

基于分叉树和 SVM 的人脸图像光照方向估计

孙雪梅 苏菲 蔡安妮

(北京邮电大学电信工程学院多媒体通信与生物特征识别实验室, 北京 100086)

摘要 根据人脸识别中光照方向类别多,类与类之间特征不明显,不易分类的问题,提出了采用 SVM 分类方法解决人脸识别中复杂的、非线性光照方向分类问题。用分叉树将复杂分类变成多级分类,并保证在 4 类以内的分类方法;同时在各级分类中,根据当前分类的类别特点,提取当前要处理的分类之间的明显特征作为特征进行分类;另外,给出了最优的 SVM 分类器的训练和构造过程。实验结果表明,该方法对解决分类特征不易于提取,类别数目多的分类问题有明显效果,分类准确率达到 89.16%。

关键词 分叉树 SVM 光照方向 人脸识别 非线性

中图分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2007)10-1885-04

Illumination Direction Estimation in Face Recognition Based on Bifurcate Tree and SVM

SUN Xue-mei, SU Fei, CAI An-ni

(School of Telecommunication Engineering, Beijing University of Posts and Telecommunications, Beijing 100086)

Abstract To solve the complex and nonlinear illumination direction classification problem in face recognition, a method based on bifurcate tree and SVM is proposed in this paper. Before using the SVM, light directions are firstly sorted into a bifurcate tree according to its intrinsic intensity distribution which assure that only a few classes are sorted at every branch. Moreover, different features are extracted for SVM at the different levels according to the lighting characteristics of the corresponding levels. Experiment results show that the proposed method is efficient to solve multi-classes classification problem and the accuracy can be achieved to 89.16%.

Keywords SVM, illumination direction, face recognition

1 引言

光照问题是人脸识别中具有挑战性的一个课题^[1]。由于光照引起人脸表面灰度的变化可能超过了不同人脸间原本的差别,因而给人脸识别造成相当的困难。到目前为止,对这个问题的解决方法主要集中于 3 个方面,即寻找对光照不敏感的特征;建立人脸在不同光照下的模型和建立光照变化模型,以及仿真人脸在不同光照下的变化情况。其中大多数方法都需要先确定光照方向,因此,判断图像

的光照方向是十分必要的。

支持向量机(SVM)是近十几年来发展起来的模式分类方法,在模式识别领域有广泛的应用。最初的 SVM 主要用于解决两类问题的分类。后来,许多学者改进了 SVM 的核心设计,采用几个二元分类器组合的方法^[2];或者运用有向无环图的 SVMS (support vector machines),进行多类分类^[3]。上述方法或者需要训练的 SVM 分类器个数太多,或者存在过拟合等问题。

由于光照变化对图像的影响是非线性过程,而 SVM 在高维空间构造最优超平面对非线性问题解

基金项目:国家自然科学基金资助项目(60472069)

收稿日期:2007-07-05;改回日期:2007-07-09

第一作者简介:孙雪梅(1973 ~),女,现为北京邮电大学通信与信息系统专业博士研究生。主要研究方向为图像处理与生物特征识别。E-mail: sxmxm123@263.net

决效果较好,因而本文选用 SVM 解决复杂光照条件下人脸图像的光照分类问题。同时,为了避开其不足,本文引入分叉树概念,将光照方向根据一定的规则逐级分类,每次分类的数目在 4 类以内;同时根据分类的类别,提取相应特征。实验结果表明,本文提出的方法可以有效地解决复杂的多类分类问题,在对人脸图像的光照方向分类中有效率可达到 89% 以上。

2 算法描述

2.1 光照方向分叉树

摄像机轴向和图像水平轴构成的平面内的光照方向变化称为水平方向变化,摄像机轴向与图像垂直轴构成的平面内的光照方向变化称为垂直方向变化。本文选用 Yale B 人脸数据库,选取水平方向光照变化范围为 $-130^\circ \sim 130^\circ$,垂直方向上变化范围为 $0^\circ \sim 90^\circ$ 的正面人脸图像组成数据集 A,并以此设计分叉树结构。

分叉树的设计原则如下:

(1) 每级分叉的分支之间要有较大的区分特征;

(2) 每级分叉时使该级分类结果在 4 类以内;

(3) 分叉时将易混淆的类别分到同一子集中;

(4) 由作者前面研究可知,水平方向偏角大于 60° 或垂直方向大于 45° 的图像在识别前需要作光照补偿处理。因此,偏光角度小于这个范围的图像不应在分叉树中进行多级分叉;

