

# 改进的 ASM 方法在人脸定位中的应用

刘爱平<sup>1</sup>, 周 焰<sup>1</sup>, 关鑫璞<sup>2</sup>

(1. 空军雷达学院自动化教研室, 武汉 430019; 2. 国防科技大学电子科学与工程学院, 长沙 410073)

**摘要:** 在传统 ASM 模型的基础上, 采用以轮廓点为中心的矩形区域的 DCT 系数进行轮廓匹配, 提出了一种改进算法。该方法充分利用了轮廓点附近的二维纹理信息, 改善了 ASM 方法采用轮廓点附近一维灰度信息导致模型匹配精度不高的问题。DCT 具有良好的去相关性能和能量集中性能, 仅采取 3.52% 的 DCT 系数却平均保留了图像 95% 的能量, 并采用八方向搜索最佳轮廓点, 提高了定位速度。实验证明该方法非常有效。

**关键词:** 主动形状模型; 主动外观模型; 主成分分析; 离散余弦变换; 人脸定位

## Application of Improved Active Shape Model in Face Positioning

LIU Ai-ping<sup>1</sup>, ZHOU Yan<sup>1</sup>, GUAN Xin-pu<sup>2</sup>

(1. Automation Faction, Airforce Radar Academy, Wuhan 430019 ;

2. School of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha 410073)

**【Abstract】** This paper presents an improvement on the conventional active shape model(ASM) for face alignment. In this method, DCT coefficients of rectangular domain centering the lineament points are used in lineament matching, and a 2-D texture pattern search-and-fitting scheme instead of an 1-D analysis is used to improve the precision of the ASM method. DCT has good capability in abandoning correlation of data and concentrating energy of images. It withholds 95% energy of all by extracting only 3.52% DCT coefficients and also enhances the orientation speed. Experiment results demonstrate the efficacy of the method in locating facial features.

**【Key words】** active shape model(ASM); active appearance model(AAM); principle component analysis(PCA); discrete Cosine transformation (DCT); face positioning

随着计算机软硬件和图像处理技术的飞速发展, 人脸定位跟踪越来越成为人们关注的研究课题。由于人脸是柔性体, 在形状、纹理、姿态、图像质量等方面有很大的不定性, 因此人脸的跟踪定位极具挑战性。

在过去的 20 多年中, 基于统计学的可变形模型在变形物体定位研究方面取得了成功。近年来的研究寻找到一些典型的可变形模型, 最有名的是 Kass 提出的主动轮廓模型(active contour model, ACM)<sup>[1]</sup>。ACM 采用限制能量最小的方法进行形状校准, 能够保证得到封闭的轮廓, 但对具有较高曲率的凸形边缘的检测效果不佳; ACM 的收缩性使得它不能检测物体内部的轮廓, 没有结合先验知识收敛速度较慢。Kirby 和 Sirovich 在 ACM 的基础上描述了利用面部灰度水平统计建模, 却忽略了脸的可变特性<sup>[2]</sup>。Cootes 提出的主动形状模型(active shape model, ASM)<sup>[3]</sup>和主动外观模型(active appearance model, AAM)<sup>[4]</sup>是目前目标物体定位研究中应用较新的方法。ASM、AAM 这 2 种方法已应用于 CT 图像定位<sup>[5]</sup>、骨骼评价定位<sup>[6]</sup>、行为识别等方面。

传统 ASM 方法仅提取了轮廓的一维信息, 限制了搜索范围, 容易收敛到错误的位置。Bemhard Froba<sup>[7]</sup>采用特征点梯度值作为 ASM 模型的统计信息, 采用特征一维信息, 定位精度没有很大改善。本文在传统 ASM 的基础上提出了一种改进算法, 该方法采用轮廓点附近的矩形区域的 DCT 系数作为匹配特征, 充分利用了轮廓点附近的纹理特征, 提高了定位精度; DCT 具有良好的去相关性能和能量集中性能, 仅采用 DCT 矩阵左上角的少量系数就可以保留图像的大部分能量, 匹配时采用八方向搜索, 提高了定位速度。

