

# 基于视频的人脸验证

庄 莉, 艾海舟, 徐光

(清华大学计算机科学与技术系, 智能技术与系统国家重点实验室, 北京 100084)

**摘要:** 本文提出了一种基于视频的人脸验证方法. 采用立体视觉方法初步将人脸区域与背景分割开, 再根据多关联模板匹配方法精确定位人脸. 对定位后的人脸区域抽取特征器官位置, 再依此裁剪出人脸样本. 从视频流中收集人脸样本, 训练支持向量机(SVM)作为验证器. 实验表明该方法在复杂的现场环境下是有效的、鲁棒的.

**关键词:** 人脸检测与跟踪; 人脸特征抽取; 人脸验证; 支持向量机(SVM)

**中图分类号:** TP391.4      **文献标识码:** A      **文章编号:** 0372-2112(2002)08-1222-04

## Video Based Face Verification

ZHUANG Li, AI Hai zhou, XU Guang-you

(Computer Science and Technology Department, Tsinghua University State Key Laboratory of Intelligent Technology and Systems, Beijing 100084, China)

**Abstract:** A video based face verification method is proposed. Stereo vision is used to roughly segment face area from its background, and then multiple related template matching method is used to locate the face area finely. The face organs are extracted from the located face area and accordingly the face sample is cut out. Face samples are collected from video sequence, and a support vector machine classifier is trained on this sample set against other face samples as his verifier. Experimental results demonstrate the feasibility and robustness of this algorithm in complex real environment.

**Key words:** face detection and tracking; facial feature extraction; face verification; support vector machines

### 1 引言

人脸识别是计算机视觉与模式识别领域关注的热点问题. 早期的研究多采用基于几何特征的方法<sup>[1]</sup>, 即抽取人脸上的特征点, 如瞳孔中心、眼角、嘴角、脸部轮廓等, 结合相关匹配方法进行识别. 近十年来, 基于模板的方法受到重视<sup>[2]</sup>, 如: 特征脸方法<sup>[3]</sup>、变形模板方法<sup>[4,5]</sup>. 绝大多数现有方法采用的训练集和测试集都是静止图象, 如: FERET<sup>[6]</sup>人脸库和 Yale 人脸库<sup>[7]</sup>. 早期的系统多采用的是每个人一张图作为训练, 后来的研究一般都采用同一个人多个姿态作为训练样本, 如: 9 个典型姿态, 希望借此较好地适应姿态变化. 但是, 由于训练样本仅包含少数几个典型姿态, 难以涵盖各种情况, 由此建立的识别算法其推广能力在本质上受到训练数据的限制, 因而目前的人脸识别算法仍然很难对各种姿态取得很好的适应性, 特别是在实际的现场环境中性能难以保证, 大大地限制了其推广应用的价值.

在人脸识别问题上, 一方面, 识别的效果很大程度上依赖于对人脸特征的描述, 而从样本中提取特征时却有可能丢失一些有利于区别不同样本的信息; 另一方面, 分类器的选取直接影响在给定特征描述下识别器的性能. 一般而言, 训练样本越充分, 训练出来的分类器就越能够取得好的适应性. 因此, 研究人脸识别问题, 一方面应该尽可能地获得更多的训练样

本, 另一方面需要选择具有良好学习能力的训练算法, 充分发挥训练样本的作用. 本论文深入研究了上述问题, 提出了基于视频的人脸验证方法, 一方面通过视频中的人脸检测与跟踪提供丰富的样本供训练之用, 直接使用人脸图象作为人脸特征的描述, 避免特征抽取时丢失信息; 另一方面采用基于统计学习理论的支持向量机(Support Vector Machines, 以下简称 SVM)训练分类器, 具有最好的推广能力. SVM 方法在模式识别领域倍受重视<sup>[8]</sup>, 已被用于人脸的跟踪与分析<sup>[9]</sup>, 在传统的人脸库上 SVM 方法取得了比其他方法优越的性能<sup>[10,11]</sup>. 我们认为视频输入可以提供更为丰富的训练样本, 充分发挥 SVM 统计学习能力, 有望进一步提高人脸验证系统的性能.

### 2 人脸验证系统的框架

人脸验证系统由双目立体摄像机作为视频输入, 从视频中获取各种姿态的人脸样本, 使用有两块 Matrox Meteor II 图象采集卡的 Pentium III 966MHz PC.

