

融合内外特征的性别分类^①

张景安, 张天刚

(山西大同大学 数学与计算机科学学院, 大同 037009)

摘要: 针对以往仅用人脸特征或头发特征来进行性别分类的片面性, 提出了将两类特征相融合的性别分类方法. 用对光照、尺度变化具有很强鲁棒性的 Gabor 小波变换提取人脸内部特征并用 PCA 方法降维. 利用最小代价原理, 将动态搜索技术用于图像空间取得头发区域, 定义了头发长度、头发表面积两种外部特征, 并提出了相应的特征提取方法. 采用模糊神经网络对三种特征进行非线性融合. 在 Essex 人脸库中进行了性别分类实验, 取得了 97.1% 的准确率.

关键词: 性别分类; Gabor 小波; 特征融合; 最小代价; 动态搜索; 模糊神经网络

Gender Classification Using Fusion of Facial Features and Hair Features

ZHANG Jing-An, ZHANG Tian-Gang

(School of Mathematics and Computer Science, Shanxi Datong University, Datong 037009, China)

Abstract: Aiming at the One-sidedness of face features or hair features to identify gender in the past, a novel method based on fusion of these features for gender classification was proposed. Face internal features were extracted by Gabor wavelet transform which are robust to the illumination change and scale variations, and feature dimensions were reduced by the method of PCA. Hair region was obtained by the method of dynamic searching in the area of face image. Two kinds of external features of hair length and hair surface area were defined and puted forward the method of corresponding feature extraction. To achieve nonlinear fusion of the three types of features with fuzzy neural network(FNN), the gender classification was completed in the Essex face database and a correct identification rate of 97.1% was obtained.

Key words: gender classification; Gabor wavelet; feature fusion; minimum cost; dynamic search; fuzzy neural network

近年来性别识别技术在计算机视觉及模式识别领域中已成为一个研究热点, 具有重要的理论研究、实用价值. 研究者们提出了多种识别算法及系统, 或者从图像提取人脸内部特征来进行性别识别^[1-5]、或者利用人脸面部以外的信息(如头发)来识别性别^[6]. 各研究虽然都取得了较高的识别率, 但分析有些片面, 没有全面把握图像信息, 即使是人, 也很难仅凭人脸面部特征或外部信息准确地判断出一个人的性别.

K.Ueki 等^[7]将人脸外部信息(头发, 衣着)和人脸面部特征相结合, 开发了一个性别分类系统, 但文中所用某些方法需进一步改进, 如文中使用 PCA 方法来提

取人脸内部特征及头发、衣着特征. 用 PCA 方法来提取人脸内部特征是十分有效的, 但用该法来提取头发、衣着等外部特征, 效果显不如前.

鉴于前述研究者的欠妥之处, 提出了融合内、外特征的性别识别方法, 即从图像依次提取人脸内部特征、头发特征, 两者融合起来, 进行性别识别. 众所周知, 这两类特征对于性别识别都至关重要.

1 人脸内部特征提取

在特征提取方面, Gabor 小波变换比其它方法具有优势, 1)其处理的数据量偏少, 能满足系统实时性要求;

^① 基金项目:山西省软科学研究计划(2014041049-1);山西省大同市政府专项研究项目(2014001)

收稿时间:2014-10-04;收到修改稿时间:2014-11-28

2)Gabor 变换对光照和尺度变化不敏感,能容忍一定程度的图像旋转和变形;3)具有良好的空间局部性和方向选择性,能精确地提取图像区域内多个方向的结构特征,具有较高的鲁棒性.故用二维 Gabor 小波变换提取人脸内部特征.

