

# 基于 Gabor 小波变换多特征向量的人脸识别鲁棒性研究

彭 辉

(浙江大学计算机学院 杭州 310018)

**摘 要** 传统的 Gabor 小波变换人脸识别技术在曲线奇异性的表达上存在着不足,难以识别包含表情的人脸信息,针对该问题,提出了结合 Gabor 小波变换和多特征向量的人脸识别算法。算法首先利用 Gabor 小波变换的频率及方向选择性来提取出人脸的多尺度、多方向上的 Gabor 特征,并组成联合稀疏模型,通过计算可以得到各个方向和尺度上 Gabor 特征的共同特征和表情特征,利用这两个特征向量可以精确重构测试图像的特征向量。仿真实验结果表明,所提出的方法能够有效提高带表情人脸图像的正确匹配率,改善识别效果。

**关键词** 人脸识别,表情识别,Gabor 小波变换,多特征向量

中图法分类号 TP391 文献标识码 A

## Research on Gabor Wavelet Transform Feature Recognition Robustness Based on Vector of Face

PENG Hui

(School of Computer, Zhejiang University, Hangzhou 310018, China)

**Abstract** There is insufficiency in expressing curve singularity for traditional Gabor wavelet transformation in face recognition technology that causes facial expression information hard to identify. This paper proposed a face recognition algorithm combining Gabor wavelet transform and multiple feature vectors. The algorithm firstly utilizes frequency and direction selectivity of Gabor wavelet transformation to extract the Gabor features of face multi-scale and direction and forms a joint sparse model in which the common features and expression characteristics of Gabor can be characterized in all directions and scales via calculation, at the same time, the test image feature vector can be accurately reconstructed using the two feature vector. Finally, the simulation results show that this method can effectively enhance the correct matching ratio of facial expression image and improve the recognition effect.

**Keywords** Face recognition, Facial expression recognition, Gabor wavelet transformation, Feature vector

## 1 引言

人脸识别技术是一项非常重要的生物特征识别技术,日常生活中我们对一个人身份的辨认,就是人脸识别的过程。我们就是要把人的这种自动识别功能转嫁于计算机,应用于人们的日常生活中,以为其提供更多的便利。它与指纹、虹膜、语音、DNA 等其他人体生物特征相比,更加直接、友好,且具有在识别过程中不需要进行接触操作、与人类的识别习惯相符、操作隐蔽特征等特点,所以在身份识别、访问控制、视频会议、档案管理、电子相册、基于对象的图像和视频检索等方面有着广泛的应用,是当前模式识别和人工智能领域的一个研究热点。研究人脸识别,不仅可以推动图像处理、模式识别理论及应用的发展,而且满足身份验证、基于内容的检索等实际需求,对推动认知科学、生理学、心理学等相关学科的研究也有积极的影响。

Yacoob<sup>[1]</sup>比较了基于统计的方法和基于几何特征的方法对表情人脸的识别结果,结果显示特征脸的方法更好一些。余冰<sup>[2]</sup>等提出了一种基于特征运动的人脸识别方法,该方法

在表情人脸识别方面优于特征脸方法。胡平、曹伟国、李华<sup>[3]</sup>等提出一种适用于三维人脸的等距不变量构造方法,但是同时也显示出很大的局限性。杜杏菁<sup>[4]</sup>针对姿态表情严重影响人脸识别准确率的问题,基于 Candide-3 模型的简化,提出了形状表情关键点拟合的人脸几何结构重建和基于三角网格模型的纹理映射的方法,该方法确定关键特征点,根据人脸的几何结构信息来确定姿态角,提取 Candide-3 模型形状表情对应点,调整模型参数,进行几何结构重建;对几何结构中每个三角网格模型进行纹理映射,得到较为逼真的特定人脸模型。近来还出现了一些其它的识别方法,如基于子空间的方法、基于形态模型的方法、基于光流模型的方法、基于共同向量的方法和基于 CS (Compressive Sensing) 理论的方法等。这些方法都对带表情人脸识别具有一定的鲁棒性,但是仍然存在一些不足之处。基于子空间的方法虽然结合了线性判别和非线性判别优势,但是当识别的种类增加时,识别复杂度会急剧上升,导致识别分类难度增加。基于形态模型的方法通过线性变形测试图像的形状达到与训练图像一致的效果,但是存在与人脸的非线性特征不符的情况。基于光流模型的方法存

