场信号预处理的扩展视频编码器设计

姚春莲 李 炜 孟庆磊

(北京航空航天大学 计算机学院, 北京 100083)

摘 要: 虽然视频标准中给出了场信号的一些编码方法,但由于宏块划分方式比较多造成运动估计非常耗时,很难得到实际应用. 为此,设计了具有运动自适应去隔行预处理功能的扩展视频编码器. 该编码器先将场信号处理成逐行信号,之后采用固定帧方式进行编码. 在将场信号处理成帧信号时,首先对相同奇、偶性的四场做运动检测,然后对具有不同运动特点的区域采用不同的插值方法得到待插值像素,最后插值场与当前场交织成逐行帧. 测试结果表明,该编码器在相同压缩比下能取得与其它编码方法相同的峰值信噪比,编码一帧时间为宏块级帧/场自适应方法的 1/3. 此外,该编码器能有效地解决场信号在逐行扫描设备上的显示问题.

关 键 词: 视频编码; 帧场自适应; 去隔行; 中值滤波

中图分类号: TN 919.8

文献标识码: A **文 章 编 号**: 1001-5965(2006)09-1046-05

Design of extended video encoder with pre-processing of interlace signal

Yao Chunlian Li Wei Meng Qinglei

(School of Computer Science and Technology, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Video coding stands gives out several coding methods for interlace signal. However, many kinds of macroblock division methods for motion estimation, which makes these methods very time-consuming and can't be used actually. So an extended video encoder was designed, motion-adaptive de-interlacing was used to convert interlace signal into frame signal, then encoding frame signal with the fixed frame method. There are three steps for signal converting. Firstly, motion detection by utilizing four field signals with the same parity to get motion status of pixel, secondly, according to the motion status of pixel, use different filter to interpolate the pixel, thirdly, interlace interpolation field and current field to get progressive frame. The experimental result shows that it can get the same peak signal noise rate as MBAFF, but the encoding time is about 1/3 of other methods. What's more, when displayed on progressive screen, it can solve the problem such as line crawling etc.

Key words: video encoding; frame/field adaptive; de-interlacing; median filter

视频信号从扫描格式上分为逐行和隔行 2 种,一帧隔行视频信号由两场组成(顶场和底场),两场采集时有一定的时间间隔. 在编码时,可以基于帧方式,也可以基于两场方式进行编码. ITU(International Telecommunication Union)和 ISO(International Organization for Standardization)组织制定的 H.26x 和 MPEG-x 系列标准中,给出了对场信号的编码方法. 主要可分为 2 类,帧/场固定和帧/场自适应方式[1-2]. 帧/场固定方式指对输入的场信号固定采用帧方式或者场方式来划分宏块,实现简单,速度快,由于不考虑视频序列内容,编码性能较低;帧/场自适应方式自适应地选择帧或场方式进行编码,可分为图像级和宏块级 2 种,属于基于图像内容的编码方式,但由于实现复杂

收稿日期: 2005-11-25

基金项目: 国家自然科学基金资助项目(60505007);教育部博士点基金资助项目(2003006034);新世纪优秀人才支持计划

作者简介: 姚春莲(1973-),女,河北唐山人,博士生,demons1217@sina.com.

而很难实际应用. 由于 2 类场信号编码方法都没有经过预先的场到帧的变换,因此不能解决场信号在逐行设备上显示时存在的行爬行、锯齿效应、闪烁等问题.

针对上述问题,设计了以运动自适应去隔行做为预处理的扩展视频编码器. 首先对场信号做运动检测,然后对运动区域和静止区域用不同的滤波处理,将场信号处理成逐行信号,最后以帧方式划分宏块实现对场信号的编码. 该编码器解决了场信号在逐行设备上的显示问题,且预处理中的滤波器能够减少输入视频的噪声,有效地提高压缩比.

1　视频标准中的场信号编码方式

视频标准针对场信号的编码方法主要有帧/场固定方式和帧/场自适应方式 2 种[3-4].

