

基于加权主动形状建模的人脸特征定位方法

冀贞海^{①②} 孙宁^{①②} 邹采荣^② 赵力^②

^①(东南大学学习科学研究中心 南京 210096)

^②(东南大学无线电工程系 南京 210096)

摘要: 主动形状建模是面部特征定位和人脸识别等模式识别领域中常用的一种方法。然而,由于受到初始情况、光照等诸多因素的影响,主动形状建模经常会陷入最优化过程中的局部最小问题,从而导致其性能下降。该文在传统主动形状模型基础上,提出了一种加权主动形状建模的方法,该方法可以有效地解决上述局部最小问题,并且更好地捕捉局部点的特征信息,从而更精确地进行面部特征定位。最后通过实验验证了上面的结论。

关键词: 人脸识别; 面部特征定位; 主动形状模型; 局部形状建模

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2007)04-0800-04

Face Alignment Based on Weighted Active Shape Models

Ji Zhen-hai^{①②} Sun Ning^{①②} Zou Cai-rong^② Zhao Li^②

^①(Research Center of Learning Science, Southeast University, Nanjing 210096, China)

^②(Department of Radio Engineer, Southeast University, Nanjing 210096, China)

Abstract: Active Shape Models (ASM) is one of powerful tools for face alignment and face recognition. However, the performance of ASM is often influenced by some factors such as the initial location, illumination and so on, which will frequently lead to the local minima in optimization. This paper proposes a weighted Active Shape Models, in which the more robust local appearance model is constructed with the local information of each landmark fully. Through the improved method the local minima problem can be solved efficiently, extracted the detailed local information of face feature points and localized the points accurately. Experiments verify the above conclusions.

Key words: Face recognition; Face alignment; Active Shape Models(ASM); Local appearance model

1 引言

人脸识别和面部表情识别是计算机视觉和模式识别领域中两个热点课题,如何准确定位和提取面部关键特征往往是其中首先需要解决的问题之一。Beinglass和Wolfson借助于子形状的每个清晰点作为参照点^[1],运用广义Hough变换,提出了一种定位特征点的方案;Kass等人则基于能量最小化准则原理,引入了主动轮廓方法^[2];Laurenz等人则借鉴了Gabor小波的理论^[3],构建了弹性图匹配的方法来进行特征定位和提取;Nastar和Ayache采用了有限元的方法来搜索类似形状的变化^[4]。

当前,最受关注的是Cootes等人提出的主动形状模型(ASM)方法^[5,6],该核心算法包括两个子模型,全局形状模型和局部纹理模型。然而,它的性能经常会受到诸如初始位置、光照和面部表情等因素的影响,上述因素会导致ASM陷入最优化过程中的局部最小问题。为此,文献[7]提出了一种新的形状评价方法,并用其解决局部最小问题;文献[8]则借助于

Gabor小波的理论来解决相应的问题。上面的两种方法立足于特征提取和形状评价的角度,本文则从ASM方法原始的思想出发,充分挖掘了标定点附近的局部纹理信息,将其扩展为更为鲁棒的模型。被扩展的局部纹理模型包括3个子模型:第1个是原来的局部纹理模型,另外两个则采样每个标定点两边附近的纹理信息,经过归一化等步骤构建而成。综合3个子模型,并对其进行相应的加权,可以更好地捕获特征点附近的纹理信息,从而为更精确地定位面部关键特征点提供更为可靠的依据。本文的剩余部分是这么安排的:第2节简要描述了传统的ASM方法,接下来则详细阐述了加权ASM方法,第4节则从3个方面对改进的方法和传统的方法进行了比较,最后是结束语。

2 主动形状模型(ASM)方法

传统的ASM方法是以点分布模型为理论基础,通过将描述面部形状的关键点的坐标作为训练样本向量,对其进行统计分析,进而获得全局形状模型和局部纹理模型来实现的。图1给出了一幅由58个关键点标定的完整人脸图像。假设给定训练样本集合为 $M = \{(I_j, s_j), s_j = \{x_1^j, y_1^j, x_2^j, y_2^j, \dots, x_N^j, y_N^j\}, j = 1, 2, \dots, K\}$,其中K为训练样本数目,N

