

# 基于改进 Gabor 滤波器的多模态特征融合技术

陈支泽,荆晓远,陈 芸,朱阳平

(南京邮电大学 自动化学院,江苏 南京 210023)

**摘要:**传统的 Gabor 滤波器具有良好的方向特性和尺度特性,然而传统的 Gabor 滤波器不能提取图像中弯曲区域的局部信息。文中首先对传统的 Gabor 滤波器加以改进,使其在具有方向和尺度特性的同时具有良好的曲率响应特性,因而对于图像中弯曲的区域能够提取丰富的边缘特征。图像在不同的 Gabor 滤波器特征下有不同的表现形式,利用 Gabor 滤波器丰富的多特征信息,可以形成包含丰富信息的多个模态。然后文中提出一个多模态学习(Multi-modal Learning)框架。在此框架内,样本集合被投影到一个公共的鉴别空间内,在这个空间里,来自不同模态的同类样本相互聚集,异类样本相互散开。文中提出的多模态学习框架能很好地利用 Gabor 滤波器的多特征信息,PolyU 掌纹数据库和 AR 彩色人脸数据库的实验结果表明了该方法的有效性。

**关键词:**Gabor 滤波器;曲率特性;多模态学习;特征提取

**中图分类号:**TP301

**文献标识码:**A

**文章编号:**1673-629X(2015)10-0107-04

doi:10.3969/j.issn.1673-629X.2015.10.023

## Multi-modal Features Fusion Technology Based on Improved Gabor Filter

CHEN Zhi-ze, JING Xiao-yuan, CHEN Yun, ZHU Yang-ping

(College of Automation, Nanjing University of Posts and Telecommunications,  
Nanjing 210023, China)

**Abstract:** Traditional Gabor filter has good characteristics of direction and scale, but cannot extract the local information of bending area for image. Firstly, improve traditional Gabor filter to make it has good curvature response based on good characteristics of direction and scale. So for the image area can extract the edge of the rich characteristics of bending. After filtering with different characteristics of Gabor filter, images have more abundant characteristic information, and contain abundant information of multiple modes. Then propose a Multi-Modal Learning (MML) framework, within this framework, samples are projected onto a common space. In this common space, samples in same class from multiple modals are close to each other, while samples in different classes from multiple modals are far away from each other. Multi-modal learning framework proposed in this paper can make good use of Gabor filter characteristic information. Experimental results with PolyU palmprint database and AR color data set show the effectiveness of the method in this paper.

**Key words:** Gabor filter; curvature response; multi-modal learning; feature extraction

## 0 引言

为了提高生物特征图像<sup>[1-5]</sup>识别的鲁棒性,国内外学者进行了大量的探索,提出了很多有效的特征提取方法。Gabor 滤波器具有良好的方向特性和尺度特性,在特征提取方面得到了广泛的应用。Gabor 滤波器<sup>[6-9]</sup>具有两个优越的特性:

(1)可以同时时在域和频域获得最佳的局部化,因此能够获取丰富的局部信息;

(2)Gabor 滤波器是带通的,与人类视觉接收场模型相吻合。

Zhang 和 Kong 等提出的 PalmCode<sup>[10]</sup>使用一个方向的 Gabor 滤波器对掌纹图进行滤波,运用过零点准则对图像进行编码,由于 PalmCode 仅仅使用一个方向的 Gabor 滤波器,因此提取的信息较为单一。Com-Code<sup>[7]</sup>利用 Gabor 滤波器的多方向特性抽取掌纹的方向信息。传统的 Gabor 滤波器不能提取图像中局部的

收稿日期:2015-01-30

修回日期:2015-05-05

网络出版时间:2015-09-23

基金项目:国家自然科学基金资助项目(61272273)

作者简介:陈支泽(1989-),男,研究生,研究方向为生物特征识别;荆晓远,教授,博士生导师,研究方向为模式识别、图像与信号处理、信息安全、机器学习与数据挖掘。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150923.1506.042.html>

弯曲特征,对于如人脸面部的鼻子、眼睛等区域的弯曲处不具有很好的表征能力。基于此动机,文中对传统的 Gabor 滤波器进行改进,改进后的滤波器具有方向、尺度以及曲率三方面的特性。相对于传统的 Gabor 滤波器和局部描述子方法<sup>[11]</sup>,文中提出的 Gabor 滤波器所提取的信息更加丰富、鲁棒性更强。

