

文章编号:1001-9081(2010)04-0960-04

隐马尔可夫后处理模型在视频人脸识别中的应用

代毅,肖国强,宋刚

(西南大学计算机与信息科学学院,重庆400715)

(hekuheku@21cn.com; gqxiao@swu.edu.cn)

摘要:现有的多数人脸识别系统都专注于如何提高人脸识别算法的性能,但缺乏一种对数据源(人脸样本)进行分析和评估的机制。针对此问题,提出了一种建立在数据源分析基础上对典型人脸识别算法进行后处理的方法。为了揭示现有典型识别算法的识别性能在无约束环境下的鲁棒性,通过建立 Lambertian 反射模型和 3D 人脸模型,对特征脸算法的识别性能随数据源的变化(人脸姿态和光照改变)而变化的情况进行了分析评估。针对“数据源灾难”问题,提出了一种基于隐马尔可夫模型(HMM)的后处理方法,该方法通过利用视频序列图像的连续性和对训练人脸库的统计分析来提高判别分析方法对无约束环境的鲁棒性。实验结果表明,该方法可以有效地提高识别算法对“数据源灾难”的鲁棒性,提高识别率。

关键词:隐马尔可夫模型;人脸识别;后处理;混淆矩阵;先验模型;置信度

中图分类号:TP391.41 **文献标志码:**A

Application of post-processing based on HMM to video face recognition

DAI Yi, XIAO Guo-qiang, SONG Gang

(College of Computer and Information Science, Southwest University, Chongqing 400715, China)

Abstract: In this paper, the rarely concerned problem of data source in face recognition was investigated, and a novel post-processing HMM-based solution was proposed. Data source problem was firstly empirically investigated through systematically evaluating the eigenfaces sensitivity to variations of pose and illumination by Lambertian reflection model and 3D face model, which revealed that the changes of pose and illumination abruptly degrade the eigenfaces system. This problem is explicitly defined as "data source disaster" for highlighting its significance. Aiming at solving this problem, combining the recognition rate with the analysis of the data sources, two methods were proposed to evaluate the overall performance of specific face recognition approach with its robustness against the low-quality data sources considered. Finally, a post-processing method was proposed to improve the robustness of the recognizer under unconstrained environment. The experimental results have impressively indicated the effectiveness of the proposed post-processing solution to tackle the "data source disaster" problem.

Key words: Hidden Markov Model (HMM); face recognition; post-processing; confusion matrix; priori model; confidence

0 引言

在视频序列中,人脸目标的变化具有帧间连续性,比如姿态或光照的改变。当这种改变造成人脸样本的变化较大时(姿态旋转过大或者光照不足等),所获得的样本包含的识别信息非常少,甚至发生人脸检测错误,最终容易导致人脸识别的失败——本文称这一现象为“数据源灾难”问题,并将引起“数据源灾难”的样本称为“低质量数据源”。现有的许多人脸识别方法都专注于如何提高人脸识别算法的性能,在理想情况下实验结果已经达到了较高的精度。然而在非约束情况下,识别性能受到了挑战。因此,关于人脸识别鲁棒性的研究成为一大热点。

为了克服非约束条件下姿态和光照变化对识别的影响,研究人员提出了诸多策略。一类策略是把样本集归一化到统一的条件下^[1-3]。其中除了基于 2D 的方法以外,研究人员对基于 3D 的方法进行了大量探索。在国外研究中,文献[2]中提出对已知人脸图像进行 3D 重建的思想,在一定程度上

