基于 Gabor 滤波器的图像目标识别方法

曾姝彦 张广军 李秀智

(北京航空航天大学 仪器科学与光电工程学院, 北京 100083)

摘 要: 为了给机器人视觉导航提供有效信息,提出一种基于图像匹配的目标识别方法. 对 CCD 采集的目标图像,由 Gabor 变换生成的二维 Gabor 滤波器有着优良的滤波器性能,不需要对图像进行分割,能适应一定的旋转、尺度、光照的变化,通过多个频率和角度的 Gabor 算子与图像的卷积,获取图像全局信息的特征描述. 分类方法采用统计学习理论基础上发展起来的一种新的机器学习方法——支持向量机(SVM, Support Vector Machines),它可解决模型选择、过学习、维数灾难等问题. 通过支持向量机进行多维特征向量的分类. 该方法可达到较高的识别率,达到实时处理的要求,可以在人脸识别、机器人视觉定位等领域得到应用.

关 键 词: 目标识别; 多通道 Gabor 滤波器; 支持向量机; 视觉定位

中图分类号: TP 391

文献标识码: A **文 章 编 号**: 1001-5965(2006)08-0954-04

Image target distinguish based on Gabor filters

Zeng Shuyan Zhang Guangjun Li Xiuzhi

(School of Instrument Science and Opto-electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: In order to provide useful information for mobile robot, the target distinguish approach based on image matching needs to recognize image pattern. 2D Gabor filters were created by Gabor transform, which had excellent performance and didn't need to segmente images catched by the single CCD camera. Gabor filters were robust to the various orientations and illumination of target images; They also satisfied the real time image processing. The multi-channel Gabor filters, with multi-frequencies and multi-angles, were used to convolute with images. The corresponding filter results included entirely information of images, from which the feature vectors of the images could be extracted. The method of classifying these feature vectors was SVM (support vector machine). SVM was a new method of machine learning developed by statistical learning theory. It resolved problems such as model select, over learning, non-linear, high dimension, etc. Experiment results indicate that the algorithm can reach up to high recognize rate. The algorithm can be applied in fields such as face recognition, robot vision localization and so on.

Key words: image target distinguish; multi-channel Gabor filter; SVM (support vector machines); vision localization

在机器人移动过程中,利用安装在机器人上的摄像头摄取目标场景的图像,与已知场景图像进行匹配,识别出机器人所处的位置,以达到视觉定位的目的[1-2]. 目前的图像匹配算法包括基于灰度的匹配技术和基于特征的匹配技术. 基于灰度的算法是根据一对图像灰度之间点点匹配或者局部区域之间的匹配定义相关函数;基于特征的匹配算法则是对图像进行某种特征的抽取,根据特征的描述选择分类方法进行匹配运算. 用来衡量匹配算法的主要性能指标是正确识别的概率和算法的耗时,目前大多数算法存在的问题是速度过慢[3]. 为了适应多样化的环境,以及同一场景由于不同的拍摄角度导致的一些变化,如获取的图像尺度伸缩、平面内旋转、空间任意旋转等变

收稿日期: 2005-10-10

基金项目: 国家自然科学基金资助项目(50375012)

作者简介: 曾姝彦(1982-),女,湖北洪湖人,硕士生,zengshuyan@126.com.

换,要求算法具有综合的识别不变性. 由 Gabor 变换生成的二维 Gabor 滤波器有着优良的滤波器性能,通过多个频率和角度的 Gabor 算子与图像的卷积,获取图像的全局信息,不需要对图像进行分割,能适应一定的旋转、尺度、光照的变化,而且 Gabor 变换是一种较快的处理算法. 它在频率域和空间域同时具有很好的局部性,从而被广泛用于图像处理、人脸识别、纹理特征识别等领域[4].

1　Gabor 滤波器法提取特征向量

1.1　Gabor 滤波器

Gabor 变换的一个重要特性就在于 Gabor 变换系数揭示了一个信号或一幅图像的局部频率的分布规律,而不像 Fourier 变换系数仅仅反映全局频率的信息. Gabor 变换在一维和二维信号处理中是非常有用的,其优点已在多个研究方向被证明,如语音识别、信号的检测、图像压缩、纹理分析、图像的分割和识别等方面. Gabor 变换的另一有趣特性在于 Gabor 基本函数具有与人类视觉基元相似的性质,尤其是其二维基本函数波形,类似于哺乳动物视觉系统的感受面波形,这就从数学上为研究动物视觉感受面的空间特性提供了一种有效的方式,并且这一点对研究图像特征检测或空间频率滤波也是极其重要的[5].

