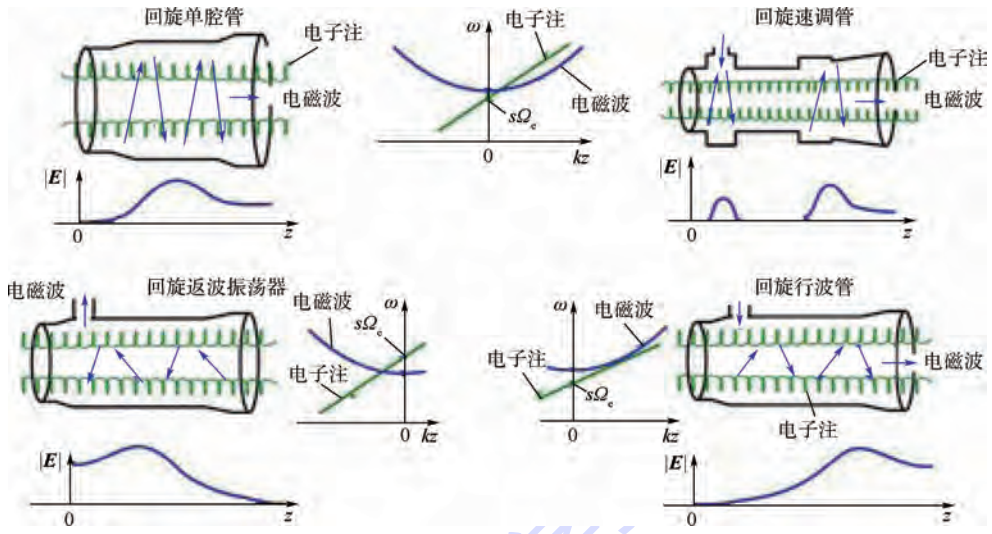


图3 4类回旋器件原理示意图<sup>[2]</sup>

Fig. 3 Schematic of principle of four kinds of gyrotron devices<sup>[2]</sup>

功率毫米波雷达、毫米波武器、受控核聚变的等离子体加热和材料处理等一系列相关领域的发展和变革,已经成为当代国防安全和科学技术研究中最活跃的研究领域之一<sup>[1-10]</sup>.

现以高功率回旋振荡器为例来说明回旋器件的结构.图4给出了高功率短毫米波回旋振荡器实物<sup>[17]</sup>和结构图<sup>[18]</sup>.

电子回旋器件采用磁控注入式电子枪产生高功率回旋电子注;在超导磁体提供的强磁场的导引下,回旋电子注与开放式谐振腔中的圆极化电磁模式相互作用,电子注的回旋动能转化为电磁

波能量;相互作用之后的电子注在磁场的导引下,被收集极截获,将剩余能量转化成为热能;高功率电磁波通过准光学模式变换系统转化为可以在自由空间传输的高斯波束,通过高功率真空窗输出,供应用系统使用.作为一种快波器件,高功率回旋器件拥有其他类型器件难以比拟的优势:无“共渡效应”、高能量转换效率、大功率容量等,因此在定向能武器<sup>[19]</sup>、高分辨率空间探测<sup>[20]</sup>、掠海飞行目标监测<sup>[21]</sup>、受控核聚变的等离子体加热<sup>[6,11]</sup>以及生物医学<sup>[22-23]</sup>等领域都有重要的应用前景,如图5所示.

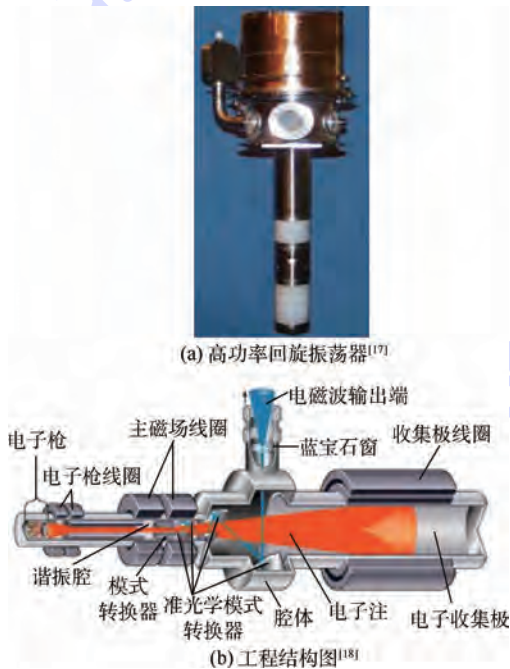


图4 高功率回旋振荡器<sup>[17]</sup>及其工程结构图<sup>[18]</sup>

Fig. 4 High-power gyrotron<sup>[17]</sup> and its engineering configuration<sup>[18]</sup>

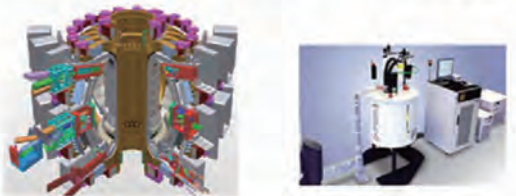


图5 基于高功率电子回旋器件的应用系统<sup>[6,11,19-23]</sup>

Fig. 5 Application system of high power cyclotron electron devices<sup>[6,11,19-23]</sup>

## 2 ECM 相互作用系统的关键问题

探索和发展更高频率的高功率电磁波源是电子器件的重要发展方向. ECM 系统能够产生高功率电磁辐射的根本原因在于,强磁场为电子注提供了固有回旋频率,使得回旋电子注能够与高阶

电磁模式互作用,电磁模式的阶数越高,对应的注-波互作用空间和功率容量越大. 伴随电子回旋器件的发展,互作用系统采用的工作模式经历了由低阶模式向高阶模式的发展. 图 6 给出了过去 20 多年时间内,高功率回旋振荡器所采用的典型工作模式的演变过程. 工作模式由早期的圆对称 TE<sub>0,2</sub> 模式,发展到近年来德国、日本和俄罗斯等国发展兆瓦级回旋管所采用的高阶腔体模式,如 TE<sub>28,8</sub>、TE<sub>31,8</sub>、TE<sub>25,10</sub> 等,腔体的空间尺度由最初的约 2 倍波长,扩展到当今的约 20 倍波长量级<sup>[1-3,6]</sup>.

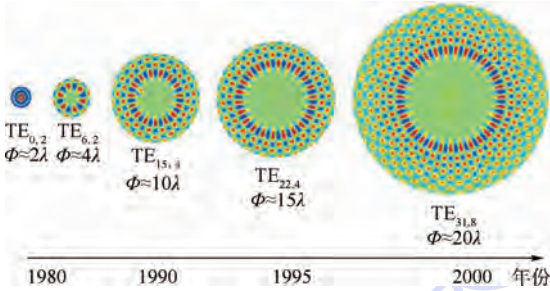
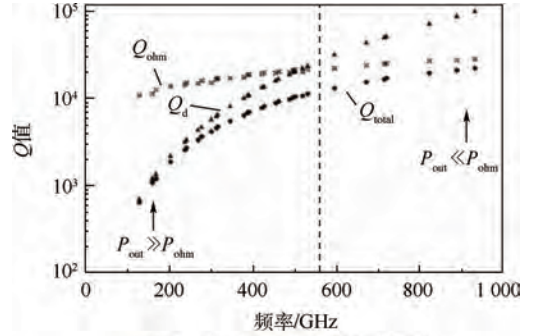


图 6 1980—2000 年,高功率回旋振荡器工作模式的演变过程

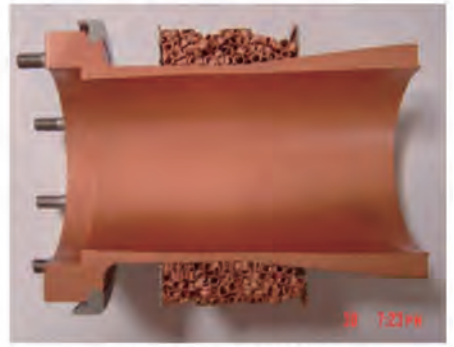
Fig. 6 Evolution process of high power gyrotron oscillator working mode from 1980 to 2000

短毫米波 ECM 系统面临严重的互作用腔体欧姆耗散问题,需要采用高阶模式工作以降低系统的损耗功率,提高有效输出功率. 对于 ECM 系统常用的开放腔体,回旋电子注交出的功率一部分辐射出腔体,一部分被欧姆损耗转化成热能. 通常采用欧姆品质因数  $Q_{ohm}$  和辐射品质因数  $Q_d$  以分别衡量腔体的欧姆耗散和辐射效率. 根据简单的经验公式可以得到:  $Q_{ohm} \propto \sqrt{f_0}$ ,  $Q_d \propto f_0^2$ , 其中:  $f_0$  为频率<sup>[24]</sup>.

在低频段通过恰当的设计很容易实现  $Q_{ohm} \gg Q_d$ , 即欧姆耗散的功率远小于输出功率;但是由于欧姆品质因数和辐射品质因数对频率的敏感性不同,随着频率从毫米波频段进入太赫兹频段,为了获得较高的相互作用效率,通常出现  $Q_{ohm} \leq Q_d$ , 即欧姆耗散功率高于输出功率<sup>[24]</sup>. 图 7(a) 给出了理论计算得到的典型太赫兹回旋管在一系列工作模式上的品质因数变化情况<sup>[24]</sup>. 可见,在毫米波频段 (< 100 GHz) 欧姆品质因素远大于辐射品质因数 ( $Q_{ohm} > 10Q_d$ ), 欧姆损耗的影响较小;当频率提高到短毫米波 300 GHz 附近,欧姆耗散功率达到输出功率一半的量级 ( $Q_{ohm} \sim 2Q_d$ );在更高的工作频率上,通常表现出欧姆耗散功率超过输出功率. 可见短毫米波 ECM 系统的欧姆耗散是十分严



(a) 一系列模式的品质因数变化<sup>[24]</sup>



(b) 高功率回旋管腔体结构图<sup>[25-27]</sup>

图 7 太赫兹回旋管中一系列工作模式品质因数变化情况<sup>[24]</sup>和高功率回旋管腔体结构图<sup>[25-27]</sup>

Fig. 7  $Q$  factors of modes in a THz cavity<sup>[24]</sup> and structure of high-power gyrotron cavity<sup>[25-27]</sup>

重的问题. 图 7(b) 给出了高功率回旋振荡器互作用腔的典型结构图<sup>[25-27]</sup>.

由图 7 可见,工程上为了降低高功率回旋管腔体欧姆损耗需要在腔体外壁加上特殊的结构,达到散热和控制热形变的目的. 因此,需要发展高阶腔体模式互作用系统,提高  $Q_{ohm}$  的绝对值,同时优化互作用系统,保证较高相互作用效率情况下尽可能降低  $Q_d$ ,以获得较高的有效输出功率.

高阶模式工作面临严重的模式竞争. 模式竞争是一个复杂的非线性物理问题<sup>[28]</sup>. ECM 器件采用高阶模式互作用腔体是增加系统的功率容量、降低欧姆耗散(提高  $Q_{ohm}$ )的重要途径. 然而高阶模式腔体同时会导致系统具有更高的模式密度,同一回旋谐波可能会激发工作频率相近的多个模式;同时也可能出现不同回旋谐波激发不同的模式,竞争模式的频率近似成倍频关系. 模式竞争问题导致 ECM 系统出现非稳态的多模竞争行为,工作效率低,频谱不纯. 此外,器件在高电压脉冲建立的过程中,工作模式起振是一个十分复杂的过程,而模式竞争问题也会影响高阶模式 ECM 系统的起振过程 (start-up scenario). 图 8 给出了美国 CPI 公司发展 140 GHz 回旋振荡器的工作模式建立过程<sup>[29]</sup>. 器件启动的过程中,电压和电流

都是缓慢升高的,系统工作产生的模式是时间的函数,器件的性能和电压调制器的工作参数决定了模式建立过程的系统参数.在每组参数下,系统可能会出现某种中间状态(模式),这种中间状态的建立有可能抑制或者帮助后续工作模式的建立.模式竞争可能有助于提高器件的效率.为了使ECM系统获得较高的相互作用效率,器件往往工作在“硬激励”区域(hard excitation region)<sup>[30]</sup>.高效工作状态通常位于稳定工作参数窗(stable parameter window)的边缘,这时候器件十分容易受到竞争模式的影响.2007年,日本原子能机构在《Nature Physics》上发表了在170 GHz TE<sub>31,8</sub>模式回旋管上利用控制电子注参数达到控制模式竞争的目的,实验获得55%的高效率<sup>[11]</sup>.模式竞争导致ECM器件产生非相干的电磁辐射,从而限制了高功率波源在电子对抗、功率合成、成像、探测和通信等领域的应用.总之,高功率短毫米波回旋器件通过采用高阶工作模式,可提高系统功率容量,降低欧姆耗散的影响.因此,发展高阶工作模式下的相互作用系统对连续波状态工作的高功率短毫米

波回旋器件有重要意义.但是,高阶腔体模式相互作用系统面临的严重模式竞争问题,使得器件的稳定性受到严重影响,并且会导致杂谱和非相干电磁辐射.在高次谐波短毫米波ECM系统中抑制模式竞争,获得高功率输出和高效率相互作用还需要解决一系列关键问题.

高次谐波ECM系统能够成倍降低磁场强度,但会导致较低的相互作用效率和更复杂的模式竞争问题.ECM原理决定了回旋管需要强磁场为电子注提供ECM作用所需的固有回旋频率.工作频率100 GHz的基波回旋管需要磁场约3.7 T,必须采用超导磁场工作;随着工作频率提高,频率为1 THz的基波回旋管,所需磁场强度达到40 T,超出了超导磁体的能力范围.高次谐波回旋管能够解决太赫兹回旋管所需的强磁场问题.因为谐波相互作用系统( $s \geq 2$ )能够将磁场强度降为基波系统的 $1/s$ .谐波相互作用模式使得采用超导磁体的回旋管可获得高频率(1 THz)连续波,从而促进了高分辨率的DNP-NMR、太赫兹成像探测等系统的应用;同时谐波相互作用使得采用永磁体或者线包磁体的低频率(94 GHz)回旋管系统的可靠性和快速启动能力大大增强,适合应用于机载、舰载和车载等高机动性的雷达或武器系统<sup>[2,19]</sup>.目前这两方面应用对高次谐波回旋管提出了直接、迫切的需求.表1给出了日本FUKUI大学开展了一系列谐波回旋管的实验测试结果<sup>[30-33]</sup>.图9给出了日本<sup>[30]</sup>、俄罗斯<sup>[34-35]</sup>和英国<sup>[36]</sup>分别基于超导、脉冲磁体和常规的线包磁体发展的3种不同类型的短毫米波谐波回旋管.此外2000年,我国电子科技大学曾经发展了基于永磁体的Ka波段三次谐波回旋管<sup>[37]</sup>,近几年发展了基于脉冲磁体<sup>[38]</sup>和超导磁体<sup>[39]</sup>的太赫兹回旋管.谐波回旋管是ECM发现以来的主要趋势之一,同时也是未来短毫米波回旋器件解决强磁场问题,获得连续波的必然发展方向.

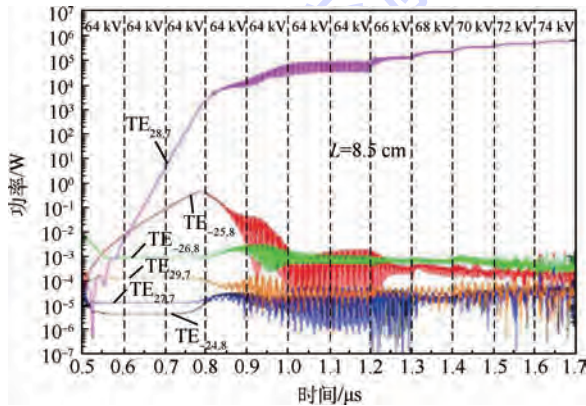


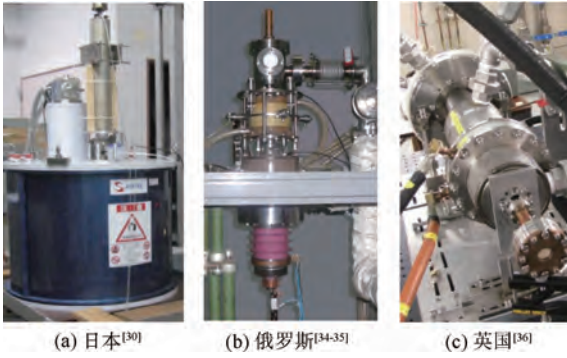
图8 美国CPI公司发展的140 GHz, TE<sub>28,7</sub>模回旋管的起振过程<sup>[29]</sup>

Fig. 8 Start-up scenario of 140 GHz, TE<sub>28,7</sub> mode gyrotron developed by CPI corporation of USA<sup>[29]</sup>

表1 日本FUKUI大学开展回旋管实验<sup>[30-33]</sup>

Table 1 Gyrotron experiments carried out in FUKUI University of Japan<sup>[30-33]</sup>

型号	频率/GHz	功率	连续波(CW)/脉冲(Pulse)	谐波次数	工作模式	发展进度
FU CW I	300	2.5 kW	CW	1st		测试完成
FU CW II	394	几百瓦(1st) 几十瓦(2nd)	CW	2nd	TE <sub>2,5</sub>	测试完成
FU CW III	1013.7	几百瓦(1st) 几十瓦(2nd)	CW	2nd	TE <sub>4,12</sub>	测试完成
FU CW IV	131.5	几十瓦	CW	1st	TE <sub>1,2</sub>	测试完成
其他 I	424	100 W	CW	2nd	TE <sub>8,2</sub>	设计中
其他 II	349/390	几十千瓦 100 kW	Pulse	2nd	TE <sub>6,5</sub> , TE <sub>8,5</sub>	测试完成



(a) 日本<sup>[30]</sup> (b) 俄罗斯<sup>[34-35]</sup> (c) 英国<sup>[36]</sup>

图 9 日本<sup>[30]</sup>、俄罗斯<sup>[34-35]</sup>和英国<sup>[36]</sup>分别基于超导磁体、脉冲磁体和线包磁体发展的谐波回旋管

Fig.9 Harmonic gyrotrons developed by Japan<sup>[30]</sup>, Russia<sup>[34-35]</sup>, and UK<sup>[36]</sup> based on superconducting magnet, pulse magnet and coil magnet, respectively

### 3 前景

未来发展低电压、紧凑型、连续波的毫米波-太赫兹回旋器件是很有前景的发展方向,能够为成像、探测和生物医学等领域带来一系列革新的解决方案.针对毫米波-太赫兹电子回旋脉塞器件所面临的 3 个关键问题,有重要价值的研究工作包括:

- 1) 发展高阶腔体模式工作的 ECM 系统,提高系统欧姆品质因数,降低欧姆耗散的影响.
- 2) 研究 ECM 系统模式竞争机理和控制技术,提高系统稳定性.
- 3) 发展高次谐波 ECM 系统,成倍地降低系统所需的磁场强度,同时获得较高的互作用效率,实现连续波高平均功率工作.通过这 3 个方面关键科学问题的研究,可为未来发展高功率太赫兹回旋器件提供前瞻性引导,从而解决回旋器件难以实现高稳定性、高功率和高效率等困难,为工程发展高功率连续波工作的回旋器件提供创新方案和技术储备.

#### 参考文献 (References)

- [1] Chu K R. The electron cyclotron maser [J]. Reviews of Modern Physics, 2004, 76 (2) : 489-540.
- [2] Du C H, Liu P K. Millimeter-wave gyrotron traveling-wave tube [M]. Berlin: Springer-Verlag, 2014: 1-192.
- [3] Hirshfield J L, Wachtell J M. Electron cyclotron maser [J]. Physical Review Letters, 1964, 12 (19) : 533-536.
- [4] Booske J H, Dobbs R J, Joye C D. Vacuum electronic high power terahertz sources [J]. IEEE Transactions on Terahertz Science and Technology, 2011, 1 (1) : 54-75.
- [5] Nusinovich G S, Thumm M K, Petelin M I. The gyrotron at 50: Historical overview [J]. Journal of Infrared Millimeter and Terahertz Waves, 2014, 35 (4) : 325-381.
- [6] Thumm M. State-of-art of high power gyro-devices and free electron masers update 2012, 7641 [R]. Karlsruhe: KIT Scientific Publishing, 2013.
- [7] SIEGEL P H. Terahertz technology [J]. IEEE Transactions on Microwave Theory and Techniques, 2002, 50 (3) : 910-928.
- [8] 刘盛纲. 太赫兹科学技术的新发展 [J]. 中国基础科学, 2006, 8 (1) : 7-12.
- [9] Liu S G. Recent development of terahertz science and technology [J]. Chinese Basic Science, 2006, 8 (1) : 7-12 (in Chinese).
- [10] 刘盛纲. 相对论电子学 [M]. 北京: 科学出版社, 1987: 1-187.
- [11] Liu S G. Relativistic electronics [M]. Beijing: Science Press, 1987: 1-187 (in Chinese).
- [12] Liu S G, Yang Z H. The kinetic theory of the ECRM with space-charge effect taken into consideration [J]. International Journal of Electronics, 1981, 51 (4) : 341-349.
- [13] Sakamoto K, Kasugai A, Minami R, et al. Achievement of robust high-efficiency 1 MW oscillation in the hard-self-excitation region by a 170 GHz continuous-wave gyrotron [J]. Nature Physics, 2007, 3 (6) : 411-414.
- [14] Granatstein V L, Levush B, Danly B G, et al. A quarter century of gyrotron research and development [J]. IEEE Transactions on Plasma Science, 1997, 25 (6) : 1322-1335.
- [15] Sirtori C. Applied physics: Bridge for the terahertz gap [J]. Nature, 2002, 420 (6912) : 131-133.
- [16] Thumm M. Free-electron masers vs gyrotrons: Prospects for high-power sources at millimeter and submillimeter wavelengths [J]. Nuclear Instruments and Methods in Physics Research A, 2002, 483 (1-2) : 186-194.
- [17] 刘濮鲲, 徐寿喜. 回旋速调管放大器及其发展评述 [J]. 电子与信息学报, 2003, 25 (5) : 683-694.
- [18] Liu P K, Xu S X. Review of gyrokystron amplifier and its development [J]. Journal of Electronics and Information Technology, 2003, 25 (5) : 683-694 (in Chinese).
- [19] 刘濮鲲, 杜朝海. 毫米波回旋波放大器的发展述评 [J]. 微波学报, 2013, 29 (5-6) : 33-42.
- [20] Liu P K, Du C H. Review of the gyrotron traveling-wave-tube amplifier development [J]. Journal of Microwaves, 2013, 29 (5-6) : 33-42 (in Chinese).
- [21] Communications & Power Industries. Microwave power products division, VGB-8095 gyrotron oscillator [EB/OL]. California: Communications & Power Industries, 2007 [2014-11-01]. <http://www.digchip.com/datasheets/parts/datasheet/888/VGB8095-pdf.php>.
- [22] Estanislao Navarro Sosa. Electron tube [DB/OL]. Chicago: Encyclopedia Britannica, Inc., 1768 (2002) [2014-11-01]. <http://global.britannica.com/technology/electron-tube>.
- [23] Mchale J. Raytheon delivers active denial system 2 to U. S. air force [J]. Military & Aerospace Electronics, 2007, 18 (12) : 8.
- [24] MIT Lincoln Laboratory Members. Haystack ultra-wideband satellite imaging radar [EB/OL]. Massachusetts: MIT Lincoln Laboratory, 2004 [2014-11-01]. <http://www.haystack.mit.edu/obs/haystack/LincolnUpgrade.pdf>.
- [25] Linde G J, Ngo M T, Danly B G, et al. WARLOC: A high-power coherent 94 GHz radar [J]. IEEE Transactions on Aerospace and Electronic Systems, 2008, 44 (3) : 1102-1117.
- [26] Ni Q Z, Daviso E, Can T, et al. High frequency dynamic nuclear

- polarization[J]. *Accounts of Chemical Research*, 2012, 46(9): 1933-1941.
- [23] Nanni E A, Barnes A B, Griffin R G, et al. THz dynamic nuclear polarization NMR[J]. *IEEE Transactions on Terahertz Science and Technology*, 2011, 1(1): 145-163.
- [24] Kao S H, Chiu C C, Chu K R. A study of sub-terahertz and terahertz gyrotron oscillators[J]. *Physics of Plasmas*, 2012, 19(2): 023112.
- [25] Thumm M. Progress in gyrotron development[J]. *Fusion Engineering and Design*, 2003, 66(8): 69-90.
- [26] Dumbrajs O, Zaginaylov G I. Ohmic losses in coaxial gyrotron cavities with corrugated insert[J]. *IEEE Transactions on Plasma Science*, 2004, 32(31): 861-866.
- [27] Kreischer K E, Temkin R J. Single-mode operation of a high-power, step-tunable gyrotron[J]. *Physical Review Letters*, 1987, 59(5): 547-550.
- [28] Kao S H, Chiu C C, Pao K F, et al. Competition between harmonic cyclotron maser interactions in the terahertz regime[J]. *Physical Review Letters*, 2011, 107(13): 135101.
- [29] Nusinovich G S, Simitsyn O V, Velikovich L, et al. Startup scenarios in high-power gyrotrons[J]. *IEEE Transactions on Plasma Science*, 2004, 32(3): 841-852.
- [30] Saito T, Tatematsu Y, Yamaguchi Y, et al. Observation of dynamic interactions between fundamental and second-harmonic modes in a high-power sub-terahertz gyrotron operating in regimes of soft and hard self-excitation[J]. *Physical Review Letters*, 2012, 109(15): 155001.
- [31] Bratman V, Glyavin M, Idehara T, et al. Review of subterahertz and terahertz gyrodevices at IAP RAS and FIR FU[J]. *IEEE Transactions on Plasma Science*, 2009, 37(1): 36-43.
- [32] Idehara T, Ogawa I, Mitsudo S, et al. Development of frequency tunable, medium power gyrotrons (Gyrotron FU series) as sub-millimeter wave radiation sources[J]. *IEEE Transactions on Plasma Science*, 1999, 27(2): 340-354.
- [33] Idehara T, Saito T, gawa I O, et al. Development of terahertz FUCW gyrotron series for DNP[J]. *Applied Magnetic Resonance*, 2008, 34(3-4): 265-275.
- [34] Glyavin M Y, Luchinin A G, Golubiatnikov G Y. Generation of 1.5-kW, 1-THz coherent radiation from a gyrotron with a pulsed magnetic field[J]. *Physical Review Letters*, 2008, 100(1): 015101.
- [35] Bratman V L, Kalynov Y K, manuilov V N. Large orbit gyrotron operation in the terahertz frequency range[J]. *Physical Review Letters*, 2009, 102(24): 245101.
- [36] He W L, Donaldson C R, Zhang L, et al. High power wideband gyrotron backward wave oscillator operating towards the terahertz region[J]. *Physical Review Letters*, 2013, 110(16): 165101.
- [37] Huang Y, Li H F, Yang S W, et al. Study of a 35-GHz third harmonic low-voltage complex cavity gyrotron[J]. *IEEE Transactions on Plasma Science*, 1999, 27(2): 368-373.
- [38] Fu W, Yan Y, Li X, et al. The experiment of a 220 GHz gyrotron with a pulse magnet[J]. *Journal of Infrared, Millimeter, and Terahertz Waves*, 2010, 31(4): 404-410.
- [39] Yan Y, Fu W J, Li X Y, et al. Experimental results of a 0.42 THz harmonic gyrotron[C]//35th International Conference on Infrared, Millimeter, and Terahertz Waves, Conference Guide. Piscataway, NJ: IEEE Press, 2010: 1-2.

## Several fundamental challenges of millimeter-to-terahertz electron cyclotron devices

DU Chaohai<sup>1</sup>, LUO Li<sup>2,3</sup>, LIU Pukun<sup>\*1</sup>

(1. School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China;

2. Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China;

3. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Science, Beijing 100049, China)

**Abstract:** It is always an important tendency of the electronic power devices to explore and develop electromagnetic radiation sources operating on higher frequency and with higher power capability. The cyclotron electron device which was developed based on the mechanism of the electron cyclotron maser and with the excellent advantage of radiating high power in millimeter-to-terahertz range was introduced. It systematically discussed about the fundamental challenges encountered by the electron cyclotron devices, including the strong ohmic dissipation problem, the mode competition problem, and the unavoidable dependency on operating with strong magnetic field. Finally, it was proposed that, on the basis of fully exploring the physical mechanism of the mode competition, to develop higher order mode circuit and higher harmonic interaction system would be very helpful to realize the high power, high efficiency and high stability cyclotron electron devices, which is of important reference for guiding devices development towards terahertz band.

**Key words:** electron cyclotron maser; gyrotron devices; mode competition; ohmic dissipation; higher harmonic

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0074

# W波段多注交错双栅行波放大器的高频特性

张慕武<sup>1,2</sup>, 阮存军<sup>\*3</sup>

(1. 中国科学院 电子学研究所, 北京 100190; 2. 中国科学院大学 电子电气与通信工程学院, 北京 100049;

3. 北京航空航天大学 电子信息工程学院, 北京 100191)

**摘 要:** 提出了一种采用平面线性排列的3个圆形电子注代替平面带状电子注的W波段交错双栅行波放大器微型化慢波高频结构,并重点对其行波传输特性和输入输出耦合特性进行了仿真设计和参数优化.结果表明,这种平面微型化高频结构的行波传输色散特性良好且有很大的工作带宽,所产生的强轴向电场分布非常有利于电子注与高频场的能量交换和相互作用.在保证平面微型结构中圆形注通道直径和带状注通道高度相同的情况下,得到的耦合阻抗是带状注交错双栅慢波高频系统2~3倍,为行波放大器高效率的注波相互作用和高功率输出提供了新的研究思路.为了与该交错双栅高频系统相匹配,提出了一种更为简单易行的输入与输出耦合结构,仅采用三周期渐变过渡段就可以实现反射系数 $S_{11}$ 在较宽频带内低于-20 dB的良好结果,更有利于行波放大器未来的工程实现与应用.

**关键词:** W波段; 交错双栅; 多电子注; 行波放大器; 耦合结构

**中图分类号:** O441.3

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1887-07

行波放大器是一种对微波和毫米波进行放大并产生大功率相干辐射的传统真空电子器件,主要利用在真空中运动的强流电子注与慢波高频系统中的行波电磁场相互作用,把电子注动能逐步转化为电磁波能量,从而实现对电磁场高频信号的放大与应用<sup>[1]</sup>.行波放大器具有宽频带、高效率和高功率等技术优势,在电子对抗、卫星通信、雷达成像等领域得到了广泛的应用<sup>[2-3]</sup>.但是,随着电磁频谱向毫米波和太赫兹等高频段发展,由于器件几何结构尺寸与频率的共渡性效应,所需的高频结构必须向微型化和高精度化发展.为此,传统的轴对称加工技术已无法满足这类高频段微型化器件尺寸及其精度的要求.近年来,随着现代平面微加工方法与技术的蓬勃发展与交叉融合,有望使得阻碍真空电子器件向高频段发展的微型

化结构及其加工技术瓶颈等问题在一定程度上得到解决.众所周知,现代微加工技术往往是基于二维的平面加工技术<sup>[4]</sup>,很难制作出圆柱轴对称结构的螺旋线、耦合腔等传统高频慢波结构.为了适应平面化微结构加工技术,折叠波导、单矩形栅、交错双栅<sup>[5-7]</sup>等新型平面慢波结构被提出并逐步成为了本领域前沿热点研究工作.在这些微型化平面慢波高频系统中,折叠波导有着宽频带和良好的阻抗匹配等特点,但其功率容量较小,难以产生高功率输出;单矩形栅结构功率容量大,易于加工,但其应用受限于窄带宽;交错双栅结构最大特点在于它既有宽频带,又有着很大的功率容量,相比较而言在高功率宽频带行波管研究领域已得到了广泛重视.交错双栅结构是在传统对称双栅的基础上,通过上下两排矩形栅结构的半周期错位,

收稿日期: 2015-02-04; 录用日期: 2015-03-13; 网络出版时间: 2015-04-21 17:24

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150421.1724.001.html

基金项目: 国家自然科学基金(61222110)

作者简介: 张慕武(1990—),男,江西上饶人,硕士研究生, zhangmuwu@126.com

\* 通讯作者: 阮存军(1974—),男,研究员, ruancunjun@buaa.edu.cn, 主要研究方向为微波、毫米波与太赫兹电子学.

**引用格式:** 张慕武, 阮存军. W波段多注交错双栅行波放大器的高频特性[J]. 北京航空航天大学学报, 2015, 41(10): 1887-1893.  
Zhang M W, Ruan C J. High frequency characteristics for W-band multiple beam staggered double-vane traveling wave tube amplifier[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1887-1893 (in Chinese).

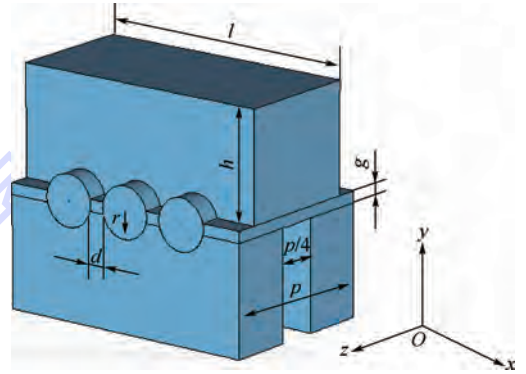
使得上下相邻单栅结构的 TE 模彼此发生  $180^\circ$  的相移. 因此, 对称栅通道中原本关于通道中心反对称的轴向电场变为对称场<sup>[8]</sup>, 从而使得传输的行波工作于基模且有较高的耦合阻抗. 近两年, Shin 等对 220 GHz 带状注交错双栅结构的研究表明, 其交错双栅高频结构冷带宽可达 0.25 个倍频程, 行波放大器峰值输出功率可达 160 W<sup>[7]</sup>. 国内学者对交错双栅高频系统也做了初步的研究<sup>[9-11]</sup>, 电子科技大学的赖剑强等在 W 波段对此结构进行了系统的设计<sup>[10]</sup>, 结果表明, 当采用带状电子注方案时, 在 W 波段可获得 1122 W 的饱和峰值输出, 3 dB 带宽约 15 GHz. 但是, 由于行波放大器高频结构较长 ( $\approx 119$  mm), 以及带状注固有的难以约束和良好聚焦等问题, 他们不得不采用高达 1.3 T 的均匀聚焦磁场来约束带状注的长距离传输. 随后, 他们又提出利用单注代替带状注应用于交错双栅系统来解决带状注聚焦困难问题<sup>[11]</sup>, 但其放大器输出功率只有 289 W. 总体而言, 平面多电子注和带状注是两种在高频段平面化微型结构中产生高功率输出的有效解决方案<sup>[12]</sup>: 平面多注不仅可以解决带状注聚焦困难等问题, 并且由于电磁波在微型结构多注通道内处于截止状态, 因此相对于带状行波放大器中复杂的输入和输出耦合系统而言, 多注的耦合结构设计将会较为相对简单和在工程实现上更为可行. 此外, 针对平面折叠波导的毫米波和太赫兹放大器研究工作, 国内外已经有一些关于多注行波管的设计实例<sup>[5, 13-14]</sup>.

基于上述分析与考虑, 本文提出了在微型化交错双栅结构中, 采用平面多注代替带状注在 W 波段实现宽频带行波放大并产生高功率输出的研究方案, 并对采用 3 个平面电子注的行波放大器高频特性进行了详细的仿真设计与分析. 结果表明, 采用这种技术方案不仅具有超宽带输出的技术优势, 而且在电子通道内可以产生较强的轴向电场分布. 在保证平面多注通道直径与带状注通道高度一致的情况下, 平面多注高频系统中间电子注通道轴线上的轴向电场强度是带状注通道中心轴线的 1.5 倍. 而在整个通带内, 其平均耦合阻抗也是带状注的 2~3 倍, 这将为平面多注行波放大器中电子注与高频场的相互作用与能量交换提供优良的技术特性. 同时, 本文也为三电子注的 W 波段行波放大器设计了一种简单易行的输入与输出耦合过渡结构, 结果表明仅仅利用 3 个周期过渡, 就可在很宽的工作频带内达到反射损耗  $S_{11} < -20$  dB 和传输损耗  $S_{21} > -1$  dB 的效果, 从

而可以在一定程度上有效缩短行波放大器的轴向高频结构长度和电子注传输距离, 这将更有利于微型化行波放大器的工程实现.

## 1 行波高频系统的设计与分析

基于交错双栅行波放大器方案, 本文提出的三电子注行波高频系统单个周期的结构如图 1 所示.



$l$ —高频结构的宽度;  $h$ —交错栅的高度;  
 $g$ —上下栅之间的距离;  $r$ —电子注半径;  
 $d$ —电子注通道边缘之间的距离;  $p$ —结构周期.

图 1 三电子注交错双栅单个周期慢波结构  
三维仿真结构图

Fig. 1 Three-dimensional simulation structure  
schematic for a single period of three beam  
staggered double-vane

### 1.1 色散特性分析

根据本文所确定的行波放大器中心工作频率  $f_0 = 94$  GHz, 同步电压  $V_0 = 21$  kV, 同步相位  $\phi = 2.7\pi$ , 由

$$V_p = \beta c = c \sqrt{1 - 1/[1 + V_0/(5.11 \times 10^5)]^2} \quad (1)$$

可以得到该放大器的同步相速  $V_p = 0.2785c$ , 其中  $c$  为真空中电磁波的传播速度;  $\beta$  为相速与光速的比值. 进而由

$$p = \phi/k_z = \phi/(\omega/V_p) = \phi V_p/(2\pi f_0) \quad (2)$$

可以确定出行波高频结构的周期  $p = 1.2$  mm.

为了对交错双栅周期高频结构的色散特性进行分析研究, 在保持其他参数不变的情况下, 通过改变高频结构的某一个几何参数, 得到的色散特性随该结构参数的变化关系如图 2 所示. 对于本文设计的三注行波放大器双栅高频结构工作在基模, 其色散特性和矩形波导  $TE_{10}$  基模类似, 横向宽度  $l$  决定了最低截止频率, 也就是横向截止频率. 随着  $l$  的增加, 其横向截止频率下降, 如图 2(b) 所示, 也导致了归一化相速的总体下移, 如图 2(a) 所示; 交错栅高  $h$  决定着高频结构的纵

向截止频率,随着栅高  $h$  的增加,纵向截止频率下降,如图 2(d) 所示,这也导致高频部分归一化相速下降而低频部分的归一化相速基本保持不变,如图 2(c) 所示;电子注通道半径  $r$  的大小既影响着高频结构的横向尺寸也影响着纵向尺寸,当  $r$  变大时,可以理解为在横向上变得更弯曲,而纵向上半径的变大将导致栅高的减小,这对应着  $l$  的增加和  $h$  的减小,也对应着归一化相速整体下移,以及高频段归一化相速增加和低频归一化相速不变,即随着  $r$  的增加,整个频带内归一化相速整体下降,但低频部分的相速下降的幅度较高频部分

幅度更大,如图 2(e) 所示;对于电子注通道边缘之间的距离  $d$  和横向宽度  $l$  一样,仅仅影响着结构的横向尺寸,随着  $d$  的增加,结构的归一化相速整体下降,如图 2(f) 所示;上下栅之间的距离  $g$  对色散的影响不大,但对其散射参数影响很大.通过详细的交错双栅高频结构仿真设计与参数优化,得到在  $l = 1.86 \text{ mm}$ ,  $h = 0.77 \text{ mm}$ ,  $r = 0.18 \text{ mm}$ ,  $d = 0.1 \text{ mm}$  和  $g = 0.1 \text{ mm}$  时,该行波结构具有更宽的带宽,最终优化设计得到的色散特性如图 3 所示,可见该行波高频系统工作在基模的一次空间谐波,其冷色散频带宽高达 24 GHz 左右.

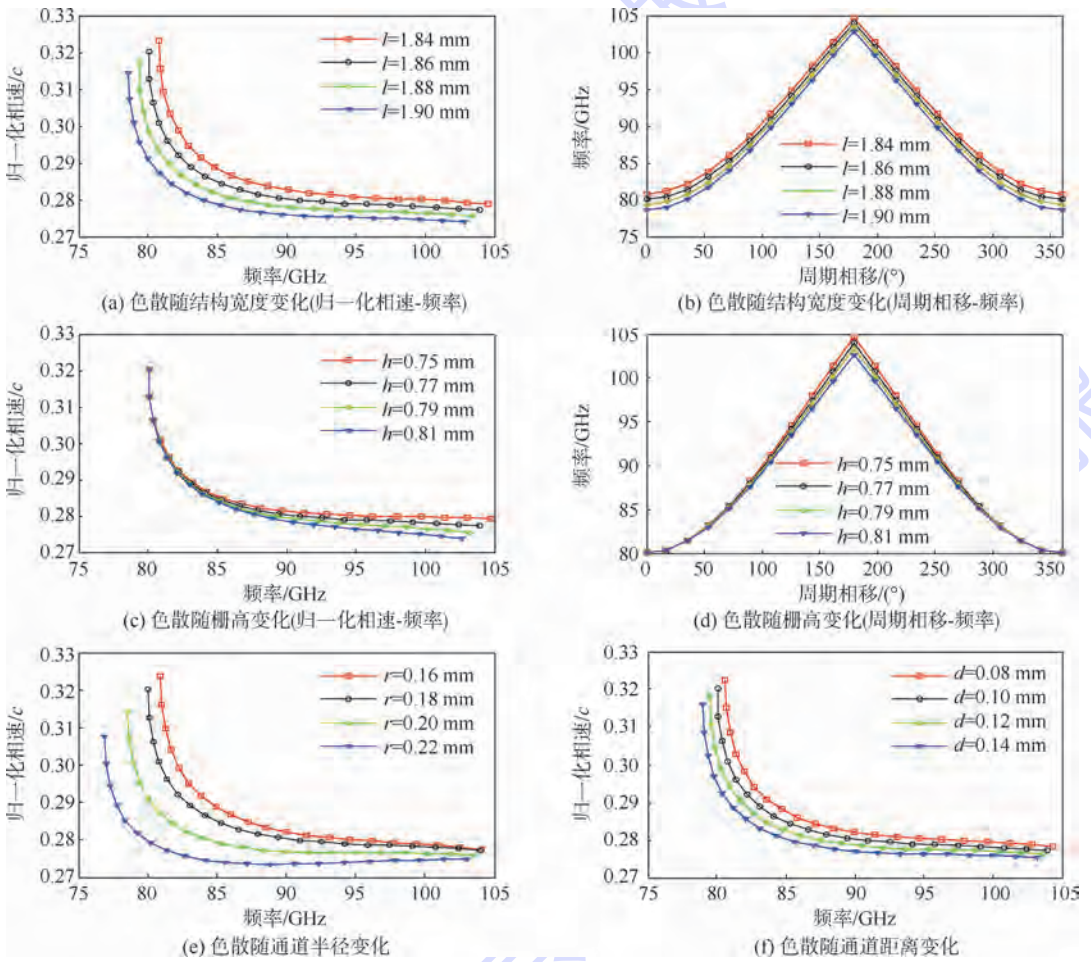


图 2 结构参数对平面多注交错双栅色散特性的影响

Fig. 2 Influences of structure parameters on dispersion characteristics of multiple beam staggered double-vane

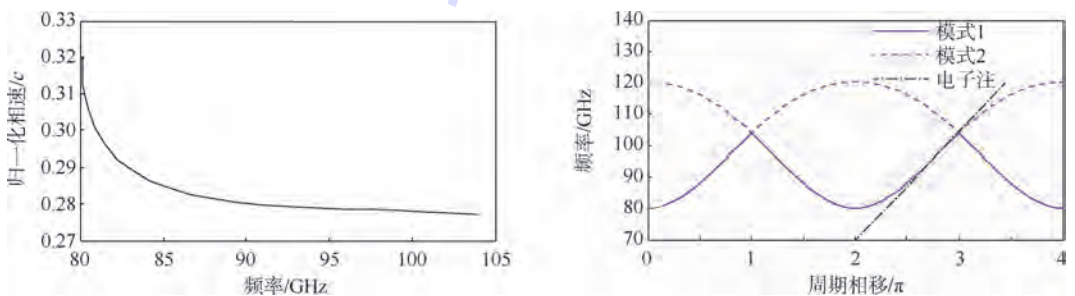


图 3 优化后结构的色散特性

Fig. 3 Dispersion characteristic of optimized structures

### 1.2 耦合阻抗

与对称双栅相比,交错双栅结构通过上下两排矩形栅的半周期错位,使得在电子通道内原本沿轴线反对称的电场变为对称的电场,从而可以大大提高该高频结构的耦合阻抗.但是,改进后交错双栅耦合阻抗还是不够高.由皮尔斯理论可知<sup>[1]</sup>,低的耦合阻抗意味着电子注和高频场的相互作用效率较低,这将导致更低的信号放大增长率和更长的相互作用距离,特别是对于带状注行波放大器,这将大大增加电子注聚焦与长距离稳定传输的难度.

为了对采用多电子注和带状注的交错双栅行波高频结构耦合阻抗进行比较与分析研究,对这两种结构分别在三维仿真软件 CST MWS 中建模.使得功率为 1 W 的 94 GHz 连续波正弦信号通过一个具有 20 周期的三注交错双栅慢波结构和一个具有 20 周期的带状注交错双栅结构,并保持三电子注通道直径和带状注通道高度相同.仿真得到的各高频结构 H 面和 E 面轴向电场分布图如图 4 所示.可以看出,和带状注交错双栅结构一样,在多注交错双栅行波高频系统中同样形成了关于中心轴对称的轴向电场.但是,通过对比图 4(a)、图 4(b)、图 4(c)和图 4(d)可以看出,多注行波高频系统内的轴向电场明显强于带状注系统内的轴向电场,特别是在电子注通道的中心处.

为了进一步进行比较,本文选取多注通道和带状注通道的截面如图 5(a)所示,其中  $2r$  为圆形电子注通道直径, $2t$  为带状注通道高度.在保证

$2r = 2t$  情况下,仍然让相同功率的 94 GHz 正弦信号分别通过带状注和多注交错双栅行波高频结构,通过观察多注通道中心和带状注通道中心对应相同位置  $a$ 、 $b$ 、 $c$ 、 $d$ 、 $e$  点的 94 GHz 轴向电场分布,结果如图 5(b)所示,其中用  $a_1$ 、 $b_1$ 、 $c_1$ 、 $d_1$ 、 $e_1$  表示多注高频结构中相对应点的轴向电场强度, $a_2$ 、 $b_2$ 、 $c_2$ 、 $d_2$ 、 $e_2$  表示带状注高频结构中相对应点的轴向电场强度.

由图 5(b)可以看出,多注结构中轴向电场明显大于带状注结构,特别是在  $c$  点和  $e$  点,这是因为交错双栅是一种表面波器件,在靠近栅表面处,电场呈指数增长.相对于带状注结构,多注结构中  $b$ 、 $c$ 、 $d$ 、 $e$  各点更靠近栅表面,特别是  $e$  点和  $c$  点.由此可见,采用多注行波结构可以把高频场约束在更小空间内从而达到了增强轴向电场的效果.当采用相同的电子注填充比 36% (圆形电子注半径取  $0.6r$ ,带状注截面取  $0.6l \times 0.6 \times 2r$ ) 时,利用 CST MWS 的本征模求解器分别计算带状注截面和圆形电子注截面上均匀分布的 75 个点的耦合阻抗,每个点的计算公式为

$$K_c = \frac{|E_{z1}|^2}{2\beta_1^2 P} \quad (3)$$

式中:  $|E_{z1}|$  为一次空间谐波在电子通道内相应点处轴向电场幅值;  $\beta_1$  为一次空间谐波在特定频率下的单位周期相移;  $P$  为通过慢波结构的总功率流.取这均匀分布的 75 个点的耦合阻抗的平均值

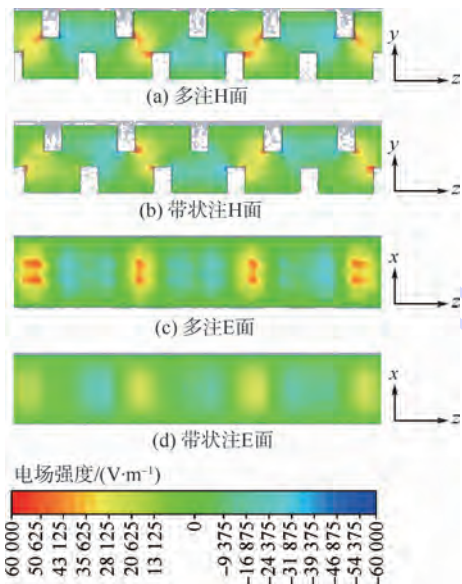


图 4 多注通道和带状注通道内 H 面和 E 面轴向电场分布

Fig. 4 Axial electric distribution on H-plane and

E-plane for multiple beam tunnel and sheet beam tunnel

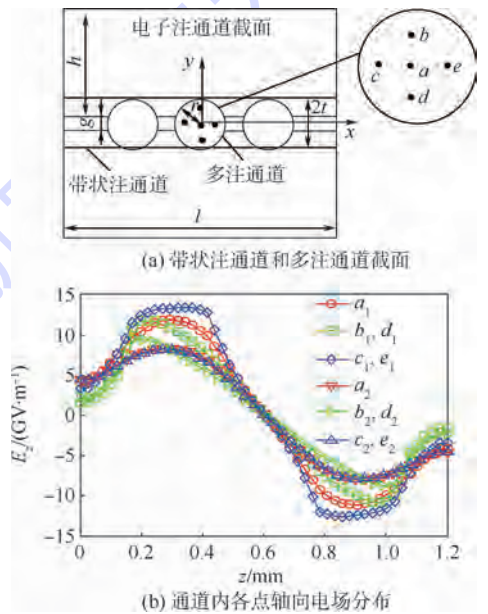


图 5 带状注通道和多注通道截面图以及各通道内轴向电场分布图

Fig. 5 Cross section of sheet beam tunnel and multiple beam tunnel and axial electric field distribution of both tunnels

就得到了高频结构的平均耦合阻抗. 计算结果如图6(a)所示,可以看出,在工作频带中心94 GHz处,多注通带内耦合阻抗为 $2.46 \Omega$ ,为带状注结构 $0.95 \Omega$ 的2.59倍,而在整个通带内,多注结构耦合阻抗近似为带状注结构的2~3倍.图6(b)也给出了这两种结构相对于各自中心频率点相速的归一化色散特性.

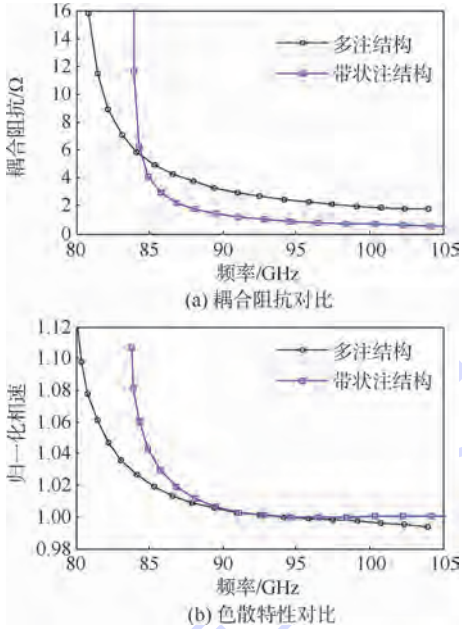


图6 带状注结构和多注结构耦合阻抗与色散特性对比图

Fig. 6 Comparison schematic on interaction impedance and dispersion characteristic for sheet beam and multiple beam structures

可以看出,多注系统有着和带状注系统近似相同的工作带宽,但其耦合阻抗却高于带状注系统,这将更有利于利用多电子注交错双栅行波放大器方案获得更高的注波相互作用效率和输出功率.

## 2 输入输出耦合系统的设计

为了保证高频信号在行波放大器输入和输出端口尽可能低反射损耗和传输损耗,需要对其输入与输出耦合结构进行详细的研究与设计,以减小其技术难度并降低工程实现复杂度.

赖剑强和 Shin 等<sup>[10,15]</sup>分别为带状注交错双栅行波放大器设计了一个5个周期的输入和输出耦合结构,设计思路是通过栅高、栅宽和槽深等几何参数的渐变过渡来达到高频场输入与输出时的良好匹配.但是,由于电磁场在电子枪区和与慢波结构连接处的带状注通道内不截止,因此他们不得不在耦合结构中设计一个连接器以防止高频场

泄露到电子枪区或收集极中,而这将大大增加高频结构的复杂度并严重影响耦合结构的传输特性.相比较而言,采用多注高频结构时,输入与输出的电磁场在电子注通道内完全截止,从而使得多注方案中的输入输出耦合结构在设计与工程实现上具有更为优良的技术优势.

基于前述的研究工作,本文从模式变换和结构渐变两种角度为三注交错双栅结构设计了一个更为简单的3个周期的输入和输出耦合结构,如图7所示.由分析可知,在交错双栅高频结构中传播的是 $z$ 向分量较强的TE和TM混合模式,为激起相应工作模式,应在交错双栅的E面上开槽并连接在输入矩形波导上,矩形波导中的 $TE_{10}$ 模的 $E_z$ 分量就自然转变为交错双栅结构中所传输的混合模式的 $E_z$ 分量.此后,通过渐变槽深 $dy = k_n - k_{n-1}$ 并调整输入波导端口宽度 $Z$ 就可以使输入波导和交错双栅达到良好的匹配.由于此耦合结构是通过矩形波导 $TE_{10}$ 模到交错双栅结构中混合模式的转变以实现匹配,所以矩形波导端口宽度 $Z$ 对匹配带宽和反射损耗 $S_{11}$ 的大小起着至关重要的作用,通过仿真设计得到的不同接口宽度 $Z$ 所对应的传输损耗、反射损耗及有效带宽如表1所示.可以看出,当输入矩形波导端口宽度 $Z = 0.50 \text{ mm}$ 、槽深渐变 $dy = 0.08 \text{ mm}$ 时,得到的有效带宽可以达到13 GHz左右.

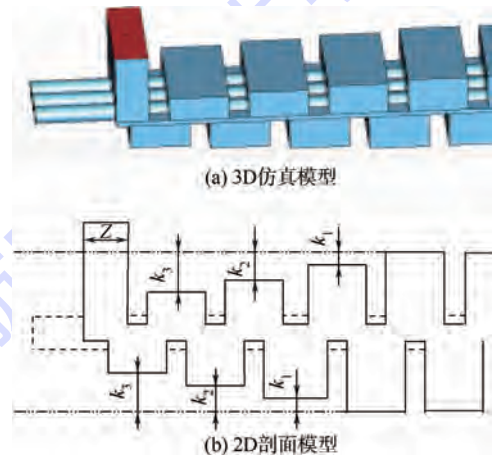


图7 输入输出耦合结构的3D模型和2D剖面模型  
Fig. 7 3D and 2D profile model of the input/output coupler structure

根据仿真设计和优化得到的输入与输出耦合结构渐变参数及其模型,本文将该模型与52个周期的交错双栅行波结构相连接,构建出了完整的行波放大器高频系统.对该完整的高频系统进行了详细的仿真设计与优化,得到的高频特性参数如图8所示.可以看出,通过简单的输入输出模式

变换和仅仅只有 3 个周期的交错双栅结构渐变, 所设计的耦合结构能在 85 ~ 98 GHz 较宽的频带内达到传输损耗  $S_{21} > -1$  dB 和反射损耗  $S_{11} < -20$  dB 的良好效果, 比前人所研究的带状注交错双栅模式变换和耦合结构更为简单和易于工程实现, 从而为 W 波段宽频带和高功率行波放大器的研究提供了新思路并奠定了基础。

表 1 不同接口宽度  $Z$  对应的传输损耗、反射损耗及有效带宽

Table 1 Transmission loss, return loss and effective bandwidth with different interface width of  $Z$

$Z/\text{mm}$	频率范围/GHz		有效带宽/GHz
	$S_{21} > -1$ dB	$S_{11} < -20$ dB	
0.40	80 ~ 100	81 ~ 83	2
0.45	81 ~ 101	83 ~ 88, 95 ~ 96	7
0.50	80 ~ 103	85 ~ 98	13
0.55	81 ~ 104	88 ~ 93	5
0.60	81 ~ 104	0	0

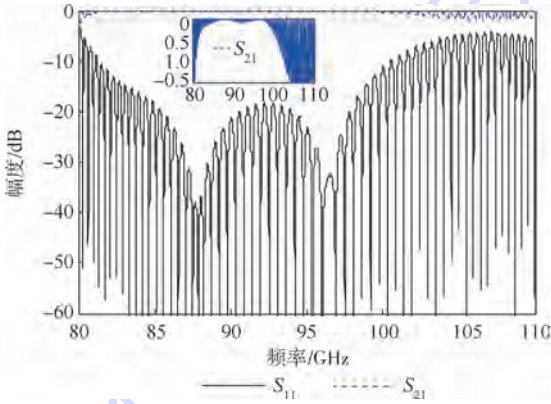


图 8 过渡结构传输特性

Fig. 8 Transmission characteristics of transition structure

### 3 结论

本文针对带状注交错双栅耦合阻抗较低、注波相互作用长度过长、输入输出耦合结构复杂以及带状注聚焦与长距离稳定传输困难等问题, 提出了一种平面线性排列的三电子注交错双栅行波放大器微型化高频系统结构, 通过在 W 波段对其进行的仿真设计和分析, 得到以下结论:

1) 该高频结构色散特性良好, 在 W 波段冷带宽可达 24 GHz, 有利于电子和电磁信号在较宽频带范围内进行持续的相互作用, 使器件具有很宽的频谱特性。

2) 该高频结构在电子注通道中存在很强的轴向电场, 在保持圆形电子注通道直径和带状注通道高度相同的情况下, 其通道轴线上的轴向电场强度可达带状注轴线上的 1.5 倍, 中心频率处的耦合阻抗可达  $2.46 \Omega$ 。在整个工作频带内,

该高频结构的耦合阻抗近似为带状注结构的 2 ~ 3 倍。这将带来更高的注波相互作用效率与饱和增益, 更大的输出功率和更短的注波相互作用长度。从而有望通过较短的行波高频结构在 W 波段获得宽频带和高功率相干辐射放大功率输出, 有效地解决带状注难以聚焦与长距离传输的难题。

3) 通过优化模式变换和渐变过渡结构, 设计了一种更为简单的输入与输出耦合器结构, 实现了通过 3 个周期渐变就可以在较宽频带达到良好的输入与输出匹配的目标。

本文提出的平面排列三电子注交错双栅行波放大器高频系统方案, 有望在 W 波段乃至太赫兹频段的宽频带高功率相干辐射放大器中得到广泛的应用。

### 参考文献 (References)

- [1] Gilmour A S. Principles of traveling-wave tubes [M]. Boston, MA: Artech House, 1994: 174-193.
- [2] Zhong H, Zhang C, Lin K I, et al. Large plasma-edge broadened magneto-optic-polar-Kerr-effect-based broadband incoherent detection of terahertz spectral frequencies [J]. Applied Physics Letters, 2005, 86(9): 091103-1 - 091103-3.
- [3] 张章. 太赫兹行波管的研究进展 [J]. 激光与红外, 2012, 42(4): 250-257.  
Zhang Z. Research progress of THz traveling wave tubes [J]. Laser & Infrared, 2012, 42(4): 250-257 (in Chinese).
- [4] Lawrence I R. Microfabrication of high-frequency vacuum electron devices [J]. IEEE Transactions on Plasma Science, 2004, 32(31): 1277-1292.
- [5] Gong Y B, Yin H R, Yue L N, et al. A 140-GHz two-beam overmoded folded-waveguide traveling-wave tube [J]. IEEE Transactions on Plasma Science, 2011, 39(3): 847-851.
- [6] Colin D J, Jeffrey P C, Garven M, et al. Microfabrication of a 220 GHz grating for sheet beam amplifiers [C] // IEEE International Conference on Plasma Science. Piscataway, NJ: IEEE Press, 2010: 187-188.
- [7] Shin Y M, Barnett L R. Intense wideband terahertz amplifier using phase shifted periodic electron-plasmon coupling [J]. Applied Physics Letters, 2008, 92(9): 091501-1-091501-3.
- [8] Shin Y M, Barnett L R, Luhmann N C. Phase-shift traveling-wave-tube circuit for ultra-wideband high-power sub-millimeter-wave generation [J]. IEEE Transactions on Electron Devices, 2009, 56(5): 706-712.
- [9] 刘青伦, 王自成, 刘濮鲲, 等. 基于场匹配法的双排矩形栅慢波结构高频特性研究 [J]. 物理学报, 2012, 61(24): 244102-1-244102-8.  
Liu Q L, Wang Z C, Liu P K, et al. Analysis of high frequency characteristics of the double-grating rectangular waveguide slow-wave-structure based on the field match method [J]. Acta Physica Sinica, 2012, 61(24): 244102-1-244102-8 (in Chinese).
- [10] Lai J Q, Gong Y B, Xu X, et al. W-band 1-kW staggered doub-

- le-vane traveling-wave tube[J]. IEEE Transactions on Electron Devices,2012,59(2):496-503.
- [11] Lai J Q,Wei Y Y,Liu Y,et al. A staggered double vane circuit for a W-band traveling-wave tube amplifier[J]. Chinese Physics B,2012,21(6):068403-1-068403-6.
- [12] Borsuk G M,Levush B. Vacuum electronics research perspective at the naval research laboratory[C]//IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press, 2010:3-4.
- [13] Li K,Liu W X,Wang Y,et al. Dispersion characteristics of two-beam folded waveguide for terahertz radiation[J]. IEEE Transactions on Electron Devices,2013,60(12):4252-4257.
- [14] Kreischer K E,Tucek J C,Basten M A,et al. 220 GHz power amplifier testing at Northrop Grumman[C]//IEEE International Vacuum Electronics Conference. Piscataway, NJ: IEEE Press,2013:1-2.
- [15] Shin Y M,Baig A,Barchfeld R,et al. Experimental study of multichromatic terahertz wave propagation through planar micro-channels [J]. Applied Physics Letters, 2012, 100 ( 15 ): 154103-1-154103-4.