到稿日期:2013-04-20 返修日期:2013-05-27 本文受浙江省高等教育课堂教学改革项目:《图形与图像处理》延伸型教学管理模式和多阶段考核方式的探索资助。

彭 辉(1980—),女,硕士,讲师,主要研究方向为图形与图像处理、数据恢复与数字取证、计算机硬件开发。

在计算量较大的问题。基于CS的方法是新兴的用于人脸识别的理论,在提取稀疏性方面具有较强的优势,对表情具有一定的鲁棒性。而之前提出的基于共同向量的方法都只是通过去除表情特征作为人脸的共同特征来用于识别,没有发挥表情特征的作用,所以识别效果并没有达到最优。

本文针对传统的 Gabor 小波变换人脸识别技术在曲线奇异性的表达上存在不足而造成难以识别包含表情的人脸信息的问题,提出了结合 Gabor 小波变换和多特征向量的人脸识别算法。算法首先利用 Gabor 小波变换的频率及方向选择性,提取出人脸的多尺度多方向上的 Gabor 特征,并组成联合稀疏模型,计算得到各个方向和尺度上 Gabor 特征的共同特征和表情特征,利用这两个特征向量可以精确重构测试图像的特征向量。最后仿真实验结果表明,本文提出的方法能够有效提高带表情人脸图像的正确匹配率,达到改善识别效果的作用。

## 2 人脸表情识别相关原理

人脸由眼睛、鼻子、嘴等部位构成,这些部位的形状发生改变时,人脸才会呈现出各种各样的表情。通常人的面部表情包括生气、厌恶、高兴、悲伤、惊讶和恐惧等,这些表情特征在非约束环境下的识别占据着很重要的地位。当采用的识别方法忽略这些表情特征或提取的表情特征不足时,在表情人脸库上的识别率将急剧下降。对于带表情人脸的识别,有两个方面很重要:1)提取的人脸特征能较好地表征图像中的表情信息;2)设计的分类器功能要强大。因为提取的特征越能精确地表征表情信息,产生的分类效果会越好;而分类器的功能越强大对于特征的要求越低,分类的正确率越高。因此,为了提高含有表情的人脸识别的效果,必须重视这两个方面的要求<sup>[8]</sup>。

Gabor 小波变换<sup>[5]</sup>是多尺度几何变换的一种,由 Gabor 变换与小波理论相结合所产生。传统的傅立叶分析中,信号完全是在频域展开的,不包含任何时域的信息虽然非常适用于平稳信号的分析,但对于非平稳信号的分析则显示出了其缺陷所在。这是因为处理非平稳信号时,不仅需要区分各种频率成分,还需要获得各个时刻附近的频率成分。于是人们尝试在傅立叶分析的基础上引入时域信息的方法<sup>[6]</sup>,把时域和频域分解为大小相等的小窗口,使得在这些“窗”内能够体现频率的信息,而窗与窗之间则反映频率随时间的变换,这样可以得到信号的局部频率信息<sup>[7]</sup>。将短时傅立叶变换的窗函数取成高斯函数,就成为了 Gabor 变换<sup>[9]</sup>。但传统 Gabor 变换的窗口尺寸一旦确定,在信号分析过程中就无法更改,即只能在一个分辨率上进行分析。为了解决这一问题,人们将 Gabor 变换与小波理论相结合,提出了 Gabor 小波变换,因此 Gabor 小波变换既具有小波变换的多分辨率特性,也具有 Gabor 函数本身所具有的局域性和方向性<sup>[10,11]</sup>。

一维 Gabor 小波变换的表达式:

$$C(x(t))(t_0, \omega) = \int_{-\infty}^{+\infty} x(t)W(t, t_0, \omega) dt \quad (1)$$