提高了系统的鲁棒性。文献[3]中提出一种利用可变形人脸模型的形状和纹理参数进行 3D 人脸重建的方法,在虚拟人脸基础上进行人脸识别。遗憾的是,此类方法需要迭代求解大量的 3D 数据,在数据量大的情况下是难以忍受的。在国内研究中,文献[4]中采用了更为方便、实用的球面谐波商图像方法来处理光照估计和光照影响消除问题,这些措施极大地降低了算法的计算复杂度。另一类策略是提取对姿态和光照不敏感的特征向量。文献[5]中提出基于 Bit-Plane 的方法,能够有效提高人脸识别的鲁棒性。该方法从人脸图片中抽取轮廓特征和纹理特征,然后将这两种特征进行融合,得到一张虚拟人脸,最后进行主成分分析(Principal Components Analysis, PCA)以达到人脸识别目的。文献[6]把主动形状模型(Active Shape Model, ASM)得到的人脸配准的结果结合到线段 Hausdorff 距离(Line Hausdorff Distance, LHD)的边缘图提取过程,对人脸不同部分的线段给予不同的权重,并且把线段对的匹配限制在给定的邻域内。实验结果表明,改进的 LHD 方法在微笑和小角度姿态变化下的识别率有明显提高。

收稿日期:2009-10-10;修回日期:2009-12-30。

基金项目:国家大学生创新实验计划项目(081063510);重庆市科委自然科学基金资助项目(2008BB2252)。

作者简介:代毅(1984-),男,四川自贡人,硕士研究生,CCF 会员,主要研究方向:数字图像处理;肖国强(1965-),男,重庆人,教授,博士,主要研究方向:信号与信息处理、无线网络通信;宋刚(1983-),男,四川乐山人,硕士研究生,主要研究方向:嵌入式系统。

上述方法都有一定局限,该局限来自于“低质量数据源”所包含的信息较少或者信息错误。上述方法都试图从“低质量数据源”中恢复出更多的信息以提高识别性能,缺乏一种对数据源进行分析和评估的机制。本文系统地分析了“数据源灾难”对人脸识别系统的影响,使用隐马尔可夫模型(Hidden Markov Model, HMM)描述了视频中人脸识别的后处理,将视频的连续性和训练集的图像观测两个随机过程结合起来,有效地应用于“数据源灾难”的估计,从而提高人脸识别的性能。

1 “数据源灾难”问题的分析

1.1 光照对人脸识别系统的影响

光照变化对许多人脸识别算法的性能都有很大影响^[7],本文以子空间算法为例,对光照变化是如何影响人脸图片在特征空间的投影系数这一问题进行了分析。假设人脸样本的光照反射面为 Lambertian 表面,一幅含目标人脸的图片 I 同目标人脸 $(p; q)$ 之间的关系可以用变反射率的 Lambertian 反射模型来表示^[8]:

$$I = \rho \frac{1 + pP_s + qQ_s}{\sqrt{1 + p^2 + q^2} \sqrt{1 + P_s^2 + Q_s^2}} \quad (1)$$

其中:目标 $(p; q)$ 和反射率 ρ 都为位置 (x, y) 的函数^[9]。光源位置用与入射光线相关的一个 slant 角和一个 tilt 倾角来表示^[7],其中 slant 角 α 表示逆入射光线方向同 z 轴正方向之间的夹角; tilt 倾角 τ 表示逆入射方向同 x - z 平面之间的夹角,同 P_s 和 Q_s 的关系为: $P_s = \tan \alpha \cos \tau$, $Q_s = \tan \alpha \sin \tau$ 。一个无限远距离的点光源表示为 $(P_s; Q_s; -1)$ 。为了便于分析,本文使用 K 表示项 $\sqrt{1 + P_s^2 + Q_s^2}$,并假设所使用的人脸样本是基于人脸中垂线对称的正面人脸。