### 1 ASM 方法简介

ASM 是一种基于模型的特征匹配方法, 它既可以灵活地改变模型的形状以适应目标形状不确定的特性, 又将形状的变化控制在模型允许的范围内, 从而保证模型改变时不会受各种因素影响而出现不合理的形状。ASM 方法主要有以下 3 个步骤:

(1) 图像采样获得形状向量和轮廓点特征信息

建立模型需要手工标定训练图像。选择  $N$  幅人脸图像(包括多个人的不同表情和姿态), 每一幅图像手工标定  $n$  个轮廓点作为训练数据。轮廓点一般标定在能够代表目标轮廓的地方, 本文选择的轮廓点标记在脸的外部轮廓和器官的边缘。标定点为

$$s_i = (x_{i1}, y_{i1}, x_{i2}, y_{i2}, \dots, x_{in}, y_{in})^T, \quad i = 1, 2, \dots, N \quad (1)$$

其中,  $(x_{ij}, y_{ij})$  代表第  $i$  幅图像的第  $j$  个轮廓点的坐标;  $n$  代表每幅图像标记的点数;  $N$  代表训练图像数目; 每一幅图像的  $s_i$  代表一个形状向量。并获取每一个标定点(轮廓点)附近的特征信息, 这些特征是进行匹配的主要依据<sup>[3,7,8]</sup>。

(2) 建立模型

边缘标定是在各自不同的样本图像上进行操作。由于各个样本图像拍摄条件、分辨率的差异, 得到形状向量的坐标具有不同的比例尺寸, 因此要对样本图像的形状向量归一化,

**基金项目:** 国家自然科学基金资助项目(94101010)

**作者简介:** 刘爱平(1982 -), 女, 硕士研究生, 主研方向: 图像处理, 模式识别; 周 焰, 副教授、博士后; 关鑫璞, 博士研究生

**收稿日期:** 2006-10-24 **E-mail:** guanxinpu@nudt.edu.cn

通过旋转、平移、缩放使得它们在同一坐标系中表示时具有一致性。图 1 表示了对样本图像进行归一化前后的比较，可以看出，对样本图像归一化之后其边缘标定具有更清晰的轮廓，标定点更加集中。

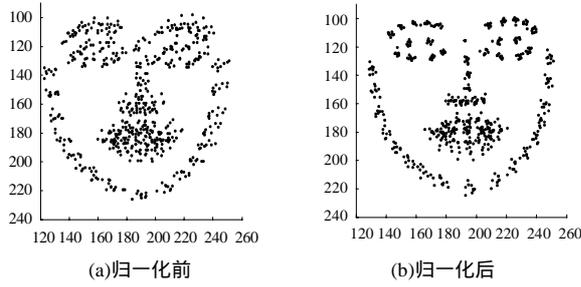


图 1 向量归一化前后比较

各个样本向量之间存在一定的相关性，且样本数量比较大，会造成计算量过大，必须进行简化处理。采用主成分分析(PCA)的思想对样本空间进行正交变换，并确定主成分，经 PCA 计算可得到主要变化参数<sup>[3,5,6,8]</sup>。从训练样本集中获得轮廓的模型统计信息：平均形状向量  $\bar{s} = \frac{1}{N} \sum_{i=1}^N s_i$ ； $N$  个形状向量的协方差阵  $C = \frac{1}{N} \sum_{i=1}^N (s_i - \bar{s})(s_i - \bar{s})^T$ ；协方差  $C$  的特征值以及特征值对应的特征向量，并且把特征值按照降序排列  $\lambda^i = \{\lambda_1, \lambda_2, \dots, \lambda_{2n}\}$ ，其中， $\lambda_i \geq \lambda_{i+1} (i = 1, \dots, 2n-1)$ ，求出特征值  $\lambda$  对应的特征向量  $P^i = \{P_1, P_2, \dots, P_{2n}\}$ 。选取  $t$  个大特征值  $\sum_{i=1}^t \lambda_i / \sum_{i=1}^{2n} \lambda_i \geq 98\%$ ，得到  $\lambda^i = \{\lambda_1, \lambda_2, \dots, \lambda_t\}$  以及对应的特征向量  $P = \{P_1, P_2, \dots, P_t\}$ ， $P$  代表主成分。