其中,  $W(t, t_0, \omega)$  为 Gabor 核函数(如图 1),是一个三角函数和一个高斯函数的乘积,即:

$$W(t, t_0, \omega) = e^{-\alpha(t-t_0)^2} e^{i\omega(t-t_0)} \quad (2)$$

将式(2)代入式(1),可以得到:

$$C(x(t))(t_0, \omega) = \int_{-\infty}^{+\infty} x(t) e^{-\alpha(t-t_0)^2} e^{i\omega(t-t_0)} dt \quad (3)$$

将上式展开得到:

$$C(x(t))(t_0, \omega) = \int_{-\infty}^{+\infty} x(t) e^{-\alpha(t-t_0)^2} \cos(\omega(t-t_0)) dt + i \int_{-\infty}^{+\infty} x(t) e^{-\alpha(t-t_0)^2} \sin(\omega(t-t_0)) dt \quad (4)$$

复数  $C(x(t))(t_0, \omega)$  即为信号  $x(t)$  经过一维 Gabor 小波变换得到的频率为  $\omega$ 、时间为  $t_0$  的频域信息。同时,这个复数也可以表示成实部和虚部两部分:

$$C(x(t))(t_0, \omega) = a_{real} + ia_{imag} \quad (5)$$

由此可得到  $C(x(t))(t_0, \omega)$  的极坐标表示形式,极坐标的幅值  $a$  和相角  $\phi$  为:

$$\begin{cases} a = \sqrt{a_{real}^2 + a_{imag}^2} \\ \phi = \arctan(a_{imag}/a_{real}) \end{cases} \quad (6)$$

另外,需要指出的是在时间轴上的平移距离和 Gabor 小波系数的相角有着很重要的关系:在时间轴上的两点非常接近时,它们之间的距离大致正比于两点小波系数的相角值之差。

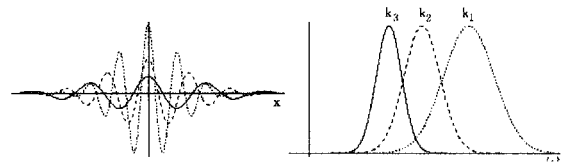


图 1 Gabor 核函数在空域和频域中的表示

## 3 Gabor 小波变换多特征向量

### 3.1 特征表示与提取

图像  $I(x, y)$  的 Gabor 滤波可以通过 Gabor 函数与图像  $I$  的卷积获得。对于图像中的一点  $z_0 = (x_0, y_0)$ , 经 Gabor 小波变换后得到图像特征表示如下:

$$T_{u,v}(z_0) = I(z_0) * \psi_{u,v}(z_0) = \int I(z) \psi_{u,v}(z_0 - z) d^2 z \quad (7)$$

其中,  $*$  表示卷积,  $T_{u,v}(z_0)$  表示该点处的卷积值, 即图像特征。为了获得图像多方向及多尺度下的局部显著特征, 通常取一组 Gabor 核函数对图像进行变换, 它们具有不同的方向参数  $u$  及尺度参数  $v(u=0, \dots, U+1; v=0, \dots, V-1)$ 。因此图像  $I$  的 Gabor 小波特征表示集合  $S = \{T_{u,v}(z) | u \in \{0, \dots, U-1\}, v \in \{0, \dots, V-1\}\}$ 。这里, 采用变换后的幅值作为人脸 Gabor 特征。之所以选择幅值作为人脸 Gabor 特征, 是因为在图像边缘附近, Gabor 系数的实部和虚部由于会产生振荡而不是一个光滑的峰值响应, 从而不利于识别。同时, 相位也会随着位置而发生改变, 在较小范围内像素点的相位差异很大, 也会对识别产生影响。而幅值信息则反映了图像的能量信息, 比较稳定, 所以比较适合用于表示人脸特征。对于尺度和方向的选择, 尺度既不能太大也不能太小, 否则都会影响识别效果。这是因为尺度太大时, 容易忽略图像的细节特征而使提取出的创新特征不能准确地表示表情信息; 尺度太小则会增加识别的运算复杂度。一般取 5 个尺度、8 个方向的 Gabor 滤波器时可以达到一个较好的识别效果; 图 2 为经过滤