子空间学习一直是特征提取技术中很重要的方法。传统的子空间学习方法如线性鉴别分析(LDA)<sup>[12]</sup>试图将原始样本集投影到新的空间,使得样本在新的空间里,同类样本集中,异类样本散开,但这些方法都只是在单一的空间内进行投影,没有利用到现实中样本的多模态信息。比如彩色图像的不同分量,多视图图像在不同视图下的不同形态,图像在不同特征下的不同表征形式,等等。近年来,一些多模态学习的方法<sup>[13-14]</sup>正在成为热点。Sharma 等<sup>[13]</sup>提出了一个一般性的多模态学习框架(GDA)。多元典型相关性分析(MCCA)<sup>[14]</sup>通过寻找一个空间,使来自各个模态间的相关性达到最大。

文中提出一种改进的 Gabor 滤波器设计方法,使其在具有方向和尺度特性的同时具有良好的曲率响应特性,因而对于图像中弯曲的区域能够提取丰富的边缘特征。图像在不同的 Gabor 滤波器特征下有不同的表现形式,利用 Gabor 滤波器丰富的多特征信息,可以形成包含丰富信息的多个模态。同时提出一个基于改进 Gabor 滤波器的多模态学习框架,将样本投影到一个公共的空间内,在这个空间里,来自不同模态的同类样本相互聚集,异类样本相互散开,从而更好地达到分类效果。

## 1 改进的 Gabor 滤波器

常用 Gabor 滤波器<sup>[5-6]</sup>的核函数形式为:

$$G(x, y, \theta, u, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x'^2 + y'^2}{2\sigma^2}\right\} \exp\{2\pi i(ux' + uy')\} \quad (1)$$

其中,  $x'$  和  $y'$  定义为:

$$\begin{cases} x' = +x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases} \quad (2)$$

此时式(1)变换为:

$$G(x, y, \theta, u, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\} \exp\{2\pi i(ux\cos\theta + uy\sin\theta)\} \quad (3)$$

其中,  $u$  可近似定义为  $u = 1/\sigma$ ;  $\sigma$  为高斯函数包络在  $x$  轴和  $y$  轴上的标准差,它们决定了高斯包络的空间扩展,反映了 Gabor 滤波器的尺度特性;  $\theta$  为正弦波的方向,根据  $\theta$  参数的不同,可以对图像提取出丰富

的方向信息。

使用式(3)与图像进行卷积便可以得到滤波之后的图像。图 1 为掌纹图像经过不同方向和尺度 Gabor 滤波器滤波后的示意图。

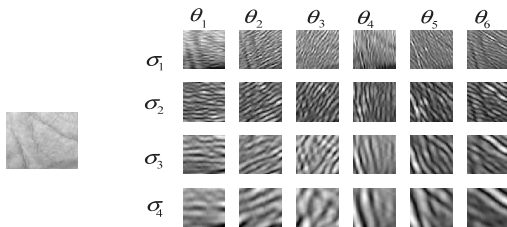


图 1 原始掌纹图的 Gabor 滤波示意图

对于给定的图像,可以想象其是由大量的非直线成分组成。从生物学的角度来看,大脑皮层对于卷曲的图像成分具有较强的刺激响应。基于此动机,文中对一般性的 Gabor 滤波器加以改进,使其具有良好的曲率响应。改进之后坐标变化:

$$\begin{cases} x' = +x\cos\theta + y\sin\theta + c(-x\sin\theta + y\cos\theta)^2 \\ y' = -x\sin\theta + y\cos\theta \end{cases} \quad (4)$$

其中,  $c$  是表征图像局部区域弯曲度的参数。

新的 Gabor 滤波器(Curvature Gabor, CGabor)不仅具有方向和尺度响应,而且具有良好的曲率特性。图 2 形象地表示了经过新的滤波器滤波之后的图像细节。在人脸的眼睛、鼻子等边缘弯曲较为明显的局部区域有强烈的响应。

通过对比多组实验(实验中  $\theta$  取三组方向  $0^\circ, 45^\circ, 90^\circ, 135^\circ$ ,  $\sigma_j = 0.259\ 2/\sqrt{2}^{j-1}$  ( $j = 1, 4$ )),发现在单一特征情况下,  $c$  取 0.1 能够达到较好的滤波效果。

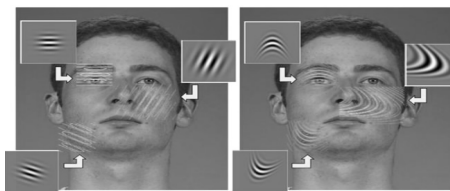


图 2 传统 Gabor 滤波器与改进后的效果对比

## 2 多模态学习框架

本节首先讨论提出的多模态学习(Multi-Modal Learning, MML)框架的思想,然后分析并给出其数学模型,最后给出求解过程。

### 2.1 多模态学习思想

传统的分类识别算法一般仅被运用于单一的形态之下,比如经典的子空间学习方法如 LDA 在单一形态下投影到子空间,而多模态学习强调在多特征、多视角的情况下进行学习,这样很好地利用图像在不同形态里所蕴含的丰富的特征信息。