## High frequency characteristics for W-band multiple beam staggered double-vane traveling wave tube amplifier

ZHANG Muwu<sup>1,2</sup>, RUAN Cunjun<sup>\*3</sup>

(1. Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China;

3. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** A new micro-structure scheme was proposed by using three parallel pencil beams instead of the sheet beam for the W-band staggered double vane traveling wave tube amplifier, and detailed analysis and optimization were given for the high frequency characteristic with its input/output couplers. The results show that the structures have a good dispersion characteristic and very broad working bandwidth, the strong component of longitudinal electric field is very benefit to the interaction and exchange of the energy between the electron beam and the high frequency field. Keep the diameter of the pencil beam tunnel as the same of the height of the sheet beam tunnel, the interaction impedance of the structures can achieve 2 – 3 times as high as the sheet beam scheme to obtain the high interaction efficiency and high output power. In order to match the staggered double-vane high frequency system, we present a more simple and easy input/output coupler structure. Just using three periods of transition structures, the input/output structures of the traveling wave tube amplifier can achieve a low reflection loss below the -20 dB in a broad bandwidth with more suitable application in the devices later.

**Key words:** W-band; staggered double-vane; multiple electron beam; traveling wave tube amplifier; coupler structure

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0776



# 航空电子设备 NSEE 试验评价方法

薛海红<sup>1</sup>, 王群勇<sup>2</sup>, 陈冬梅<sup>\*2</sup>, 陈宇<sup>2</sup>, 阳辉<sup>2</sup>, 李红军<sup>2</sup>

(1. 中航工业第一飞机设计研究院, 西安 710089; 2. 北京圣涛平试验工程技术研究院有限责任公司, 北京 100089)

**摘 要:** 本文定义了航空电子设备大气中子单粒子效应(NSEE)硬失效、硬错误与软错误等大气中子辐射可靠性表征参数,提出了一套航空电子设备抗大气中子单粒子效应危害能力的试验评价方法,包括试验应力终止条件与起始条件的确定方法、 $M_{NSEE}$  预估方法、试验程序、试验评价方法等,并通过某航空用卫星导航接收机试验案例证明了该方法在工程应用中的可操作性与有效性,通过掌握数字信号处理(DSP)、静态随机存储器(SRAM)、现场可编程门阵列(FPGA)的NSEE敏感特性,可为航空用卫星导航接收机NSEE危害防控提供针对性的技术支持. 本文为我国航空电子产品大气中子单粒子效应试验评价奠定了理论与应用基础,同时也为其他电子信息产品在大气中子单粒子效应试验评价方面提供了重要依据和参考.

**关键词:** 大气中子单粒子效应(NSEE); NSEE 试验评价方法;  $M_{NSEE}$  预估方法; 辐射可靠性; 航空用卫星导航接收机; 硬失效; 硬错误; 软错误

**中图分类号:** V21

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1894-08

单粒子效应是指单个高能粒子作用于半导体器件引发的翻转、锁定、烧毁、栅穿等现象<sup>[1]</sup>. 以往研究认为单粒子效应主要发生在航天领域,但国外近年来的研究表明,各类飞机在3~20 km的自然空间环境中,同样会遭遇单粒子效应. 只不过,航天领域诱发单粒子效应的高能粒子辐射源主要为重离子或质子,而在航空领域诱发单粒子效应的高能粒子辐射源主要为大气中子. 大气中子辐射应力一般用注量率来表征,其国际典型值表述为:在高度12.16 km、纬度45°、能量10 MeV以上的大气中子每平方米每小时6 000个. 在高度12.16 km、纬度45°、能量1 MeV以上的大气中子每平方米每小时9 200个<sup>[1]</sup>.

这些大气中子穿透力强<sup>[2]</sup>,金属材料几乎没有阻挡作用,因此会穿透机舱蒙皮,打在航空电子设备的核心指令控制单元或关键数据存储单元上,产生大气中子单粒子效应( Neutron Single

Event Effect, NSEE),从而引发软错误、硬错误甚至硬失效,导致飞行控制系统、航电系统等出现黑屏、死机、复位、重启、数据丢失、命令错误等故障现象. 有的会直接影响飞机的安全性与可靠性<sup>[3]</sup>,有的则会误导飞机驾驶员产生错误判断与错误操作,从而间接影响飞机的安全性与可靠性.

NSEE对复杂航空电子系统的危害影响,当前在国际航空领域内已经达成共识. 美国联邦航空局<sup>[4]</sup>、欧洲航空安全局<sup>[5]</sup>、波音<sup>[6]</sup>、空客<sup>[7]</sup>、GE、Honeywell、EMBRAER等均认为,自1992年首次证明飞机航空电子系统对宇宙射线诱发的软错误敏感以来,自然环境大气辐射效应已经成为航空电子系统安全性与可靠性的关键影响因素. 各国政府与用户规范逐渐增加了大气中子诱发航空电子设备单粒子效应率<sup>[8]</sup>的评价要求,已经将NSEE防护设计和试验评价要求具体纳入了工程化标准<sup>[9-12]</sup>或适航要求<sup>[13-15]</sup>.

收稿日期: 2014-12-11; 录用日期: 2015-02-25; 网络出版时间: 2015-03-31 14:54

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150331.1454.003.html

作者简介: 薛海红(1976—),男,陕西绥德人,高级工程师, xuehh0601@163.com

\* 通讯作者: 陈冬梅(1969—),女,重庆人,高级工程师, chendm@stea.cn, 主要研究方向为试验工程.

**引用格式:** 薛海红, 王群勇, 陈冬梅, 等. 航空电子设备 NSEE 试验评价方法[J]. 北京航空航天大学学报, 2015, 41(10): 1894-1901. Xue H H, Wang Q Y, Chen D M, et al. Neutron single event effects testing and evaluation method for avionics[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1894-1901 (in Chinese).

我国航空界在飞行器设计过程中,从航空电子设备可靠性及安全性角度,虽然提出了 NSEE 防护设计与试验评价的要求,但是因研究基础较弱,目前国内尚无相应的 NSEE 试验评价方法来支撑相应的设计工作。

本文提出了一套较为完整的航空电子设备抗 NSEE 危害能力的试验评价方法.并以某航空用卫星导航接收机为例,验证了此试验评价方法的可操作性和有效性。

## 1 航空电子设备 NSEE 故障类型及试验评价方法

### 1.1 故障类型

航空电子设备 NSEE 故障类型定义如下:

第 1 类故障为硬失效,其特征为断电重启不能恢复正常功能,需要维修。

第 2 类故障为硬错误,其特征为断电重启才能恢复正常功能。

第 3 类故障为软错误,其特征为软复位可恢复正常功能。

### 1.2 试验评价方法

#### 1.2.1 目的

航空电子设备 NSEE 试验评价的主要目的是获得航空电子设备 NSEE 故障率或平均 NSEE 故障间隔时间,从而支撑判断 NSEE 对航空电子设备安全性与可靠性的危害。

#### 1.2.2 航空电子设备 NSEE 故障率计算方法

参考 IEC 62396-1《航空电子设备大气中子单粒子效应对策策略》,航空电子设备 NSEE 故障率计算方法如下:

$$\lambda_{\text{NSEE}} = \lambda_{\text{NSEE-HF}} + \lambda_{\text{NSEE-HE}} + \lambda_{\text{NSEE-SE}} \quad (1)$$

式中: $\lambda_{\text{NSEE-HF}}$  为航空电子设备 NSEE 硬失效效率,个/h; $\lambda_{\text{NSEE-HE}}$  为航空电子设备 NSEE 硬错误率,个/h; $\lambda_{\text{NSEE-SE}}$  为航空电子设备 NSEE 软错误率,个/h。

#### 1.2.3 航空电子设备平均 NSEE 故障间隔时间计算方法

航空电子设备平均 NSEE 故障间隔时间计算方法如下:

$$M_{\text{NSEE}} = \frac{1}{\lambda_{\text{NSEE}}} = \frac{1}{\lambda_{\text{NSEE-HF}} + \lambda_{\text{NSEE-HE}} + \lambda_{\text{NSEE-SE}}} \quad (2)$$

#### 1.2.4 航空电子设备三类 NSEE 故障的平均故障间隔时间计算方法

航空电子设备三类 NSEE 故障的平均故障间隔时间计算方法如下:

$$M_{\text{NSEE-}i} = \frac{1}{\lambda_{\text{NSEE-}i}} \quad (3)$$

式中: $i$  分别表示 HF(硬失效)、HE(硬错误)、SE(软错误)。

#### 1.2.5 航空电子设备三类 NSEE 故障率计算方法

航空电子设备三类 NSEE 故障率计算方法如下:

$$\lambda_{\text{NSEE-}i} = f_{\text{NSEE-}i} \times \sigma_{\text{NSEE-}i} \quad (4)$$

式中: $f_{\text{NSEE-}i}$  为任务环境下诱发航空电子设备相应 NSEE 故障的大气中子注量率,个/( $\text{cm}^2 \cdot \text{h}$ ),通常采用国际典型值 6 000 个/( $\text{cm}^2 \cdot \text{h}$ ); $\sigma_{\text{NSEE-}i}$  为航空电子设备相应 NSEE 故障的敏感截面, $\text{cm}^2/\text{板}$  或  $\text{cm}^2/\text{Mbit}$ 。

#### 1.2.6 航空电子设备三类 NSEE 故障敏感截面计算方法

航空电子设备三类 NSEE 故障敏感截面计算方法如下:

$$\sigma_{\text{NSEE-}i} = \frac{N_{\text{NSEE-}i}}{F_{\text{NSEE-}i} \times N_{\text{s NSEE-}i}} \times A_{\text{NSEE-}i} \quad (5)$$

式中: $N_{\text{NSEE-}i}$  为整个试验过程中记录的航空电子设备相应 NSEE 故障总数量,个; $F_{\text{NSEE-}i}$  为整个试验过程中记录的诱发航空电子设备相应 NSEE 故障的累积中子注量,个/ $\text{cm}^2$ ; $N_{\text{s NSEE-}i}$  为整个试验过程中记录的诱发航空电子设备相应 NSEE 故障的试验件样品数量,个或 Mbit; $A_{\text{NSEE-}i}$  为在 14 MeV 中子源辐照下航空电子设备相应 NSEE 故障的敏感截面与在真实飞行任务环境辐照下相应 NSEE 故障的敏感截面的差异修正系数,无量纲,一般情况下取 1。

### 1.3 试验方法

#### 1.3.1 目的

航空电子设备 NSEE 试验的主要目的是获得航空电子设备三类 NSEE 故障的敏感截面  $\sigma_{\text{NSEE-}i}$ 。

#### 1.3.2 试验应力终止条件计算方法

参考 IEC 62396-2《航空电子系统单粒子效应试验指南》,地面模拟试验应当证明航空电子设备在整个寿命周期能够经受住任务环境大气中子辐射的危害影响,因此,地面模拟试验应力应当远远高于航空电子设备整个寿命周期所遭受的大气中子辐射注量。

航空电子设备 NSEE 试验应力终止条件计算方法如下:

$$F_{\text{T}} \gg F_{\text{M}} = T \times f \quad (6)$$

式中: $F_{\text{T}}$  为航空电子设备 NSEE 地面模拟试验应力,也即整个试验过程中给航空电子设备施加的累积中子注量,个/ $\text{cm}^2$ ; $F_{\text{M}}$  为航空电子设备整个

寿命周期所遭受的大气中子辐射累积注量或辐射应力,个/cm<sup>2</sup>;T为航空电子设备的寿命周期,飞行小时;*f*为航空电子设备在整个寿命周期在飞行任务环境下的平均大气中子注量率,个/(cm<sup>2</sup>·h),通常采用国际典型值6000个/(cm<sup>2</sup>·h).

1.3.3 试验应力起始条件确定方法

初始中子注量率可在10<sup>4</sup>~10<sup>6</sup> n/(cm<sup>2</sup>·s)附近调节,其目的是通过设定恰当而稳定的试验应力起始条件,使式(5)中在航空电子设备试验件样品数量*N<sub>s</sub>*已知固定的前提下,可观测记录的NSEE故障数*N<sub>NSEE-i</sub>*与试验施加的累积注量*F<sub>NSEE-i</sub>*(试验累积应力)初步形成线性关系,然后启动试验,通过具有统计意义的试验数据*N<sub>NSEE-i</sub>*获得*N<sub>NSEE-i</sub>*与累积注量*F<sub>NSEE-i</sub>*比较精确的线性关系,从而寻找到航空电子设备NSEE的本征敏感特性,即NSEE故障敏感截面σ<sub>NSEE-i</sub>.

航空电子设备NSEE试验应力起始条件计算公式如下:

$$F_0 = f_0 \times t_0 \quad (7)$$

式中:*F<sub>0</sub>*为地面模拟试验初始累积注量,个/cm<sup>2</sup>; *f<sub>0</sub>*为试验启动前调试确定的恰当而稳定的中子注量率,个/(cm<sup>2</sup>·h),通常调节范围为10<sup>4</sup>~10<sup>6</sup>个/(cm<sup>2</sup>·s); *t<sub>0</sub>*为试验调试时间,s,通常约为1000s.

航空电子设备NSEE试验应力起始条件确定方法是一个复杂的调试过程,简述如下:

- 1) 在10<sup>4</sup>~10<sup>6</sup>个/(cm<sup>2</sup>·s)附近调试出某一*f<sub>0</sub>*.
- 2) 由式(7)知在一定时间*t<sub>0</sub>*后,可以观察记录获得*F<sub>0</sub>*,即*F<sub>NSEE-i</sub>* = *F<sub>0</sub>*.

3) 由式(5)知可以观察获得*N<sub>NSEE-i</sub>*.

4) 向上或向下调试*f<sub>0</sub>*,观察航空电子设备试验件某些最为敏感的NSEE效应的*N<sub>NSEE-i</sub>*与*F<sub>NSEE-i</sub>*可以形成初步的线性关系时,稳定此时的*f<sub>0</sub>*;通常,这就是试验应力的初始条件.

5) 此时,可以将*N<sub>NSEE-i</sub>*与*F<sub>NSEE-i</sub>*的计数均清零,重新启动试验,并重新从零开始记录相应*N<sub>NSEE-i</sub>*与*F<sub>NSEE-i</sub>*,进入正式试验.

1.3.4 试验时间计算方法

由

$$F_T = f_0 \times t_T \quad (8)$$

推导出试验时间计算方法如下:

$$t_T = \frac{F_T}{f_0} \quad (9)$$

式中:*t<sub>T</sub>*为试验时间,h;*F<sub>T</sub>*为航空电子设备NSEE地面模拟试验应力,也即整个试验过程中给航空电子设备施加的累积中子注量,个/cm<sup>2</sup>,即式(6)的计算结果;*f<sub>0</sub>*为第1.3.3节的结果.

1.3.5 *M<sub>NSEE</sub>* 预估

参考IEC 62396-1《航空电子设备大气中子辐射单粒子效应应对策略》中的评价原理,分析航空电子设备NSEE敏感器件清单,结合工程经验判断,考虑资源使用率,故障传递率,预估航空电子设备NSEE故障敏感截面σ<sub>NSEE-i</sub>,代入式(1)~式(4),预估航空电子设备NSEE故障率λ<sub>NSEE</sub>和平均NSEE故障间隔时间*M<sub>NSEE</sub>*,如图1所示.其中,由NSEE敏感器件的敏感截面汇总计算航空电子设备NSEE软错误敏感截面的计算方法如下(NSEE硬错误或NSEE硬失效敏感截面的计算方法可以参照计算):

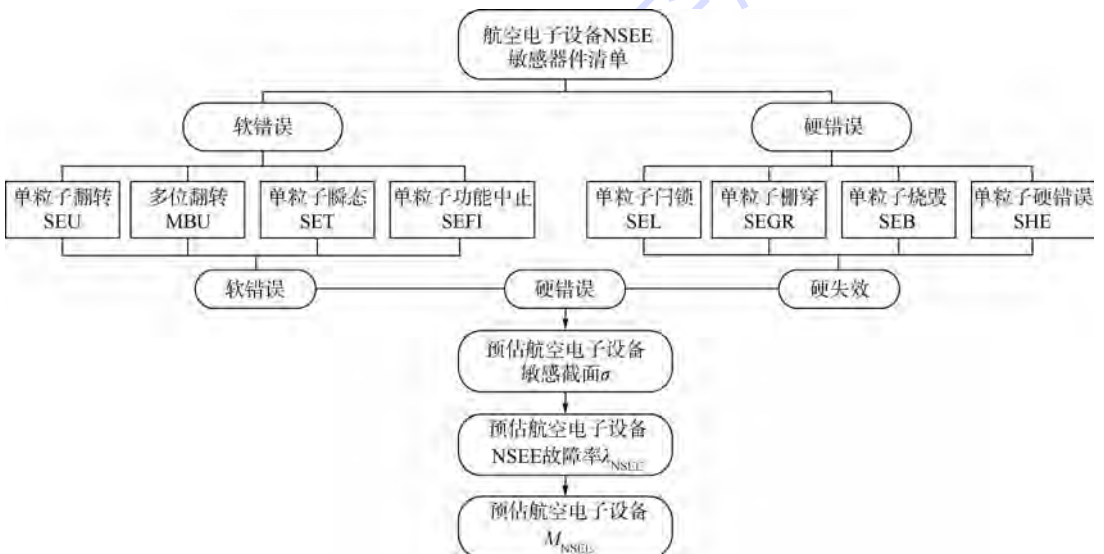


图1 航空电子设备 *M<sub>NSEE</sub>* 预估

Fig.1 Estimation of *M<sub>NSEE</sub>* of avionics

$$\sigma_{NSEE-SE} = \sum_{i=1}^n \sigma_i \times n_i \times \varepsilon_{i1} \times \varepsilon_{i2} \quad (10)$$

式中: $\sigma_i$ 为航空电子设备中诱发 NSEE 软错误的各个敏感器件的敏感截面, $\text{cm}^2/\text{器件}$ 或 $\text{cm}^2/\text{Mbit}$ ;  $n_i$ 为航空电子设备中诱发 NSEE 软错误的敏感器件的器件数量或 Mbit 数量; $\varepsilon_{i1}$ 为航空电子设备中诱发 NSEE 软错误的敏感器件的位资源利用率,无量纲,通常由用户提供; $\varepsilon_{i2}$ 为航空电子设备中诱发 NSEE 软错误的敏感器件将软错误从器件级传递至航空电子设备形成设备级软错误的传递率,无量纲,通常可按 10% 计算. 而 NSEE 硬错误与硬失效的传递率通常按 100% 计算.

### 1.3.6 NSEE 试验实施程序

NSEE 试验实施程序如下:

- 1) 试验系统搭建(包括:固定试验件位置、安置测试平台、连接试验件及测试设备).
- 2) 加载测试用例,进行联调联试,保证试验件能够正常工作.
- 3) 通知高压倍加器的操作者所需的初始中子注量率(初始值约为  $10^4$  个/ $(\text{cm}^2 \cdot \text{s})$ )左右,中子注量率调试期间,试验件暂不加电.
- 4) 在获得所需中子注量率后,试验件加电,开始辐照试验前调试.
- 5) 循环遍历测试用例,调出合适的中子注量

率,在注量率调试阶段出现的故障现象仅记录,作为调试数据,但不作为正式试验数据处理输入.

6) 注量记录仪清零,试验件断电重启并同步启动注量记录仪,重新累计注量并记录.

7) 辐照期间,循环遍历测试用例,如果试验件出现故障现象,必要时及时暂停辐照,进行分析并记录出现的故障现象及中子注量.

8) 故障处理分析和记录完成后,根据需要试验件断电重启,重复步骤 7),直至符合试验终止条件.

9) 按所确定的试验终止条件终止试验.

10) 依据式(5)进行试验数据分类处理,计算航空电子设备 NSEE 故障敏感截面  $\sigma_{NSEE-i}$ .

11) 将步骤 10) 试验计算结果替换预估值,代入本文第 1.3.5 节所述的相应计算流程,计算航空电子设备平均 NSEE 故障间隔时间  $M_{NSEE}$ .

## 2 试验案例

### 2.1 试验件

本文所选试验件为某航空用卫星导航接收机. 航空用导航接收机接受天线接收的卫星信号,经过下变频处理、数字基带处理、软件解算,可实时给出导航设备的位置和速度信息. 其功能框图如图 2 所示.

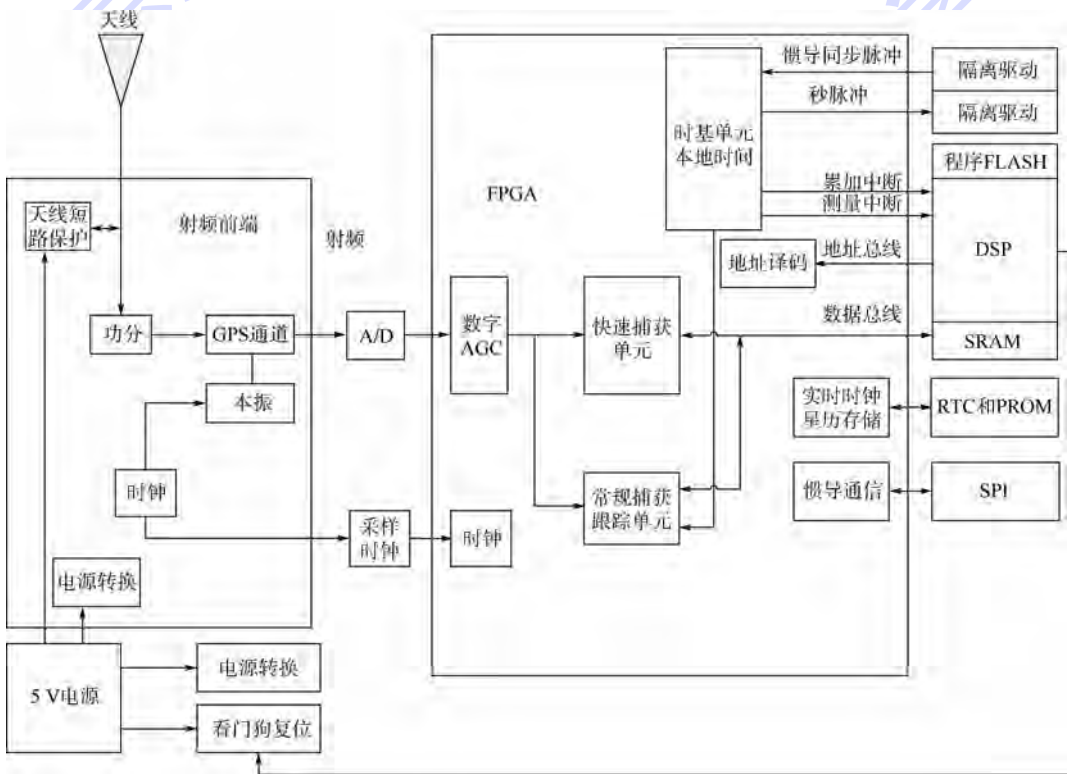


图 2 航空用卫星导航接收机功能框图

Fig. 2 Aircraft satellite navigation receiver function block diagram

## 2.2 试验应力确定

### 2.2.1 试验源

本试验采用国内 14 MeV 中子辐照源。

### 2.2.2 试验应力 $F_T$ 终止条件计算

试验件产品总寿命  $T$  为 30000 h, 大气中子注量率国际典型值  $f$  为 6000 个/( $\text{cm}^2 \cdot \text{h}$ ), 依据式(6), 计算得出

$$F_M = 30000 \times 6000 = 1.8 \times 10^8 \text{ 个}/\text{cm}^2$$

因此,  $F_T$  取  $1 \times 10^9$  个/ $\text{cm}^2$ 。

### 2.2.3 试验应力 $f_0$ 起始条件与试验时间 $t_T$ 确定

依据本文第 1.3.3 节确定试验应力起始条件  $f_0$ , 依据本文第 1.3.4 节确定本次试验时间  $t_T$ 。初始注量率调试到  $f_0$  时, NSEE 故障数与累积注量成初步线性关系, 然后, 启动开展正式试验。试验约至  $t_T$  时注意观察累积注量, 满足终止条件时, 停止试验。

## 2.3 $M_{\text{NSEE}}$ 预估

### 2.3.1 试验件 NSEE 敏感器件清单

经分析, 该试验件在 14 MeV 中子源辐射下,

DSP、SRAM、FPGA 为大气中子单粒子效应主要敏感器件, 各敏感器件在试验件中的主要功能分别为: DSP 主要完成软件解算功能; SRAM 完成软件数据和程序代码的存储功能; FPGA 完成卫星信号基带处理功能。

### 2.3.2 试验件 NSEE 敏感器件敏感截面 $\sigma_{\text{NSEE}}$ 估算

该试验件中 DSP、SRAM、FPGA 会产生单粒子翻转效应 (SEU), 其敏感截面典型值约为  $1 \times 10^{-14} \text{ cm}^2/\text{bit}$ ; 其中 DSP 与 FPGA 还会出现单粒子功能中止效应 (SEFI), 其敏感截面典型值约为  $1 \times 10^{-16} \text{ cm}^2/\text{bit}$ 。这些敏感器件产生的 SEU 与 SEFI 效应可能会导致卫星导航接收机的部分或全部功能丧失。

### 2.3.3 试验件 NSEE 敏感截面 $\sigma_{\text{NSEE-i}}$ 估算

依据本文第 1.3.5 节, 预估试验件板级试验中由于器件级 SEU 引发的板级软错误敏感截面约为  $6.750 \times 10^{-8}/\text{板}$ , 由于器件级 SEFI 引发的板级硬错误敏感截面约为  $6.278 \times 10^{-9}/\text{板}$ , 如表 1 所示。

表 1 NSEE 主要敏感器件清单

Table 1 List of main atmospheric NSEE sensitive devices

序号	器件	器件数量	试验测试比特位数	主要敏感效应	预估敏感截面/( $\text{cm}^2 \cdot \text{bit}^{-1}$ )	资源利用率/%	故障传递率/%	预估敏感截面( $\text{cm}^2 \cdot \text{器件}^{-1}$ )
1	DSP	1	RAM 8388608	SEU	$1 \times 10^{-14}$	100	10	$8.389 \times 10^{-9}$
			整个器件	SEFI	$1 \times 10^{-16}$	100	100	$8.389 \times 10^{-10}$
2	SRAM	1	RAM 18874368	SEU	$1 \times 10^{-14}$	25	10	$4.719 \times 10^{-9}$
			CLB 33909664	SEU	$1 \times 10^{-14}$	70	10	$2.374 \times 10^{-8}$
3	FPGA (1#)	1	BRAM 4939776	SEU	$1 \times 10^{-14}$	70	10	$3.458 \times 10^{-9}$
			整个器件	SEFI	$1 \times 10^{-16}$	70	100	$2.719 \times 10^{-9}$
4	FPGA (2#)	1	CLB 33909664	SEU	$1 \times 10^{-14}$	70	10	$2.374 \times 10^{-8}$
			BRAM 4939776	SEU	$1 \times 10^{-14}$	70	10	$3.458 \times 10^{-9}$
	试验件合计	4	4 个器件约 100 Mbit	SEU 导致软错误	$1 \times 10^{-14}$	64	10	$6.750 \times 10^{-8}/\text{板}$
			3 个器件约 82 Mbit	SEFI 导致硬错误	$1 \times 10^{-16}$	73	100	$6.278 \times 10^{-9}/\text{板}$

### 2.3.4 试验件 NSEE 故障率 $\lambda_{\text{NSEE-i}}$ 估算

依据本文式(4), 采用大气中子注量率国际典型值 6000 个/( $\text{cm}^2 \cdot \text{h}$ ), 预估该试验件在整个寿命周期 NSEE 软错误率约为  $4.050 \times 10^{-4}/\text{板}$ , NSEE 硬错误率约为  $3.767 \times 10^{-5}/\text{板}$ , NSEE 软硬错误率约为  $4.427 \times 10^{-4}/\text{板}$ 。

### 2.3.5 试验件平均 NSEE 故障间隔时间 $M_{\text{NSEE-i}}$ 估算

依据本文式(3), 试验件平均 NSEE 软错误间隔时间  $M_{\text{NSEE-SE}}$  为 2469 h; 试验件平均 NSEE 硬错误间隔时间  $M_{\text{NSEE-SE}}$  为 26546 h。因此, 试验件平均 NSEE 软硬错误间隔时间  $M_{\text{NSEE}}$  为 2259 h。

## 2.4 试验实施步骤

NSEE 试验实施步骤如下:

1) 试验系统搭建: 试验系统主要由 14 MeV 中子加速器、屏蔽板、辐照电路板、试验件供电电源、供电及输入输出信号电缆、测试和控制系统等组成。辐照前, 将试验件固定在靶前至少 80 cm 处, 试验件垂直于试验件中心与靶心的连线; 在靶与试验件之间竖立一块 4 mm 厚的铝屏蔽板, 模拟真实的机舱蒙皮厚度; PC 机、电源放置在测试间内; 试验件与 PC 机、电源之间通过 25 m 电缆连接, 试验系统布局如图 3 所示。

2) 试验环境: 试验环境温度为 15 ~ 35℃; 相对湿度为 20% ~ 80%; 标准环境大气压约为 760 mm 汞柱。

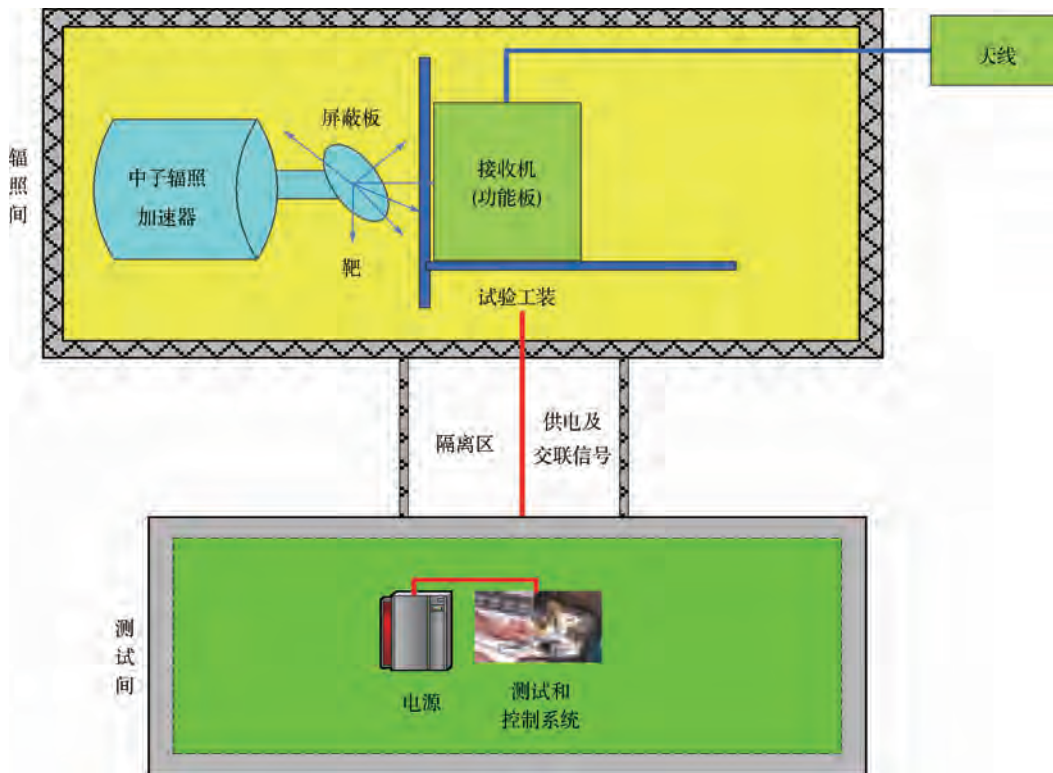


图 3 试验构架示意图

Fig.3 Testing architectural layout schematic

3) 开展联调联试:测试设备交联关系如图 4 所示. 重点关注以下事项:①故障名称与故障分类的定义确认;②测试用例的版本与标识;③测试用例的资源利用率、故障传递率的初步分析与判断.

4) 依据本文第 1.3.6 节开展试验. 当注量达到  $10^9$  个/cm<sup>2</sup> 时终止辐照, 依据式(5), 经试验数据处理可得出试验件 NSEE 软硬错误敏感截面  $\sigma$ , 结果详见表 2, 数据拟合结果如图 5 所示, NSEE 软硬错误数与累积中子注量成良好线性关系, 验证了本试验方法的有效性.

2.5  $M_{NSEE}$  的试验数据计算结果

由表 2 可见, 试验件 NSEE 硬错误率为  $4.154 \times 10^{-5}$  次/(板·h), NSEE 软错误率为  $1.246 \times 10^{-4}$  次/(板·h), NSEE 软硬错误率为  $1.661 \times 10^{-4}$  (次/板·h). 试验件平均 NSEE 硬错

误间隔时间  $M_{NSEE-HE}$  为 24073 h; 试验件平均 NSEE 软错误间隔时间  $M_{NSEE-SH}$  为 8024 h; 试验件平均软错误间隔时间  $M_{NSEE}$  为 6018 h.

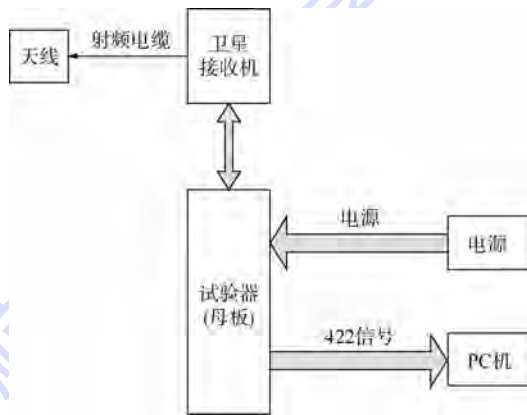


图 4 测试设备交联关系示意图

Fig.4 Testing equipment connection overview schematic

表 2 试验件 NSEE 试验数据计算结果

Table 2 Calculated NSEE test results for tested board

序号	故障现象	故障次数/次	累积注量/( $10^9$ 个·cm <sup>-2</sup> )	试验件样品数量/个	板级敏感截面/( $10^{-9}$ cm <sup>2</sup> ·板 <sup>-1</sup> )	故障类别	故障次数	任务环境中子注量率/(个·(cm <sup>2</sup> ·h) <sup>-1</sup> )	故障率/( $10^{-5}$ 次·(板·h) <sup>-1</sup> )	平均故障间隔时间/h
1	死机	7	1.011	1	6.924	硬错误	7	6000	4.154	24073
2	无定位	18	1.011	1	17.804	软错误	21	6000	12.463	8024
3	定位超差	3	1.011	1	2.967	软硬错误	3	6000	0.5	2000
合计		28	1.011	1	27.695	软硬错误	28	6000	16.617	6018

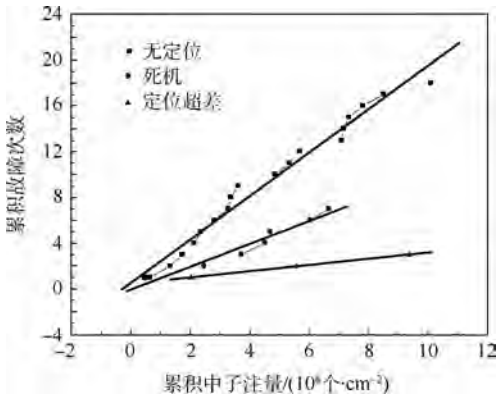


图5 试验件 NSEE 试验拟合结果

Fig. 5 Test fitting results for NSEE tested board

与表 1 中预计结果相比,硬错误结果差异较小,主要差异在于软错误的差异.有可能是试验期间实际资源利用率偏小,这有待于进一步研究.3 倍以内的误差在国际上已经是比较理想的数据.

### 3 结论

本文提出了一套航空电子设备抗大气中子单粒子效应危害能力的试验评价方法,即航空电子设备大气中子辐射可靠性参数  $M_{NSEE}$  的计算方法,并通过某航空用卫星导航接收机试验案例证明了该方法在工程应用中的可操作性与有效性,具体结论如下:

1) 依据试验实施程序,可以分类获得航空电子设备 NSEE 硬错误、软错误的原始数据;依据式(5),可以获得航空电子设备 NSEE 故障敏感截面  $\sigma_{NSEE}$ .

2) 依据第 1 节所述的航空电子设备 NSEE 试验评价方法可以计算获得航空电子设备平均 NSEE 故障间隔时间  $M_{NSEE}$ .

3) 依据航空电子设备  $M_{NSEE}$  预估方法第 1.3.5 节计算结果与依据航空电子设备  $M_{NSEE}$  试验评价方法第 1.3.6 节计算结果在 3 倍以内.

本文  $M_{NSEE}$  预估方法中资源利用率与故障传递率是两个比较复杂的参数,试验实施程序中的监测方法也是一项复杂的影响因素,均需要进一步研究以提高  $M_{NSEE}$  预估结果与试验评价结果的一致性.

致谢 感谢北京圣涛平试验工程技术研究院

航空电子设备大气中子单粒子效应数据库维护组白桦、孙旭朋、罗星等的大力帮助.

### 参考文献 (References)

- [1] IEC62396-1. Process management for avionics: Atmospheric radiation effects. Part 1: Accommodation of atmospheric radiation effects via single event effects within avionics electronic equipment[S]. Geneva: IEC, ,2012:14,20.
- [2] Slayman C. Acceleration factors for neutron beam soft error testing in commercial and avionic environments [M]. Piscataway, NJ: IEEE Press, 2010:11.
- [3] Matthews DC, Dion MJ. NSEU impact on commercial avionics [M]. Cedar Rapids: Rockwell Collins, Inc. ,2009:181.
- [4] Vranish K. The growing impact of atmospheric radiation effects on semiconductor devices and the associated impact on avionics suppliers [M]. Kentwood: KVA Engineering, Inc. ,2011:4-5.
- [5] EASA Proposed CM-AS-004. Single event effects (SEE) caused by atmospheric radiation[S]. Brussels: European Aviation Safety Agency, 2014:10.
- [6] Normand E. Single event effects in avionics [M]. Seattle: Boeing Defense & Space Group 1995:1.
- [7] Colas G. Neutrons atmosphériques en avionique [M]. Vélizy: Thalès Avionics, 2002:22.
- [8] IEC 62239. Process management for avionics: Preparation of an electronic components management plan [S]. Geneva: IEC, Edition 2.0, 2008:14.
- [9] IEC62396-2. Process management for avionics: Atmospheric radiation effects. Part 2: Guidelines for single event effects testing for avionics systems[S]. Geneva: IEC, 2008:1-27.
- [10] IEC62396-3. Process management for avionics: Atmospheric radiation effects. Part 3: Optimizing system design to accommodate the single event effects (SEE) of atmospheric radiation [S]. Geneva: IEC, 2008:1-25.
- [11] IEC 62396-4. Process management for avionics: Atmospheric radiation effects. Part 4: Guidelines for designing with high voltage aircraft electronics and potential single event effects [S]. Geneva: IEC, 2008:1-16.
- [12] IEC 62396-5. Process management for avionics: Atmospheric radiation effects. Part 5: Guidelines for assessing thermal neutron fluxes and effects in avionics systems [S]. Geneva: IEC, 2008:1-18.
- [13] RTCA/DO-297. Integrated modular avionics (IMA) design guidance and certification considerations [S]. Washington, D. C. : RTCA Inc. ,2005:33.
- [14] RTCA/DO-254. Design assurance guidance for airborne electronic hardware [S]. Eurocae: RTCA Inc. ,2000:15,85,106.
- [15] RTCA/DO-178C. Software considerations in airborne systems and equipment certification [S]. Washington, D. C. : RTCA Inc. ,2011:32,72,119.

## Neutron single event effects testing and evaluation method for avionics

XUE Haihong<sup>1</sup>, WANG Qunyong<sup>2</sup>, CHEN Dongmei<sup>\*2</sup>, CHEN Yu<sup>2</sup>, YANG Hui<sup>2</sup>, LI Hongjun<sup>2</sup>

(1. AVIC the First Aircraft Institute, Xi'an 710089, China;

2. Beijing San-talking Testing Engineering Academy, Co. Ltd, Beijing 100089, China)

**Abstract:** The atmospheric neutron radiation reliability parameters, such as hard failure, hard error and soft error for avionics neutron single event effect (NSEE) were defined. A set of testing and evaluation method to demonstrate the avionics NSEE-resistance capability was presented and verified by a case study on some aircraft satellite navigation receiver NSEE testing and evaluation. The testing and evaluation method included test stress determination method on ending condition and starting condition,  $M_{\text{NSEE}}$  prediction method, test procedures, testing and evaluation method, etc. This case study verifies that the NSEE testing and evaluation method for avionics is practical, operational and effective in the engineering application. The NSEE sensitivity features of DSP, SRAM and FPGA in aircraft satellite navigation receiver may provide pinpointed technical support for NSEE mitigation and protection for the related NSEE failure. The theoretical and application foundation is paved for atmospheric neutron single event effects testing and evaluation for avionics in China. In the meantime, it provides important and valuable reference for the atmospheric NSEE testing and evaluation for other electronic and information products in China.

**Key words:** atmospheric neutron single event effect (NSEE); NSEE testing and evaluation method;  $M_{\text{NSEE}}$  prediction method; radiation reliability; aircraft satellite navigation receiver; hard failure; hard error; soft error

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0030

# 融合粗糙集与 D-S 证据理论的航空装备故障诊断



孙伟超<sup>1</sup>, 李文海<sup>\*2</sup>, 李文峰<sup>1</sup>

(1. 海军航空工程学院 研究生管理大队, 烟台 264001; 2. 海军航空工程学院 科研部, 烟台 264001)

**摘 要:** 针对航空电子装备故障诊断中出现的多源诊断信息存在冲突的情况, 基于粗糙集与证据理论在处理不确定问题时的优势, 提出了一种融合粗糙集与证据理论的故障诊断方法. 该方法利用粗糙集将信息源给出的诊断数据转化为证据理论中的 mass 函数, 进行结果融合. 同时, 该方法给出边界粗糙熵的定义, 并基于边界粗糙熵获得反映各信息源在诊断融合过程中重要度的动态权重参数, 提出一种新的证据理论的冲突合成规则. 仿真实验表明, 该方法可以有效地提升诊断信息融合结果的准确性, 在航空电子装备故障诊断方面有较好的实用价值.

**关键词:** 边界粗糙熵; 粗糙集; D-S 证据理论; 冲突证据融合; 故障诊断

**中图分类号:** TP181; V24

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1902-08

随着航空电子装备的复杂程度不断提高, 其故障模式呈现出多样化与复杂化的特点, 现有检测手段大多都利用单一信息源对航空电子装备某一类故障进行诊断, 缺乏多源多维信息综合处理能力, 在诊断的可靠性和实用性方面也都存在不同程度的缺陷. 因此, 需要采用多传感器数据进行融合, 进而提升诊断的准确度.

近年来, 多源信息融合技术得到了迅速发展<sup>[1]</sup>, 已广泛应用于信息安全、图像处理、传感器网络、智能系统等多个领域<sup>[2]</sup>. 其研究主要集中于信息融合模型、动态融合方法、数据的预处理、冲突数据合成、数据维数转换、数据解相关等方向. 在不同的研究方向上, 国内外学者提出了多种融合算法, 其中包括模糊推理、贝叶斯估计、卡尔曼滤波、粗糙集、D-S (Dempster-Shafer) 证据理论、DSmT (Dezert-Smarandache 理论)、随机集以及基于人工智能的方法等. 目前, 软数据/硬数据融

合<sup>[3-4]</sup>、自适应融合<sup>[5-6]</sup>以及多融合算法的联合应用成为信息融合新的研究方向.

在信息融合算法中, 粗糙集理论可以有效地对不确定、不完备数据进行处理, 目前广泛应用于模式识别、机器学习及故障诊断等领域. 其无需任何先验知识, 通过数据集本身的不可分辨关系, 即可消除冗余属性, 完成数据分类与特征提取. 属性约简作为粗糙集理论的核心, 目前大量研究围绕其展开, 文献[7-10]给出了不同数据结构下的多种属性约简算法. 在未来一段时期内, 高效约简算法、海量数据处理及粗糙集与其他算法的结合使用仍会作为粗糙集理论研究的重点. 在故障诊断应用中, 粗糙集多用于建立诊断数据与故障类型的映射关系, 对高维故障数据进行降维处理, 减少后续计算复杂度. 文献[11-13]将粗糙集与神经网络结合, 减少了网络的输入维数, 提高了神经网络的训练速度; 文献[14]使用支持向量机(SVM)对

收稿日期: 2015-01-15; 录用日期: 2015-04-17; 网络出版时间: 2015-06-24 16:46

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150624.1646.008.html

作者简介: 孙伟超(1986—), 男, 山东烟台人, 博士研究生, ben\_phoenix@163.com

\* 通讯作者: 李文海(1969—), 男, 山东无棣人, 教授, ythylwh@vip.163.com.cn, 主要研究方向为复杂装备故障诊断.

**引用格式:** 孙伟超, 李文海, 李文峰. 融合粗糙集与 D-S 证据理论的航空装备故障诊断[J]. 北京航空航天大学学报, 2015, 41(10): 1902-1909. Sun W C, Li W H, Li W F. Avionic devices fault diagnosis based on fusion method of rough set and D-S theory[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1902-1909 (in Chinese).

经粗糙集处理后的数据进行分类;文献[15]则将粗糙集与灰色理论相结合提出新的模型对故障进行预测.粗糙集与模糊理论以及多种优化算法的结合应用在此不再赘述.

D-S 证据理论作为决策级信息融合算法因其具有对不确定信息表示、度量与组合的能力,在多属性决策问题中得到广泛的应用.为了解决数据存在冲突的问题,目前针对证据理论的改进也在不断进行,主要集中于两方面,通过改变冲突部分的分配方式对组合规则进行改进<sup>[16]</sup>以及通过对各证据权重的计算对证据本身进行修正<sup>[17-19]</sup>.故障诊断作为一种多属性决策问题,是 D-S 证据理论的主要应用方向之一.文献[20]将其用于齿轮箱的故障诊断并引入信息熵对证据进行修正,文献[21]在故障诊断过程中通过对证据权重进行动态修正得到更为精确的结果;而更为常见的是 D-S 证据理论与其他算法的联合应用,与粗糙集不同的是,证据理论在其中负责诊断结果的融合.文献[22]将神经网络的输出作为证据进行合成,对涡轮机故障进行诊断;文献[23]用 SVM 构建分类器,对于多个分类器的结果使用 D-S 证据理论进行合成;文献[24]在对旋转机械进行并发故障诊断时用人工免疫算法对故障数据进行聚类,避免证据合成时焦元过多造成的计算困难.

由于航空电子装备诊断过程中采集到的多源信息存在冗余以及信息间存在冲突,考虑到粗糙集在提取关键信息与证据理论在冲突信息处理方面的优势,本文提出一种新的基于粗糙与证据理论的故障诊断方法.利用粗糙集构建故障诊断系统,减小属性维数;利用证据理论对约简后的测试结果进行融合.对于冲突证据的处理,在定义边界粗糙熵,作为在粗糙集框架下一种新的粗糙度衡量方法对证据进行修正,提出一种新的冲突证据合成规则,并将其应用于实际的诊断过程中.

## 1 故障诊断决策模型

### 1.1 基于粗糙集的数据处理

有序数组  $S = \{U, A, V, f\}$  表示一个信息系统.其中,  $U$  为关于对象的非空有限集合,  $A$  为属性的非空有限集合,  $V$  为属性的值域,  $f: U \times A \rightarrow V$  为信息函数,它为每一个对象的每个属性赋予一个属性值.

在信息系统  $S = \{U, A, V, f\}$  中,每个属性的子集  $B \subseteq A$  决定了一个不可分辨关系  $I(B)$ .

对于信息系统  $S = \{U, A, V, f\}$ ,若属性集可以被分为条件属性集  $C$  和决策属性集  $D$ ,即  $A = C \cup$

$D$  且  $C \cup D \neq \emptyset$ ,则该信息系统被称为一个决策系统.

在决策系统  $L = \{U, C \cup \{d\}, V, f\}$  中,若  $R \subseteq C$  满足:

$$1) P_{S(R)}(\{d\}) = P_{S(C)}(\{d\}).$$

2) 不存在  $r \in R$ ,使得  $P_{S(R-\{r\})}(\{d\}) = P_{S(R)}(\{d\})$  成立.

则称  $R$  为条件属性集  $C$  相对于决策属性  $d$  的约简,简称  $C$  的相对约简.

在决策系统  $L = \{U, C \cup \{d\}, V, f\}$  中,对于  $\forall a \in C$ ,对象  $x_i$  在属性  $a$  上的属性值频率为

$$p[f_a(x_i)] = \frac{|f_a(x_i)|}{|V_a|} \quad (1)$$

式中:  $|f_a(x_i)|$  为决策表中某一决策属性下与属性  $a$  列上取值为  $f_a(x_i)$  的相对应的对象个数;  $|V_a|$  为决策表中所有决策属性下与属性  $a$  列上取值为  $f_a(x_i)$  的相对应的对象个数.

### 1.2 基于证据理论的冲突融合

在 D-S 证据理论中,对于一个判决问题,所有可能答案的完备集合用  $\Theta$  表示.辨识框架  $\Theta = \{\theta_1, \theta_2, \dots, \theta_N\}$ .其中,  $\theta_i$  称为辨识框架  $\Theta$  的一个元素,  $N$  为元素个数,  $i = 1, 2, \dots, N$ .

在辨识框架  $\Theta$  下,基本信任分配函数是从集合  $2^\Theta$  到  $[0, 1]$  区间的映射,  $G$  表示识别框架  $\Theta$  的任一子集,记做  $G \subseteq \Theta$ ,并满足:  $m(\emptyset) = 0$  且  $\sum_{G \subseteq \Theta} m(G) = 1$ .其中,  $\sum_{G \subseteq \Theta} m(G)$  称为事件  $G$  的基本信任分配函数,也叫做 mass 函数,表示证据对  $G$  的信任程度.对于任一子集  $G$ ,只要有  $m(G) > 0$ ,则称  $G$  为焦元.

D-S 合成方法可以表述为,假设辨识框架  $\Theta$  下的两个证据  $E_{v_1}$  和  $E_{v_2}$ ,其相应的基本信任分配函数为  $m_1$  和  $m_2$ ,焦元分别为  $A_i$  和  $B_j$ ,则 D-S 合成规则为

$$m(G) = \begin{cases} \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - K} & A \neq \emptyset \\ 0 & A = \emptyset \end{cases}$$

式中:  $K = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)$ ,它反映出各证据之间的冲突程度.

## 2 航空电子装备故障诊断方法

### 2.1 基于边界粗糙熵的决策重要度计算

粗糙集可以对知识的不确定性进行有效刻画.知识的不确定性主要由两方面原因引起:一是来自论域上的二元关系,其产生知识模块越大,则

相对于近似空间的知识就越不确定,这种不确定性使用粗糙熵来刻画;另一原因来自于给定的近似空间的粗集边界,边界越大则知识越模糊,粗糙集引入粗糙度来实现测量.为了衡量每一属性的不确定程度,本文在考虑以上两方面原因的基础上提出边界粗糙熵的概念.

**定义 1** 给定一个决策系统  $L = \{U, C \cup \{d\}\}$ , 集合  $X \subseteq U, R$  是  $U$  上的等价关系, 定义两个子集:  $\underline{R}X = \{x \in U \mid [x]_R \subseteq X\}$ ,  $\bar{R}X = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$ . 其中,  $\underline{R}X$  和  $\bar{R}X$  分别称为  $X$  关于  $R$  的下近似和上近似. 定义集合  $B_{nR}(X) = \bar{R}X - \underline{R}X$  称为  $X$  的  $R$  边界域.

**定义 2** 对于决策系统  $L = \{U, C \cup \{d\}\}$ , 集合  $X \subseteq U, Y \subseteq U, R \subseteq C$ . 知识  $R$  对论域  $U$  的划分记为  $U/R = \{X_1, X_2, \dots, X_m\}$ , 决策  $\{d\}$  对论域  $U$  的划分记为  $U/\{d\} = \{Y_1, Y_2, \dots, Y_n\}$ . 如果记  $Y_i = d^{-1}(\{v_d\}) = \{k \in U \mid d(k) = v_d\}$ , 其中  $v_d$  为决策属性值, 则定义论域  $U$  在知识  $R$  下的划分对于决策分类  $Y_i$  的决策边界域为  $B_{nR}(Y_i) = \bar{R}Y_i - \underline{R}Y_i$ .

**定义 3** 在决策系统  $L = \{U, C \cup \{d\}\}$  中, 定义知识  $R$  的边界粗糙熵为

$$E_B(R) = \sum_{i=1}^n r_i \left[ \frac{|B_{nR}(Y_i)|}{|U|} \lg |U| \right] - \sum_{j=1}^m \frac{|X_j|}{|U|} \lg \frac{1}{|X_j|} \tag{2}$$

式中:  $r_i = |Y_i|/|U|$ ,  $r_i$  为等价类  $Y_i$  的基数与  $U$  的基数之比.

边界粗糙熵的定义公式同时考虑了由不可分辨关系  $R$  所引起的粗糙集不确定性的两个因素.

公式的前半部分  $\sum_{i=1}^n r_i \left[ \frac{|B_{nR}(Y_i)|}{|U|} \lg |U| \right]$  表示决策表中的对象集合在  $R$  划分下相对于决策类产生的粗集边界所引起的知识不确定性, 公式的后半部分  $-\sum_{j=1}^m \frac{|X_j|}{|U|} \lg \frac{1}{|X_j|}$  则定义了  $R$  划分下产生的知识模块大小引起的不确定性.

**定理 1** 给定一决策系统  $L = \{U, C \cup \{d\}\}$ , 集合  $X \subseteq U, P$  和  $Q$  是  $U$  上的等价关系, 若  $I(P) \subseteq I(Q)$ , 则有  $E_B(P) \leq E_B(Q)$ .

该定理表明边界粗糙熵满足熵的条件. 熵作为系统不确定性的一种度量手段, 随着等价类划分越细, 系统知识不确定性减少, 熵逐渐变小.

要证明定理 1, 首先需要证明引理 1.

**引理 1** 对决策系统  $L = \{U, C \cup \{d\}\}$ , 集合

$X \subseteq U, P$  和  $Q$  是  $U$  上的等价关系,  $U/\{d\} = \{Y_1, Y_2, \dots, Y_n\}$ . 若  $I(P) \subseteq I(Q)$ , 则有  $B_{nP}(Y_i) \subseteq B_{nQ}(Y_i)$ .

**证明** 先证明  $\underline{P}Y_i \supseteq \underline{Q}Y_i, \bar{P}Y_i \subseteq \bar{Q}Y_i$ .

对于  $\forall x \in \underline{Q}Y_i$ , 则  $[x]_Q \subseteq Y_i$ , 由于  $I(P) \subseteq I(Q)$ , 所以  $[x]_P \subseteq [x]_Q, [x]_P \subseteq Y_i$ , 有  $x \in \underline{P}Y_i$ , 故  $\underline{P}Y_i \supseteq \underline{Q}Y_i$ .

对于  $\forall x \in \bar{P}Y_i$ , 则  $[x]_P \cap Y_i \neq \emptyset$ , 由于  $[x]_P \subseteq [x]_Q$ , 所以  $[x]_Q \cap Y_i \neq \emptyset, x \in \bar{Q}Y_i$ , 故  $\bar{P}Y_i \subseteq \bar{Q}Y_i$ .

又  $B_{nP}(Y_i) = \bar{P}Y_i - \underline{P}Y_i$ , 所以  $B_{nP}(Y_i) \subseteq B_{nQ}(Y_i)$ . 证毕

下面证明定理 1.

**证明** 令  $E_B^1(R) = \sum_{i=1}^n r_i \left[ \frac{|B_{nR}(Y_i)|}{|U|} \lg |U| \right]$ ,

$$E_B^2(R) = - \sum_{j=1}^m \frac{|X_j|}{|U|} \lg \frac{1}{|X_j|}$$

先证明  $E_B^1(P) \leq E_B^1(Q), E_B^2(P) \leq E_B^2(Q)$ .

