

# 用于人脸对准的快速形状学习方法

柴秀娟<sup>1</sup>, 山世光<sup>2</sup>, 高文<sup>1,2</sup>

(<sup>1</sup> 哈尔滨工业大学 计算机学院 智能接口实验室, 哈尔滨, 150001)

(<sup>2</sup> 中国科学院计算所 ICT-ISVISION 面像识别联合实验室, 北京, 100080)

( {xjchai, sgshan, wgao}@jdl.ac.cn )

**摘要:** 人脸配准是人脸识别中的一个关键问题。本文提出了一种新的用于人脸特征对准的基于样本的形状学习策略(ESL)。该方法是基于下面结论提出的: 人脸图像差和形状差之间存在一种近似的线性关系——即相似的人脸图像蕴涵着相似的形状。因此, 给定一个人脸图像的学习集, 其中人脸图像的特征点都已手工标定。则任意新的人脸图像的形状可以如下抽取: 估计该人脸图像和训练集中训练图像的相似度, 并将同样的相似度用于该人脸图像形状的重建。实验表明, 本算法相对于其他传统的形状提取算法具有较高的精度和较快的速度。

**关键词:** 人脸识别、人脸配准、基于样本的形状学习

## 1. 概述

人脸识别在多模式人机接口、商业监控、法律实施等诸多方面有很广泛的应用前景。阻碍人脸识别系统真正实用化有很多瓶颈问题, 包括: 姿态问题, 表情问题, 光照问题等[1,2]等, 这些问题都要以解决精确的人脸配准为基础。

人脸配准在通常意义下可被划分为三个层次: 仿射变换; 稀疏的特征对应以及基于像素的密集的特征对应。

目前的很多识别系统都是基于仿射配准这一层次的对应。但是, 这样粗糙的特征配准对于复杂的人脸分析任务来说过于简单, 因此, 稀疏的特征对应目前在人脸研究领域里吸引了广泛的注意。ASM(Active Shape Models)和 AAM(Active Appearance Models)[3,4]是目前解决稀疏特征配准问题的主流方法。Beymer 等人[5,7]提出的基于光流的向量化技术, 可计算不同人脸图像之间的密集的对。但是由于受到光流计算精度的限制, 因此在实际应用中密集的对是很难精确计算得到的。

本文研究了人脸图像和人脸形状的关系。通过实验发现, 在图像差和形状差之间存在着近似的线性关系, 即相似的人脸图像蕴涵着相似的形状。由此结论, 我们提出了基于样本的形状学习(ESL)策略来解决人脸配准问题。对于任意一幅新的人脸图像的形状都可以如下抽取得到: 估计新图像和训练集中图像的相似度, 并将同样的相似度用于图像的重建, 图像的重构系数即作为形状的重构系数。

本文的组织结构如下: 第二部分, 在人脸图像和人脸形状之间存在线性关系的前提下, 对基于样本的形状学习算法进行详细的描述; 形状提取的实验结果及算法评测在第三部分给出; 最后, 得出本文的结论。

## 2. 基于样本的人脸形状学习算法

本方法中, 人脸形状由预先定义的  $n$  个特征点的  $x, y$  坐标连接而成的向量表示。即:

$S = (x_1, y_1, \dots, x_n, y_n)^T$ , 人脸灰度纹理也同样表达为向量形式, 本系统中,  $n=103$ 。

为找到人脸图像和形状之间的关系，我们设计了如下实验：手工标注训练图像形状特征点。得到  $m$  幅人脸图像  $I_1, I_2, \dots, I_m$  以及对应的形状  $S_1, S_2, \dots, S_m$ 。所有训练图像通过仿射变换归一到同尺度下。然后随机选择一幅图像为参考图像，用向量间的欧氏距离计算训练集中每个图像和参考图像之间的形状差和灰度差，得到图像差  $\Delta I_j$  和形状差  $\Delta S_j$  的集合。

将这些数据分布可视化后，可以发现，对于固定的参考图像，图像差和形状差之间存在着一种近似的线性关系，而对于不同的参考图像来说，这种线性关系的斜率近似统一。

从此实验中可得到如下结论：在图像差和形状差之间存在着近似的线性关系，相似的人脸图像蕴涵着相似的形状。基于这个结论，我们可以如下抽取人脸形状：给定一个新的人脸图像，首先通过最小化重建图像和原图像的残差来得到样本人脸图像的最佳线性组合系数，然后将同样的线性系数应用于对应的训练图像的样本形状来重构新人脸图像的形状。

为使重构人脸图像的线性系数更合理，本方法中所有的人脸图像都被通过固定双眼的位置来进行标准化[9,10]。人脸图像亮度值被标准化到 0 均值，1 方差，以减小光照的影响。