波后的 Gabor 幅值图像。

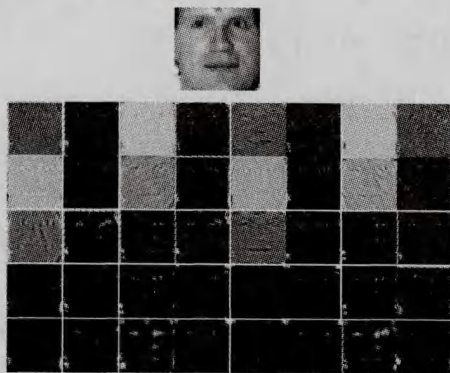


图2 样本图像及其 Gabor 幅值

由于图像库中一个人通常包含多幅带表情的图像,如果直接存储这些图像用于识别将会占据大量的存储空间,而且会增加识别的难度。为了解决这个问题,本文采用两个特征向量,即共同特征向量和创新特征向量来表示同一人的多幅人脸图像信息。其中,共同特征向量包含了该个体的整体面貌信息,而创新特征向量则包含了这些图像的所有表情信息。由于一幅图像经 Gabor 小波变换后得到 40 个 Gabor 幅值特征,为了从多幅图像中提取其稀疏的特征向量,考虑把同一个人的多幅带表情图像的相应 Gabor 特征级联组成一个公式化的联合稀疏模型进行计算,并提取其共同特征和创新特征,这样就分别得到 40 个共同特征和 40 个创新特征,单个 Gabor 特征的共同特征和创新特征的提取过程如下:

假如有  $C$  类,每类  $n$  幅图像,图像的大小为  $N_1 \times N_2$ ,则第  $k(k=1,2,\dots,C)$  类的第  $m(m=1,2,\dots,40)$  个 Gabor 特征可被看作一个整体  $\{g_{k,j}^m\} \in R^{N_1 \times N_2} (j=1,2,\dots,n)$ :

$$g_k^m = [g_{k,1}^m \ g_{k,2}^m \ \dots \ g_{k,n}^m], k=1,2,\dots,C \quad (8)$$

对于每一个  $g_{k,j}^m$  可分解为一个共同部分  $w_k^m$  ( $c$  代表 common) 和一个表示自身表情的创新部分  $w_{k,j}^m$  ( $i$  代表 innovation),即

$$g_{k,j}^m = w_k^m + w_{k,j}^m = E^T w_k^m + E^T w_{k,j}^m; w_k^m, w_{k,j}^m \in R^{N_1 \times N_2} \quad (9)$$

$$W_k = [w_k^m \ w_{k,1}^m \ \dots \ w_{k,n}^m] \quad (10)$$

由式(13)~式(15)得到公式化的联合稀疏模型,即:

$$g_k^m = \Phi W_k \quad (11)$$

其中,  $\Phi = [\Phi_1 \ \Phi_2]$ ,  $\Phi_1 = [E^T \ E^T \ \dots \ E^T]^T \in R^{(nN_1) \times N_1}$ ,  $E \in R^{N_1 \times N_1}$  是单位矩阵,  $\Phi_2 = \text{diag}(\Phi_1) \in R^{(nN_1) \times (nN_1)}$ 。利用 GPSR(Gradient Projection for Sparse Reconstruction) 算法求解出  $W_k$ , 得到的  $w_k^m$  与  $w_{k,j}^m (j=1,2,\dots,n)$  之和就是第  $k$  类第  $m$  个 Gabor 特征的共同特征和创新特征。为了得到第  $k$  类的共同特征和创新特征,把得到的 40 个 Gabor 特征的共同特征和创新特征进行级联。这种方法简单易用,识别效果也较好。但是这样得到的人脸特征向量的维数极高,处理起来比较困难,所以识别运算效率很低。假如有一幅大小  $64 \times 64$  的图像,经过 Gabor 小波变换后,提取得到的特征维数高达  $64 \times 64 \times 40 = 163840$ ,所以需要采用降维方法进行处理。常用的降维方法有下采样方法和 PCA 方法,但下采样存在图像对齐问题,而 PCA 是无监督的子空间方法,在样本最优重建方面具有优势,但降维后的特征却不能取得最优的分类效果。而 LDA 作为一种有监督的子空间方法,比较适合用于分类。