二维 Gabor 小波核函数为:

$$\psi_{u,v}(z) = \frac{\|K_{u,v}\|^2}{\sigma^2} e^{-\frac{\|K_{u,v}\|^2 \|z\|^2}{2\sigma^2}} [e^{iK_{u,v}z} - e^{-\frac{\sigma^2}{2}}] \quad (1)$$

式中 u 和 v 为 Gabor 核的方向和尺度, $z = (x, y)$ 为图像中给定位置的坐标,绝对值为模操作,波向量 $K_{u,v}$ 为滤波器的中心频率:

$$K_{u,v} = K_v e^{i\phi_u} = \begin{pmatrix} K_v \cos \phi_u \\ K_v \sin \phi_u \end{pmatrix},$$

其中 $K_v = K_{\max} / f^v$, $\phi_u = \pi u / 8$, f 是限定频域中核函数距离的间隔因子,通常取 $\sqrt{2}$, K_{\max} 为最大频率,通常取 $\pi/2$. ϕ_u 为滤波器的方向选择性,通常取 8 个不同的方向($u=0,1,2,\dots,7$)和 5 个不同的尺度($v=0,1,2,3,4$). 参数 σ 通常取 $2\pi \cdot \frac{\|K_{u,v}\|^2}{2\sigma^2}$ 用来补偿图像中由频率决定的能量谱衰减. 高斯包络函数 $e^{-\frac{\|K_{u,v}\|^2 \|z\|^2}{2\sigma^2}}$ 用来约束平面波. $e^{iK_{u,v}z}$ 为复数平面波,实部是余弦波,虚部是正弦波. 直流量补偿 $e^{-\frac{\sigma^2}{2}}$ 用来消除图像的直流成份对 Gabor 小波变换的影响.

Gabor 小波特征提取是通过计算图像与每个 Gabor 小波核函数(共 40 个)的卷积来实现的,其公式为:

$$G = \psi_{u,v} * I(x, y) = \iint \psi_{u,v} I(x, y) dx dy \quad (2)$$

式中, $I(x, y)$ 为对应坐标图像的像素值.

在特征提取中,截取人脸有效部分,并将图像归一为两眼间距为 72,大小为 128×160 ,由于采用 8 个不同的方向和 5 个不同的尺度,故整个图像区域向量的维数为 $128 \times 160 \times 5 \times 8 = 819200$,为了避免维数灾难、降低后续计算量,利用主成分分析(Principal Component Analysis, 简称 PCA)对上述特征向量进行降维,获得最佳人脸描述特征.

主成分分析法的理论与计算是较成熟的,但在解决实际问题中,主成分分析法的应用并没有达到较成熟状

态.需明确:何时使用主成分分析法?原始数据没有正向化,有何影响?如何正向化?

主成分分析法的优点是,对原始变量具有综合性的降维能力;如果系数的绝对值往 0 或 1 靠近得较多,即主成分命名、解释原始变量清晰,同时主成分解释原始变量的信息误差达到最小,使用主成分分析法最好.

主成分分析法是一种综合评价方法,是通过样品的相对位置,比较找出样品的优势、不足、差距状况及其原因,如果指标体系方向不是正向化的,便得不出有效结论.因此,分析中必须对指标体系中的强度逆向指标、适度指标进行正向化.

强度逆向指标 x_j 正向化公式:

$$\begin{cases} 1/x_j & x_j > 0 \\ 1/(\max|x_j| + x_j + 1) & x_j \leq 0 \end{cases} \quad (3)$$

适度指标 x_j 正向化公式: $1 / (|x_j - E| + 1)$, E 为理想值. 这里 x_{ij} 为第 i 个样品第 j 个指标的观测值.

2 头发特征提取

文献[8]定义了 6 种有效的头发特征:长度、表面积、对称度、颜色、分头、纹理,文献[6]提出了相应的特征提取方法,但方法较为简单.就人类视觉而言,判定一个人性别往往依据的是脸部特征和头发长度、头发表面积.故从 6 种头发特征中,选取两个典型的特征:长度、表面积,用于本文的性别识别.

提取有效头发特征的前提是进行头发边缘检测.在边缘检测中,最大的难题是:1)对于和脸部等高的头发区域,当头发颜色与背景颜色相接近时,边缘提取困难;2)对于头发的下面部分在与肩部、衣服重叠的区域,当头发颜色与背景颜色相近时,边缘提取愈加困难.无论头发颜色与背景颜色多接近,其分界线两侧像素的灰度值总是有所差异,故本文借助各像素的灰度值差异,利用最小代价通道^[9]来搜索头发的边界.