人脸图片 I 可以用子空间成分表示为^[10] $I = I_A + \sum_{i=1}^m \alpha_i \Phi_i$,其中 I_A 表示平均图片, Φ_i 为特征向量, α_i 为投影系数。选择某一特定的人在标准条件下(即式(1)中, $P_s = 0$; $Q_s = 0$)的正面人脸 I_p 作为原型,通过将 I_p 与一个新的人脸样本 \tilde{I} 对比来达到分析目的,其中 \tilde{I} 为在光照 $(P_s; Q_s; -1)$ 下,与 I_p 属于同类的人脸样本。 I_p 对应的子空间系数 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_m]^T$ 和 \tilde{I} 对应的子空间系数 $\tilde{\alpha} = [\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_m]^T$ 可表示为:

$$\alpha_i = I_p \ominus \Phi_i - I_A \ominus \Phi_i \quad (2)$$

$$\tilde{\alpha}_i = \tilde{I}_p \ominus \Phi_i - I_A \ominus \Phi_i$$

其中 \ominus 表示求两向量的内积。若把人脸图像和特征脸基于中轴线分成左右两部分,则有:

$$\alpha_i = I_p^L \ominus \Phi_i^L + I_p^R \ominus \Phi_i^R - I_A \ominus \Phi_i \quad (3)$$

$$\tilde{\alpha}_i = \tilde{I}_p^L \ominus \Phi_i^L + \tilde{I}_p^R \ominus \Phi_i^R - I_A \ominus \Phi_i$$

式(3)中上标 R 和 L 分别代表左边和右边部分。注意到人脸和特征脸都具有左右对称的性质,结合式(1),得到:

$$\alpha_i = 2I_p^L[x, y] \ominus \Phi_i^L[x, y] - I_A \ominus \Phi_i \quad (4)$$

$$\tilde{\alpha}_i = \left(\frac{2}{K} \right) (I_p^L[x, y] + I_p^L[x, y]q^L[x, y]Q_s) \ominus \Phi_i^L[x, y] - I_A \ominus \Phi_i$$

最终推导出下面的关系式:

$$\tilde{\alpha} = \left(\frac{1}{K} \alpha \right) + \frac{Q_s}{K} [f_1^a, f_2^a, \dots, f_m^a]^T - \frac{K-1}{K} \alpha_A \quad (5)$$

其中 $f_i^a = 2(I_p^L[x, y]q^L[x, y]) \ominus \Phi_i^L[x, y]$ 和 α_A 为平均图片 I_A 的投影系数。式(5)表明:光照改变会明显地影响基于子空间算法的性能。

1.2 姿态对人脸识别系统的影响

在分析姿态变化是如何影响人脸识别系统性能这一问题中,需要收集大量的不同姿态的样本,实际上这是非常困难的。如图1所示,通过正面的中性人脸图片,本文使用三维可变形模型^[2-3]较精确地恢复了人脸的三维模型,通过旋转三维模型能够得到不同姿态的人脸样本。图2所示为一组正面人脸样本与不同姿态人脸之间的特征脸距离^[11]。在曲面的中心,人脸的旋转角度为0,因此它们之间的特征脸距离也为0。随着人脸姿态的改变,特征脸距离也发生了明显的变化:如果人脸向右旋转10°,相应的特征脸距离为10.4%;如果向右旋转20°,特征脸距离变为14.6%。

对于人脸识别来说,系统倾向于不同类人脸间的类间距离更大,而同类人脸间的类内距离则更小^[12]。从图2中可以看出,姿态变化会增大类内距离,从而造成分类器的错误输出。下一章所介绍的后处理模型,将对“低质量数据源”进行分析和估计,以获得更好的识别结果。



图1 原始图片以及其3D人脸模型

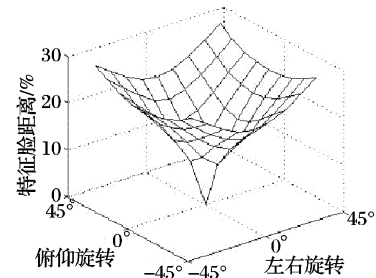


图2 原人脸与各姿态人脸之间的特征脸距离

2 基于HMM的后处理模型

人脸识别系统框架如图3所示。

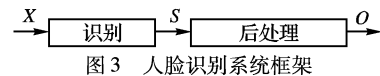


图3 人脸识别系统框架

设 $X = x_1, x_2, \dots, x_{len}$ 为一个待测人脸视频序列,即观测值序列,其中 x_i 表示在输入序列 X 中的第 i 帧图像, len 表示视频序列长度。