人脸验证算法的流程图如图 1 所示. 采用立体视觉分割出人脸所在区域, 再通过多模板匹配方法精确定位人脸区域, 抽取人脸特征点. 在此基础上, 剪裁出几何标准化的人脸图片并进行光照矫正, 得到规格化人脸样本. 将规格化人脸样本作为 SVM 分类器的输入. SVM 的训练使用 SMO(Sequential Minimal Optimization) 算法<sup>[12]</sup>.

### 3 人脸样本的采集

#### 3.1 基于立体视觉的人脸区域分割

假设待测试人站在摄像机前,他在视差图上具有最大的视差<sup>[13]</sup>.如图 2 所示,将视差最大的前景区域作为人脸的粗分割.在双目匹配过程中,对图象进行金字塔分解,采用由粗到精的相关匹配方法,快速而鲁棒地将人脸区域从复杂的背景中分割出来.

#### 3.2 人脸定位与跟踪

在立体视觉得到的人脸分割的基础上,采用基于多关联模板匹配的人脸检测算法<sup>[14]</sup>进一步精确确定人脸位置.

##### (1) 多关联模板

在 50 个手工标定眼睛和嘴角的标准证件照上,根据特征点裁剪出人脸样本并进行几何校正和灰度标准化,得到  $20 \times 20$  的平均脸模板和  $20 \times 8$  的平均眼模板.如图 3 所示.

##### (2) 人脸检测匹配算法

给定灰度平均值为  $\mu_T$ 、方差为  $\sigma_T^2$  的模板  $T[M][N]$ , 在一个灰度平均值  $\mu_R$ 、方差  $\sigma_R^2$  的图象窗口  $R[M][N]$  中,  $T$  和  $R$  的相关系数  $r(T, R)$  如下计算:

$$r(T, R) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (T[i][j] - \mu_T)(R[i][j] - \mu_R)}{M \cdot N \cdot \sigma_T \cdot \sigma_R}$$

在立体视觉匹配得到的分割中,首先用平均眼模板搜索候选人脸,再用人脸模板匹配进一步筛选,仅保留超过阈值的最大值的位置作为结果.根据可能的人脸大小范围,通过对图象反复作等比例压缩重采样并在压缩后的图象中进行搜索的方法实现多尺度人脸的定位<sup>[14]</sup>.如图 2 所示,外框是立体视觉匹配的结果,内框是多关联人脸模板匹配的结果.

#### 3.3 人脸特征检测

用水平 Sobel 算子对人脸分割图进行卷积,人脸器官的位置(眼睛、鼻子、嘴巴)是灰度变化明显的区域,如图 4(b).再进行直方图统计,根据人脸器官所占面积与人脸面积的大致比例,取累计直方图(图 4(c))中灰度值最大的 10% 的象素点为 1,其余象素点为 0,得到人脸器官的分割图(图 4(d)).

根据器官分割图,去掉所有几何位置、面积不符合要求的连通域,得到如图 5 左图的单眼候选区域.将几何位置和尺度上可能成为一对眼睛的单眼候选区域配对得到图 5 中图所示的双眼候选对.采用依据主分量分析的特征眼方法从双眼候选对中挑出真正的人眼<sup>[15]</sup>,得到图 5 右图所示人眼(瞳孔)位置.

从图 4(b) 根据眼睛的位置、尺度,框定鼻子和嘴大致位置,如图 6 左图虚线框所示.对该块区域沿着  $x$  方向计算灰度累积,第一个峰指示了鼻子的水平位置  $Y_{nose}$ .在  $[Y_{nose} - \sigma, Y_{nose} + \sigma]$  沿着  $y$  方向计算灰度累积,可以确定鼻子最宽部分左右边界.进而根据鼻子位置在图 4(d) 上检测左右嘴角位置.

在实验中,我们测试了 156 段视频序列,每段视频序列为 50 帧,其中所有特征点(眼、鼻、嘴巴中心)全部正确检测 7690 帧,错误检测 110 帧,正确检测率  $7690/7800 = 98.59\%$ .