图像中任何边界是由若干边缘元素构成,像素 a 和 b 之间的边界用竖线(或横线)表示,示意如图 1.

由边缘像素 a 和 b 确定的代价函数为:

$$c(a,b) = H - [f(a) - f(b)] \quad (4)$$

其中, H 为图像中最大灰度值, $f(a)$ 、 $f(b)$ 为像素 a 和 b 的灰度值.该代价函数的值与像素间的灰度差值成反

比, 灰度差值越大则代价越小.

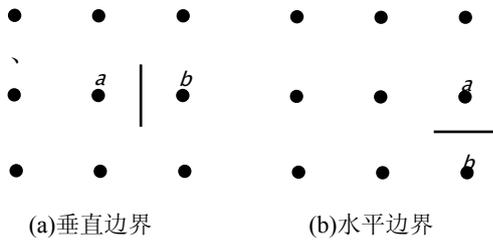


图 1 边缘元素及边界

利用最小代价原理, 将动态搜索技术用于图像空间可得到精确的头发边缘, 主要步骤如下:

1)将彩色图像转换为灰度图像, 方法采用文献[10]中的最佳灰度转换. 公式为:

$$Gray = (9798 R + 19235G + 3735B)/32768 \quad (5)$$

其中 R 、 G 、 B 分别代表图像中每一个像素点的红色分量、绿色分量和蓝色分量, $Gray$ 表示转换后的灰度值.

2)采用文献[11]中所用的方法旋转图像, 使两眼位于同一水平线上.

3)取过两眼连线中点的竖线将图像分为左右两部分, 并分别手工标定该竖线与头发上部内、外边缘的交点 Q 、 P .

4)在图像左半部分自点 P 开始, 借助动态搜索技术逐行向下搜索, 搜索过程中要保持各点的连通性, 得到左半部分头发的外边缘.

5)在图像左半部分自点 Q 开始, 重复步骤 4), 得到左半部分头发的内边缘.

6)在图像右半部分依次重复步骤 4)、5), 分别得到右半部分头发的外、内边缘.

四条曲线交于另外两点, 即取得头发的轮廓, 如图 2.



(a)原图 (b)头发边缘检测效果示例

图 2 最小代价通道搜索头发边界

头发区域确定后, 即可进行头发特征提取.

2.1 头发长度

通常情况下男性头发较短, 整个头发区域在耳部以上, 即使头发稍长者, 在图像上的体现也仅仅是小面积覆

盖耳部; 而女性则不然, 一般头发都较长, 通常全部覆盖耳部, 即使图像中外露耳部, 多数情况是头发向后扎起的缘故. 在图像中, 因头发长度覆盖水平、垂直及其弯曲时的三个范围, 导致计算头发的绝对长度存在一定的困难, 故取两耳连线下头发区域在垂直方向的长度值作为相对长度来作为性别识别特征之一, 相对长度与绝对长度对于本文识别效果等同.

多数情况下图像中耳部被头发遮挡, 致使图中无法做出两耳连线, 再者耳部与眼睛几乎等高, 故将求两耳连线下头发区域在垂直方向的最大长度值进一步转化为求两眼连线下头发区域在垂直方向的最大长度值 L_{Max} , 作为本文的头发长度特征. 当然 $L_{Max} >= 0$, 其值越小越接近男性.

2.2 头发表面积

现实中, 部分女性虽头发较长, 但向后上方扎起, 导致其在两眼连线下头发长度为 0, 该情况, 给人的直观感觉是: 头发区域所占的表面积较大; 对于男性, 恰好相反. 故本文把头发表面积也作为评判性别的另一特征.

头发边缘检测后, 得到不规则的头发区域在二维平面上的面积定义为头发表面积, 因头发内、外边缘很难用准确的函数式表达, 因而用定积分定义去近似求头发的表面积. 以两眼连线的中点为极点、以极点到右眼的射线为极轴建立极坐标系.