1.1　帧/场固定方式

固定帧方式编码场信号时,在宏块划分方式、预测帧选择方面与编码逐行信号相似,16×16 的宏块以帧间或帧内 2 种方式编码,作为帧内宏块时可用 16×16 和 4×4 划分方式,作为帧间宏块时可以选择 7 种(16×16,16×8, 8×16, 8×8, 8×4, 4×8, 4×4)划分方式之一,运动预测时不能采用本帧内的顶场或底场做参考. 固定场方式下,每个 16×16 的宏块首先分成 2 个场块,即顶场块和底场块,2 个场块允许有 5 种划分方式,即 8×16, 8×8, 8×4, 4×8, 4×4. 帧/场固定方式编码不考虑视频信号的内容,容易实现,速度快, 只需在编码的序列头写上帧或场标识即可.

1.2　帧/场自适应编码方式

帧/场自适应从粒度上可分为图像级和宏块级帧/场自适应方式. 图像级自适应相当于固定帧/场 2 种方式的结合,分别用帧方式和场方式进行编码,获得 2 种方式下的位率(bitrate)和失真(distortion),根据 cost = distortion + λbitrate 选择代价(cost)最小的方式编码. 相当于进行 2 遍编码且需要为每一帧写入编码的帧或场标识.

宏块级帧/场自适应编码方式,在宏块级为每个宏块选用 cost 最小的划分方式进行编码. 图 1 为宏块级帧/场自适应与图像级帧/场自适应相结合的编码框图.

图 1 中左半部分为帧方式编码,每个宏块用帧/场 2 种划分方式;编码右半部分为场方式编码,宏块只能以场方式划分. 最后在图像级选择帧或场代价最小的方式进行编码. 从图 1 中可以看出,宏块级帧/场自适应编码方式可以得到每个宏块的最佳编码模式,但由于涉及多种块模式的运动估计以及率失真计算,编码时间约为固定帧/场编码方式的 3 倍左右,此外每个宏块都额外需要帧/场宏块类型的标识,以及参考帧/场的标识.

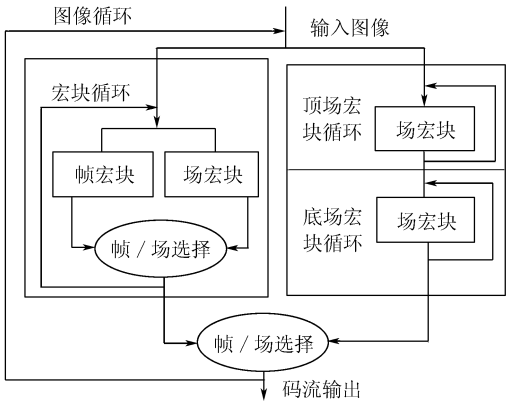


图 1　宏块级帧/场适应编码框图

2　运动自适应去隔行的场信号编码

视频标准中的帧/场自适应编码方式虽然是考虑到视频内容的编码方式,但是其编码比较耗时,需要增加额外码流造成压缩性能下降,并且不能解决场信号在逐行设备显示时存在的场效应问题. 为此,本文设计了以去隔行为预处理的场信号编码器,首先将场信号经过运动自适应去隔行处理成逐行信号,结合固定帧方式进行编码.

2.1　运动自适应去隔行预处理

2.1.1　去隔行原理与方法

去隔行的原理就是将隔行场图像转换为逐行帧图像的处理方法[5],其形式化描述为

$$F_o(\boldsymbol{x},n)=\begin{cases}F(\boldsymbol{x},n) & \text{if } y \bmod 2 = n \bmod 2\\ F_i(\boldsymbol{x},n) & \text{其它}\end{cases} \quad (1)$$

式中,$\boldsymbol{x}=(x,y)^{\mathrm{T}}$ 为像素位置;$F_o(\boldsymbol{x},n)$ 为输出逐行帧;$F(\boldsymbol{x},n)$ 为当前场;$F_i(\boldsymbol{x},n)$ 为待插值场;n 为场号. 去隔行处理方法主要分为 3 类[5]:场内插值、场间插值、场内与场间结合插值. 场内插值相当于空域滤波,这类方法计算量小,因只利用了当前场的信息,图像垂直分辨率较低. 场间插值相当于时域滤波,在两场间无运动的情况下,该方法能获得高的图像垂直分辨率,有运动时,会出现行爬行现象. 场内与场间结合插值相当于时/空结合滤波,这类方法应用广泛,获得高垂直分辨率的同时能有效解决行爬行问题和边缘模糊等现象.