$S = s_1, s_2, \dots, s_{len}$ 为人脸识别器的输出人脸序列(每个输出 s_i 为前 K 个候选样本之一),即状态序列,所有的候选样本组成状态集 C 。即 $x_i \rightarrow s_i \in \{c_{ij} | j = 1, 2, \dots, k\}, \{c_{ij} | j = 1, 2, \dots, k\}$ 为 s_i 的候选子集,其中 $c_{ij} \in C$ 。 $O = o_1, o_2, \dots, o_T$ 为后处理器的输出序列。

每一个 s_i 包含 K 个候选样本,则共 K^{len} 种组合。后处理的任务就是从这 K^{len} 种组合中选择最大可能输出。这是典型HMM的第二个问题:选择最佳状态转移序列问题。这里:

状态转移概率分布矩阵 $A = (a_{ij})$, $a_{ij} = P(S_j | S_i)$ 。

观察值概率分布矩阵 $B = (b_j(k))$, $b_j(k) = P(X_k | S_j)$ 。

初始状态分布 $\pi = (\pi_1)$, $\pi_i = P(q_1 = S_i)$ 。

基于HMM的后处理过程如下:

$$O = \arg \max_s P(S|X) = \arg \max_s P(S) * P(X|S) \quad (6)$$

其中:先验概率 $P(S)$ 描述人脸序列的统计概率分布; $P(X|S)$ 描述人脸的观测图像概率分布,表示图像观测这一随机过程,由人脸识别模型决定。

典型 HMM 中的模型参数由 Baum-Welch 算法^[13] 迭代得到,但是它需要大量的观测值序列进行训练。在视频人脸识别后处理中,图像观测序列是极其有限的,因此 Baum-Welch 算法在这里不适用。实际上, A 矩阵和初始状态分布 π 可以通过将概率求解问题转换为能量函数求解问题来解决。 B 矩阵的元素为条件概率,难以求解。但是,通过下面的理论分析,条件概率可转化为后验概率,从而解决了 B 矩阵求解问题。

由于人脸识别器对每一帧孤立的人脸图像进行识别,这种识别显然不依赖于前后帧的关系。故式(6)中的条件概率 $P(X|S)$ 可表示为:

$$P(X|S) = P(x_1 \cdots x_{len} | s_1 \cdots s_{len}) = \prod_{i=1}^{len} P(x_i | s_i) \quad (7)$$

$P(x_i | s_i)$ 为单样本的条件概率(HMM 中的 B 矩阵参数), s_i 为对应的候选样本 c_{ij} 之一($j = 1, 2, \dots, 10$), 所以

$$P(X|S) = \prod_{i=1}^{len} P(x_i | s_i) = \prod_{i=1}^{len} \frac{P(s_i | x_i) * P(x_i)}{P(s_i)} \quad (8)$$

这里, $P(x_i)$ 对 x_i 的每个候选样本 c_{ij} 是相等的, $P(c_{ij})$ 为人类出现的先验概率,在人脸识别器中假定各类出现的概率是相等的。故在实际计算中, $P(x_i)$ 与 $P(c_{ij})$ 这两项可不考虑,式(7)代入式(6)中可得:

$$O = \operatorname{argmax}_s P(S) * \prod_{i=1}^{len} p(s_i | x_i) \quad (9)$$

其中 $P(s_i | x_i)$ 代表了候选样本 c_{ij} 的置信度。本文使用逻辑回归模型来对其进行估计。

这样,利用 HMM 就能将视频人脸识别的后处理问题转化为求解最大后验概率估计问题。但是,求解后验概率属于组合优化问题,直接求解难度很大。通常使用最大期望算法(Expectation-Maximization, EM)、模拟退火等优化算法估计该概率。