对称的 2D-Gauss 函数定义为 $g(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$,其中 σ 为标准差,函数中心为(0,0)[6]. 沿 x 轴的复变正弦波定义为 $e^{j\omega x}$,根据欧拉公式展开为

$$e^{j\omega x}=\cos(\omega x)+j\sin(\omega x)=\cos\frac{2\pi x}{l}+j\sin\frac{2\pi x}{l}$$

2D-Gabor 函数是在 Gauss 包络下的一个沿 x 轴的复变正弦波:

$$\mathrm{Gabor}(x,y)=g(x,y)\cdot e^{j\omega x}=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\left(\cos\frac{2\pi x}{l}+j\sin\frac{2\pi x}{l}\right)$$

根据 3σ 原理,取 $l=6\sigma/f$,f 为中心频率. 将 Gabor (x,y) 中的 x,y 坐标以 θ 角度进行方向坐标变换:

$$x'=x\cdot\cos\theta+y\cdot\sin\theta$$

$$y'=-x\cdot\sin\theta+y\cdot\cos\theta$$

Gabor 函数的实部见图 1,其大小为

$$\mathrm{Gabor}_R(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\cos\frac{2\pi(x\cos\theta+y\sin\theta)}{l}$$

Gabor 函数的虚部见图 2,其大小为

$$\mathrm{Gabor}_I(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\sin\frac{2\pi(x\cos\theta+y\sin\theta)}{l}$$

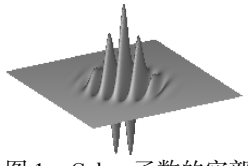


图 1　Gabor 函数的实部

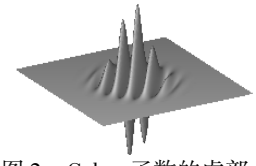


图 2　Gabor 函数的虚部

决定 Gabor_R 和 Gabor_I 的参数有 3 个:f,θ,σ. f 选择 2,4,8 三个尺度,方向 θ 选择 0°,22.5°,45°,67.5°,90°,112.5°,135°,157.5°八个方向,σ 越大能量越分散,越小越集中,选择 $\sigma=\pi$,共计 24 个滤波器. 将图像 $F(x,y)$ 与滤波器进行卷积:

$$Q(x,y)=[(\mathrm{Gabor}_R(x,y)*F(x,y))^2+(\mathrm{Gabor}_I(x,y)*F(x,y))^2]^{1/2}$$

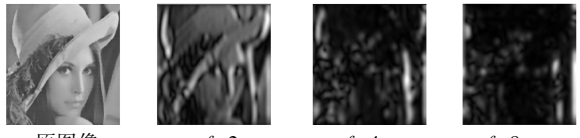
1.2　全局特征向量

每一幅 $N\times N$ 的场景图像 $F(x,y)$ 进行 Gabor 变换后得到 24 幅图像 $Q(x,y)$,对这 24 幅图的特征进一步提取,计算其均值和方差,构成一个 48 维的特征向量,作为图像的全局特征信息.

均值　$$E=\frac{\sum_x\sum_y Q(x,y)}{N^2}$$

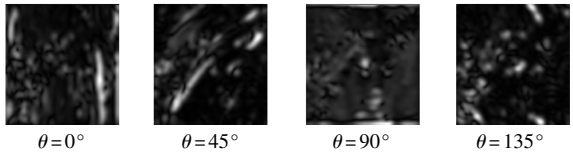
方差　$$D=\sqrt{\sum_x\sum_y\frac{[Q(x,y)-E]^2}{N^2}}$$

图 3 是原图像与 $f=2,4,8,\theta=0°$ 的 Gabor 变换结果,当 f 取值小时获得的是图像的细节信息,当 f 取值大时获得的是全局信息,可以获得不同尺度的全局信息. 图 4 是 $f=2$,不同 θ 角度方向的 Gabor 算子进行卷积后的结果,可以获得不同方向上的图像信息.



原图像　$f=2$　$f=4$　$f=8$

图 3　不同中心频率的 Gabor 变换图像($\theta=0°$)



$\theta=0°$　$\theta=45°$　$\theta=90°$　$\theta=135°$

图 4　不同角度的 Gabor 变换图像($f=2$)

Gabor 算法可用快速傅里叶卷积实现,是一种快速算法. 同时,为了保证数据处理的实时性,由微视凌志公司生产的 MOKA-C10 图像采集卡获取的图像大小设置为 128 像素×128 像素;24 个 Gabor 滤波器的数据可以作为数据文件保存,避免重复计算浪费时间和资源.

2 SVM支持向量机

2.1 SVM理论

支持向量机(SVM, Support Vector Machines)是在统计学习理论基础上发展起来的一种新的机器学习方法.统计学习理论对有限样本情况下模式识别中的一些根本性问题进行了系统研究,很大程度上解决了模型选择与过学习、非线性和维数灾难、局部极小点等问题,因此成为目前研究的热点[7].SVM方法的主要优点有:①其目标是得到现有信息下的最优解,而不仅仅是样本数趋于无穷大时的最优值;②算法最终将转化为一个二次型寻优问题,全局最优点解决了在神经网络方法中无法避免的局部极值问题;③算法将实际问题通过非线性变换转换到高维的特征空间(feature space),在高维空间中构造线性判别函数来实现原空间中的非线性判别函数特殊性质,能保证机器有较好的推广能力,同时它巧妙地解决了维数问题,其算法复杂度与样本维数无关.SVM是从线性可分情况下的最优分类面发展而来的,基本思想可用图5的二维情况说明[8].