面积求法如下:

1) 取任一微元 $[\theta, \theta + \Delta\theta]$, 如图 3, 中心角 $\Delta\theta$ 的值尽量小, 该微元的面积为:

$$\Delta A = \frac{r_2^2 \Delta\theta}{2} - \frac{r_1^2 \Delta\theta}{2} = \frac{\Delta\theta}{2} (r_2 + r_1)(r_2 - r_1) \quad (6)$$

其中, r_2 为头发外轮廓的极径, r_1 为头发内轮廓的极径.



图 3 用于计算头发表面积的极坐标及参数

2) 扫描线从极角 $\theta = 0$ 出发, 利用式(5)求第一个

微元 $[0, \Delta\theta]$ 的面积 ΔA_1 .

3) 扫描线极角 θ 每次增加一个单位 $\Delta\theta$, 重复步骤 2), 直到极角 $\theta = 2\pi$, 得面积特征向量 $Surface=(\Delta A_1, \Delta A_2, \Delta A_3, \dots, \Delta A_N)$. $N = 2\pi/\Delta\theta$, 本文取 $\Delta\theta = \pi/90$, 即 $N=180$. 当然, 也可取 $N=360, 720 \dots N$ 值越大, 理论上计算面积更接近实际面积, 但经大量实验验证, 其值与 $N=180$ 时相接近, 再者, 计算量也翻倍增长, 故取 $N=180$, 来满足本文的头发面积特征提取.

4) 头发表面积近似值:

$$Area = \Delta A_1 + \Delta A_2 + \Delta A_3 + \dots + \Delta A_N \quad (7)$$

很明显, $Area$ 值越大, 越接近于女性.

3 模糊神经网络的特征融合

长度、面积与人脸特征不是同源数据, 也不具有线性相关性, 因而不能对其进行较底层的特征融合. 由于 Takagi-Sugeno 模糊模型^[12]具有良好的逼近能力, 能以任意精度逼近, 具有很强的可解释性, 鲁棒性和抗干扰性, 可以借助于线性系统的方法实现对非线性复杂系统的分析, 故本文采用模糊神经网络(T-S 模型算法)系统来进行性别识别.

T-S 模型算法简单, 易于工程化, 可以任意精度拟合非线性函数, 但在实现时仍存在问题: 隶属函数的调整比较困难, 不具备自适应性, 不能保证参数的最优性. 因此, 下面借助神经网络的自学习功能来解决以上问题. 设计出如图 4 所示的多输入单输出模糊神经网络系统.

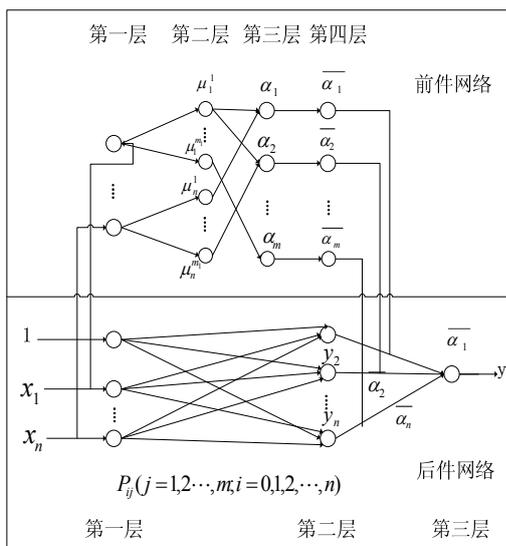


图 4 基于 Takagi-Sugeno 模型的模糊神经网络

由图 4 可见, 该网络由前件网络和后件网络两部分组成, 前件网络用来匹配模糊规则的前件, 后件网络用来产生模糊规则的后件, 即产生输出.

1) 前件网络

前件网络由四层组成. 第一层为输入层, 它的每个节点直接与输入向量的各分量 x_i 连接, 它起着将输入值 $x = [x_1, x_2, \dots, x_n]^T$ 传到下一层的作用. 该层的节点数为 $N_1 = n$.