2.1 混淆矩阵

假设训练集中包含 M 类人,每人 L 帧图像, N 为训练集的大小,即 $N = M \times L$ 。当输入某一帧待测样本时,在训练集中所对应正确的人脸类为 w_i ; w_j 表示经人脸识别器得到的识别结果(候选集中的第一序列为 w_j), $w_i, w_j \in M$ 。对应于一个特定人脸识别器的易混淆矩阵可以表示为: $P_M = (p_{ij})_{M \times M}$ 可用频度近似求得:

$$p_{ij} = n_{ij} / n_i; \quad n_j = \sum_{i=1}^n n_{ij}$$

其中 n_{ij} 表示在训练集中 w_i 被人脸识别器分类为 w_j 的次数。

在一般情况下,人脸的类内距离比类间距离小,一类人脸被误识为错误类的次数是有限的。因此 P_M 具有以下特点:1) 当 w_i 与 w_j 为同类时,其 p_{ij} 值较大;2) 当 w_i 与 w_j 为不同类时,即表明类间差异大于类内差异,其发生的概率较小,因此 p_{ij} 值较小,甚至为 0。

因为混淆矩阵 P_M 可以经过大量的训练得到,所以可把它看作此人脸识别器的先验知识,利用此信息能够提高系统的性能。

2.2 先验概率 $P(S)$

直接获取一个候选集的先验概率是十分困难的,实际应用中,对先验模型的研究往往转化为对能量函数的研究^[13]。

为此,一个序列 $S = s_1, s_2, \dots, s_{len}$ 的先验概率使用如下的能量函数来描述:

$$U(S) = \exp\left(-\sum_{i=1}^{T-1} \omega(s_i, s_{i+1})\right) / Q \quad (10)$$

其中: $\omega(s_i, s_{i+1})$ 为 P_M 中所对应的元素; Q 为归一化常数。从式(10)可以看出,当 s_i 和 s_{i+1} 为同类时, $\omega(s_i, s_{i+1})$ 取值较大,也就是说,其相应的先验概率较大;反之,当 s_i 和 s_{i+1} 不为同类时, $\omega(s_i, s_{i+1})$ 取值较小,其相应的先验概率就较小,这遵循了视频中的序列具有连续性的规律。

2.3 置信度计算

c_1, c_2, \dots, c_k 为对应一幅人脸图片 x 的前 K 个候选样本,它们的距离值分别为 d_1, d_2, \dots, d_k 。逻辑回归模型^[14] 可以直接将候选集的距离值转换为它们的置信度:

$$p(c_k | x) = \left(1 + \exp\left(\beta_0^k + \sum_{i=1}^y \beta_i^k d_i\right)\right)^{-1}; \quad 1 \leq k \leq K \quad (12)$$

其中: β_i^k 为回归系数,通常是通过分析训练集的识别结果,利用最大似然概率估计(Maximum Likelihood Estimation, MLE)对它进行估计, y 为回归模型的序号。在本文中, d_i ($1 \leq i \leq K$) 被归一化为 0 ~ 100 的值。

实验使用了平均识别率为 85.4% 的 480 组样本集去估计回归系数 β_i^k 。例如,对于第一候选样本,可以得到:

$$p(c_1 | x) = \left(1 + \exp(-0.647 + 0.43d_1 - 0.303d_2 - 0.01d_3)\right)^{-1} \quad (13)$$

式(13)表示了当 d_1 为一个比较小的值时(而 d_2 和 d_3 相对较大),第一候选样本的置信度更高。候选集其余子序列的置信度满足以下公式:

$$p(c_k | x) = \left(1 + \exp(\beta_0^k + \beta_1^k d_1 + \beta_k^k d_k)\right)^{-1}; \quad \beta_1^k < 0, \beta_k^k > 0, k \geq 2 \quad (14)$$