PCA 分析后，任意的人脸形状向量  $S$  可以表示成一个平均形状向量与  $t$  个形状向量  $P_i$  的线性组合： $S = \bar{s} + Pb$ ，其中， $b = (b_1, b_2, \dots, b_t)^T$ ，因为  $P$  是标准正交的： $PP^T = I$ ，所以  $b$  可以表示成： $b = P^T(S - \bar{s})$ 。 $b$  是形状参数，控制前  $t$  个模式的系数，不同的  $b$  对应不同的形状。

### (3) 定位轮廓

在 ASM 中一般采用基于轮廓的灰度匹配法，主要思想是对训练样本中的每个轮廓点沿着轮廓的法线方向采样作为灰度特征。 $g_{ij}$  表示第  $i$  幅图像的第  $j$  个采样点的灰度特征。所有图像第  $j$  个点的灰度特征平均值为

$$\bar{g}_j = \sum_{k=1}^M g_{kj} \quad (2)$$

协方差矩阵为

$$G_j = \frac{1}{M} \sum_{i=1}^M (g_{ij} - \bar{g}_j)(g_{ij} - \bar{g}_j)^T \quad (3)$$

灰度平均值和协方差作为该点的匹配特征。对每一个特征点作同样的采样分析可得整个轮廓的统计特征。这些统计信息呈高斯分布，可采用 Mahalanobis 距离作为匹配函数，如式(4)：

$$f(g_{si}) = (g_{si} - \bar{g}_i)^T G_i^{-1} (g_{si} - \bar{g}_i) \quad (4)$$

ASM 利用轮廓的灰度特征进行一个迭代过程，在每一步迭代中通过调整参数  $b$  改变当前模型的位置和形状完成模型与测试图像轮廓的匹配<sup>[3,5,6]</sup>。

## 2 基于 DCT 特征的 ASM 方法

选取什么样的特征作为 ASM 各个特征点的统计特征直接关系到迭代的最终模型能否逼近真实目标。传统的 ASM 方法只是采用轮廓点附近的一维灰度信息作为匹配特征<sup>[3,5]</sup>，信

息量少，会使人脸某些轮廓点的灰度信息相近，不利于下一步的目标配准。本文采用以特征点为中心的矩形区域  $B$  的 DCT 系数作为匹配特征。

### 2.1 光照度修正(光照补偿)

特征点的局部纹理信息常常受到光照不均的影响，这些都会影响获取特征点的局部统计特征信息，对于小矩形区域  $B$ ，计算出像素值  $P(x, y)$  的均值  $\mu_B$ 、方差  $\sigma_B$ ，通过式(5)来减少光照的影响：

$$P_N(x, y) = \frac{P(x, y) - \mu_B}{\sigma_B} \quad (5)$$

### 2.2 DCT 特征的计算

离散余弦变换(DCT)是一种常用的正交变换，可以降低图像的空间冗余，广泛应用于图像压缩处理中。没有经过处理的图像数据量非常大、数据间的相关性较强。采用 DCT 变换，可以去掉无关数据且不影响关键的属性和特征，使得能量集中在矩阵左上角的低频分量附近。

DCT 是酉变换的一种，其变换前后的信号熵和能量不变，其变换核可分离的正、逆 DCT 分别定义如下：

$$C(u, v) = a(u)a(v) \times \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (6)$$

$$u, v = 0, 1, \dots, N-1$$

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \{a(u)a(v)C(u, v) \times \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]\} \quad (7)$$

$$x, y = 0, 1, \dots, N-1$$

其中，

$$a(u) = \begin{cases} \sqrt{1/N} & \text{当 } u = 0 \\ \sqrt{2/N} & \text{当 } u = 1, 2, \dots, N-1 \end{cases}$$