#### 3.4 人脸样本的标准化

将双眼及嘴的中心共三个点对应到  $48 \times 48$  图片上 3 个标准分布的眼、嘴位置上,形成一个仿射变换.根据仿射变换将原图规范并裁剪为  $48 \times 48$  样本.对于裁剪后的样本,减掉一个根据图象灰度插值的灰度平面作光线矫正,再进行灰度分布归一化<sup>[14]</sup>.最后将人脸图象的 4 个角部分掩模掉,得到一个 2112 维的向量,如图 7 所示.

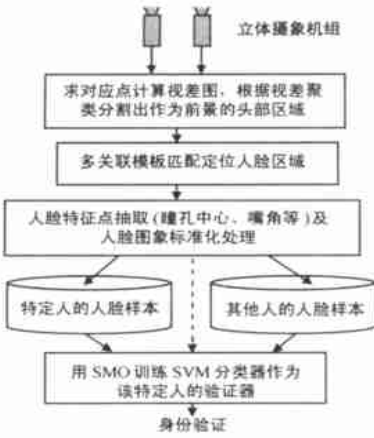


图 1 人脸验证算法流程



图 2 立体视觉及其模板匹配人脸分割图



图 3 多关联模板

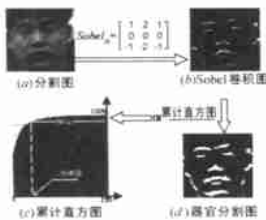


图 4 利用水平 Sobel 卷积和自动阈值选取得到人脸器官分割图



图 6 根据眼睛位置框定鼻子区域、双向灰度累积图

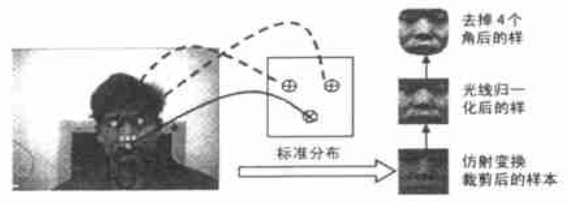


图 7 人脸样本采集和标准化过程

### 4 基于 SVM 的人脸验证器

#### 4.1 SVM 分类器基本原理

SVM 方法是基于结构风险最小化原理的统计学习方法<sup>[16]</sup>,它是一类基于最大间隔距离的分类器设计方法,对于

线性可分的问题,表达为求解最佳分割超平面的形式,通过使用 Lagrange 乘子法,优化过程转化为一个二次规划问题.对于线性不可分的问题,可以通过引入松弛变量和惩罚函数的方法推广最优超平面的概念.对于更为复杂的情况,需要引入非

线性映射将输入向量映射到高维特征空间,使原问题在高维空间中得以用线性方法解决,而根据积分算子空间的核函数理论,上述映射无须具体知道,只要通过选择核函数在原输入空间就可以求解 SVM 给出最佳分类超平面。

Platt 提出了 SMO 算法<sup>[12]</sup>,将 SVM 的二次规划问题分解为最小的求解两个 Lagrange 乘子的优化问题,通过解析的方法逐步求解.该方法使 SVM 的训练过程简便易行。

#### 4.2 用 SVM 分类器进行人脸验证

我们建立的立体视觉人脸库中包含 140 人的视频序列,每个人有 1 到 10 段视频序列不等,每段视频长度为 50 帧.这些视频序列在前后 5 个月的时间内采集完成.视频采集时人在摄像头前转动头部一周,以便采集到各个不同姿态.经过人脸检测定位、特征抽取、基于几何校验的样本挑选之后,每段人脸视频大约可以保留 30–50 帧图象.对每帧抽取的特征点位置进行微小扰动,进行几何、光照的标准化后,一帧图象得到 3 个剪裁样本,再左右镜像变换,得到 6 个标准化后的样本.因此,每段视频得到大约 180–300 个标准化后的样本。

实验中,任选 100 个人作为实验的考察对象.对每个考察对象,训练一个支持向量机(SVM)分类器.对于任一考察对象,将从本人的一段视频中得到的人脸样本集作为类内对象,将其他 99 人各一段视频中得到的所有人脸样本集作为类外对象,以此来训练 SVM 分类器.选择高斯径向基函数作为 SVM 的核函数,取界限参数  $C = 100$ ,训练一个考察对象的 SVM 分类器大约需要 30 分钟左右.表 1 给出了 5 个考察对象的 SVM 分类器训练情况。

表 1 5 个考察对象的训练情况

编 号	类 内 样本数	类 外 样本数	类内支持 向量数	类外支持 向量数	错分类 数 目
1	300	29187	26	84	0
2	300	29181	18	58	0
3	276	29211	34	78	0
4	300	29187	28	86	0
5	300	29187	32	75	0
平均	295.2	29190.6	27.6	76.2	0

由上表可以看出,支持向量的数目,特别是类外的支持向量的数目,比样本数据大大减少.图 8 中是部分类内支持向量(图 8(a))和类外支持向量(图 8(b)).