式(14)表示当 d_1 较小(而 d_k 相对较大)的情况下,候选子序列的置信度更小。

3 实验与分析

研究基于视频的人脸识别技术的挑战之一就是缺少一个完整的数据库^[15]。文献[16]中建立的数据库包含样本较少,并且主要只包含姿态的变化。文献[17]中描述的数据库仅包含了在固定光照下较大角度的姿态变化,以及在固定姿态下光照的变化。对于我们的实验来说,一个理想的测试数据库是类似于 CMU_PIE^[18],但是离散的姿态被连续的视频序列所取代,并且在一段视频中存在着不同类的人脸交替出现。因此本文从 BBC NEWS 中收集到满足上述要求的一共 24 段视频作为实验数据库,部分视频的信息如表 1 所示。在实验数据库中一共包含了 12 类人脸,表 2 显示了每类人脸在视频中的显示情况。选择库中的 6 段视频作为训练视频,8 段视频用于混淆矩阵的训练,剩下的 10 段视频作为测试视频。

3.1 训练阶段

实验的训练阶段包括以下两个步骤。

1) 构建训练样本库。使用人脸检测和跟踪方法提取人脸图像^[19],尺寸归一化为 64×64 像素。每类人抽取 60 幅图像加入训练集,即 $L = 60$,一共 12 个人,训练集的总大小为 $N = 720$ 。

表1 实验测试库

视频序列	内容	长度	帧数	镜头数	帧率/(f/s)	尺寸
新闻1	News of Science	24'33"	36 825	154	25	CIF(352×288)
新闻6	Weekly news	27'16"	40 900	95	25	CIF(352×288)
新闻11	News of Economic	16'12"	24 300	102	25	CIF(352×288)

表2 人脸在视频中的显示

人物	视频序列	包含该人物镜头数/总镜头数
人物1	新闻1	24/154
人物2	新闻1	45/154
人物3	新闻1	34/154
人物4	新闻1	22/154
人物5	新闻1	8/154
人物6	新闻6	4/95
人物7	新闻6	7/95
人物8	新闻6	14/95
人物9	新闻6	22/95
人物10	新闻11	7/102
人物11	新闻11	8/102
人物12	新闻11	20/102
镜头总数		215/1456

2) 构建混淆矩阵。如1.1节所述,将待测图像输入人脸识别器进行计算得到识别结果,构建训练集的混淆矩阵作为此识别器的先验知识。

3.2 测试阶段

在测试视频序列中,每隔 h 帧提取人脸样本,组成待测人脸序列 X 。选择经识别器分类后获得的前 K 个样本,构成候选集 S 。可以注意到,在视频序列较长的情况下,构成候选集所需时间复杂度较高。为了提高效率,实验采用图像检索技术中的vocabulary tree^[20],将训练集建成树状结构。当视频序列 X 较长时,实验中通过设置滑动窗口将其拆分为较短的序列。假设原视频序列长度为 w ,取窗口宽度为 h ,步长 l 。每次后处理时窗口往前滑动 l 步长,得到一个新的长度为 h 的子序列。最后在求解式(6)过程中采用EM优化算法,得到最大概率输出。

3.3 实验结果

在测试中采用了两类不同的人脸识别器:基于特征脸方法^[10]和基于线性鉴别分析(Linear Discriminant Analysis, LDA)方法^[21]。表3所示为加入后处理器与否情况下的实验结果。

表3 实验结果

是否使用 后处理	算法	正确的人脸类被列为第 i 候选的百分比/%		
		$i=1$	$i=2$	$i=3$
未使用	特征脸	85	73	45
	LDA	86	58	41
后处理 $K=5$	特征脸	92	86	47
	LDA	94	88	53
后处理 $K=10$	特征脸	93	87	56
	LDA	94	89	66