$E_B^1(P) - E_B^1(Q) = \sum_{i=1}^n r_i \frac{\lg |U|}{|U|} [ |B_{nP}(Y_i)| - |B_{nQ}(Y_i)| ]$ , 因为  $B_{nP}(Y_i) \subseteq B_{nQ}(Y_i)$ , 有  $|B_{nP}(Y_i)| \leq |B_{nQ}(Y_i)|$ , 故  $E_B^1(P) \leq E_B^1(Q)$ .

设  $X_l^P \subseteq U/I(P), X_k^Q \subseteq U/I(Q)$ .

对于  $\forall x \in X_l^P$ , 总存在  $X_k^Q$  使得  $x \in X_k^Q, |X_l^P| \leq |X_k^Q|$ .

因此,  $\sum |X_l^P| \lg |X_l^P| \leq \sum |X_k^Q| \lg |X_k^Q|$ .

故  $E_B^2(P) \leq E_B^2(Q)$ .

$E_B(R) = E_B^1(R) + E_B^2(R)$ , 故  $E_B(P) \leq E_B(Q)$ . 证毕

**性质 1** 在决策系统  $L = \{U, C \cup \{d\}\}$  中, 当  $U$  中的所有元素均可分辨, 即  $X_j = \{x_j\}, 1 \leq j \leq |U|$  时,  $E_B(R) = 0$ .

**性质 2** 在决策系统  $L = \{U, C \cup \{d\}\}$  中, 对于任意等价关系  $R$ , 粗糙边界熵满足  $0 \leq E_B(R) \leq \left( \sum_{i=1}^n \frac{|Y_i|^2}{|U|^2} + 1 \right) \lg |U|$ .

从 D-S 证据理论角度考虑, 在冲突信息融合阶段, 约简结果中的属性将作为证据进行融合, 而属性的决策重要度将作为反映该属性相对于结果重要程度的权重, 决定着冲突信息的分配. 所以要求属性的决策重要度包含两部分的内容, 分别反映该属性在约简结果的不可分辨关系中所起的作用, 以及该属性的分类效果与决策类的一致性.

而定义的边界粗糙熵则可以很好地体现了这两方面内容. 在最优约简结果中, 去掉某一属性将会导致边界粗糙熵的增加, 因此, 本文采用边界粗糙熵的增量定义该属性的决策重要度.

**定义 4** 给定一个决策系统  $L = \{U, C \cup \{d\}\}$ ,  $R \in C$ , 对于  $\forall a \in R$ , 属性  $a$  的决策重要度为

$$s(a) = E_B(R \setminus \{a\}) - E_B(R) \quad (3)$$

由上面定义的决策重要度经过归一化处理后就可以得到属性的相对权重:

$$w(a) = s(a) / \sum_{x \in R} s(x) \quad (4)$$

### 2.2 基于决策重要度的冲突证据合成方法

本文针对如何在航空电子装备的故障诊断应用中有效地分配冲突信息这一问题, 结合诊断信息处理过程所得到的决策重要度提出了一种基于决策重要度的冲突证据合成方法 (Conflicting Evidence Combination method based on Significance of decision, CECS). 其核心思想是: 冲突信息量的部分是可以利用的, 在合成过程中需要把证据冲突信息量按照证据对各个命题的支持程度加权进行平均, 而冲突信息分配取决于所获得证据的证据群体可信度和单个证据可信度这两方面的因素.

对于诊断数据进行粗糙集约简, 所得到的往往不止一个约简结果. 每一个约简结果  $R_i$  都可以作为一个独立的证据群, 证据的群体可信度反映出每个证据群  $R_i$  对决策结果的支持能力. 为了定义证据群体可信度, 首先定义证据群  $R_i$  的群体支持度:

**定义 5** 在决策系统  $L = \{U, C \cup \{d\}\}$  中,  $R_i \in C$  是决策表属性约简结果,  $R_0$  是  $U$  上的等价关系, 且满足  $U/R_0 = U$ , 则  $R_i$  作为证据群的群体支持度为

$$s_p(R_i) = E_B(R_0) - E_B(R_i) \quad (5)$$

显然, 当  $R_i$  的分类能力越好, 对决策结果的支持度越高, 粗糙边界熵  $E_B(R_i)$  会越小, 则  $R_i$  的群体支持度越高.

由以上定义, 将群体支持度进行归一化处理得到单个证据群  $R_i$  的群体可信度:

$$e(R_i) = \frac{s_p(R_i)}{\sum s_p(R_i)} \quad (6)$$

证据合成过程中的单个证据对应着约简结果  $R$  中的属性. 单个证据可信度反映了该证据在证据群中的重要程度, 可信度值越大, 则重要程度越高, 在冲突部分分配的过程中应获得越高的比例. 综上, 单个证据可信度由之前定义的相对权重  $w(a)$  给出.

基于冲突信息部分可以利用的思想, 冲突信息分配方案由 CECS 方法给出, 即证据组合规则如下:

**定义 6** 设  $m_1, m_2, \dots, m_n$  是同一辨识框架  $\Theta$  上证据  $E_1, E_2, \dots, E_n$  的  $n$  个基本信任分配函数, 所对应的焦元分别为  $G_i, i = 1, 2, \dots, n$ , 则合成公式为

$$\begin{aligned} m(\emptyset) &= 0 \\ m(G) &= \sum_{\cap A_i = A} \prod_{1 \leq i \leq n} m_i(G_i) + K \cdot e(R) \cdot \sum_{i=1}^n (m_i(G) \cdot w(a_i)) \quad G \neq \emptyset, \Theta \\ m(\Theta) &= \sum_{\cap A_i = \Theta} \prod_{1 \leq i \leq n} m_i(G_i) + K \cdot e(R) \cdot \sum_{i=1}^n (m_i(\Theta) \cdot w(a_i)) + K \cdot (1 - e(R)) \end{aligned}$$

式中:  $K = \sum_{\cap A_i = \emptyset} \prod_{1 \leq i \leq n} m_i(G_i)$ , 它反映了各个证据之间总的冲突程度.

### 2.3 基于 CECS 的航空电子装备故障诊断方法

在对航空电子装备进行测试的过程中, 各种传感器模块实时的采集现场数据. 产生的各种测试信息, 从不同的侧面反映出装备的当前状态. 需要对众多的诊断数据, 获得诊断规则, 给出测试结果. 具体方法如下:

- 1) 采集装备测试数据, 选取故障信息, 构建粗糙集诊断决策表  $L = \{U, C \cup \{d\}, V, f\}$ .
- 2) 对决策表中数据进行预处理. 根据测试指标正常区间, 按照“高于”、“低于”、“属于”该区间将测试数据进行离散化.
- 3) 按照粗糙集方法对决策表进行约简, 获得约简结果  $R$ .
- 4) 若得到约简结果不唯一, 通过式(2)计算边界粗糙熵  $E_B(R_i)$ , 选取值最小的约简结果  $R_i$ .
- 5) 根据式(1)计算  $R_i$  中的每条属性的值频率  $p[f_a(x_i)]$ , 构建证据的基本概率赋值函数  $m_i$ .
- 6) 根据式(6)计算  $e(R_i)$  得到证据源  $R_i$  的群体可信度, 并通过式(4)计算  $w(a_i)$  得到每条证据的权重.
- 7) 通过 CECS 算法对证据进行组合.

CECS 算法针对选定的证据群, 为了界定可分配部分比例, 引入群体可信度  $e(R_i)$ , 通过步骤 6) 计算每个证据的支持度  $w(a_i)$  来决定冲突信息在证据间的分配比例, 在证据合成的过程中将其作为权重因子进行证据修改. 通过以上步骤, CECS 算法使得故障诊断结果具有更大的可信性和合理性.

### 3 故障诊断实例

#### 3.1 高度表故障分析

某型机载高度表收发机属于机载导航设备的重要组成部分.其中收发机为系统核心部件,担负发射脉冲、接收回拨、解算距离等任务.高度表收发机作为独立的外场可更换单元(LRU),包括5个内场可更换单元(SRU):电源板、发射机、接收机、总线板和距离计算器.

在实际的使用过程中,“无高度数据显示”故障较为常见.本文选取该故障为例,其涉及到的故障征兆:启动时间、电源功耗、工作电流、基准电压、总线电压、总线电流、自检电压,共7个方面;其涉及到的故障模块包括3个SRU:电源板、总线版和距离计算器.

#### 3.2 高度表故障诊断

对故障征兆所对应的5个测试项目进行数据采样,选取3种故障模式对应的典型样本,以此方式构建诊断决策表(见表1).表中典型故障对象列举18条,条件属性为5个,决策属性为3个.首先运用粗糙集方法对数据进行离散化并对决策表进行属性约简.

表1 离散化故障诊断决策表

Table 1 Discrete fault diagnostic decision table

编号	C					D
	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	
1	1	0	1	0	0	d <sub>1</sub>
2	1	0	0	0	0	d <sub>1</sub>
3	2	0	2	2	0	d <sub>1</sub>
4	0	0	1	1	1	d <sub>1</sub>
5	1	2	0	0	2	d <sub>1</sub>
6	0	0	1	2	0	d <sub>1</sub>
7	2	1	0	0	2	d <sub>2</sub>
8	1	1	1	0	0	d <sub>2</sub>
9	1	2	2	1	0	d <sub>2</sub>
10	1	0	0	0	1	d <sub>2</sub>
11	2	2	1	0	0	d <sub>2</sub>
12	1	0	2	1	0	d <sub>2</sub>
13	2	0	2	0	1	d <sub>3</sub>
14	0	0	2	0	0	d <sub>3</sub>
15	2	1	2	1	0	d <sub>3</sub>
16	0	2	2	1	1	d <sub>3</sub>
17	2	2	2	0	0	d <sub>3</sub>
18	1	2	0	0	0	d <sub>3</sub>

在对决策表进行约简之后,得到两组约简结果: $R_1 = \{C_1, C_2, C_3, C_4\}$ ,  $R_2 = \{C_1, C_2, C_3, C_5\}$ .  $C_1 \sim C_5$ 分别代表着:启动时间、电源功耗、基准电压、总线电压、自检电压;而  $d_1 \sim d_3$  代表3种模块:电源板、总线版、距离计算器.为了选定一组最

优约简作为融合的证据群,分别计算粗糙熵得: $E_B(R_1) = 1.1827$ ,  $E_B(R_2) = 1.7532$ .因此选定约简结果  $R_1$  作为证据群,通过式(6)计算其作为证据群的群体可信度为: $e(R_1) = 0.5349$ .

将  $R_1$  中的每一个属性  $C_i$  作为一条证据,而将不同的故障模块  $d_i$  作为辨识框架.在不同的属性值下会得到证据不同的基本信任分配函数,这些基本概率分配函数如表2所示.

表2 基本概率分配函数表

Table 2 Table of basic probability assignment function

属性	$C_i$	$m_i(d_1)$	$m_i(d_2)$	$m_i(d_3)$
$i=1$	0	0.5	0	0.5
	1	0.375	0.5	0.125
	2	0.167	0.333	0.5
$i=2$	0	0.556	0.222	0.222
	1	0	0.667	0.333
	2	0.167	0.333	0.5
$i=3$	0	0.286	0.428	0.286
	1	0.6	0.2	0.2
	2	0.167	0.333	0.5
$i=4$	0	0.272	0.364	0.364
	1	0.2	0.4	0.4
	2	1	0	0

由此可得到每条证据在不同的情况下基本概率分配函数,再根据计算每个证据作为条件属性时在约简结果中决策重要度的公式,算出  $s(a)$  值,进而得到单个证据可信度  $w(a)$  作为权重,得到的结果如表3所示.

表3 单个证据决策重要度与权重

Table 3 Decision significance and weight of evidence

特征	$m_1$	$m_2$	$m_3$	$m_4$
$s(a_i)$	0.608	0.830	0.719	0.763
$w(a_i)$	0.208	0.284	0.246	0.262

选取实测的4组故障数据作为验证样本.对故障数据进行处理,将其离散化后,按照对应约简  $R_1 = \{C_1, C_2, C_3, C_4\}$  取得4个验证样本: $T_1 = \{2, 0, 2, 2\}$ ,  $T_2 = \{0, 1, 1, 2\}$ ,  $T_3 = \{2, 1, 1, 0\}$ ,  $T_4 = \{2, 2, 2, 2\}$ .通过表2可获得验证样本中每条证据的基本信任分配函数.

文中分别采用经典 D-S 方法、Yager<sup>[25]</sup>方法、文献[26]方法、文献[27]方法、基于证据距离的冲突合成方法和本文方法进行比较.其中,基于证据距离的冲突合成方法定义为用基于文献[28]所提出的证据间距离得出每条证据的重要度,替代本文 CECS 算法中的单个证据可信度作为证据的权重,其余参数设置与本文 CECS 算法相同.利用以上6种方法对上述4个验证样本中证据进行,得到的结果如表4所示.

表 4 不同故障样本下融合结果对比

Table 4 Fusion results comparison of different diagnostic samples

测试样本	合成方法	D-S 方法	Yager <sup>[25]</sup> 方法	文献[26]方法	文献[27]方法	基于证据距离方法 <sup>[28]</sup>	本文方法
$T_1 = \{2,0,2,2\}$	$m(d_1)$	1	0.016	0.481	0.249	0.234	0.282
	$m(d_2)$	0	0	0.218	0.109	0.130	0.115
	$m(d_3)$	0	0	0.301	0.150	0.179	0.156
	$m(\Theta)$	0	0.984	0	0.492	0.457	0.457
$T_2 = \{0,1,1,2\}$	$m(d_1)$		0	0.525	0.263	0.299	0.275
	$m(d_2)$		0	0.217	0.109	0.095	0.128
	$m(d_3)$		0	0.258	0.130	0.141	0.132
	$m(\Theta)$		1	0	0.498	0.465	0.465
$T_3 = \{2,1,1,0\}$	$m(d_1)$	0	0	0.253	0.126	0.131	0.132
	$m(d_2)$	0.572	0.016	0.396	0.206	0.216	0.226
	$m(d_3)$	0.428	0.012	0.351	0.132	0.198	0.190
	$m(\Theta)$	0	0.972	0	0.486	0.452	0.452
$T_4 = \{2,2,2,2\}$	$m(d_1)$	1	0.005	0.378	0.164	0.141	0.210
	$m(d_2)$	0	0	0.249	0.106	0.158	0.131
	$m(d_3)$	0	0	0.373	0.159	0.278	0.196
	$m(\Theta)$	0	0.995	0	0.571	0.463	0.463

从融合结果中可以看出,经典 D-S 合成方法在对样本  $T_1$ 、 $T_3$ 、 $T_4$  进行融合的过程中可以得出结果.但是样本  $T_2$  其中有两条证据完全冲突,证据间的总冲突  $K=1$ ,此时,D-S 合成方法失效,不能有效识别故障,暴露出 D-S 合成规则存在的不足.

Yager 合成方法<sup>[25]</sup>将冲突的部分全部分配在辨识框架上,由融合结果可以看出,证据源为 4 个时,效果并不理想,且该方法在处理低冲突证据时的效果好于处理高冲突证据的效果,当处理样本  $T_2$  中的高冲突证据时,Yager 合成方法将全部信任都赋予辨识框架上,导致方法失效.

文献[26]的方法在样本  $T_1$ 、 $T_2$ 、 $T_3$  的融合过程中得到了正确结果,但是在样本  $T_4$  上基本失效,无法做出决策,该方法认为冲突信息可以全部利用,并将证据总冲突的部分对单个证据源进行平均分配,这是一种不考虑外在因素冒险决策,同时,认为单个证据源拥有同样的可信度不符合实际情况,会造成误差.

文献[27]的方法针对以上问题,对合成规则进行修改,认为冲突信息的一部分可以利用,其比例取决于可信度函数  $\varepsilon$ ,该方法在融合过程中取得了较好的效果,但其同样认为所有单个证据源具有相同的可信度,将冲突部分信息在证据源间进行平均分配,所以在样本上的融合结果也不理想.

基于证据距离的方法与本文提出的方法都针对冲突部分信息在证据源间的分配方面进行了改进,区别在于基于证据距离的方法基于证据间距离定义单个证据源的重要度,将其作为融合时的证据权重;而本文方法应用可以从决策表中获知

的决策重要度得出单个证据源的可信度,将其作为权重.从融合结果上看,两种方法在样本  $T_1$  上得到了一致的结果,本文方法要优于基于证据距离的方法;在样本  $T_4$  上,基于证据距离的方法得到的融合结果错误.其根本原因在于证据间距离的计算只考虑了证据间的相似性,而没有像将之前蕴含在数据集中的知识体现在融合过程中,造成融合失效.为了说明这一点,表 5 给出了基于证据距离方法中与本文方法中的证据权重对比.

表 5 证据权重对比

Table 5 Comparison of weight of evidence

方法	样本	$w(a_1)$	$w(a_2)$	$w(a_3)$	$w(a_4)$	
本文方法	$T_1 \sim T_4$	0.208	0.284	0.246	0.262	
	$T_1$	0.278	0.276	0.278	0.168	
	基于证据 距离方法	$T_2$	0.286	0.169	0.322	0.223
		$T_3$	0.270	0.230	0.218	0.282
	$T_4$	0.298	0.298	0.298	0.106	

从表 5 中可以看出,基于证据距离的方法在样本  $T_2$  上之所以取得效果较本文方法好,主要由于第 1 条和第 3 条基本概率分配相似而得到较多权重,而这两条证据明显支持决策.基于证据距离的方法在样本  $T_4$  融合过程中没有考虑到样本中第 4 条证据拥有较高的决策权而仅因其与其他 3 条证据有差异而赋予极低的权重导致融合结果错误.

通过样本验证,可以看出本文所提出的融合粗糙集与改进证据理论的故障诊断方法可以充分挖掘出隐藏在数据集中的知识,衡量出每条证据对决策的重要程度,对证据理论合成公式进行改进.在冲突证据的融合效果、适用性等方面存在优势.

## 4 结论

针对本文在对多源信息融合技术进行分析后,提出了一种基于粗糙集与证据理论的航空装备故障诊断方法.经实验验证表明:

1) 在粗糙集框架内实现故障诊断决策模型构建.

2) 对冗余故障信息进行约简,所提出的边界粗糙熵可以有效衡量每条测试项目对于决策的重要程度,提升了后期融合精度.

3) 在冲突证据的合成应用中,有效集成了证据的先验知识,得到诊断结果的准确度要好于目前存在的证据合成方法.

为了能对多故障类型、多测试属性故障,仍需对方法的计算量进行优化.

### 参考文献 (References)

- [1] 潘泉,王增福,梁彦,等.信息融合理论的基本方法与进展(II)[J].控制理论与应用,2012,29(10):1233-1244.  
Pan Q, Wang Z F, Liang Y, et al. Basic methods and progress of information fusion (II) [J]. Control Theory & Applications, 2012, 29(10): 1233-1244 (in Chinese).
- [2] Khaleghi B, Khamis A, Karray F O, et al. Multisensor data fusion: A review of the state-of-the-art [J]. Information Fusion, 2013, 14(1): 28-44.
- [3] Pravia M A, Babko-Malaya O, Schneider M K, et al. Lessons learned in the creation of a data set for hard/soft information fusion [C] // 12th International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2009: 2114-2121.
- [4] Gross G A, Nagi R, Sambhoos K, et al. Towards hard + soft data fusion: Processing architecture and implementation for the joint fusion and analysis of hard and soft intelligence data [C] // 15th International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2012: 955-962.
- [5] Hossain M A, Atray P K, El Saddik A. Learning multisensor confidence using a reward-and-punishment mechanism [J]. IEEE Transactions on Instrumentation and Measurement, 2009, 58(5): 1525-1534.
- [6] David R P, Sampaio-Neto R, Medina C A. A linear adaptive algorithm for data fusion in distributed detection systems [C] // 11th International Symposium on Wireless Communications Systems. Piscataway, NJ: IEEE Press, 2014: 370-374.
- [7] Meng Z, Shi Z. Extended rough set-based attribute reduction in inconsistent incomplete decision systems [J]. Information Sciences, 2012, 204(10): 44-69.
- [8] Lu Z, Qin Z, Zhang Y, et al. A fast feature selection approach based on rough set boundary regions [J]. Pattern Recognition Letters, 2014, 36(1): 81-88.
- [9] Zheng K, Hu J, Zhan Z, et al. An enhancement for heuristic attribute reduction algorithm in rough set [J]. Expert Systems with Applications, 2014, 41(15): 6748-6754.
- [10] Shu W, Shen H. Incremental feature selection based on rough set in dynamic incomplete data [J]. Pattern Recognition, 2014, 47(12): 3890-3906.
- [11] Wang J S, Song J D, Gao J. Rough set-probabilistic neural networks fault diagnosis method of polymerization kettle equipment based on shuffled frog leaping algorithm [J]. Information, 2015, 6(1): 49-68.
- [12] Gao S, Wang J, Zhao N. Fault diagnosis method of polymerization kettle equipment based on rough sets and BP neural network [J]. Mathematical Problems in Engineering, 2013, 7(2): 91-103.
- [13] Liu C, Wu X, Wu N, et al. Structural damage identification based on rough sets and artificial neural network [J]. The Scientific World Journal, 2014, 7(11): 102-111.
- [14] Chen R C, Cheng K F, Chen Y H, et al. Using rough set and support vector machine for network intrusion detection system [C] // 2009 1st Asian Conference on Intelligent Information and Database Systems. Piscataway, NJ: IEEE Press, 2009: 465-470.
- [15] Niu W, Cheng J, Wang G, et al. Fast fault prediction model based on rough sets and grey model [J]. Journal of Computational and Theoretical Nanoscience, 2013, 10(6): 1460-1464.
- [16] Han D Q, Dezert J, Tacnet J M, et al. A fuzzy-cautious OWA approach with evidential reasoning [C] // 15th International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2012: 278-285.
- [17] Yang J, Huang H Z, Miao Q, et al. A novel information fusion method based on Dempster-Shafer evidence theory for conflict resolution [J]. Intelligent Data Analysis, 2011, 15(3): 399-411.
- [18] Li Y B, Kang J, Xie H. The algorithm aiming at conflict to improve DS evidence theory [J]. Information Technology Journal, 2011, 10(9): 1779-1783.
- [19] Hu B, Shen B, Liu Q. An experience-feedback algorithm of DS evidence theory [C] // 2013 International Conference on Information Science and Computer Applications. Paris: Atlantis Press, 2013.
- [20] Zhu H, Ma Z, Sun H, et al. Information correlation entropy based DS evidence theory used in fault diagnosis [C] // 2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE). Piscataway, NJ: IEEE Press, 2012: 336-338.
- [21] Wang J F, Zhang Q L, Zhi H L. Fault diagnosis and optimization for agent based on the ds evidence theory [M] // Advances in Swarm Intelligence. Berlin: Springer, 2012: 535-542.
- [22] Xu C, Zhang H, Peng D, et al. Study of fault diagnosis of integrate of DS evidence theory based on neural network for turbine [J]. Energy Procedia, 2012, 16(10): 2027-2032.
- [23] Zhang L, Dong Y. Research on diagnosis of ac engine wear fault based on support vectormachine and information fusion [J]. Journal of Computers, 2012, 7(9): 2292-2297.
- [24] Zhang Q H, Hu Q, Sun G, et al. Concurrent fault diagnosis for rotating machinery based on vibration sensors [J]. International Journal of Distributed Sensor Networks, 2013, 4(5): 37-47.

- [25] Yager R R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking [J]. IEEE Transactions on Systems, Man and Cybernetics, 1988, 18(1):183-190.
- [26] 李弼程, 钱曾波. 一种有效的证据理论合成公式 [J]. 数据采集与处理, 2002, 17(1):33-36.
- Li B C, Qiang Z B. An efficient combination rule of evidence theory [J]. Journal of Data Acquisition and Processing, 2002, 17(1):33-36 (in Chinese).
- [27] 邓勇, 施文康. 一种改进的证据推理组合规则 [J]. 上海交通大学学报, 2003, 37(8):1275-1278.
- Deng Y, Shi W K. An advanced combination rules of evidence [J]. Journal of Shanghai Jiao Tong University, 2003, 37(8):1275-1278 (in Chinese).
- [28] Jousselme A L, Grenier D, Bossé É. A new distance between two bodies of evidence [J]. Information Fusion, 2001, 2(2):91-101.

## Avionic devices fault diagnosis based on fusion method of rough set and D-S theory

SUN Weichao<sup>1</sup>, LI Wenhai<sup>\*2</sup>, LI Wenfeng<sup>1</sup>

(1. Graduate Student's Brigade, Naval Aeronautical Engineering University, Yantai 264001, China;

2. Department of Scientific Research, Naval Aeronautical Engineering University, Yantai 264001, China)

**Abstract:** In order to solve the conflict of multi-sources information in the fault diagnosis process of avionics electric equipment, a method based on rough set theory and evidence theory for fault diagnosis was proposed. Because both rough set theory and evidence theory had advantages in dealing with uncertainty problems. The method proposed converted diagnostic data to mass function which was needed in evidence theory in order to fuse results with rough set theory. Meanwhile, the method defined boundary rough entropy, got dynamic weight parameters which reflected the significance of every information source used in fusion process with the entropy and improve the rule for conflicting evidence combination. The experiment shows that the method improves the fusion results' accuracy of diagnostic information effectively and has a good practical value in process of avionics electric fault diagnosis.

**Key words:** boundary rough set entropy; rough set; D-S evidence theory; conflicting evidence fusion; fault diagnosis

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0706

# 基于子目标进化的高维多目标优化算法

雷宇曜, 姜文志\*, 刘立佳, 马向玲

(海军航空工程学院 兵器科学与技术系, 烟台 264001)



**摘 要:** 多目标优化问题是工程应用中的常见问题, 已有的方法在解决 3 个目标以上的高维优化问题时效果欠佳. 如何进行有效的个体选择是求解高维多目标优化问题的关键. 针对该问题, 提出了求解高维多目标优化问题的子目标进化算法. 从理论上证明了多目标优化问题 Pareto 非支配解的求取, 可通过子目标函数值排序, 先行选择进化种群中部分非支配解; 然后, 根据排序信息有选择性地比较进化种群中的元素, 减少了比较次数, 从而快速获得非支配解集. 同时, 提出归一化函数差值的 Minkowski 距离“ $k$  近邻”距离计算方法, 在进化过程中应用到密度函数中, 加速了收敛速度. 同当前求解高维多目标优化的算法, 在对标准测试函数的计算性能上进行比较, 统计结果显示了所提算法在性能上的优势.

**关键词:** 高维多目标优化; 子目标进化算法; Pareto 非支配解集; Minkowski 距离; 遗传算法

中图分类号: TP181

文献标识码: A 文章编号: 1001-5965(2015)10-1910-08

当前在解决目标数超过 3 个的高维多目标优化问题 (Many-objective Optimization Problems, MOPs) 中, 由于目标函数数量的增加, 传统的基于 Pareto 占优关系生成 Pareto 最优解时, 互不占优解的数目急剧增多, 使得 Pareto 最优解的选择压力降低, 难以有效地逼近 Pareto 前沿<sup>[1-2]</sup>. 一些经典算法: NSGA-II<sup>[3]</sup>、SPEA2<sup>[4]</sup> 和 PESA-II<sup>[5]</sup> 对于高维多目标优化问题都不能很好地适应<sup>[6-7]</sup>. Saxena 和 Deb<sup>[8]</sup> 分别运用 PCA 和最大变化伸展 PCA 以及相关熵 PCA 的方法来减少冗余目标, 但是多个目标函数不存在冗余时是无效的. 文献 [9] 提出了 Pareto 自适应  $\varepsilon$  占优机制, 结合 Pareto 前沿面的几何形状和前沿面的不同部分的占优强度信息, 自适应地调节  $\varepsilon$  矢量中元素的值. 实现了高维多目标的降维, 但是这种方法会改变部分解

的支配关系, 求解有些高维目标优化问题时解的质量较差. 文献 [10] 基于 GDDR (Group Divided Dimensional Reduction) 算法实现了多目标降维处理, 虽然性能上有一定的优势, 但是该方法在目标函数进行分组时具有较大的随意性, 不同的分组方法将导致不同的优化结果. 文献 [11] 应用文献 [12] 提出的基于聚集函数的方法, 使用加权的最小最大法以及结合对偶优化的权角距离的度量方法, 能用于解决 MOPs 问题, 但是处理较低维数时效果欠佳. 文献 [13] 提出的 efficiency of order 方法能够将 Pareto 解集中的每个解的目标向量向  $K$  维空间进行投影. 该方法能够找到更逼近 Pareto 前沿且分布均匀的非支配解, 在解决低维的多目标优化问题时有一定的优势. 文献 [14] 利用超体积、分布度和延展度将高维多目标优化问题转化

收稿日期: 2014-11-17; 录用日期: 2015-02-13; 网络出版时间: 2015-03-23 15:18

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150323.1518.003.html

基金项目: 国防预研项目 (2014CX-C201-FW); 国家自然科学基金青年科学基金 (61002006)

作者简介: 雷宇曜 (1984—), 男, 陕西丹凤人, 博士研究生, 2669116897@qq.com

\* 通讯作者: 姜文志 (1964—), 男, 山东莱州人, 教授, ytjwz@sohu.com, 主要研究方向为作战指挥一体化技术.

**引用格式:** 雷宇曜, 姜文志, 刘立佳, 等. 基于子目标进化的高维多目标优化算法 [J]. 北京航空航天大学学报, 2015, 41 (10): 1910-1917. Lei Y Y, Jiang W Z, Liu L J, et al. Many-objective optimization based on sub-objective evolutionary algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41 (10): 1910-1917 (in Chinese).

为具有 3 个目标函数的多目标优化问题,虽然从理论上证明了可应用于求解高维多目标优化,但是设计高性能的集合进化策略往往是十分困难的,这给此种方法的应用造成了很大的限制。

本文从缩小搜索空间入手,在理论上证明了通过子目标函数值排序进行 Pareto 最优解求取的可行性,能够有效减少比较次数,加快求解速度。同时,基于 SPEA2 (Strength Pareto Evolutionary Algorithm2) 算法框架,对密度评估策略中的距离和密度计算公式进行重新定义。提出的 SOEA (Sub-Object Evolutionary Algorithm) 算法同文献 [15-17] 中的高维多目标优化算法,在测试函数 DTLZ<sup>[2,18]</sup> 上关于收敛性指标<sup>[17]</sup>、覆盖率指标<sup>[19]</sup> 进行比较,显示出一定的优势。

## 1 多目标优化相关表述

$n$  个决策变量和  $m$  个目标变量的多目标优化问题<sup>[2]</sup> 可以表述为

$$\begin{aligned} \min \quad & \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})) \\ \text{s. t.} \quad & g_i(\mathbf{x}) \leq 0 \quad i=1, 2, \dots, G \\ & h_j(\mathbf{x}) = 0 \quad j=1, 2, \dots, H \end{aligned}$$

式中:  $\mathbf{x}$  为  $n$  维决策矢量,  $\mathbf{x} \in \mathbf{R}^n$ ;  $\mathbf{f}(\mathbf{x})$  为  $m$  维目标矢量,  $\mathbf{f}(\mathbf{x}) \in \mathbf{R}^m$ ,  $\mathbf{f}(\mathbf{x})$  定义了  $m$  个由决策空间向目标空间映射的函数;  $g_i(\mathbf{x}) \leq 0 (i=1, 2, \dots, G)$  定义了  $G$  个不等式约束;  $h_j(\mathbf{x}) = 0 (j=1, 2, \dots, H)$  定义了  $H$  个等式约束。

**定义 1** Pareto 占优<sup>[3]</sup>。若  $x_a, x_b \in X_f$  是多目标优化问题的两个可行解,则称  $x_a$  是比  $x_b$  Pareto 占优的。

1) 若对于  $\forall i=1, 2, \dots, m; f_i(x_a) \leq f_i(x_b)$  且  $\exists j=1, 2, \dots, m; f_j(x_a) < f_j(x_b)$ , 则记作  $x_a >_s x_b$ , 也称为  $x_a$  强支配  $x_b$ 。

2) 若对于  $\forall i=1, 2, \dots, m; f_i(x_a) \leq f_i(x_b)$  记作  $x_a >_w x_b$ , 也称为  $x_a$  弱支配  $x_b$ 。文中的支配概念, 均是该定义下的弱支配。

## 2 高维优化 SOEA 算法

### 2.1 理论可行性证明

在证明相关命题之前,需要对  $\mathbf{R}^m$  空间 ( $m$  维欧氏空间) 上的多目标优化问题,进行如下的假设:

1) 多目标优化问题的各个子目标函数在解空间上是连续的。若存在可去间断点和阶跃间断点,则利用目标函数在解空间连续的方法进行讨论,而不影响问题求解结果。

2) 解空间是  $\mathbf{R}^n$  上的有界闭子空间。

**引理 1** 多目标优化问题的可行解集是半序集。

**证明** 设  $x_1, x_2 \in X_f$ , 对应的目标函数为  $\mathbf{F}_1 = (f_1(x_1), f_2(x_1), \dots, f_m(x_1))$ ,  $\mathbf{F}_2 = (f_1(x_2), f_2(x_2), \dots, f_m(x_2))$ ,  $\mathbf{F}_1, \mathbf{F}_2 \in \mathbf{R}^m$ , “ $\leq$ ” 表示实数域条件下的小于等于含义。①对于  $x_1$ , 有  $f_i(x_1) \leq f_i(x_1), i=1, 2, \dots, m$  成立, 由解支配的定义即得  $x_1 > x_1$ 。②对  $x_1, x_2$ , 若有  $f_i(x_1) \leq f_i(x_2), f_i(x_2) \leq f_i(x_1)$ , 即  $f_i(x_2) = f_i(x_1)$  成立。由①的结论可得  $x_1 > x_2 (x_2 > x_1)$ 。③对  $x_1, x_2, x_3 \in X_f$ , 若有  $f_i(x_1) \leq f_i(x_2), f_i(x_2) \leq f_i(x_3)$ , 即有  $f_i(x_1) \leq f_i(x_3) (i=1, 2, \dots, m)$  成立, 则  $x_1 > x_3$ 。所以, 在可行解集上定义的 Pareto 占优关系, 即 Pareto 支配“ $>$ ”, 是集合  $X_f$  上的半序。 **证毕**

**引理 2** 子目标的最优解集是全序集。

**证明** 设  $\{x_p\}$  是子目标  $f_i(x)$  可行解的集合, 则  $\{x_p\} \subset X_f$ 。由于  $X_f \neq \emptyset$ 。对  $\forall x_1, x_2 \in \{x_p\}, x_1 \neq x_2$ , 有  $f_i(x_1) \leq f_i(x_2)$  或  $f_i(x_2) \leq f_i(x_1)$ , 则  $\{x_p\}$  是全序集。由于  $f_i(\mathbf{x})$  是解空间上的连续函数, 解空间是  $\mathbf{R}^n$  上的有界闭子空间, 故  $\{x_p\}$  存在上界。由 Zorn 引理<sup>[20]</sup> 可知  $X_f$  至少存在一个极大元。 **证毕**

**引理 3** 极大元是非支配解。

**证明** 对于多目标优化问题的可行解集  $X_f$ ,  $X_f$  已被证明为半序集。若  $\mathbf{x}^*$  是  $X_f$  的极大元, 则由极大元的定义<sup>[20]</sup>, 对  $\forall \mathbf{x} \in X_f$ , 若  $\mathbf{x} > \mathbf{x}^*$ , 则  $\mathbf{x}^* = \mathbf{x}$ 。故对  $\forall \mathbf{x} \in X_f$ , 有  $\mathbf{x}^* > \mathbf{x}$  成立。也即:  $\neg \exists \mathbf{x} \in X_f$  使得  $\mathbf{x} > \mathbf{x}^*$  成立, 则  $\mathbf{x}^*$  是多目标优化问题的非支配解。 **证毕**

**定理 1** 设  $\{f_1, f_2, \dots, f_m\}$  是优化问题的目标函数集,  $X_f$  为其解空间,  $\mathbf{x}^{**}$  为其 Pareto 非支配集;  $x_1^*, x_2^*, \dots, x_m^*$  是使得各目标函数取得极值的解, 则集合  $Y = \{x_1^*, x_2^*, \dots, x_m^*\}$  是非支配解集, 且  $Y$  是  $\mathbf{x}^{**}$  的子集。

**证明** 对  $\forall x_i \in X_f$ , 取集合  $Y$  中的任一元素  $x_j^*$ , 由定义,  $x_j^*$  使得  $f_j(\cdot)$  取得极值, 则由支配的定义, 对  $\forall x_i \in X_f$ , 使得  $f_j(x_i) \leq f_j(x_j^*)$  不成立, 故  $x_i > x_j^*$  不成立, 即  $x_j^*$  是极大元。由引理 3 知  $x_j^*$  是非支配解, 故集合  $Y = \{x_1^*, x_2^*, \dots, x_m^*\}$  是非支配解集。

往证,  $Y$  是  $\mathbf{x}^{**}$  的子集。对  $\forall x_j^* \in Y$ , 由于  $Y$  是非支配解集, 故  $x_j^* \in \mathbf{x}^{**}$ 。若有  $x_j^{*'} \in \mathbf{x}^{**}$ , 则对  $\forall x_i \in X_f, \exists l, l \in \{1, 2, \dots, m\}$ , 使得  $f_l(x_i) \geq f_l(x_j^{*'}) \geq f_l(x_j^*)$  成立, 则  $f_l(x_j^{*'})$  不是  $f_l(\cdot)$  上的极值, 故  $x_j^{*'} \notin Y$ , 则  $Y \subset \mathbf{x}^{**}$ 。 **证毕**

**定理 2** 多目标优化 Pareto 非支配解  $\mathbf{x}^{**}$  的求取,可分解为  $f_1, f_2, \dots, f_m$  个子目标进行,且此种求取非支配解的方法是完备的.

**证明** 对  $\forall x_i, x_j \in X_j$ ,若对  $\forall l = 1, 2, \dots, m$ ,有  $f_l^i \leq f_l^j$  成立,则  $x_i > x_j$ . 若  $f_l^i \leq f_l^j$  成立,则记  $R(f_l^i, f_l^j) = 1$ , 否则,记  $R(f_l^i, f_l^j) = 0$ .  $\bigwedge_{l=1}^m R(f_l^i, f_l^j) = 1 \Leftrightarrow x_i > x_j$ , 可见  $>$  关系可分解为  $m$  个子目标函数值比较结果的逻辑与关系. 若  $\bigwedge_{l=1}^m R(f_l^i, f_l^j) = 0$ , 则  $x_i > x_j$  不成立.

对于  $x_i^*$  ( $i = 1, 2, \dots, m$ ) 是使得各目标函数取得极值的解,则集合  $Y = \{x_1^*, x_2^*, \dots, x_m^*\}$  是非支配解集,且  $Y$  是  $\mathbf{x}^{**}$  的子集,已由定理 1 得证. 若  $x_i^{*'}$  是在目标函数  $f_i(\cdot)$  上取值仅次于的极值点  $x_i^*$  的解. 由于  $x_i^*$  是 Pareto 非支配解,  $x_i^{*'}$  在目标函数  $f_i(\cdot)$  上的取值仅次于  $x_i^*$ , 故只需比较  $j \neq i$  上的  $f_j(x_i^*)$  与  $f_j(x_i^{*'})$  的取值大小,若  $f_j(x_i^{*'}) \leq f_j(x_i^*)$  成立,则  $x_i^* > x_i^{*'}$  不成立,由于  $x_i^*$  是非支配解,故  $x_i^{*'}$  也是非支配解. 同理,若  $x_i^{*''}$  是在目标函数  $f_i(\cdot)$  上的取值次于  $x_i^{*'}$  的解. 若有  $f_k(x_i^{*''}) \leq f_k(x_i^{*'})$ 、 $f_j(x_i^{*''}) \leq f_j(x_i^{*'})$ ,  $j \neq i, k \neq j$  成立,则  $x_i^{*' > x_i^{*''}$ ,  $x_i^* > x_i^{*''}$  不成立,则  $x_i^{*''}$  也是非支配解. 继续此过程求取更多 Pareto 非支配解  $x_i^{*'''} \dots, x_i^{*(n')}$ , 记这些非支配解构成的集合为  $\{x_i^{*(n')}\}$ .

往证,完备性. 对  $\forall \mathbf{x} \in \mathbf{x}^{**}$ , 若  $\mathbf{x} \in Y$ , 则  $\exists i \in \{1, 2, \dots, m\}$  使得  $f_i(\mathbf{x}) = \min f_i(\cdot)$ , 则通过子目标排序的方法一定能找到  $\min f_i(\cdot)$ , 即能找到非支配解  $\mathbf{x}$ . 若  $\mathbf{x} \in \mathbf{x}^{**}$ ,  $\mathbf{x} \notin Y$ , 不妨设在某子目标  $i \in \{1, 2, \dots, m\}$  上有  $f_i(\mathbf{x}) \geq f_i(\mathbf{x}^{*''}) \geq f_i(\mathbf{x}^{*'})$ , 由于  $f_i(\mathbf{x}^{*'}) = \min f_i(\cdot)$ , 故  $\mathbf{x}^{*'}$  必定被找到, 且  $\mathbf{x}^{*' > \mathbf{x}^{*''}$  不成立, 由于  $\mathbf{x}^{*''}$  在  $f_i(\cdot)$  上的取值仅次于  $\mathbf{x}^{*'}$ , 故  $\mathbf{x}^{*''}$  必定被找到, 依此法  $\mathbf{x}$  必定被找到, 则对  $\forall \mathbf{x} \in \mathbf{x}^{**}$  都可被找到. **证毕**

## 2.2 SOEA 算法

### 2.2.1 进化过程

SOEA 算法是以进化种群中子目标函数值排序为基础. 根据求取的是目标函数的最大(最小)值,对子目标函数值进行排序. 通过子目标值的排序,快速找出非支配解. 以子目标函数  $f_1$  为例,要求取目标函数的最小值. 进化过程如图 1 所示,通过排序,得到使  $f_1$  取最小值的解空间中的点  $A$ ,  $A$  为通过目标函数  $f_1$  获得的第 1 个非支配解. 继续该过程得到  $C, C_1$  在  $f_1$  上的函数值仅比  $A$  大,但

是  $C, C_1$  在子目标  $f_2$  上的函数值比  $A$  小,故  $C, C_1$  相对于  $A$  为非支配解. 但是  $C$  支配  $C_1$ , 故  $C$  为非支配解. 由此可见,子目标进化算法的方向性很强.

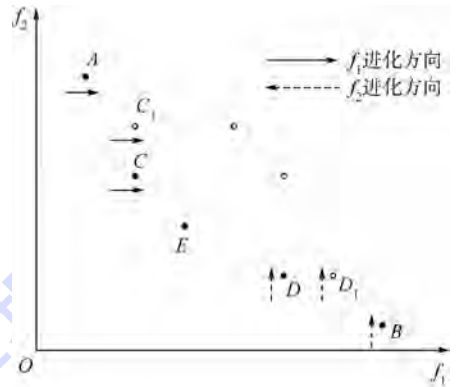


图 1 进化过程示意图

Fig.1 Schematic of evolution process

SOEA 算法利用了 SPEA2 算法的框架,与 SPEA2 算法不同之处有:利用子目标函数值排序,快速获得非支配解集,而不需要通过进化种群中的解之间两两比较来获得非支配解;利用归一化函数差值的 Minkowski 距离进行解之间的距离计算,通过“ $k$  近邻”距离来评判密度值的大小. 而不是通过计算进化种群中解之间的欧氏距离. 利用从外部种群中获取下一代进化的解,而不是在外部集合和进化种群中选取下一代的进化种群.

### 2.2.2 算法流程

输入值:进化种群的大小  $M$ , 外部集合的势  $\bar{M}$ , 最大进化代数  $P$ .

输出值:非支配解集  $U_{\text{non}}$ .

**第 1 步** 初始化,产生初始种群  $U_0$ , 并且产生一个空的外部集合  $\bar{U}_0 = \emptyset$ , 设置进化代数  $p = 0$ .

**第 2 步** 利用子目标优化方法求取进化种群  $U_p$  中的 Pareto 非支配解.

**第 3 步** 将产生的非支配解加入外部集合  $\bar{U}_p$ , 如果非支配解的个数大于外部集合的大小, 则利用截断方法进行操作; 如果产生的解的个数小于外部集合的大小, 则利用适应度的值, 选择满足条件的解进入外部集合.

**第 4 步** 如果达到最大进化代数  $P$ , 则停止计算. 将外部集合  $\bar{U}_p$  中的非支配解输出到  $U_{\text{non}}$  中作为最终的非支配解集.

**第 5 步** 计算外部集合  $\bar{U}_p$  中解的适应度, 对外部集合利用锦标赛的方法进行选择, 当被选

择的解的个数等于进化种群的大小  $M$  时停止选择.

**第 6 步** 利用交叉、变异操作对进化种群中的解进行操作,产生新一代种群  $U_{p+1}$ ,转入第 2 步.

### 2.2.3 非支配解获取

在求取外部集合的解时,设置一个临时集合用于存放由进化种群中选出的解,这些解在满足一定的条件时将被选入外部集合,临时集合作为求取过程中的辅助存储空间.根据外部集合  $\bar{U}_p$  的势  $\bar{M} = |\bar{U}_p|$ ,对于有  $m$  个子目标的多目标优化问题,首先由每个子目标对外部集合提供  $M_i = \bar{M}/m$  个元素.利用子目标快速选择方法生成非支配解,生成的非支配解的数目为  $|U_{pi}|$ ,当  $|U_{pi}| = \bar{M}/m$  时,暂时中止在此子目标上生成 Pareto 非支配解.将以上生成的非支配解加入临时集合,同时将这些非支配解从进化种群中删去.依据子目标  $f_1, f_2, \dots, f_m$  依次生成非支配解集  $U_{p1}, U_{p2}, \dots, U_{pm}$ .按最后一个子目标生成非支配解时,如果产生的非支配解的个数已经等于外部集合的大小.则依次选择进化种群中的解,按照该解中子目标的值最小(或者最大)的子目标,与临时集合中相应的非支配解进行比较,直至进化种群中所有的解比较完毕.此时,进化种群中的全部非支配解被选入临时集合.在选择非支配解进入临时集合的过程中,如果进化种群中的解选择完毕,也没有达到外部集合的个数,则利用适应度值选择支配解进入外部集合.

### 2.2.4 基于归一化 Minkowski 距离的“ $k$ 近邻”方法

解的选择是以适应度为准则进行评判的,解的适应度计算由解的强度和密度两部分组成.强度由解之间的支配关系确定;密度评估以解的距离计算为前提.密度值小,说明该解与空间中其他解的距离大,从而优先考虑被保留.关于距离的计算也有很多不同的计算方法,NSGA II (Non-dominated Sorting Genetic Algorithm II) 在子目标函数值排序的基础上,通过相邻的两个函数值的差值求取绝对值,归一化处理后,对每个子目标上的这些差值进行求和,从而得出个体的排挤距离. SPEA2 算法中,通过计算两个非支配解之间的欧氏距离,然后对与本非支配解相关的距离进行排序,利用“ $k$  近邻”的方法,求取在排序中的第  $k$  个距离值的大小,通过第  $k$  个距离值的大小计算密度值. Pareto 存档进化策略 (Pareto Archived Evolution Strategy, PAES) 算法利用空间超格机制来保

持种群的多样性.超格的数目需要预先设定,但是当目标的数目增多时,超格的数目难以确定,并没有很好的方法可循.以上的密度评估策略应用在高维多目标优化问题上时,获得的效果并不理想.

根据以上方法的不足,提出了更为广义的基于归一化 Minkowski 距离的“ $k$  近邻”方法.流程如下.

**第 1 步** 对于临时集合中的解,找出每个子目标上的最大值和最小值,计算两个解的同一子目标上函数值的差值,对这一差值进行归一化处理.

**第 2 步** 利用 Minkowski 距离计算公式,对两个解的子目标进行归一化处理后的差值进行距离计算.

**第 3 步** 对于临时集合中的每个解,对计算所得的与其他解之间的 Minkowski 距离按升序进行排列.

**第 4 步**  $k$  值取外部集合大小的方根,  $k = \sqrt{\bar{M}}$ . 记与解  $X_i$  距离为  $k$  的解的距离为  $\sigma_i^k$ . 密度评估按  $D(i) = \frac{1}{\sigma_i^k + 2}$  计算.

子目标函数差值的归一化是对每个子目标函数  $f_i$ , 在临时集合中分别求取最大值  $f_i^{\max}$  和最小值  $f_i^{\min}$ , 则两个解  $X_k, X_l$  在此子目标函数上的差值归一化后的形式为

$$\Delta_i(x_k, x_l) = \frac{f_i^k - f_i^l}{f_i^{\max} - f_i^{\min}} \quad (1)$$

Minkowski 距离计算公式为

$$d(x_k, x_l) = \sqrt[p]{\sum_{i=1}^m \left| \frac{f_i^k - f_i^l}{f_i^{\max} - f_i^{\min}} \right|^p} \quad (2)$$

当  $p$  值较小时,对差值的影响不明显,相当于均衡考虑各个目标;当  $p$  值较大时,差值大的目标对距离的影响明显.可以看出,NSGA II 和 SPEA2 中的距离计算公式是式(2)在  $p$  取值为 1 和 2 时的特例.在高维多目标优化时,非支配解的数目增长很快,均衡的进化方法势必导致收敛到 Pareto 前沿的速度过慢,因此适当地增大  $p$  的取值可以加速收敛速度.本文中采用随机的  $p$  值选取方法,在  $[1, H]$  的范围内随机选取一个  $p$  值,  $H$  为某正整数,作为每一代进化的 Minkowski 距离计算公式中的幂次,在同一代进化中  $p$  取相同的值.

### 2.2.5 外部集合解的选取

1) 非支配解个数小于外部集合时解的选取. 利用子目标快速选择方法中产生的非支配解

的数目不足  $\bar{M}$  时,通过计算进化种群中的支配解的适应度,选择适应度值小的支配解进入外部集合,直至满足外部集合的大小要求. 适应度计算公式为

$$F(i) = R(i) + D(i)$$

式中: 
$$R(i) = \sum_{\substack{j \in M, j \in \bar{M} \\ j > i}} S(j)$$

$$S(i) = |\{j | i \in \bar{M}, j \in M \wedge i > j\}|$$

$D(i)$  的计算按归一化 Minkowski 距离的“ $k$  近邻”方法进行.

2) 非支配解个数大于外部集合时解的选取.

由于此时临时集合中所有的解都是非支配解,适应度值计算公式  $F(i) = R(i) + D(i)$ ,退化为  $F(i) = D(i)$ . 将密度评估值进行升序排列,取密度值小的前  $\bar{M}$  个解进入外部集合. 非支配解  $x_{k1}, x_{k2}$  到  $x_i$  的距离相同时,即  $\sigma_i^{k1} = \sigma_i^{k2}$ . 统计非支配解  $x_{k1}, x_{k2}$  在各个子目标函数上的差值的正负. 若在  $f_j(x_{k1}) - f_j(x_{k2}) (j = 1, 2, \dots, m)$  中,取正值的子目标函数占多数,则当优化是求取最小化各目标时,选取  $x_{k2}$  进入外部集合;反之,选取  $x_{k1}$ .

### 2.2.6 时间复杂度分析

对于含有  $m$  个子目标,进化种群中有  $M$  个解,外部集合的势为  $\bar{M}$  的多目标进化问题. 初始化的时间复杂度是  $O(M)$ ,求取非支配解的时间复杂度是  $O(KM)$ ;求取外部集合的时间复杂度是  $O(L^2) + O(L \log L)$ ,其中,  $L$  是临时存储集合  $0 < L \leq M$ ;交叉、变异及种群更新时间复杂度是  $O(M)$ . 总的时间复杂度为:  $O(M) + O(KM) + O(L^2) + O(L \log L) + O(M)$ ,按照时间复杂度的运算规则,最差的时间复杂度应为:  $\max\{O(KM), O(L^2)\}$ .

对于 SPEA2 算法,初始化的时间复杂度是  $O(M)$ . 非支配解求取由适应度计算决定,适应度计算包含强度和距离计算. 计算强度  $R(i)$  的时间复杂度为  $O((M + \bar{M})^2)$ ;计算距离  $D(i)$  的时间复杂度为  $O((M + \bar{M})^2 \log(M + \bar{M}))$ . 求取外部集合的时间复杂度为  $O((M + \bar{M})^2 \log(M + \bar{M}))$ ;交叉、变异以及种群更新的时间复杂度为  $O(M)$ . 所以, SPEA2 算法的平均时间复杂度为:  $O((M + \bar{M})^2 \log(M + \bar{M}))$ . 由时间复杂度的分析可以看出,本文提出的算法在时间复杂度方面与 SPEA2 算法相比具有明显的优势.

解决高维多目标优化问题的基于子集覆盖实

现目标减少的近似  $k$ -MOSS 算法,时间复杂度为  $O(m^3 L^2)$ ,其中  $m$  为子目标的个数,  $L$  为临时集合的大小. 基于特征选择的目标减少算法,通过选择最拥挤的邻居群,保留该邻居群的中心,然后删除其余的目标,算法的时间复杂度为  $O(m^2 L)$ .

## 3 测试结果与分析

### 3.1 测试函数及评价指标

1) 测试函数.

测试函数使用高维多目标进化领域被广泛使用的 DTLZ1 ~ DTLZ7 函数,函数的具体形式可参见文献[2,18],在此不再赘述.

2) 收敛性指标<sup>[17]</sup>.

令  $G^* = \{g_1, g_2, \dots, g_{|G^*|}\}$  是理想的均匀分布的 Pareto 最优解集,  $Q = \{a_1, a_2, \dots, a_{|Q|}\}$  是通过进化算法得到的 Pareto 最优解集. 则对于集合  $Q$  中的解  $a_i$ ,通过式(3)得到该解距离  $G^*$  的最小欧氏距离:

$$d_i = \min_{j=1}^{|G^*|} \sum_{m=1}^k (f_m(a_i) - f_m(g_j))^2 \quad (3)$$

IGD 评价方法用于衡量所求的 Pareto 前沿面到真实的 Pareto 前沿面的收敛程度,该指标的值越小,表明算法得到的解的收敛性越好,越接近理想 Pareto 前端,定义为

$$IGD = \sqrt{\frac{|Q|}{\sum_{i=1}^{|Q|} d_i^2}} \quad (4)$$

3) 覆盖率指标  $C$ <sup>[19]</sup>.

$E, F$  是优化问题的两个解集,覆盖率指标  $C$  定义为

$$C(E, F) = \frac{|\{f \in F; \exists e \in E: e > f\}|}{|F|} \quad (5)$$

式中:  $>$  表示 Pareto 支配(弱支配);  $C(E, F) = 1$ ,表示所有在  $F$  中的解均可由  $E$  中的解支配;  $C(E, F) \in [0, 1]$ ,  $C(E, F)$  取值越大,表示  $E$  集合中的解越优于  $F$ .

### 3.2 测试结果及分析

设置进化种群的规模为 600,外部种群的大小为 200,进化代数为 400 代,每个算法独立运行 20 次,评价指标为 20 次运行的平均值. 为了测试 SOEA 算法的性能,选择类似的解决高维多目标优化问题的算法进行比较,分别是由 Jaimes 和 Coello 提出的基于特征选择的算法(记为 JC),以及由 Brockhoff 和 Zitzler 提出的基于最小子集覆盖的目标减少算法  $k$ -MOSS(记为 BZ). 利用以上 3 种解决高维多目标优化问题的算法对 DTLZ1 ~ DTLZ7 测试函数进行计算,比对各算法对不同的

测试函数的计算性能. 由于以上 3 种算法对 7 个测试函数进行求解后, 求解结果表现出一定的相似性, 选取 SOEA 算法、BZ 算法、JC 算法对 DTLZ2、DTLZ5、DTLZ7 测试函数的测试结果在文中列出.

SOEA 算法在进化的过程中, 并没有减少目标函数的个数, 所以在进化过程中解之间的支配关系不会发生变化. 求得的非支配解必定是原来的多目标优化问题的 Pareto 最优解, 从表 1 中的数据也能说明这一点. 可以看出 SOEA 算法的 IGD 值比 BZ 算法和 JC 算法得到的值要小, 说明 SOEA 算法获得的 Pareto 前沿与最优 Pareto 前沿接近程度更好. JC 算法由于删除了过多的目标函数, 解的支配情况与 BZ 算法相比改变程度会大一些, 得到的 IGD 值也比 BZ 算法得到的要大.

由表 2 可见, 当目标数为 12 维时, 通过对获得的非支配解进行统计, 得出利用 SOEA 算法求得的非支配解, 可以支配 BZ 算法求得的非支配解, 占 BZ 算法求得解的 0.5443, 而 BZ 算法中可支配 SOEA 算法中的解, 只占到 SOEA 算法求得解的 0.3962. 可以得出利用 SOEA 算法所求得的非支配解集比 BZ 算法的质量要高, 与 JC 算法比较结果的统计数据也显示出这一结论.

表 1 对 DTLZ2 的收敛性指标比较

**Table 1 Convergence metric comparison on DTLZ2**

目标数量	BZ 算法	JC 算法	SOEA 算法
6	0.3946	0.4793	0.1843
12	0.5025	0.6252	0.3638
18	0.4237	0.3846	0.3741
24	0.3882	0.4027	0.2569
30	0.5849	0.6019	0.4362

表 2 对 DTLZ2 的覆盖率指标比较

**Table 2 Coverage metric comparison on DTLZ2**

目标数量	$C(\text{SOEA}, \text{BZ})$	$C(\text{BZ}, \text{SOEA})$	$C(\text{SOEA}, \text{JC})$	$C(\text{JC}, \text{SOEA})$
6	0.5384	0.3837	0.5479	0.3174
12	0.5443	0.3962	0.5604	0.3283
18	0.5602	0.3985	0.5693	0.3359
24	0.5728	0.4036	0.5685	0.3472
30	0.5105	0.3451	0.5334	0.3196

通过图 2 可见, JC 算法的运行时间比另两个算法占优, 这是因为该算法是直接目标函数上进行处理, 通过删除冗余的目标函数, 来达到降低维数的目的, 算法的时间复杂度相当于处理低维的目标优化问题. SOEA 算法和 BZ 算法均需要进行支配关系的比较, 时间复杂度与外部集合的

大小相关. 可以看出 SOEA 算法相比较于 BZ 算法, 基本上是正比例相关的. SOEA 算法在时间上要优于 BZ 算法.

通过对 3 个算法在测试函数 DTLZ5、DTLZ7 上相应值的统计结果, 显示了与在 DTLZ2 上类似的结论, 如表 3 ~ 表 6 和图 2 所示.