2.1.2　运动自适应去隔行方法

图 2 为本文运动自适应去隔行处理算法框图. 首先对相同奇、偶性的四场做运动检测,并对检测结果进行腐蚀、扩张等形态学滤波处理;然后

对运动区域采取基于边缘的插值(ELA),静止区域直接取前一场的信息,即场合并方法;最后插值场与当前场交织成逐行帧.

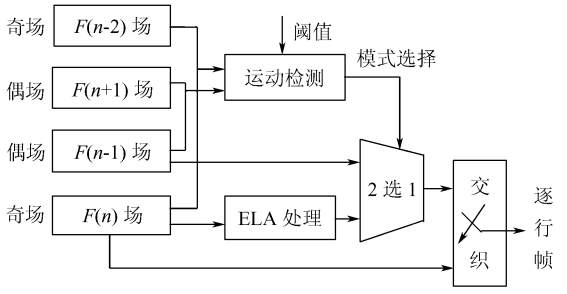


图 2 运动自适应去隔行算法功能框图

由于前后 2 场扫描线位置不同,缓冲 1 场或 2 场数据做运动检测会对噪声比较敏感,而多场缓冲[6-7]能够较好地解决运动检测的准确性问题. 图 3 中缓冲当前场前面的 2 场和后面 1 场,求取当前场与前前场的 2 个差,前 1 场与后 1 场的差,取 3 个差的中间值与阈值做比较,得到当前场中待插值像素(m,n)的运动情况.

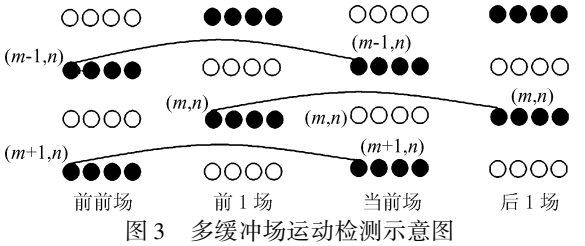


图 3 多缓冲场运动检测示意图

$$d_1(m,n) = |f_f(m,n) - f_b(m,n)| \quad (2)$$

$$d_2(m-1,n) = |f_c(m-1,n) - f_{ff}(m-1,n)| \quad (3)$$

$$d_3(m+1,n) = |f_c(m+1,n) - f_{ff}(m+1,n)| \quad (4)$$

$$d(m,n) = \begin{cases} 1 & \text{if}(\text{med}(d_1(m,n),d_2(m-1,n),d_3(m+1,n)) > V_{th}) \\ 0 & \text{其它} \end{cases} \quad (5)$$

式(2)中,$f_f(m,n)$,$f_b(m,n)$分别为前 1 场与后 1 场中位于(m,n)处的像素亮度值;$d_1(m,n)$表示二者之差. 式(3)中,$f_c(m-1,n)$,$f_{ff}(m-1,n)$分别为当前场与前前场中位置在$(m-1,n)$的像素亮度值;$d_2(m-1,n)$表示二者之差;$f_c(m+1,n)$,$f_{ff}(m+1,n)$分别为当前场与前前场位置在$(m+1,n)$的像素亮度值;$d_3(m+1,n)$表示二者之差. 式(5)中的$d(m,n)$表示当前场待插值像素(m,n)的运动情况,取$d_1(m,n)$,$d_2(m-1,n)$,$d_3(m+1,n)$的中值与阈值做比较,大于阈值说明待插值像素(m,n)为运动像素点,小于阈值说明为静止像素点,式(5)中的V_{th}为检测阈值,由高斯检测方法获得[8].

运动物体的边缘处会出现锯齿效应,需要正确地检测边缘并沿边缘方向滤波,本文采用多方向检测以提高检测精度,共 5 个方向±60°,±45°,0°(见图 4).