然而, LDA 也有一定的缺陷,如在实际应用中,样本的个数往往远小于样本的特征维数,类内散度矩阵  $S_w$  奇异,出现“小样本”问题。解决这个问题一般有 3 种处理方法:1)用  $S_w$  的伪逆代替  $S_w$  的逆,避免直接求解  $S_w$  的逆;2)用  $S_w + \sigma I$  ( $I$  为单位矩阵,  $\sigma$  为一个极小的整数)代替  $S_w$ ;3)采用 PCA+LDA 方法。本文采用最后一种方法,首先利用 PCA 把原始特征投影到低维空间,然后在降维后的子空间进行 LDA 分析,得到人脸特征向量即共同特征向量和创新特征向量。

为了便于识别,求出训练集中每一类的共同特征向量和创新特征分量,组成共同特征向量集  $\{V_k^c\} (k=1,2,\dots,C)$  和创新特征向量集  $\{V_k^i\} (k=1,2,\dots,C)$ ,从而可以用它们取代原始图像训练集用于分类识别,这样不仅节约了大量的存储空间,而且降低了识别难度,提高了运算效率。

### 3.2 分类识别

通过基于 Gabor 小波变换的共同向量法提取得到训练集中所有类的共同特征向量和创新特征向量,组成共同特征向量集和创新特征向量集。由这两个特征集就可以取代训练集图像进行测试图像的识别。共同特征向量由于包含了该类的本质特征,而且去除了表情特征的影响,相当于是对表情鲁棒的人脸特征表示,因此可以直接用于识别,利用最近邻分类法可以快速地判别出测试图像类别。

对于有  $C$  个类别、每类  $n$  个样本的训练集,测试图像  $x$  经 Gabor 小波变换和 LDA 降维后得到相同维数的特征向量  $v$ ,则其所属类别的判别函数为:

$$h_k(x) = \min_k \|v - V_k^c\|, k=1,2,\dots,C \quad (12)$$

其中,下标  $k$  表示类别。对于测试样本  $x$ ,只需比较其特征向量  $v$  与  $\{V_k^c\} (k=1,2,\dots,C)$  中的各类别的共同特征向量之间的欧氏距离,并决策与它最近的样本类别为所属类。虽然只利用共同特征向量就可以快速地进行识别,但这样做没有考虑把表情信息也用于识别的过程,因而表情变化较小的情况下识别效果较好,表情变化大的情况下则由于缺失的信息过多而导致识别效果不佳。因此,要结合共同特征向量和含有表情信息的创新特征向量进行识别,以此来实现较好的识别效果。

用每一类的共同特征向量  $V_k^c$  和创新特征向量  $V_k^i$  重构测试图像特征向量  $v$ ,  $v$  与第  $k$  类的共同特征向量  $V_k^c$  的差值为  $r_k$ :

$$r_k = v - V_k^c, (k=1,2,\dots,C) \quad (13)$$

然后从  $r_k$  和  $V_k^i$  中提取共同向量  $f_k$ , 重构  $v$ :

$$\hat{v} = V_k^c + f_k \quad (14)$$

重构误差最小的就是该测试图像所归属的类别  $\omega$ :

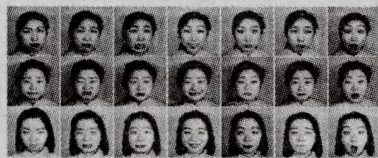
$$\omega = \arg \min_k (\|\hat{v} - v\|) \quad (15)$$

## 4 实验结果与分析

### 4.1 实验数据

本文提出的方法是基于 Gabor 小波变换的共同向量法,所以也可称为 CVBG(Common Vector method based on Gabor wavelet transform)方法。这里我们选取 3 个数据库进行实验:(1) JAFFE 人脸表情库;(2) CMU-AMP 人脸表情库;