第二层每个节点代表一个语言变量值, 作用是计算各输入分量属于各语言变量值模糊集合的隶属度 μ_i^j , 即:

$$\mu_i^j = \mu_{A_j^i}(x_i) \quad (8)$$

式中, $i=1, 2, \dots, n; j=1, 2, \dots, m_i; n$ 是输入量的维数; m_i 是 x_i 的模糊分割数.

第三层的每个节点代表一条模糊规则, 作用是用来匹配模糊规则的前件, 计算出每条规则的适应度, 即:

$$a_j = \min\{\mu_1^{i_1}, \mu_2^{i_2}, \dots, \mu_n^{i_n}\} \quad (9)$$

式中, $i_1 \in \{1, 2, \dots, m_1\}, i_2 \in \{1, 2, \dots, m_2\}, \dots, i_n \in \{1, 2, \dots, m_n\}$

第四层实现的是归一化计算, 即

$$\bar{a}_j = \frac{a_j}{\sum_{i=1}^n a_i} \quad (10)$$

$j=1, 2, \dots, m$

2) 后件网络

子网络的第一层是输入层, 它将输入变量传送到第二层. 输入层中第 0 个节点的输入值 $x_0 = 1$, 作用是提供模糊后件中的常数项.

子网络的第二层共有 m 个节点, 每个节点代表一条规则, 该层的规则是计算每一条规则的后件, 即

$$y_{ij} = p_{j0}^i + p_{j1}^i x_1 + \dots + p_{jn}^i x_n \quad (11)$$

$$j=1, 2, \dots, m; i=1, 2, \dots, r$$

子网络的第三层是系统的输出, 即

$$y_i = \sum_{j=1}^m \bar{a}_j y_{ij} \quad (12)$$

$$j=1, 2, \dots, m; i=1, 2, \dots, r$$

3) 学习算法

需要学习的参数主要是后件网络的连接权

$p_{ji}^k (j=1,2, \dots, m; i=1,2, \dots, n; k=1,2, \dots, r).$

设误差代价函数为:

$$E = \frac{1}{2} \sum_{i=1}^r (t_i - y_i)^2 \quad (13)$$

t_i 和 y_i 分别代表期望输出和实际输出. 参数 p_{ji}^k 的学习算法为

$$\frac{\partial E}{\partial p_{ji}^k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial y_{kj}} \frac{\partial y_{kj}}{\partial p_{ji}^k} = -(t_k - y_k) \overline{a_j x_i} \quad (14)$$

$$p_{ji}^k(l+1) = p_{ji}^k(l) - \beta$$

$$\frac{\partial E}{\partial p_{ji}^k} = p_{ji}^k(l) + \beta (t_k - y_k) \overline{a_j x_i} \quad (15)$$

式中, $j=1,2, \dots, m; i=1,2, \dots, n; k=1,2, \dots, r$

4 实验与分析

4.1 人脸库的选择

为了验证融合特征在性别分类中的高效性, 在 Essex 人脸图像库中进行了实验.

Essex 人脸库是由英国的 University of Essex 收集的, 库中有 72 人, 每人 20 张, 共 1440 幅彩色图像. 图像格式为 jpg, 男性、女性皆有, 图像头发保留完整, 照明没有太大变化, 头部有微小的转动、俯仰, 部分图像如图 5.



图 5 Essex 人脸库中男、女图像示例

4.2 实验结果对比

首先利用上述方法将彩色图像转换为灰度图像, 性别分类流程见图 6.

任选 8 位女性、8 位男性, 每人各 5 幅共 80 幅图像作为训练样本, 来训练模糊神经网络. 用训练好的模糊神经网络分类器对库中剩下的 1360 幅图像进行测试, 结果错判 2 人(即每人 20 幅图像, 共 40 幅), 正确率为 97.1%. 错判 2 人, 是由于图像库中有三个男性戴帽子, 其中一人有

胡须, 识别结果正确, 另两人误求其帽子的表面积当头发表面积的缘故.