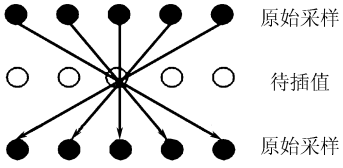


图 4 多方向边缘检测

$$\sigma(\theta) = \begin{cases} -2 & \theta = -60° \\ -1 & \theta = -45° \\ 0 & \theta = 0° \\ 1 & \theta = 45° \\ 2 & \theta = 60° \end{cases}$$

$$\Delta(\theta) = \alpha \mid F((x+\sigma(\theta)),(y-1)) - F((x-\sigma(\theta),(y+1)) \mid \quad (6)$$

$$\theta_{\min} = \min(\Delta(\theta) \mid \theta = 60°,45°,0°,-45°,-60°)$$

$$V_{ela} = (F((x+\sigma(\theta_{\min})),(y-1)) + F((x-\sigma(\theta_{\min})),(y+1)))/2 \quad (7)$$

$\sigma(\theta)$代表了 5 个不同边缘方向角度,式(6)中$\Delta(\theta)$代表不同方向的上下 2 行像素差值,$F((x+\sigma(\theta)),(y-1))$,$F((x-\sigma(\theta)),(y+1))$为待插值像素上/下行对应于角度$\theta$的像素点亮度值,$\theta_{\min}$为差值最小的方向. 为了减少插值误差,式(6)中参数α的值随方向而变化,当$\theta=\pm45°$时,参数$\alpha=1/\sqrt{2}\approx0.7$;$\theta=\pm60°$时,参数$\alpha=1/2$;$\theta=0°$时,参数$\alpha=1$. 式(7)中$V_{ela}$为待插值点的亮度值.

$$F_0(\boldsymbol{x},n) = \begin{cases} F(\boldsymbol{x},n) & \text{if } y \bmod 2 = n \bmod 2 \\ \text{med}\begin{pmatrix} V_{ela} \\ F(\boldsymbol{x}-\boldsymbol{u}_y,n) \\ F(\boldsymbol{x}+\boldsymbol{u}_y,n) \end{pmatrix} & \text{其它} \end{cases} \quad (8)$$

将V_{ela}值和当前场待插值行的上、下行做为中值滤波的输入,中值滤波器可以避免突发噪声.

2.2 带去隔行预处理的编码器

图 5 为带去隔行预处理的扩展视频编码器,将输入的场序列进行 2.1 节的运动自适应去隔行预处理后,转换成帧序列,然后以帧方式划分宏块,输出压缩码流.

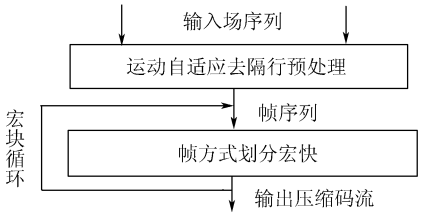


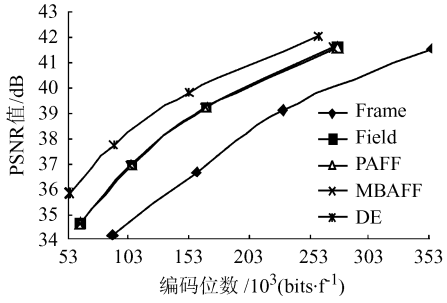
图 5　带去隔行预处理的扩展视频编器

3　性能测试与分析

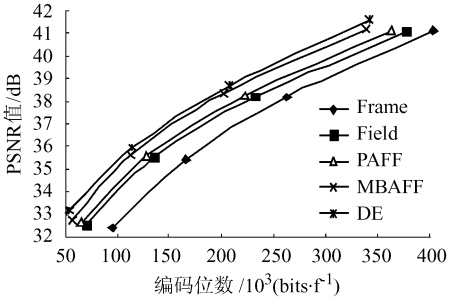
本文将 2.2 节的扩展视频编码器在 H.264 标准的参考软件 JM9.0 上实现,使用峰值信噪比 PSNR(Peak Signal Noise Rate)值作为衡量指标.从标准序列中选取相对静止的 Panio,运动程度中等的 Stefan,运动剧烈的 Football,纹理复杂的 Flower 4 个具有代表性的场序列.通过调整量化参数 QP 为 20,24,28,32,测得各序列在 5 种不同编码方式下的编码位数和 PSNR 值,其中 Frame 为固定帧方式编码,Field 为固定场方式编码,PAFF 为图像级帧/场自适应方法,MBAFF 为宏块级帧/场自适应方法,De _ Encoder 为本文的扩展视频编码器.根据表 1 的结果,得到序列的率失真曲线,见图 6.