(3) Yale 人脸表情库,如图 3 所示。JAFFE 库则共选取了 200 幅图像(10 个人,每人 20 幅),每个人分别取  $J=2,3,4,5$  幅图像,对每个  $J$  分别做 10 次实验。CMU 人脸库共包含 975 幅图像(13 个人,每人 75 幅),每个人分别取  $J=4,5,6,7,8$  幅图像,对每个  $J$  分别做 5 次实验。Yale 库共包含 165 幅图像(15 个人,每人 11 幅图像),每个人分别取  $J=3,4,5$  幅图像,对每个  $J$  分别做 5 次实验。同时,这 3 个库的图像均采用  $64 \times 64$  大小的图像。



(a) JAFFE 人脸表情库



(b) CMU 人脸表情库



(c) Yale 人脸表情库

图 3 实验数据库

#### 4.2 实验分析

本文主要将提出的 CVBG 方法与 BJSM 方法和 SRC 方法进行比较。SRC 方法是基于 CS 理论的稀疏思想提出的,它把所有样本图像作为一组基,通过求解最优化问题找到测试图像在该基上的稀疏表示形式。SRC 方法对光照和表情都具有较好的鲁棒性。从表 1 和表 2 得到,当  $J=5$  时 SRC 方法在 JAFFE 和 CMU 人脸库上的最高识别率分别达到了 98% 和 100%,由此可以证明稀疏性在人脸识别方面的有效性。SRC 方法虽然已经是当下比较有效的具有鲁棒性的识别方法,但是仍存在存储空间和运行时间上的局限性。之后提出的 BJSM 方法在这方面得到了改善,它是基于 DCT 变换的 CS 的方法,提取出的特征较 SRC 方法更具稀疏性,所以在两个图像库上当  $J=5$  时,最高识别率分别达到了 98.67% 和 100%。但是 DCT 变换是单一尺度的全局变换,因此在特征提取上稍显粗略,而小波变换的出现弥补了这个缺陷。Gabor 小波变换作为其中之一,继承了小波变换的所有优点,所以基于 Gabor 小波变换的 CVBG 方法提取的人脸特征更有效性,表征表情信息的能力更强,识别效果也更好一些。在 JAFFE 和 CMU 人脸库上,当  $J=5$  时,CVBG 方法的最高识别率均达到了 100%。另外,从图 4 和图 5 可以看出,在 JAFFE 库上 CVBG 方法的平均识别率分别比 BJSM 方法和 SRC 方法高出约 0.41% 和 2.24%。JAFFE 库和 CMU 库由于测试集中含有与训练集相同表情的图像,因此识别率普遍较高,而 Yale 库则由于测试集中不含与训练集相同表情的图像,因此识别率相对较低。但是,从表 3 中可知,CVBG 方法的识别效果仍好于其它两种方法。

表 1 CVBG、BJSM 和 SRC 方法在 JAFFE 库上的识别结果(单位:%)

J	CVBG			BJSM			SRC		
	High	Low	Avg	High	Low	Avg	High	Low	Avg
2	96.11	83.89	92.13	95.56	83.89	90.94	94.44	83.33	89.24
3	98.23	93.53	95.97	97.65	91.76	95.88	97.65	91.18	93.32
4	99.38	95.63	96.78	98.12	93.75	96.62	97.50	91.25	94.78
5	100	96.67	97.99	98.67	96.67	97.80	98.00	96.00	96.57

表 2 CVBG、BJSM 和 SRC 方法在 CMU 库上的识别结果(单位:%)

J	CVBG			BJSM			SRC		
	High	Low	Avg	High	Low	Avg	High	Low	Avg
4	100	99.24	99.67	100	97.94	99.50	99.89	97.50	98.41
5	100	100	100	100	99.67	99.93	100	98.90	99.80
6	100	100	100	100	100	100	100	99.22	99.87
7	100	100	100	100	100	100	100	98.30	99.72
8	100	100	100	100	100	100	100	99.89	99.93

表 3 CVBG、BJSM 和 SRC 方法在 Yale 库上的识别结果(单位:%)