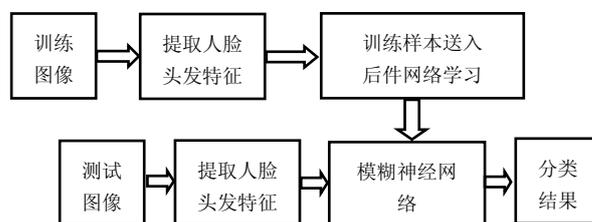


图 6 性别分类流程图

另外, 将融合特征与单独使用人脸内部特征、头发特征进行了对比, 所有特征均使用模糊神经网络分类器进行分类. 测试结果对比如图 7, 融合特征识别的准确率远高于单独使用人脸内部特征、头发特征.

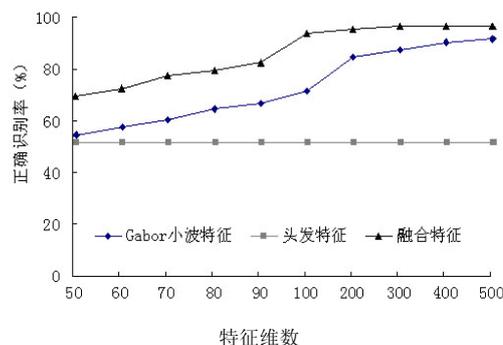


图 7 不同特征的分类准确率比较

从实验结果看到, 随着特征维数的增大, 识别准确率在逐渐上升, 当特征维数达到一定数值后(如融合特征为 100, Gabor 小波特征为 200), 上升缓慢. 由于本文提取的头发特征不涉及特征维数, 故头发特征识别率为水平直线, 与预期相符, 但基于头发特征的准确率远低于文献[6]取得的结果, 这是因为: 文献[6]采用 6 种头发特征去进行性别识别, 而本文只用到其中的 2 种头发特征.

5 结语

本文提出了融合人脸内部特征和头发特征的性别分类方法, 在 Essex 人脸库上的实验表明, 该方法具有良好的鲁棒性.

但文中还有诸多欠妥之处. 更为有效的头发特征提取方法, 高效的特征融合、压缩技术, 都需要进一步研究.

参考文献

- 1 Hayashi J, Yasumoto M, Ito H, Koshimizu H. Age and Gender Estimation based on Wrinkle Texture and Color of Facial Images. Proc. of the 16th Int. Conf. on Pattern Recognition. 2003. 405–408.
- 2 Iga R, Izumi K, Hayashi H, Fukano G, Ohtani T. A gender and age estimation system from face images. Proc. of the SICE 2003 Annual Conference. 2003. 275–281.
- 3 朱文球, 刘强. 融合 AdaBoost 和启发式特征搜索的人脸性别分类. 计算机工程, 2007, 33(2): 171–173.
- 4 孙鹤, 吕宝粮. 基于局部 Gabor 二值映射和 SVM 的性别分类. 计算机工程, 2009, 35(2): 210–213.
- 5 王振花, 穆志纯. 基于独立分量分析和遗传算法的人脸性别分类. 计算机科学, 2010, 37(4): 197–200.
- 6 刘爽, 谢金融, 吕宝粮. 基于头发信息的性别分类. 计算机仿真, 2009, 26(2): 212–215.
- 7 Ueki K, Komatsu H, Imaizumi S, Kaneko K, et al. A method of gender classification by integrating facial, hairstyle, and clothing images. Proc. of the 17th International Conference on Pattern Recognition. 2004, (4). 105–110.
- 8 Yacoob LD. Detection, Analysis and Matching of Hair. Proc. of International Conference on Computer Vision, Warsaw. 2004. 103–106.
- 9 章毓晋. 图像处理和分析. 北京: 清华大学出版社, 1999.
- 10 陆丽. 基于人脸图像的性别识别与年龄估计研究[博士学位论文]. 上海: 上海交通大学, 2010.
- 11 胡澜. 基于图像的年龄估计与人脸年龄图像重构[博士学位论文]. 长沙: 中南大学, 2007.
- 12 吴今培. 现代数据分析. 北京: 机械工业出版社, 2006.