在文中提出的多模态学习框架中,来自不同模态下的样本通过投影变化映射到统一的鉴别子空间中,

在这个子空间里,来自不同模态的同类样本相互聚集,异类样本相互分开。为了构造信息丰富的模态,巧妙地利用第一节所构造出的新的 Gabor 滤波器丰富的特征信息,可以构造出基于不同特征的模式。在每个模态内,固定其中一个特征,然后融合其他不同的特征信息可以构造出较之于原始图像特征信息更加丰富的单独模态。

## 2.2 数学模型

令  $x_{ijk}$  为第  $k$  模态内第  $i$  类的第  $j$  样本,其中  $k = 1, 2, \dots, \nu, i = 1, 2, \dots, c, j = 1, 2, \dots, n_i, y_{ijk} = \omega_j^T x_{ijk}$  为投影之后的样本,  $\mu$  为相应图像的均值。试图寻找投影变换  $\nu$ , 把初始样本投影到一个公共鉴别子空间,使得样本在这个公共空间内,来自所有模态的样本类内散度最小,类间散度最大。因此可以把模型描述为以下数学形式:

$$(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_\nu) = \arg \max_{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_\nu} \frac{\text{Tr}(\mathbf{S}_B)}{\text{Tr}(\mathbf{S}_W)} \quad (5)$$

其中,投影之后低维子空间的类内散度矩阵和类间散度矩阵分别定义为:

$$\mathbf{S}_W = \sum_{i=1}^c \sum_{j=1}^{n_i} \sum_{k=1}^{\nu} (y_{ijk} - \mu_i) (y_{ijk} - \mu_i)^T \quad (6)$$

$$\mathbf{S}_B = \sum_{i=1}^c n_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (7)$$

从上式可以看出,这里类内散度信息和类间散度信息都来源于多个模态,而不是单一模态,既考虑到了模态内的图像信息,又考虑到了模态间的图像信息。

## 2.3 MML 的求解分析

式(5)中的类内散布矩阵可以转化为:

$$\mathbf{S}_W = [\mathbf{w}_1^T, \mathbf{w}_2^T, \dots, \mathbf{w}_\nu^T] \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1\nu} \\ S_{21} & S_{22} & \dots & S_{2\nu} \\ \vdots & \vdots & \ddots & \vdots \\ S_{\nu 1} & S_{\nu 2} & \dots & S_{\nu\nu} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_\nu \end{bmatrix} = \mathbf{W}^T \mathbf{S}_W \mathbf{W} \quad (8)$$

其中,  $\mathbf{W} = [\mathbf{w}_1^T, \mathbf{w}_2^T, \dots, \mathbf{w}_\nu^T]^T$ 。这里  $S_{kr}$  定义为:

$$S_{kr} = \begin{cases} \sum_{i=1}^c \left( \sum_{j=1}^{n_i} x_{ijk} x_{ijk}^T - \frac{n_{ik} n_{ir}}{n_i} \mu_{ik} \mu_{ir}^T \right), & k = r \\ - \sum_{i=1}^c \frac{n_{ik} n_{ir}}{n_i} \mu_{ik} \mu_{ir}^T, & k \neq r \end{cases} \quad (9)$$

同理,类间散度矩阵可以进行如下形式的变换:

$$\mathbf{S}_B = [\mathbf{w}_1^T, \mathbf{w}_2^T, \dots, \mathbf{w}_\nu^T] \begin{bmatrix} D_{11} & D_{12} & \dots & D_{1\nu} \\ D_{21} & D_{22} & \dots & D_{2\nu} \\ \vdots & \vdots & \ddots & \vdots \\ D_{\nu 1} & D_{\nu 2} & \dots & D_{\nu\nu} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_\nu \end{bmatrix} = \mathbf{W}^T \mathbf{D} \mathbf{W} \quad (10)$$

这里  $D_{kr}$  定义为:

$$D_{kr} = \left( \sum_{i=1}^c \frac{n_{ik} n_{ir}}{n_i} \mu_{ik} \mu_{ir}^T \right) - \frac{1}{n} \left( \sum_{i=1}^c n_{ik} \mu_{ik} \right) \left( \sum_{i=1}^c n_{ir} \mu_{ir} \right)^T \quad (11)$$

