

基于脸部和步态特征融合的身份识别

侯晓慧, 刘志镜

(西安电子科技大学 计算机学院, 西安 710071)

(xh.hou@hotmail.com)

摘 要:提出了一种将脸部和步态特征相结合,应用于智能监控系统进行远距离视频流中身份识别的新方法。该方法首先分别采用隐马尔可夫模型(HMM)和Fisherfaces方法进行步态和脸部的识别,之后将这两个分类器得到的结果进行匹配级的融合。对从不同方向采集的31个人的视频序列进行分析实验,结果表明将脸部和步态特征相结合进行身份识别具有很好的鲁棒性,其识别性能也优于只采用脸部或步态单一特征的识别方法。

关键词:步态识别;脸部识别;多生物特征的融合

中图分类号:TP391.41 **文献标志码:**A

Combination of face and gait for human recognition

HOU Xiao-hui, LIU Zhi-jing

(School of Computer Science and Technology, Xidian University, Xi'an Shaanxi 710071, China)

Abstract: Focusing on the application of intelligent surveillance, this paper proposed a new approach in which the combination of face and gait was used for human recognition at a distance in video sequences. Hidden Markov Model (HMM) and Fisher faces method were primarily applied for gait and face recognition, respectively. And then, the results obtained from the two classifiers were utilized and integrated at match score level. The system was tested on video sequences of 31 individuals collected from different directions. The results show that the combination of face and gait provides a more robust recognition strategy, and it has better recognition performance compared with face-only or gait-only method.

Key words: gait recognition, face recognition, multi-biometrics fusion

0 引言

近年来,智能监控在公众安全领域发挥着越来越重要的作用。如果一个智能监控系统能有效检测并发现可疑的人或行为,就可以及时发出报警信号,从而减少犯罪的发生。动态场景中的可视化智能监控作为计算机视觉领域的一个研究热点,可以实现对视频流中目标的监测、跟踪和识别,以及进一步地对目标的理解和描述。其中的身份识别技术更是以其广泛的研究前景和巨大的经济价值,吸引着世界上越来越多机构和研究人员的关注。

基于人脸图像的身份识别是一个相对成熟的技术,特别是在一些室内环境的应用中性能非常可靠。相比于指纹、虹膜这些人们比较熟悉的生物特征,脸部特征具有几个优势:无需直接接触,具有非入侵性,应用成本较低,且容易采集。然而,采用脸部特征进行身份识别也存在一定的弊端,它很容易受到年龄、表情、化妆等因素影响。另外,人脸是短距离的生物特征,只有当目标与摄像机距离足够近的情况下才能获得有效的脸部细节用于识别过程。而步态是中长距离生物特征,在距离很远的情况下或质量较低的图像序列中也可获得可靠的数据。而且在识别细小区别方面,步态比其他生物特征具有更好的鲁棒性和不易伪装性。当然,步态作为生物特征也存在缺点:由于是一种行为特征,步态会受到诸如衣物、年龄、伤病、环境及其他提取可视化线索时常遇问题的影响。

鉴于以上原因,将脸部特征和步态特征相结合,发挥各自优势进行身份识别可以使系统更具鲁棒性,获得更好的性能。

人脸和步态都是视觉线索,而且都可以从大多数的视频监控系统中提取出来。它们又是采用相互独立的人的特征:人脸识别系统采用相对细节的面部特征,而步态识别方法采用相对粗略的身体形状变化作为研究对象^[1]。因此,某些对人脸识别系统性能影响很大的因素,比如灯光和面部表情,对步态影响却很小或者根本没有影响。类似地,某些对步态识别准确率影响很大的因素,如衣服、鞋子,根本不会影响脸部识别的效果。因此有理由相信,将脸部和步态特征相结合用于身份识别将使系统的识别率有较大提高。将不同特征结合进行身份识别的研究已经获得了很好的结果^[2-4]。本文将通过实验对比单纯采用步态特征、脸部特征和将二者结合进行身份识别的系统性能。实验结果证明了将视频中采集到的面部和步态特征相结合进行身份识别是一个非常有效的方法。