2005-12-26收到,2006-06-26改回

国家自然科学基金项目(60503023)和教育部博士点基金(20050286001)资助项目



图 1 手工标定58个关键点的人脸图像

为预先定义的关键特征点的数目。 M 中的每个形状向量 s_j 是由训练图像 I_j 上预先定义并手工标定的 N 个关键特征点的横纵坐标串接而成。由于每幅图像的标定点的坐标取自于各自的坐标框架，如果将它们不加以对齐地描绘出来，是一幅意义不大的散图，如图 2 左图所示。此时，不同训练图像对应的标定点的坐标没有可比性。运用广义对齐算法将它们统一到同一坐标框架之下，重新描绘训练集中标定点的分布图，如图 2 右图所示，从图上可以看出，对齐后的图像能明显地体现出脸部形状的主轮廓，不同部位的点只在相应主轮廓点的附近“簇拥”分布。对齐后的形状向量记为 $\tilde{S}_j = \{\tilde{x}_1^j, \tilde{y}_1^j, \tilde{x}_2^j, \tilde{y}_2^j, \dots, \tilde{x}_N^j, \tilde{y}_N^j\}, j = 1, 2, \dots, K$ 。接着对其进行主分量分析便可获得 ASM 的全局形状模型，具体过程如下：

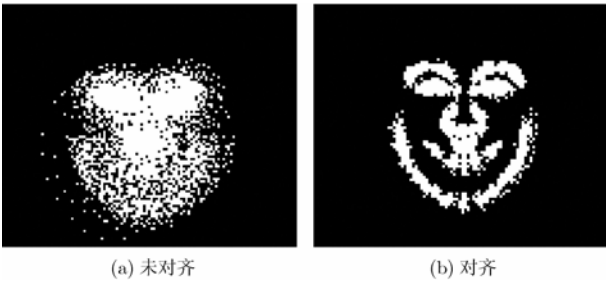


图 2 未对齐和对齐后标定点分布

(1) 计算对齐后形状向量的平均形状：

$$\bar{s} = \frac{1}{K} \sum_{j=1}^K \tilde{S}_j \quad (1)$$

(2) 计算对齐后形状向量的协方差矩阵：

$$\mathbf{V} = \frac{1}{K-1} \sum_{j=1}^K (\tilde{S}_j - \bar{s})(\tilde{S}_j - \bar{s})^T \quad (2)$$

该矩阵的维数为 $2N \times 2N$ ，其中上标 T 代表向量转置；

(3) 对协方差矩阵 \mathbf{V} 进行特征分解，其特征向量 \mathbf{p}_m 满足下式：

$$\mathbf{V}\mathbf{p}_m = \lambda_m \mathbf{p}_m \quad (3)$$

其中 λ_m 表示 \mathbf{V} 的第 m 个特征值， $\lambda_m \geq \lambda_{m+1}$ ，并且 $\lambda_m \neq 0$ ， $\mathbf{p}_m^T \mathbf{p}_m = 1, m = 1, 2, \dots, D$ ；

(4) 取前 t 个特征值，满足 $\sum_{m=1}^t \lambda_m / \sum_{m=1}^D \lambda_m \geq \theta$ ， θ 为一设定阈值，表征能量截断的比例，一般为 95%~98%，同时构造