综合式(8) ~ (11),式(5)可以转化为:

$$(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_\nu) = \arg \max_{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_\nu} \text{Tr} \left( \frac{\mathbf{W}^T \mathbf{D} \mathbf{W}}{\mathbf{W}^T \mathbf{S} \mathbf{W}} \right) \quad (12)$$

从此式可以看出,该问题的求解最终可以转化成求解特征值分解<sup>[12]</sup>的问题。

## 2.4 单模态数据集构造

文中利用改进后的 Gabor 滤波器的多特征信息,构造出每个单独的模态。具体方法为:在方向模态内,固定参数  $\theta$ , 然后调整  $c = 0, 0.1$  和  $\sigma_j = 0.259 \cdot 2^{\sqrt{2}^{-j-1}}$  ( $j = 1, 4$ ), 分别获得不同尺度和曲率下的图像特征,对于每种特征下的样本  $x_{ijk}^m$ , 进行特征融合从而形成新的特征向量  $((x_{ijk}^1)^T, (x_{ijk}^2)^T, \dots, (x_{ijk}^n)^T)^T$ ; 在尺度模态内,固定  $\sigma$ , 调整  $c = 0, 0.1$  和  $\theta = 0, \pi/3, 2\pi/3$ , 然后对各个参数组合下滤波后的样本进行特征融合; 在曲率模态内,固定  $c$ , 分别调整  $\theta$  和  $\sigma$  以获得不同的特征信息, 然后进行融合。通过以上步骤可以获得信息量更为丰富的单独模态。下面给出文中多模态学习框架的算法。

训练:

输入参数:数据集  $X$ , 类别标签  $c$

输出参数:投影矩阵  $\mathbf{W}$

(1) 利用 2.4 构造每个单独的模态;

(2) 根据式(8) ~ (11) 计算矩阵  $\mathbf{D}$  和  $\mathbf{W}$ ;

(3) 利用拉格朗日乘子法求解前  $d$  个最大特征值所对应的特征向量; 利用这些特征向量构建投影矩阵  $\mathbf{W}$ 。

测试:

(1) 将测试样本集投影到  $\mathbf{W}$  所在的投影空间;

(2) 运用最近邻分类器进行分类识别。

## 3 实验结果及分析

为了公平比较文中方法的有效性,选取 PolyU<sup>[7]</sup> 掌纹数据库和 AR<sup>[14]</sup> 彩色人脸数据库进行实验,这两个数据库中的图像均包含大量弯曲的局部区域。在 PolyU 数据库与传统的掌纹识别方法相比较,在 AR 彩色数据库上与传统的多模态方法比较。

PolyU 掌纹数据库有 3 780 张图像,这些图像来自 189 个人,每个人 20 张掌纹图像,这些掌纹图像包括光照,掌纹位置和掌纹纹理的变化,图像原始大小为 384×284。为提高计算效率,每张图像提取其中的有用部分,然后压缩为 60×60 大小。图 3 表示 PolyU 数据库中一个人的样本图像。在实验中随机选取其中 8

个作为训练样本(剩下的作为测试样本),并打乱样本的顺序运行20次,得到文中和对比方法的20次识别率结果,如图4所示。

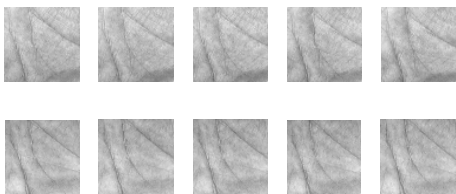


图3 PolyU 掌纹数据库图像示意图

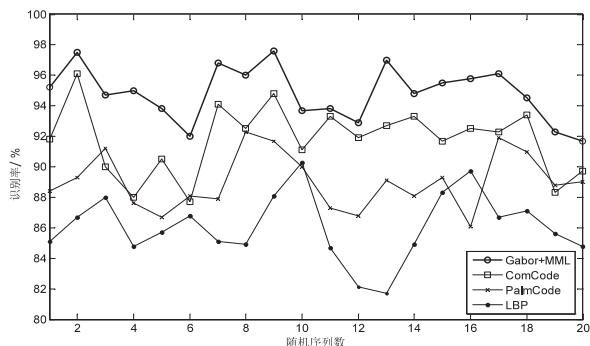


图4 PolyU 掌纹数据库实验识别率对比图

从图4可以看出,文中提出的CGabor+MML学习方法相对于其他三种方法能至少提高2.62。



图5 AR 彩色人脸数据库示意图

AR 彩色人脸数据库包含102个类,每类平均26个人脸图像。文中实验将原始的AR数据库中的图像压缩成60×60,图5为AR彩色人脸数据库其中一个人脸的图像。在AR彩色人脸数据库上,随机选取其中10个作为训练样本,剩下的作为测试样本,图6显示了该数据库的实验结果。