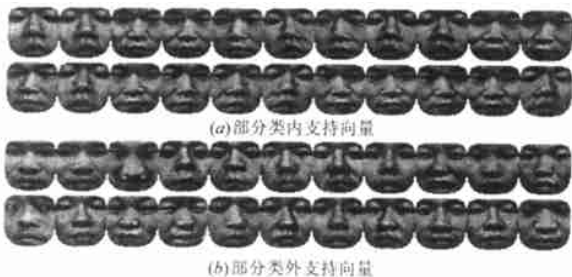


图 8 某特定人的 SVM 验证器的支持向量

## 5 实验结果

用错误接受率(FAR: False Acceptance Rate)和错误拒绝率

(FRR: False Rejection Rate)来评价算法的性能。

### 5.1 错误接受率(False Acceptance Rate)

错误接受就是错误地把冒充者(imposter)当作用户(client)接受了.用 50 段非训练视频段来测试错误接受率,实验结果见表 2。

表 2 错误接受率测试

样本编号	视频段数目	采样样本数目	错误接收数目	FAR(%)
1	50	14776	0	0.00
2	50	14776	18	0.12
3	50	14776	24	0.16
4	50	14776	30	0.20
5	50	14776	8	0.05
总数	250	73880	80	0.11

### 5.2 错误拒绝率(False Rejection Rate)

错误拒绝就是错误地把用户(client)当作冒充者(imposter)拒绝了.采集测试样本时,通过特征点之间的几何约束自动地抛弃那些过于侧面的样本,使测试集中在较正面的样本上.每段视频取得 170–300 个正面/准正面采集样本,对非训练集中的 27 个视频段进行测试,实验结果见表 3。

表 3 错误拒绝率结果

样本编号	视频段数目	采样样本数目	错误拒绝	FRR(%)
1	6	240	0	0.00
		300	0	0.00
		272	16	5.88
		300	0	0.00
		300	0	0.00
		176	12	6.82
2	3	276	0	0.00
		270	0	0.00
		276	0	0.00
3	10	300	18	6.00
		300	16	5.33
		300	0	0.00
		300	0	0.00
		300	0	0.00
		300	14	4.67
		300	0	0.00
		300	0	0.00
		300	0	0.00
		300	0	0.00
		300	12	4.00
		300	0	0.00
4	5	300	34	11.33
		300	0	0.00
		300	2	0.67
		300	10	3.33
		258	22	8.53
5	3	300	0	0.00
		294	0	0.00
		300	12	4.00
总数	27	7762	168	2.16

### 5.3 总的错误率和验证策略

根据表 2 和表 3,我们得到人脸验证实际正确率 97.73%

( $1-FAR-FRR$ )。该结果是对每一个采样样本计算得到的。考虑实际系统中的验证策略,对一段多帧的视频序列,可以通过投票的方式:当在这段视频序列中通过认证的样本数目超过一定比例的时候,就认为对这个人的认证成功。如果按这种验证策略,在测试集上我们的错误率则降为 0,达到完全正确的理想情况。由此,极大地提高了系统的验证性能。单纯的验证时间大约是 60ms/frame,接近实时要求。

## 6 总结与讨论

本文介绍了一个基于视频序列的人脸验证原型系统,采用自动的从视频流中收集训练样本的方式,训练 SVM 分类器作为人脸验证器,取得了十分有说服力的实验结果。提出了一整套从视频中采集人脸样本的方法,包括:立体视觉粗分割,多关联模板匹配精确定位,人脸器官特征点检测及以此为基本的人脸样本剪裁和标准化过程,实验表明了其有效性和鲁棒性。