矩阵 $\mathbf{P} = [\mathbf{p}_1 \mathbf{p}_2 \dots \mathbf{p}_t]$ ，最终获得全局形状模型为：

$$\mathbf{s} \approx \bar{\mathbf{s}} + \mathbf{P}\mathbf{b} \quad (4)$$

其中 $\bar{\mathbf{s}}$ 表示平均形状， \mathbf{b} 为主分量参数， \mathbf{P} 为主成分特征

向量构成的变换矩阵。图 3 显示了当分别变动前两个最大主分量参数，从 0 变动到各自 ± 3 倍的方差时所引起的面部形状变化。

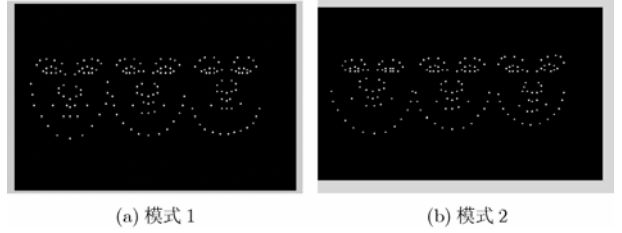


图 3 前两个最大主分量引起的脸形变化
(从左至右：-3 倍方差;平均形状;+3 倍方差)

ASM 方法中另一个重要模型是局部纹理模型，它与每一个标定点对应，通过平均纹理和协方差矩阵两个参数来具体刻画该点的灰度分布情况。该模型的构建过程大致如下：

(1) 假设给定标定点 p ，在其对应的法线方向上，以 p_{start} 为起点， p_{end} 为终点，取长度为 n_p 个像素的局部纹理向量 \mathbf{g}_{pj} ，并对其进行差分处理，该向量中的元素可以表示为：

$$g_{pj} = I_j(y_{p(k+1)}) - I_j(y_{pk}), k = 1, 2, \dots, n_p \quad (5)$$

其中 y_{pk} 表示沿着标定点 p 所在法线方向上的第 k 个点， $I_j(y_{pk})$ 表示该点在相应的图像 I_j 中的纹理灰度， y_{pk} 计算公式为：

$$y_{pk} = p_{start} + \frac{k-1}{n_p-1}(p_{end} - p_{start}) \quad (6)$$

(2) 为了降低光照的影响，对上面得到的向量进行归一化处理

$$\mathbf{g}_{pj}^* = \frac{\mathbf{g}_{pj}}{\sum_{k=1}^{n_p} |g_{pj,k}|} \quad (7)$$

(3) 类似全局模型构建中的过程(1)和(2)进行计算，最终获得标定点 p 对应的平均纹理 $\bar{\mathbf{l}}_p$ 和协方差矩阵 Σ_p ，此即为该点的局部纹理模型。

运用该模型在未知人脸图像中搜索某个特征点 p 的最佳候选点 q 时，用式(8)计算相应的马氏距离即可^[9]。

$$d(\mathbf{l}_q) = (\mathbf{l}_q - \bar{\mathbf{l}}_p)^T \Sigma_p^{-1} (\mathbf{l}_q - \bar{\mathbf{l}}_p) \quad (8)$$

其中 \mathbf{l}_q 是由未知图像 q 点附近采样得到的归一化纹理向量，上标-1表示求逆运算， $d(\mathbf{l}_q)$ 最小值对应的点 q 就是 p 的最佳候选点。

3 加权主动形状模型

3.1 原始局部纹理模型的扩展

尽管 ASM 有较好的性能，然而有时它的性能会受到诸多因素的影响，如初始化状态、光照和面部表情等。上述因素会经常导致 ASM 陷入最优化过程中的局部最小问题。本文在传统的 ASM 方法基础上，将其中的局部纹理模型进行

了推广。根据标定点附近的局部纹理信息,另外构建了两个相关的子模型,三者结合起来共同描述局部纹理的特性,从而在搜索未知目标人脸图像时能够更精确地定位面部关键特征点。