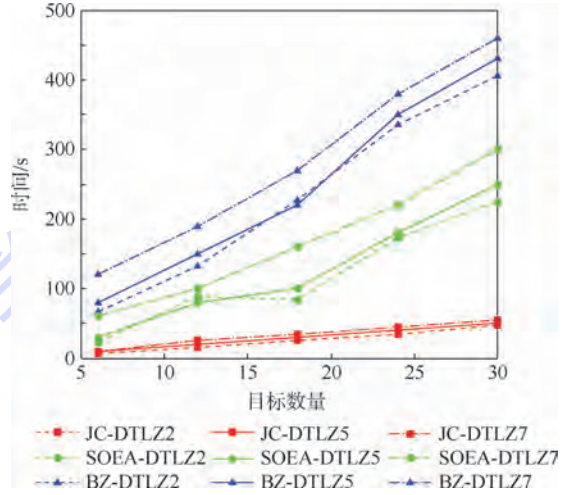


图 2 求解 DTLZ2 ~ DTLZ7 时各算法运行时间

Fig. 2 DTLZ2-DTLZ7 computation time of every algorithm

表 3 对 DTLZ5 的收敛性指标比较

**Table 3 Convergence metric comparison on DTLZ5**

目标数量	BZ 算法	JC 算法	SOEA 算法
6	0.1487	0.2942	0.1382
12	0.3642	0.5775	0.1549
18	0.4054	0.2912	0.2361
24	0.5068	0.1956	0.3622
30	0.3132	0.3803	0.2857

表 4 对 DTLZ5 的覆盖率指标比较

**Table 4 Coverage metric comparison on DTLZ5**

目标数量	$C(\text{SOEA}, \text{BZ})$	$C(\text{BZ}, \text{SOEA})$	$C(\text{SOEA}, \text{JC})$	$C(\text{JC}, \text{SOEA})$
6	0.4526	0.3572	0.4873	0.3249
12	0.4732	0.3655	0.5016	0.3364
18	0.5147	0.3824	0.5357	0.3381
24	0.5372	0.3897	0.5628	0.3453
30	0.5578	0.3925	0.5954	0.3275

表 5 对 DTLZ7 的收敛性指标比较

**Table 5 Convergence metric comparison on DTLZ7**

目标数量	BZ 算法	JC 算法	SOEA 算法
6	0.4384	0.5649	0.4275
12	18.79	18.93	11.48
18	58.96	58.37	24.72
24	72.48	73.06	36.44
30	85.15	87.17	78.53

表6 对DTLZ7的覆盖率指标比较

Table 6 Coverage metric comparison on DTLZ7

目标数量	$C(\text{SOEA}, \text{BZ})$	$C(\text{BZ}, \text{SOEA})$	$C(\text{SOEA}, \text{JC})$	$C(\text{JC}, \text{SOEA})$
6	0.5682	0.2875	0.5974	0.2256
12	0.5834	0.2928	0.6127	0.2286
18	0.5976	0.2939	0.6384	0.2483
24	0.6014	0.2951	0.6528	0.2463
30	0.6327	0.2949	0.6573	0.2459

## 4 结论

在分析高维多目标优化问题特点和研究现有多目标优化求解算法的基础上,得出如下结论:

1) 高维多目标优化问题 Pareto 非支配解  $x^*$  的求取,可以通过分解为  $f_1, f_2, \dots, f_m$  个子目标,并在子目标函数值排序的基础上获得,且此法获得的 Pareto 非支配解完备。

2) 解的适应度值求取时,在基于归一化函数差值的 Minkowski 距离计算方法中采用随机在  $[1, H]$  范围内选取  $p$  值的方法可以加快算法的收敛速度。

3) 将本文提出的 SOEA 算法与 BZ、JC 算法比较,得出 SOEA 算法在收敛性和覆盖率两个指标上均优于 BZ、JC 算法。运行的时效性 JC 算法最优,SOEA 算法次之,BZ 算法最差。

## 参考文献 (References)

- [1] 孔维健,丁进良,柴天佑. 高维多目标进化算法研究综述[J]. 控制与决策,2010,25(3):321-326.  
Kong W J, Ding J L, Chai T Y. Survey on large-dimensional multi-objective evolutionary algorithms [J]. Control and Decision, 2010, 25(3):321-326 (in Chinese).
- [2] 公茂果,焦李成,杨咚咚,等. 进化多目标优化算法研究[J]. 软件学报,2009,20(2):271-289.  
Gong M G, Jiao L C, Yang D D, et al. Evolutionary multi-objective optimization algorithm [J]. Journal of Software, 2009, 20(2):271-289 (in Chinese).
- [3] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2):182-197.
- [4] Zitzler E, Laumanns M, Thiele L. SPEA2: Improving the strength Pareto evolutionary algorithm, TIK-Rep103 [R]. Lausanne: Swiss Federal Institute of Technology, 2001.
- [5] Corne D W, Jerram N R, Knowles J D, et al. PESA-II: Region-based selection in evolutionary multi-objective optimization [C] // Proceeding of the Genetic and Evolutionary Computation Conference, GECCO 2001. San Francisco: Morgan Kaufmann Publishers, 2001:283-290.
- [6] Khare V, Yao X, Deb K. Performance scaling of multi-objective evolutionary algorithms [C] // Proceeding of the 2nd Conference on Evolutionary Multi-Criterion Optimization, EMO 2003. Berlin: Springer-Verlag, 2003:376-390.

- [7] Bandyopadhyay S, Bhattacharya R. Solving multi-objective parallel machine scheduling problem by a modified NSGA-II [J]. Applied Mathematical Modelling, 2013, 37(10-11):6718-6729.
- [8] Saxena D K, Deb K. Non-linear dimensionality reduction procedure for certain large-dimensional multi-objective optimization problems; Employing correntropy and a novel maximum variance unfolding [C] // Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization. Berlin: Springer-Verlag, 2007:772-787.
- [9] Hernández-Díaz A G, Santana-Quintero L V, Coello Coello C A, et al. Pareto-adaptive  $\epsilon$ -dominance [J]. Evolutionary Computation, 2007, 15(4):493-517.
- [10] 刘立佳,李相民,颜骥. 解决高维多目标优化的分组进化算法[J]. 四川大学学报,2013,45(1):118-122.  
Liu L J, Li X M, Yan J. Group divided dimensional reduction evolutionary algorithm for multi-objective optimization [J]. Journal of Sichuan University, 2013, 45(1):118-122 (in Chinese).
- [11] Wagner T, Beume N, Naujoks B. Pareto-, aggregation-, and indicator-based methods in many-objective optimization [C] // Lecture Notes in Computer Science. Berlin: Springer-Verlag, 2007:742-756.
- [12] Hughes E J. Multiple single objective Pareto sampling [C] // Congress on Evolutionary Computation (CEC03). Piscataway, NJ: IEEE Press, 2003:2678-2684.
- [13] di Piero F. Many objectives evolutionary algorithms and applications to water resources engineering [D]. Exeter: University of Exeter, 2006.
- [14] 巩敦卫,季新芳,孙晓燕. 基于集合的高维多目标优化问题的进化算法[J]. 电子学报,2014,42(1):77-83.  
Gong D W, Ji X F, Sun X Y. Solving many-objective optimization problems using set-based evolutionary algorithms [J]. Acta Electronica Sinica, 2014, 42(1):77-83 (in Chinese).
- [15] Brockhoff D, Zitzler E. Are all objectives necessary? On dimensionality reduction in evolutionary multi objective optimization [C] // Parallel Problem Solving from Nature. Berlin: Springer, 2006:533-542.
- [16] Brockhoff D, Zitzler E. Objective reduction in evolutionary multi objective optimization: Theory and applications [J]. Evolutionary Computation, 2009, 17(2):135-166.
- [17] Jaimes A L, Coello C A C, Chakraborty D, et al. Objective reduction using a feature selection technique [C] // Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 08). Atlanta: ACM Press, 2008:673-680.
- [18] Deb K, Thiele L, Laumanns M, et al. Scalable multi objective optimization test problems [C] // Congress on Evolutionary Computation (CEC02). Piscataway, NJ: IEEE Press, 2002:825-830.
- [19] Knowles J, Corne D. On metrics for comparing non-dominated sets [C] // Proceeding of the IEEE Congress on Evolutionary Computation. Piscataway, NJ: IEEE Press, 2002:711-716.
- [20] 时宝,王兴平,盖明久. 泛函分析引论及其应用 [M]. 北京:国防工业出版社,2009:104-110.  
Shi B, Wang X P, Gai M J. Introduction to functional analysis with application [M]. Beijing: National Defense Industry Press, 2009:104-110 (in Chinese).

# Many-objective optimization based on sub-objective evolutionary algorithm

LEI Yuyao, JIANG Wenzhi\*, LIU Lijia, MA Xiangling

(Department of Ordnance Science and Technology, Naval Aeronautical and Astronautical University, Yantai 264001, China)

**Abstract:** many-objective optimization is widely used in engineering area. There are some flaws to deal with many-objective optimization problem which the number of objectives exceeded three. The method which could chose proper individual solution is very crucial to solve high-dimension many-objective optimization problem. A sub-objective evolutionary algorithm (SOEA) was put forward to solve this problem. It was given in an abstract way to get the non-dominance solutions of high-dimension many-objective optimization problem. Firstly, the value of sub-objective function was sorted, and then partial Pareto non-dominance solutions of evolutionary set were obtained quickly. By using the information of sorting, it could reduce the times of solution comparison in evolutionary set and could get the solutions quickly. A uniform difference Minkowski distance algorithm and “ $k$ -neighbor” strategy were applied to compute fitness function. By using this method, it could improve the convergence speed to approach Pareto non-dominance solutions. Compared with the algorithms which can solve many-objective optimization problem for computing standard testing functions, it was showed the better performance of the SOEA algorithm.

**Key words:** many-objective optimization; sub-objective evolutionary algorithm; Pareto non-dominance solution set; Minkowski distance; genetic algorithm

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0715

# 基于带宽优化的载波跟踪算法

牟伟清, 刘荣科\*, 宋青平, 段瑞枫

(北京航空航天大学 电子信息工程学院, 北京 100191)

**摘 要:** 在现有的高动态微弱信号的载波跟踪算法中, 针对锁频环(FLL)辅助锁相环(PLL)载波跟踪算法, 环路调整不连续, 易出现跟踪失锁的问题, 给出了FLL与PLL的更优组合算法, 并确定了组合环路状态转换的过程以及环路状态转换的门限值, 从而优化环路性能; 针对带宽调整不准确影响环路跟踪性能的问题, 分析得出最优带宽值, 确定了试探法带宽调整策略, 对环路带宽进行实时调整, 最终在信噪比为3 dB, 且存在加加速度分量时, 环路的跟踪误差达到3 Hz左右。

**关键词:** 高动态; 低信噪比; 载波跟踪; 锁频环; 锁相环

**中图分类号:** V443+.1

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1918-07

在航天测控系统中, 由于通信传输距离远及载体相对运动剧烈, 使得接收信号具有低信噪比高动态的特点<sup>[1]</sup>. 在目前测控系统中, 接收机的射频部分尚不能完全实现数字化, 一般采用数字中频方案<sup>[2]</sup>. 载波跟踪是对频率及相位的精细估计, 是决定中频接收机性能的重要环节, 常见的算法有极大似然估计法(MLE)<sup>[3]</sup>、扩展卡尔曼滤波法(EKF)<sup>[4-5]</sup>、叉积自动频率控制环(CPAFC)<sup>[6]</sup>、锁频辅助锁相环法(FLL-PLL)<sup>[7-8]</sup>等. MLE是对频率及相位的精细估计, 其计算复杂度高且实时性差. EKF是非线性方程的线性化处理, 对输入信号当前状态迭代估计, 估计精度高, 但其结构复杂, 不利于高速运算, 对相位跟踪时, 收敛速度慢. CPAFC是基于鉴频器的负反馈系统, 对高动态信号有较强适应性, 但跟踪精度较低, 对微弱信号跟踪性能较差. FLL-PLL性能稳定, 既发挥了FLL动态跟踪能力, 又利用了PLL跟踪精度高的优势,

但在目前的FLL-PLL组合算法中, FLL-PLL算法存在环路状态转变过渡不连续的问题, FLL-FPLL算法存在跟踪的锁定后存在精度相对较低的问题; 并且两种算法中未针对状态转换后进行带宽调整; 未分析状态转换的过渡门限; 在实际的应用中影响环路跟踪的性能. 本文针对FLL-PLL组合算法提出了更优的组合算法, 分析算法中状态转变时机; 为了得到更优跟踪性能, 确定了环路的最优带宽值以及环路调整策略.

## 1 FLL-PLL 载波跟踪算法

### 1.1 FLL 和 PLL 鉴别器的设计

根据信号的体制(DS\_BPSK)确定所需要的鉴别器, 同时需要抑制载波, 采用二象限反正切鉴别器, 可以在高低信噪比下都能接近最佳,  $I(I_{ps}(k))$ 、 $Q(Q_{ps}(k))$ 两路信号在进入鉴别器前

收稿日期: 2014-11-18; 录用日期: 2014-12-19; 网络出版时间: 2015-02-25 15:32

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150225.1532.003.html

基金项目: 国家“863”计划(2012AA120602); 教育部新世纪优秀人才支持计划(NCET-12-0030); 国家自然科学基金重大研究计划(91438116)

作者简介: 牟伟清(1986—), 男, 山东烟台人, 博士研究生, weiqing0017@163.com

\* 通讯作者: 刘荣科(1973—)男, 教授, rongke\_liu@buaa.edu.cn, 主要研究方向为空天信息传输与处理、无线多媒体通信、专用集成电路设计.

**引用格式:** 牟伟清, 刘荣科, 宋青平, 等. 基于带宽优化的载波跟踪算法[J]. 北京航空航天大学学报, 2015, 41(10): 1918-1924.  
Mu W Q, Liu R K, Song Q P, et al. Carrier tracking algorithm based on bandwidth optimization[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1918-1924 (in Chinese).

需要对积分结果作叉积 ( $P_{\text{Cross}}$ ) 和点积 ( $P_{\text{Dot}}$ ), 得到鉴频器和鉴相器的结构图如图 1 所示。

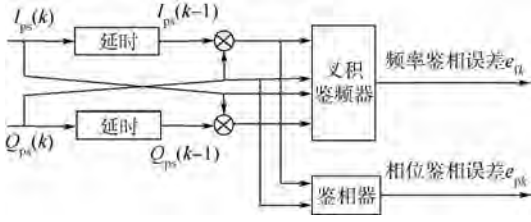


图 1 鉴频器和鉴相器的结构

Fig. 1 Structure of frequency discriminator and phase discriminator

鉴相过程中得出的叉积 ( $P_{\text{Cross}}$ ) 和点积 ( $P_{\text{Dot}}$ ) 误差分别为

$$\begin{cases} P_{\text{Cross}} = 0.25A^2 D(k) D(k-1) \cdot R(\varepsilon(k)) (\varepsilon(k-1)) \cdot \text{sinc}(\Delta f_d(k) \cdot \pi T) \cdot \text{sinc}(\Delta f_d(k-1) \cdot \pi T) \cdot \sin(\varphi_k - \varphi_{k-1}) \\ P_{\text{Dot}} = 0.25A^2 D(k) D(k-1) \cdot R(\varepsilon(k)) (\varepsilon(k-1)) \cdot \text{sinc}(\Delta f_d(k) \cdot \pi T) \cdot \text{sinc}(\Delta f_d(k-1) \cdot \pi T) \cdot \cos(\varphi_k - \varphi_{k-1}) \end{cases} \quad (1)$$

式中:  $A$  为正余弦信号幅度;  $D(k)$  为调制数据;  $\varepsilon(k)$  为伪码相位估计偏差;  $R$  为伪随机码相关函数;  $\Delta f_d(k)$  为多普勒频移估计残差;  $T$  为积分时间;  $\varphi_k$  为相位。

采用二象限反正切误差鉴别器对叉积和点积结果鉴别得到输出的频率误差为

$$\begin{aligned} e_{fk} &= \arctan(P_{\text{Cross}}/P_{\text{Dot}}) = \\ \theta_k - \theta_{k-1} &= 2\pi \cdot \Delta f_d(k) \cdot T \end{aligned} \quad (2)$$

式中:  $\theta_k$  为输出相位误差。

可见, 误差量与单位时间内的频率变化成正比, 可以控制载波 NCO 达到频率跟踪的目的。斜率与信号幅度无关, 频率鉴别范围为  $[-1/4T, 1/4T]$ 。

对于鉴相器, 由于其不受数据跳变的影响, 因此可采用四象限反正切鉴相算法, 得到输出的相位误差为

$$e_{pk} = \arctan 2(Q_{ps}, I_{ps}) = \theta_k \quad (3)$$

式中:  $I_{ps}$  为同相支路;  $Q_{ps}$  为正交支路。

在输入误差为  $\pm \pi$  范围内保持为线性。在高低信噪比下均为最佳, 斜率与信号幅度无关。

## 1.2 环路滤波器的设计

高动态的信号一般包括加速度和加加速度分量, 为了克服这两项动态应力误差, 需要对环路的阶数进行确定。

对于二阶锁频环, 其传递函数<sup>[9]</sup>为

$$H(s) = \frac{\sqrt{2} b_1 s + b_1^2}{s^2 + \sqrt{2} b_1 s + b_1^2} \quad (4)$$

式中:  $b_1^2 = \Delta \omega_n / \sqrt{2}$ ,  $\Delta \omega_n$  为自然角频率的阶跃量;  $b_1 = 4\sqrt{2}/3B_n \approx 1.886 B_n$ ,  $B_n$  为环路带宽。

加速度的数学模型为频率的一次斜升信号, 据此二阶锁频环 FLL 的稳态相位误差为

$$\theta_e(\infty) = \lim_{s \rightarrow 0} (s \cdot \theta_e) = \lim_{s \rightarrow 0} s \cdot H_e(s) \cdot \theta_1(s) =$$

$$\lim_{s \rightarrow 0} s \cdot \frac{s^2}{s^2 + 2b_1 s + b_1^2} \cdot \frac{2\pi f'_d}{s^3}$$

又:  $b_1 = 4\sqrt{2}/3B_n$

$$\theta_e(\infty) = \frac{9\pi}{16} \cdot \frac{f'_d}{B_n^2} \quad (5)$$

可见, 二阶锁频环跟踪频率斜升信号时, 存在频率稳态相差, 其大小与环路带宽平方成反比, 与频率变化率成正比。

对于三阶锁相环, 环路滤波器的闭环传递函数<sup>[10]</sup>:

$$H(s) = \frac{2bs^2 + 2b^2s + b^3}{s^3 + 2bs^2 + 2b^2s + b^3} \quad (6)$$

当输入信号存在加速度分量, 三阶锁相环 PLL 的稳态误差为 0。

当输入信号存在加加速度分量, 其数学模型为频率的二次斜升信号, 假设信号的二次变化率为  $f''_d$ , 输入信号的相位的变化为:  $\theta_2(t) = \frac{1}{3} \pi f''_d t^3$

$$\theta_e(\infty) = \frac{2\pi f''_d}{b^3} \quad (7)$$

又:  $b = 1.2B_n$  所以稳态相位误差:

$$\theta_e(\infty) = \frac{3.636 f''_d}{B_n^3} \quad (8)$$

因此, 三阶锁相环可以很好地克服加速度的动态应力误差, 对于加加速度误差会有一个稳定的相位误差, 因此在环路的跟踪锁定过程中, FLL 单独工作时需要将跟踪误差调整至接近稳态误差, 就认为锁定加速度信号分量, 在 PLL 单独工作、接近稳态误差时, 认为锁定加加速度分量。根据以上锁频环和锁相环的设计过程的推导, 可以确定二阶 FLL 与三阶 PLL 组合的方式, FLL 的输出为频率分量, 需要在环路滤波器的接入点提前一个积分环节, 以便使频率积分为相位, 添加至环路迭代中, 环路滤波器的结构如图 2 所示。

参数定义如下:

$$\begin{aligned} \omega_{nf} &= B_{nf}/0.53 & f_{c1} &= 1.414\omega_{nf} & f_{c2} &= \omega_{nf}^2 T \\ \omega_{np} &= B_{np}/0.7845 & p_{c1} &= 2.4\omega_{np} \\ p_{c2} &= 1.1\omega_{0p}^2 & p_{c3} &= \omega_{0p}^3 T \end{aligned}$$

式中:  $B_{np}$  为相位噪声带宽;  $B_{nf}$  为频率噪声带宽。

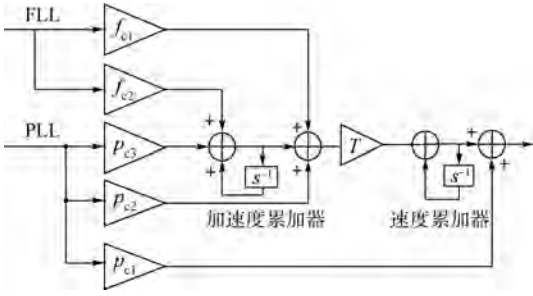


图2 二阶FLL辅助三阶PLL环路的滤波器结构

Fig. 2 Loop filter structure of second-order FLL auxiliary third-order PLL

FLL auxiliary third-order PLL

## 2 FLL-PLL 组合算法的优化

FLL-PLL的载波跟踪算法,既要保证锁定速度,又要提高环路的跟踪精度,FLL-PLL的载波跟踪方案的优化主要从组合环路的状态确定和状态转变的门限值确定进行.FLL-FPLL-PLL组合算法具有较优性能的环路闭合过程是:以FLL的形式牵引至频率锁定,然后过渡到FLL辅助PLL的方式,至相位锁定,之后再转换为PLL单独工作的精确跟踪模式,直到相位锁定丢失.FLL-FPLL-PLL组合环路状态之间的状态转换门限是根据蒙特卡罗仿真实验以及借鉴工程实现的经验值确定的,根据门限值进行状态转换可以增加跟踪环路转换连续性和跟踪精确度,减少环路失锁的概率.

### 2.1 组合环路的状态分析

针对目前的FLL-PLL组合跟踪算法,对环路跟踪曲线进行对比验证其合理性.

FLL-PLL的跟踪方案,即为FLL先牵引至锁定状态,然后PLL单独工作的方式.由于FLL是在PLL的前一个积分环节进入环路,从FLL的跟踪直接切换至PLL跟踪状态,显然环路状态的过渡不连续,会影响环路跟踪性能,甚至出现环路失锁的问题.

FLL-FPLL跟踪方案,FLL牵引而后转换为FLL和PLL同时工作的组合方式FPLL.由于二阶FLL的相位跟踪精度低于三阶PLL,并且在环路滤波器的工作过程中,是将FLL输出的频率误差和PLL输出的相位误差进行融合处理,会引入较大的环路误差,因此FLL和PLL同时工作的方式会使环路的误差增大.FLL-FPLL跟踪方案切换至FPLL状态后,在低信噪比的信道环境下融入噪声过大易使环路失锁.FLL-FPLL-PLL的跟踪过程中,当环路处于PLL单独精确跟踪阶段,相比前一阶段的FPLL环路误差会明显减小.

### 2.2 状态转换门限的确定

为了保证3种状态能够较优的切换,切换时机的判决门限尤为重要.如果FLL断开较早,跟踪环路无法锁定;如果FLL断开较晚,使环路的鉴相误差的抖动范围变大,影响跟踪精确度.

对于FLL牵引状态转换至FLL辅助PLL的锁定状态的门限值可以设为FLL的动态应力误差极限值,当考虑存在加速度分量时,根据式(5)的推导过程可知FLL剩余一个固定频差:

$$f_e = \frac{9\pi}{16} \cdot \frac{f'_d}{B_n^2} = 1.7671 \frac{d^2 R/dt^2}{B_n^2} \quad (9)$$

式中: $R$ 为发射机与接收机之间的视线距离. $f_e$ 存在一个极限值,当接近极限值时,可以增加PLL进入环路,环路进入FLL辅助PLL的锁定状态,图2中的 $p_{e1}$ 、 $p_{e2}$ 、 $p_{e3}$ 误差值添加到环路中,可以有较宽的跟踪范围,保证环路锁定速度.

对于FPLL状态转换至PLL的精确跟踪状态的检测门限设为锁相环的快捕带宽.锁相环的快捕带宽是环路不经过周期跳变而直接快速锁定输入信号范围<sup>[11]</sup>.工程应用中,设计信号的快捕方案时,要求快捕得到的信号相位估计值与输入信号的实际相位差处于锁相环的快捕带宽内,才能保证锁相环快速进入跟踪状态.锁相环的快捕带宽与环路的带宽相关,经过多次蒙特卡罗仿真实验,以及借鉴工程实现的经验值,考虑小于锁相环的单边最小带宽的2.5倍,检测门限值为

$$\Delta\omega_{\text{PLL}} = 2.5B_n \quad (10)$$

当达到门限值,图2中的 $f_{e1}$ 和 $f_{e2}$ 设为0,将FLL断开.环路进入PLL精确跟踪状态,实时地调整环路带宽,提高跟踪精度.单边最小带宽定义为<sup>[10]</sup>

$$B_n = \frac{\omega_n}{8} \left( \frac{22\xi^4 + 10\xi^2 + 1}{3\xi^3 + \xi} \right) \quad (11)$$

式中: $\omega_n$ 为自然震荡角频率; $\xi$ 为环路的阻尼系数,取值为0.707.

## 3 FLL-FPLL-PLL跟踪环路的优化

跟踪环路的优化主要是对环路带宽的实时调整,分为状态转换时刻的带宽调整和PLL单独工作时的带宽调整两方面.

为了在环路状态转变后环路的过渡更加连续,同时尽量提高跟踪精度,状态转换时刻的带宽选择至关重要,根据前面的分析,可按照以下规则进行:

1) FLL在牵引跟踪状态,带宽要尽量设置得宽,提高锁定速度.

2) 当转换至 FPLL 状态, FLL 环的环路带宽减少, PLL 带宽在调整范围内选择最大值.

环路状态切换至 PLL 单独工作的精确调整状态时, 环路已经进入锁定状态, 可以根据环路误差的来源, 优化环路带宽, 使环路误差达到最小值. 锁相环的相位测量误差源包括相位抖动误差和动态应力误差, 这两者存在矛盾, 当带宽变大, 相位抖动误差增大, 而动态应力误差减少. 而造成相位抖动的误差源又主要分为热噪声、振荡器引起的振荡器颤动以及阿伦偏差引起的振荡器颤动 3 种<sup>[12]</sup>.

PLL 的总误差函数关系式为

$$\delta_{\text{PLL}} = \sqrt{\delta_{\text{iPLL}}^2 + \delta_{\text{v}}^2 + \delta_{\text{A}}^2} + \frac{\theta_{\text{e}}}{3} \quad (12)$$

式中:  $\delta_{\text{iPLL}}$  为热噪声误差;  $\delta_{\text{v}}$  为由振荡器引起的振荡器颤动;  $\delta_{\text{A}}$  为由阿伦偏差引起的振荡器颤动.

由 PLL 振动源引起误差可以忽略, PLL 误差只考虑热噪声误差和动态应力误差两方面<sup>[12]</sup>. 将其代入式(12)得

$$\delta_{\text{PLL}} = \frac{360}{2\pi} \sqrt{\frac{B_{\text{n}}}{C/N_0} \left(1 + \frac{1}{2TC/N_0}\right)} + \frac{d^3 R/dt^3}{3\omega_{\text{n}}^3} \quad (13)$$

式中:  $C$  为输入环路的信号功率;  $N_0$  为单边噪声功率谱密度;  $\omega_{\text{n}} = B_{\text{n}}/0.7845$ ;  $d^3 R/dt^3$  为视线上加加速度, 设其动态值为  $30 \text{ g/s}$ , 载波频率为  $1.6 \text{ GHz}$ .

根据二象限反正切鉴别器, PLL 的跟踪门限<sup>[12]</sup>

$$\delta_{\text{PLL}} \leq 15^\circ \quad (14)$$

根据式(12)可知:

$$B_{\text{n}} > 0 \quad (15)$$

跟踪环路带宽一定要选得尽可能小以提供足够的环路信噪比  $\rho$ , 一般需要保证在  $15 \sim 20 \text{ dB}$  范围之内<sup>[13]</sup>:

$$15 \leq \rho = \frac{CS_{\text{L}}}{N_0 B_{\text{n}}} \leq 20 \text{ dB} \quad (16)$$

式中:  $S_{\text{L}}$  为环路的平方损耗, 一般取  $-3 \text{ dB}$  左右<sup>[13]</sup>.

由载噪比  $C/N_0$  和信噪比  $E_{\text{b}}/N_0$  之间的关系公式可知<sup>[14-15]</sup>:

$$C/N_0 = E_{\text{b}}/N_0 + 10 \lg R_{\text{b}} \quad (17)$$

式中:  $R_{\text{b}}$  为符号速率, 取  $1 \text{ kb/s}$ ;  $E_{\text{b}}/N_0$  取  $3 \sim 5 \text{ dB}$  的动态范围(微弱信号的  $E_{\text{b}}/N_0$  一般小于  $5 \text{ dB}$ ), 由此根据式(16)和式(17)可知:

$$10 \text{ Hz} \leq B_{\text{n}} \leq 50 \text{ Hz} \quad (18)$$

$$33 \text{ dB} \leq C/N_0 \leq 35 \text{ dB} \quad (19)$$

因此带宽优化问题变成了有约束的二元非线性优化问题:

Minimize:

$$\delta_{\text{PLL}} = \frac{360}{2\pi} \sqrt{\frac{B_{\text{n}}}{C/N_0} \left(1 + \frac{1}{2TC/N_0}\right)} + \frac{d^3 R/dt^3}{\omega_{\text{n}}^3}$$

Subject to:

$$\delta_{\text{PLL}} \leq 15$$

$$33 \text{ dB} \leq C/N_0 \leq 35 \text{ dB}$$

$$B_{\text{n}} > 0$$

$$10 \leq B_{\text{n}} \leq 50$$

采用 active-set 法求多变量函数的最小值点, 其计算过程如表 1 所示.

表 1 有效集算法的优化过程

Table 1 Optimization process of active-set algorithm

迭代次数	函数值	最大约束	方向导数	一阶最优值
0	7.13486	0		
1	6.48274	-4.844	-0.854	0.505
2	5.13631	0	-0.687	0.204
3	5.06308	0	-0.204	0.185
4	4.88276	0	-0.166	0.0543
5	4.85194	0	-0.0543	0.0208
6	4.84546	0	-0.0208	0.0043
7	4.84512	0	-0.00443	0.000479
8	4.84511	0	-0.000479	$1.27 \times 10^{-5}$

计算结果如下:

$$\delta_{\text{PLL.min}} = 4.8451$$

$$B_{\text{n}} = 16.354797$$

$$C/N_0 = 35.0000$$

根据优化结果可以得出如下结论:

1) 目标函数的最小误差为  $4.8451$ , 远小于 PLL 的失锁极限值  $15^\circ$ , 表明信号存在加加速度分量的情况下, 二阶 FLL 辅助三阶 PLL 的工作方式可以稳定的工作.

2) 带宽为  $16 \text{ Hz}$ 、载噪比为  $35 \text{ dB}$  时误差取得最小值, 验证了载噪比越高 PLL 的跟踪精度越高的结论.

3) 在带宽调整优化的过程中, 由于相位抖动误差和动态应力误差存在矛盾性, 并不是带宽值越小相位误差越少, 而是存在最优值. 因此在高动态的航天测控通信环境下, 可以根据其实际情况计算带宽最优值.

当进入 PLL 单独工作的 PLL 精确跟踪状态时, 已经进入锁定状态, 不需考虑锁定速度, 为了提高环路的跟踪精度, 按照带宽最优值调整带宽. 如果立即调整带宽至最优值, 容易出现环路失锁的问题; 并且由于环路的存在瞬时误差值过大的情况, 会造成带宽误调整, 因此需要增加环路调整

的约束条件,根据实际测试过程中提出了试探调整的办法.

试探法是指在一段时间内取环路滤波器的相位误差的平均值,当其均值稳定,可以试探减少带宽,在带宽减少后仍然观测相差的变化情况,如果相差变大就将带宽还原为调整前的值,如果相差变小,就将带宽保持为调整后的值.

鉴于环路会出现瞬时噪声过大的情况,通过取多次相差均值的方式,降低错误调整的机率.其数学模型为

$$B_{ns} = B_n + \alpha \frac{1}{n} \sum_{i=1}^n |\theta_{ei}| \quad (20)$$

式中: $B_{ns}$ 为环路动态调整带宽; $\alpha$ 为带宽调整系数,取值范围为 $[0, 1]$ .

根据以上的试探法的要求可以通过调整系数 $\alpha$ 的值,调整带宽,使带宽逐步地达到最优值.

### 4 仿真验证

为了验证本文提出的载波跟踪优化算法的性能,对信号的跟踪过程进行仿真.

**仿真 1** 为了验证 FLL-FPLL-PLL 载波跟踪优化算法的可靠性,与 FLL-PLL 和 FLL-FPLL 两种算法进行对比,在  $E_b/N_0 = 3$  dB,捕获结束后的剩余频偏  $|f_{rem}| = 200$  Hz、加速度  $|a_{rem}| = 50$  Hz/s 相同的条件下进行仿真,并且跟踪过程中未进行带宽调整,容易出现环路失锁问题.图 3 是 FLL-PLL 算法仿真结果,开始阶段 FLL 环已经锁定频率,但是在转变为 PLL 跟踪状态时,由于直接跳转至 PLL 跟踪阶段,出现过渡不连续的情况造成环路失锁;图 4 中环路进入 FPLL 状态后,并未断开 FLL<sup>[7,16]</sup>.由于 FLL 的调整误差较大,环路中引入噪声过大,使环路调整不稳定,在锁定一段时间后失锁.对比发现,在只有加速度的情况会出现失锁问题,当存在加加速度的动态环境下,环路失锁问题变得更频繁,因此在高动态低信噪比的环境下,FLL-FPLL-PLL 组合方案优势会更明显.

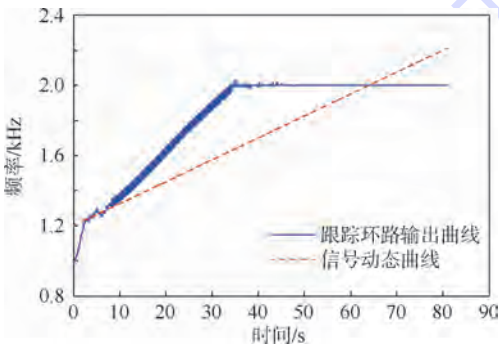


图 3 FLL-PLL 不连续环路滤波器输出

Fig. 3 Output of discontinuity FLL-PLL loop filter

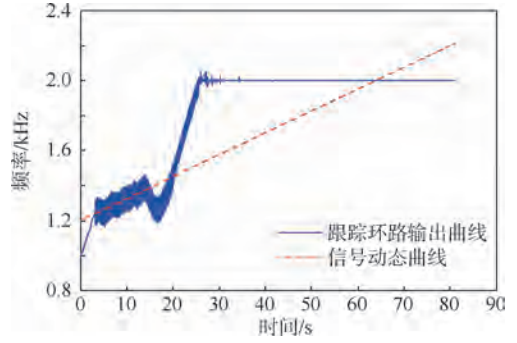


图 4 FLL-FPLL 的环路滤波器输出

Fig. 4 Output of FLL-FPLL loop filter

FLL-FPLL-PLL 载波跟踪优化算法的环路滤波器频率输出如图 5 所示,可以看出环路的状态转变和带宽调整过程:FLL 牵引状态可以快速地消除捕获阶段剩余的多普勒频移,快速达到锁定状态,转换至 FLL 辅助 PLL 状态,误差减少至快捕带宽,转变至 PLL 精确跟踪状态,进而根据提出的环路优化调整策略对带宽进行了调整,将带宽调整至最优值,可以看出环路滤波器的输出误差值不断减少.

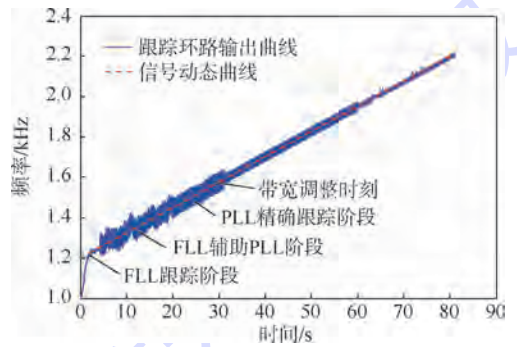


图 5 FLL-FPLL-PLL 环路输出

Fig. 5 Output of FLL-FPLL-PLL loop filter

**仿真 2** 为了验证最优带宽的存在性,以及有效集算法解算的带宽最优值的可靠性,在载噪比分别为 33 ~ 35 dB 的情况下,仿真带宽与总的噪声误差之间的关系如图 6 所示.

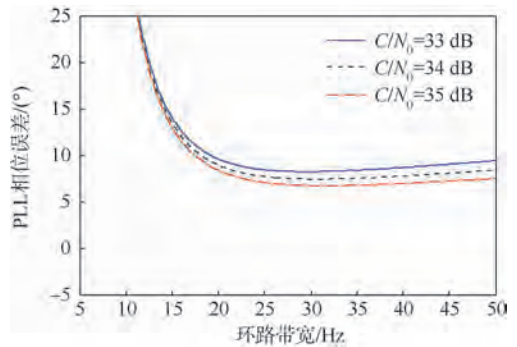


图 6 不同  $C/N_0$  带宽与噪声误差的关系

Fig. 6 Relationship between bandwidth and noise error under different  $C/N_0$

Fig. 6 Relationship between bandwidth and noise error under different  $C/N_0$

**仿真3** 为了验证跟踪更高动态信号的可靠性,以及 FLL-FPLL-PLL 算法和带宽调整策略的可靠性.以火星探测器进入阶段的接收信号为例,从变化多普勒频偏为  $-300 \sim 300$  kHz,变化率  $a$  为  $-800 \sim 800$  Hz/s,经过载波粗捕获后达到  $f_{\text{rem}} \leq 200$  Hz,  $a_{\text{rem}} \leq 50$  Hz/s,加加速度  $|a'_{\text{rem}}| = 5$  Hz/s<sup>2</sup>,中频中心频点为 70 MHz 的情况下,进行了蒙特卡罗仿真.图7给出了带宽优化算法跟踪误差仿真结果:环路进入 PLL 精确调整状态后,需要根据实际的跟踪环境,调整带宽使之达到最优值,在此过程中,会出现带宽调整的误操作,试探法调整带宽的策略就会调整至原带宽值,减少误调整的机率.图8给出跟踪过程中,环路的误差仿真曲线,可以看出环路误差在带宽的调整过程中不断地减少,相比传统的跟踪环路提高了环路跟踪精度.经过多次蒙特卡罗仿真验证锁相环的锁定时间大约为700 ms,与普通的锁相环的锁定时间相当.

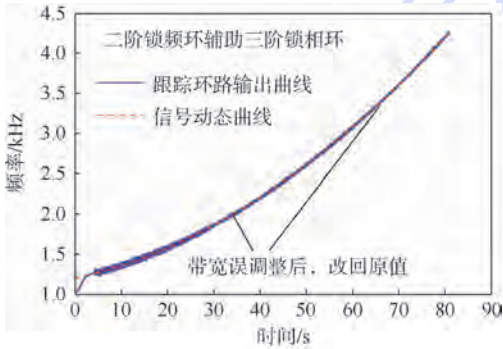


图7 带宽试探调整算法的验证

Fig. 7 Testing of tentative fitting bandwidth algorithm

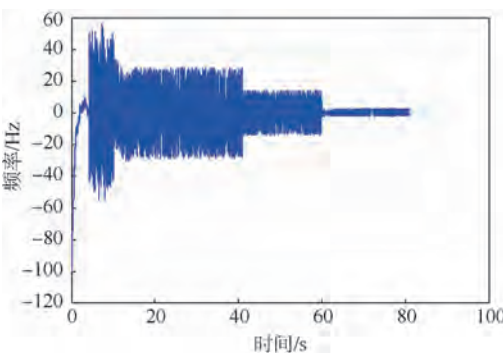


图8 跟踪环路误差曲线

Fig. 8 Loop tracking error curve

## 5 结论

1) FLL 和 PLL 的组合方式具有最佳组合方式,并分析确定了状态之间的转换时机,使状态转换更为连续.

2) 提出了试探法调整带宽和优化环路的方法,在高动态低信噪比的情况下,环路调整误差低

至 3 Hz,提高了跟踪精度,实现了高动态微弱信号精确跟踪.

## 参考文献 (References)

- [1] Fountain G H, Kusnierkiewicz D Y, Hersman C B, et al. The new horizons spacecraft [M]. New York: Springer, 2009: 23-47.
- [2] 唐友喜, 邵士海. 改进的基于二维扩频 WCDMA 非 Rake 接收机 [J]. 通信学报, 2005, 36(11): 93-99.  
Tang Y X, Shao S H. Improved non-Rake receiver based on two dimensional spread spectrum for WCDMA systems [J]. Journal on Communications, 2005, 36(11): 93-99 (in Chinese).
- [3] Li Y, Fu H, Kam P Y. Improved, approximate, time-domain ML estimators of chirp signal parameters and their performance analysis [J]. IEEE Transactions on Signal Processing, 2009, 57(4): 1260-1272.
- [4] Li W, Liu S, Zhou C, et al. High dynamic carrier tracking using Kalman filter aided phase-lock loop [C] // Wireless Communications, Networking and Mobile Computing. Piscataway, NJ: IEEE Press, 2007: 673-676.
- [5] Miao J F, Sun Y R, Liu J Y, et al. A Kalman filter based tracking loop in weak GPS signal processing [C] // International Conference on Fuzzy Systems and Knowledge Discovery. Piscataway, NJ: IEEE Press, 2009: 438-442.
- [6] 卢辉斌, 王伟伟, 谷青川. 低信噪比高动态条件下的载波同步技术 [J]. 信息与控制, 2010, 39(4): 451-454.  
Lu H B, Wang W W, Gu Q C. Carrier synchronization technique for low SNR and high dynamic condition [J]. Information and Control, 2010, 39(4): 451-454 (in Chinese).
- [7] 王俊, 李加琪, 吴嗣亮. 锁频环辅助下锁相环的跟踪误差分析 [J]. 北京理工大学学报, 2011, 31(7): 838-843.  
Wang J, Li J Q, Wu S L. Tracking error analysis of FLL-assisted-PLL [J]. Transactions of Beijing Institute of Technology, 2011, 31(7): 838-843 (in Chinese).
- [8] 程乃平, 任宇飞, 吕金飞. 高动态扩频信号的载波跟踪技术研究 [J]. 电子学报, 2003, 31(12A): 2147-2150.  
Cheng N P, Ren Y F, Lu J F. Study of carrier tracking for high-dynamic spread spectrum signals [J]. Acta Electronica Sinica, 2003, 31(12A): 2147-2150 (in Chinese).
- [9] 占巍, 张晓林, 常啸鸣. 跳频对 DS/FH 接收机载波跟踪精度影响分析 [J]. 北京航空航天大学学报, 2012, 38(10): 1358-1361.  
Zhan W, Zhang X L, Chang X M. Impact analysis on frequency hopping to carrier tracking precision in DS/FH hybrid spread spectrum receiver [J]. Journal of Beijing University of Aeronautics and Astronautics, 2012, 38(10): 1358-1361 (in Chinese).
- [10] 张厥盛, 郑继禹, 万心平. 锁相技术 [M]. 西安: 西安电子科技大学出版社, 2004: 38-39.  
Zhang J S, Zheng J Y, Wan X P. Phase-locked technology [M]. Xi'an: Xidian University Press, 2004: 38-39 (in Chinese).
- [11] 卢满宏, 周三文, 谌明, 等. 深空测控通信技术专题研究 [J]. 遥测遥控, 2007, 28(增刊): 11-16.  
Lu M H, Zhou S W, Chen M, et al. Research on deep space TT&C and communication [J]. Journal of Telemetry, Tracking and Command, 2007, 28(Sup.): 11-16 (in Chinese).

- [12] 寇艳红. GPS原理与应用[M]. 2版. 北京: 电子工业出版社, 2007: 131-141.
- Kou Y H. Understanding GPS principles and applications[M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2007: 131-141 (in Chinese)
- [13] 闫春生. 用于深空的自主软件无线电接收机[M]. 北京: 清华大学出版社, 2010: 180.
- Yan C S. Autonomous software-defined radio receiver for deep space applications [M]. Beijing: Tsinghua University Press, 2010: 180 (in Chinese).
- [14] Vilmrotter V A, Hinedi S, Kumar R. Frequency estimation techniques for high dynamic trajectories[J]. IEEE Transactions on Aerospace and Electronic Systems, 1989, 25(4): 559-577.
- [15] Satorius E, Estabrook P, Wilson J. Direct-to-earth communications and signal processing for mars exploration rover entry, descent, and landing, IPN Progress Report 42-153[R]. Washington, D. C.: NASA, 2003: 1-35.
- [16] 匡鸿博, 茅旭初, 王永凯. 一种用于高动态环境的 GPS 信号跟踪方法[J]. 上海交通大学学报, 2012, 46(6): 854-858.
- Kuang H B, Mao X C, Wang Y K. A tracking method for GPS signal under high dynamic condition[J]. Journal of Shanghai Jiao Tong University, 2012, 46(6): 854-858 (in Chinese).

## Carrier tracking algorithm based on bandwidth optimization

MU Weiqing, LIU Rongke\*, SONG Qingping, DUAN Ruifeng

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** In the existing high dynamic weak signal carrier tracking algorithm, frequency locked loops (FLL) auxiliary phase locked loops (PLL) carrier tracking algorithm has the problem that the loop adjusted discontinuously, and easily lost lock. More optimal combination FLL with PLL algorithm and transition process of combination loop were given, threshold for state transition was determined. Aiming at the problem that bandwidth adjustment is not accurate, the optimal PLL loop bandwidth and the bandwidth adjustment strategy were confirmed, the loop bandwidth was adjusted in real-time. In the presence of jerk cases and signal-to-noise ratio is 3 dB, the tracking error is around 3 Hz.

**Key words:** high-dynamic; low signal-to-noise ratio; carrier tracking; frequency locked loops; phase locked loop

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0771

# 汉语双模情感语音数据库标注及一致性检测

景少玲, 毛峡\*, 陈立江, 张娜娜

(北京航空航天大学 电子信息工程学院, 北京 100191)



**摘 要:** 对缺少含有丰富情感标注信息的情感语音数据库问题, 建立了一个包含语音和电声门图仪(EGG)信息的汉语双模情感语音数据库, 并对其进行了标注和一致性检测. 首先, 根据情感语音数据库的特色制定了详细的标注规则和方法, 由5名标注者按照制定的标注规则对情感语音数据库进行标注. 其次, 为了确保情感语音数据库的标注质量和测试标注规则的完整性, 标注者在正式标注之前先进行了测试性标注, 测试语音包含280条语音(7种情感×2名说话人×20条语音). 最后, 根据语音标注规则设计了相应的一致性检测算法. 结果表明, 在5 ms的时间误差范围内, 5名标注者对相同语音标注的一致性平均可以达到60%以上, 当误差范围增大至8 ms和10 ms时, 一致性平均可提高5%和8%. 实验说明5名标注者对语音的理解较一致, 制定的标注规则比较完整, 情感语音数据库的质量也较高.

**关键词:** 汉语; 双模; 情感语音数据库; 语音标注; 一致性检测

**中图分类号:** TP391.4

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1925-10

情感语音中含有丰富的情感交流信息, 在人们日常生活中起着非常重要的作用, 对其研究可以进一步了解人类相应的心理状态, 情感语音识别在情感计算和智能交互领域有重要意义. 情感语音数据库是情感语音识别、合成等研究的基础. 由于情感的复杂性、数据收集的困难性、录制标准的不统一性、情感分类目的和任务的差异性等, 导致目前大多数研究机构都是按照各自的科研需求建立相应的情感语音数据库<sup>[1-2]</sup>. 情感化的智能交互要求机器具有接近人类的语音情感识别和表达能力<sup>[3]</sup>, 为了实现该目标, 建立一个真实自然并且拥有完整情感标注信息的优质情感语音数据库是十分重要的.

电声门图仪(Electroglottography, EGG)是测

量说话人声门开启与闭合状态以及声门振动方式的仪器. 客观地反应了发元音时声带开闭的状态, 对情感语音识别研究有重要作用. 目前大多数情感语音数据库都不含EGG信息, 而本文的汉语双模情感语音数据库中采用了EGG信息.

## 1 情感语音数据库概况

2003年在Ververidis和Kotropoulos总结的情感语音数据库中, 和汉语相关的仅列出了微软公司从电视剧截取录音片段的数据库<sup>[4]</sup>, 表1总结了语音情感识别研究中常用的一些情感语音数据库的基本情况.

从表1中可以看出, 目前的情感语音数据库

收稿日期: 2014-12-08; 录用日期: 2015-01-16; 网络出版时间: 2015-03-18 14:37

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150318.1437.002.html

基金项目: 高等学校博士学科点专项科研基金(20121102130001); 中央高校基本科研业务费专项资金(YWF-14-DZXY-015)

作者简介: 景少玲(1987—), 女, 山西永济人, 博士研究生, jingshaoling2013@163.com

\* 通讯作者: 毛峡(1952—), 女, 浙江义乌人, 教授, moukyou@buaa.edu.cn, 主要研究方向为人工智能、模式识别、情感计算、人机交互及红外目标检测、跟踪、识别和评价等.

**引用格式:** 景少玲, 毛峡, 陈立江, 等. 汉语双模情感语音数据库标注及一致性检测[J]. 北京航空航天大学学报, 2015, 41(10): 1925-1934. Jing S L, Mao X, Chen L J, et al. Annotations and consistency detection for Chinese dual-mode emotional speech database[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1925-1934 (in Chinese).

存在的共同问题是:①包含 EGG 信息的情感语音数据库非常少,汉语中尚未有;②大多数库规模小,如 NATURAL、ESMBS;③录制标准不统一,如情感类别多样;④一些库中记录的话语质量差,如 KISMET,采样率为 8 kHz;⑤一些库不提供音标等标注信息,如 BabyEars,对提取语言内容等信息造成困难<sup>[5]</sup>.

一个完整的情感语音数据库除了包括基本的语音文件外,还应该有相应的标注文件,详细的情感语音标注可以为情感语音合成和识别提供可靠

的训练集和测试集,因此对情感语音数据库进行标注就显得极其重要. 目前国内的情感语音数据库尚未有与情感有关的标注信息,国外有一些情感语音数据库进行了与情感有关的标注,如英语情感语音数据库<sup>[6]</sup>,德语音视频自然语音库<sup>[7]</sup>和日语 JST/CREST ESR 库<sup>[8]</sup>;马来西亚语音库<sup>[9]</sup>收集了对话语音,标注了不平滑以及填充间断等语言现象,但未标注情感信息;日本 UU (Utsumomiya University) 语音库<sup>[10]</sup>标注了副语言信息<sup>[11]</sup>.

表 1 部分情感语音数据库总结

Table 1 Summary of parts of emotional speech database

情感语音数据库	语言	说话人	容量/条	使用权	EGG	情感类型
CASIA <sup>[12]</sup>	汉语	4 名非专业人员	9 600	非公开	否	高兴,悲伤,生气,惊吓,中性
ACCorpus_SR <sup>[2,13]</sup>	汉语	50 名非专业人员	50 000	非公开	否	高兴,生气,恐惧,悲伤,中性
ESMBS <sup>[14]</sup>	汉语和缅甸语	6 名中国人和 6 名缅甸人	720	非公开	否	高兴,生气,伤心,嫌恶,恐惧,惊奇
NATURAL <sup>[15]</sup>	汉语	11 名人员呼叫中心对话	388	非公开	否	生气,中性
Beihang University <sup>[16]</sup>	汉语	7 名非专业人员	2 100	部分公开	否	生气,高兴,伤心,沮丧,惊奇
CLDC <sup>[17]</sup>	汉语	4 名非专业人员	1 200	非公开	否	高兴,生气,惊奇,恐惧,中性,伤心
Hu 等 <sup>[18]</sup>	汉语	8 名非专业人员	1 600	非公开	否	高兴,生气,伤心,恐惧,中性
Berlin emotional database <sup>[19]</sup>	德语	10 名专业人员	800	公开	是	高兴,生气,伤心,恐惧,嫌恶,厌烦,中性
FERMUS III <sup>[20]</sup>	德语和英语	13 名人员;汽车噪音环境	2 829	公开	否	高兴,生气,伤心,嫌恶,惊奇,中性
Danish emotional database <sup>[21]</sup>	丹麦语	4 名专业人员	500 多	公开	否	高兴,生气,伤心,惊奇,中性
SUSAS <sup>[22]</sup>	英语	32 名非专业人员	16 000	公开	否	生气,恐惧,沮丧,焦虑,害怕等情绪
KISMET <sup>[23]</sup>	英语和美图	3 名非专业人员	1 002	非公开	否	赞成,关心,禁止,吮吸,中性
BabyEars <sup>[24]</sup>	英语	12 名非专业人员	509	非公开	否	赞成,关心,禁止

本文结合中国社会科学院(简称社科院)制定的标注符号和研究需要,主要做了以下工作:首先根据北京航空航天大学(简称北航)录制的双模情感语音数据库制定了详细的语音数据库标注规则,考虑到标注的规范化,采用了社科院设计的一部分标注符号;然后对 5 名标注者进行培训,培训之后标注者开始标注语音数据库中的语音;最后为了确保标注质量,根据标注规则设计了一致性检测算法,检验几名标注者标注结果的一致性好坏. 本文的工作与其他语音数据库不同之处在于:①录制的情感语音数据库为双模的,包括语音波形图和 EGG 信息,信息更丰富;②除了标注基本的汉字转换和音节等信息外,还标注了与清音、静音、浊音、情感、副语言信息和重音等相关的信息,目前国内其他情感语音数据库中未有如此完整的标注信息;③设计了一致性检测算法并对标注过的语音进行了一致性测试,实验结果表明不同标注者标注的一致性较高,说明语音数据库的质量较高.

## 2 汉语双模情感语音数据库的设计

为了获得高质量的情感语音数据库,本文建

立了北航汉语双模情感语音数据库 (Chinese Dual-mode Emotional Speech Database, CDESD), 并对数据采集过程以及对被试者的诱导方式进行了设计.

### 2.1 实验条件

双模是指语音录制过程中同时采集了语音信息和电声门图信息. 说话人包括 20 名年龄在 21~23 岁的大学生(13 男,7 女),每位说话人根据预先设定好的 20 条语句进行发音,每句话表达 7 种不同的情感,分别是悲伤、高兴、害怕、惊奇、平静、生气、嫌恶,每条语句发音 3 到 5 遍不等,后期经过拆分和删选,每条保留 3 遍,共得到 8 400 条语句. 表 2 描述了北航情感语音数据库录音文本,要求文本无某一方面的情感倾向,有较高的情感自由度,能施加各种情感进行分析比较,并且是口语化的陈述句. 录音时间选择在较安静的晚上,地点是在密闭空旷的实验室. 录音设备采用笔记本电脑,外接创新 Audigy 2 NX USB 接口声卡,台式麦克风,录音软件采用 GoldWave V5.14,采样率为 44 100 Hz,16 b 量化.

表 2 北航情感语音数据库录音文本

Table 2 Recorded text of Beihang emotional speech database

语句编号	语句内容
1	啊,你可真伟大呀
2	快点干
3	这下完了
4	啊,下雨了
5	太棒了
6	我真的以为你是这个意思
7	我在论文上看到你的名字了
8	AC 米兰赢球了
9	我这次考试刚刚通过
10	今天是星期天
11	你这人
12	电话铃响了
13	他就快来了
14	路上人真多啊
15	明天我要搬家了
16	这件事是他干的
17	你这段时间变瘦了
18	过两天学校就要开学了
19	昨天晚上我做了一个梦
20	有一辆车向我们开过来了

## 2.2 数据采集过程的设计

由工作人员为被试者带上电声门图仪的采集工具,面前放置麦克风。①采集前先进行信号的测试,尤其注意电声门图和语音信号是否有削顶现象,适当地调整电脑音量或者电声门图和语音采集的放大倍数,声音大小也可以通过调整麦克与人的距离实现。②同时保证测试环境的绝对安静,测试时告知被试者不能随便移动,周围不能发出声响。③被试者以尽量真实的情感表达出 7 种不同的、区分度较高的情感状态,不需要过分夸张的成分。④测试中,需要与被试者多进行沟通,引导其感情,舒缓其情绪,使其放松,当被试者表达不理想时,重复 4~6 次甚至更多。⑤实验中,应让被试者有足够的休息时间,期间也需要尽量多的沟通交流,调节被试者状态。

## 2.3 提示的设计

为了获得被试者的 7 种情感数据,实验前对表 2 中的每一条语音及相应的情感,都进行了情景设置,建立了一个完整的含有 140 场情景的表。例如对第 4 条语句“啊,下雨了”,当表达悲伤情感时,设置情景为:“一位诗人,想起怀才不遇,看见阴沉的天空,淅沥的小雨,有感而发:啊,下雨了!”;当表达高兴情感时,设置情景为:“入春以来,整个地区没有掉过一滴雨,以土地为生的农民们躲在家中感慨春雨贵如油。突然有一天,外边雷声大作,你马上跑入雨中,高兴地说:老天开眼,

啊,下雨了!”。

## 3 情感语音数据库标注规则

社科院语音研究室根据国际语音学者团体制定的可机读的音段标注体系 SAMPA (Speech Assessment Methods Phonetic Alphabet) 和韵律标注体系 ToBI (Tones and Break Indices) 设计了适用于汉语标注的 SAMPA-C 和 ToBI-C 等标注符号体系,并根据指定的符号标注了很多语音数据库,例如 DFEIC (Database of Functional and Emotional Intonation in Chinese) 语音数据库<sup>[25]</sup>,该库收集了电视节目和谈话节目的语音数据,主要用于情感中的声调识别,但是仅标注了两层信息,分别是音节层和韵律层;自然口语库<sup>[26]</sup>总共标注了 7 层信息,包括韵律和语言学标注,主要用于语音合成和韵律识别。本文的汉语双模情感语音数据库不但包含 EGG 信息,而且加入了情感信息的标注,为情感识别提供丰富的信息。

根据 CDESD,制定了详细的标注规则。标注规则包括标注的一致性、连贯性,标注符号的易记性,但同时还需要遵循的一条原则是允许标注的不确定性和差异性存在,即允许不同的标注者对同一条语音中的情感、重音、声调等有不同的理解,避免向用户提供错误信息。标注软件使用 Praat 5.3.59,标注内容一共包括 8 层:第 1 层是文字转换层(HZ),第 2 层是音节层(PY),第 3 层是声韵母层(SY),第 4 层是清音静音浊音层(SUV),第 5 层是副语言信息层(PARAL),第 6 层是情感层(EMO),第 7 层是重音指数层(ST),第 8 层是语句功能层(FU)。下面详细介绍每一层的标注内容和标注规则。

### 3.1 文字转换层(HZ)

文字转换层的主要工作就是将语音转化为文字,也是语音标注中最基本的标注内容,文字标注必须包括文字信息和基本的副语言信息,文字标注的要点是汉字和副语言信息的准确标注,其中副语言信息采用通用的副语言符号表示。本文的副语言符号参考社科院语言研究所语音研究室制定的一些基本的副语言符号<sup>[27]</sup>。

### 3.2 音节层(PY)

音节层主要标注正则的拼音和声调,声调标注在音节之后,汉语中的声调包括轻声、一声、二声、三声和四声,分别用 0、1、2、3、4 表示,声调标在音节之后。例如:a0、tian1、ren2、guo3、fang4。标注时标注者通过听语音将原始语音划分为音节,音节的划分也是标注中的难点,为了保证不同标

注者的标注误差在 5 ms 内,同时也规定了更加详细的标注细则.