本文利用 DCT 变换提取人脸轮廓点的特征。图像经 DCT 变换后，能量集中在矩阵左上角的低频分量部分，因此，本文采用 zigzag 扫描方式，扫描 DCT 变换矩阵的左上角，使能量按照从大到小的顺序排列。图 2 为使用 zigzag 扫描某图像块的 DCT 系数矩阵，箭头指明了扫描方向。

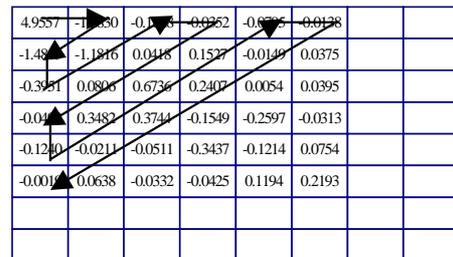


图 2 使用 zigzag 扫描矩阵

本文抽取以特征点为中心的  $M \times M$  矩形区域进行 DCT 变换，利用矩阵左上角的低频分量作为这个轮廓点的匹配特征。如果  $M$  取值较小不利于获得丰富的特征信息，会降低匹配精度； $M$  取值较大能提高轮廓点的信息量，但一方面信息量的提高有饱和趋势且占用空间大，另一方面搜索匹配特征点时，运算量增大使得速度受限。本文考虑到轮廓点信息量、计算速度、存储空间，选取了  $32 \times 32$  的矩形区域进行 DCT 变换。

图 3 是以某轮廓点为中心的  $32 \times 32$  矩形图像块 DCT 变换后，按照 zigzag 扫描方法得到的矩阵左上角 120 个 DCT 系数。其中，横坐标代表这 120 个 DCT 系数，纵坐标代表 120 个 DCT 系数的数值分布。从图 3 中看到，距离零点越近，数值幅度越大；距离零点越远，数值逐步趋于 0，且能量主要集中在矩阵左上角。

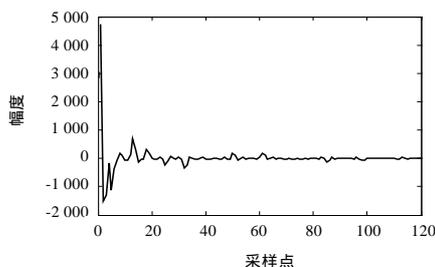


图3 DCT系数

计算以这个特征点为中心的矩形区域 DCT 变换后的能量  $E=31\ 267\ 000$ ，以及矩阵左上角 36 个 DCT 系数的能量  $E=31\ 060\ 000$ ，即矩阵的 3.52% 数据集中了图像 99.34% 的能量。由于各图像块局部性质各不相同，集中图像能量的能力并不完全一样，但总的规律是：各变换系数块中能量主要集中在低频区域。本文采取的矩阵数据平均集中了图像 95% 的能量。

图 4 为 8 组取主要能量变换后的图像与原图的对比，第 1 行是以某些轮廓点为中心的矩形区域原图；第 2 行是相对应的第 1 行图像经 DCT 变换并且按照前面提到的规则取 3.52% DCT 系数反变换后的图像。通过 8 组图像内部比较可以看出 DCT 变换不仅可以提取轮廓点附近的纹理信息，使轮廓点的特征信息更加丰富，还可以去掉不必要的噪声。