(5) 水平方向偏光角度较小,垂直方向偏光较大的图像主要受垂直方向偏光的影响,因此,此时水平方向的偏光对图像的影响忽略不计,将这类图像归为单一偏光源类,不再进行水平方向的分类;

(6) 因为偏离摄像机轴线的光源使图像上靠近光源一侧慢慢变亮,另一侧则变暗,左侧和右侧的光源对图像的影响截然不同,因此,细分光照方向之前先完成判断光源在图像哪一侧的分类;

(7) 由于水平方向上同样间隔的偏光变化比垂直方向对识别影响大,所以对水平和垂直方向都有偏角的光照方向分类时先以水平方向为主,分类角度间隔不应太大,以减少错分类对光照补偿的影响,然后再对分类结果按垂直方向偏角进一步分类。

按上述原则,数据集 A 被分为如图 1 所示结构。

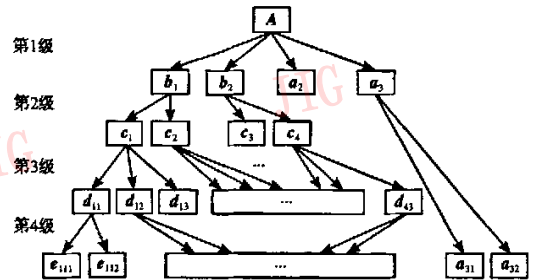


图 1 人脸图像光照方向分叉树结构图

Fig. 1 Structure of bifurcate tree

第 1 级 不需要光照补偿的子集 a_2 的图像,即垂直光照方向小于 45° ,且水平方向小于 60° 的图像;水平偏角小于 60° ,垂直偏角大于 45° 的图像单独归类为子集 a_3 ;水平偏角大于 60° 且垂直偏角大于 45° 的,归于子集 b_1 ;最后是水平偏角大于 60° 且垂直偏角小于 45° ,归为 b_2 。

第 2 级 将 b_1, b_2 按光源在人脸图像的左侧或右侧进一步分为子集 c_1, c_2 和 c_3, c_4 ,其中, c_1 和 c_3 是从左侧照射的图像, c_2, c_4 是从右侧照射的图像。

第 3 级 水平方向以大约 20° 为间隔, c_1, c_2, c_3, c_4 各自被分为 3 个子集,如图 1 所示。这样每类内图像受到的偏光影响较为相似;另外,可能发生分类错误的边缘方向在被错分时落入附近类别的可能较大,这样按照被错分类的补偿方法对该图像进行补偿不会对识别效果影响太大。

第 4 级 对第 3 级的每个子集和子集 a_3 按垂直方向偏角分类,每 25° 为 1 类,则每个子集只包含两类,于是第 4 级分类后得子集 $e_{111} \dots e_{432}, a_{31}, a_{32}$;

最后得到一个分叉分类树,与其他 SVM 多类分类器体系相比,分叉树结构简单,对于 k 类分类问题至多构造 $k-1$ 个分类器,所以分类速度较快。

2.2 各级分类特征

由于图像的灰度值分布特征在一定程度上反映了光照角度变化情况,根据光照方向分叉树结构,每级分类是根据该级各类的特点提取不同的分类特征。

2.2.1 第 1 级分类特征

这一级中提取 4 种特征为分类特征。

(1) 第 1 种,整个图像的灰度平均值, $G_0 = \sum_{i=1}^M \sum_{j=1}^N g(i, j)$, 其中 $g(i, j)$ 为第 (i, j) 点的像素值, M 是图像行数, N 为图像列数;

(2) 将图像平均分为 16 块,以每块的灰度平均

值为第 2 种分类特征, G_1, \dots, G_{16} ;

(3) 统计整个图像的灰度直方图作为第 3 种特征, 并把每 16 个灰度级合为一个灰度级, 得到 16 个灰度直方图分类特征, H_0, \dots, H_{15} ;

(4) 由于垂直方向偏角较大的图像鼻子与双眼的灰度值明显不同, 按人脸五官自然结构划分双眼和鼻子区域, 提取鼻子分别与左眼、右眼区域的平均灰度差 $V_{n,l} = |G_n - G_l|^2$ 和 $V_{n,r} = |G_n - G_r|^2$ 为第 4 种分类特征, 以区分 a_3 类, 其中 G_n, G_l 和 G_r 分别为鼻子、左眼和右眼的平均灰度值。