表 1　各种编码方法的 PSNR 值和位率比较

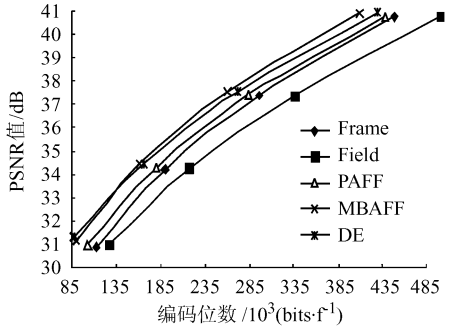
序列	QP	Frame		Field		PAFF		MBAFF		De _ Encoder	
		PSNR 值/dB	编码位数/(bit·f^{-1})	PSNR 值/dB	编码位数/(bit·f^{-1})	PSNR 值/dB	编码位数/(bit·f^{-1})	PSNR 值/dB	编码位数/(bit·f^{-1})	PSNR 值/dB	编码位数/(bit·f^{-1})
Football	20	41.58	355 659	41.59	276 775	41.60	276 596	41.61	273 542	42.05	260 745
	24	39.15	231 406	39.25	168 012	39.25	168 130	39.26	166 159	39.80	153 221
	28	36.68	159 729	36.96	105 706	36.96	105 362	36.96	104 814	37.76	91 136
	32	34.17	90 143	34.61	63 305	34.61	63 111	34.62	63 020	35.84	53 831
Flower	20	40.77	448 066	40.75	500 453	40.78	438 022	40.89	408 892	40.94	429 580
	24	37.40	295 285	37.38	335 856	37.42	284 732	37.55	260 771	37.53	271 084
	28	34.23	190 170	34.26	217 779	34.30	179 542	34.43	161 057	34.46	165 703
	32	30.88	112 926	30.98	126 684	31.00	103 097	31.15	90 458	31.34	87 323
Stefan	20	41.13	403 647	41.06	378 569	41.11	363 381	41.17	339 557	41.63	342 234
	24	38.22	262 879	38.21	232 519	38.26	221 901	38.34	202 460	38.67	207 293
	28	35.42	165 633	35.54	136 104	35.57	128 115	35.65	112 781	35.96	114 825
	32	32.42	94 960	32.57	70 222	32.64	64 792	32.76	56 377	33.17	54 231
Panio	20	40.50	577 903	40.05	534 100	40.27	537 685	40.28	496 328	41.14	519 897
	24	37.33	298 357	37.26	295 718	37.32	279 700	37.40	248 530	38.10	299 606
	28	35.01	193 711	34.87	188 201	35.0	175 866	35.12	152 302	35.60	175 422
	32	32.25	117 630	32.04	110 460	32.22	101 726	32.39	85 712	32.74	96 276



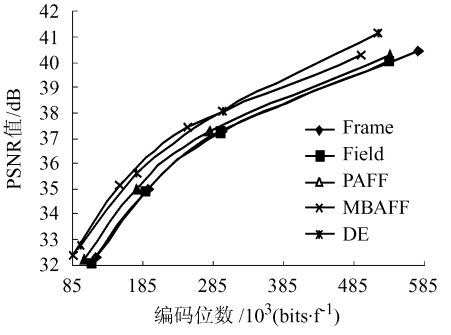
a　Football 序列



b　Stefan 序列



c　Flower 序列



d　Panio 序列

从表 1 的测试数据和图 6 的率失真曲线可以看出,PAFF 算法相当于 Frame/Field 算法的最优结合,其性能比 Frame/Field 算法要好,MBAFF 算法相当于 Frame/Field/PAFF 3 个算法的最优组合,压缩性能比 3 种方法都好. 本文算法的压缩性能高于 Frame/Field/PAFF 3 个算法,对于 Football 和 Stefan 序列,在相同压缩比下,PSNR 值比 MBAFF 算法高 0.2 ~ 1 dB,对于 Flower 和 Panio 序列取得了与 MBAFF 算法相当的结果.

在 PIV 1.8 GHz,内存为 256 MB 的机器上,编码 Panio 序列一帧,5 种算法所用的时间分别为 Frame:31.81 s, Field:42.65 s, PAFF:72.82 s, MBAFF:111.88 s, De _ Encoder:31.65 s, MBAFF 是本文方法时间的 3 倍多.