第1个子模型是原始的局部纹理模型,它的详细构建过程已经在第2节进行了阐述,为了与后面两个子模型名称一致,将其重新命名为中间局部纹理模型。剩余的两个子模型分别被称作内部局部纹理模型和外部局部纹理模型,两者的构建过程与中间局部纹理模型类似,差别仅在于采样点的位置选择不同。内部纹理模型主要在面部图像的“内部”进行采样,外部纹理模型主要在面部图像的“外部”进行采样,而中间纹理模型主要在面部轮廓上进行采样。图4简要给出3个子模型的采样位置以及区域,其中假设 p_0 是面部轮廓上的点,相对于点 p_0 而言, p_1 被称作是面部图像“内部”的点,而 p_2 则被称作是面部图像“外部”的点。 λ_1 和 λ_2 分别表示点 p_0 到点 p_1 和 p_2 之间的距离。分别以 p_0 , p_1 , p_2 为中心,以一定像素长度为半径,在相应的法线方向上采样,可得3个相应的向量,采用相同的方法对训练集中每幅图像的标定点做类似处理之后,最终可以获得与某给定点 p 相对应的3个平均向量和3个协方差矩阵,分别记作 $\bar{l}_p^m, \bar{l}_p^i, \bar{l}_p^e$, $\Sigma_p^m, \Sigma_p^i, \Sigma_p^e$,其中 \bar{l}_p^m 和 Σ_p^m 表示点 p 的中间局部纹理模型; \bar{l}_p^i 和 Σ_p^i 表示点 p 的内部局部纹理模型; \bar{l}_p^e 和 Σ_p^e 表示点 p 的外部局部纹理模型。至此,已经将原始的局部纹理模型推广为一个由3个子模型组成的更鲁棒的模型。

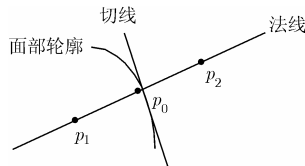


图4 3个子模型的采样位置示意图

经过对局部纹理模型扩充之后,与式(8)对应的的准则函数可以推广为下面更一般的形式:

$$q_{\text{opt}} = \arg \min_q [\alpha (\mathbf{l}_q^i - \bar{\mathbf{l}}_p^i)^\top \Sigma_p^{i-1} (\mathbf{l}_q^i - \bar{\mathbf{l}}_p^i) + \beta (\mathbf{l}_q^m - \bar{\mathbf{l}}_p^m)^\top \Sigma_p^{m-1} (\mathbf{l}_q^m - \bar{\mathbf{l}}_p^m) + \gamma (\mathbf{l}_q^e - \bar{\mathbf{l}}_p^e)^\top \Sigma_p^{e-1} (\mathbf{l}_q^e - \bar{\mathbf{l}}_p^e)] \quad (9)$$

其中 $\mathbf{l}_q^i, \mathbf{l}_q^m, \mathbf{l}_q^e$ 分别表示在 q 点附近采样得到的内部局部纹理向量,中间局部纹理向量以及外部局部纹理向量, α, β, γ 是相应的加权参数,满足条件 $\alpha + \beta + \gamma = 1$ 和 $\alpha, \beta, \gamma \geq 0$ 即可。

对比式(8)和式(9)我们可以得到下面两条重要的结论:

(1)式(8)是式(9)的一个特例,只需满足下面特定的条件即可。

$$\left. \begin{aligned} \alpha &= \gamma = 0 \\ \beta &= 1 \end{aligned} \right\} \quad (10)$$

(2)在定位面部关键特征点时,加权的ASM方法涵盖该点更多的信息量,因此,使用公式(9)作为评价准则则会获得更优的候选点。

在上述构建推广的局部纹理模型过程中,涉及到5个参数 $\lambda_1, \lambda_2, \alpha, \beta, \gamma$ 的选择问题。在本文中,相对于内部和外部局部纹理模型而言,由于中间局部纹理模型着重于在面部轮廓上采样,应该扮演着更为重要的角色,而内部和外部局部纹理模型地位几乎均等,因此, α, β, γ 3个值分别设定为0.25, 0.5, 0.25。 λ_1 和 λ_2 表征了点 p_0 到点 p_1 和 p_2 之间的距离,如果它们的值太小,则搜索的范围不能包含更多点局部信息量,不能充分发挥扩充ASM方法的作用;反之,如果所取值过大,则点 p_0 和点 p_1 及 p_2 偏离过远,由此违背了扩充ASM方法的初衷,鉴于上述的分析,我们采取了折中的办法,本文中, λ_1 和 λ_2 的值均选取为4个单位像素长度。