2.2.2 第 2 级分类特征

由于光源从摄像轴两侧的照射对图像的影响截然不同, 因此这一级分类不需用 SVM, 由 V_L 和 V_R , 将 b_1 分成 c_1, c_2 ; b_2 分成 c_3, c_4 。

$$V_L = \sum g(i, j) \quad i = 0, \dots, M; j = 0, \dots, N/2 - 1 \quad (1)$$

$$V_R = \sum g(i, j) \quad i = 0, \dots, M; j = 0, \dots, N/2 - 1 \quad (2)$$

式中, V_L 是以图像轴线为中心左半部分的灰度平均值, V_R 是轴线右半部分的灰度平均值; 比较 V_L 和 V_R , 则值大的部分对应于光源所在侧。

2.2.3 第 3 级分类特征

第 3 级分类细分水平偏角为 $60^\circ \sim 130^\circ$ 的图像。随着偏角的增加, 图像上大部分区域的灰度平均值越来越小, 较少区域的灰度平均值越来越大; 所有区域的纹理层次差别越来越小; 而整体的灰度平均值保持下降趋势。因此, c_1, c_2, c_3 和 c_4 提取 3 种特征作为分类特征。

(1) 第 1 种特征, 整体灰度平均值 G_0 ;

(2) 将图像平均分为 16 块, 以每个分块灰度平均值为第 2 种特征: G_1, \dots, G_{16} ;

(3) 计算每个分块的方差作为第 3 种特征: $F_0, \dots, F_{15}, F_m = (g(i, j) - G_m)^2$, 其中, $g(i, j)$ 是第 m 块中第 (i, j) 点的像素值, G_m 是第 m 块的灰度平均值。

2.2.4 第 4 级分类特征

这一级分类以垂直方向每 25° 为间隔, 对第 3 级中得到的子集和 a_3 子集分类。因为垂直方向的光偏对人脸图像造成的影响直观地反映在双眼部分随偏角的增加而变暗, 鼻子部分则相对变亮, 所以提取下列特征进行分类:

(1) 鼻子部分的平均灰度值 G_n , 左眼部分的平均灰度值 G_l , 右眼部分的平均灰度值 G_r ;

(2) 鼻子与左眼部分的灰度值差的平方 $V_{n,l} = |G_n - G_l|^2$, 鼻子与右眼部分的灰度值差平方 $V_{n,r} = |G_n - G_r|^2$;

(3) 左眼部分与右眼部分灰度平均值之差的平方 $V_{r,l} = |G_r - G_l|^2$ 。

3 SVM 的训练和构造

本文选用 Yale B 和 Harvard 库。因为本文的研究主要是对光照方向分类, 因此只选所有人的 640 张正面人脸图像作为实验数据库 C。并和 Harvard 库中 7 个人共 462 张图像组成实验库。

对于特定问题, 选取哪种 SVM^[3] 核函数及其参数, 目前没有较好理论方法为依据, 最优参数选择只能凭经验、大范围搜寻进行寻优。本文根据 libsvm^[4] 软件提供的默认参数, 在向上、向下放缩 1 000 倍的范围内搜索, 逐步缩小范围, 最后确定识别率最高时对应的参数为最佳参数。

下列 4 个核函数是常用到的:

(1) 线性核函数: $K(x_i, x_j) = x_i^T x_j$

(2) 多项式核函数: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$

(3) 径向基函数: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$

(4) Sigmoid 函数: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$

这里, r, d, γ 是核参数, r 默认值为零。

为了选择合适的核函数, 将 C 中 yale01-05 和 Harvard 中任意 4 个人的所有图像作为训练集, Harvard 中剩余图像为评价核函数性能的测试集, 提取特征作为 SVM 训练和测试的矢量数据, 对 64 种光照方向分类。则用上述 4 种核函数训练 SVM, 得到核函数的最佳设置参数, 如表 1 所示。

表 1 4 种核函数的最佳参数

Tab. 1 Best parameters of four kinds of kernel

核函数	最佳参数		
	第 1 级	第 3 级	第 4 级
线性核函数	$C=10, \gamma=1.2$	$C=3, \gamma=0.3$	$C=5, \gamma=0.5$
多项式核函数	$C=80, \gamma=0.015, d=3$	$C=180, \gamma=0.3, d=3$	$C=280, \gamma=0.6, d=3$
径向基函数	$C=20, \gamma=0.02$	$C=5, \gamma=0.016$	$C=15, \gamma=0.8$
Sigmoid 函数	$C=5, \gamma=2$	$C=20, \gamma=0.03$	$C=30, \gamma=0.08$