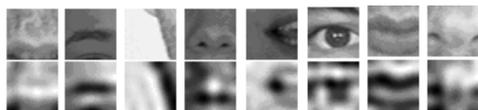


图4 原图与 3.52% 的 DCT 系数反变换后的图像

### 2.3 特征点配准 (轮廓点匹配)

候选点和各特征点的统计特征间的差别反映候选点与实际点的接近程度，因此，采用下面的匹配公式：

$$F(B, \bar{B}) = \sum_{i=1}^K [C_i(B) - C_i(\bar{B})]^2 \quad (6)$$

其中， $C_i(B)$  代表某个轮廓点周围矩形区域  $B$  的第  $i$  个 DCT 系数； $C_i(\bar{B})$  代表所有训练图像某个轮廓点平均矩形区域  $\bar{B}$  的第  $i$  个 DCT 系数； $\bar{B}$  的  $K$  个 DCT 系数作为匹配特征信息。本文只提取了每个轮廓点矩形区域的 3.52% 的数据，即实际参与计算的 DCT 系数的数量  $K \times M \times M$ 。

传统搜索最佳点的方法是在法线方向上搜索，当最佳匹配点不在这条法线方向上时，模型就不能很好地逼近目标。为了降低这种失误率，本文采用八方向搜索，在 8 个方向上各搜索  $m$  个点，利用匹配准则寻找最佳的特征点。具体的匹配过程如下：

- (1) 初始化参数  $b$ ，生成模型实例  $S = \bar{S} + Pb$ ；
- (2) 找到当前模型在测试图像中各个轮廓点 DCT 特征；
- (3) 对每一点  $(x_i, y_i)$  进行八方向搜索，找到它的最佳匹配点  $(x'_i, y'_i)$ ，对每一点都进行同样的搜索匹配得到一个新的  $S'$ ；
- (4) 使用  $b = P^T(S - \bar{S})$  更新参数  $b$ ，产生新的模型实例；
- (5) 重复(2)~(4)直到模型实例的变化小到一定的范围。

通过调整参数  $b$  改变当前模型的位置和形状，产生新模型实例并使用新模型替换原来的模型，重复匹配过程，使模型不断向目标逼近。为了保证形状的变化受控，通常要限制参数  $b$ ，一般调整参数  $b$  满足： $-3\lambda_i \leq b_i \leq 3\lambda_i$ 。

### 3 仿真结果与分析

应用上述基于 DCT 特征的 ASM 模型对 BioID 人脸数据

库中的部分图像进行了训练、测试。总共选择了分辨率为  $384 \times 286$  的 200 幅人脸图像，120 幅图像作为训练样本，建立先验模型。选择其余的 80 幅图像以及本文自拍摄的部分图像作为测试图像。训练样本采样需要对每一幅训练图像标定 68 个轮廓点，标定点选取的是否准确直接关系到特征提取和最终匹配的效果。

图 5 为改进 ASM 方法与传统 ASM 方法对 BioID 人脸数据库中部分图像仿真结果进行的比较。从图 5 中可以看到，按传统 ASM 方法定位第 1 幅图，外轮廓以及鼻梁没有很好地逼近，第 2 幅图中嘴巴、鼻子的定位效果不佳。本文改进的 ASM 算法能对传统 ASM 定位效果不好的图像精确定位。



图5 传统 ASM 方法与本文改进方法的定位比较

图 6 为使用改进的 ASM 方法对自采集图像定位前后比较图。上一行图像为初始状态，下一行图像为最终定位效果。



图6 改进 ASM 方法定位效果

为了检验改进 ASM 模型的匹配精度，本文采用轮廓形状模型距离作为标准，使用 ASM 算法匹配后的各轮廓点与手动标记各点之间的距离取平均，检测匹配精度：

$$d = \frac{1}{n} \sum_{k=1}^n \sqrt{(x_k - x'_k)^2 + (y_k - y'_k)^2} \quad (7)$$

其中， $n$  代表每幅图像的标定点数； $(x_k, y_k)$  代表标定坐标； $(x'_k, y'_k)$  代表使用模型匹配后的坐标。对于每一幅测试图像都利用式(7)分别计算 ASM 定位结果与标定轮廓距离，以及使用改进 ASM 算法定位结果与标定轮廓距离。图 7 横坐标代表匹配结果和手工标定结果的平均点误差像素值，纵坐标代表对应 X 轴的实例百分数。从图 7 中可以看到，平均误差点

(下转第 241 页)