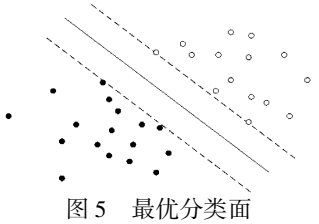


图5 最优分类面

图5中实心点和空心点代表2类样本,实线为分类线,2条虚线是2条平行于分类线的直线,分别穿过2类中离分类线最近的那些标本,虚线之间的距离叫做分类间隔(margin).所谓最优分类线就是要求分类线不但能将2类样本正确分开(训练错误率为0),而且使分类间隔最大.

算法最终将转化为一个二次型寻优问题,从理论上说得到的将是全局最优点.分类器由图5中被圈起来的点决定,这些点就被称为支持向量.如果样本不是线性可分,不可能满足以上的条件,可以引入一个惩罚参数C,得到推广的最大间隔法.

2.2 核函数的选择

在支持向量机中,需要选择函数$K(.,.)$,或者说需要选择一个映射$\boldsymbol{\Phi}(.)$,把x所在的输入空间$\boldsymbol{X}$映射到另一个空间$\boldsymbol{H}$.这个空间可以是有限维空间,也可以是无穷维空间.分类问题的求解,依赖于对相似性和相似程度的估价.而在SVM中,相似性和相似程度是用内积进行估价的.这些内积强烈依赖于映射的选择,选择不同的映射就意味着对相似性和相似程度的不同估价标准.实际上,甚至不需要知道具体的映射是什么,只要选定函数$K(.,.)$就够了.将任意2个标本x和x'映射并内积,有一些常用的核函数,它们是:

1)径向基核(RBF,Radial Basis Function):

$$K(x,x') = \exp\left(-\delta \| x - x' \|^2\right)$$

2)多项式核:

$$K(x,x') = \left((x \cdot x') + c\right)^d$$

其中,$c \geqslant 0$;d是任意的正整数.

3)Sigmoid核:

$$K(x,x') = \tanh(v(x \cdot x') + c)$$

实验表明,由RBF构造的分类器,分类正确率较高,因此RBF是经常被选用的核函数.

2.3 C-支持向量机的训练算法

1)设已知训练集$T = \{(x_1, y_1), \cdots, (x_l, y_l)\} \in (\chi \times \gamma)^l$.其中,$x_i \in \chi = \mathbf{R}^n$;$y_i \in \gamma = \{1, -1\}$;$i = 1, \cdots, l$;

2)选择适当的核函数$K(x,x')$和适当的惩罚参数$C>0$,构造并求解最优化问题:

$$\min_{\boldsymbol{\alpha}} \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i \cdot x_j) - \sum_{j=1}^{l}\alpha_j$$

$$\text{s.t.} \sum_{i=1}^{l} y_i \alpha_i = 0 \quad C \geqslant \alpha_i \geqslant 0, i = 1, \cdots, l$$

得最优解$\boldsymbol{\alpha}^* = [\alpha_1^*, \cdots, \alpha_l^*]^{\mathrm{T}}$;

3)选择$\boldsymbol{\alpha}^*$的一个小于C的正分量α_j^*,并据此计算阈值$b^* = y_j - \sum_{i=1}^{l} y_i \alpha_i^* K(x_i \cdot x_j)$;

4)构造决策函数计算

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i^* y_i K(x,x_i) + b^*\right)$$

2.4 svm_light构造分类器

采用svm_light进行SVM分类,svm_light中采用了分解算法,针对固定工作样本集的方法,存在的主要问题有:①工作样本集的大小的确定;②如何确定工作样本集;③如何确定终止条件.因此,在工作样本集的选择上,svm_light中是沿着最速下降可行方向d,d仅含q个非零元素,由非零元素对应的q个优化变量构成工作样本集.已经证明了只要最速下降可行方向d存在,则用相应子集构成的子问题就可以进一步优化,而子问题的可行解也是原问题的可行解.这就解决了工作样本集不能包括所有支持向量的问题.

2.5 多类分类器

SVM构造出的是2类分类器:一类标记为1,

另一类标记为-1. 为了实现 n 个场景的识别,将某一类场景样本和所有非此类的样本作为训练样本集,构造与每一类场景相对应的 n 个分类器. 识别的过程是将待分类图像的特征向量代入这 n 个决策函数,将决策函数取得最大值的那一类,作为该样本对应的分类结果.