根据参数构建 4 种 SVM 分类器。为测试其性

能,分别用它们对光照方向分类,实验数据见表 2。由表 2 可以看出,核函数是径向基函数时效果最好。

表 2 4 种核函数的测试结果

Tab. 2 Testing results of four kinds of kernel

级数	准确率(%)			
	线性核函数	多项式核函数	径向基函数	Sigmoid 函数
第 1 级	90.204 1	91.428 6	93.571 4	90.625 6
第 3 级	96.923 1	94.285 7	96.923 1	95.714 3

4 算法实验结果及评价

用错分率和错分距离评价算法。定义如下:

错分率 = 错分图像个数/测试图像个数;

水平错分距 $D_h = \sum |\alpha_{h1} - \alpha_{h2}|/P$;

垂直错分距 $D_v = |\alpha_{v1} - \alpha_{v2}|/P$;

错分距为错分图像所属类与被错分类的平均方向角度之间的差距。其中, α_{h1} 是该图像真实的水平方向光照偏角, α_{h2} 是图像被分成的水平平均光偏角, α_{v1} 是图像真实的垂直方向光偏角, α_{v2} 是图像被分成的垂直平均光偏角, P 为错分图像个数。错分率反映分类错误率,错分距说明错分的严重程度。错分距小,即图像被错分到附近类,光照补偿后对识别影响较小。

用 C 中剩余图像为测试集评估算法,并与一对一方法比较,两种算法的训练图像和提取特征相同。在一对一方法中测试集分成 8 类:不需光照补偿为第 1 类,与本文算法第 1 级对应;水平偏转方向每隔 20° 为一类,共 6 类,与本文算法第 2 和第 3 级对应;水平偏转小于 60° ,垂直偏转大于 45° 的为第 8 类。实验结果见表 3。

表 3 测试集中图像的光照方向分类结果

Tab. 3 Classifying results of illumination directions

		1 级	2 级	3 级	4 级	总正确率 (%)
本文 算法	正确率 (%)	93.57	100	96.92	97.14	
	水平错分距 ($^\circ$)	40	0	15	10	89.16
	垂直错分距 ($^\circ$)	10	0	20	20.5	
— 对 —	正确率 (%)	78		89.62	90	
	水平错分距 ($^\circ$)	16		50	0	75.0
	垂直错分距 ($^\circ$)	37.2		20	40	

由表 3 可知,两种方法相比,本文算法比一对一算法有明显提高。其原因在于一对一算法的正负类样本数量不对称,影响了分类器的训练效果;且对不同类别但相邻的图像,可能几个分类器的输出值都接近最大值而导致误判。而本文算法将光照方向按最近类别归类,使每次分类时可以针对差别最大的几种特征分类,避免多类分类时明显特征在不同类别之间交叉跨越问题,使分类更有针对性。同时,由于容易分错的方向大多是处于类别边缘的图像,即使分错所属类别,建立在分叉树结构下的分类,也使分类的错分距较小,从而减少了错分类对光照补偿的影响。

5 结 论

本文采用基于分叉树和 SVM 的光照方向估计方法,通过建立分叉树,将复杂的多类分类变成简单的多级分类,并减少每级分类个数,使分类时可以针对每级的特点提取不同的分类特征。实验结果表明,这种方法较好地解决了人脸识别中复杂的、非线性光照方向分类问题,对分类特征不易于提取,类与类之间特征交叉跨越的分类有明显效果。关于改善 SVM 核的设计,进一步提高 SVM 多类分类能力的工作将是今后研究的主要内容。

致 谢 本文中应用了 Yale B 和 Harvard 人脸库。

参考文献 (References)

- Guo G, Li S Z, Chan K. Face recognition by support vector machines [A]. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition [C], Hawaii, USA, 2001: 511 ~ 518.
- Pontil M, Verri A. Support vector machines for 3-D object recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(6): 637 ~ 646.
- Vpanik V. The Nature of Statistical Learning Theory [M]. New York: Springer Verlag, 1995.
- Chang C C, Lin C J. LIBSVM: a library for support vector machine [EB/OL]. <http://www.csis.ntu.edu.tw/~cjlin/libsvm>. 2001.