我们认为现有的人脸识别方法在根本上受制于人脸建模所用到的样本集,通常这样的样本集对于每个特定人而言只包含数量很少的人脸样本,在此基础上研究出来的识别方法,尽管采用了十分复杂的方法,仍难以取得很好的推广能力。考虑到在实际的应用中,通常输入的是视频信号,可以提供非常丰富的样本集,在此基础上,采用基于数据的学习方法,比如 SVM 方法,在理论上可以保证获得更好的识别及验证性能。

### 参考文献:

- [ 1 ] Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: a survey [ J ]. Proc. Of The IEEE, 1995, 83( 5 ): 705- 740.
- [ 2 ] Brunelli R, Poggio T. Face recognition: feature versus templates [ J ]. IEEE Trans, 1993, PAMI- 15( 10 ): 1042- 1052.
- [ 3 ] Moghaddam B, Pentland A. Beyond Linear Eigenspaces: Bayesian Matching for Face Recognition [ M ]. In: Wechsler H et al, eds. Face Recognition from Theory to Applications. New York: Springer, 1998. 230- 243.
- [ 4 ] Ullman S, Basri R. Recognition by linear combinations of models [ J ]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1991, ( 13 ): 992- 1006.
- [ 5 ] C Kotropoulos, A Tefas, I Pitas. Frontal face authentication using variants of dynamic link matching based on mathematical morphology [ A ]. IEEE Conference on Computer Vision and Pattern Recognition [ C ]. Santa Barbara, California: 1998. 814- 819.
- [ 6 ] P J Phillips, M Hyeonjoon, et al. The FERET evaluation methodology for face recognition algorithms [ J ]. IEEE Trans, 2000, PAMI - 22 ( 10 ): 1090- 1104.

- [ 7 ] The Yale Database [ DB/ OL ]. <http://cvc.yale.edu/projects/yalefaces.html> 1999.
- [ 8 ] Lanitis A, Taylor C J, Cootes T F. Automatic face identification system using flexible appearance models [ J ]. Image and Vision Computing, 1995, 13( 5 ): 393- 401.
- [ 9 ] Kumar V P, Poggio T. Learning based approach to real time tracking and analysis of faces [ A ]. IEEE Conf. on Automatic Face and Gesture Recognition Proceedings [ C ]. Grenoble, Frans: 2000. 96- 101.
- [ 10 ] K Jonsson, J Matas, et al. Learning support vectors for face verification and recognition [ A ]. IEEE . Conf. on Automatic Face and Gesture Recognition Proceedings [ C ]. Grenoble, France: 2000. 208- 213.
- [ 11 ] G Guo, S Z Li, K Chan. Face recognition by support vector machines [ A ]. IEEE. Conf. on Automatic Face and Gesture Recognition Proceedings [ C ]. 2000. 196- 201.
- [ 12 ] J C Platt. Sequential minimal optimization: A fast algorithm for training support vector machines [ R ]. Technical Report MSR-TR 98 14, 1998.
- [ 13 ] T Darrell, G Gordon, et al. Integrated person tracking using stereo, color, and pattern detection [ A ]. IEEE Conf. on CVPR' 98 [ C ]. Santa Barbara: 1998. 601- 609.
- [ 14 ] 梁路宏, 艾海舟, 何克忠, 张钺. 基于多关联模板匹配的人脸检测 [ J ], 软件学报, 2001, 12( 1 ): 94- 102.
- [ 15 ] Zhenyuan Peng, Linmi Tao, Guangyou Xu, Hongjiang Zhang. Detecting facial features on images with multiple faces [ A ]. ICMF 2000 [ C ], Beijing: Springer Publish, 2000. 191- 198.
- [ 16 ] Gunn S R. Support Vector Machines for Classification and Regression [ R ]. Technical Report, Image Speech and Intelligent Systems Research group, University of Southampton, 1997.

### 作者简介:



庄莉女, 1978年1月生于四川成都, 2000年毕业于清华大学计算机系, 获学士学位, 同年攻读该系计算机应用专业硕士学位, 研究方向为人脸识别。



艾海舟男, 1964年1月生于江苏丹阳, 1985、1988、1991年分别获得清华大学学士、硕士、博士学位, 1994.9-1996.8在比利时布鲁塞尔自由大学作博士后研究, 现为清华大学计算机系副教授, 研究方向为计算机视觉、模式识别。

徐光男, 1940年2月生于上海, 清华大学计算机系教授、博士生导师, 研究方向为计算机视觉, 人机交互技术和多媒体技术。