J	CVBG			BJSM			SRC		
	High	Low	Avg	High	Low	Avg	High	Low	Avg
3	81.67	76.67	79.88	81.67	75.83	79.67	81.67	73.33	78.93
4	84.76	76.19	79.67	84.76	76.19	79.43	83.81	76.19	79.12
5	86.67	74.44	80.06	85.56	74.44	79.56	84.76	73.33	79.19

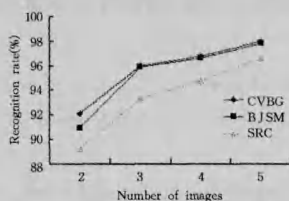


图 4 JAFFE 库上平均识别率

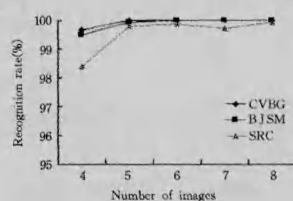


图 5 CMU 库上平均识别率

**结束语** 本文提出了结合 Gabor 小波变换和多特征向量的人脸识别算法。该算法首先利用 Gabor 小波变换的频率及方向选择性,提取出人脸的多尺度多方向上的 Gabor 特征,并组成联合稀疏模型,通过计算可以得到各个方向和尺度上 Gabor 特征的共同特征和表情特征,利用这两个特征向量可以精确重构测试图像的特征向量。最后仿真实验结果表明,本文提出的方法能够有效提高带表情人脸图像的正确匹配率,达到改善识别效果的作用。

#### 参考文献

- [1] Yacoob Y, Lam H, Davis L S. Recognizing faces showing expressions[A]//International Workshop on Automatic Face and Gesture Recognition[C]. Zurich, 1995: 278-283
- [2] Yu Bing, Jin Lian-fu, Chen Ping. Expression-invariant Face Recognition Based on Eigenmotion[J]. Journal of Image and Graphics, 2002, 7(A): 1140-1143
- [3] Hu Ping, Cao Wei-guo, Li Hua. A Novel Isometric Invariant and its Applications in 3D Face Recognition[J]. Journal of Computer-Aided Design & Computer Graphics, 2010, 22(12): 2090-2094
- [4] 杜杏菁. 基于 Candide-3 模型的姿态表情人脸识别研究[J]. 计算机工程与设计, 2012, 33(3): 1018-1020
- [5] Peng Hui. A method based on wavelet transform and SVM of ECG human identification [J]. Microelectronics & Computer, 2012, 3: 152-155
- [6] Hu D, Feng G, Zhou Z. Two-dimensional locality preserving projection with its application to palmprint recognition[J]. Pattern Recognition, 2007, 40(1): 339-342

(下转第 316 页)

从表 3 和表 4 可以看出,发髻的检索效率下降较快,检索效率最低,头冠的检索效率最高,僧帽的检索效率居中。发髻的检索效率低的原因是发髻与其他类型头饰的特征有一定的相似性。根据轮廓图求得傅里叶描述子特征,进行相似性排序,得到发髻的检索结果,如图 7 所示。其中,图 7(a)为检索原图,图 7(b)~(k)为按相似性排序的检索结果,图 7(b)和原图一样,排在最前面。从检索结果看,图 7(e)、(f)、(i)出现了检索错误,本应属于发冠。由表 3、表 4 和图 7 可知,基于内容的检索虽在一定程度上可以满足检索的要求,但检索的准确率还有待进一步提高。

表 3 轮廓图的检索结果

图像名称	头饰类型	前 10 幅相似率	前 20 幅相似率	前 30 幅相似率
阿弥陀佛	发髻	50%	25%	20%
白玛哈嘎拉	头冠	100%	100%	100%
宗喀巴大师 8	僧帽	100%	100%	80%

表 4 初始分割图的检索结果

图像名称	头饰类型	前 10 幅相似率	前 20 幅相似率	前 30 幅相似率
阿弥陀佛	发髻	60%	35%	23%
白玛哈嘎拉	头冠	60%	80%	67%
宗喀巴大师 8	僧帽	100%	70%	57%