### 3.3 声母/韵母层 (SY)

汉语中,一个完整的音节一般由声母和韵母组成.声母是在韵母之前的辅音,韵母是除去声母和音调之外的部分.声母、韵母标注时,对应于拼音层,将每个音节的声母、韵母、声母和韵母中间的过渡发音进行标注,声调标注在韵母之后.为了区分普通话中正则的声调,声韵母层的声调标注为实际发音的声调,分别用\_1、\_2、\_3、\_4 表示,轻声用“\_0”表示.例如:shuo 标注为:声母段为 sh + u,韵母段为 uo\_1.

### 3.4 清音静音浊音层 (SUV)

清音、静音和浊音层的标注目的在于划分语

音中的清音 S (Silence)、静音 U (Unvoiced)、浊音 V (Voiced). 如图 1 所示,图中包含 3 个子图,从上至下依次代表:语音波形图 (channel 1)、电声门图 (channel 2)、语谱图.判断静音、清音和浊音的一般简易方法是:观察电声门图波形,电声门图中有波形的部分代表浊音,电声门图中无波形而语音波形图中有波形的部分则代表清音,语音波形图和电声门图中均无波形的部分代表静音.一般地,电声门图中有波形的部分,对应的语谱图中也会对应地显示基频曲线(不包括基频错误点),即浊音,否则为清音或静音,进一步判断清音和静音主要通过观察语音波形图,清音有语音波形,而静音无语音波形.当然标注者还要结合语音知识进行仔细划分.

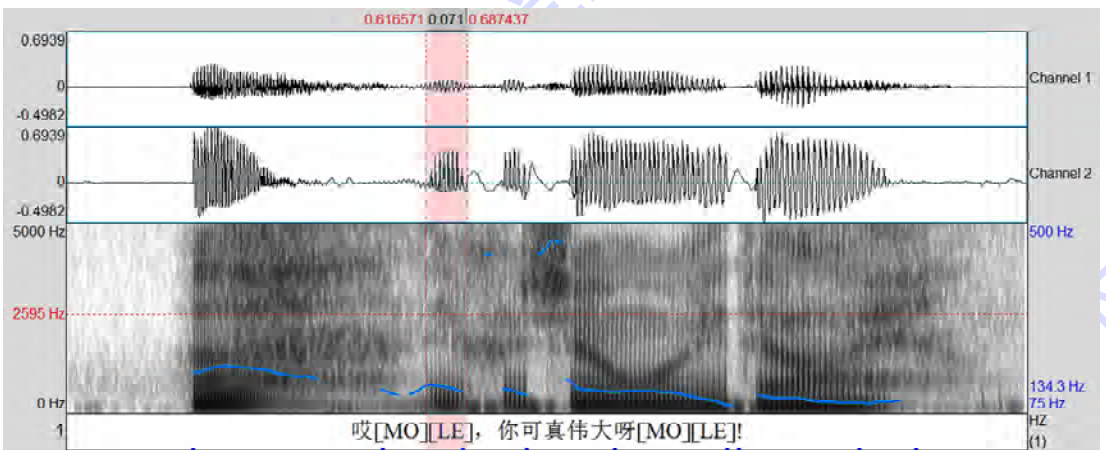


图 1 语音波形图和电声门图  
Fig. 1 Speech waveforms and EGG

### 3.5 副语言信息层 (PARAL)

副语言信息 (Paralinguistic Information) 指语言中非音段特征的信息,如拉长音、喘气声、语气词和咳嗽声等.副语言信息是以情感为导向的情感标签,描述了听话人对语音中情感的理解,在情感语音合成以及情感语音交互研究中至关重要<sup>[23]</sup>,同时副语言信息可以为语音情感识别提供丰富的特征,Divillers 和 Vidrascu<sup>[28]</sup>、Truong 等<sup>[29]</sup>、Arimoto 等<sup>[30]</sup> 都进行了副语言信息的研究.表 3 为一部分基本的副语言标注符号<sup>[31]</sup>.

### 3.6 情感层 (EMO)

情感层是标注该条语句表达的情感类型及情感的强烈程度.语音数据库中包括的 7 种情感类型,分别为:b:悲伤;s:生气;j:惊奇;h:害怕;g:高兴;x:嫌恶;p:平静.标注者根据 7 种情感标注出实际表达的情感,同时,将说话人表达的情感强烈程度划分为 3 级,以打分的形式进行,满分 5 分,分数越高,代表情感表达越充分.具体的对应关系是:情感难以辨别:1、2 分;有情感但情感表达不明

表 3 副语言信息标注符号

Table 3 Annotation symbols of paralinguistic information

No.	副语言学现象	符号	
		开始	结束
1	拉长(拖音) lengthening	LE <	LE >
2	喘气 breathing	BR <	BR >
3	笑声 laughing	LA <	LA >
4	哭声 crying	CR <	CR >
5	咳嗽 coughing	CO <	CO >
6	犹豫(间断) disfluency	DS <	DS >
7	口误 error	ER <	ER >
8	静音(沉默) silence	SI <	SI >
9	含混音 murmur/uncertain segment	UC <	UC >
10	语气词 modal/exclamation	MO <	MO >
11	啞嘴音 smack	SM <	SM >
12	非汉语词汇 non-Chinese	NC <	NC >
13	吸气(吸鼻音) sniffle	SN <	SN >
14	打哈欠 yawn	YA <	YA >
15	叠接(重叠发音) overlap	OY <	OY >
16	吞咽声 deglutition	DE <	DE >
17	消噪(清嗓子) hawk	HA <	HA >
18	打喷嚏 sneezes	SE <	SE >
19	填充停顿 filled pause	FP <	FP >
20	颤音 trill	TR <	TR >

显:3分;情感表达良好:4、5分。为了能更准确地反应语音中的情感,标注时采用对每一种情感都打分的形式进行,情感所得的分数越高,代表该种情感越明显,例如一条语句标注为:b4 g1 h1 j1 p1 s1 x1,说明该条语句表达的是悲伤的情感,并且情感表达比较强烈。

### 3.7 重音指数层(ST)

重音指数是指根据每一个音节发音的轻重程度赋予其一个数值。本文将重音指数划分为4级,分别用1、2、3、4表示,从1到4代表重音越来越强。

### 3.8 语句功能层(FU)

根据语音数据库中的情感类型,目前语句功能层主要标注4种语句类型,分别是:陈述句 declarative、疑问句 interrogative、感叹句 exclamatory、祈使句 imperative。图2为一条完整的语音标注示例,包含4层主要信息,从上往下依次为:语音波形图、电声门图、语谱图、1~8层标注信息。在标注信息中,最左侧竖列的数字代表标注的层数,最右侧代表标注层的名字,中间代表标注内容。例如:第1层为HZ层,标注内容为“哎[MO][LE],你可真伟大呀[MO][LE]!”。

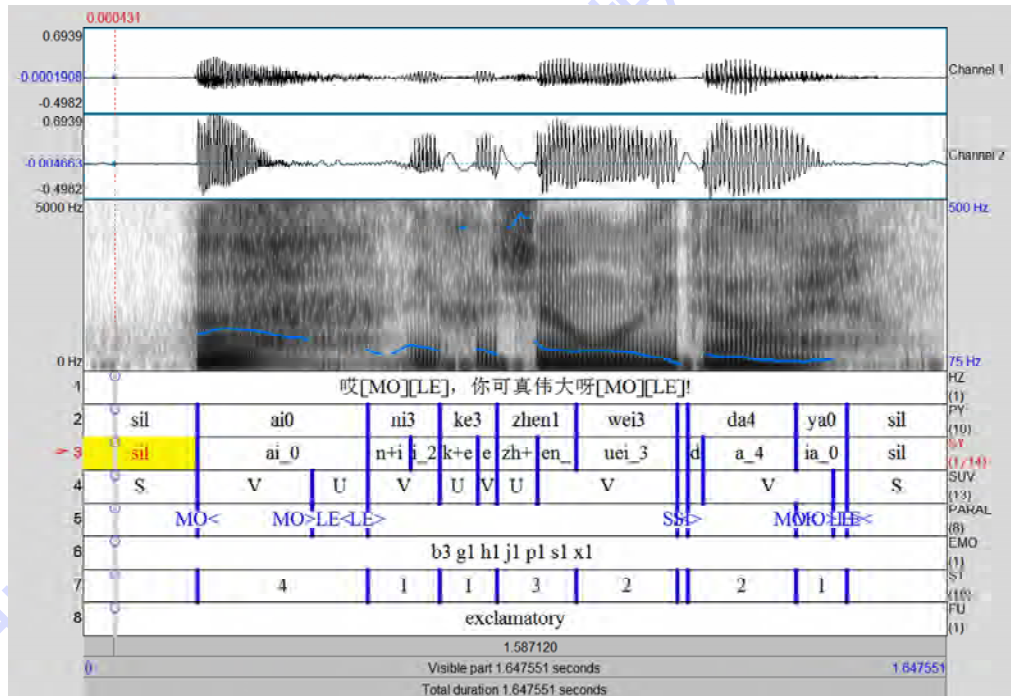


图2 语音标注示例

Fig. 2 Speech annotation example

## 4 一致性检测算法

一致性检测<sup>[31]</sup>是指不同的语音标注者标注相同的语音,然后对标注的结果进行对比,观察和分析标注结果的相同性和差异性。语音手工标注的一致性评价是评价语音质量好坏和标注系统是否完备的重要指标。为了保证情感语音数据库的标注质量,同时检验标注规则的完整性,通过做一致性检测,可以更好地发现标注中存在的问题,从而及时完善标注规则,也可以更好地理解语音的特点和语音数据库的质量。

针对标注规则,设计了相应地一致性检测算法。对于汉字转换层,由于语音文本是预先设定好的,所以HZ层没有必要进行一致性检测。另外,语句功能层受标注者主观情绪影响较大,暂定不

做一致性检测。通过对文献[31]中一致性检测的理解,结合本文制定的标注规则,对其他6层分别设计了对比算法。

1) PY层、SY层。PY层的标注内容是正则的音节和声调,SY层标注的是声母、韵母和实际的声调。所以在这2层的对比中涉及到是否考虑声调一致的问题,因为不同标注者对语句中表达的情感理解有差异,对声调的理解也有差异,所以无论声调相同与否,只要音节相同时,即判断为标注内容相同。设两个标注者为A和B,则具体对比过程是:将待对比的当前层中标注信息理解为由多段组成,每一段的信息包括该音节的起始时间、截止时间和标注内容。则两人标注的一致性可以通过式(1)进行计算:

$$p = \frac{N_{\text{count}} - 2}{2 \times [\min(A_{\text{size}}, B_{\text{size}}) - 1]} \times 100\% \quad (1)$$

式中： $p$  代表一致性结果； $N_{\text{count}}$  代表标注者 A、B 标注结果一致的分割点数； $A_{\text{size}}$  和  $B_{\text{size}}$  分别代表标注者 A、B 在该层标注的音节段数。

图 3 为一致性检测的算法流程图。具体的对比过程为

$$\text{令} \begin{cases} D_{\text{StartTime}} = |A_{\text{StartTime}}(i,j) - B_{\text{StartTime}}(i,j)| \\ D_{\text{EndTime}} = |A_{\text{EndTime}}(i,j) - B_{\text{EndTime}}(i,j)| \end{cases}$$

则对于同一条语音， $A_{\text{StartTime}}(i,j)$  代表标注者 A 标注的语音中第  $i$  层第  $j$  段的起始时间； $A_{\text{EndTime}}(i,j)$  代表标注者 A 标注的语音中第  $i$  层第  $j$  段的结束时间； $B_{\text{StartTime}}(i,j)$  代表标注者 B 标注的语音中第  $i$  层第  $j$  段的起始时间； $B_{\text{EndTime}}(i,j)$  代表标注者 B 标注的语音中第  $i$  层第  $j$  段的结束时间。将标注者 A 和标注者 B 在该层的每一段进行遍历对比，包括音节和声调的对比，然后进行下面的判断：

- ① 若  $\begin{cases} D_{\text{StartTime}} \leq T_0 \\ D_{\text{EndTime}} \leq T_0 \end{cases}$ ，则 A 和 B 有 2 个相同的分割点，即相同的音节有两个一致点。

- ② 若  $\begin{cases} D_{\text{StartTime}} \leq T_0 \\ D_{\text{EndTime}} > T_0 \end{cases}$ ，则 A 和 B 有一个相同的分割点。

- ③ 若  $\begin{cases} D_{\text{StartTime}} > T_0 \\ D_{\text{EndTime}} \leq T_0 \end{cases}$ ，则 A 和 B 有一个相同的分割点。

- ④ 若  $\begin{cases} T_0 < D_{\text{StartTime}} \leq T_1 \\ T_0 < D_{\text{EndTime}} \leq T_1 \end{cases}$ ，则 A 和 B 无相同的分割点。

其中  $T_0$  代表两人标注的误差大小，标注过程中，规定 2 名标注者的时间误差限制在 5 ms 内； $T_1$  用来区分两个不同音节，此处限制在 20 ms 内，避免一句话中有同音字时，一个同音字与另一个同音字判断为一致。

2) SUV 层。SUV 层的内容不包括声调，所以无需判断音调是否一致，其余判断流程和计算公式与 PY 层一样。

3) PARAL 层。标注副语言信息时，符号是标

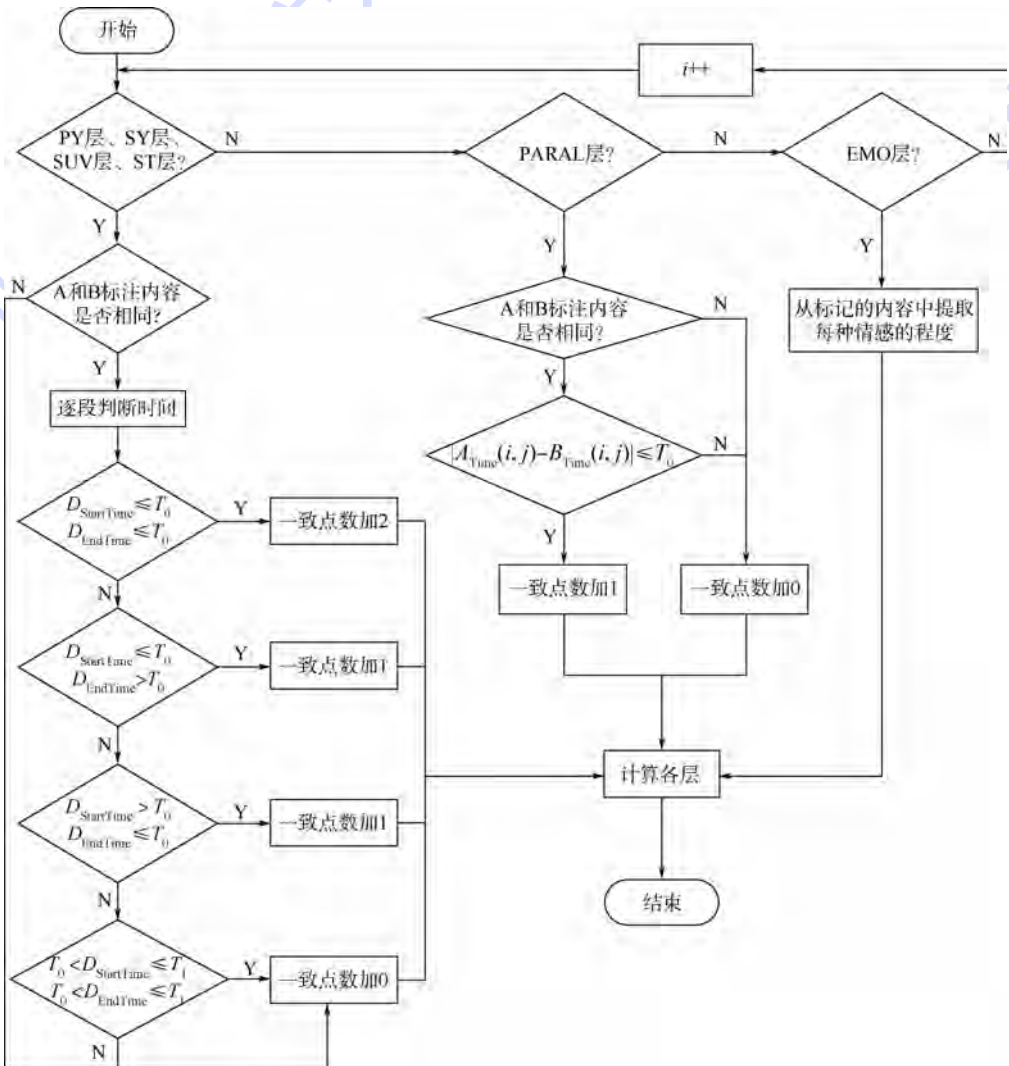


图 3 一致性检测算法流程图

Fig. 3 Algorithm flowchart of consistency detection

在某个时间点上,而非一个时间间隔内,所以该层判断时只需要判断某一个时间点以及该点上的标注内容即可.具体的判断方法如下:

若  $|A_{\text{Time}}(i,j) - B_{\text{Time}}(i,j)| \leq T_0$ ,且标注内容相同时,则 A 和 B 标注一致,否则认为 A 和 B 标注不一致,其一致性计算公式为

$$p = \frac{N_{\text{count}}}{\min(A_{\text{size}}, B_{\text{size}})} \times 100\% \quad (2)$$

4) EMO 层.由于每一句话的情感层标注内容包括 7 种情感类型和其对应的程度,所以应用欧氏距离公式比较其一致性,为了表示高的分数代表一致性高,低的分数代表一致性低,将欧氏距离公式表示为

$$p = \frac{\sqrt{d_{\max} - \sum_{k=1}^7 (A_k - B_k)^2}}{\sqrt{d_{\max}}} \times 100\% \quad (3)$$

式中: $d_{\max} = (a - b)^2 \times c$ ,代表两人标注的结果完全相反时的分数, $a$  代表最高分, $b$  代表最低分, $c$  代表情感种类总数,在本次标注中, $a = 5, b = 1, c = 7$ ;  $A_k$  和  $B_k$  分别代表相应的情感所得分数,即情感程度,其中  $k = 1, 2, \dots, 7$ ,代表 7 种情感类型.

5) ST 层. ST 层标注的是重音的程度,当说话人根据录音文本准确表达并且发音清晰时,则该层的重音段数应该是固定的,但是实际上说话人表达感情时会自发地增加、减少或者改变一些语气词,导致该层的重音段数不一致.因此该层采用与 PY 层相同的对比和计算方法,不同的是标注内容是否相同的判断,认为 1 和 2 的重音轻重程度相同,3 和 4 的重音轻重程度相同.

## 5 一致性检测结果

为了保证情感语音数据库标注的质量,在正式标注之前进行了标注测试,标注结果的优劣采用第 3 节中的一致性算法进行检测,测试语音由一名男生的语音和一名女生的语音组成,每人的语音均包括 7 种情感,每种情感 20 句话,一共是 280 条语句.表 4~表 9 所示为一致性检测结果, A、B、C、D 和 E 代表 5 名不同的标注者. A、B、C 3 名标注者分别对同一男生语音标注,标注误差  $T_0$  取 5、8、10 ms 时,分别对标注结果进行一致性检测,对比结果如表 4、表 5 和表 6 所示. C、D、E 3 名标注者分别对同一女生语音标注,标注误差  $T_0$  取 5、8、10 ms 时,分别对标注结果进行一致性检测,对比结果如表 7、表 8 和表 9 所示.

表 4  $T_0 = 5 \text{ ms}$  的男生语音一致性对比结果

Table 4 Comparison results of consistency for male speech when  $T_0 = 5 \text{ ms}$

标注者	p/%				
	PY	SY	SUV	PARAL	ST
A-B	61.72	63.01	73.01	66.32	63.55
A-C	60.49	61.42	72.84	69.86	63.11
B-C	61.41	62.86	74.06	69.92	64.52
平均	61.21	62.43	73.30	68.70	63.73

表 5  $T_0 = 8 \text{ ms}$  的男生语音一致性对比结果

Table 5 Comparison results of consistency for male speech when  $T_0 = 8 \text{ ms}$

标注者	p/%				
	PY	SY	SUV	PARAL	ST
A-B	67.41	69.22	77.54	69.40	70.23
A-C	66.00	65.40	75.71	73.06	69.21
B-C	68.85	69.54	78.97	74.17	72.36
平均	67.42	68.05	77.41	72.21	70.60

表 6  $T_0 = 10 \text{ ms}$  的男生语音一致性对比结果

Table 6 Comparison results of consistency for male speech when  $T_0 = 10 \text{ ms}$

标注者	p/%				
	PY	SY	SUV	PARAL	ST
A-B	70.46	72.10	78.88	71.37	73.83
A-C	69.70	69.01	78.04	74.65	73.01
B-C	72.29	72.96	81.17	75.93	76.32
平均	70.82	71.36	79.36	73.98	74.39

表 7  $T_0 = 5 \text{ ms}$  的女生语音一致性对比结果

Table 7 Comparison results of consistency for female speech when  $T_0 = 5 \text{ ms}$

标注者	p/%				
	PY	SY	SUV	PARAL	ST
C-D	57.01	59.92	72.21	65.98	61.49
C-E	50.69	55.41	65.41	63.08	52.47
D-E	50.40	55.12	66.46	63.62	55.96
平均	52.70	56.82	68.03	64.23	56.64

表 8  $T_0 = 8 \text{ ms}$  的女生语音一致性对比结果

Table 8 Comparison results of consistency for female speech when  $T_0 = 8 \text{ ms}$

标注者	p/%				
	PY	SY	SUV	PARAL	ST
C-D	63.27	65.31	76.45	69.23	68.34
C-E	58.70	61.85	72.87	69.17	64.69
D-E	58.70	61.85	72.87	69.17	64.69
平均	60.22	63.01	74.06	69.19	65.91

表 9  $T_0 = 10 \text{ ms}$  的女生语音一致性对比结果

Table 9 Comparison results of consistency for female speech when  $T_0 = 10 \text{ ms}$

标注者	p/%				
	PY	SY	SUV	PARAL	ST
C-D	66.50	68.16	78.52	70.27	71.70
C-E	61.17	64.11	72.07	66.50	63.74
D-E	62.05	64.36	75.36	71.37	68.14
平均	63.24	65.54	75.32	69.38	67.86

观察对比结果,可以得出以下结论:

1) 对于男生语音,  $T_0 = 5\text{ ms}$  时,各层的一致性综合平均为 65.87%,  $T_0 = 8\text{ ms}$  时,一致性平均为 71.19%,  $T_0 = 10\text{ ms}$  时,一致性平均为 73.98%. 对于女生语音,  $T_0 = 5\text{ ms}$  时,一致性平均为 59.68%,  $T_0 = 8\text{ ms}$  时,一致性平均为 66.48%,  $T_0 = 10\text{ ms}$  时,一致性平均为 68.27%.

2) 对比表 4、表 5 和表 6,可以看出,当  $T_0$  从 5 ms 增大至 8 ms 时,各层的一致性都有提高,最低可提高 4%,最高可提高 7%,平均提高了 5.5%;当  $T_0$  从 5 ms 增大至 10 ms 时,对比各层的一致性,最低可提高 5%,最高可提高 11%,平均提高了 8%. 从表 7、表 8 和表 9 中可以看出,当  $T_0$  从 5ms 增大至 8 ms 时,各层的一致性最低可提高 4%,最高可提高 9%,平均提高了 6.5%;当  $T_0$  从 5 ms 增大至 10 ms 时,各层的一致性最低可提高 5%,最高可提高 11%,平均提高了 8%. 说明 2 名标注者之间的标注误差增大时,标注的一致性提高,限制的时间误差大小直接影响着一致性对比结果.

3) 观察表 4 ~ 表 9,可以发现 SUV 层的一致性最高, PY 层和 SY 层的一致性较低. 因为 SUV 层存在的差异性最小,容易区分, PY 层和 SY 层的标注需要的语音知识较多,造成音节和声韵母的划分难度比较大. ST 层重音程度的判断受标注者主观因素的影响比较大,例如标注时间、标注者心情、个人情感理解等. C-E、D-E 的一致性结果较低,因为标注者 E 的语音基础知识较薄弱,直接影响其对语音的理解. 另外,当长时间进行标注工作时,难免出现烦躁和难以判断的情形,这些都会影响标注的一致性.

4) 表 4 ~ 表 9 整体的一致性比文献 [31] 中的低,分析造成整体一致性较低的原因有 3 点: ①一致性检测的具体算法不同,虽然基本的检测标准都是标注符号和时间点误差,但是标注层数及每层具体的一致性算法不同,本文的一致性检测算法还考虑了分割段数,这可能造成一致性偏低. 因为本文的一致性检测算法原理是一致点的个数占总个数的比重,所以当语句较短、音节较少时,每个一致点占的比重就会较大,每个点的错误都会导致一致性百分比大大降低. ②一致性检测算法中规定的时间误差较短,这对标注者的专业要求非常高,实验结果中可以看出,将对误差从 5 ms 扩大到 10 ms 时,一致性会大大提高. ③一致性对比算法设计中,对标注准确性的要求严格程度不同.

设计的一致性检测算法除了可以反映出 2 个人标注的一致性结果,同时还可以将 2 个人标注不一致的时间点以及内容准确地显示出来,返回的内容以 txt 文件存储,方便标注人员检查标注错误的原因. txt 文件中显示的信息包括标注人员、文件序号、文件名、标注层、一致性检测结果、不一致的时间点及对应的不一致内容. 表 10 为标注者 A 和 B 标注结果对比的详细信息示例,文件名为 bM0C03011.TextGrid,引号内无内容时代表标注内容为空,从表 10 中可以看出,在 SY 层的 0.2189 s 时刻,一个人标记为“ai\_0”,另一个人标记为“ai\_4”,后面的\_0 和\_4 代表实际发音中的声调,属于主观判断,标注者的判断受外界环境等因素影响比较严重,所以在检测的时候认定音节相同而仅仅声调不同的情况为一致,不影响检测的分数,而 1.6422 s 时刻和 1.64755 s 时刻相差 5.35 ms,仅比 5 ms 大 0.35 ms,检测结果为不一致,由此也可以发现一致性检测中对时间的要求十分苛刻.

表 10 标注中不一致点详情示例

Table 10 Example of inconsistency details in annotation

标注层	不一致时间点/s	不一致内容	一致性结果/%
PY	xminA:1.6422	A:“ ”	94.44
	xmaxA:1.4563	B:“sil”	
	...	...	
SY	xminA:0.2189	A:“ai_0”	96.15
	...	B:“ai_4”	
	...	...	
SUV	xminA:1.6422	A:“ ”	95.83
	xminB:1.6475	B:“S”	
	...	...	
PARAL	numberA:0.2189	markA:“MO < ”	71.43
	...	markB:“MO < SI < ”	
	...	...	
ST	xminA:0.68044	A:“1”	83.303
	...	B:“4”	
	...	...	

## 6 结 论

本文总结了国内外情感语音数据库的发展现状,据此做了以下工作:

1) 建立了一个汉语双模情感语音数据库,与其他情感语音数据库主要不同在于包含 EGG 信息.

2) 制定了一套详细的情感语音数据库标注规则,并对汉语双模情感语音数据库进行标注.

3) 根据规则设计了相应的一致性检测算法,综合对男生和女生语音检测结果表明:

①  $T_0 = 5 \text{ ms}$  时,各层的一致性平均为 62.77%,  $T_0 = 8 \text{ ms}$  时,一致性平均为 68.84%,提高了 5%;  $T_0 = 10 \text{ ms}$  时,一致性平均为 71.13%;提高了 8%。

② 算法可以反映出标注错误的位置和内容,便于标注者修改错误之处。

③ 算法可以有效地对标注结果进行对比,5名标注者对语音的理解较一致,证明了语音数据库的质量较好,为后续标注工作提供了保障,也为其他语音数据库的标注提供了有效的标注和检验方法。

本文还有需要进一步改善的内容:

1) 在保证标注时间误差相同的条件下,通过改进一致性检测算法,使新的一致性检测算法能更合理、更好地反映出真实的一致性;通过加强标注者的专业培训,保证不同标注者对相同语音有更一致的理解,使标注符号更一致,提高一致性。

2) 由于标注过程中对 EMO 层的标注规则进行了修改,部分标注内容尚未修改完,所以尚未标出 EMO 层的一致性,作为下一步的工作重点。

全部语音标注完成后,可以建立一个完整的情感语音数据库,为语音情感识别的模型建立提供了可靠的汉语双模情感语音数据库,对国内外情感语音数据库进行了有益补充。

### 参考文献 (References)

[1] 韩文静,李海峰.情感语音数据库综述[J].智能计算机与应用,2013,3(1):5-7.  
Han W J, Li H F. A brief review on emotional speech databases [J]. Intelligent Computer and Applications, 2013, 3(1): 5-7 (in Chinese).

[2] 徐露,徐明星,杨大利.面向情感变化检测的汉语情感语音数据库[J].清华大学学报:自然科学版,2009,49(S1):1413-1418.  
Xu L, Xu M X, Yang D L. Chinese emotional speech database for the detection of emotion variations[J]. Journal of Tsinghua University: Natural Science, 2009, 49(S1): 1413-1418 (in Chinese).

[3] 薛雨丽,毛峡,张帆. BHU 人脸表情数据库的设计与实现[J].北京航空航天大学学报,2007,33(2):224-228.  
Xue Y L, Mao X, Zhang F. Design and realization of BHU facial expression database [J]. Beijing University of Aeronautics and Astronautics, 2007, 33(2): 224-228 (in Chinese).

[4] Ververidis D, Kotropoulos C. A state of the art review on emotional speech databases [C] // Proceedings of 1st Richmedia Conference. Lausanne: The European Association for Signal Processing, 2003: 109-119.

[5] El Ayadi M, Kamel M S, Karray F. Survey on speech emotion recognition: Features, classification schemes, and databases [J]. Pattern Recognition, 2011, 44(3): 572-587.

[6] Greasley P, Setter J, Waterman M, et al. Representation of prosodic and emotional features in a spoken language database [C] // Proceedings of the 13th International Congress of Phonetic Sciences. Paris: IPA, 1995: 242-245.

[7] Grimm M, Kroschel K, Narayanan S. The Vera am Mittag German audio-visual emotional speech database [C] // IEEE International Conference on Multimedia and Expo. Piscataway, NJ: IEEE Press, 2008: 865-868.

[8] Campbell N. The JST/CREST ESP project-a mid-term progress report [C] // 1st JST/CREST Intl. Workshop Expressive Speech Processing. Baixas: ISCA, 2003: 61-70.

[9] Chong T Y, Xiao X, Tan T P, et al. Collection and annotation of malay conversational speech corpus [C] // International Conference on Speech Database and Assessments (Oriental COCOSDA 2012). Piscataway, NJ: IEEE Press, 2012: 30-35.

[10] Mori H, Satake T, Nakamura M, et al. Constructing a spoken dialogue corpus for studying paralinguistic information in expressive conversation and analyzing its statistical/acoustic characteristics [J]. Speech Communication, 2011, 53(1): 36-50.

[11] Mori H, Hitomi T. Annotating conversational speech for corpus-based dialogue speech synthesizer-a first step [C] // International Conference on Speech Database and Assessments (Oriental COCOSDA 2012). Piscataway, NJ: IEEE Press, 2012: 135-140.

[12] CASIA. Database of Chinese emotional speech [EB/OL]. Beijing: Chinese Linguistic Data Consortium, 2008 (2010-10-09) [2014-12-8]. [http://www.chineseldc.org/resource\\_info.php?rid=76](http://www.chineseldc.org/resource_info.php?rid=76).

[13] Pan Y C, Xu M X, Liu L Q, et al. Emotion-detecting based model selection for emotional speech recognition [C] // IMACS Multiconference on Computational Engineering in Systems Applications. Piscataway, NJ: IEEE Press, 2006, 2: 2169-2172.

[14] Nwe T L, Foo S W, de Silva L C. Speech emotion recognition using hidden markov models [J]. Speech Communication, 2003, 41(4): 603-623.

[15] Morrison D, Wang R, de Silva L C. Ensemble methods for spoken emotion recognition in call-centres [J]. Speech Communication, 2007, 49(2): 98-112.

[16] Fu L, Mao X, Chen L. Speaker independent emotion recognition based on SVM/HMMS fusion system [C] // International Conference on Audio, Language and Image Processing, 2008. Piscataway, NJ: IEEE Press: 61-65.

[17] Zhou J, Wang G, Yang Y, et al. Speech emotion recognition based on rough set and SVM [C] // IEEE International Conference on Cognitive Informatics. Piscataway, NJ: IEEE Press, 2006, 1: 53-61.

[18] Hu H, Xu M X, Wu W. GMM Supervector based SVM with spectral features for speech emotion recognition [C] // 2007 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2007: 413-416.

[19] Burkhardt F, Paeschke A, Rolfes M, et al. A database of German emotional speech [C] // Interspeech 2005. Baixas: ISCA, 2005, 5: 1517-1520.

[20] Schuller B. Towards intuitive speech interaction by the integration of emotional aspects [C] // IEEE International Conference on Systems, Man and Cybernetics. Piscataway, NJ: IEEE Press,

2002,6:6-12.

- [21] Engberg I S, Hansen A V. Documentation of the danish emotional speech database [R]. Denmark: Aalborg University, 1996.
- [22] Hansen J H L, Bou-Ghazale S E, Sarikaya R, et al. Getting started with SUSAS: A speech under simulated and actual stress database [C] // EUROSpeech 1997. Baixas: ISCA, 1997, 97(4):1743-1746.
- [23] Breazeal C, Aryananda L. Recognition of affective communicative intent in robot-directed speech [J]. Autonomous Robots, 2002, 12(1):83-104.
- [24] Slaney M, Mcroberts G. BabyEars: A recognition system for affective vocalizations [J]. Speech Communication, 2003, 39(3):367-384.
- [25] Wang M, Li Y, Lin M, et al. The development of a database of functional and emotional intonation in Chinese [C] // International Conference on Speech Database and Assessments (Oriental COCOSDA 2011). Piscataway, NJ: IEEE Press, 2011:136-141.
- [26] Li A J. Chinese prosody and prosodic labeling of spontaneous speech [C] // International Conference on Speech Prosody 2002. Baixas: ISCA, 2002.
- [27] 刘亚斌. 汉语自然口语的韵律分析和自动标注研究 [D]. 北京: 中国社会科学院研究生院, 2003.
- Liu Y B. Prosodic analysis and automatic prosodic-labeling for Chinese spontaneous speech [D]. Beijing: Graduate School of Chinese Academy of Social Sciences, 2003 (in Chinese).
- [28] Devillers L, Vidrascu L. Real-life emotions detection with lexical and paralinguistic cues on human-human call center dialogs [C] // Interspeech 2006. Baixas: ISCA 2006:801-804.
- [29] Truong K P, Neerinx M A, van Leeuwen D A. Assessing agreement of observer-and self-annotations in spontaneous multimodal emotion data [C] // Interspeech 2008. Baixas: ISCA, 2008:318-321.
- [30] Arimoto Y, Kawatsu H, Ohno S, et al. Emotion recognition in spontaneous emotional speech for anonymity-protected voice chat systems [C] // Ninth Annual Conference of the International Speech Communication Association. Baixas: ISCA, 2008:322-325.
- [31] 李爱军, 陈肖霞, 孙国华, 等. CASS: 一个具有语音学标注的汉语口语语音库 [J]. 当代语言学, 2002, 4(2):81-89.
- Li A J, Chen X X, Sun G H, et al. CASS: A Chinese annotated spontaneous speech corpus [J]. Contemporary Linguistics, 2002, 4(2):81-89 (in Chinese).

## Annotations and consistency detection for Chinese dual-mode emotional speech database

JING Shaoling, MAO Xia\*, CHEN Lijiang, ZHANG Nana

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** To solve problem of lacking emotional speech database with rich emotion annotation information, a Chinese dual-mode emotional speech database which contained speech and Electroglottography (EGG) information was established. Annotation and consistency detection for the established database were conducted. Firstly, we designed detailed annotation rules and methods according to characteristics of emotional speech database and selected 5 annotators labeling emotional speech database in accordance with the rules. Secondly, in order to ensure annotation quality of emotional speech database and test the integrity of annotation rules, annotators labeled parts of utterances as a test before the official annotation, the test material comprises 280 sentences (seven emotions  $\times$  two actors  $\times$  twenty sentences). Finally, according to the speech annotation rules, we designed corresponding consistency detection algorithm. The results show that within the time error range of 5 ms, the annotation consistency for the same utterances which labeled by 5 annotators reaches more than 60% on average. When the time error range increased to 8 ms and 10 ms, consistency can be increased by 5% and 8% on average. The experiment indicates that 5 annotators are more consistent in understanding speech. The annotation rules we designed are more complete. The quality of emotional speech database is higher.

**Key words:** Chinese; dual-mode; emotional speech database; speech annotation; consistency detection

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0826

# 混合点状和非规则军标的在线手绘识别

邓维<sup>1</sup>, 吴玲达<sup>\*1</sup>, 张友根<sup>2</sup>, 赵志鹏<sup>3</sup>

(1. 装备学院 复杂电子系统仿真实验室, 北京 101416; 2. 国防信息学院 信息系统系, 武汉 430010;  
3. 桂林电子科技大学信息科技学院, 桂林 541004)



**摘 要:** 当前对在线手绘军标图符识别的研究只针对单一类型的手绘点状军标或非规则军标, 分别使用不同方法进行识别. 但在特殊应用中二者常混合输入, 当待识别军标图符的类型未知时, 如何识别是一个重要问题. 提出一种基于最小生成树(MST)覆盖模型的混合识别方法, 训练阶段, 分别对点状和非规则军标样本建立 MST 覆盖模型, 并训练一个二分类支持向量机(SVM)分类器; 识别阶段, 先通过几何和结构信息粗判断军标类型, 再通过置信度估计和融合的方法确定未知军标的类型. 在 113 类点状军标和 36 类非规则军标的数据集集中实验, 军标类型区分准确率为 94.7%, 最终识别率为 91.6%, 且能满足实时要求.

**关键词:** 草图识别; 点状军标; 非规则军标; 最小生成树(MST); 分类

**中图分类号:** TP391

**文献标识码:** A **文章编号:** 1001-5965(2015)10-1935-08

在线草图符号识别是笔式交互系统的关键技术之一, 旨在将用户手绘输入的草图形状自动识别成预定义的标准形状. 高效的草图识别算法在人机交互、图像检索、工业设计以及儿童教育等领域有广泛的应用<sup>[1-2]</sup>. 在军事领域, 态势图是辅助指挥员理解战场态势, 拟定作战计划和制定决策的重要工具, 军队标号(也称军标)是态势图的关键元素, 是战场态势可视化的工具和图形语言<sup>[3]</sup>. 当前, 态势标绘系统仍广泛采用 WIMP (Windows, Icon, Menu, Point device) 风格的人机交互方法. 与之相比, 基于手绘草图的态势标绘方法具有简明直观、易学易用、更好地利用屏幕空间等优点, 逐渐受到重视和研究<sup>[3-7]</sup>. 其核心问题是识别用户手绘的军标图符.

军标图符按标示功能和几何特点不同, 可分为规则军标和非规则军标 2 类. 前者指不依比例尺标绘的点状军标, 其外观形状、线画构成和各比

例关系保持不变; 后者包括半依比例尺的线状军标和依比例尺的面状军标, 用于标示行军路线、障碍分布、覆盖范围等信息, 图形线条的走向根据具体地形和作战意图等因素确定, 线条长度、方向等几何特征的变化性大, 没有规则、固定的形状.

此外, 点状军标和非规则军标的识别目标不同. 在态势标绘系统的实现过程中, 点状军标能预先制成标准化模板文件, 如二维矢量模型或图像; 而非规则军标需要根据若干控制点, 常通过曲线拟合的方式动态生成. 因此, 手绘点状军标的识别目标仅仅是模式分类, 而对非规则军标不仅要获得形状类别, 还需获得形状参数, 如各线条图元的类别和控制点.

由于形状特点的差异和识别目标的不同, 手绘点状和非规则军标难以采用统一的识别方法. 当前对手绘军标识别的研究都只针对单一类型的军标图符, 没有考虑两类图符的类型区分和混合

收稿日期: 2014-12-30; 录用日期: 2015-04-10; 网络出版时间: 2015-06-17 17:10

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150617.1710.008.html

基金项目: “核高基”国家科技重大专项(2013ZX01045-004)

作者简介: 邓维(1986—), 男, 湖北武汉人, 博士研究生, dengwei@whu.edu.cn

\* 通讯作者: 吴玲达(1962—), 女, 上海人, 教授, wld@nudt.edu.cn, 主要研究方向为空间信息获取与处理.

**引用格式:** 邓维, 吴玲达, 张友根, 等. 混合点状和非规则军标的在线手绘识别[J]. 北京航空航天大学学报, 2015, 41(10): 1935-1942. Deng W, Wu L D, Zhang Y G, et al. Online sketch recognition for mixed point and irregular military symbols[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1935-1942 (in Chinese).

识别问题. 在之前的工作中,曾提出一种基于图元匹配的手绘非规则军标识别方法<sup>[4]</sup>和一种基于方向特征的手绘点状图符识别方法<sup>[8]</sup>(文献[8]是点状图符的通用识别方法,不只针对军标). 然而在实际应用中,点状和非规则军标通常是混合输入,对一个类型未知的手绘军标进行识别是一个必须解决的问题. 否则,基于草图的态势标绘系统将难以投入实际应用.

由于军标图符是一个相对小众的研究对象,对混合点状和非规则军标的手绘识别问题的研究还不多. 与此类似的研究包括手绘图文笔画分类<sup>[9]</sup>、印刷体和手写体分类<sup>[10]</sup>、混合多语种字符识别<sup>[11-12]</sup>等. 针对混合输入点状和非规则军标时的手绘图符识别问题,本文提出一种基于最小生成树(MST)覆盖模型和置信度转化的识别方法.

# 1 混合识别方案设计

## 1.1 算法概述

首先形式化定义以下术语:类型(type),指军标图符是点状军标或非规则军标;类别(class),指具体类属性,如指挥所、进攻箭头等. 获取图符类别是军标识别的最终任务. 手绘军标图符混合识别的难点在于非规则军标的线条可变性大,难以提取稳定的几何、图像特征,同点状军标的几何性质差异较大. 图1(a)所示的进攻箭头和双箭头图符都属于非规则军标,箭头可随着进攻方向移动,箭身部分将被自动拉伸或弯曲,而图1(b)所示的点状军标的各图元关系(如长度比值)固定不变.

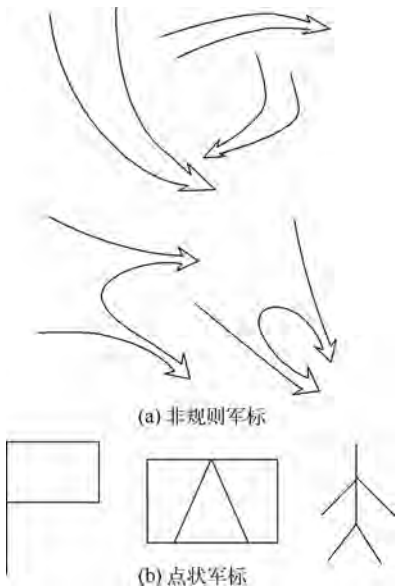


图1 2种不同类型的军标图符

Fig.1 Two different types of military symbols

本文算法的框架如图2所示. 对类型未知的待识别军标图符,整合几何、结构和识别信息,对图符类型进行判定,然后对不同类型的图符输入各自识别模块,最后得到识别结果. 本文将文献[8]和文献[4]的方法分别作为点状和非规则军标识别模块. 军标类型的粗判定利用几何和结构信息实现,它适用于特征明显的非规则军标图符的快速类型确认. 点状军标覆盖模型和非规则军标覆盖模型分别从全体点状军标和全体非规则军标的角度评估待识别模式的类型,二者都属于识别信息. 通过置信度估计的方法融合这两类识别信息,精确判别待识别图符类型. 需要说明的是,文献[4]在识别非规则军标时作出假设:每条笔画仅属于一个图元,它可能是图元的一部分,也可能是一个完整图元,即不能用一个笔画绘制多个图元. 本文仍沿用该假设,用户绘制非规则军标时每条笔画仅属于一个图元,但对点状军标的绘制方式没有限制.



图2 提出算法的流程图

Fig.2 Flowchart of proposed algorithm

## 1.2 军标类型粗区分

同一幅态势图中,点状军标的几何大小具有相近性,而非规则军标由于用来表示地域或部队行为,几何大小不固定,有的显著大于点状军标,有的和点状军标大小相近. 且非规则军标的图元组成不同于点状军标,是有用的结构信息. 因此,本文通过几何和结构信息分析,快速判断明显是非规则军标的图符.

### 1) 几何信息分析.

图符几何特征可用多种形式表示,如包围盒面积、包围盒对角线、笔画总长度、凸壳面积等. 由于在手写汉字的分割和识别<sup>[13]</sup>以及草图识别研究中,包围盒几何特征简单有效<sup>[14]</sup>,本文以此作为图符几何特征. 若满足

$$\frac{D}{\max(D_p)} > \lambda_1 \text{ 或 } \frac{R}{\max(R_p)} > \lambda_2 \quad (1)$$

则待识别模式为非规则军标.式中: $D$ 和 $R$ 分别为待识别图符有向包围盒对角线长度和有向包围盒长宽比; $\max(D_p)$ 为点状军标训练样本的有向包围盒对角线长度的最大值; $\max(R_p)$ 为点状军标训练样本的有向包围盒长宽比的最大值; $\lambda_1$ 和 $\lambda_2$ 为较宽松的阈值,取 $\lambda_1 = 2.0, \lambda_2 = 3.0$ .

## 2) 结构信息分析.

非规则军标具有较稳定的结构信息,可用于区别点状军标,如非规则军标的图元包括实心小圆或小齿、长螺旋线、长折线、闭合猪腰形等,而点状军标没有这类图元.它们作为非规则军标的专有图元,具有很强的指示性.如果检测到待识别图符具有上述图元,则也可直接判定为非规则军标.识别这类图元可以用启发式规则<sup>[15-16]</sup>,笔画特征包括长度、折线拟合误差、累计转向角度、笔迹密度、反向计数(相邻3个采样点夹角为锐角的个数,能检测填充区域)、闭合度、外接矩形长宽比等.需要说明的,本节的图元识别只用来实现图符类型粗区分,并不用来识别非规则军标的类别,和文献[4]的图元定义不同.

综上所述,若待识别图符满足式(1)或检测出非规则军标专有图元,则直接判定为非规则军标,输入非规则军标识别模块,得到最终识别结果.

## 2 点状军标覆盖模型及置信度估计

点状军标覆盖模型及置信度估计的目的是将所有类别的点状军标看作一个整体,从全局的角度,度量待识别军标属于点状军标类型的可能性.

本文采用可拒绝分类问题的思想.它是模式识别中的一类特殊问题,主要研究训练同类样本在高维空间中分布形成的复杂几何形体并对其构建合理的覆盖描述模型.其理论基础是同类样本的流形分布具有连续性:同类样本分布在同一个光滑流形上,不存在孤立的样本点<sup>[17]</sup>.胡正平和冯凯<sup>[17]</sup>提出一种基于最小生成树(MST)的可拒绝分类算法,本文以此为借鉴并作改进,建立点状军标覆盖模型.不同于文献[17]的0-1判定,本节通过点状军标覆盖模型提取待识别模式相对于全体点状军标的特征,并训练一个分类器,最后将分类器输出转换为后验概率.考虑到支持向量机(SVM)在解决小样本、非线性以及高维模式识别中具有许多特有的优势<sup>[18]</sup>,本文采用SVM分类器.本节主要有3个步骤:①点状军标图符特征提取;②建立MST覆盖模型;③待识别图符类型置信度估计.

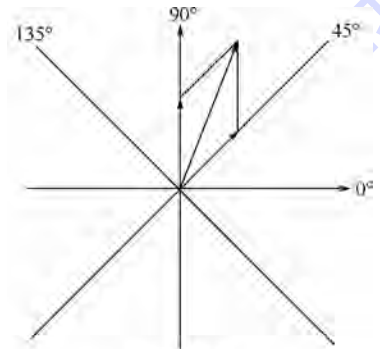
## 2.1 点状军标图符特征提取

手绘点状军标和手写汉字有类似之处,可以借鉴汉字识别的相关成果,方向特征对汉字识别非常有效,因此点状军标图符的特征提取方法步骤如下:

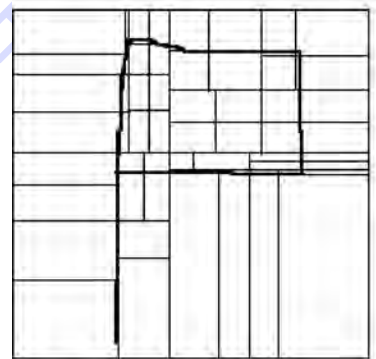
首先,将图符线性缩放到 $100 \times 100$ 范围内并重采样(保持宽高比),提取采样点局部方向特征:对每一个笔画,任取一采样点 $P_i$ ,除笔画的最后一点外,都有一个后继点 $P_{i+1}$ ,将局部向量 $(P_i, P_{i+1})$ 按平行四边形法则分解到 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 共4个方向轴,如图3(a)所示,得到4个响应强度值,作为点 $P_i$ 的4方向特征.

然后,对图符区域进行 $8 \times 8$ 网格划分:先计算所有采样点几何重心,经过重心点画出一对水平和垂直网格线,它们将采样点分成4个子区域,然后在每一个子区域内采用同样的方法迭代计算2次,即得到 $8 \times 8$ 网格,如图3(b)所示,网格可吸收草图的局部形变.

最后,对每个网格单元,累加其内部所有采样点的4方向特征,共得到 $8 \times 8 \times 4 = 256$ 维的特征矢量,作为点状图符的特征描述.



(a) 方向分解



(b) 网格区域划分

图3 点状军标图符特征提取示意图

Fig. 3 Feature extraction schematic of point military symbols

## 2.2 建立MST覆盖模型

提取点状军标训练样本的特征矢量后,借鉴

文献[17]的方法建立点状军标的自适应覆盖模型,简述如下:

首先,对由点状军标训练样本组成的高维特征空间的点集,计算一个 MST. 然后,根据同类样本的连续性规律,采用 MST 刻画高维空间样本点的分布:将 MST 的边(不含端点)作为新增虚拟样本,该 MST 作为覆盖模型的骨架,并对每个样本点(包括真实训练样本和虚拟样本)计算一个局部覆盖半径,从而“认识”该类别. 图 4 为一组二维空间样本点组成的 MST 覆盖模型示意图,图 4(a)为 MST 组成的覆盖模型骨架,其中直线端点表示真实样本,图 4(b)为计算覆盖半径后的覆盖模型. 设点状军标样本的类别一共有  $C$  类,对每一类别建立一个 MST 覆盖模型,它们的并集作为点状军标的覆盖模型.

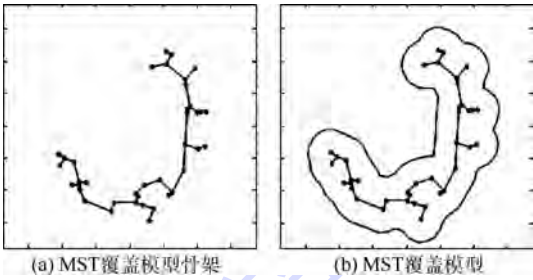


图 4 MST 覆盖模型图<sup>[17]</sup>

Fig. 4 Minimum spanning tree (MST) coverage model diagram<sup>[17]</sup>

MST 覆盖模型的关键是计算样本点(记为  $x_p$ )的覆盖半径,它与以下 2 个因素有关: $x_p$  周围的真实样本点分布密度,密度越大,则  $x_p$  的覆盖半径越大; $x_p$  所在边的长度越长,则  $x_p$  覆盖半径越小. 计算细节请参考文献[17].

### 2.3 基于点状军标 MST 覆盖模型的置信度估计

置信度是衡量模式分类结果可靠性的指标,常用于多方案融合<sup>[19]</sup>,广泛应用于手写文本的分割和识别<sup>[13]</sup>. 本节的目的是根据点状军标的 MST 覆盖模型,估计待识别图符的类型置信度.

文献[17]建立的基于 MST 的可拒绝分类算法没有考虑拒识决策的置信度,即测试样本拒识与否是一个 0-1 判断,具体表现是:建立  $C$  个类别 MST 覆盖模型后,计算测试样本到每一类 MST 的距离,如果没有 MST 覆盖区域包含该测试样本,则说明其不属于  $C$  类中任何一类,拒识;否则,样本属于  $C$  类中的某一类,再进一步对测试样本进行分类. 由于草图信息的模糊性,这种二值描述难以达到鲁棒性和精确性的统一. 本文结合点状军标和非规则军标的混合识别问题,将军标图符类型判别问题以置信度的形式输出,以便融合多种

信息. 为此,提出的改进方法如下:

将点状军标训练样本分为 2 部分,记为集合  $A_1, A_2$ ,记非规则军标训练样本集合为  $B$ . 用  $A_1$  建立点状军标的 MST 覆盖模型,记为集合  $\{MST_1, MST_2, \dots, MST_C\}$  ( $MST_i$  为第  $i$  类点状军标的 MST 覆盖模型).

对  $A_2 \cup B$  中的每个样本  $x$ ,记

$$F_i = \min_{P_{ij} \in MST_i} (d(x, P_{ij}) - Rad(P_{ij})) \quad i = 1, 2, \dots, C \quad (2)$$

式中: $P_{ij}$  为第  $i$  类 MST 上的第  $j$  个样本点(包括真实样本和虚拟样本); $d(\cdot, \cdot)$  表示两点间的距离; $Rad(\cdot)$  表示局部覆盖半径.

记  $(i^*, j^*) = \arg \min_{i,j} (F_i)$ ,取

$$f = [d(x, P_{i^*j^*}) \quad Rad(P_{i^*j^*})] \quad (3)$$

作为  $x$  的特征描述,这是一个二维矢量,然后建立二分类问题:

$$y = \begin{cases} 1 & x \text{ 为点状军标} \\ 0 & x \text{ 为非规则军标} \end{cases} \quad (4)$$

据此,文献[17]的决策方法实际是

$$y = \begin{cases} 1 & d(x, P_{i^*j^*}) - Rad(P_{i^*j^*}) < 0 \\ 0 & \text{其他} \end{cases} \quad (5)$$

本文不采用该决策方法,所提出的方法是:用  $A_2 \cup B$  训练一个 SVM 分类器,其中样本特征描述如式(3)所示,训练样本的标记为

$$y = \begin{cases} 1 & x \in A_2 \\ 0 & x \in B \end{cases} \quad (6)$$

标准 SVM 分类器的决策函数  $f(x)$  为样本到决策边界的距离,可利用 Logistic 回归模型进一步将  $f(x)$  转换为后验概率:

$$\begin{cases} P(y=1 | x) = \frac{\exp(af(x) + b)}{1 + \exp(af(x) + b)} \\ P(y=0 | x) = \frac{1}{1 + \exp(af(x) + b)} \end{cases} \quad (7)$$

式中: $a, b \in \mathbf{R}$  为模型参数,可用极大似然估计方法计算,记

$$\varphi(x_i) = P(y=1 | x_i) \quad (8)$$

则似然函数为

$$L = \prod_{i=1}^N (\varphi(x_i))^{y_i} (1 - \varphi(x_i))^{1-y_i} \quad (9)$$

式中: $N$  为集合  $A_2 \cup B$  中样本的总数. 对数似然函数为

$$\ln L = \sum_{i=1}^N [y_i (af(x_i) + b) - \ln(1 + \exp(af(x_i) + b))] \quad (10)$$

采用梯度下降法对  $\ln L$  求极大值,即可估计出参数  $a, b$ .

### 3 非规则军标覆盖模型及置信度估计

同第 2 节类似,非规则军标覆盖模型的目的是将所有类别的非规则军标看作一个整体,从全局的角度,度量待识别图符属于非规则军标的可能性。

要建立非规则军标的 MST 覆盖模型,首先要度量任意 2 个非规则军标的距离,采用基于图元匹配的度量方法<sup>[4]</sup>,简述如下:

记  $U$  和  $T$  为 2 个手绘图符,图符表示为

$$\begin{cases} U = \{u_1, u_2, \dots, u_{|u|}\} \\ T = \{t_1, t_2, \dots, t_{|t|}\} \end{cases} \quad (11)$$

式中: $u$  和  $t$  分别为  $U$  和  $T$  的图元; $|u|$  为  $U$  中图元个数; $|t|$  为  $T$  中图元个数.建立  $U$  和  $T$  的图元匹配代价矩阵为

$$C = [c_{i,j}] = [c(u_i, t_j)] \quad (12)$$

式中: $c(u_i, t_j)$  表示图元  $(u_i, t_j)$  的匹配距离或代价.图元匹配的任务可以转化为一个指派问题:对  $U$  的每一个图元,在  $T$  中指派一个图元与之一一对应,使总匹配代价最小,可用匈牙利算法(Hungary algorithm)求解,最坏情况下复杂度是  $O(N^3)$ ,  $U$  和  $T$  的距离即为图元匹配的总代价. $c(u_i, t_j)$  与图元类型、图元间空间拓扑关系和图元长度有关,细节请参考文献[4].