实验结果表明,后处理能够明显地提高整个系统的性能。更为明显的是能够提高正确的人脸出现在候选集第2和第3位的可能性。比如使用LDA算法时,正确人脸被列在第2位的几率仅仅为58%,而在采用后处理后,正确人脸出现几率提高到88%。因此本文提出的基于HMM后处理与基于多个

最相似人脸的识别器配合使用时效果将更加明显,这是未来需要进一步研究的内容。

4 结语

为了揭示现有典型识别算法的性能对姿态和光照改变的鲁棒性,本文通过建立光照反射模型和3D人脸模型分析了特征脸算法的识别性能随光照和姿态改变而变化的情况,并进一步提出了一种对人脸识别器进行基于HMM后处理的“数据源灾难”解决方案。与传统的基于图片的人脸识别技术相比,基于视频的人脸识别可以利用的信息更多,比如多幅同一个人的多幅人脸图像,视频中人脸在时间和空间上的连续性等,因此如何充分利用视频中人脸的时间和空间信息克服视频中人脸分辨率低,尺度变化范围大,光照、姿态变化比较剧烈以及时常发生遮挡等困难是未来研究的重点。

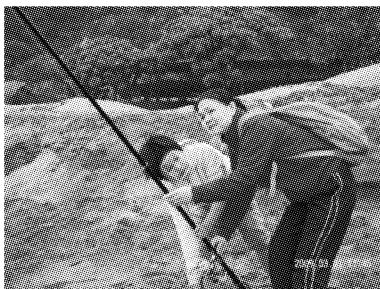
参考文献:

- [1] VASILESCU M, TERZOPOULOS D. Multilinear independent components analysis[C]// CVPR 2005: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2005, 1: 547-553.
- [2] LAM K-M, YAN HONG. An analytic-to-holistic approach for face recognition based on a single frontal view[J]// IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(7): 673-686.
- [3] BLANZ V, BOMDHANI S, VETTER T. Face identification across different poses and illuminations with a 3D morphable model[C]// Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. Washington, DC: IEEE Computer Society, 2002: 202.
- [4] 卿来云, 山世光, 陈熙霖, 等. 基于球面谐波基图像的任意光照下的人脸识别[J]. 计算机学报, 2006, 29(5): 760-768.
- [5] WANG HUIYUAN, LENG YAN, WANG ZENG FENG, et al. Application of image correction and bit-plane fusion in generalized PCA based face recognition[J]. Pattern Recognition Letters, 2007, 28(16): 2352-2358.
- [6] 杜成, 苏光大, 林行刚, 顾华; 改进的线段 Hausdorff 距离人脸识别方法[J]. 光子·激光, 2005, 16(1): 89-93.
- [7] CHELLAPPA R, ROY-CHOWDHURY A K, ZHOU S K. Recognition of humans and their activities using video[M]. [S. l.]: Morgan & Claypool, 2005.
- [8] ZHAO W. Robust Image Based 3D Face Recognition[D]. College Park, USA: University of Maryland, 1999.
- [9] HORN B K P, BROOKS M J. Shape from shading[M]. Cambridge: MIT Press, 1989.
- [10] TORRES L, VILÀ T. Automatic face recognition for video indexing applications[J]. Pattern Recognition, 2002, 35(3): 615-625.
- [11] ARTIKLAR M, HASSOUN M H, WATTA P. Application of a post-processing algorithm for improved human face recognition[C]// Proceedings of IEEE International Joint Conference on Neural Networks. [S. l.]: IEEE, 1999, 5: 3280-3283. (下转第970页)

数。图6是采用所提算法得到的部分实验结果。

表1 实验结果

检测算法	训练样本集	误检数	检测率/%	平均速度 ms/帧
Viola 和 Jones 算法 ^[7]	MIT + CMU	645	82.6	113
本文算法	MIT + CMU	186	86.9	38
本文算法	自建样本集	109	91.9	40