我们采用和线性对象类[6,8]相似的思想来描述图像和形状，即：
$$I = \sum_{j=1}^N \omega_j I_j, \quad S = \sum_{j=1}^N \omega_j S_j,$$

对一个给定的人脸图像，一旦我们得到了用于图像重构的线性组合系数，我们就可以得到他的人脸形状。因此，关键问题就是计算能够最优的线性重构输入图像的恰当的系数。其形式化描述如下：

对于给定的新图像，在尺度和灰度级上对它进行归一化，得到包含  $m$  个像素的图像，如下表示：

$I = (b_1, b_2, \dots, b_m)^T$ 。训练集中每一个训练图像表示为： $I_j = (a_{j1}, a_{j2}, \dots, a_{jm})^T$ 。因此，我们需要

$$\text{求解方程组: } \begin{cases} a_{11}\omega_1 + a_{21}\omega_2 + \dots + a_{N1}\omega_N = b_1 \\ a_{12}\omega_1 + a_{22}\omega_2 + \dots + a_{N2}\omega_N = b_2 \\ \dots \\ a_{1m}\omega_1 + a_{2m}\omega_2 + \dots + a_{Nm}\omega_N = b_m \end{cases} \quad (1)$$

通常， $m \gg n$ 。采用最小二乘算法， $A$  表示训练图像矩阵， $x = (\omega_1, \omega_2, \dots, \omega_N)^T$  是系数向量，那

么  $Y = \sum_{i=1}^N \omega_i I_i = Ax$  就是给定输入图像  $I$  的线性近似，称之为重建图像，且  $Y$  和  $I$  之间有一误差项：

$$E = \|I - Y\|^2 = \|I - Ax\|^2 = \left\| I - \sum_{i=1}^N \omega_i I_i \right\|^2。因此，解(1)就相当于解下面这个最小化问题： $x^* = \min_x E$ 。$$

假设由向量  $I_1, I_2, \dots, I_N$  张成的空间是  $L$ ，那么解这个最小化问题就相当于在  $L$  中找到一个向量  $Y$ ，使之最接近于输入的向量  $I$ 。假设  $Y = Ax$  就是这个向量。那么  $C = I - Y = I - Ax$  一定与  $L$  空间正交，也就是满足： $(C, I_1) = (C, I_2) = \dots = (C, I_N) = 0$ 。于是，可得到： $I_1' C = 0, I_2' C = 0, \dots, I_N' C = 0$ ，

重写成矩阵形式如下： $A'(I - Ax) = 0$ 。这个问题可以如下求解： $x = A^\perp I$ 。

其中,  $A^\perp$  是  $A$  的伪逆转置矩阵,  $A^\perp = (A'A)^{-1}A'$ 。得到了系数向量  $x$ , 输入图像  $I$  的形状向量  $S$  就可以由下式直接计算得到:  $S = (S_1, S_2, \dots, S_N)x$ 。最后, 通过将标准化后的形状特征点的坐标转换到原始的图像坐标系, 就可以得到原始输入图像的形状。

### 3. 实验结果

实验是在包含 300 幅 240x320 大小的准正面人脸图像库上进行的, 其中大部分的人脸图像具有中立表情, 所有测试图像都手工标定好特征点。为准确评价我们的算法, 使用 leave-one-out 策略来分开训练图像和测试图像。因此, 任一随机选择的人脸图像的形状向量就利用剩余的 299 幅图像做训练集, 使用我们的形状学习算法来计算得到。最后, 可用我们的算法得到所有 300 幅图像的形状向量, 然后计算由本方法抽取的形状特征点结果和真正的形状向量之间的欧氏距离。得到的平均误差为 1.95 个像素, 证明了此基于样本的形状提取算法的有效性。

表 1. ESL 和 ASM 方法的性能比较

方法	时间 (ms/frame)	平均误差(像素)	误差方差
平均形状	/	2.48	3.45
ASM*	70	2.14	2.23
ESL	10	1.95	1.85

\*注: 该 ASM 是我们的实现结果, 性能可能与原作者的实现有一定差距。

由于伪逆矩阵可被脱机计算, 我们的方法只需要很少的时间, 尤其是与诸如 ASM, AAM 这种需迭代计算的方法相比。表 1 给出了本方法与 ASM 方法的性能比较: ASM 方法也是按照固定两眼位置而进行初始化。可以看到, 本文提出的形状提取方法比 ASM 方法速度快很多, 而具有与 ASM 方法相似的精度。因此, 我们认为基于样本学习的形状提取算法是一种有效的人脸形状提取方法。