图 7 基于内容的检索结果示意图

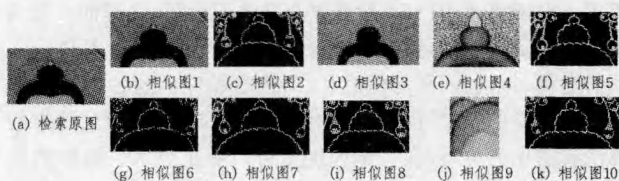


图 8 基于语义的检索结果示意图

为进一步提高检索效率,考虑语义检索。根据对唐卡内容的标注,可以检索出库中同类头饰的唐卡。此外,结合一定的特征,语义检索还可以实现各类唐卡的相似度排序。图 8 所示为语义检索时结合傅里叶描述子排序的结果,其中图 8

(a)为检索原图,图 8(b)~(k)为按相似性排序的检索结果。由实验结果可知,基于语义的检索比基于内容的检索效率高。

**结束语** 本文通过提取唐卡头饰的欧拉数、傅里叶描述子特征实现了头饰的自动分类和自动标注。设计的唐卡检索系统,不仅可以满足不同类型检索的需求,还为进一步研究其他特征(如手势、法器)的自动标注以及结合其他特征(如 Hu 距、zernike 距等)的基于内容的检索提供依据。除欧拉数、傅里叶描述子等特征外,颜色也可以作为一个特征加以考虑,这有待进一步研究。

## 参考文献

- [1] 王维兰,唐世喜,钱建军,等. 基于内容的唐卡图像数据库检索系统[J]. 湛江师范学院学报, 2008, 29(3): 91-95
- [2] Qian Jian-jun, Wang Wei-lan. Main feature extraction and expression for religious portrait Thangka image[C]//The 9th International Conference for Young Computer Scientists, 2008. Los Vaqueros Circle: IEEE Computer Society Press, 2008: 803-807
- [3] 钱建军. 基于语义的唐卡图像标注于检索研究[D]. 兰州: 西北民族大学, 2010
- [4] 王维兰,钱建军,杨且春,等. 基于频率谱变化量的唐卡图像特征提取与表示[J]. 计算机工程与应用, 2011, 47(22): 183-187
- [5] 李玲,尚文文. 基于迭代算法对 CT 肝脏感兴趣区域的提取[J]. 医疗卫生装备, 2010, 31(7): 29-32
- [6] 孙君顶,赵珊. 图像低层特征提取与检索技术[M]. 北京: 电子工业出版社, 2009: 92-93
- [7] Person E, Fu K S. Shape discrimination using Fourier descriptors[J]. IEEE Transactions on Systems, Man and Cybernetics, 1977, 7(3): 170-179
- [8] 潘崇,朱红斌. 改进 k-means 算法在图像标注和检索中的应用[J]. 计算机工程与应用, 2010, 46(4): 183-185
- [9] Kunttu I, Lepistö L, Rauhamaa J, et al. Multiscale Fourier descriptors for defect image retrieval[J]. Pattern Recognition Letters, 2006, 27(2): 123-132
- [10] Belongie S, Malik J, Puaicha J. Shape matching and object recognition using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522
- [7] Bartlett M S, Movellan J R, Sejnowski T J. Face recognition by independent component analysis[J]. IEEE Transactions on Neural Networks, 2002, 13(6): 1450-1464
- [8] 赵宏伟. 基于 PCA 针对表情变化的人脸识别研究[D]. 西安: 西安电子科技大学, 2009: 37-40
- [9] Shen L L, Bai L. Gabor wavelets and general discriminant analysis or face identification and verification[J]. Image and Vision Computing, 2007, 25(5): 553-563
- [10] He C, Dong J Y, Yuan F. Object tracking using the Gabor wavelet transform and the golden section algorithm[J]. IEEE Transactions on Multimedia, 2002, 4(4): 528-538
- [11] Candes E J, Donoho D L. New tight frames of curvelets and optimal representations of objects with C2 singularities[J]. Communications on Pure and Applied Mathematics, 2004, 57(2): 219-266
- [12] 吕翊,林贺宇,赵辉. 基于 sym8 小波和部分 hadmard 矩阵的深空图像压缩编码[J]. 重庆邮电大学学报: 自然科学版, 2012, 24(5): 646-651