(a) 部分实验结果1



(b) 部分实验结果2

图6 部分实验结果

3 结语

在多姿态人脸检测方面,提出基于人脸特征和 AdaBoost 算法相结合的检测算法。与其他文献相比,算法经过肤色检测、眼睛和嘴巴的搜索,快速有效地缩小人脸检测的范围,提高了检测速度。通过眼睛、嘴巴的几何特征确定人脸方向,按人脸方向分割出大致正向的人脸候选区域,在不扩展 Harr-like 特征或增加姿态检测器等情况下,利用 AdaBoost 算法较好地解决了包括人脸旋转在内的多姿态人脸检测问题。同

时,按长宽比为 1.2 选取样本,提高了样本的有效性。实验表明,算法提高了检测率,具有较高的检测速度和较少的误检数,对光照、遮挡和表情具有较强的鲁棒性。

参考文献:

- [1] YANG M H, KRIEGMAN D J, AHUJA N. Detecting faces in images: A survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(1): 34–58.
- [2] 黄福珍, 苏剑波. 人脸检测[M]. 上海: 上海交通大学出版社, 2006.
- [3] YANG J, LU W, WAIBEL A. Skin-color modeling and adaptation [C]// *Third Asian Conference on Computer Vision*. London: Springer-Verlag, 1998: 687–694.
- [4] KAKUMANU P, MAKROGIANNIS S, BOURBAKIS N. A survey of skin-color modeling and detection methods [J]. *Pattern Recognition*, 2007, 40(3): 1106–1122.
- [5] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2001: 511–518.
- [6] LIENHART R, MAYDT J. An extended set of haar-like features for rapid object detection [C]// *Proceedings of IEEE International Conference on Image Processing*. Piscataway: IEEE Press, 2002: 900–903.
- [7] JONES M, VIOLA P. Fast multi-view face detection [EB/OL]. [2009-02-16]. <http://www.merl.com/reports/docs/TR2003-96.pdf>.
- [8] YANG J, WAIBEL A. Real-time face tracker [C]// *Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision*. Washington, DC: IEEE Computer Society, 1996: 142–147.
- [9] HSU R L, ABDEL-MOTTALEB M, JAIN A K. Face detection in color images [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 696–706.
- [10] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of online learning and an application to boosting [C]// *Proceedings of the Second European Conference on Computational Learning Theory*. London: Springer-Verlag, 1995: 23–37.

(上接第963页)

- [10] MOGHADDAM B, JEBARA T, PENTLAND A. Bayesian face recognition [J]. *Pattern Recognition*, 2000, 33(11): 1771–1782.
- [11] LI Y X, TAN C L, DING X. Contextual post-processing based on the confusion matrix in offline handwritten Chinese script recognition [J]. *Pattern Recognition*, 2004, 37(9): 1901–1912.
- [12] HO K T, HULL J J, SRIHARI S A. Decision combination in multiple classifier systems [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(1): 66–75.
- [13] ZHAO W, CHELLAPPA R, PHILLIPS P J, *et al.* Face recognition: A literature survey [J]. *ACM Computing Surveys*, 2003, 35(4): 399–458.
- [14] LEE K C, HO J, YANG M H, *et al.* Video-based face recognition using probabilistic appearance manifolds [C]// *CVPR'03: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2003, 1: 313–320.
- [15] O' TOOLE A J, HARMS J, SNOW S L, *et al.* A video database of moving faces and people [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(5): 812–816.
- [16] SIM T, BAKER S, BSAT M. The CMU pose, illumination, and expression database [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(12): 1615–1618.
- [17] VILAPLANA V, MARQUES F, SALEMBIER P, *et al.* Region-based segmentation and tracking of human faces [C]// *EUSIPCO-98: European Signal Processing Conference*. [S. l.]: IEEE, 1998: 311–315.
- [18] NISTER D, STEWENIUS H. Scalable recognition with a vocabulary tree [C]// *CVPR'06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2006: 2161–2168.
- [19] BELLHUMEUR P N, HESPANHA J P, KRIENGMAN D J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711–720.