非规则军标的 MST 覆盖模型及其置信度的计算方法与 2.2 节和 2.3 节类似,不同之处仅仅是 2 个图符距离度量方法,不再赘述。

### 4 军标类型区分

文献[20]指出,将多个分类器的输出转换为置信度后,融合多个分类器,能取得比单个分类器更好的效果.融合规则有固定规则(fixed rules)和训练规则(trained rules)2类.固定规则简单易用,训练规则较为复杂,常用的固定规则就能取得较好效果,如 Max-rule, Median-rule, Sum-rule 等,因为 Max-rule 的效果几乎是固定规则中最好的,因此,本文采用 Max-rule. Max-rule 可具体表述为:对  $M$  类的分类问题,使用  $K$  个分类器,分类器输出为  $d_{kj}(x_k)$ ,其含义是采用第  $k$  个分类器时,待识别模式  $x$  到第  $j$  类的距离.转换后相应的置信度记为  $z_{kj}, k=1, 2, \dots, K; j=1, 2, \dots, M$ .融合后的置信度为

$$y_j(z) = \max(z_{kj}) \quad j=1, 2, \dots, M$$

设由点状军标覆盖模型计算出的待识别图符  $x$  的类型置信度为

$$P_{p1} = P(y=1 | x, MST_p)$$

$$P_{p0} = P(y=0 | x, MST_p)$$

由非规则军标覆盖模型计算出的类型置信度为

$$P_{i1} = P(y=1 | x, MST_i)$$

$$P_{i0} = P(y=0 | x, MST_i)$$

其中: $y$  的含义同式(4); $MST_p$  和  $MST_i$  分别表示点状和非规则军标的 MST 覆盖模型.按 Max-rule 的思想,融合方法为

$$\begin{cases} P(y=1 | x) = \max(P_{p1}, P_{i1}) \\ P(y=0 | x) = \max(P_{p0}, P_{i0}) \end{cases} \quad (13)$$

经置信度融合后,将待识别模式判别到置信度大的类型中去,然后再输入相应的识别模块进行分类,得到最终识别结果。

### 5 实验与分析

实验用的点状图形样本来自文献[21],该数据集包含了 113 类美军军标图符,由 8 个志愿者绘制而成,每人绘制 5 组,共有  $113 \times 5 \times 8 = 4520$  个样本.由于当前学术界对手绘非规则军标的研究很少,没有可供使用的公开数据,本文采集了以下数据集:共包含 36 种非规则军标,其中线状 23 种,面状 13 种,采集设备为 Wacom 绘图板 CTL-660,基本性能参数为采样分辨率 100 线/mm,精确度 0.25 mm,笔读取速度 133 点/s,由 5 个志愿者绘制,每人绘制 5 组,共得  $36 \times 5 \times 5 = 900$  个样本.各志愿者均具有军事学背景,能熟练使用计算机,年龄为 20~35 周岁.每个志愿者按照个人习惯绘制图符,每个样本绘制工整,不能过于随意.部分样本如图 5 所示。

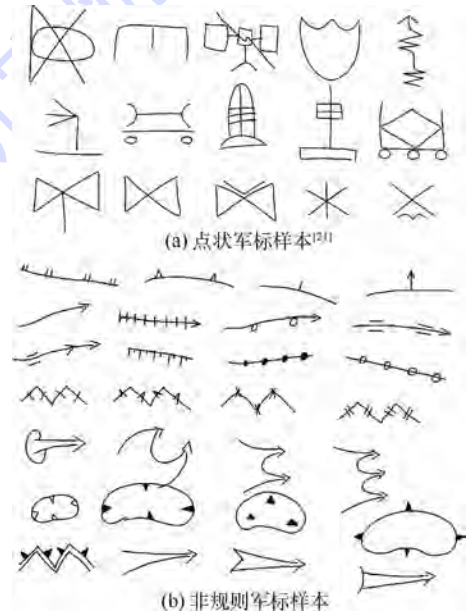


图 5 部分实验样本示例

Fig. 5 Subset of part symbols tested in experiments

实验代码采用 MATLAB 编写, SVM 分类器采用 LibSVM 工具包, 核函数采用 RBF 函数, 计算机配置: CPU: Interl Core i5-2400 3.10 GHz, RAM: 4 GB.

### 5.1 单一类型的军标草图识别

单一类型军标草图识别算法的有效性是混合识别的基础. 首先验证单一类型的军标草图识别率. 采用 5 重交叉法, 针对点状军标样本, 基于方向特征和 SVM 分类器<sup>[8]</sup> 的识别率为 98.3%; 针对非规则军标, 基于图元匹配的结构化识别方法<sup>[4]</sup> 的识别率为 96.2%. 证明了两个单一类型的军标识别模块的有效性.

### 5.2 基于 MST 覆盖模型的军标类型区分

为验证本文提出的基于 MST 覆盖模型的军标类型区分方法的效果, 设计如下 3 组实验.

第 1 组: 由于缺少对手绘军标混合识别问题的直接对比方法, 本文用文献[17]的方法分别建立点状和非规则军标的覆盖模型, 进行军标类型区分.

第 2 组: 分别单独采用第 2 节和第 3 节的方法建立点状军标和非规则军标的覆盖模型后, 使用 SVM 区分军标类型, 它们属于在文献[17]的基础上改进了决策规则.

第 3 组: 在第 2 组的基础上采用融合规则. 实验结果如表 1 所示.

表 1 军标类型区分准确率

Table 1 Accuracy of symbol type identification

方法	准确率/%	
	点状军标覆盖模型	非规则军标覆盖模型
文献[17]	85.0	91.1
本文方法(不融合)	87.3	93.2
本文方法(融合)	94.7	94.7

表 1 的结果表明:

1) 基于 MST 覆盖模型的军标类型区分方法有效, 直接利用文献[17]的方法, 就能取得较高的准确率, 且非规则军标覆盖模型的效果较好. 原因在于, 非规则军标采用结构匹配方法计算 2 个图符间的距离, 能较好地聚类同类样本, 且样本绘制较为工整, 不同类的样本混淆较少.

2) 当只使用一个覆盖模型时(不融合置信度), 本文方法效果好于直接使用文献[17]中的方法. 因为后者在军标类型区分时实际使用了线性判别函数, 如式(5)所示, 而本文通过 SVM 分类, 效果较好, 实验验证了这一点.

3) 融合置信度后的准确率有一定上升, 验证了同时利用点状和非规则军标覆盖模型的效果比

单一覆盖模型的效果好.

类型区分时易出错的图形如图 6 所示, 但错误图形较为集中, 在后续工作中, 可以有针对性地提取外围特征加以验证, 降低错误率.

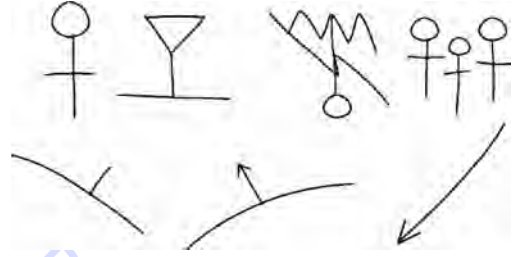


图 6 类型区分错误的图形

Fig. 6 Symbols that failed to get right types

### 5.3 混合点状和非规则军标的图符识别

采用融合规则确定待识别图符的军标类型后, 再用相应的识别算法确定具体类别, 最终识别率为 91.6%. 由于单一类型的军标草图识别已经能达到较高的精度, 识别错误的原因主要在于军标类型的区分错误. 若类型错误, 则类别无法识别正确.

识别速度方面, 由于 MST 覆盖模型可在训练阶段预先生成, 判别类型时只需计算待识别模式到 MST 的距离. 识别一个图符的平均时间为 167 ms. 由于特征提取可在用户交互绘制过程中在后台并发处理, 具有较大优化空间, 能满足交互的实时性.

此外, 融合 2 个覆盖模型(如 5.2 节的第 3 组实验)相比于使用单一覆盖模型(如 5.2 节的第 2 组实验), 在军标类型区分阶段虽然时间消耗较多, 但由于最终的图符识别仍能达到实时要求, 且军标类型区分准确率有所提高, 所以这种额外消耗是值得的.

## 6 结 论

在基于草图的态势标绘系统中, 对类型未知的手绘军标进行识别是一个必须解决的问题, 但缺少文献对此进行研究. 本文提出一种混合点状和非规则军标图符的手绘识别方法, 分别建立点状和非规则军标的 MST 覆盖模型, 并估计待识别模式的类型置信度, 经置信度融合后确定军标类型, 最后通过相应的军标识别模块进行识别. 实验验证了本文方法的有效性.

本文的缺点是, 军标类型区分和具体类别识别是 2 个独立过程, 若类型区分错误, 则该图符必然分类错误. 下一步的工作主要是将以上 2 个过

程统一在整体框架中,采用整体置信度的形式输出分类结果。

虽然本文的研究对象是军标图形,但标图作业在消防、电力、交通、公共安全、灾难救援等领域也有广泛的应用.不同领域具体的图形符号集不同,但在构成特点、标图方式上有很大相似之处,如广泛采用了点状、线状和面状图符.本文算法经适当修改,即可在其他领域中得到应用。

### 参考文献 (References)

- [1] Sun Z B, Wang C H, Zhang L Q, et al. Query-adaptive shape topic mining for hand-drawn sketch recognition [C] // Proceedings of 20th ACM International Conference on Multimedia. New York: ACM, 2012: 519-528.
- [2] 巩应奎, 范玉青, 吴志方. 笔式草图的增量识别 [J]. 北京航空航天大学学报, 2006, 32(5): 621-624.  
Gong Y K, Fan Y Q, Wu Z F. Incremental recognition method for pen-based sketch [J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(5): 621-624 (in Chinese).
- [3] 陈鸿, 汤晓安, 杨耀明, 等. 基于位移映射的非规则军队标号绘制算法 [J]. 计算机辅助设计与图形学学报, 2011, 23(5): 797-804.  
Chen H, Tang X A, Yang Y M, et al. Using displacement mapping for irregular military symbols rendering [J]. Journal of Computer-Aided Design & Computer Graphics, 2011, 23(5): 797-804 (in Chinese).
- [4] 邓维, 吴玲达, 张友根, 等. 采用快速图元匹配的手绘非规则军标识别 [J]. 计算机辅助设计与图形学学报, 2014, 26(10): 1835-1843.  
Deng W, Wu L D, Zhang Y G, et al. Using fast primitive matching to recognize sketched irregular military marking symbols [J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(10): 1835-1843 (in Chinese).
- [5] 吴玲达, 张友根, 邓维, 等. 面向手绘军标图形的旋转自由识别方法 [J]. 中国图象图形学报, 2014, 19(3): 456-462.  
Wu L D, Zhang Y G, Deng W, et al. Rotation free recognition of hand-drawn military marking symbols [J]. Journal of Image and Graphics, 2014, 19(3): 456-462 (in Chinese).
- [6] Hammond T, Logsdon D, Peschel J, et al. A sketch recognition interface that recognizes hundreds of shapes in course-of-action diagrams [C] // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI). New York: ACM, 2010: 4213-4218.
- [7] Cummings D, Fymat S, Hammond T. Sketch-based interface for interaction with unmanned air vehicles [C] // Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, 2012.
- [8] Deng W, Wu L D, Yu R H, et al. On-line sketch recognition using direction feature [C] // Proceedings of Human-Computer Interaction (INTERACT 2013). Heidelberg: Springer Verlag, 2013: 259-266.
- [9] Delaye A, Liu C L. Contextual text/non-text stroke classification in online handwritten notes with conditional random fields [J]. Pattern Recognition, 2014, 47(3): 959-968.
- [10] Malakar S, Das R K, Sarkar R, et al. Handwritten and printed word identification using gray-scale feature vector and decision tree classifier [J]. Procedia Technology, 2013, 10: 831-839.
- [11] Song Y H, Xiao G L, Zhang Y L, et al. A handwritten character extraction algorithm for multi-language document image [C] // Proceedings of 11th International Conference on Document Analysis and Recognition (ICDAR). Piscataway, NJ: IEEE Press, 2011: 93-98.
- [12] Rani R, Dhir R, Lehal G S. Script identification of pre-segmented multi-font characters and digits [C] // Proceedings of 12th International Conference on Document Analysis and Recognition (ICDAR). Piscataway, NJ: IEEE Press, 2013: 1150-1154.
- [13] 付强, 丁晓青, 蒋焰. 基于多信息融合的中文手写地址字符串切分与识别 [J]. 电子与信息学报, 2008, 30(12): 2916-2920.  
Fu Q, Ding X Q, Jiang Y. Segmentation and recognition algorithm for Chinese handwritten address character string [J]. Journal of Electronics & Information Technology, 2008, 30(12): 2916-2920 (in Chinese).
- [14] Blagojevic R, Chang S H H, Plimmer B. The power of automatic feature selection: Rubine on steroids [C] // Proceedings of the Seventh Sketch-Based Interfaces and Modeling Symposium (SBIM). Goslar: Eurographics Association, 2010: 79-86.
- [15] Paulson B, Hammond T. PaleoSketch: Accurate primitive sketch recognition and beautification [C] // Proceedings of 13th International Conference on Intelligent User Interfaces (IUI2008). New York: ACM, 2008: 1-10.
- [16] Paulson B. Rethinking pen input interaction: Enabling freehand sketching through improved primitive recognition [D]. State of Texas: Texas A&M University, 2010.
- [17] 胡正平, 冯凯. 高维空间多分辨率最小生成树模型的自适应一类分类算法 [J]. 自动化学报, 2012, 38(5): 769-775.  
Hu Z P, Feng K. An adaptive one-class classification algorithm based on multi-resolution minimum spanning tree model in high-dimensional space [J]. Acta Automatic Sinica, 2012, 38(5): 769-775 (in Chinese).
- [18] 谢毓湘, 梁柔道, 陈丹雯, 等. 一种基于局部不变特征的图像特定场景检测方法 [J]. 国防科技大学学报, 2013, 35(6): 78-83.  
Xie Y X, Luan X D, Chen D W, et al. A method of specific image scene detection based on local invariant features [J]. Journal of National University of Defense Technology, 2013, 35(3): 78-83 (in Chinese).
- [19] Liu C L, Hao H W, Sako H. Confidence transformation for combining classifiers [J]. Pattern Analysis and Applications, 2004, 7(1): 2-17.
- [20] Liu C L. Classifier combination based on confidence transformation [J]. Pattern Recognition, 2005, 38(1): 11-28.
- [21] Costagliola G, de Rosa M, Fuccella V. Recognition and auto-completion of partially drawn symbols by using polar histograms as spatial relation descriptors [J]. Computers & Graphics, 2014, 39(2): 101-116.

# Online sketch recognition for mixed point and irregular military symbols

DENG Wei<sup>1</sup>, WU Lingda<sup>\*1</sup>, ZHANG Yougen<sup>2</sup>, ZHAO Zhipeng<sup>3</sup>

(1. Science and Technology on Complex Electronic System Simulation Laboratory, Academy of Equipment, Beijing 101416, China;

2. Department of Information Systems, Academy of National Defense Information, Wuhan 430010, China;

3. Institute of Information Technology of Guilin, Guilin 541004, China)

**Abstract:** Most of current research on online sketched military symbols recognition concerns only one type of symbols, point symbols or irregular symbols, using different methods to recognize separately. But in practical applications the two types of symbols are mixed. It becomes a major issue to find a way to recognize a type-unknown military symbol. A minimum spanning tree (MST) covering model-based mixed recognition method was proposed. In the training phase, two MST-based covering models were built for point and irregular symbols respectively. And then a two-class support vector machine (SVM) classifier was trained. In the recognition phase, the coarse type identification was accomplished by using the geometrical and structural information firstly. Then the confidence estimations were calculated and integrated to identify the type of the unknown symbol. Different types of symbols were classified by two existing modules. The algorithm was tested on 113 classes of point symbols and 36 classes of irregular symbols. The accuracy rate of symbol type identification was 94.7%, and the final recognition rate was 91.6% in real time.

**Key words:** sketch recognition; point military symbol; irregular military symbol; minimum spanning tree (MST); classify

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0040

# 基于块平均预处理的 GNSS-R 延迟映射处理方法

王峰, 杨东凯\*, 李伟强, 张彦仲

(北京航空航天大学 电子信息工程学院, 北京 100191)



**摘 要:** 为了提高导航卫星反射信号信噪比, 提出了一种基于块平均预处理 (BAP) 的反射信号处理方法, 该方法首先以伪码周期为块尺寸对接收信号进行分块, 并对各分块进行叠加、平均, 然后对平均后的信号进行相关处理. 在此基础上, 理论分析了块平均预处理对导航反射信号的影响, 推导了镜面反射点反射信号处理增益的数学表达式, 以及相关处理后得到的一维相关功率, 对比分析了该方法与传统处理方法的计算复杂度, 分析表明该方法对反射信号具有空域滤波效果, 提高了镜面反射点反射信号的信噪比且有效降低了计算复杂度. 最后, 基于块平均预处理的软件接收机被开发, 并对课题组实采的岸基 GPS 数据进行了处理, 结果表明: 与未块平均预处理的传统方法相比, 在块平均预处理时间为 2、5、10、15 以及 20 ms 时, 该方法所得相关功率的峰值信噪比分别提高了约 3.5、6.7、10.2、10.6 和 10.4 dB, 且在块平均预处理时间为 10 ms 时, 处理等长数据的时间缩短了近 2.2 倍.

**关键词:** 全球导航卫星系统 (GNSS); 反射信号; 块平均; 信号处理; 信噪比 (SNR); 处理时间

**中图分类号:** TN965.5

**文献标识码:** A **文章编号:** 1001-5965 (2015) 10-1943-07

利用导航卫星反射信号进行地表参数反演是卫星导航应用的一个新型分支, 该技术通过岸/陆基、机载或星载的特殊装置接收, 处理经地表反射的导航卫星信号, 可进行海面高度测量<sup>[1-2]</sup>、海风探测<sup>[3-4]</sup>以及土壤湿度测量<sup>[5]</sup>, 是一种被动式探测技术, 也是无源双基/多基雷达的一种新手段, 具有接收装置轻、开发成本低以及应用面广等优势, 近 20 年得到了国内外学者的广泛关注.

延迟映射接收机 (Delay Mapping Receiver, DMR) 是接收、处理导航卫星反射信号的特殊装置. 文献[6]系统地描述了一种硬件 GPS 延迟/多普勒映射装置, 并在岸基和机载条件下对该装置进行了试验验证. 在国内, 北京航空航天大学电子信息工程学院基于现场可编程门阵列 (Field-

Programmable Gate Array, FPGA) 自主研发了 GRrSv. 2GPS 反射信号处理装置, 并利用该装置进行了海面风场、有效波高和土壤湿度测量等应用研究<sup>[7]</sup>. 在反射信号接收装置设计中, 如何提高信噪比一直是研究重点之一, 传统方法是进行非相干累加<sup>[8]</sup>, 但随着非相干累加次数的增多, 一方面平方损耗的增加限制了信噪比的提高, 另一方面计算复杂度线性增加. 针对上述问题, 本文提出了一种基于块平均预处理 (BAP) 的 GNSS 反射信号延迟映射处理方法, 该方法首先以伪码周期为块尺寸对接收信号进行分块, 对各分块进行叠加、平均, 然后对平均后的信号进行相关处理. 在此基础上, 理论分析了块平均预处理对反射信号的影响、镜面反射点反射信号的处理增益以及

收稿日期: 2015-01-22; 录用日期: 2015-04-17; 网络出版时间: 2015-05-22 15:31

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150522.1531.002.html

基金项目: 国家自然科学基金 (61171070); 国家“863”计划 (2011AA120501, 2013AA122402)

作者简介: 王峰 (1989—), 男, 甘肃兰州人, 博士研究生, wangf.19@163.com

\* 通讯作者: 杨东凯 (1972—), 男, 山东莱芜人, 教授, yangdongkai@sina.com, 主要研究方向为 GNSS 应用及室内定位.

**引用格式:** 王峰, 杨东凯, 李伟强, 等. 基于块平均预处理的 GNSS-R 延迟映射接收处理方法[J]. 北京航空航天大学学报, 2015, 41(10): 1943-1949. Wang F, Yang D K, Li W Q, et al. GNSS-R delay mapping processing method based on block averaging pre-processing[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1943-1949 (in Chinese).

一维时延相关功率模型,对比分析了该处理方法和未经块平均预处理的传统方法的计算复杂度.最后,利用2种处理方法对课题组实采的岸基GPS数据进行了处理验证,结果表明:与未经块平均预处理的传统方法相比,在块平均预处理时间为2、5、10、15以及20 ms时,该方法提高峰值信噪比约3.5、6.7、10.2、10.6和10.4 dB,且有效降低了计算复杂度,在块平均预处理时间为10 ms时,处理等时长数据的时间缩短了近2.2倍.

### 1 导航卫星反射信号模型

如图1所示,假设每个散射单元对应的散射信号相互独立,散射单元(x,y)对应的散射信号载波剥离后可以表示为

$$s_r(t, (x, y)) = A(x, y) \cdot \text{prn}(t - \tau(x, y)) \cdot \exp[-j(2\pi\Delta f(x, y)t + \varphi(x, y))] \quad (1)$$

式中:A(x,y)为散射单元(x,y)((0,0)为镜面反射点)对应散射信号的幅度;\tau(x,y)、\Delta f(x,y)和\varphi(x,y)分别为相对于发射时刻的时间延迟、载波剥离后的残余频率和相位;\text{prn}(t)为导航卫星信号的伪码.

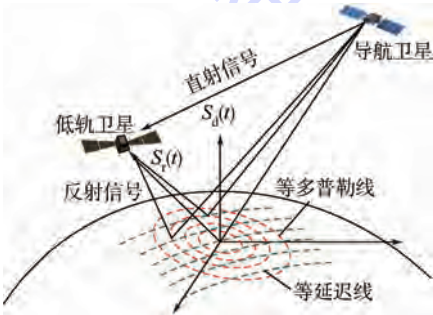


图1 观测区域反射示意图

Fig.1 Diagram of reflection in observation region

反射信号接收装置接收到的信号由各散射单元的散射信号共同作用<sup>[9]</sup>,可表示为

$$S_r(t) = \iint_S A(x, y) \cdot \text{prn}(t - \tau(x, y)) \cdot \exp[-j(2\pi\Delta f(x, y)t + \varphi(x, y))] dx dy + n(t) \quad (2)$$

式中:n(t)为信号噪声,假设为均值为0、方差为\sigma\_n^2的高斯白噪声;S为信号闪耀区.

### 2 基于BAP的反射信号处理方法

#### 2.1 基于块平均预处理的反射信号模型

块处理在导航数据的处理中已得到了广泛应用<sup>[10]</sup>,文献[11-12]论述了基于块平均的弱信号快速捕获算法.对导航数据进行块处理时,块尺寸

通常为伪码的整周期,本文设置块尺寸为伪码周期,对于GPS L1 CA码,块尺寸为1 ms.根据伪码周期性,任意时刻的伪码可表示为

$$\text{prn}(t) = \text{prn}(t' + nT) \Big|_{0 \leq t' \leq T} \quad (3)$$

式中:T为伪码周期;n为自然数.如图2所示,对L ms的反射信号S\_r(t)分块平均得

$$\bar{S}_r(t') = \frac{1}{L} \sum_{i=0}^{L-1} S_r(t' + iT) = \iint_S A(x, y) \cdot \text{prn}(t' - \tau(x, y)) \cdot \exp[-j(2\pi\Delta f(x, y)t' + \varphi(x, y))] \cdot \left\{ \frac{1}{L} \sum_{i=0}^{L-1} \exp[-j(2\pi\Delta f(x, y)iT)] \right\} dx dy + \bar{n}(t') \quad (4)$$

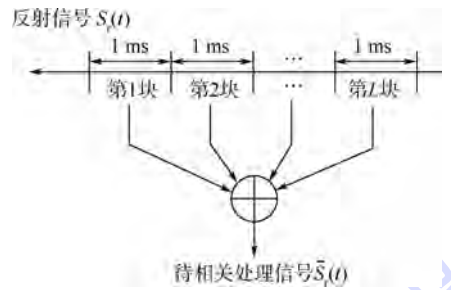


图2 L ms反射信号的块平均预处理示意图

Fig.2 Block averaging pre-processing diagram of L ms reflected signals

令H\_L(f) = \frac{1}{L} \left\{ \sum\_{i=0}^{L-1} \exp[-j(2\pi f iT)] \right\}, 则式(4)可表示为

$$\bar{S}_r(t') = \iint_S A(x, y) \cdot \text{prn}(t' - \tau(x, y)) \cdot \exp[-j(2\pi\Delta f(x, y)t' + \varphi(x, y))] \cdot H_L(\Delta f(x, y)) dx dy + \bar{n}(t') \quad (5)$$

H\_L(f)为块平均预处理响应函数,其表达式为

$$H_L(f) = \frac{1}{L} \sum_{i=0}^{L-1} \exp[-j(2\pi f iT)] = \frac{1 - \exp(-j\pi f TL)}{1 - \exp(-j\pi f T)} = \frac{1}{L} \frac{\sin(\pi f TL)}{\sin(\pi f T)} \exp[-j\pi f T(L-1)] \quad (6)$$

式中:L为分块数目,考虑到导航电文数据位的符号跳变,一般分块数目L不大于导航电文数据宽度与伪码宽度的比值,对于GPS L1 CA码,L不大于20.H\_L(f)是一个梳状滤波器,其幅频特性如图3所示,滤波器通带峰值位于i kHz,其中i=0,1,2,\dots,过零带宽为1/L kHz,通带频率间隔固定为1 kHz,与分块数目无关,但是通带宽度与分块数目相关,随着分块数目的增多,通道宽度变窄.

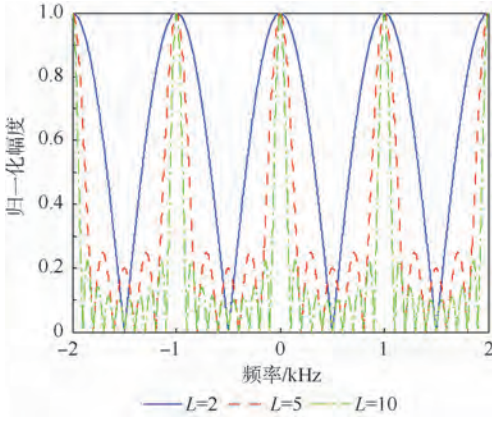


图 3 块平均预处理的幅频特性

Fig.3 Amplitude-frequency characteristics of block averaging pre-processing

### 2.2 基于块平均预处理的反射信号处理结构

基于块平均预处理的反射信号处理结构如图 4 所示,直射和反射信号分别由直射通道(点划线上部)和反射通道(点划线下部)处理.直射通道完成直射信号的捕获和跟踪,得到的载波频率和码相位为反射通道提供参考信息<sup>[13]</sup>.反射通道由载波生成模块、码生成模块、块平均预处理模块以及相关处理模块组成.

反射通道载波生成模块产生频率为  $f_a + \Delta f$

的本地载波,对反射信号进行载波剥离,其中,  $f_a$  为直射通道载波环路输出的直射信号多普勒频率,  $\Delta f$  为频率补偿量;码生成模块产生初始码相位为  $\tau_a + \Delta \tau$  的本地伪码,其中,  $\tau_a$  为直射通道码跟踪环路输出的直射信号码相位,  $\Delta \tau$  为码补偿量;块平均预处理模块对载波剥离后的反射信号进行分块、平均处理;相关处理模块对块平均预处理后的信号和本地伪码进行相关运算得到一维时延相关功率.

由于反射信号在相关窗内进行处理,首先需对本地码和反射信号进行粗对齐,使处理后的反射信号相关功率峰值处于相关窗内,此过程被称谓码补偿.通常,码补偿量  $\Delta \tau$  定义为反射信号镜面反射点处的信号分量相对于直射信号在码延迟上发生的偏移量,可由式(7)得到其粗略的估计值<sup>[14]</sup>.

$$\Delta \tau = \frac{2H \sin \theta}{c} f_c \quad (7)$$

式中:  $H$  为接收平台高度;  $\theta$  为卫星仰角;  $c$  为光速;  $f_c$  为伪码码率,对于 GPS L1 CA 码而言,其值为 1.023 MHz.

本地载波频率和反射信号频率未对齐会产生相关功率的损失,因此,需在直射通道载波跟踪环

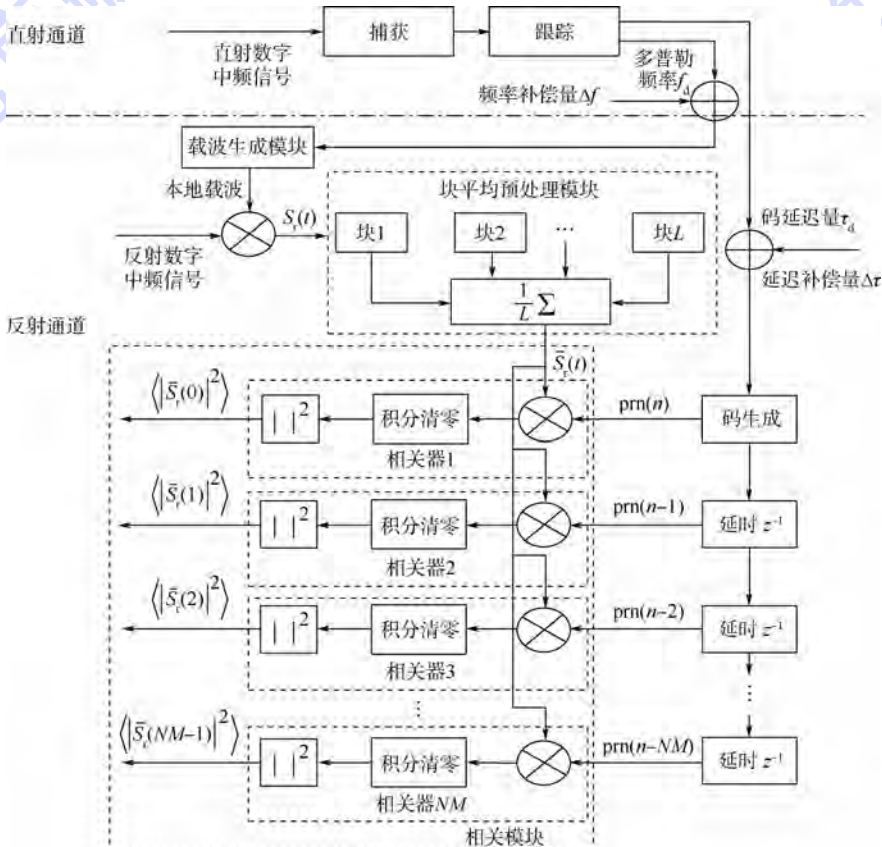


图 4 基于块平均预处理的反射信号处理结构图

Fig.4 Structure diagram of processing reflected signals based on the block averaging pre-processing

输出结果的基础上进行频率补偿使相关功率的损失尽量减小. 频率补偿量  $\Delta f$  可由式(8)<sup>[7]</sup>得到:

$$\lambda \Delta f = \mathbf{v}_r \cdot \mathbf{u}_{rs} + \mathbf{v}_i \cdot \mathbf{u}_{st} - (\mathbf{v}_i - \mathbf{v}_r) \cdot \mathbf{u}_{tr} \quad (8)$$

式中:  $\lambda$  为载波波长;  $\mathbf{v}_i$  和  $\mathbf{v}_r$  分别为导航卫星和接收平台速度;  $\mathbf{u}_{tr}$ 、 $\mathbf{u}_{rs}$  和  $\mathbf{u}_{st}$  分别为导航卫星到接收平台的单位矢量、接收平台到镜面反射点的单位矢量以及镜面反射点到导航卫星的单位矢量. 在岸基条件下, 由于反射信号相对于直射信号的码延时和载波频率的变化很小, 因此, 码补偿量和频率补偿量可设为 0.

### 2.3 基于块平均预处理的反射信号一维相关功率

假设反射通道载波生成模块产生的本地载波的频率与镜面反射点的反射信号频率相等, 根据文献[15], 可得基于块平均预处理的一维时延相关功率为

$$\begin{aligned} \langle |\bar{S}_i(\tau)|^2 \rangle &= \int_s A^2(x, y) \cdot \Lambda^2(\tau - \tau(x, y)) \cdot \\ &\text{sinc}^2(\pi(\Delta f(0, 0) - \Delta f(x, y))T_{\text{coh}}) \cdot \\ &|H_L(\Delta f(0, 0) - \Delta f(x, y))|^2 dx dy \quad (9) \end{aligned}$$

式中:  $\Lambda(\tau)$  为伪码自相关函数;  $\text{sinc}(x) = \sin x/x$ . 从式(9)可知, 基于块平均预处理的反射信号经历了 3 次空域滤波: ①相关积分时以  $\text{sinc}$  函数进行空域滤波; ②相关积分时以伪码自相关函数进行空域滤波<sup>[15]</sup>; ③块平均预处理时以  $H_L(f)$  进行空域滤波.

## 3 性能分析

### 3.1 镜面反射点处理增益

假设在块平均预处理前, 散射单元  $(x, y)$  对应散射信号的信噪比为  $R_{in}(x, y)$ , 块平均预处理后, 对应信号的信噪比为  $R_{out}(x, y)$ , 定义块平均预处理增益为  $G_{\text{BAP}}(x, y)$ , 其表达式为

$$\begin{aligned} G_{\text{BAP}}(x, y) &= \frac{R_{out}(x, y)}{R_{in}(x, y)} = \\ \frac{\frac{\sqrt{2}}{2} A^2(x, y) |H_L(\Delta f(x, y))|^2 / \bar{\sigma}_n^2}{\frac{\sqrt{2}}{2} A^2(x, y) / \sigma_n^2} &= \\ \frac{|H_L(\Delta f(x, y))|^2 \sigma_n^2 / \bar{\sigma}_n^2}{\sigma_n^2} &\quad (10) \end{aligned}$$

式中:  $\bar{\sigma}_n^2$  为块平均预处理后的噪声功率, 根据白噪声的特点, 有如下关系:

$$\bar{\sigma}_n^2 = \frac{1}{L} \sigma_n^2 \quad (11)$$

因此,  $G_{\text{BAP}}(x, y)$  可表示为

$$G_{\text{BAP}}(x, y) = 10 \lg(L |H_L(\Delta f(x, y))|^2) \quad (12)$$

假设通过频率补偿使反射通道的本地载波频率与镜面反射点的反射信号频率相等, 即  $\Delta f(0, 0) = 0$ , 则满足<sup>[12]</sup>

$$G_{\text{BAP}}(0, 0) = 10 \lg L \quad (13)$$

因此, 镜面反射点反射信号的处理增益为

$$G_{\text{spe}} = G_{\text{BAP}}(0, 0) + G_{\text{coh}} + 10 \lg m - l(m) \quad (14)$$

式中:  $G_{\text{coh}}$  为相关解扩增益, 对于 GPS L1 CA 而言, 其值为 30 dB;  $m$  为非相干累加次数,  $l(m)$  为平方损耗<sup>[8]</sup>. 假设反射信号相关功率的输出率为  $f_R$ , 则存在

$$L \cdot m = \frac{1}{T \cdot f_R} \quad (15)$$

将式(15)代入式(14)可得在反射信号相关功率输出率一定的条件下, 镜面反射点的处理增益为

$$G_{\text{spe}} = G_{\text{coh}} - 10 \lg T \cdot f_R - l(m) \quad (16)$$

当一维相关功率输出率  $f_R$  一定时, 随着块平均预处理次数  $L$  的增加, 非相干累加次数  $m$  减小, 平方损耗减小, 镜面反射点反射信号的处理增益增大. 当反射信号相关功率输出率为 1 Hz 时, GPS L1 CA 码的镜面反射点增益随  $L$  的变化结果如表 1 所示, 其中,  $L = 1$  表示未经过块平均处理的方法.

表 1 一维时延相关功率输出率为 1 Hz 时, GPS L1 CA 码的镜面反射点信号处理增益

Table 1 Processing gain of signals reflected by specular points for GPS L1 CA code when the output rate of one-dimensional correlation power is 1 Hz

$L$	$m$	$G_{\text{spe}}/\text{dB}$	与传统方法增益差/dB
1	1000	50.0	0
2	500	51.4	1.4
5	200	53.3	3.3
10	100	54.6	4.6
20	50	55.8	5.8

### 3.2 计算复杂度

假设相关功率输出率为  $f_R$ , 码周期为  $T$ , 进行  $L$  ms 数据的块平均预处理, 一个码周期内的采样点数为  $N$ , 进行  $M$  个不同时延的一维相关功率计算, 其中  $f_R/L = S_2, f_R/T_c = S_1$  ( $S_1, S_2$  为整数). 相关积分的乘法计算量为  $2N$ , 加法计算量为  $2(N - 1)$ , 进行  $M$  个时延相关值的乘法计算量为  $2MN$ , 加法计算量为  $2M(N - 1)$ , 计算相关功率的乘法计算量为  $2M$ , 加法计算量为  $M$ . 表 2 给出了未块平均预处理的传统方法和基于块平均预处理方法的计算复杂度. 对于 GPS L1 CA 码, 当  $L = 10$ , 且计算 64 个时延点的一维相关功率时, 与未经块平

均预处理的传统方法相比,乘法计算量缩小了 10 倍,加法计算量缩小了 9.3 倍。

表 2 计算复杂度

Table 2 Computational complexity

计算类型	$L=1$	$L>1$	缩小倍数
乘法	$2MS_1(N+1)$	$2MS_2(N+1)$	$S_1/S_2$
加法	$MS_1(2N-1)$	$NS_2(2M+L-1)$	$S_1/S_2(1+L/2M)$

## 4 岸基数据处理

本文对课题组 2013 年 8 月 2 日 9 时 38 分在广东省阳江市海洋气象观测站采集的岸基 GPS L1 CA 码数据进行了处理。观测点相对于海平面高度约为 134 m,观测区域位于观测点东南方向,可观测卫星高度角范围为  $10^\circ \sim 35^\circ$ 。14 号卫星反射信号在反射天线覆盖范围内且信号最强,因此,反射通道对 14 号卫星反射信号进行处理。

图 5 为 2 种处理结构处理得到的反射信号一维时延相关功率,从图 5 可知:①相比于传统处理方法,块平均预处理时间为 2、5、10、15 以及 20 ms 时,所得一维时延相关功率的峰值高了 3.4、6.8、11.7、12.1 和 11.8 dB,定义峰值信噪比为

$$R_{\text{peak}} = \frac{\max\{\langle |S_r(\tau)|^2 \rangle\}}{P_{\text{noise}}} \quad (17)$$

式中: $P_{\text{noise}}$  为相关功率底噪功率,可通过额外配置噪声通道的方法测得<sup>[16]</sup>,信噪比对比结果如表 3 所示,块平均预处理为 2、5、10、15 以及 20 ms 的处理较传统处理方法的峰值信噪比分别提高了 3.5、6.7、10.2、10.6 和 10.4 dB,但是,由于峰值功率为镜面反射点与周围散射单元信号相关功率的功率和,因此,峰值信噪比的增益差较表 1 理论仿真分析的镜面反射点信号处理增益的增益差要大;②随着块平均预处理的时间增大,峰值信噪比增加,但是在块平均预处理为 10、15 和 20 ms 时,峰值信噪比基本相同,主要是由于导航电文存在数据位比特跳变<sup>[16]</sup>,消除了部分累加效果,导致信噪比无法持续增大,且块平均预处理时间越接近导航电文数据位宽度,处于数据符号跳变的概率越大,导致信噪比下降,如图 5 所示的块平均预处理时间为 20 ms 时比 15 ms 时的峰值信噪比低。综上分析,对于存在导航电文的导航信号,并不是块平均预处理时间越长越好,块平均预处理时间应该小于导航电文数据位宽度;对于无导航电文的信号(例如:L5C 导频信号),块平均预处理时间可以设置较长。

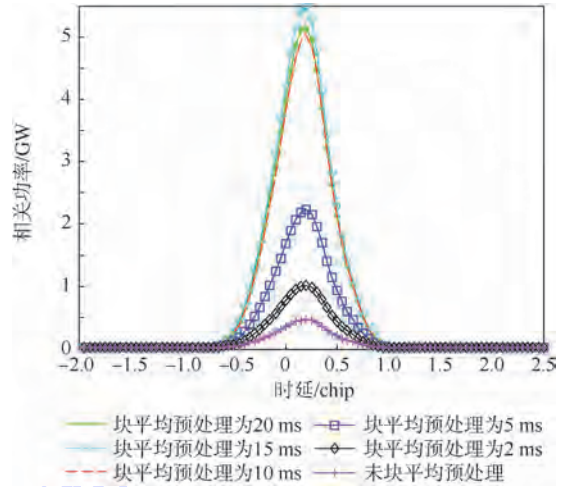


图 5 反射信号一维时延相关功率

Fig. 5 One-dimensional delay correlation power of reflected signals

表 3 信噪比对比结果

Table 3 Comparable results of signal-to-noise ratio

$L$	峰值功率/dB	底噪/dB	信噪比/dB
1	86.7	73.9	12.8
2	90.1	73.8	16.3
5	93.5	73.9	19.5
10	97.0	74.0	23.0
15	97.4	74.0	23.4
20	97.1	73.9	23.2

表 4 为块平均预处理时间为 10 ms 时,2 种处理结构的软件实现方式下,处理等长数据所用时间的对比结果,程序运行平台为 ThinkCenter M8300 台式计算机,其处理器为 Inter(R) Core(TM) i5-2400 CPU @ 3.10 GHz,内存 3 GB,操作系统为 Microsoft Windows XP Professional. 运行结果表明:处理等长度数据,基于块平均预处理的处理结构所用时间明显比传统处理结构所用时间短,缩短了约 2.2 倍。但是,由于程序实际运行时间不仅取决于算法设计,还受输入数据规模、编译器、处理器速度等诸多复杂的软硬件因素的影响,因此,处理时长的缩短比例与前面分析的计算复杂度缩短比例不符,且不同的处理平台和程序编写风格,处理时长的缩短比例不同。

表 4 2 种处理结构的数据处理时长对比

Table 4 Comparison of data-processing time for two processing structures

处理数据长度/ms	软件处理时长/s		缩短倍数
	$L=1$	$L=10$	
1000	10.2	4.6	2.22
60000	609.2	287.5	2.12
120000	1235.2	562.8	2.19

## 5 结 论

本文通过对基于块平均预处理的 GNSS-R 反射信号处理方法的理论分析表明:相比于传统的未经块平均预处理的处理方法相比,该方法:

1) 在伪码自相关函数和 sinc 函数空域滤波的基础上,增加了一次空域滤波。

2) 在一维时延相关功率输出率恒定的条件下,减小了非相干累加次数,进而减小平方损耗,有效提高了镜面反射点反射信号增益。

3) 降低计算复杂度的倍数近似为块平均预处理的分块数。

通过该方法和传统处理方法对实采数据处理,对比分析处理结果表明,该方法:

1) 有效提高了反射信号一维相关功率的峰值信噪比,对于存在导航电文的信号,块平均预处理时间受到限制,当与导航电文比特位宽相近时,峰值信噪比增加受到限制,甚至开始下降,对于无导航电文存在的信号,块平均预处理时间可以设置较长。

2) 有效缩短了处理等时长数据的时间,本文中的处理时长缩短了近 2.2 倍。

对于我国独立研发的全球卫星导航系统,其 B1 信号中进行了 NH (Neumann-Hoffman) 码的二次编码,如何提高镜面反射点处的信噪比还需要进一步研究,这将是本文后续的研究工作。

### 参考文献 (References)

- [1] D'Addio S, Martin-Neira, di Bisceglie M, et al. GNSS-R altimeter based on Doppler multi-looking[J]. IEEE Journal of Selected Topic in Applied Earth Observation and Remote Sensing, 2014, 7(5):1452-1460.
- [2] Valencia E, Zavorotny V U, Akos D M, et al. Using DDM asymmetry metrics for wind direction retrieval from GPS ocean-scattered signals in airborne experiments[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(7):3924-3936.
- [3] Garrison J L, Komjathy A, Zavorotny V U, et al. Wind speed measurement using forward scattered GPS signals[J]. IEEE Transactions on Geoscience and Remote Sensing, 2002, 40(1):50-65.
- [4] Li W Q, Yang D K, Fabra F, et al. Typhoon wind speed observation utilizing reflected signals from BeiDou GEO satellites[C]// China Satellite Navigation Conference (CSBC). Berlin: Springer-Verlag, 2014:191-200.
- [5] Rodriguez-Alvarez N, Bosch-Lluis X, Camps A, et al. Soil moisture retrieval using GNSS-R techniques: Experimental results over a are soil field[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(11):3616-3624.
- [6] Stephen T L, Peter K, Garth F, et al. A Delay/Doppler-mapping receiver system for GPS-reflection remote sensing[J]. IEEE Transactions on Geoscience and Remote Sensing, 2002, 40(5):1150-1163.
- [7] Li W Q, Yang D K, Zhang B. Real-time processing of reflected GNSS signal for remote sensing: System and experiments[J]. The Journal of Navigation, 2011, 64(S1):127-140.
- [8] 杨东凯, 张其善. GNSS 反射信号处理基础与实践[M]. 北京: 电子工业出版社, 2012:90-91.  
Yang D K, Zhang Q S. GNSS reflected signal processing: Fundamental and application[M]. Publishing House of Electronics Industry, 2012:90-91 (in Chinese).
- [9] 张建军, 袁洪, 王宪. 基于 GNSS 散射信号陆地高度计的建模与分析[J]. 系统仿真学报, 2009, 21(10):2810-2814.  
Zhang J J, Yuan H, Wang X. Modeling and analysis of land surface altimeter based on GNSS scattered Signal[J]. Journal of System Simulation, 2009, 21(10):2810-2814 (in Chinese).
- [10] Feng G, van Graas F. GPS receiver block processing[C]// Proceedings of the ION Conference GPS'99, 1999:307-316.
- [11] Mohamed S, Moeness G A. Acquisition of weak GNSS signals using a new block averaging pre-processing[C]// Position Location and Navigation IEEE Symposium. Piscataway, NJ: IEEE Press, 2008:1362-1372.
- [12] 鲁郁. GPS 全球定位接收机——原理与软件实现[M]. 北京: 电子工业出版社, 2009:77-78.  
Lu Y. GPS receiver: Principles and software realization[M]. Beijing: Publishing House of Electronics Industry, 2009:77-78 (in Chinese).
- [13] 杨东凯, 丁文锐, 张其善. 软件定义的 GNSS 反射信号接收机设计[J]. 北京航空航天大学学报, 2009, 35(9):1048-1051.  
Yang D K, Ding W R, Zhang Q S. Software defined GNSS reflected receiver design[J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(9):1048-1051 (in Chinese).
- [14] 路勇, 熊华刚, 杨东凯. GNSS-R 海洋遥感原始数据采集系统研究与实现[J]. 哈尔滨工程大学学报, 2009, 30(6):644-648.  
Lu Y, Xiong H G, Yang D K. A raw data acquisition system for detecting ocean wind-fields[J]. Journal of Harbin Engineering University, 2009, 30(6):644-648 (in Chinese).
- [15] 张益强. 基于 GNSS 反射信号的海洋微波遥感技术[D]. 北京: 北京航空航天大学, 2008.  
Zhang Y Q. Ocean microwave remote sensing using GNSS reflected signals[D]. Beijing: Beijing University of Aeronautics and Astronautics, 2008 (in Chinese).
- [16] 谢刚. GPS 原理与接收机设计[M]. 北京: 电子工业出版社, 2011:358-364.  
Xie G. Principle of GPS and receiver design[M]. Beijing: Publishing House of Electronics Industry, 2012:358-364 (in Chinese).

## GNSS-R delay mapping processing method based on block averaging pre-processing

WANG Feng, YANG Dongkai\*, LI Weiqiang, ZHANG Yanzhong

(School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

**Abstract:** To improve signal-to-noise ratio of global navigation satellite system reflected signals, a method of processing navigation satellites reflected signal based on block averaging pre-processing (BAP) was presented, which divided the received signals into blocks with period of pseudocode and achieved the mean values of blocks by superposition and averaging, and then performed the correlation process for the mean values. On the basis, the model of reflected signal which processed by BAP was theoretically analyzed; the mathematical expression of processing gain of signal reflected by specular point and correlation power model were derived; computational complexity of this method was compared with one of traditional method. The conclusion is that this method conducts the spatial filtering for reflected signals, improves signal-to-noise ratio and effectively reduces computational complexity. At last, a software receiver was developed to process coastal raw GPS data collected. The results show that when the BAP time are 2, 5, 10, 15 and 20 ms, the method improves about 3.5, 6.7, 10.2, 10.6 and 10.4 dB on signal-noise ratio of peak and reduces the processing time by about 2.2 times, with BAP being 10 ms, compared with the traditional method.

**Key words:** global navigation satellite system (GNSS); reflected signal; block averaging; signal processing; signal-to-noise ratio (SNR); processing time

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0100



# 基于均方根容积粒子的 SMC-PHD 算法

刘哲<sup>1,2</sup>, 王祖林<sup>\*1</sup>, 徐迈<sup>1</sup>, 刘景贤<sup>1</sup>, 杨蓝<sup>1</sup>

(1. 北京航空航天大学 电子信息工程学院, 北京 100191; 2. 中北大学 信息与通信工程学院, 太原 030051)

**摘 要:** 传统的序贯蒙特卡罗概率假设密度(SMC-PHD)算法采用状态转移密度作为重要性采样函数. 当目标非线性运动时, 少数粒子将具有较大的权值, 导致估计精度低、结果发散. 针对上述问题, 提出了一种基于均方根容积卡尔曼滤波(SCKF)和统计门限技术的重要性采样函数设计方法. 在重要性采样函数估计时, 首先利用 SCKF 对重要性采样函数的均值和协方差阵进行预测, 而后利用统计门限技术提取与重要性采样粒子相关联的量测. 通过相应的权值对所提取的量测进行合并, 更新重要性采样函数的均值和协方差阵. 在此基础上将设计的重要性采样函数应用于 SMC-PHD 的强度预测和更新, 最终实现多目标状态和数目的估计. 实验表明, 本算法在非线性的多目标跟踪中具有精度高、估计结果稳定的优点.

**关键词:** 序贯蒙特卡罗; 概率假设密度; 重要性采样; 均方根容积卡尔曼滤波; 统计门限

中图分类号: TP242

文献标识码: A 文章编号: 1001-5965(2015)10-1950-09

多目标跟踪是指利用传感器的量测对目标的状态(包括位置、速度等)进行估计<sup>[1-2]</sup>. 目前多目标跟踪已广泛应用于遥测遥感<sup>[3]</sup>、飞行器跟踪<sup>[4]</sup>、可视化目标跟踪<sup>[5]</sup>以及分布式融合<sup>[6]</sup>等领域. 以最近邻域卡尔曼滤波(Nearest Neighbor Kalman Filter, NNKF)、多假设前提(Multiple Hypothesis Tracking, MHT)和联合概率数据关联(Joint Probabilistic Data Association, JPDA)为代表的传统多目标跟踪方法通过目标和量测之间明确的关联关系进行跟踪. 但是, 多目标跟踪场景下目标的状态和数目通常随时间的变化而变化. 此外受漏检、虚警和杂波等影响, 传感器量测同样也随时间变化. 由于目标状态、数目、量测的不确定性, 在实际中, 状态和量测集合的明确关联关系难以建立.

近年来, 随机有限集(Random Finite Sets, RFS)理论在多目标跟踪领域得到了广泛的应用.

Mahler<sup>[7]</sup>基于随机有限集理论, 提出了概率假设密度(PHD)滤波算法. 该算法将目标的状态和量测分别用随机有限集进行表示, 通过强度函数的传播, 完成对目标状态和数目的估计. PHD 算法无需目标的状态和量测的关联, 其计算复杂度低于传统的多目标跟踪算法<sup>[8]</sup>. 但是 PHD 算法在估计过程中涉及多维积分, 在实际中, 这类多维积分通常没有解析解. Vo 等<sup>[9]</sup>将序贯蒙特卡罗(SMC)算法的思想应用于 PHD 滤波, 提出了序贯蒙特卡罗 PHD(SMC-PHD)滤波算法. 该算法将 PHD 方程中多维积分转换成多个具有一定权值的样本(也称为粒子)的求和, 避免了多维积分的求解. 由于该方法具有实现简单、估计精度高的优点, 因此引起了国内外众多学者的关注.

文献[10-12]使用状态转移密度函数作为重要性采样函数, 极大简化了 SMC-PHD 的计算. 但

收稿日期: 2015-02-17; 录用日期: 2015-04-30; 网络出版时间: 2015-06-24 16:46

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150624.1646.005.html

作者简介: 刘哲(1981—), 男, 山西运城人, 博士研究生, liuzhe201@buaa.edu.cn

\* 通讯作者: 王祖林(1965—), 男, 湖北潜江人, 教授, wzulin@buaa.edu.cn, 主要研究方向为雷达与电子对抗.

引用格式: 刘哲, 王祖林, 徐迈, 等. 基于均方根容积粒子的 SMC-PHD 算法[J]. 北京航空航天大学学报, 2015, 41(10): 1950-1958. Liu Z, Wang Z L, Xu M, et al. SMC-PHD algorithm based on squared cubature particles[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1950-1958 (in Chinese).

是这类算法忽略了当前量测集对目标状态估计的影响,当目标非线性运动时,会导致少量粒子具有大的权值,造成估计误差增大. Punithakumar 等<sup>[13]</sup>通过使用多个运动模型混合作为目标运动模型,将模型转移概率函数作为重要性采样函数,提高了非线性运动目标的估计精度.但是该算法跟踪性能依赖于多模型的选择.文献[14]将辅助粒子滤波的思想引入 SMC-PHD,通过在 SMC-PHD 算法的重要性采样函数设计中引入辅助变量,使得重要性采样函数抽样产生的粒子权值具有最小的方差.与传统的 SMC-PHD 算法相比,基于辅助粒子滤波的 SMC-PHD 算法提高了跟踪的精度,但是在严重的非线性或强状态噪声场景下该算法会出现很大的跟踪误差. Yoon 等<sup>[15]</sup>将无迹信息滤波(UIF)应用于 SMC-PHD,提高了非线性目标的跟踪精度.但是 UIF 中 Sigma 点的选择对算法跟踪性能影响较大. Li 等<sup>[16]</sup>将 Sigma 门引入 SMC-PHD 的更新,提出了高速 Sigma-gating SMC-PHD 算法.该算法仅使用目标周围的量测进行粒子更新,减少了 SMC-PHD 的计算量.

综上所述,目前常用的 SMC-PHD 算法多采用状态转移密度作为重要性采样函数,忽略了当前量测集对重要性采样函数的影响.当目标非线性运动时,这类算法呈现粒子退化、跟踪误差大的缺点.因此本文利用 SCKF 算法<sup>[17]</sup>和统计门限技术,提出了一种基于均方根容积粒子的序贯蒙特卡罗 PHD(SC-SMC-PHD)算法.本算法主要包括 2 个部分:

1) 针对传统 SMC-PHD 在非线性的多目标跟踪精度不高、结果易发散问题,本算法将近年来出现的 SCKF 算法引入 SMC-PHD 的重要性采样函数设计,提出了基于 SCKF 的重要性采样函数设计方法.通过利用 SCKF 在非线性目标跟踪下良好的准确性和稳定性,提高了 SMC-PHD 的跟踪精度.

2) 针对重要性采样函数设计时重要性采样粒子状态和量测关联关系未知的问题,本文提出了一种基于统计门限技术的量测合并方法.通过计算预测量测与量测集中量测的统计距离,并设定阈值,提取与预测量测接近的量测.在此基础上,对所提取的量测按照统计距离的大小分配不同的权值,从而为重要性采样函数设计提供逼真的量测.

和传统的 SMC-PHD 算法相比,本文所提出的 SC-SMC-PHD 算法解决了传统的 SMC-PHD 算法在非线性的多目标跟踪时存在的误差大、粒子退化等问题,具有精度高、结果稳定的优点.

# 1 SMC-PHD

## 1.1 PHD

令  $\mathcal{X}$  表示状态空间.假设  $k$  时刻,目标的状态表示为  $\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,N}$ ,  $N$  为目标的目标数目.显然  $\mathbf{x}_{k,j} \in \mathcal{X}$ .根据随机有限集理论,目标的状态可以用式(1)所示的随机有限集来表示:

$$\mathbf{X}_k = \{\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,N}\} \in \mathcal{F}(\mathcal{X}) \quad (1)$$

式中:  $\mathcal{F}(\mathcal{X})$  为状态空间  $\mathcal{X}$  中所有闭合的子集.  $\mathbf{x}_{k,j}$  ( $j=1, 2, \dots, N$ ) 与顺序无关,因此  $\mathbf{X}_k$  为随机有限集.类似的,假设  $k$  时刻单个目标的量测用  $\mathbf{z}_{k,i}$  ( $i=1, 2, \dots, M$ ) 表示,多目标的量测可以用式(2)来表示:

$$\mathbf{Z}_k = \{\mathbf{z}_{k,1}, \mathbf{z}_{k,2}, \dots, \mathbf{z}_{k,N}\} \in \mathcal{F}(\mathcal{Z}) \quad (2)$$

式中:  $\mathcal{F}(\mathcal{Z})$  为量测空间  $\mathcal{Z}$  中所有闭合的子集.

PHD 滤波器利用一阶量(也称强度)的传递进行多目标状态和数目的估计. PHD 滤波算法进行迭代时,主要包括预测和更新 2 个步骤.为描述这 2 个步骤,假定  $D_{k-1|k-1}(\mathbf{x}_{k-1} | \mathbf{Z}_{1:k-1})$  为  $k-1$  时刻的后验强度,  $\mathbf{Z}_{1:k-1}$  为  $k-1$  时刻之前的所有量测.

1) 预测. 假设  $p_s(\mathbf{x}_k)$  为状态  $\mathbf{x}_k$  的存活概率,则预测强度函数  $D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$  表示为

$$D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) = \underbrace{\int_{\mathcal{X}} p_s(\mathbf{x}_k) f_{k|k-1}(\mathbf{x}_k | \mathbf{x}) D_{k-1|k-1}(\mathbf{x} | \mathbf{Z}_{1:k-1}) d\mathbf{x}}_{\text{存活粒子强度}} + \gamma_k(\mathbf{x}_k) \quad (3)$$

式中:  $f_{k|k-1}(\cdot)$  为单目标的状态转移密度;  $\gamma_k(\cdot)$  为新生目标强度.

2) 更新. 实际测量过程中,由于漏检、虚警以及杂波等影响,真实量测与理想量测存在差异.考虑这些不利情况,假设  $k$  时刻检测概率为  $p_d(\mathbf{x}_k)$ , 杂波密度为  $\kappa_k(\cdot)$ , 后验强度函数为  $D_{k|k}(\mathbf{x}_k | \mathbf{Z}_{1:k})$ , 如式(4)所示:

$$D_{k|k}(\mathbf{x}_k | \mathbf{Z}_{1:k}) = (1 - p_d(\mathbf{x}_k)) D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) + \sum_{\mathbf{z} \in \mathcal{Z}_k} \frac{p_d(\mathbf{x}_k) g(\mathbf{z} | \mathbf{x}_k)}{\kappa_k(\mathbf{z}) + C_z} D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) \quad (4)$$

式中:  $g(\cdot)$  为似然函数;  $C_z$  可通过式(5)计算得到:

$$C_z = \int_{\mathcal{X}} p_d(\mathbf{x}_k) g(\mathbf{z} | \mathbf{x}_k) D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) d\mathbf{x}_k \quad (5)$$

由式(3)和式(4)可以看出, PHD 滤波器无需考虑状态集与量测集的关联,仅涉及单个目标

在状态空间的积分. 由于单个目标状态通常为多维向量, 式(3)和式(4)为多维积分, 这类积分在实际中很难得到闭合解. 针对这一问题, SMC-PHD算法作为一种有效的 PHD 实现算法应运而生.