3　算法的 Matlab 实现

算法采用 Matlab7. 0 开发,运算实现对每幅图像的 24 次 Gabor 变换、48 维特征向量的提取,并利用特征训练 SVM 分类器,最后测试分类器的正确识别率. 该程序的功能实现由 3 大部分组成:①图像特征提取:打开图像,人工选择场景的类型,用 Gabor 变换,提取出每幅场景的 48 维特征向量,与选定的场景类型一起保存;②训练 SVM:每一类场景都对应一个 SVM 网络,选定一个场景类型,读取出包含有特征向量和场景类型的数据,用 svm _ light 训练出该类场景所对应的 SVM;实验数据表明,径向基核 RBF 训练出的分类器识别率明显较高;③识别图像:打开一幅场景图像,计算出特征向量,带入 n 个决策函数,判别出属于哪一类场景,并显示处理时间.

4　实验结果

好的训练样本集应具有相容性、遍历性和致密性及尽可能少的训练样本数,实验选择的训练样本和识别图像的比例为 7 :3. 实验中每幅图像的平均处理时间是 2 s.

实验 1 采用的是 4 幅图在一定范围的角度、尺度、光照的变化下拍摄,每张 50 幅共 200 幅图像,将其中的 140 幅图像作为训练样本,剩下的 60 幅图像作为测试图像. 实验的结果是 60 幅图像全部能正确识别,识别率为 100%.

实验 2 采用的是 ORL 人脸库,40 人每人 10 张照片,包括表情变化、微小姿态变化、20% 以内的尺度变化. 40 个人的 400 幅图像,选用 280 幅图像进行训练,120 幅进行识别. 有 5 幅图像未能正确识别,见图 6,识别率为 95. 8%.

由实验结果分析,分类器的设计导致识别的局限性:①大样本集和少类别数的情况下识别率较高;当样本数量有限,场景变化较大,或场景类别增多时,正确率下降;② 计算 SVM 分类器的核函数选择和参数设置对识别结果有很大影响.



图 6　未能识别的人脸图像

5　结 束 语

本文研究了基于 Gabor 函数目标识别算法,构造了多通道 Gabor 滤波器,提出了对场景图像和人脸图像的特征提取方法,可有效计算出图像的特征向量. 另外,构造 SVM 分类器,实现了多维特征向量的分类,初步的实验结果证明了方法的有效性. 传统的灰度匹配法依据像素逐个比较,计算量大且精度差,本文采用的算法是一种快速算法,且有较高的识别率,其识别率和实时性,能够保证其应用在移动机器人的视觉定位中识别出场景的类型,为机器人的全局路径规划提供可靠的数据.

参考文献 (References)

[1] Ulrich I, Nourbakhsh I. Appearance-based place recognition for topological localization [C]//Proc IEEE Int Conf Rob Autom. Piscataway, NJ: IEEE, 2000, 2: 1023-1029

[2] Zhou Chao, Wei Yucheng, Tan Tieniu. Mobile robot self-localization based on global visual appearance feature [C]//Proc IEEE Int Conf Rob Autom. Piscataway, NJ: IEEE, 2003, 1: 1271-1276

[3] Jason J A, White M W, Chase T A, et al. Pattern analysis for autonomous vehicles with the region-and feature-based neural network: global self-localization and traffic sign recognition[C]// Proc IEEE Int Conf Rob Autom. Piscataway, NJ: IEEE, 1996, 4: 3598-3604

[4] Bovic A C, Clark M, Geisler W S. Multichannel texture analysis using localized spatial filters[J]. IEEE Trans Pattern Anal Mach Intell, 1990, 12(1): 55-73

[5] Rui Y, Huang T S, Chang S F. Image retrieval: past, present, and future [J]. Journal of Visual Communication and Image Representation, 1999, 10 (1): 1-23

[6] 韦娜,耿国华,周明全. 利用 Gabor 滤波器的基于内容的图像检索[J]. 计算机工程,2005,31(8):10-11
Wei Na, Geng Guohua, Zhou Mingquan. Content-based image retrieval using Gabor filters [J]. Computer Engineering, 2005, 31(8): 10-11(in Chinese)

[7] 邓乃扬,田英杰. 数据挖掘中的新方法:支持向量机[M]. 北京:科学出版社,2004
Deng Naiyang, Tian Yingjie. New method of data mining: support vector machines[M]. Beijing: Science Publishing Company, 2004(in Chinese)

[8] 王晓丹,王积勤. 支持向量机训练和实现算法综述[J]. 计算机工程与应用,2004(13):75-78
Wang Xiaodan, Wang Jiqin. A survey on support vector machines training and testing algorithms[J]. Computer Engineering and Applications, 2004(13): 75-78(in Chinese)

(责任编辑:文丽芳)