1 多特征融合应用于身份识别

本文的研究重点是脸部特征和步态特征的融合对身份识别系统性能的影响,因此无需将过多精力放在选择最优的识别算法上,接下来将分别采用隐马尔可夫模型(Hidden Markov Model, HMM)和Fisherfaces方法进行基于步态和脸部特征的身份识别,之后再将两个分类器的结果数据在匹配级上进行融合。

1.1 步态识别算法

由于HMM具有离散统计特性,可以反映步态随时间的变化,因此被广泛应用于身份识别中。典型的HMM常表示为 $\lambda = (N, M, \pi, A, B)$ 。其中, N 代表模型中的状态数目, M

收稿日期:2009-02-19;修回日期:2009-04-07。 基金项目:广东省教育部省部产学研联合开发项目(5003-00703P07)。

作者简介:侯晓慧(1980-),女,内蒙古包头人,助教,硕士,主要研究方向:视觉计算、身份识别; 刘志镜(1957-),男,陕西西安人,教授,主要研究方向:数据挖掘、视觉计算。

代表每个状态下的观测值数目, π 为初始状态的概率分布, A 为转换概率矩阵, B 是观测值矩阵。

将人体行为分成一系列的图像序列,并用其中一段连续帧来表示一个周期的规则行为^[5]。图 1 是一个正在行走的人的步态周期。每个周期的特征向量进一步划分为同样大小的片段,利用每个片段取一个样本,定义如下:

$$e_n = \frac{1}{N_{n f_t \in C_n}} \sum f_t \quad (1)$$

其中 f_t 是第 t 帧的特征向量, C_n 代表第 n 个片段, N_n 是第 n 个片段中帧的数目。样本集表示为 $E = \{e_1, e_2, \dots, e_N\}$ 。

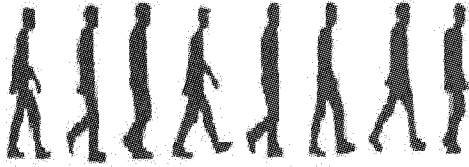


图 1 一个行走周期的人体轮廓

一个行走周期中的每个特征向量 f 和样本 e 之间的距离可以用式(2)进行定义:

$$D(f, e) = 1 - \frac{f^T e}{\sqrt{f^T f e^T e}} \quad (2)$$

用实验误差的方法对转换概率矩阵 A 进行初始化。初始观测值矩阵 $B = \{b_j(k)\}$ 的定义如下:

$$b_n(f_i) = 0.01 \delta_n e^{-\delta_n \times D(f_i, e_n)} \quad (3)$$

$$\delta_n = \frac{N_n}{\sum_{f_i \in C_n} D(f_i, e_n)} \quad (4)$$

HMM 经过反复地迭代训练,在一个周期上采用 Viterbi 算法获得最大可能路径 $Q = \{q_1^{(i)}, q_2^{(i)}, \dots, q_r^{(i)}\}$, 其中 $q_i^{(i)}$ 是第 i 次迭代后在时间 t 的状态。新样本 $E(i)$ 可以利用式(1)从最大可能路径中获得, $B(i)$ 可以利用式(2) ~ (4) 获得。下一个 $A(i)$ 和 $\pi(i)$ 可以利用 Baum-Welch 算法进行更新^[6]。训练的结果用样本和 HMM 参数表示,通常几次迭代后可以得到一个较好的估算值。

待测的步态图像序列采用和训练样本序列同样的方法进行预处理,利用训练序列的样本计算获得待测序列的观测值矩阵 B 。本文采用 Viterbi 算法计算待测序列和训练序列的概率值,并将这些值作为相似度。