4 结束语

国际标准中针对场视频信号提出了宏块级、图像级帧/场自适应编码方法,但由于引入大量的运算,且没有考虑到场图在逐行显示设备上出现的一系列问题,无法得到实际的应用. 因此,为提高压缩的质量,以运动自适应去隔行处理做为视频编码器的预处理,构成扩展视频编码器,该编码器中的去隔行处理降低了输入图像的噪声,提高运动估计和运动补偿的匹配精度,因此其压缩性能比压缩标准所提供的算法高,且运算速度快.

参考文献 (References)

[1] Wang L, Panusopone K, Gandhi R, et al. Interlace coding tools for H.26L video coding[C/OL]// Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Meeting, Thailand, VCEG-O37.doc, 2001. http://ftp3.itu.ch/av-arch/video-site

[2] Wang L, Panusopone K, Gandhi R, et al. Adaptive frame/field coding for JVT video coding[C/OL]// JVT Meeting, Geneva, JVT-B-071.doc, 2002. http://ftp3.itu.ch/av-arch/video-site

[3] Ajeev Gandhi, Wang Limin, Panusopone K, et al. Macroblock adaptive frame/field coding for interlacer sequences[C/OL]// JVT Meeting, Austria, JVT-D108.doc, 2002. http://ftp3.itu.ch/av-arch/video-site

[4] Borgwrdt P. Core experiment on macroblock adaptive frame/field encoding[C/OL]// JVT Meeting, Geneva, JVT-B117.doc, 2002. http://ftp3.itu.ch/av-arch/video-site

[5] Gerard De, ErWin Br. De-interlacing-an overview[C]// Proceedings of IEEE, 1998,86(9):1839-1856

[6] Koivunen Tero. Motion detection of an interlaced video signal [J]. IEEE J Consumer Electronics, 1994,40(3):753-759

[7] Lin Shyhfeng, Chang Yuling, Chen Lianggee. Motion adaptive interpolation with horizontal motion detection for de-interlacing [J]. IEEE J Consumer Electronics, 2003,49(4):1256-1265

[8] Chien Shaoyi, Huang Yuwen, Hsieh Bingyu, et al. Fast video segmentation algorithm with shadow cancellation, global motion compensation, and adaptive threshold techniques [J]. IEEE J Multimedia, 2004,6(5):732-748

(责任编辑:娄 嘉)

(上接第 1014 页)

[2] 郦能敬. 预警机系统导论[M]. 北京:国防工业出版社, 1998:140-144
Li Nengjing. Early warning aircraft system theory [M]. Beijing: National Defense Industry Press, 1998:140-144(in Chinese)

[3] 张有为,李少洪. 雷达系统分析[M]. 北京:国防工业出版社, 1981: 184-194
Zhang Youwei, Li Shaohong. Radar system analysis [M]. Beijing:National Defense Industry Press, 1981:184-194(in Chinese)

[4] Blake Lamont V. Radar range-performance analysis [M]. Lexington, Mass:Lexington Books,1980:293-296

[5] Morchin William C. Airborne early warning radar[M]. Boston: Artech House, 1989:140-145

[6] 王永良,彭应宁. 空时自适应信号处理[M]. 北京:清华大学出版社, 2000:14-25
Wang Yongliang, Peng Yingning. Space-time adaptive processing [M]. Beijing:Tsinghua University Press, 2000:14-25(in Chinese)

[7] 王永良,陈建文,吴志文. 现代 DPCA 技术研究[J]. 电子学报,2000,28 (6):118-121
Wang Yongliang, Chen Jianwen, Wu Zhiwen. Research on modern DPCA techniques[J]. ACTA Electronica Sinica, 2000, 28 (6):118-121(in Chinese)

[8] 保铮,张玉洪,廖桂生,等. 机载雷达空时二维信号处理[J]. 现代雷达,1994,2(1):38-48
Bao Zheng, Zhang Yuhong, Liao Guisheng, et al. Space-time signal processing for airborne radar [J]. Modern Radar, 1994,2 (1):38-48(in Chinese)

[9] 杨凤凤,郭汉伟. 二单元 DPCA 在机载相控阵雷达 AMTI 中的应用 [J]. 现代雷达,2004,26(3):18-20
Yang Fengfeng, Guo Hanwei. Application of two sectional DPCA in airborne phased array radar AMTI [J]. Modern Radar, 2004, 26(3):18-20(in Chinese)

(责任编辑:娄 嘉)

图 6 5 种算法对应的率失真曲线