图 1 给出了本方法的一些实验结果, 显示了其良好的性能。而且, ESL 方法对于不同的种族, 不同的光照条件, 甚至是有遮挡的情况也具有很大的可扩展的潜力。

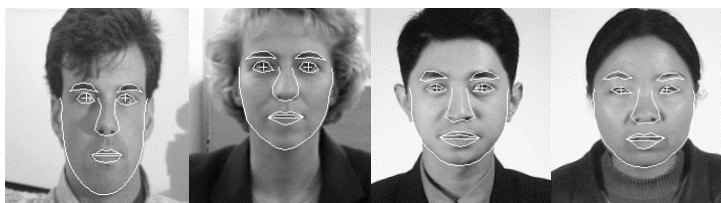


图 1. ESL 方法的一些结果

### 4. 结论

本文里, 我们讨论了 2D 人脸形状配准的问题。提出了一种基于样本的人脸形状学习策略来直接从新的人脸图像中抽取形状信息。基于相似的人脸图像蕴涵相似的形状这一实验结论, 认为人脸形状可以通过训练形状的线性组合映射计算得到, 其线性系数即为对应的重建新图像的训练图像的线性组合系数。实验证明了这个基于学习的形状提取方法具有很好的性能。

虽然我们仅仅研究了简单的最小二乘方法，但是这种学习策略已经表明了其作为解决特征定位，配准问题的方法的有效性。在今后的工作中，我们将研究更为复杂的学习策略，以期会有更好的性能。

## 致谢

该论文的研制工作得到了国家 863 计划项目（资助合同号 2001AA114190，2002AA118010）以及多媒体与智能软件技术北京市重点实验室开放课题（KP0706200377）的资助，同时还得到了银晨智能识别科技有限公司的大力资助。作者对这些资助一并表示诚挚的谢意。

## 参考文献

- [1] R.Chellappa, C.L.Wilson ect.1995.5 “Human and Machine Recognition of Faces: A survey”, Proc. of the IEEE, 83(5), pp705-740
- [2] P.J.Phillips, H.Moon, etc. 2000, “The FERET Evaluation Methodology for Face-Recognition Algorithms,” IEEE TPAMI, Vol.22, No.10, pp1090-1104
- [3] T.F.Cootes, C.J.Taylor,D.Cooper, and J.Graham. 1995, “Active Shape Models--Their Training and Application,” Computer vision and image understanding, 61(1): pp38-59
- [4] T.F.Cootes, G.J.Edwards and C.J.Taylor, 1998, “Active Appearance Models,” Proc. European Conf. Computer Vision, vol. 2, pp. 484-498
- [5] D.Beymer and T.Poggio, 1995, “Face Recognition from One Example View,” Proc. Int'l Conf. Computer Vision, pp. 500-507
- [6] D.Beymer, T.Poggio, 1995, “Image Representations for Visual Learning, Science,” Vol.272, pp1905-1909
- [7] D.Beymer, 1995.9, “Vectorizing Face Images by Interleaving Shape and Texture Computation,” A.I.Memo No. 1537
- [8] T.Vetter, T.Poggio, 1997, “Linear Object Classes and Image Synthesis from a Single Example Image,” IEEE Trans. On PAMI, Vol.19, No. 7, pp733-742
- [9] J.Miao, B.C.Yin, K.Q.Wang, et al, 1999, “A Hierarchical Multiscale and Multiangle System for Human Face Detection in a Complex Background using Gravity-Center Template,” Pattern Recognition, 32(7)
- [10] B.Cao, S.Shan, W.Gao, 2002, “Localizing the Iris Center by Region Growing Search,” Proceeding of the ICME2002. Vol. 2, pp129-132

# Novel Example-Based Shape Learning For Fast Face Alignment

Xiujuan Chai<sup>1</sup>, Shiguang Shan<sup>2</sup>, Wen Gao<sup>1,2</sup>

(<sup>1</sup>Vilab, Computer College, Harbin Institute of Technology, Harbin, China, 150001)

(<sup>2</sup>ICT-YCNC FRJDL, Institute of Computing Technology, CAS, Beijing, China, 100080)

({xjchai, sgshan, wgao}@jdl.ac.cn)

**Abstract:** In this paper, a novel Example-based Shape Learning (ESL) strategy has been proposed for facial feature alignment. The method is motivated by an intuitive and experimental observation that there exists an approximate linearity relationship between the image difference and the shape difference, that is, similar face images imply similar face shapes. Therefore, given a learning set of face images with their corresponding face landmarks labeled, the shape of any novel face image can be learned by estimating its similarities to the training images in the learning set and applying these similarities to the shape reconstruction of the novel face image. Our experiments have convincingly shown the effectiveness and efficiency of the proposed approach in both speed and accuracy performance compared with other methods.

**Key words:** Face recognition, Face alignment, Example-based shape learning(ESL)