1.2 脸部识别算法

目前有很多成熟的脸部识别算法,本文采用 Fisherfaces 方法。该方法利用 FLD (Fisher Linear Discriminant) 简化主成分分析 (Principal Component Analysis, PCA) 子空间,从而使分类具有较高的可靠性^[7]。Fisherfaces 方法简要描述如下:

令 C 和 N_1, N_2, \dots, N_c 分别表示样本类的数目和每个样本类中人脸图像的数量。令 $\overline{M}_1, \overline{M}_2, \dots, \overline{M}_c$ 和 \overline{M} 表示每类的平均值和简化的主成分分析 (Principal Component Analysis, PCA) 子空间中总的平均值。代表一个人脸图像的向量 \overline{X}_i^T 可以用映射矩阵 W_{PCA} 映射为向量 $\overline{Y}_i^T = W_{PCA}^T \overline{X}_i^T$, 由此可得 $\overline{M}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \overline{Y}_{ij}^T = W_{PCA}^T (\frac{1}{N_i} \sum_{j=1}^{N_i} \overline{X}_{ij}^T)$, 其中 \overline{X}_{ij}^T 代表第 i 个样本类中的第 j 个人脸图像。 \overline{Y}_{ij}^T 的类间离散度矩阵 S'_b 和类内离散度矩阵 S'_w 表示为:

$$S'_b = W_{PCA}^T S_b W_{PCA}, S'_w = W_{PCA}^T S_w W_{PCA} \quad (5)$$

其中 S_b 和 S_w 分别代表 \overline{X}_i^T 的类间离散度矩阵和类内离散度矩

阵^[8]。将令 $\frac{|W_{PCA}^T S'_b W_{PCA}|}{|W_{PCA}^T S'_w W_{PCA}|}$ 取得最大比率的映射矩阵 W_{fld} 作为最佳映射。 $(S'_w)^{-1} S'_b$ 的级数为 $(C-1)$, 将 W_{fld} 的列数计算做矩阵 $(S'_w)^{-1} S'_b$ 中第 $(C-1)$ 个特征向量。因此,给定输入人脸图像 \overline{X}_{ij}^T , 得到映射 $\overline{V}_{ij}^T = W_{fld}^T W_{PCA}^T \overline{X}_{ij}^T$, 并通过与从一系列训练图计算获得的向量 \overline{V}_{ij}^T 进行对比分类。



图 2 3 组规范化正脸图

1.3 步态和脸部特征的融合

通过前面的讨论,将两个分类器的结果进行融合将提高系统的识别效率和准确率。但这些结果的取值范围及类型并不相同,因此不可以直接结合。脸部识别算法估算的是人脸图的相似度,是一个概率,而步态识别算法测量待测样本和最佳匹配特征向量之间的距离。因此,这些结果必须转化到同一范围,使其具有可比性之后才可以结合应用。

选用基于 z 标准化的高斯模型进行规范化^[9]。对于某一给定的待测值 p , 可以计算出它与训练集中所有元素的相似度值 $S(p, g_i)$, 其中 $i = 1, 2, \dots, N$ 。计算得到相似度的平均值 μ_p 和标准差 σ_p 之后, p 和 g_i 之间的相似度便可以规范化为:

$$S'(p, g_i) = \frac{S(p, g_i) - \mu_p}{\sigma_p} \quad (6)$$

分别将步态分类器和脸部分类器的结果进行规范化,之后采用 SUM, PRODUCT, MAX 和 MIN 法则对两个结果值进行匹配级的融合^[3]。其中,在高噪声和高模糊性情况下,采用 SUM 法则可以获得较好的融合效果,该方法将步态和脸部分类器的结果进行简单地相加。PRODUCT 法则基于条件统计独立性的假设,相比 SUM 法则而言,对每个输出的结果比较敏感。