## 1.2 SMC-PHD 简介

SMC-PHD 滤波算法借鉴粒子滤波的思想, 利用重要性采样函数生成大量的粒子, 将 PHD 滤波中的强度函数更新预测转换成粒子集合的预测更新, 避免了 PHD 算法中的多维积分求解. 为描述 SMC-PHD 算法基本原理, 定义  $k-1$  时刻的粒子集为  $\{\mathbf{x}_{k-1}^{(i)}, w_{k-1|k-1}^{(i)}\}_{i=1}^{L_{k-1}}$ , 其中,  $\mathbf{x}_{k-1}^{(i)}$  为由状态  $\mathbf{x}_{k-1}$  采样生成的第  $i$  个粒子状态,  $w_{k-1|k-1}^{(i)}$  为其对应的权值. 设  $\pi(\cdot)$  和  $q(\cdot)$  分别为新生粒子与存活粒子对应的重要性采样函数. 在此基础上, SMC-PHD 算法的执行流程为<sup>[8]</sup>

1) 预测. 令  $J_k$  表示新生粒子的数目, 预测强度  $D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$  表示为

$$D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) = \sum_{i=1}^{L_{k-1}+J_k} w_{k|k-1}^{(i)} \delta(\mathbf{x} - \mathbf{x}_k^{(i)}) \quad (6)$$

式中:

$$\begin{cases} \mathbf{x}_k^{(i)} \sim \\ q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{Z}_k) & i=1, 2, \dots, L_{k-1} \\ \pi(\mathbf{x}_k^{(i)}, \mathbf{Z}_k) & i=L_{k-1}+1, L_{k-1}+2, \dots, L_{k-1}+J_k \end{cases} \quad (7)$$

$$w_{k|k-1}^{(i)} = \begin{cases} \frac{p_s(\mathbf{x}_{k-1}^{(i)}) \cdot f(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{Z}_k)} & i=1, 2, \dots, L_{k-1} \\ \frac{\gamma(\mathbf{x}_k^{(i)})}{J_k \cdot \pi(\mathbf{x}_k^{(i)}, \mathbf{Z}_k)} & i=L_{k-1}+1, L_{k-1}+2, \dots, L_{k-1}+J_k \end{cases} \quad (8)$$

式中:  $w_{k|k-1}^{(i)}$  为第  $i$  个粒子的预测权值;  $J_k = \nu \int \gamma(\mathbf{x}) d\mathbf{x}$ ,  $\nu$  为单个新生目标对应的粒子个数.

由式(6)~式(8)可以看出, 预测强度  $D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$  的计算可以通过粒子的权值与状态的预测来实现.

2) 更新. 本步骤利用当前量测集  $\mathbf{Z}_k$  对预测强度  $D_{k|k-1}(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$  函数进行更新, 更新得到的后验强度函数为

$$D_{k|k}(\mathbf{x}_k | \mathbf{Z}_{1:k}) = \sum_{i=1}^{L_{k-1}+J_k} w_{k|k}^{(i)} \delta(\mathbf{x} - \mathbf{x}_k^{(i)}) \quad (9)$$

式中:

$$w_{k|k}^{(i)} =$$

$$\left( 1 - p_d(\mathbf{x}_k^{(i)}) + \sum_{z \in \mathbf{Z}_k} \frac{p_d(\mathbf{x}_k^{(i)}) g_k(z | \mathbf{x}_k^{(i)})}{\kappa(z) + C_z} \right) w_{k|k-1}^{(i)} \quad (10)$$

其中:

$$C_z = \sum_{i=1}^{L_{k-1}+J_k} p_d(\mathbf{x}_k^{(i)}) g_k(z | \mathbf{x}_{k,i}) w_{k|k-1} \quad (11)$$

通过上述推导, 得到了更新后的粒子集  $\{\mathbf{x}_k^{(i)}, w_{k|k}^{(i)}\}_{i=1}^N$ . 为提取目标的状态, 传统的 SMC-PHD 滤波通常借助 EM (Expectation-Maximum)、K-means 等算法<sup>[9]</sup>, 对粒子集  $\{\mathbf{x}_k^{(i)}, w_{k|k}^{(i)}\}_{i=1}^N$  进行聚类, 形成对目标状态和数目的估计. 此外, 为避免多次迭代造成粒子退化, 通常在更新之后引入重采样策略, 减少粒子退化对估计造成的影响.

## 2 均方根容积 SMC-PHD 算法

### 2.1 问题描述

假设目标的非线性状态和量测方程为

$$\begin{cases} \mathbf{x}_k = \phi(\mathbf{x}_{k-1}) + \boldsymbol{\varepsilon}_{k-1} \\ \mathbf{z}_k = \varphi(\mathbf{x}_k) + \boldsymbol{\eta}_k \end{cases} \quad (12)$$

式中:  $\phi(\cdot)$  和  $\varphi(\cdot)$  为非线性函数;  $\boldsymbol{\varepsilon}_{k-1}$  和  $\boldsymbol{\eta}_k$  为  $k-1$  和  $k$  时刻的状态噪声与量测噪声. 本文假设  $\boldsymbol{\varepsilon}_{k-1}$  和  $\boldsymbol{\eta}_k$  均为零均值高斯分布, 协方差阵分别用  $\mathbf{Q}_{k-1}$  和  $\mathbf{R}_k$  表示.

由式(12)易知,  $\mathbf{x}_k$  和  $\mathbf{z}_k$  均为高斯分布. 为将当前量测引入重要性函数的设计, 本文选择后验概率作为式(7)中的重要性采样函数, 即

$$q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{Z}_{1:k}) = N(\mathbf{x}_k^{(i)}; \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}) \quad (13)$$

$$\pi(\mathbf{x}_k^{(i)}, \mathbf{Z}_k) = N(\mathbf{x}_k^{(i)}; \mathbf{m}_k^{(i)}, \mathbf{P}_k^{(i)}) \quad (14)$$

式中:  $\mathbf{m}_k^{(i)}$  和  $\mathbf{P}_k^{(i)}$  分别为第  $i$  个粒子在  $k$  时刻的状态均值和协方差矩阵;  $N(\cdot)$  表示高斯分布;  $i$  的取值范围与式(7)相同. 根据式(13)和式(14), 重要性采样函数  $q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{Z}_{1:k})$  和  $\pi(\mathbf{x}_k^{(i)}, \mathbf{Z}_k)$  的计算问题可以简化为参数  $\mathbf{m}_k^{(i)}$  和  $\mathbf{P}_k^{(i)}$  的计算. 为解决这一问题, 本文利用 SCKF 算法, 通过容积点的更新对  $\mathbf{m}_k^{(i)}$  和  $\mathbf{P}_k^{(i)}$  进行估计.

为描述方便, 本文使用  $\mathbf{m}_k$  和  $\mathbf{P}_k$  代表  $\mathbf{m}_k^{(i)}$  和  $\mathbf{P}_k^{(i)}$ .

### 2.2 基于 SCKF 的重要性采样函数估计

由前面的讨论, 式(13)和式(14)中重要性采样函数可以通过  $\mathbf{m}_k$  和  $\mathbf{P}_k$  的估计来完成. 为实现这一目标, 本节提出了一种基于 SCKF 算法的重要性采样函数估计方法. 本方法通过容积点的预测和更新, 对当前时刻粒子的均值和协方差进行估计. 估计过程包括预测和更新 2 个步骤.

1) 预测. 设  $k-1$  时刻的状态和协方差分别

为  $\mathbf{m}_{k-1}$  和  $\mathbf{P}_{k-1}, \mathbf{S}_{k-1}$  为  $\mathbf{P}_{k-1}$  的均方根(可以通过 cholesky 分解计算). 状态预测时,首先利用三阶容积准则<sup>[17]</sup>计算容积点,通过容积点传播计算预测状态  $\mathbf{m}_{klk-1}$  和预测均方根  $\mathbf{S}_{klk-1}$ . 预测过程如下:

① 容积点计算

根据三阶容积准则,容积点集  $\{\alpha_j, \rho_j\}$  可以用式(15)计算:

$$\begin{cases} \alpha_j = \sqrt{n} [\mathbf{1}]_j \\ \rho_j = \frac{1}{2n} \end{cases} \quad j=1, 2, \dots, 2n \quad (15)$$

式中:  $n$  为待估计目标状态的维数,  $\alpha_j$  和  $\rho_j$  分别为第  $j$  个容积点的状态及权值; 向量  $[\mathbf{1}]_j$  为集合  $[\mathbf{1}]$  的第  $j$  个向量, 其中集合  $[\mathbf{1}]$  为

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix} \right\}$$

② 计算第  $j$  个容积点

$$\mathbf{x}_{k-1,j} = \mathbf{S}_{k-1} \alpha_j + \mathbf{m}_{k-1} \quad (16)$$

③ 容积点传播

$$\mathbf{x}_{k-1,j}^* = \phi(\mathbf{x}_{k-1,j}) \quad (17)$$

④ 状态预测

$$\mathbf{m}_{klk-1} = \sum_{j=1}^{2n} \rho_j \mathbf{x}_{k-1,j}^* \quad (18)$$

⑤ 预测误差协方差矩阵均方根估计

$$\mathbf{S}_{klk-1} = \text{Tria}([\mathbf{C}_{klk-1,j}, \mathbf{S}_{Q,k-1}]) \quad (19)$$

式中:  $\text{Tria}(\cdot)$  为矩阵的三角分解(本文采用 QR 分解);  $\mathbf{S}_{Q,k-1}$  为状态噪声协方差阵  $\mathbf{Q}_{k-1}$  的均方根, 矩阵  $\tau_{klk-1,j}^*$  可通过式(20)计算:

$$\begin{aligned} \mathbf{C}_{klk-1,j} &= \frac{1}{\sqrt{2n}} \cdot \\ &[\mathbf{x}_{k-1,1}^* - \mathbf{m}_{klk-1} \quad \mathbf{x}_{k-1,2}^* - \mathbf{m}_{klk-1} \quad \dots \quad \mathbf{x}_{k-1,2n}^* - \mathbf{m}_{klk-1}] \end{aligned} \quad (20)$$

通过上面的计算,便可以得到预测状态  $\mathbf{m}_{klk-1}$  及其对应的误差协方差矩阵均方根  $\mathbf{S}_{klk-1}$ .

2) 更新. 设  $\mathbf{z}_{k,i}$  为预测的量测,  $\mathbf{z}_{k,i}$  为当前量测集  $\mathbf{Z}_k$  的第  $i$  个量测,  $\mathbf{S}_{R,k}$  为量测噪声的协方差矩阵均方根,  $K$  为卡尔曼滤波增益. 与状态预测类似, 量测更新时, 利用前面得到的  $\mathbf{m}_{klk-1}$  和  $\mathbf{S}_{klk-1}$ , 通过容积点在量测空间的传播, 计算后验状态  $\mathbf{m}_k$  及其对应的误差协方差矩阵  $\mathbf{S}_k$ .

① 容积点计算

第  $j$  个预测容积点可以表示为

$$\mathbf{x}_{klk-1,j} = \sqrt{\mathbf{P}_{klk-1}} \alpha_j + \mathbf{m}_{klk-1} \quad (21)$$

② 容积点传播

将第  $j$  个预测容积点代入式(12)的量测方程, 可得

$$\mathbf{x}_{klk-1,j}^* = \varphi(\mathbf{x}_{klk-1,j}) \quad (22)$$

③ 量测预测

$$\mathbf{z}_{klk-1} = \sum_{j=1}^{2n} \rho_j \mathbf{x}_{klk-1,j}^* \quad (23)$$

④ 新息均方根协方差矩阵均方根估计

$$\mathbf{S}_{klk-1}^{zz} = \text{Tri}([\mathbf{V}_{klk-1,j}, \mathbf{S}_{R,k}]) \quad (24)$$

式中:

$$\mathbf{V}_{klk-1,j} = \frac{1}{\sqrt{2n}} \cdot$$

$$[\mathbf{x}_{klk-1,1} - \mathbf{z}_{klk-1} \quad \mathbf{x}_{klk-1,2} - \mathbf{z}_{klk-1} \quad \dots \quad \mathbf{x}_{klk-1,2n} - \mathbf{z}_{klk-1}] \quad (25)$$

⑤ 互协方差矩阵估计

$$\mathbf{P}_{klk-1}^{mz} = \mathbf{E}_{klk-1,j} \mathbf{V}_{klk-1,j}^T \quad (26)$$

式中:

$$\mathbf{E}_{klk-1,j} = \frac{1}{\sqrt{2n}} \cdot$$

$$[\mathbf{x}_{klk-1,1} - \mathbf{m}_{klk-1} \quad \mathbf{x}_{klk-1,2} - \mathbf{m}_{klk-1} \quad \dots \quad \mathbf{x}_{klk-1,2n} - \mathbf{m}_{klk-1}] \quad (27)$$

⑥ 卡尔曼滤波增益估计

$$\mathbf{K}_k = (\mathbf{P}_{klk-1}^{mz} / (\mathbf{S}_{klk-1}^{zz})^T) / \mathbf{S}_{klk-1}^{zz} \quad (28)$$

⑦ 状态及协方差矩阵均方根估计

$$\mathbf{m}_k = \mathbf{m}_{klk-1} + \mathbf{K}_k (\mathbf{z}_k - \mathbf{z}_{klk-1}) \quad (29)$$

$$\mathbf{S}_k = \text{Tri}([\mathbf{V}_{klk-1,j}, \mathbf{S}_{R,k}]) \quad (30)$$

通过预测与更新, 最终得到  $\mathbf{m}_k$  和  $\mathbf{P}_k$  ( $\mathbf{P}_k = \mathbf{S}_k \mathbf{S}_k^T$ ).

2.3 基于统计门限技术的量测合并

由式(29)计算第  $i$  个粒子的后验均值  $\mathbf{m}_k$  时, 需使用与其关联的量测  $\mathbf{z}_k$  进行更新. 但是在多目标跟踪场景下, 当前的量测集往往由多个目标的量测和杂波组成. 在标准的 SMC-PHD 算法中, 第  $i$  个粒子状态和量测  $\mathbf{z}_k$  的关联关系通常是未知的. 此外, 由于杂波在量测空间随机分布, 杂波可能出现在目标量测附近, 导致估计精度降低.

针对上述问题, 本文提出了一种基于统计门限技术的量测合并算法. 该算法首先计算重要性采样粒子的预测量测与当前量测集中量测的统计距离, 在此基础上, 通过设定的阈值, 提取与重要性采样粒子相关联的量测集. 而后根据统计距离的大小对门限内的量测分配不同的权值, 最后对不同权值的量测进行合并, 得到重要性采样粒子对应的量测, 并用于重要性采样函数的均值更新.

在标准的 PHD 模型中, 每个目标至多只能产生一个量测, 反之亦然. 所以在 SMC-PHD 算法

中,属于同一目标的粒子至多只能由一个量测与之关联.因此,本文借鉴最近邻法的思想,将统计意义上与粒子量测预测 $z_{k|k-1}$ 接近的量测作为可能与粒子关联的量测.首先,定义 $z_{k|k-1}$ 与 $z_{k,j}$ ( $z_{k,j} \in Z_k$ )的统计距离为

$$d_j^2(z_{k,j}, z_{k|k-1}) = (z_{k,j} - z_{k|k-1})(P_k^{zz})^{-1}(z_{k,j} - z_{k|k-1})^T \quad (31)$$

根据式(31),通过选取特定的阈值,与第*i*个粒子关联的量测可以用式(32)的统计门限提取:

$$\hat{Z}_k = \{z_{k,j} | d_j^2(z_{k,j}, z_{k|k-1}) < T_h, z_{k,j} \in Z_k\} \quad (32)$$

式中: $T_h$ 为事先设定的阈值; $P_k^{zz}$ 为新息协方差阵; $P_k^{zz} = S_k^{zz}(S_k^{zz})^T$ .

由式(32)得到第*i*个粒子相关联的量测集,为降低门限内杂波对状态更新的影响,本文提出依据门限内量测统计距离的大小,为每个量测分

配相应的权值并进行合并,从而抑制关联量测附近杂波对目标状态估计的影响.

令 $\hat{z}_i \in \hat{Z}_k$ ,其对应的权值可以表示为

$$\rho_i = \frac{1}{d_i^2} \times \frac{1}{D} \quad (33)$$

式中: $D = \sum_{i=1}^{|\hat{Z}_k|} \frac{1}{d_i^2}$ 为归一化因子,显然 $\sum_{i=1}^{|\hat{Z}_k|} \rho_i = 1$ .

由式(33)可知, $z_{k|k-1}$ 所对应的量测 $z_k$ 可以表示为

$$z_k = \sum_{i=1}^{|\hat{Z}_k|} \rho_i \hat{z}_i \quad (34)$$

通过上述的推导,便得到了合并后的量测 $z_k$ .将 $z_k$ 代入式(29),便可以计算粒子的状态均值 $m_k$ .最后将 $m_k$ 和 $P_k$ 代入式(13)~式(14),执行式(7)~式(10)的SMC-PHD算法,整体流程如图1所示.

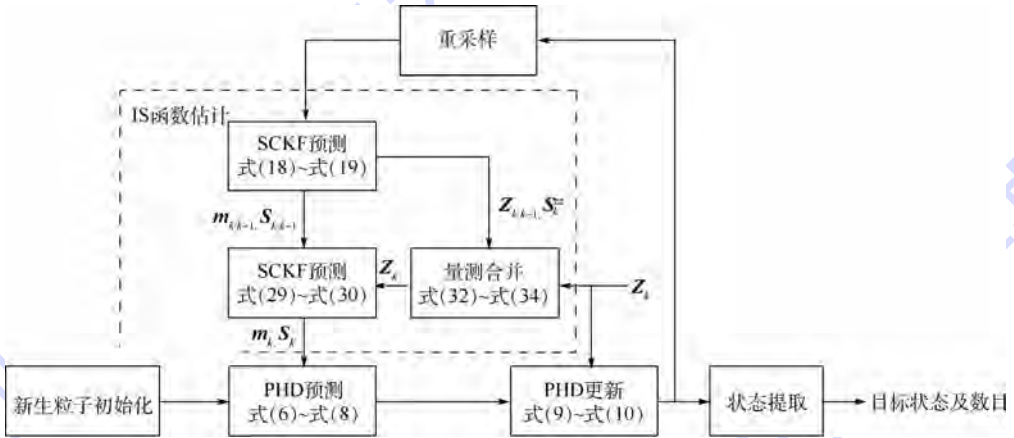


图1 SC-SMC-PHD算法执行流程

Fig.1 Execution process of SC-SMC-PHD algorithm

图1中,重采样采用多项式重采样,新生粒子强度初始化和状态提取使用文献[12]的提取算法.

另外, $T_h$ 的设定可通过查 $\chi^2$ 分布表得到.过小的 $T_h$ 将导致大量的量测无法用于更新,增加估计误差;过大的 $T_h$ 则会造成运算量过大.

### 3 仿真实验

本节将对本文所提出的SC-SMC-PHD算法与文献[12]中的IBSMC-PHD算法进行比较.为验证算法的有效性,本文选择最优次模式分配(OSPA)距离<sup>[18]</sup>和均方根误差(RMSE)作为度量标准.

#### 3.1 实验场景

本文采用文献[8]所使用的运动场景.设*k*时刻目标的状态向量为 $x = [x_k \ y_k \ \dot{x}_k \ \dot{y}_k \ \alpha_k]^T$ ,其中 $(x_k, y_k)$ 为目标的位置; $(\dot{x}_k, \dot{y}_k)$ 为速度分量;

$\alpha_k$ 为*k*时刻的转弯速率.目标的运动方程为

$$x_k = \begin{bmatrix} 1 & \frac{\sin(\alpha_{k-1}T)}{\alpha_{k-1}} & 0 & -\frac{1 - \cos(\alpha_{k-1}T)}{\alpha_{k-1}} & 0 \\ 0 & \cos(\alpha_{k-1}T) & 0 & -\sin(\alpha_{k-1}T) & 0 \\ 0 & \frac{1 - \cos(\alpha_{k-1}T)}{\alpha_{k-1}} & 1 & \frac{\sin(\alpha_{k-1}T)}{\alpha_{k-1}} & 0 \\ 0 & \sin(\alpha_{k-1}T) & 0 & \cos(\alpha_{k-1}T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot x_{k-1} + \begin{bmatrix} \frac{T^2}{2} & 0 & 0 \\ T & 0 & 0 \\ 0 & \frac{T^2}{2} & 0 \\ 0 & T & 0 \\ 0 & 0 & 1 \end{bmatrix} \epsilon_{k-1} \quad (35)$$

量测方程为

$$z_k = \begin{bmatrix} \arctan\left(\frac{y_k}{x_k}\right) \\ \sqrt{x_k^2 + y_k^2} \end{bmatrix} + \boldsymbol{\eta}_k \quad (36)$$

仿真所采用的电脑 CPU 为 Intel 双核 2.93 GHz,内存 2 GB, MATLAB 版本为 2013a. 场景参数设定如下:目标的状态如表 1 所示,目标运动轨迹如图 2 所示(杂波数为 10). 目标存活概率  $p_s(\mathbf{x}_k) = 0.99$ , 检测概率  $p_d(\mathbf{x}_k) = 0.95$ , 统计距离门限  $T_h = 16$ , 每个存活目标和新生目标对应的粒子数分别为 5 和 100.

表 1 目标初始状态设置  
Table 1 Target initial states setting

目标序号	起始状态	出现时间/s	消失时间/s
1	[325 5 375 5 0]	2	30
2	[500 5 300 5 0]	15	40
3	[400 5 475 -5 0]	25	55
4	[300 5 450 -5 0]	44	70
5	[275 -5 375 5 0]	44	70

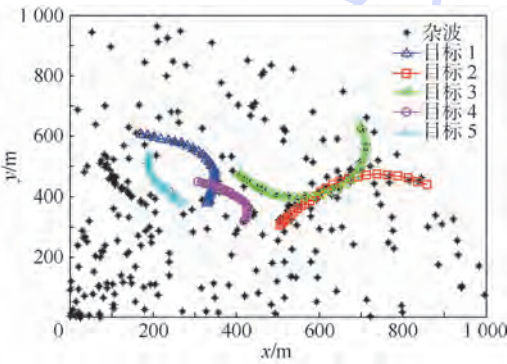


图 2 目标的真实航迹

Fig. 2 Ground-truth tracks of targets

### 3.2 仿真及分析

仿真 1 不同门限下的算法性能.

为验证阈值对算法性能的影响,根据  $\chi^2$  分布表,本文将统计门限距离分别取 4、16、25、49 以及 81. 为分析跟踪精度,本文使用 OSPA 距离和估计数目的 RMSE 等作为衡量指标.

本文提出的 SC-SMC-PHD 算法经 500 次蒙特卡罗结果如表 2 所示.

表 2 不同门限下的性能对比

Table 2 Performance comparison under different thresholds

门限	平均 OSPA/m	平均 RMSE
4	26.56	0.28
16	25.34	0.26
25	25.91	0.27
49	26.51	0.28
81	27.52	0.29

由表 2 可知,随着门限的增大,平均 OSPA 和 RMSE 的增幅不大. 这说明 SC-SMC-PHD 算法采用统计门限合并的方法,具有良好的稳定性.

仿真 2 不同量测误差的算法性能.

改变量测噪声中距离均方根偏差,使用 SC-SMC-PHD 算法跟踪图 2 的目标,500 次蒙特卡罗仿真结果如图 3 和图 4 所示.

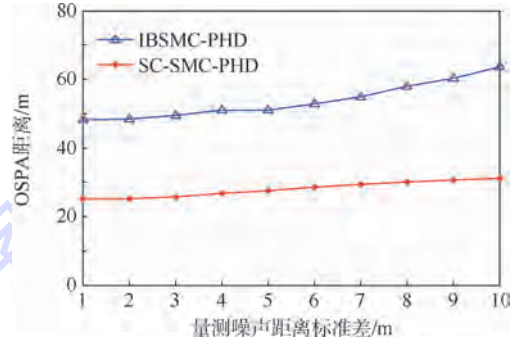


图 3 不同量测噪声距离标准差下的 OSPA 距离

Fig. 3 OSPA distances under different range standard deviations of observation noise

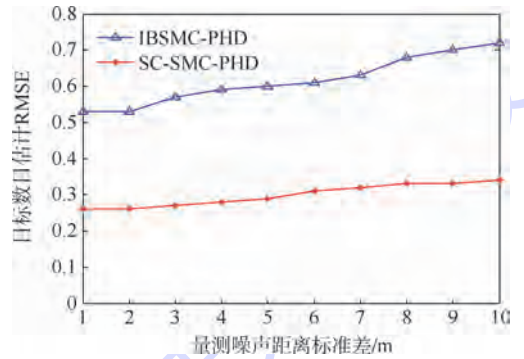


图 4 不同量测噪声距离标准差下的 RMSE

Fig. 4 RMSEs under different range standard deviations of observation noise

由图 3 和图 4 可知,当量测噪声标准差从 1 m 变化到 10 m 时,SC-SMC-PHD 算法的变化趋势小于 IBSMC-PHD 算法. 这说明本算法在量测噪声增大时仍能保持良好的跟踪精度.

仿真 3 固定杂波下多目标跟踪性能对比.

对图 2 的目标分别使用 SC-SMC-PHD 算法与 IBSMC-PHD 算法进行估计,每种算法各进行 500 次蒙特卡罗仿真,本算法执行所需时间为 0.14 s, 而 IBSMC-PHD 为 0.13 s. 由于 IBSMC-PHD 使用状态转移函数  $f(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})$  作为重要性采样函数  $q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{Z}_k)$ , 避免了式(8)的重要性函数估计,因此其时间复杂度低于本算法.

图 5 和图 6 分别为 IBSMC-PHD 和本文提出的 SC-SMC-PHD 算法 500 次蒙特卡罗仿真的估计航迹. 从图 5 和图 6 可以看出,IBSMC-PHD 算法

在对目标 1, 目标 4 和目标 5 估计时, 多数的估计值和真实目标位置重合. 但是对具有强非线性运动的目标 2 和目标 3 估计时, IBSMC-PHD 算法出现了大量的目标丢失. 相比而言, 本文提出的 SC-SMC-PHD 算法对上述 5 个目标均具有良好的估计效果, 如图 6 所示.

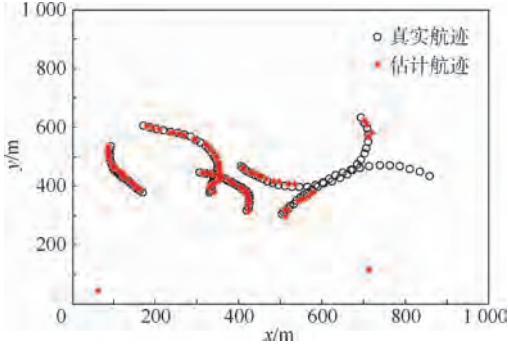


图 5 IBSMC-PHD 算法估计航迹

Fig. 5 Estimated tracks of IBSMC-PHD algorithm

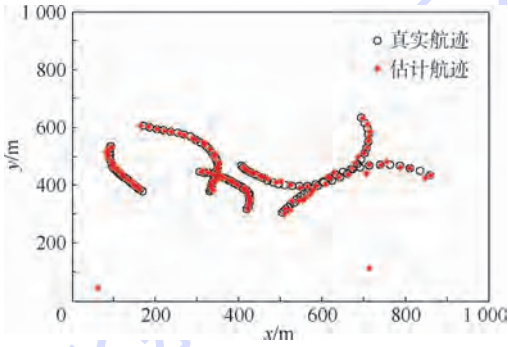


图 6 SC-SMC-PHD 算法估计航迹

Fig. 6 Estimated tracks of SC-SMC-PHD algorithm

首先分析 SC-SMC-PHD 和 IBSMC-PHD 算法对起始航迹的跟踪效果, 这里使用文献 [19] 所定义的虚假航迹起始概率和目标正确起始概率作为衡量标准. SC-SMC-PHD 的虚假起始概率为 0.13, IBSMC-PHD 为 0.41, 目标的正确起始概率如表 3 所示. 由表 3 可以看出, 对各目标的正确起始概率而言, SC-SMC-PHD 算法均高于 IBSMC-PHD 算法. 尤其是对同时出现目标 4 和目标 5 时而言, IBSMC-PHD 算法几乎无法获取二者起始航迹. 相比而言 SC-SMC-PHD 算法仍能保持良好的识别概率.

表 3 目标正确起始概率对比

Table 3 Comparison of target correct initial probability

目标序号	SC-SMC-PHD	IBSMC-PHD
1	0.97	0.89
2	0.84	0.75
3	0.88	0.57
4	0.83	0.09
5	0.87	0.12

为进一步对比二者的跟踪性能, 图 7 为二者的 OSPA 距离, 其中 SC-SMC-PHD 算法的平均 OSPA 距离为 27.87 m, 而 IBSMC-PHD 算法的平均 OSPA 距离为 47.82 m. 由图 7 可以看出, 从第 26 s 到第 55 s 内, IBSMC-PHD 算法的 OSPA 距离具有较大的波动, 其余时间内二者 OSPA 距离接近, 这是因为从第 26 s 到第 55 s 内, 具有强非线性运动的目标 2 和目标 3 的存在造成了 IBSMC-PHD 算法的估计误差增大. 由于本文所提出的 SC-SMC-PHD 算法将当前的量测引入重要性采样函数的设计, 并使用 SCKF 与 Gating 技术进行重要性函数的估计, 因此在整个跟踪时间内, 本算法具有相对稳定的跟踪效果. 此外, 从图中还可以看出, 在 2、15、25 和 45 s 处, SC-SMC-PHD 算法的 OSPA 曲线出现 4 个峰值, 这代表在这些时刻有新的目标出现, 因此 SC-SMC-PHD 算法可以准确地检测目标的出现.

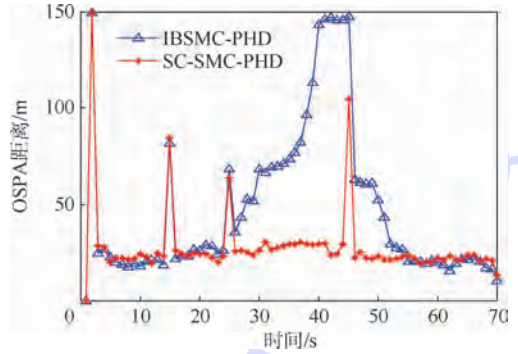


图 7 OSPA 距离

Fig. 7 OSPA distances

从图 8 的目标数目估计曲线可以看出, 和 IBSMC-PHD 算法相比, SC-SMC-PHD 算法的目标数目估计曲线更加接近真实目标数目曲线. 图 9 为图 8 目标数目估计的 RMSE 曲线, 图中 SC-SMC-PHD 算法的平均 RMSE 为 0.26, IBSMC-PHD 算法为 0.53. 与图 7 类似, 第 26 s 到第 55 s 内, IBSMC-PHD 算法的 RMSE 曲线出现较大的波

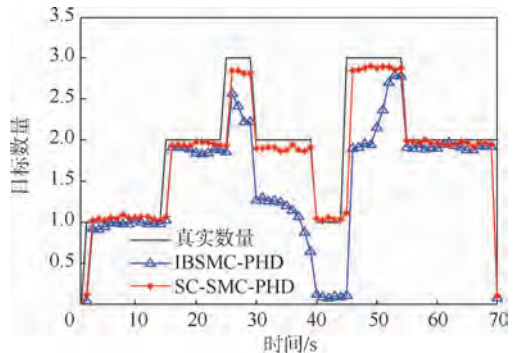


图 8 目标数目估计

Fig. 8 Estimated target numbers

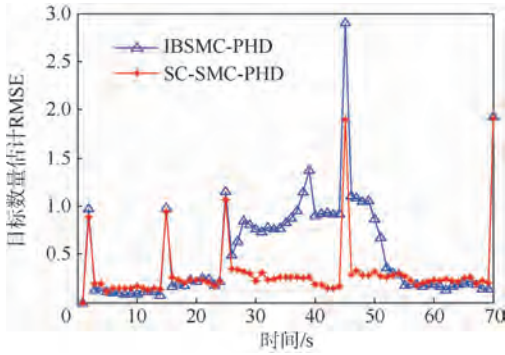


图 9 目标数目估计 RMSE

Fig. 9 RMSEs of estimated target numbers

动,而 SC-SMC-PHD 算法的估计曲线在 1~70 s 比较稳定.从图 7 到图 9,可以看出,通过本文所提出的重要性采样函数设计方法,SC-SMC-PHD 算法明显优于 IBSMC-PHD 算法,但其时间复杂度略高于 IBSMC-PHD 算法.

仿真 4 不同杂波下多目标跟踪性能对比.

为验证算法在不同杂波下跟踪的稳定性,改变图 2 中的杂波数目,对 SC-SMC-PHD 与 IBSMC-PHD 分别进行 500 次蒙特卡罗仿真.表 4 为各算法的平均 OSPA 距离和 RMSE.从表 4 中可以看出,随着杂波数目的增加,IBSMC-PHD 的平均 OSPA 距离和 RMSE 急剧增大,而 SC-SMC-PHD 算法变化比较缓慢.这表明,本文所提出的 SC-SMC-PHD 算法在不同杂波环境下具有良好的稳定性.这是因为 SC-SMC-PHD 算法在重要性采样函数设计时引入了当前量测估计重要性采样函数.由于 SCKF 在非线形估计中良好的跟踪性能,SC-SMC-PHD 算法具有较小的 OSPA 距离.此外,由于本算法使用了基于 Gating 技术的量测合并算法,减小了杂波对估计精度的影响,因此在不同的杂波环境下具有良好的稳定性.

表 4 不同杂波数目下的跟踪性能对比

Table 4 Comparison of tracking performance under different clutter numbers

杂波数目	平均 OSPA			平均 RMSE		
	IBSMC-PHD/m	SC-SMC-PHD/m	提升/%	IBSMC-PHD	SC-SMC-PHD	提升/%
1	44.57	21.62	51.59	0.49	0.21	57.14
10	47.92	24.64	48.58	0.53	0.25	52.83
20	52.54	30.95	41.09	0.53	0.34	35.85
30	57.45	39.08	31.98	0.61	0.43	29.51
40	63.62	47.07	26.01	0.68	0.55	19.12

4 结 论

本文将 SCKF 和统计门限技术引入 SMC-PHD 的重要性采样函数设计,提出了一种 SC-SMC-PHD 算法.仿真实验表明:

1) 在固定/不同杂波数目下,和传统的 SMC-PHD 算法相比,本算法具有较高的准确性.当杂波数在 1~50 变化时,本算法的精度提升达到 30% 以上,但时间复杂度略高于传统 SMC-PHD 算法.

2) 在不同的门限和量测噪声水平下,本算法能保持良好的稳定性.

3) 本算法使用了传统的多项式重采样算法,仍存在粒子退化的问题,后续的研究将对其进行完善.

参考文献 (References)

[1] Bocquel M, Driessen H, Bagchi A. Multitarget tracking with interacting population-based MCMC-PF [C] // IEEE 15th International Conference on Information Fusion (FUSION). Piscataway, NJ: IEEE Press, 2012: 74-81.  
 [2] Ristic B. Particle filters for random set models [M]. New York: Springer, 2013.

[3] Mallick M, Vo B N, Kirubarajan T, et al. Introduction to the issue on multitarget tracking [J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(3): 373-375.  
 [4] Vermaak J, Godsill S J, Perez P. Monte carlo filtering for multi target tracking and data association [J]. IEEE Transactions on Aerospace and Electronic Systems, 2005, 41(1): 309-332.  
 [5] Wu J, Hu S, Wang Y. Adaptive multifeature visual tracking in a probability-hypothesis-density filtering framework [J]. Signal Processing, 2013, 93(11): 2915-2926.  
 [6] Uney M, Clark D E, Julier S J. Distributed fusion of PHD filters via exponential mixture densities [J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(3): 521-531.  
 [7] Mahler R P S. Multitarget Bayes filtering via first-order multitarget moments [J]. IEEE Transactions on Aerospace and Electronic Systems, 2003, 39(4): 1152-1178.  
 [8] 杨峰, 王永齐, 梁彦, 等. 基于概率假设密度滤波方法的多目标跟踪技术综述 [J]. 自动化学报, 2013, 39(11): 1944-1956. Yang F, Wang Y Q, Liang Y, et al. A survey of PHD filter based multi-target tracking [J]. Acta Automatica Sinica, 2013, 39(11): 1944-1956 (in Chinese).  
 [9] Vo B N, Singh S, Doucet A. Sequential Monte Carlo methods for multitarget filtering with random finite sets [J]. IEEE Transactions on Aerospace and Electronic Systems, 2005, 41(4): 1224-1245.  
 [10] Mahler R P S. Statistical multisource-multitarget information fusion [M]. Norwood: Artech House, Inc., 2007.

- [11] Ristic B, Clark D, Vo B N. Improved SMC implementation of the PHD filter [C] // 13th Conference on Information Fusion (FUSION). Piscataway, NJ: IEEE Press, 2010: 1-8.
- [12] Ristic B, Clark D, Vo B N, et al. Adaptive target birth intensity for PHD and CPHD filters [J]. IEEE Transactions on Aerospace and Electronic Systems, 2012, 48(2): 1656-1668.
- [13] Punithakumar K, Kirubarajan T, Sinha A. Multiple-model probability hypothesis density filter for tracking maneuvering targets [J]. IEEE Transactions on Aerospace and Electronic Systems, 2008, 44(1): 87-98.
- [14] Baser E, Efe M. A novel auxiliary particle PHD filter [C] // IEEE 15th International Conference on Information Fusion (FUSION). Piscataway, NJ: IEEE Press, 2012: 165-172.
- [15] Yoon J H, Kim D Y, Yoon K J. Efficient importance sampling function design for sequential Monte Carlo PHD filter [J]. Signal Processing, 2012, 92(9): 2315-2321.
- [16] Li T, Sun S, Sattar T P. High-speed sigma-gating SMC-PHD filter [J]. Signal Processing, 2013, 93(9): 2586-2593.
- [17] Arasaratnam I, Haykin S. Cubature Kalman filters [J]. IEEE Transactions on Automatic Control, 2009, 54(6): 1254-1269.
- [18] Schuhmacher D, Vo B T, Vo B N. A consistent metric for performance evaluation of multi-object filters [J]. IEEE Transactions on Signal Processing, 2008, 56(8): 3447-3457.
- [19] 汤琦, 黄建国, 杨旭东. 航迹起始算法及性能仿真 [J]. 系统仿真学报, 2007, 19(1): 149-152.  
Tang Q, Huang J G, Yang X D. Algorithm of track initiation and performance evaluation [J]. Journal of System Simulation, 2007, 19(1): 149-152 (in Chinese).

## SMC-PHD algorithm based on squared cubature particles

LIU Zhe<sup>1,2</sup>, WANG Zulin<sup>\*1</sup>, XU Mai<sup>1</sup>, LIU Jingxian<sup>1</sup>, YANG Lan<sup>1</sup>

(1. School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China;

2. School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

**Abstract:** Most of the conventional sequential Monte Carlo probability hypothesis density (SMC-PHD) approaches adopt the state transition density as importance sampling function. When targets are with nonlinear motions, such a selection makes few particles with large weights, leading to inaccurate estimation and particle divergence. To avoid such problems, a novel importance sampling function approximation approach with the squared cubature Kalman filter (SCKF) and statistical gating method was proposed. To design such an importance sampling function, the mean and covariance of importance sampling function were predicted at first. Then, the statistical gating method were utilized to extract observations associated with the importance sampling particle from the current observation set. Merging the extracted observations with corresponding weights, the mean and covariance of importance sampling function were updated. Using the designed importance sampling function, the intensity of particles can be predicted and updated, according to the conventional SMC-PHD method. At last, the states and number of multi-target can be approximated by the intensity of particles. The simulation results demonstrate that the proposed approach has the advantages of high accuracy and stable estimation in nonlinear target tracking.

**Key words:** sequential Monte Carlo; probability hypothesis density; importance sampling; squared cubature Kalman filter; statistical gating

http://bhxb.buaa.edu.cn      jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2014.0708

# 基于可编程控制网关的一体化网络体系结构

刘锋<sup>\*1,2</sup>, 王渝斐<sup>1</sup>

(1. 北京航空航天大学 电子信息工程学院, 北京 100191; 2. 地球空间信息技术协同创新中心, 武汉 430079)

**摘 要:** 当前 IP 网络和非 IP 网络长期共存, 协议体系和通信体制互不相同, 异构网络的互联互通困难. 针对这个问题, 首先从异构网络间连接的关键节点——网关出发, 通过将网关底层的数据转发和控制功能分离, 并将网关上层的管控功能和应用服务功能抽取出来, 提出了一种基于可编程控制网关的一体化网络体系结构. 该系统结构将网络分为一体化网络协同管控单元、子网间可编程控制网关节点和异构子网 3 层, 设计了基于典型异构网络的互联协议栈, 以集成不同的网络协议. 进一步在一体化网络构架的基础上, 给出并分析了异构网络的抽象与描述、基于可编程网关的异构子网智能适配技术、一体化网络的资源管理与协同优化理论等关键技术. 最后基于民用航空通信给出了一个异构互联示例, 表明一体化网络体系结构的可行性.

**关键词:** 一体化网络; 体系结构; 可编程控制网关; 异构互联; 网络协议; 智能适配

**中图分类号:** V221+.3; TB553

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1959-07

随着 Internet、移动通信和卫星通信技术的高速发展, 涌现出了大量不同类型的通信网络, 其中典型的网络有互联网、地面无线通信网络以及各种专用网络. 其中互联网和地面无线通信网络以 IP (Internet Protocol) 技术为基础逐步走向融合. 而非 IP 技术的专用网络, 如战术数据链网络、基于国际空间数据系统咨询委员会 (CCSDS) 的空间通信网络和航空电信网 (ATN) 等, 采用不同于 IP 的通信体制与协议体系, 为用户提供各种不同的业务支持和服务. 尤其是在专业应用领域方面, 针对各种特殊应用场景、不同目标用户、技术特点和特殊通信传输等需求和应用问题, 专用网络发挥了不可替代的巨大作用. 这样形成了当前 IP 网络与非 IP 网络 (如战术数据链、空间通信网络、航空电信网等) 共存的异构网络环境.

当前, IP 网络与非 IP 网络长期共存发展. 由于 IP 网络较大的时延和对用户服务质量支持有限等

问题, 在未来很长的一段时期内, 部分专用网络和协议, 如空间测控通信的 CCSDS 网络协议、实时可靠通信的数据链和空管领域的航空电信网络等, 因为应用的特殊性和不同技术需求难以 IP 化.

因此, 一体化网络技术的研究, 尤其是以 IP 技术为主的网络 (IP 网络) 和以非 IP 技术为主的网络 (非 IP 网络) 的一体化和融合技术的研究, 迫切需要借鉴吸收现有的网络架构理论与技术, 设计一个具有良好兼容性与可扩展性的未来网络体系架构<sup>[1]</sup>.

世界上许多国家都十分重视未来互联网体系结构和创新环境的研究与建设<sup>[2]</sup>. 美国在 2005 年先后提出了著名的 GENI (Global Environment for Network Innovations)<sup>[3]</sup> 和 FIND (Future Internet Design)<sup>[4]</sup> 计划, 主要研究目标是设计新的网络体系和服务, 使得网络服务更加普适化; 日本在 2006 年启动了新一代网络架构 AKARI<sup>[5]</sup> 项目; 欧盟在

收稿日期: 2014-11-17; 录用日期: 2014-12-26; 网络出版时间: 2015-01-28 16:52

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150128.1652.002.html

基金项目: 国家自然科学基金 (61231013, 61221061, 91438206); 中央高校基本科研业务费专项资金 (YWF-14-DXZY-027); 新世纪优秀人才支持计划 (NCET-09-0025); 空天宽带网络技术协同创新实验室基金

\* 通讯作者: 刘锋 (1970—), 男, 湖北麻城人, 教授, liuf@buaa.edu.cn, 主要研究方向为网络通信理论与技术、空天信息传输和复杂网络.

**引用格式:** 刘锋, 王渝斐. 基于可编程控制网关的一体化网络体系结构[J]. 北京航空航天大学学报, 2015, 41(10): 1959-1965.

Liu F, Wang Y F. Architecture of unified network based on programmable control gateway [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1959-1965 (in Chinese).

2008年开启了 FIRE (Future Internet Research and Experimentation)<sup>[6]</sup> 计划, 研究未来互联网。

同时, 专业应用机构和组织也在研究专用网络与 IP 网络的互联互通与信息共享<sup>[7]</sup>. 国际民航组织 (ICAO) 提出建立一个全球化无缝隙的互联网络——航空电信网, 制定出将 IP 网作为 ATN 子网的隧道方案, 即 ATN over IP 方案. 国际空间数据系统咨询委员会<sup>[8]</sup> 在 2010 年 4 月发布了 CCSDS 702.1-R-4 及相关文本, 建立起 CCSDS 空间链路上传递 IP\_PDU 的实践规范. 美军在 2001 年提出了基于 IP 的嵌入到联合战术无线电系统 (JTRS) 的战术目标瞄准网络技术 (TTNT)<sup>[9]</sup>. 但是, 以上计划和方案都是针对于单一应用网络展开研究, 其可扩展性和灵活性需要进一步深入研究。

本文在多年现代空管信息网络和空天地网络融合技术研究的基础上, 结合互联网的最新发展技术和应用现状, 提出一种基于可编程控制智能网关的一体化网络体系结构, 分析了其中的关键技术, 基于民用航空给出了一个异构网络的一体化互联示例。

## 1 基于网关的一体化网络体系

### 1.1 对象及定位

针对未来互联网、无线移动网络 (航空电信网)、空间网络 (国际空间数据系统 CCSDS) 等一体化组网和综合应用实际需求<sup>[10-11]</sup>, 设计一种能够兼容已有的 IP、CCSDS、ATN 和数据链等典型协议的一体化网络体系结构, 而且能够向前灵活演进和发展, 满足不同网络用户在一体化网络中的个性化需求和专业应用。

### 1.2 主要问题和解决思路

对于异构网络的互联互通问题, 传统的解决方案之一是采用网关进行处理. 但是传统网关主要面向具体应用, 以协议转换和数据转发为主要功能, 采用静态的、固定的转换方式, 缺乏灵活性和可扩展性<sup>[12-13]</sup>. 而且由于异构网络之间的通信方式、传输体制、协议体系等子网特征和能力的不同, 带来了异构子网间的速率 (带宽) 不匹配和服务质量差别 (误码率、时延) 等传输问题, 从而影响网络的实际效用, 同时导致网络传输效率下降, 这也成为一体化网络发展的瓶颈和难点。

本文的主要思路是将异构网络之间网关底层的数据转发与控制功能分离, 形成一个可编程控制的智能网关单元; 并且在上层将网关的管理控制功能提取出来, 形成一个可以全局规划与协同优化的一体化网络管控单元。

网关是异构子网互联互通的中间节点, 是异构网络之间连接的桥梁, 是一体化网络的关键节点, 也是一体化网络的管理服务和优化控制的关键执行环节. 同时由于异构网络间的网关节点数不多, 方便一体化网络的集中管理和控制。

### 1.3 一体化网络的体系结构

通过把网络的控制功能从网络的转发功能中解耦出来, 可以将异构网络之间协议和数据结构等转换的复杂逻辑简化剥离, 从网关连接点等设备中抽离出来, 从而实现通过编辑软件程序对硬件进行数据监控和管理。

针对一体化网络的架构方案, 以网关节点为主要研究对象. 将异构子网间的传输问题抽象、映射为网关间的传输问题. 即异构网络通过可编程网关控制节点实现互联互通和相互融合, 实现一体化。

通过赋予网关节点可编程性<sup>[14]</sup>, 即子网间可编程控制网关节点负责将数据转发和控制功能分离, 将网关的管理决策功能转移到由网关节点所形成的网络系统的管控单元, 将网络的控制功能从设备内部迁移到设备外部. 针对异构子网和专用通信协议复杂多样的问题, 可编程控制网关通过将网关的控制功能与转发功能分离, 来实现多种网络的智能适配, 其适配算法既可以事先设定, 也可以实时动态调整和远程控制, 成为一个可编程控制的智能网关. 这样就可以避免传统网关提前定制, 缺乏动态灵活性的问题. 可以针对子网网络状态的实时变化, 动态设置和调整网关参数, 进行子网间的动态适配, 提高网络的运行效率。

一体化网络的体系结构主要分为 3 层, 自上向下分别是协同管控单元、可编程控制网关和异构子网, 各层之间由通用接口 API 连接, 如图 1 所示。

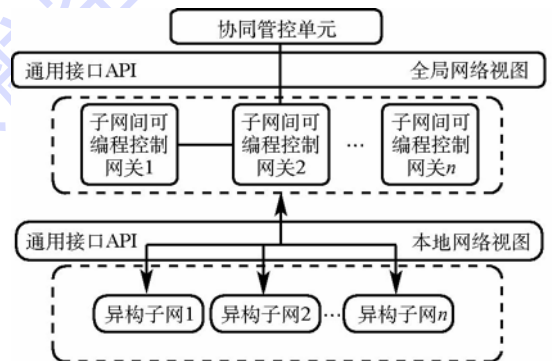


图 1 一体化网络的 3 层架构

Fig. 1 Framework of unified network with three layers

### 1.4 系统组成与功能模块

可编程控制网关节点: 可编程控制网关内部分为与底层设备之间的底层硬件接口模块、数据协议解析与分类模块、异构子网协议转换与地址

映射模块、子网间适配模块以及与管控单元的网关接口与操作规范模块. 可编程控制网关提供了多种数据转换、协议转换技术以及本地视图的 API 接口, 从而能够以多种能力支持新体系和新协议. 在可编程控制网关节点之上, 将数据的转发控制功能分离, 同时为协同管控单元提供了全局网络视图和统一的控制接口.

**协同管控单元:** 协同管控单元分为拓扑控制模块、路径计算模块和安全控制模块等. 根据可编程控制网关节点提供的全局抽象拓扑视图及子网性能, 通过子网需求分析与资源分配算法, 进行全局路径规划与流量优化. 基于拓扑感知, 实施预测和更新网络拓扑视图. 根据子网用户的动态资源请求, 通过指令数据库实现自动化调度.

由网关节点所形成的协同管控单元负责对整个异构网络的全局路由计算、流量规划、资源分配、拓扑控制与网络管理. 通过将网络的控制功能从传统的网关节点中转发功能中解耦出来, 使得网络的控制功能位置发生变化, 从网络设备内迁移到网络设备外. 那么网络管理者能够在一体化网络架构的基础上, 实现多种数据格式的统一控制和网络资源的统一管理.

**通用接口:** 可编程网关与异构子网间的接口

屏蔽了异构子网之间通信方式和协议体系之间的差异, 克服了传统网关提前定制而引起的通用性差和兼容性差等问题. 协同管控单元与可编程网关之间的接口, 为协同管控单元提供了全局网络视图, 可以实现异构网络的资源共享和资源整合. 同时, 为可编程控制网关提供了编程接口, 通过支持在线网络的动态配置和编程, 任何网关节点可以根据用户子网的类型和实时状态来进行任务状态的改变和适配<sup>[15]</sup>.

本文提出的一体化网络体系架构可以实现异构网络有效的互联互通与信息共享, 能够从面向应用的角度满足用户和业务需求的多样性, 实现多功能、多体制以及多服务的融合.

## 2 一体化网络的关键技术

一体化网络实现方案的总体思路为: 在异构子网内部进行子网自治, 避免改变现有的网络结构, 将异构子网之间传输问题映射为网关间的传输问题. 其实现方案如图 2 所示. 图中: AS (Autonomors System) 为自治系统; TCP (Transmission Control Protocol) 为传输控制协议; AFDX (Avionics Full-Duplex Switched Ethernet) 为新一代航空总线协议.

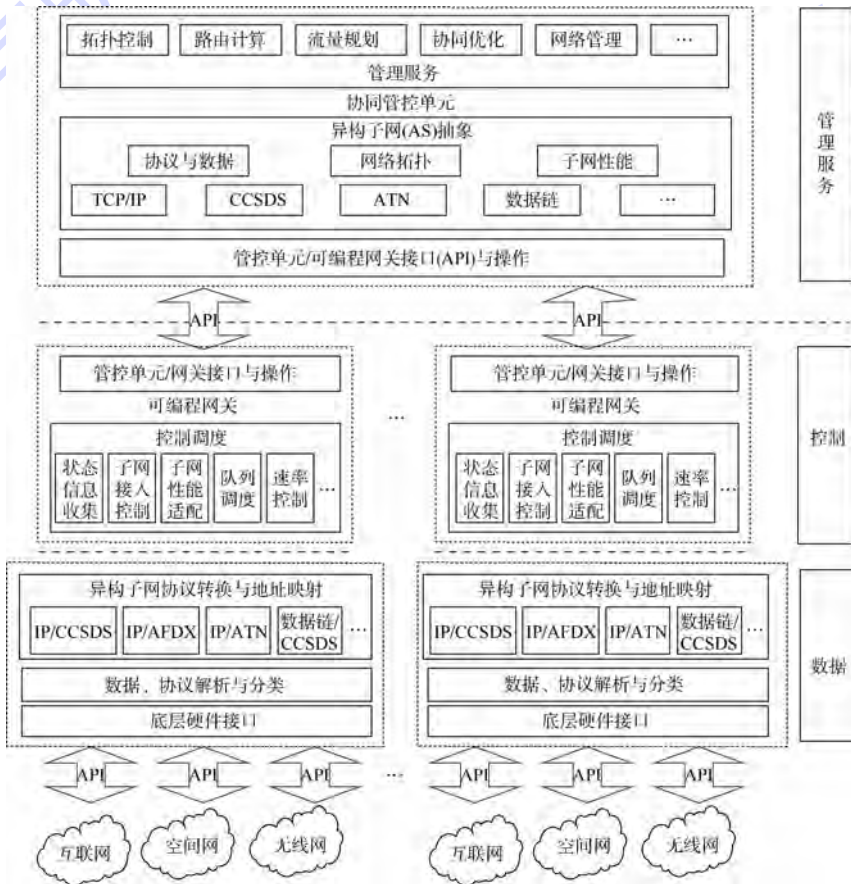


图 2 一体化网络的实现方案

Fig. 2 Implementation scheme of unified network

将一体化组网和综合应用变为网关节点之间的联通和协同组网应用,即通过网关协议转化、地址映射和智能适配实现异构网络的互联互通,通过网关节点的接入与优化控制实现异构网络的随遇接入,通过网关协同调度与优化控制实现异构网络的资源共享与资源整合。

## 2.1 抽象与描述

异构网络之间使用的不同协议体系和通信体制以及能力差异构成了网络互联互通的障碍并且增大了管理控制的难度,其中网络协议与子网地址的多样性增加了协议转换、地址映射和智能适配的难度,而且子网的移动性更是增加了统一编址和协同优化的难度.针对这些问题,将异构网络抽象化并统一描述,来消除网间协议差异影响,提高信息共享效能。

网络资源的统一描述,即将所有的网络资源都分配抽象的标识,使其拥有全局范围内的唯一的子网标识号(ID)和全局唯一标识号(GUID).GUID由子网类型ID+IP地址构成.这样就保证了该资源在一体化网络架构下能够被使用和访问.同时还需要对资源的功能进行形式化的统一描述,获得资源的可用性情况.一般来说,对于一体化网络中的节点资源和链路资源,可以通过资源发现机制进行定量描述。

协议与数据的统一描述,即可编程控制网关将流入其中的所有数据视为统一的流,忽略一体化网络中多种协议与数据结构共存的复杂情况.可编程控制网关能够解析数据源的子网类型、协议类型和数据结构.通过数据包头分析,与通配域进行匹配,最后执行相应的操作规则,包括直接转发、协议转换、链路建立和链路释放等.通过网关控制层和数据层分离,网络资源的管理转移到控制层进行统一调度,可编程控制网关节点不再关注路由决策和接入控制等,只负责数据的转发。

## 2.2 异构互联协议栈

图3是一个以局域网-异构网络-互联网为示例的异构网络互联协议栈.图中:FTP(File Transportation Protocol)为文件传输协议;HTTP(Hyper Text Transportation)为超文本传输协议.由该协议栈可以看出,异构互联的关键在于网络适配层的协议转换和地址映射,该地址即为每个对象全局范围内的GUID.异构子网在内部使用独立的链路层协议,在外部互联互通时,则通过子网间可编程控制网关的子网智能适配模块完成地址映射与协议转换.该协议栈可以有效整合当前典型的CCSDS、ATN异构子网和IP网络协议,实现这些异构子网的互联互通和信息共享.该协议栈采用模块化的设计思想,具有一定的典型性和普适性。

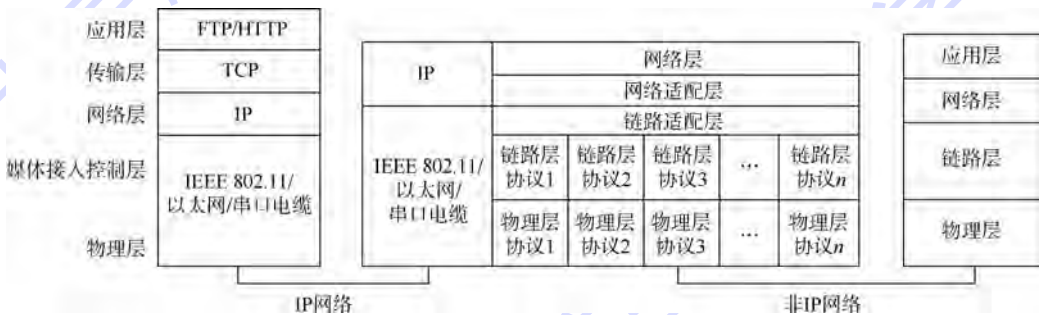


图3 一体化网络的协议栈示例

Fig. 3 Examples of protocol stack for unified network

## 2.3 异构子网间的智能适配技术

为了实现一体化网络资源的统一管理调度,子网间可编程控制网关必须消除因为异构子网间性能差异而造成的影响,实现网间速率与资源的匹配.针对上述问题,有效解决途径是赋予网关节点可编程性,通过支持在线网络动态配置和编程,任何网关节点可以根据用户子网的类型和实时状态来进行任务状态的改变和适配.异构子网的适配技术包括异构子网网关传输的速率匹配与控制、时隙和带宽资源的匹配.异构子网的适配技术是用户业务跨网传输性能的匹配和保障,保证了网关管理资源的有效性。

当异构子网进行通信时,网关节点利用子网的性能信息,通过队列调度算法,动态安排数据在网关内的进出,实现速率适配.将流入网关接口的分组,送入不同的转发接口,等待发送.IP网络和非IP网络间通信时,还要考虑带宽与时隙的对应关系.可编程控制网关通过带宽资源等效原则策略,决定了带宽与时隙之间的映射,实现带宽与时隙之间的转换,包括带宽/时隙的可用性、带宽分配和时隙分配。

## 2.4 一体化网络的资源管理和协同优化

协同管控单元包括流量规划模块、路由计算模块、拓扑控制模块和协同优化模块等.网络管理

者可以通过协同管控单元为子网提供区分性服务,并能够统一调度底层资源,动态调整资源的分配方案.一体化网络的资源管理与协同优化负责为实现全局统一规划,避免局部优化,提高全网资源的整合效能.

协同管控单元通过可编程控制网关的抽象接口对多元抽象数据操作和异构子网开放网络设备进行管理.同时能够为新体系和新协议的随时加入提供统一的控制平面接口,降低一体化网络互联互通的复杂性.