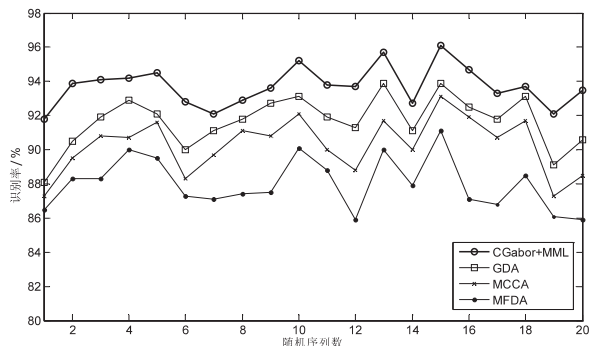


图6 AR 彩色人脸数据库实验识别率对比图

从图6可以看出,在AR彩色人脸数据库的实验对比中,文中提出的CGabor+MML学习方法相对于其他三种方法能至少提高5.93。

## 4 结束语

文中提出了一种改进的Gabor滤波器设计方法,新的滤波器对于图像中弯曲的区域能够提取丰富的边缘特征。同时提出一个多模态的学习框架,样本集合被投影到一个公共的空间内,使来自不同模态的同类样本相互聚集,异类样本相互离散,这样可以更好地达到分类效果。PolyU掌纹数据库和AR彩色人脸数据库的实验结果证明了该方法的有效性。

## 参考文献:

- [1] 赵振勇,王保华,王力,等.人脸图像的特征提取[J].计算机技术与发展,2007,17(5):221-224.
- [2] 尹飞,冯大政.基于PCA算法的人脸识别[J].计算机技术与发展,2008,18(10):31-33.
- [3] 王李冬.一种新的人脸识别算法[J].计算机技术与发展,2009,19(5):147-149.
- [4] 李强,裘正定,孙冬梅,等.基于改进二维主成分分析的在线掌纹识别[J].电子学报,2005,33(10):1886-1889.
- [5] 岳峰,左旺孟,张大鹏.掌纹识别算法综述[J].自动化学报,2010,36(3):353-365.
- [6] Kong A, Zhang D, Kamel M. Palmprint identification using feature-level fusion[J]. Pattern Recognition, 2006, 39(3): 478-487.
- [7] Kong A W K, Zhang D. Competitive coding scheme for palmprint verification[C]//Proc of international conference on pattern recognition. [s. l.]: IEEE, 2004: 520-523.
- [8] Pan X, Ruan Q Q. Palmprint recognition using Gabor-based local invariant features[J]. Neurocomputing, 2009, 72(7-9): 2040-2045.
- [9] Sun Z, Tan T. Ordinal measures for iris recognition[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31(12): 2211-2226.
- [10] Zhang D D, Kong W A, You J, et al. Online palmprint identification[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(9): 1041-1050.
- [11] Mu M, Ruan Q, Guo S. Shift and gray scale invariant features for palmprint identification using complex directional wavelet and local binary pattern[J]. Neurocomputing, 2011, 74(17): 3351-3360.
- [12] Belhumeur P, Hespanha J, Kriegman D. Eigenfaces vs. fisherfaces: recognition using class specific linear projection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- [13] Sharma A, Kumar A, Daume H, et al. Generalized multiview analysis: a discriminative latent space[C]//Proc of IEEE conf on computer vision and pattern recognition. [s. l.]: IEEE, 2012: 2160-2167.
- [14] Rupnik J, Shawe-Taylor J. Multi-view canonical correlation analysis[C]//Proc of international conference on data mining and data warehouses. [s. l.]: [s. n.], 2010: 1-4.



# 基于改进Gabor滤波器的多模态特征融合技术

作者:

[陈支泽](#), [荆晓远](#), [陈芸](#), [朱阳平](#), [CHEN Zhi-ze](#), [JING Xiao-yuan](#), [CHEN Yun](#),  
[ZHU Yang-ping](#)

作者单位:

[南京邮电大学 自动化学院, 江苏 南京, 210023](#)

刊名:

[计算机技术与发展](#) 

英文刊名:

[Computer Technology and Development](#)

年, 卷(期):

[2015\(10\)](#)

引用本文格式: [陈支泽](#). [荆晓远](#). [陈芸](#). [朱阳平](#). [CHEN Zhi-ze](#). [JING Xiao-yuan](#). [CHEN Yun](#). [ZHU Yang-ping](#) [基于改进Gabor滤波器的多模态特征融合技术](#)[期刊论文]-[计算机技术与发展](#) 2015(10)