3.2 加权ASM算法

基于上述加权ASM方法,我们可以通过迭代的方法来实现未知人脸图像的关键特征点的搜索,每步的迭代过程如下:

(1)假设当前的全局面部形状为 s_{t-1} ;

(2)用式(9)搜索每一个标定点对应的最佳候选点,由此获得图像框架下的新的面部形状,记为 s_t' ;

(3)借助于相似变换,将新的全局面部形状从图像框架投影到坐标框架下,获得坐标框架下对应的形状 $s_{t+1} = s''$, $s'' = G^{-1}(\bar{s} + P\mathbf{b}, \Theta)$,其中 G^{-1} 代表相关的逆相似变换, Θ 为对应的相似变换参数;

(4)比较相邻两次迭代结果 s_t 和 s_{t+1} 之间的差别,如果它们之间足够小,则宣布算法收敛,否则,转到过程(2),继续新的迭代。

4 实验分析

实验所用样本包括160幅近乎正面的人脸图像^[10],其中涵盖了40个不同的人物,每人4幅,图像的大小为 120×160 。每幅图像总共由58个特征点手工标定,分别代表了眉毛、眼睛、鼻子、嘴巴和面部轮廓等关键部位,如图1所示。选择其中120幅作为训练样本,分别来自于30个人,剩余的40幅作为测试样本。另外,从YALE B选择40幅作为测试集的补充^[11],该库中的图象涵盖了不同的光照背景差别。

4.1 局部最小问题的解决

利用加权ASM模型,通过内部纹理模型和外部纹理模型面部特征点的强制约束,可以有效地解决局部最小问题,图5显示了两种方法比较的结果,从图中可以看出,传统的ASM方法有时在面部特征定位时不能奏效,利用加权ASM可以将这些点拉回到更准确的位置。

4.2 搜索精度性能的比较

实验的过程大致如下:首先将测试集中的每一幅图像对应的准确标定点的坐标分别沿 x 坐标和 y 坐标偏移不同的幅值,最高可达 ± 8 个像素,每幅图像偏移多次,然后以此偏

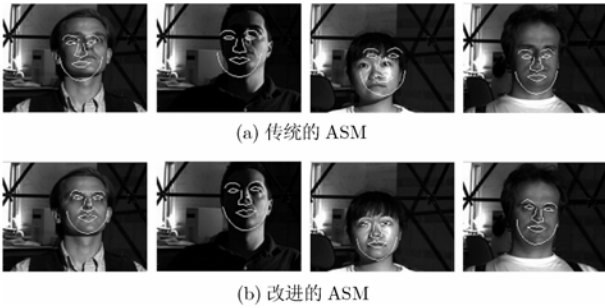


图5 局部最小问题的解决

移作为初始位置，用 ASM 方法进行迭代收敛，以期偏移点重新回到原来准确的位置。图 6 显示了比较的结果，其中水平坐标表示收敛得到的点和手工标定点之间的距离，垂直坐标表示在偏移当前水平坐标值的情况下，收敛点所占的比例。从图中可以看出，在收敛点和准确标定点距离相差较小的时候(水平坐标值小的情况)，改进算法对应的纵轴百分比要高于原来的算法，这说明此时的收敛点的数目多，和原来准确点拟合的程度好；相反，在水平坐标值大的情况，改进算法对应的纵轴百分比要略低于原来的算法，这说明偏离准确位置的点少。两者的比较表明，改进的加权 ASM 算法能更精确的定位面部的关键特征点。