2 实验结果

该实验中,我们采用 3 部摄像机获取 31 个人的 186 个视频序列。较采用 1 部摄像机而言,设置 3 部摄像机从不同方向进行拍摄有助于我们更好地获取正脸图和侧面的步态序列。每个人在同一天不同时间从摄像机前通过两次,以 25 fps 的速度拍摄,图像序列的平均长度为 94 帧。

为了确保侧面轮廓的完整性,采用背景减除和时间差分法获得侧面轮廓后,采用形态学腐蚀和扩张的理论结合区域融合对结果进行处理^[10]。成功提取到人体的侧面轮廓后,选用上面 1/6 部分作为人脸区域。选取 2 个行走周期用作步态识别,其中一个作为训练集,另一个作为测试序列。相应地,得到每个人的两幅脸部图像,提取到的人脸图像如图 2 所示。步态序列图的尺寸标准化为 80×196 , 脸部图标准化为 48×72 。

基于步态特征、脸部特征以及二者融合的身份识别方法的识别率对比如表 1 所示。根据以上实验结果,采用 SUM, PRODUCT 和 MAX 法则均可以获得较采用单一生物特征进行身份识别更好的性能。其中 SUM 法则获得的识别率最高,为 96.5%, 接下来是 MAX 和 PRODUCT 法则。MIN 法则的识别率

(下转第 2091 页)

4 结语

本文提出的人体运动跟踪算法,主要做了以下改进:针对 Snake 模型一般需手动确定初始轮廓位置的难题,利用角点检测结果确定的人体外接矩形作为 Snake 模型的初始轮廓,既实现跟踪自动化,又将初始轮廓尽可能地靠近人体真实轮廓,减少迭代次数,并降低 Snake 收敛到局部极值的概率;针对 Snake 模型跟踪不够稳定、容易出现跟踪丢失问题,结合了 KLT 光流法,选取当前帧所得到的轮廓点中的强特征点,对这些特征点进行光流估计,将估计结果作为下一帧 Snake 模型的初始轮廓,有效地解决上述问题,同时很好地解决光流算法大计算量的问题,达到实时跟踪的目的。

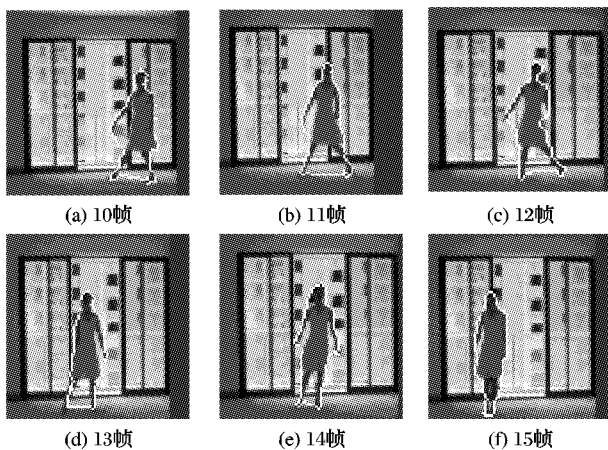


图 4 本文算法用于人体快速运动跟踪结果

(上接第 2088 页)

相对较低,为 70.4%。造成这一结果的原因可能是融合过程中某一个特征的不可靠降低了另一个较好的特征的性能。

表 1 基于步态、脸部和特征融合的身份识别结果对比

识别方法		识别率/%
步态识别		89.0
脸部识别		72.6
脸部和 步态的 融合	SUM	96.5
	PRODUCT	93.1
	MAX	91.0
	MIN	70.4

3 结语

本文提出了一种视频中基于多生物特征融合的身份识别方法。由于脸部特征是短距离生物特征,而步态是长距离生物特征。而且某些对其中一种生物特征造成明显影响的因素可能对另一种生物特征的影响很小。因此,采用将两种生物特征融合的方法进行身份识别相对于采用单一特征的识别方法而言,更具鲁棒性。实验结果也证明,将步态和脸部生物特征进行融合对于提高视频中的身份识别性能是一个很好的策略。下一步的研究将包括复杂场景下的基于多生物特征融合的身份识别及其在监控系统中的应用。

参考文献:

[1] JAFRI R, ARABNIA H. Fusion of face and gait for automatic human recognition[C]// Proceedings of the 5th International Conference on Information Technology. Washington, DC: IEEE Computer Society, 2008: 167 - 173.