根据网络资源的 GUID,接入可编程控制网关的子网资源被统一分配抽象标识,可以保证资源在一体化框架下能够被使用和访问.网络资源包括链路资源和节点资源.根据资源探测机制,获得底层资源的本地视图.经过资源描述与资源发现机制,可以获得全局的资源视图,再利用协同管控单元的资源分配模块,实现路径规划与全局优化.

在一体化环境下,底层资源的位置、服务提供能力、负载和网络连通性实时发生变化.因此对于子网的加入和退出,需要实时预测和更新网络拓扑视图,并利用接入控制机制,提高网络的鲁棒性.协同管控单元同时能够根据子网用户的动态资源请求,实现资源的重定向.

### 2.5 统一接口标准

为了屏蔽异构网络协议数据、通信体制以及底层网络资源之间的差异,需要为异构子网与可编程控制网关之间的通用接口 API 制定标准的接

口规范,满足支持 Ethernet、AFDX、ECL 和 LVDS 等多种通用接口,接口可以支持从 Kbps 到 Gbps 级的转发速率.

### 3 一体化网络在民航的互联示例

基于提出的一体化网络体系构架,这里针对民用航空通信应用给出了如图 4 所示的异构网络一体化互联应用示例.即航空用户通过机载可编程控制网关,使用 CCSDS 或者空地数据链,再经过地面可编程控制网关与地面互联网通信.

图 4 描述的是一个以太网(Ethernet)和一个广域网(Internet)通过卫星链路或者空地通信链路进行互联互通的网络系统.图 4 所示的系统主要包括机载网络、空间链路/子网和地面子网 3 部分.机载以太网和地面 Internet 支持语音、视频和数据等多种业务.其中卫星链路采用 CCSDS 的 AOS(Advanced Orbiting Systems)协议,空地通信采用 L-DACS(L-band Digital Aeronautical Communications System)数据链协议,AOS 和 L-DACS 数据链都是典型的非 IP 网络协议.该系统可以模拟适用于标准以太网通过数据链、CCSDS 卫星链路或者其他传输网络与 Internet 互联互通的通用网络体系架构和可编程控制网关功能,也就是 IP 网络和非 IP 网络的互联.

图 5 给出了基于可编程网关的网络互联系统.其中网关的主要功能模块有数据转发模块、协

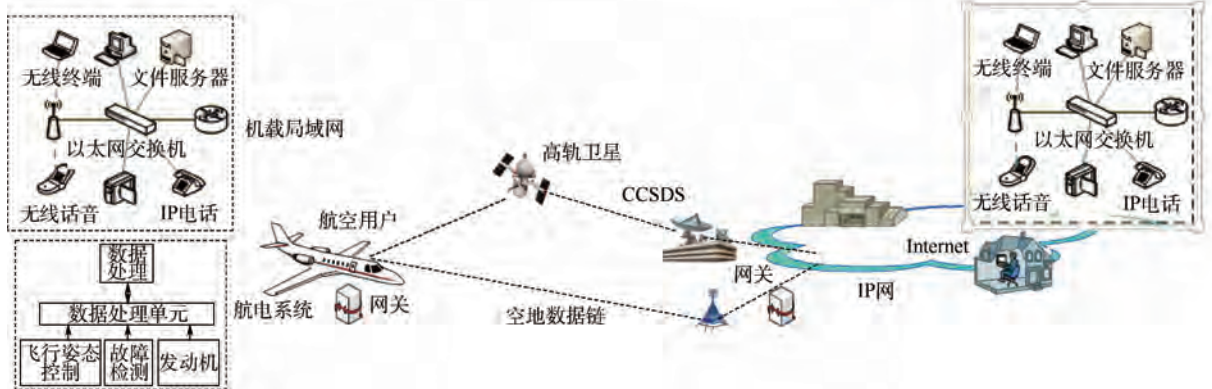


图 4 异构网络的一体化互联应用示例

Fig. 4 Application case of unified interconnection network for heterogeneous networks

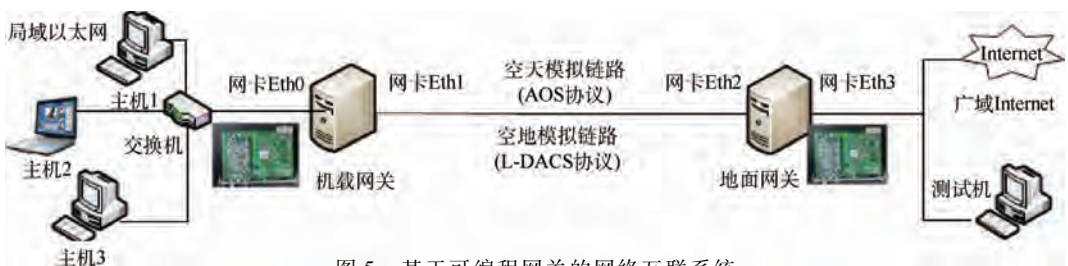


图 5 基于可编程网关的网络互联系统

Fig. 5 Interconnection networks based on programmable gateway

议转换模块、编址和地址映射模块、速率匹配模块、资源调度与可编程管控模块。网关的协议转换模块主要包括 IP over AOS 和 IP over L-DACS, 网关的控制管理功能可以通过编程实现。

可编程控制网关可以在基于 Linux 的软件系统上进行软件实现, 或者在嵌入式平台 PowerPC 和 FPGA 的硬件板卡上进行硬件实现。

在基于 Linux 的软件网关系统中, TCP 数据包的传输效率可以达到 80% 以上 (84.7 Mbps/百兆网卡), 而无网关的直接传输效率大约为 95% (96 Mbps/百兆网卡), 表明 TCP over CCSDS 的软件实现效率大约为 88%, 即性能损失在 12% 左右。而 UDP (User Datagram Protocol) 的传输性能大约为 20% (20 Mbps/百兆网卡), 无网关测试速率为 32% (32 Mbps/百兆网卡), 说明 UDP over CCSDS 的效率大约为 62%, 即性能损失在 38% 左右。千兆网卡的软件实现结果类似。这些数据表明了该体系架构的可行性和可编程控制网关的可用性。在基于嵌入式平台 PowerPC 和 FPGA 的硬件板卡上, 其系统性能得到进一步的提升, 接近硬件系统的数据转发速率。当传输速率达到数 Gbps 时, 可以进一步通过提高硬件芯片能力来提高异构网络传输的性能。

## 4 结 论

1) 针对 IP 网络和非 IP 网络长期共存发展、互联互通与信息共享困难的现状, 提出一种针对典型网络的基于可编程控制网关的一体化网络的架构方法与总体框架。通过将网关的数据转发和控制功能分离, 从而在网络上层具有更大的灵活性和集中管控能力, 使得该体系结构相较于现有的网络体系结构更具有灵活性和可扩展性。

2) 分析并提出了实现一体化网络的关键技术。通过数据、协议和资源等的抽象与统一描述解决异构网络多种协议数据共存和地址多样的问题; 通过智能适配技术解决网间性能差异、资源与速率不匹配的问题; 通过一体化网络的资源管理和协同优化解决局部优化和资源利用率不足的问题。

3) 给出了异构网络一体化互联示例。通过基于可编程网关的网关互联系统的仿真实验测试了 UDP 和 TCP 的端口速率, 验证了体系架构可行性和可编程控制网关的可用性及有效性。

4) 提出的一体化网络架构具有良好的管控能力、可扩展性和融合性, 其思想方法可以广泛应

用于陆地、空中、太空和海洋等各类异构网络用户的信息共享、资源整合及互联互通, 为中国未来一体化网络的建设和发展提供了重要的理论技术基础和工程指导。

## 参考文献 (References)

- [1] 张军. 面向未来的空天地一体化网络技术 [J]. 国际航空, 2008 (9): 36-39.  
Zhang J. Future space-ground interconnection network [J]. International Aviation, 2008 (9): 36-39 (in Chinese).
- [2] Greenberg A, Hjalmytsson G, Maltz D A, et al. A clean slate 4D approach to network control and management [J]. SIGCOMM Computer Communication Review, 2005, 35 (5): 41-54.
- [3] Global environment for network innovations (GENI) project [EB/OL]. (2012-11-15) [2015-01-04]. <http://www.geni.net/>.
- [4] NSF NeTS FIND initiative [EB/OL]. (2012-09-18) [2015-01-04]. <http://www.nets-find.net>.
- [5] AKARI architecture design project [EB/OL]. (2012-12-08) [2015-01-04]. <http://akari-project.nict.go.jp/eng/index2.htm>.
- [6] Greene K. Software-defined networking [J]. MIT Technology Review, 2009, 38 (8): 66-69.
- [7] Chowdhury N M M K, Boutaba R. A survey of network virtualization [J]. Computer Networks, 2010, 54 (8): 862-876.
- [8] Consultative committee for space data systems (CCSDS) [EB/OL]. (2015-01-01) [2015-01-04]. <http://public.ccsds.org/default.aspx>.
- [9] Tactical targeting network technology and connectivity [EB/OL]. (2009-11-09) [2015-01-04]. <http://www.sldinfo.com/wp-content/uploads/2009/11/TTNTWhitePaper.pdf>.
- [10] NSF future internet architecture project [EB/OL]. (2012-09-18) [2015-01-04]. <http://www.nets-fia.net/>.
- [11] Bavier A, Feamster N, Huang M. In VINI veritas: Realistic and controlled network experimentation [C] // ACM Annual Conference of the Special Interest Group on Data Communication. Pisa: ACM, 2006: 3-14.
- [12] 沈荣骏. 我国天地一体化航天互联网构想 [J]. 中国工程科学, 2006 (10): 19-30.  
Shen R J. Some thoughts of chinese integrated space-ground network system [J]. Engineering Science, 2006 (10): 19-30 (in Chinese).
- [13] Liu Z J, Li Y, Cui B, et al. Grain flow: Enable testing for future internet architectures by per-bit customization [J]. Computer Networks, 2014, 69: 121-132.
- [14] Shanmugalingam S. Programmable mobile core network [C] // Proceedings of IEEE Symposium on Computers and Communication. Piscataway, NJ: IEEE Press, 2014: 1-7.
- [15] Anwere M B, Feamster N. Building a fast virtualized data plane with programmable hardware [C] // VISA09 Proceedings of the 1st ACM Workshop on Virtualized Infrastructure Systems and Architectures. New York: ACM, 2009: 1-8.

# Architecture of unified network based on programmable control gateway

LIU Feng<sup>\*1,2</sup>, WANG Yufei<sup>1</sup>

(1. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China;

2. Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China)

**Abstract:** Due to the long-term coexistence of Internet Protocol (IP) and non-IP networks, there exist interconnection and interworking difficulties in heterogeneous networks for the difference of network architectures and protocols. Focusing on the gateway, the key node of heterogeneous networks interconnection, a unified network architecture based on programmable gateway is proposed to solve this problem, which separated the control plane from the data forwarding plane and extracted the combination of the management and control functionality and application service. The unified network architecture divides the network into three layers, which were unified network collaborative management and control unit, programmable gateway, and heterogeneous subnet. The interconnection protocol stack for the typical heterogeneous networks is designed to integrate different protocols. Then the key techniques and theory for the programmable gateway are given and analyzed based on the unified network architecture, including the abstraction and description of heterogeneous networks, intelligent adaptation technology based on programmable gateway, resource management and optimization of unified network, etc. Finally, a typical application case of interconnection networks in civil aviation is given as an illustration to demonstrate the proposed network architecture.

**Key words:** unified network; network architecture; programmable control gateway; heterogeneous networks; network protocol; intelligent adaptation

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0010



# 构建广义立方体感知网络安全态势

文志诚<sup>1,2</sup>, 陈志刚<sup>\*1</sup>

(1. 中南大学 信息科学与工程学院, 长沙 410083; 2. 湖南工业大学 计算机与通信学院, 株洲 412007)

**摘 要:** 针对大多方法感知范围局限、信息来源单一、空间时间复杂性高及准确性偏差较大等问题,提出了分层感知模型与构建广义立方体感知网络安全态势的方法.将监测到的连续型态势因子数据经“ $3\sigma$ 法则”离散化预处理后,聚合在所构建的广义立方体格中,纵向上融合成组件的安全态势,横向上对组件安全态势采用统计的方法融合成网络的安全态势,为增强网络安全性提供可靠的参照依据.利用网络实例数据对所提出的网络安全态势感知模型和算法进行验证,表明了该方法的正确性.

**关键词:** 网络安全; 态势感知; 网络管理; 信息融合; 广义立方体

**中图分类号:** TP311

**文献标识码:** A

**文章编号:** 1001-5965(2015)10-1966-09

网络安全态势感知是态势感知(Situation Awareness, SA)技术在网络安全方面的应用.1999年,网络态势感知(Cyberspace Situational Awareness, CSA)概念最初由Bass提出,他指出“基于融合的网络态势感知将成为未来网络管理的发展方向”<sup>[1]</sup>.网络态势是指由各种多源异构网络设备目前运行状况、网络行为以及各用户之间交互行为等要素所构成的整个网络状况与趋势.

网络安全态势感知具有网络态势感知的基本性质,在大规模网络中,对能使安全态势发生变化的各要素进行获取、理解、显示以及预测未来发展趋势,并综合考虑各方面的安全因素,从整体动态上把握网络安全状况,一旦完成了安全态势感知,决策几乎可以根据态势自动生成.在网络安全系统中,网络安全态势感知可以对网络做到“监、管、控”于一体,简化各种应用处理环节,提升网络安全防护级别.

安全态势感知从全局动态上把握网络安全状

态,主要运用信息融合与实时处理技术,包括防火墙、入侵检测和病毒等日志,以及非法外联、网络扫描、各类设备运行状态及报警等在内的多源异构信息.需要将来自各类异构信息源的网络安全事件数据关联分析,分析发生在不同时空和各层次上关联的安全事件,挖掘时空上分散的协同多步攻击事件,识别出真正的风险与威胁.

网络安全态势感知将态势感知成熟的理论应用于网络管理之中,从整体动态上反映网络目前的安全状况,但目前还没有明确的定义.

其中,网络安全态势感知模型与框架是态势感知的基础与前提条件.目前已经提出了许多种数据融合模型<sup>[2-3]</sup>,具有影响力的当属JDL(the Join Director of Laboratories)数据融合模型<sup>[4]</sup>.文献[5-6]在分析现有网络管理存在的不足以及发展需求的基础上,提出了一个网络态势感知研究框架,指出了研究重点以及存在的问题,着重评价了每种评估方法的基本思路,评估过程和优缺点,并进行了对比分析.

收稿日期: 2015-01-05; 录用日期: 2015-04-10; 网络出版时间: 2015-05-21 15:05

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150521.1505.005.html

基金项目: 国家自然科学基金(61073186,61073104,60903058); 中南大学博士后基金

作者简介: 文志诚(1972—),男,湖南东安人,副教授,zewen@mail.shu.edu.cn

\* 通讯作者: 陈志刚(1964—),男,湖南益阳人,教授,czg@mail.csu.edu.cn,主要研究方向为网络计算与分布式处理.

引用格式: 文志诚, 陈志刚. 构建广义立方体感知网络安全态势[J]. 北京航空航天大学学报, 2015, 41(10): 1966-1974.

Wen Z C, Chen Z G. Constructing general cube to be aware of network security situation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(10): 1966-1974 (in Chinese).

感知方法是重点,目前提出了许多种关于网络安全态势的评估方法,主要有证据理论<sup>[7]</sup>、灰色系统理论、粗糙集理论<sup>[8]</sup>、决策表<sup>[9]</sup>、神经网络、模糊逻辑<sup>[10]</sup>和专家系统等。

文献[11]提出了层次化威胁评估量化方法,利用IDS(Intrusion Detection Systems)的报警信息和性能指标并结合漏洞信息等,对主机、服务和网络进行层次化定量评估,得到直观的网络安全态势图。文献[12]通过对IDS与IPS(Intrusion Prevention System)等系统的报警信息进行融合处理,生成网络态势,文献[13]提出了建立一个能应用于其他领域的下一代网络态势感知系统,目的是提高决策者选择行为的过程,确保网络管理人员的真正需求。文献[14]引入D-S(Dempster-Shafer)证据理论,融合多源异构信息,计算经过态势因子融合与节点态势融合的安全态势。文献[15]提出一种基于马尔可夫博弈分析的网络安全态势感知方法,提供的加固方案可有效抑制威胁的扩散。文献[16]给出了基于神经网络的网络安全态势感知方法。

针对目前安全态势感知方法中的感知范围局限、信息来源单一、时空开销大及可信度不高等问题,本文将综合考虑影响网络安全态势的各方面因素,通过构建广义立方体,快速高效地融合异构多源多层次信息,利用全方位感知网络安全态势的方法,提供宏观的安全状况视图,对整个网络具有理解与控制力。

## 1 安全态势感知模型

### 1.1 整体感知模型

网络系统结构中存在着大量的主机、服务器、路由器、防火墙和IDS等网络组件,网络安全态势感知整体感知过程如下:通过单个组件上的传感器实时监测系统实时的运行状况,采集大量原始数据,由组件安全态势感知模块经信息融合成本地的安全态势;再由 $N$ 个组件安全态势向上融合成整个网络安全态势。网络安全态势预测与决策支持以安全态势感知为基础,整体网络安全态势感知模型如图1所示。本文主要研究网络安全态势感知。为了区分组件和网络安全态势,加相应的下标 $c$ 和 $n$ 以示区别,表示组件和网络。

### 1.2 多源多层次信息融合分层感知框架

定义网络安全态势由网络基础运行性(runnability)、网络脆弱性(vulnerability)和网络威胁性(threat)三维组成,从3个不同的维度以直观的

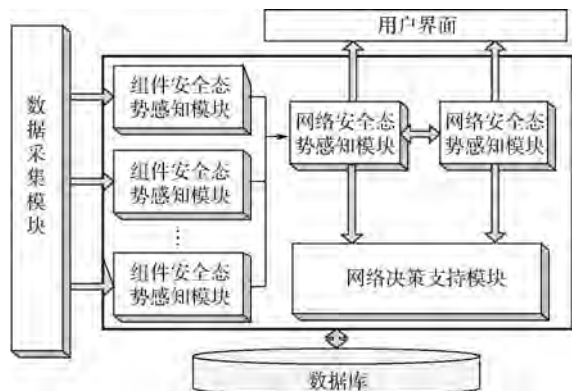


图1 整体网络安全态势感知模型

Fig. 1 Overall network security situational awareness model

形式向用户展现出网络整体的当前安全态势  $SA = (runnability, vulnerability, threat)$ 。对于每个维度都有组件和网络之分,如基础运行性有组件基础运行性和网络基础运行性,而网络基础运行性则由 $N$ 个组件基础运行性融合而成,其他两个维度类似。

多源多层次信息融合分层感知框架如图2所示。由观测指标也称作态势因子向上首先融合成组件安全态势,再向上融合成网络的安全态势。根据实际情况,态势因子可分为3大类:与运行信息相关的态势因子、与配置信息相关的态势因子和与IDS系统日志库中相关的态势因子。

与运行信息相关的指标有CPU利用率、占用内存、子网流量变化率、子网数据流总量及子网内不同大小数据包的分布等,构成组件的基础运行性;与配置信息相关的指标有漏洞、系统配置、防护软件是否安装、关键设备漏洞数及级别、子网安全设备数、子网各关键设备端口的总量等,构成组件的脆弱性;与IDS系统日志库中相关的指标有DDOS(Distributed Denial of Service)、蠕虫攻击、木马和普通病毒、子网带宽使用率、子网数据流入量和子网流入量增长率等,构成组件的威胁性。

基础运行性,是指目前组件运行状况,动态描述组件目前运行情况,由工作性能和服务性能构成,是否能正常运行和提供服务,称为组件安全态势的外在表现性。

脆弱性是可以被组件威胁利用的系统薄弱环节。单个脆弱性可用  $vulnerability = (idv, namev, typev, ida, idt, pv, \xi)$  来描述。其中, $idv$ 为系统脆弱性标识; $namev$ 为脆弱性名称; $typev$ 为脆弱性类型; $ida$ 为脆弱性所在组件的标识; $idt$ 为利用该脆弱性威胁的标识; $pv$ 为脆弱性能被利用的可能性; $\xi$ 为脆弱性具有的影响。

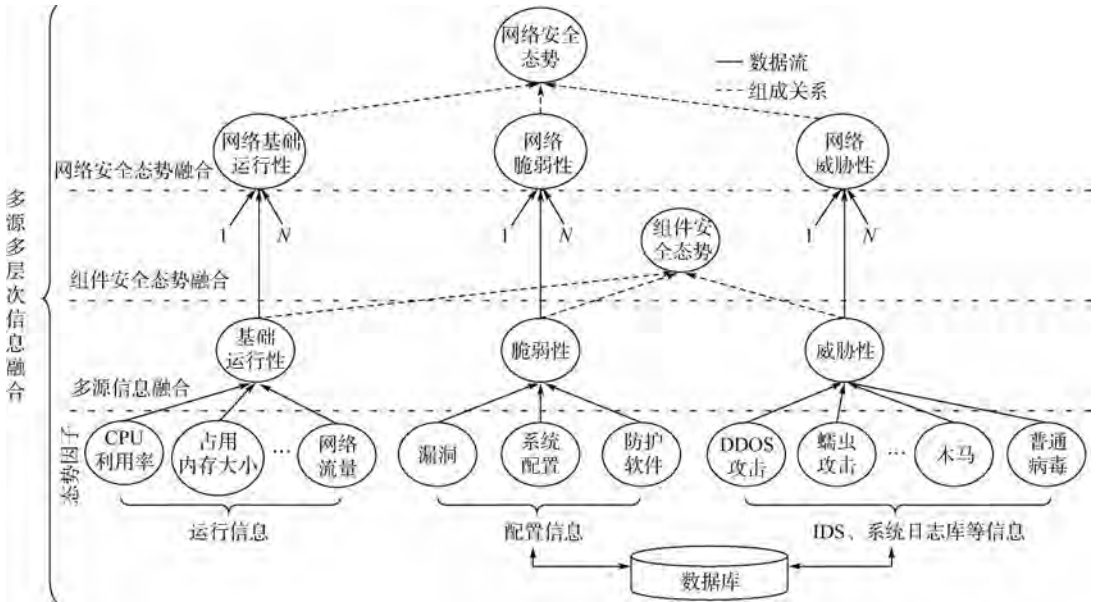


图2 多源多层次信息融合分层感知框架

Fig. 2 Multi-source and multi-level information fusion layered awareness framework

威胁性是能对组件造成损害的因素. 单个威胁可用  $threat = (id_t', namet, typet, ida', idv', pt)$  来描述. 其中,  $id_t'$  为威胁标识;  $namet$  为威胁名称;  $typet$  为威胁类型;  $ida'$  为威胁所在组件的标识;  $idv'$  为威胁利用的脆弱性的标识;  $pt$  为威胁将发生的概率.

网络的基础运行性、脆弱性和威胁性与组件的相应维度意义相近. 安全态势感知将从纵向和横向两个不同的方向进行信息融合, 所谓纵向信息融合是指组件态势因子的多源信息融合: 由运行信息的态势因子融合成基础运行性, 由配置信息的态势因子融合成脆弱性, 由 IDS 系统日志的态势因子融合成威胁性, 如图 2 所示; 横向信息融合是指网络基础运行性由各组件的基础运行性融合而成, 网络脆弱性由各组件的脆弱性融合而成, 网络威胁性由各组件的威胁性融合而成.

在融合方式上, 从态势因子融合成组件的 3 个维度是最重要的一步, 将使用广义立方体模型及相关算法; 从组件 3 个维度分别融合成网络的 3 个维度, 将使用基于统计的算法.

## 2 安全态势感知前提与基础

安全态势感知需要解决以下 4 个方面问题.

### 2.1 态势因子的遴选

有必要遴选出一些具有代表性的指标, 剔除与安全态势感知不相关的、冗余的指标. 判断观测指标  $x_i$  和  $x_j$  是否存在冗余, 可通过计算其相关系数  $\rho$  来判定.

计算观测指标  $x_i$  和  $x_j$  的相关系数:

$$\rho_{x_i x_j} = \text{Cov}(x_i, x_j) / \sqrt{D(x_i)D(x_j)}$$

式中:  $\text{Cov}(x_i, x_j)$  为  $x_i$  和  $x_j$  的协方差;  $D(x_i)$  和  $D(x_j)$  分别为  $x_i$  和  $x_j$  的方差. 有

$$\text{Cov}(x_i, x_j) = E[(x_i - E(x_i))(x_j - E(x_j))] = E(x_i x_j) - E(x_i)E(x_j)$$

式中:  $E(x_i)$  和  $E(x_j)$  分别为  $x_i$  和  $x_j$  的期望. 给定一个实数  $\varepsilon (0 < \varepsilon < 1)$ , 若相关系数  $|\rho_{x_i x_j}| > \varepsilon$ , 可认为  $x_i$  和  $x_j$  非常相关, 只保留一个起主要作用的观测指标, 即态势因子.

### 2.2 态势因子值离散化

态势因子可取离散型和连续型两种观测值, 为了便于原始数据在广义立方体聚合及统计算法中的应用, 把连续型取值离散化, 按实际意义, 可取“高、中高、中、中低、低”或“0、1、2、3、4”5 个等级值. 若态势因子本来就是离散型取值, 则不需离散化.

**引理 1** 设连续型随机变量  $X \sim N(\mu, \sigma^2)$ , 则  $Z = (X - \mu) / \sigma \sim N(0, 1)$ . 其中,  $\mu$  为随机变量  $X$  的数学期望;  $\sigma^2$  为方差.

可把随机变量  $X$  观测值划分为 5 个互不相交的区域  $SS_i: (-\infty, \mu - 3\sigma) \cup (\mu + 3\sigma, +\infty)$ 、 $(\mu - 3\sigma, \mu - 2.5\sigma) \cup (\mu + 2.5\sigma, \mu + 3\sigma)$ 、 $(\mu - 2.5\sigma, \mu - 2\sigma) \cup (\mu + 2\sigma, \mu + 2.5\sigma)$ 、 $(\mu - 2\sigma, \mu - \sigma) \cup (\mu + \sigma, \mu + 2\sigma)$  和  $[\mu - \sigma, \mu + \sigma]$ .

经计算, 5 个区域  $SS_i (i = 0, 1, 2, 3, 4)$  对应的概率  $PS_i$  分别为 0.26%、0.98%、3.32%、27.18% 和 68.26%. 事实上, 组件或网络在绝大多数情况下运行是正常的, 不正常的情况非常少, 因此这 5 个层次取值符合实际网络运行的情况, 不正常的

情况一般会落在  $3\sigma$  区域之外,  $\mu$  和  $\sigma$  都是常量.

这就是本文所提出的“ $3\sigma$  法则”, 基于大样本观测数据. 利用“ $3\sigma$  法则”指导构建广义立方体融合异构多源多层次信息以及在融合算法的区间概率分级上, 具有严格的理论基础. 离散化态势因子  $x_i$  的一般步骤如下:

- 1) 采集  $x_i$  的  $n$  个大样本历史数据, 计算其平均值  $\bar{x}_i$ , 代替总体  $x_i$  的数学期望  $E(x_i) = \mu = \bar{x}_i$ .
- 2) 同理, 计算  $x_i$  的方差  $S_i^2$ , 代替总体  $x_i$  的方差  $D(x_i) = \sigma^2 = S_i^2$ .
- 3) 按照上述方法把  $x_i$  划分为 5 个区间  $SS_i$ .
- 4) 当获得  $x_i$  的一个具体值时, 根据落在哪个区间  $SS_i$  来取相应的离散值  $i$ .

### 2.3 数据库的构建

对于第 1 类基础运行性态势因子观测数据的获取与离散化比较容易, 通过监测软件可直接采集并按“ $3\sigma$  法则”离散化为五等. 而后面两大类的态势因子数据采集及离散化稍微复杂些: 首先构建一个数据库, 存贮配置信息和 IDS 日志库信息相关联态势因子的历史信息, 也即把相关的信息按脆弱性 (vulnerability) 和威胁性 (threat) 格式以记录形式存入, 如图 1 所示.

数据库构建完毕, 起两个作用: 一是为了离散化连续型态势因子; 二是当实时监测到数据如漏洞  $x$  时, 通过查找其数据库中的记录, 获得  $\xi$  值, 对照已划分好的 5 个区间  $SS_i$ , 便于获得相应的离散值  $i$ .

### 2.4 信息融合方法

在纵向上, 由态势因子向上融合成组件的 3 个维度, 将采用广义立方体作为信息融合的模式以及关联的融合算法; 在横向上, 由各组件的 3 个维度融合成网络的 3 个相应的维度, 将采用基于统计的算法.

## 3 安全态势感知方法

### 3.1 构建广义立方体

立方体一般是由三维构成, 而广义立方体是由  $m$  维构成. 数据立方体是一种在立方体格上的聚合数据, 通过数据的聚合向上融合成某方面的特征, 常用支持多维多层次数据分析的模型. 通过该数据模型可从不同的视角考察所研究数据集的特性. 在网络安全态势感知中, 各个态势因子构成不同的维度, 各个离散值构成不同的抽象层次. 对各个维度、层次上的数据分析, 从不同角度反映了当前网络的运行状况. 下面给出数据立方体模型

的相关定义.

**定义 1** 给定一个数据集  $S$ 、一组维度  $A = \{A_1, A_2, \dots, A_m \mid m > 0\}$  和各个维度的层次集合  $H = \{H_1, H_2, \dots, H_L \mid L > 0\}$  ( $H_i$  表示维度  $A_i$  层次集合,  $H_i^j$  表示维度  $A_i$  的第  $j$  个层次). 在  $H$  的分量  $\{H_1, H_2, \dots, H_L\}$  中各取一个层次进行组合, 所有这样的组合  $\{H_1^1, H_2^2, \dots, H_m^m\}$  对数据集  $S$  执行分组合操作, 得到的一组立方体单元格将构成一个数据立方体  $(A, H)$ .

**定义 2** 一个立方体单元格可表示成二元组  $(A_H, M)$ , 其中:  $A_H = \{H_1^1, H_2^2, \dots, H_m^m\}$ ,  $H_i^j \in H_i$  为在维度  $A_i$  上选定的一个具体层次  $j$ ;  $M$  为对  $A_H$  数据集  $S$  经“ $3\sigma$  法则”离散化后进行聚合操作下, 落在本单元格中的个数. 立方体由立方体格组成:  $(A, H) = (A_H, M)$ .

例如, 取  $a$  为 CPU 利用率,  $b$  为占用内存大小,  $c$  为网络流量, 数据集  $S$  是包含此 3 个观测指标的样本, 经离散化五等后, 则一个三维广义立方体格如图 3 所示. 经过聚合操作后, 每个立方体格上具有相应的值  $((a_i, b_j, c_k), M_{ijk})$ , 其中  $M_{ijk}$  为数据集  $S$  经离散化处理后落在立方格  $(a_i, b_j, c_k)$  中的个数.

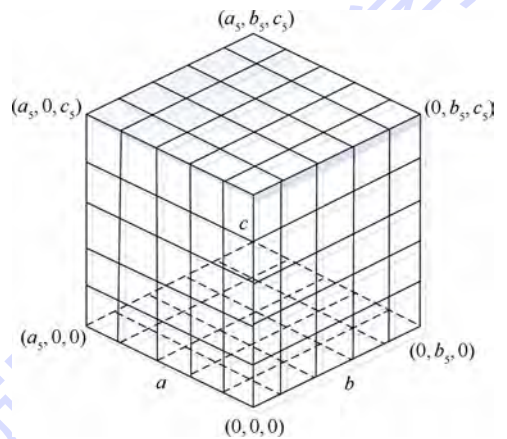


图 3 三维广义立方体格

Fig. 3 3D generalized cubic physique

本文至少需要建立: 基础运行性、脆弱性和威胁性 3 个广义立方体. 若广义立方体维度过高, 在信息融合上造成时间和空间复杂度过高, 可把态势因子再细分为几个小组, 每个小组构造一个广义立方体, 几个小立方体向上再合并成一个大立方体, 采取分而治之降低维度.

对同一类或一组的  $m$  个态势因子  $x_i$ , 根据历史运行信息, 采集  $n$  个大样本数据 (每个数据为  $m$  元组, 构成一个样本) 建立一个广义立方体  $(A, H)$ . 例如对于基础运行性广义立方体, 每个态势

因子  $x_i$  作为一个维度,每个维度若取连续型则离散化 5 个层次:

$$\hat{\mu}_i = \bar{x}_i = \frac{1}{n}(x_{i1} + x_{i2} + \dots + x_{in})$$

$$\hat{\sigma}_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}$$

在安全态势感知过程中,采集到某个态势因子实时原始数据时,与此因子 5 个取值范围  $SS_i$  相比较,可得到它相应的抽象层次  $i$ . 当本组所有的态势因子抽象层次确定后,说明这组态势因子  $x_i (i=1, 2, \dots, m)$  此时的原始数据  $S(m$  元组)就会聚合在此立方格中.

### 3.2 组件安全态势感知

组件安全态势感知是网络安全态势感知的基础,首先纵向上对组件进行安全态势感知,再横向上对网络进行安全态势感知. 由于安全态势感知具有时段性,即在某间隔时间段感知安全态势,所有采集的数据只能在此时间段内有效.

对于一组态势因子  $x_1, x_2, \dots, x_m$ , 在同一时间段内采集到  $n$  个  $m$  元组原始数据  $S$  (整体上看来是  $n$  个数据,  $n$  足够大), 每个分量  $x_i$  按照“ $3\sigma$  法则”后,则可构成  $n$  个  $m$  元组离散化后的数据. 具有 5 个层次的  $m$  个维度广义立方体,可达到  $5^m$  个立方格,则这  $n$  个  $m$  元组(每分量只取 0, 1, 2, 3, 4) 数据  $S$  分布在这  $5^m$  个立方格中,所有立方格中数据的个数之和为  $n$ , 即有

$$(A_n, M) = ((H_1^1, H_2^2, \dots, H_m^m), M_{j_1 j_2 \dots j_m})$$

$$\sum_{j_1=0}^4 \sum_{j_2=0}^4 \dots \sum_{j_m=0}^4 M_{j_1 j_2 \dots j_m} = n$$

式中:  $M_{j_1 j_2 \dots j_m}$  取正整数,立方体格值  $M_{j_1 j_2 \dots j_m}$  表示这  $n$  个  $m$  元组原始数据  $S$  落在本格中的个数.

**定义 3** 设  $M_i (i=0, 1, 2, 3, 4)$  表示立方体格值  $M_{j_1 j_2 \dots j_m}$  边际之和, 即当固定  $m-1$  个下标取值为  $i$  时, 对另一个下标  $j$  求和, 下标  $j$  取值从  $i$  到 4,  $j$  可处于任何  $m$  个位置中的一个. 直观上说,  $M_i$  表示广义立方体中取  $i$  等时原始数据的个数, 即有

$$M_i = \sum_{j=i}^4 M_{i, i, \dots, i, j, i, \dots, i} - (m-1)M_{i, i, \dots, i}, j \text{ in } j_1, j_2, \dots, \text{ or } j_m$$

固定  $m-1$  个下标为  $i$ , 只有一个下标  $j$  从  $i$  到 4 求和, 当  $m$  个下标依次求和完毕后, 所得之和为  $M_i$ , 注意  $M_{i, i, \dots, i}$  不能重复求和.

当样本量足够大时, 频率接近概率, 落在第  $i$  ( $i=0, 1, 2, 3, 4$ ) 等的概率为

$$P_{c_i} \approx f_{c_i} = \frac{M_i}{n}$$

正常情况下, 在第  $i$  等中的概率  $P_{c_i}$  应该与“ $3\sigma$  法则”五等相应区间  $SS_i$  概率  $PS_i$  相差不大. 若组件安全态势异常时,  $P_{c_i}$  有些可能比正常概率  $PS_i$  要大得多, 也有可能比正常相应第  $i$  个区间概率  $PS_i$  要小很多. 令

$$L_c = \arg \max_i \{ CR_i \mid CR_i = \frac{|P_{c_i} - PS_i|}{PS_i}, i = 0, 1, 2, 3, 4 \}$$

式中:  $PS_i$  为经“ $3\sigma$  法则”划分的第  $i$  个区间的概率;  $CR_i$  为第  $i$  等中的概率  $P_{c_i}$  与正常概率  $PS_i$  的偏差的比率, 从 5 个  $CR_i$  中选择最大者, 有理由认为本组  $m$  个态势因子所感知的组件安全态势  $L_c$  属于第  $i$  等, 差异显著.

### 3.3 网络安全态势感知

通过构建 3 个广义立方体感知组件安全态势, 每个组件安全态势由三元组 ( $runnability_c, vulnerability_c, threat_c$ ) 构成, 而网络安全态势也由相应的三元组 ( $runnability_n, vulnerability_n, threat_n$ ) 构成. 一个网络中存在  $N$  个组件, 这  $N$  个组件的安全态势向上融合成整个网络的安全态势.

对于  $N$  个组件的分量如  $runnability_c$  要融合成网络的分量如  $runnability_n$ , 都取 0, 1, 2, 3, 4 的 5 个离散值, 将采用统计的方法. 令

$$P_{n_i} \approx f_{n_i} = \frac{|L_{c_i}|}{N}$$

式中:  $|L_{c_i}|$  为  $N$  个组件取第  $i$  等的个数 ( $i=0, 1, 2, 3, 4$ );  $P_{n_i}$  为  $N$  个组件中取值为  $i$  等的概率, 用频率逼近计算之. 令

$$L_n = \arg \max_i \{ NR_i \mid NR_i = \frac{|P_{n_i} - PS_i|}{PS_i}, i = 0, 1, 2, 3, 4 \}$$

式中:  $L_n$  为  $N$  个组件一个分量融合成网络相应的分量所取等级  $i$ . 3 个分量融合计算类似.

### 3.4 查找异常组件

当得到网络安全态势后, 三维分量中任何一维度分量级别较低, 说明整个网络的安全态势可能存在着隐患, 可方便逆向查找异常组件.

例如对于网络安全态势的基础运行性分量, 其等级为中或更低, 表示某些组件的基础运行性出现异常. 由于网络的基础运行性  $runnability_n$  由  $N$  个组件的基础运行性  $runnability_c$  经统计计算所得, 则容易查找到组件基础运行性比较低的组件, 其他两个分量类似.

对于处于停止服务或处于调试阶段的少数组件, 在感知组件安全态势时可能会出现“异常”, 但在大样本情况下, 对整个网络安全态势影响甚

微,而且可以反向查到此类组件,了解到异常组件的真正原因。

### 3.5 安全态势感知算法

包括3个算法:一是对同一类的态势因子根据历史运行信息,通过“ $3\sigma$ 法则”离散化各连续型态势因子,建立广义立方体;二是把所采集态势因子数据聚合在广义立方体格中,融合成组件安全态势  $SA_c$ ;三是从  $N$  个组件安全态势  $SA_c$  融合成网络的安全态势  $SA_n$ 。

#### 算法1

输入:同一类态势因子历史运行数据  $x_{ij}$

输出:每个连续型态势因子五等区间  $SS_i$  及广义立方体  $(A, H)$

1) for every situation factor  $x_i$  do

$$2) \mu \leftarrow \hat{\mu}_i = \bar{x}_i = \frac{1}{n}(x_{i1} + x_{i2} + \dots + x_{in})$$

$$3) \sigma \leftarrow \hat{\sigma}_i = \sqrt{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 / (n - 1)}$$

4) constructing five sections  $SS_i$  for this factor

5) end do

6) constructing general cube  $(A, H)$  according to the definition 1 and definition 2

算法1可把同一类中的  $m$  个连续型态势因子经“ $3\sigma$ 法则”划分5个区间  $SS_i$ , 每个区间的概率用  $PS_i$  表示,根据定义1和定义2对此类态势因子构建一个广义立方体  $(A, H)$ 。当3类态势因子依照算法1,则所有的连续型态势因子都会被划分为5个区间,可构建3个广义立方体。注意,对不同的态势因子  $x_i$  和  $x_j$ ,它们第  $k$  个相应区间的概率是相同的,统一用  $PS_k$  ( $k = 0, 1, 2, 3, 4$ ) 表示。

#### 算法2

输入:在同一个时段内所采集的态势因子  $n$  个  $m$  元组数据大样本  $S$

输出:组件安全态势  $SA_c$

1) discretizing the monitored  $n \times m$  data

2) the discretization  $n \times m$  data are aggregated into the general cube  $(A, H)$

3) for  $i = 0$  to 4 do

$$4) M_i \leftarrow \sum_{j=1}^4 M_{i,i,\dots,i,j,i,\dots,i} - (m - 1)M_{i,i,\dots,i,j}$$

in  $j_1, j_2, \dots, \text{or } j_m$

$$5) P_{c_i} \leftarrow f_{c_i} = \frac{M_i}{n}$$

$$6) L_c \leftarrow \arg \max_i \left\{ CR_i \mid CR_i = \frac{|P_{c_i} - PS_i|}{PS_i} \right\}$$

7) gaining a part of the component's situation  $SA_c$ , that the value is  $L_c$

8) end do

9) for the three category situation factors, repeating from 3) to 8)

10) gaining the situation of this component  $SA_c = (\text{runnability}_c, \text{vulnerability}_c, \text{threat}_c)$

在算法2计算算法1中的  $M_i$  和  $P_{c_i}$  的值,得到组件安全态势的一个分量的值  $L_c$ ,当3个分量计算完毕,得到组件的安全态势  $SA_c = (\text{runnability}_c, \text{vulnerability}_c, \text{threat}_c)$ 。

#### 算法3

输入:在同一个时段内所采集  $N$  个组件的安全态势  $SA_c$

输出:网络安全态势  $SA_n$

1) for  $j = 1$  to 3

2) for  $i = 0$  to 4 do

$$3) P_{n_i} \leftarrow f_{n_i} = \frac{|L_{c_i}|}{N}$$

4) end do

$$5) L_n \leftarrow \arg \max_i \{ NR_i \mid NR_i = \frac{|P_{n_i} - PS_i|}{PS_i} \}$$

6) gaining a part of the network's situation  $SA_n$ , that the value is  $L_n$ .

7) end do

8) gaining the situation of network  $SA_n = (\text{runnability}_n, \text{vulnerability}_n, \text{threat}_n)$ .

算法3采集同时间段内  $N$  个组件的安全态势  $SA_c$ ,根据统计的方法,计算  $N$  个组件中取值为  $i$  等的概率  $P_{n_i}$  值,再计算偏差最大级别确定为本分量级别  $L_n$ ,经过3次计算,得出网络的总安全态势  $SA_n = (\text{runnability}_n, \text{vulnerability}_n, \text{threat}_n)$ 。

算法1的时间复杂度为  $O(n \cdot m)$ ;算法2的时间复杂度为  $O(n \cdot m \cdot l)$ ,  $l$  表示立方体个数;算法3的时间复杂度为  $O(l \cdot k \cdot N)$ ,  $k$  为等级数。

## 4 仿真实验

采用 MATLAB 7.0 仿真实验,实验过程中主要采用了两类数据:一类是通过开发一个软件监测得到的实时数据;一类来自 Snort 入侵检测数据,将各类恶意网络流量数据按照一定规则注入到背景流量中,获得异常数据。

本实验测试部署在拥有 90 台计算机的局域网(操作系统为 Windows 7.0,内存为 4 G;CPU 为奔腾 3.4 GHz,其他计算机配置均与此相同)。

本局域网具有一台主机用于整个网络的安全态势感知,每台机器安装监测软件和 Snort 入侵检测软件,在本地机器上感知组件安全态势,定时向主机发送本地机器上融合数据,由主机感知整体网络的安全态势.

正常情况下,组件取连续型变量如 CPU 利用率和内存使用基本上呈高斯分布. 异常情况下,组件不安装任何防病毒软件,且对此组件实施木马和蠕虫等病毒攻击,会对各类态势因子产生影响,采样如图 4 所示,CPU 利用率及内存使用情况明显增加.

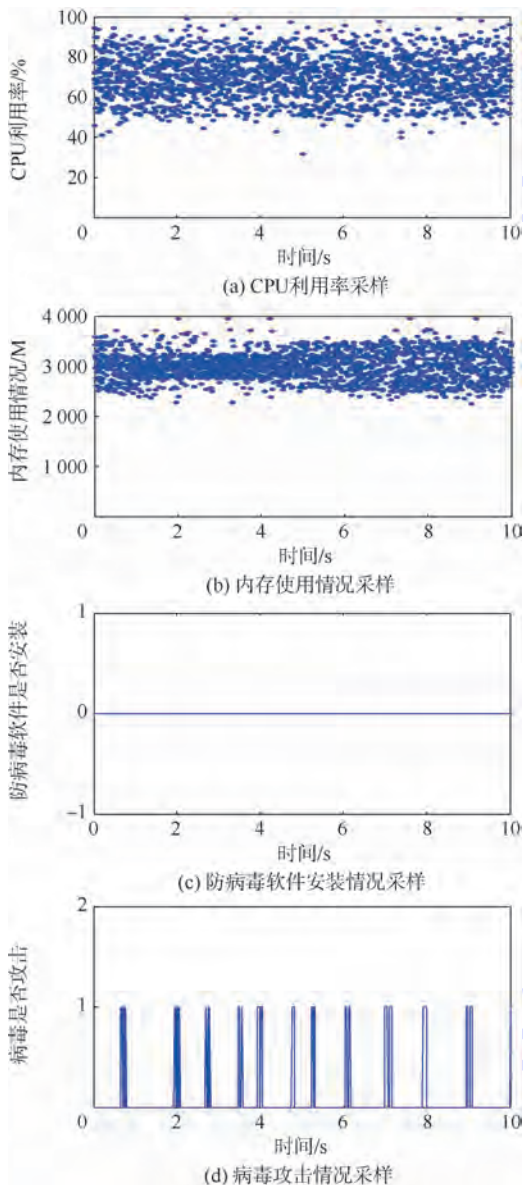


图 4 组件异常情况采样

Fig. 4 Component anomalies sample

首先为每个连续型态势因子划分 5 个离散取值区域  $SS_i$ , 再为每类的态势因子建立广义立方体,把每类相关的态势因子离散化后的值聚集在此立方体中,计算其落在每个立方体格中的频率,

近似其概率. 当组件没安装防病毒软件及受到攻击后,以“基础运行性”的广义立方体为例,五等  $f_{e_i} (i = 0, 1, 2, 3, 4)$  的频率(近似概率)之和为 1, 当样本量达到 4000 以上,逐渐平稳收敛,广义立方体中样本聚集如图 5 所示.

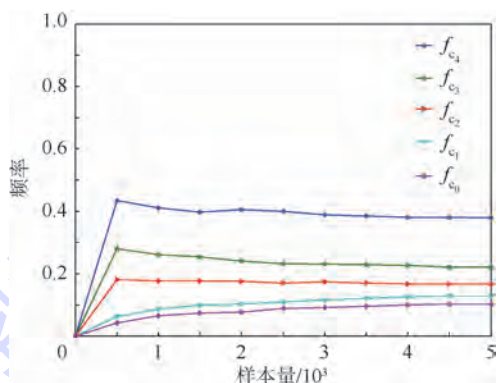


图 5 广义立方体中样本聚集

Fig. 5 Generalized cube samples gathered

所采集的态势因子大样本数据聚集在广义立方体中,得到落在每个立方体格中样本个数  $M_{j_1 j_2 \dots j_m}$ , 计算其边际之和  $M_i$ , 表示本广义立方体中取  $i$  等时原始数据的个数,用相应的五等频率  $f_{e_i} = M_i/n$  近似其第  $i$  等的概率  $P_{e_i}$ , 再从 5 个比率  $CR_i$  中选择最大者,与正常情况差异显著,有理由认为本组  $m$  个态势因子所感知的组件安全态势  $L_c$  属于第  $i$  等,如图 6 所示. 可知,基础运行性中第 0 等差异显著,认为此时取 0 等(最差);脆弱性第 0 等差异显著,此时也应取 0 等;威胁性第 4 等差异显著,此时应取第 4 等(最高,表示威胁性最高).

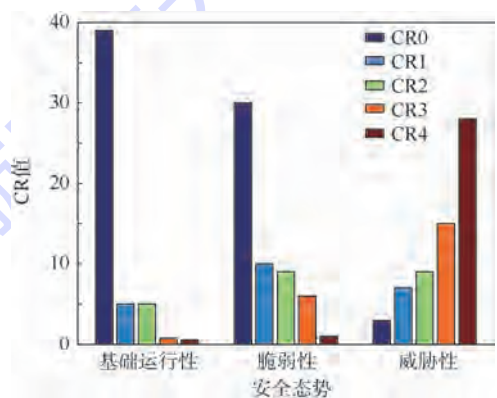


图 6 组件安全态势对比

Fig. 6 Components security situation contrast

本次实验中,对某些未安装防病毒软件的组件,人为施加病毒攻击,得到此组件的 5 次安全态势感知如图 7(a) 所示. 当获得所有组件的安全态势后,按照上述的方法可得到网络的安全态势. 同样对某些组件加以病毒攻击,得到本网络的 5 次安全态势感知图,如图 7(b) 所示.

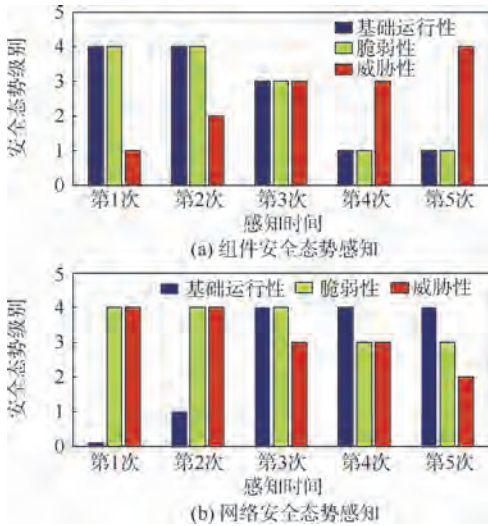


图 7 组件和网络安全态势感知

Fig. 7 Components and network security situational awareness

从上述仿真实验中可以看出,本文算法具有高效性与正确性.图 5 表明当样本量达 3 000,基本上收敛在某个平稳状态,图 6 表明安全态势受多个态势因子影响,之间具有关联性.

## 5 结束语

本文提出了构建广义立方体模型融合异构多源多层次信息,感知网络安全态势的方法,为网络安全与管理提出解决方案.在提高态势感知准确性的基础上,同时也在时空开销等因素之间进行了折衷.主要的创新为:广义立方体模型能很好地融合异构多源多层次信息,高效组织各类信息源;“ $3\sigma$ 法则”指导构建广义立方体融合模型,“ $3\sigma$ 法则”在融合算法的区间概率分级上具有严格的理论基础与实践意义.

今后的研究工作包括完善网络安全态势感知模型及量化其感知方法,进一步提高算法的效率,研究更全面的安全态势因子及其表示方法,在量化算法上突破传统的利用权重进行加权的方法;研究网络安全态势的多维可视化问题,并改进各类网络安全事件的关联分析算法,从而更加准确地预测网络安全态势发展趋势及安全威胁来源.

## 参考文献 (References)

[1] Bass T. Multi-sensor data fusion for next generation distributed intrusion detection systems [C] // Proceedings of the '99 IRIS National Symposium on Sensor and Data Fusion. Piscataway, NJ: IEEE Press, 1999: 24-27.

[2] Mazur J, Kaderali L. The importance and challenges of bayesian parameter learning in systems biology [J]. Model Based Param-

ter Estimation Contributions in Mathematical and Computational Sciences, 2013, 4: 145-156.

- [3] 黄同庆,庄毅.一种实时网络安全态势预测方法[J].小型微型计算机系统,2014,35(2):303-306.  
Huang T Q, Zhuang Y. An approach to real-time network security situation prediction [J]. Journal of Chinese Computer Systems, 2014, 35(2): 303-306 (in Chinese).
- [4] Blasch E P, Plano S. JDL level 5 fusion model “user refinement” issues and applications in group tracking [C] // Proceedings of the Signal Processing, Sensor Fusion, and Target Recognition XI, Spie. Bellingham, WA: SPIE, 2002: 270-279.
- [5] 龚正虎,卓莹.网络态势感知研究[J].软件学报,2010,21(7):1605-1619.  
Gong Z H, Zhuo Y. Research on cyberspace situational awareness [J]. Journal of Software, 2010, 21(7): 1605-1619 (in Chinese).
- [6] Bradshaw J M, Carvalho M, Bunch L, et al. Sol: An agent-based framework for cyber situation awareness [J]. KI-Künstliche Intelligenz, 2012, 26(1): 127-140.
- [7] Digioia G, Foglietta C, Oliva G, et al. Aware online interdependency modeling via evidence theory [J]. International Journal of Critical Infrastructures, 2013, 6893: 74-92.
- [8] Bazan J G, Bazan-Socha S, Buregwa-Czuma S, et al. Classifiers based on data sets and domain knowledge: A rough set approach [J]. Intelligent Systems Reference Library, 2013, 43: 93-136.
- [9] Sample C, Schaffer K. An overview of anomaly detection [J]. IT Professional, 2013, 15(1): 8-11.
- [10] 王宏,龚正虎.一种基于信息熵的关键流量矩阵发现算法 [J].软件学报,2009,20(5):1377-1383.  
Wang H, Gong Z H. Algorithm based on entropy for finding critical traffic matrices [J]. Journal of Software, 2009, 20(5): 1377-1383 (in Chinese).
- [11] 陈秀真,郑庆华,管晓宏,等.层次化网络安全威胁态势量化评估方法 [J].软件学报,2006,17(4):885-897.  
Chen X Z, Zheng Q H, Guan X H, et al. Quantitative hierarchical threat evaluation model for network security [J]. Journal of Software, 2006, 17(4): 885-897 (in Chinese).
- [12] Gömritz N, Kloft M, Rieck K, et al. Toward supervised anomaly detection [J]. Journal of Artificial Intelligence Research, 2013, 46: 235-262.
- [13] Erbacher R F, Frinckeb D A, Wongb P C, et al. A multi-phase network situational awareness cognitive task analysis [J]. Information Visualization, 2010, 9(3): 204-219.
- [14] 韦勇,连一峰,冯登国.基于信息融合的网络安全态势评估模型 [J].计算机研究与发展,2009,46(3):353-362.  
Wei Y, Lian Y F, Feng D G. A network security situational awareness model based on information fusion [J]. Journal of Computer Research and Development, 2009, 46(3): 353-362 (in Chinese).
- [15] 张勇,谭小彬,崔孝林,等.基于 Markov 博弈模型的网络安全态势感知方法 [J].软件学报,2011,22(3):495-508.  
Zhang Y, Tan X B, Cui X L, et al. Network security situation awareness approach based on Markov game model [J]. Journal of Software, 2011, 22(3): 495-508 (in Chinese).
- [16] 谢丽霞,王亚超,于巾博.基于神经网络的网络安全态势感

知[J]. 清华大学学报: 自然科学版, 2013, 53(12): 1750-1760.

Xie L X, Wang Y C, Yu J B. Network security situation aware-

ness based on neural networks[J]. Journal of Tsinghua University: Science & Technology, 2013, 53(12): 1750-1760 (in Chinese).

## Constructing general cube to be aware of network security situation

WEN Zhicheng<sup>1,2</sup>, Chen Zhigang<sup>\*1</sup>

(1. School of Information Science and Engineering, Central South University, Changsha 410083, China;

2. College of Computer and Communication, Hunan University of Technology, Zhuzhou 412007, China)

**Abstract:** Concerning the problems of limited current network security situation assessment scope, single information source, higher time and space complexity and larger deviation of the accuracy, a method was put forward to construct general cube, which can be aware of the network security situation. The continuous situation factor data monitored can be pretreated by discretizing by “ $3\sigma$  rule” and aggregated in the general built cube, that fused into component security situation vertically and merged into the network security situation from component security situation using statistical methods horizontally. It can provide reliable reference to enhance network security. Finally, making full use of network data, the network security situation awareness model and algorithm proposed are verified and the experimental results show correctness of this method.

**Key words:** network security; situation awareness; network manager; information fusion; general cube

## 《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

### 1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

1.3 本刊只接收学术及科研论文,不刊登研究简报、综述文章和快讯。

### 2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码,中文摘要(包括目的、方法、结果及结论),中文关键词(5~8 个),中图分类号,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词,引言,正文,参考文献。首页下角注明基金项目名称及编号,作者信息(包括姓名、出生年、性别、籍贯、职称或学生身份、E-mail 地址)。

2.5 作者请登录本刊网页进行在线投稿。

### 3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定录用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投它刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经刊登,即赠送单行本。

3.4 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100191 北京市海淀区学院路 37 号 北京航空航天大学学报编辑部

办公地点:北京航空航天大学办公楼 404,405,407,409 房间

电 话:(010)82315339,82338922,82314839,82315426,82315594,82317826

E-mail: jbuua@buaa.edu.cn

http://bhxb.buaa.edu.cn

http://www.buaa.edu.cn

《北京航空航天大学学报》  
第五届编辑委员会

主任(主编): 赵沁平

(以下按姓氏笔划为序)

副主任(副主编): 丁希仑 王少萍 孙志梅 李秋实 李焕喜 杨嘉陵  
苗俊刚 相艳 徐立军 钱德沛 曹晋滨

编委: 马殿富 王琪 王聪 邓小燕 王青云 王荣明 刘宇  
刘红 江洁 刘强 闫鹏 朱天乐 刘铁钢 齐铂金  
陈万春 邹正平 苏东林 杨世春 沈成平 邱志平 宋知人  
杨树斌 张晓林 杨晓奕 杨继萍 李惠峰 吴新开 张瑞丰  
杨照华 宋凝芳 周锐 林宇震 林贵平 战强 姚仰平  
胡庆雷 赵秋红 段海滨 赵巍胜 席平 郭宏 徐洁  
徐世杰 郭洪波 康锐 翟锦 熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

(原《北京航空学院学报》)

(月刊 1956年创刊)

第41卷第10期 2015年10月

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS  
(JBUA)

(Monthly, Started in 1956)

Vol.41 No.10 October 2015

主管单位 中华人民共和国工业和信息化部  
主办单位 北京航空航天大学  
主编 赵沁平  
编辑出版 《北京航空航天大学学报》  
编辑部  
邮编 100191  
地址 北京市海淀区学院路37号  
印刷 北京科信印刷有限公司  
发行 北航文化传媒集团  
发行范围 国内外发行  
联系电话 (010) 82315594 82338922  
82314839  
电子信箱 jbuaa@buaa.edu.cn

Administrated by Ministry of Industry and Information  
Technology of the People's Republic of China  
Sponsored by Beijing University of Aeronautics  
and Astronautics (BUAA)  
(Beijing 100191, P. R. China)  
Chief Editor Zhao Qinpeng  
Edited and Published by Editorial Board of JBUA  
Printed by Beijing Kexin Printing Co., Ltd.  
Distributed by BUAA Culture Media Group Limited  
Telephone (010) 82315594 82338922  
82314839  
E-mail jbuaa@buaa.edu.cn  
http://bhxb.buaa.edu.cn

刊号 ISSN 1001-5965  
CN 11-2625/V  
国内定价 20.00元/期