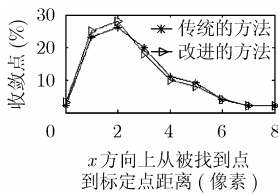


图 6 精度性能比较

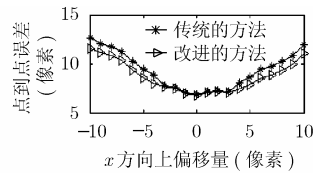


图 7 搜索范围比较

4.3 搜索的范围比较

将每一幅测试集中的准确标定点偏离它原来的准确位置(通常偏移到平均形状的位置)，然后以此为起始位置，采用 ASM 算法进行搜索，使其尽量回到原来的最佳位置，计算搜索得到的关键点和原来手工标定点之间的距离差值。图 7 显示了搜索范围比较的结果，从两条曲线上可以看出，同样误差的情况下，改进的算法搜索范围相对宽一些。

5 结束语

作为统计建模方法中很重要的一种，主动形状建模在面部特征定位和人脸识别等模式识别领域中有着广泛的应用。然而，由于受到初始情况、光照等诸多因素的影响，主动形状建模的性能会有所下降，经常会陷入最优化过程中的局部最小问题。本文基于传统的主动形状模型，从 ASM 最原始的思想出发，构建了一种加权的 ASM 方法。实验结果表明，加权 ASM 方法不仅能够有效地解决了最优化过程中的局部最小问题，而且在定位面部关键特征点上精度更高，搜索范

围也比传统的方法更宽。

参考文献

- [1] Beinglass A and Wolfson H J. Articulated object recognition, or: How to generalize the generalized Hough transform[C]. In proceedings IEEE computer society conference on computer vision and pattern Recognition, Hawaii, 1991: 461-466.
- [2] Kass M, Witkin A, and Terzopoulos D. Snakes: Active contour models[C]. 1st International Conference on Computer Vision, London, 1987: 259-268.
- [3] Wiskott L, Fellous J M, and Kruger N, *et al.* Face recognition by Elastic Graph Matching[C]. Interlligent Biometric Techniques in Fingerprint and Face Recognition, Eds. Jain L.C. *et al.*, Washington DC.,pub1. CRC Press, 1999: 355-396.
- [4] Nastar C and Ayache N. Fast segmentation, tracking and analysis of deformable objects[C]. International Conference on Computer Vision, Berlin, 1993: 275-279.
- [5] Cootes T F and Taylor C J, *et al.* Active shape models – their training and application[J]. *Computer Vision and Image Understanding*, 1995, 61(1): 38-59.
- [6] Cootes T F, Edwards G J, and Taylor C J. Active appearance models[C]. European Conference On Computer Vision. Berlin, 1998, vol.2: 484-498.
- [7] Zhao M, Li S Z, Chen C, and Bu J J. Shape evaluation for weighted active shape models[C]. Asian Conference on Computer Vision, Korea, 2004 vol.2: 1074-1079.
- [8] Jiao F, and Li S Z, *et al.* Face alignment using statistical models and wavelet features[C]. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Wisconsin, 2003, Vol.1: 321-327.
- [9] Cootes T F, and Hill A, *et al.* The use of active shape models for locating structures in medical images[J]. *Image & Vision Comput.*, 1994, 12(6): 355-366.
- [10] Stegmann M B, Ersböll B K, and Larsen R. FAME—A flexible appearance modeling environment[J]. *IEEE Trans. on Medical Imaging*, 2003, 22(10): 1319-1331.
- [11] Athinodoros S G and Peter N B. From few to many: Illumination cone models for face recognition under variable lighting and pose[J]. *IEEE Trans. on PAMI*, 2001, 23(6): 643-660.

冀贞海：男，1976 年生，博士生，研究方向为图像处理和模式识别。
 孙 宁：男，1978 年生，博士生，研究方向为图像处理和模式识别。
 邹采荣：男，1963 年生，博士，教授，博士生导师，研究方向为多维数字信号处理。
 赵 力：男，1958 年生，博士，教授，博士生导师，研究方向为语音信号处理和情感信息处理。