参考文献:

[1] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564 - 577.

[2] ROSALES R, SELAROFF S. Improved tracking of multiple humans with trajectory prediction and occlusion modeling[EB/OL]. [2008 - 12 - 20]. <http://www.cs.bu.edu/techreports/pdf/1998-007-tracking-multiple-humans.pdf>.

[3] JU S, BLACK M, YACCOB Y. Cardboard people: A parameterized model of articulated image motion[EB/OL]. [2008 - 12 - 20]. http://reference.kfupm.edu.sa/content/c/a/cardboard_people__a_parameterized_model__44982.pdf.

[4] PARAGIOS N, DERICHE R. Geodesic active contours and level sets for the detection and tracking of moving objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(3): 266 - 280.

[5] KASS M, WITKIN M A, TERZOPOULOS D. Snakes: Active contour models[EB/OL]. [2008 - 12 - 20]. <http://www.cs.ualberta.ca/~jag/papersVis2/levsetReadGr/snakesTerzopoulos.pdf>.

[6] SMITH S M, BRADY J M. SUSAN: A new approach to low level image processing[EB/OL]. [2008 - 12 - 20]. <https://eprints.kfupm.edu.sa/67707/1/67707.pdf>.

[7] BIRCHES D. KLT: An implementation of the Kanade - Lucas - Thomas feature tracker[EB/OL]. [2008 - 12 - 20]. <http://vision.stanford.edu/birch/klt>.

[8] SHI J, TOMASI C. Good feature to track[EB/OL]. [2008 - 12 - 20]. <http://www.cs.ucf.edu/courses/cap6411/Lecture-14.pdf>.

[2] 王风华, 韩九强, 姚向华. 一种基于虹膜和人脸的多生物特征融合方法[J]. 西安交通大学学报, 2008, 42(2): 133.

[3] SHAKHNAROVICH G, LEE L, DARRELL T. Integrated face and gait recognition from multiple views[EB/OL]. [2008 - 12 - 15]. <http://people.csail.mit.edu/gregory/papers/cvpr2001.pdf>.

[4] ZHOU X, BHANU B. Feature fusion of side face and gait for video-based human identification[J]. Pattern Recognition, 2008, 41(3): 778 - 795.

[5] SUNDARESAN A, ROYCHOWDHURY A, CHELLPPA R. A hidden Markov model based framework for recognition of humans from gait sequences[EB/OL]. [2008 - 12 - 15]. <http://www.umi.acs.umd.edu/~rama/Conf.pdf-files/icip18.pdf>.

[6] CHEN C, LIANG J, ZHAO H. Gait recognition using hidden markov model[C]// The 2nd International Conference on Natural Computation. Berlin: Springer, 2006: 399 - 407.

[7] YI J, YANG H, KIM Y. Enhanced fisherfaces for robust face recognition[C]// Proceedings of the 1st IEEE International Workshop on Biologically Motivated Computer Vision. London: Springer-Verlag, 2000: 711 - 720.

[8] 高秀梅, 杨静宇, 杨健. 一种最优的核 Fisher 鉴别分析与人脸识别[J]. 系统仿真学报, 2004, 16(2): 2864.

[9] PHILLIPS P J, GROTH P, MICHEALS R J, et al. Face recognition vendor test 2002[EB/OL]. [2008 - 12 - 15]. <http://www.image2003.com/paper/down/19216829107200767813223824837.pdf>.

[10] Intel Corporation. Open source computer vision reference manual [EB/OL]. [2008 - 12 - 15]. <http://www.intel.com/research/ml/research/opencv/>, 2000.