

# Adaptive Particle Filtering for Facial Feature Tracking

Xiaorong Pu, Peng Li, Luping Ji

School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu

Email: puxiaor@uestc.edu.cn

Received: Aug. 31st, 2011; revised: Sep. 27th, 2011; accepted: Oct. 15th, 2011.

**Abstract:** Active shape model (ASM) has been widely used to track a face from a video sequence. However, it is usually limited to frontal view or the cases of small-scale head movement. ASM may fail in condition of cluttered background, adverse circumstances and partial occlusion. An enhanced ASM is proposed to meet those challenges based on particle filtering tracking algorithm. Experiments demonstrate the effectiveness and flexibility of the proposed algorithm in tracking the target objects undergoing significant variation of the object's appearance or surrounding illumination.

**Keywords:** Visual Tracking; Particle Filter; Incremental PCA; ASM

## 基于自适应粒子滤波的 ASM 新技术

蒲晓蓉, 李 鹏, 纪禄平

电子科技大学计算机科学与工程学院, 成都

Email: puxiaor@uestc.edu.cn

收稿日期: 2011 年 8 月 31 日; 修回日期: 2011 年 9 月 27 日; 录用日期: 2011 年 10 月 15 日

**摘 要:** ASM 已经被广泛应用于视频流中的人脸跟踪, 但大多局限于头部运动较缓慢的正面人脸跟踪, 很难精确跟踪复杂背景、恶劣光照和部分遮挡条件下的人脸。本文结合粒子滤波预测跟踪人脸, 改进传统 ASM 方法, 实现人脸特征点的精确定位。实验表明, 该方法能有效提高 ASM 搜索的精确度和鲁棒性。

**关键词:** 人脸跟踪; 增量 PCA; 粒子滤波; ASM 搜索

### 1. 引言

人脸跟踪和特征提取是人脸识别和机器视觉的关键技术。主动形状模型(Active Shape Model, ASM)是由 Cootes<sup>[1]</sup>等人于 1995 年提出的一种基于点分布模型算法, 已经被广泛应用于单幅人脸或视频流的人脸特征提取和跟踪。ASM 基于统计方法, 将一系列人脸图像进行标点训练得到原始形状向量, 再进行对齐操作和 PCA 分析建立形状建模<sup>[2]</sup>。其中需要设置形状模型参数, 约束其变化范围, 确保模型能适应目标形状的不确定性<sup>[3]</sup>。

ASM 模型的初始化过程非常重要, 如果不能准确定位目标图像中的人脸位置, 后续的 ASM 搜索过程将会失败。传统人脸检测技术包括基于肤色<sup>[4]</sup>、Harr

特征<sup>[5]</sup>等方法, 大多局限于良好背景、均匀光照和人脸无遮挡等约束环境, 无法检测视频流中快速头部移动的人脸。为了自动确定跟踪目标, 本文方法仅视频流的第一帧利用人脸自动检测捕获目标, 获得人脸位置、大小和旋转等参数, 并将人脸的位置作为粒子滤波人脸跟踪算法的初始化窗口。此后采用粒子滤波方法定位特定目标的人脸位置, 以检测并跟踪复杂背景、光照变化、局部遮挡, 以及快速移动的人脸图像。

ASM 大多采用 PCA 方法建立形状模型, 需要大量数据集构建特征基。但利用传统 PCA 不能准确定位复杂状况下的人脸目标, 且无法适应目标特征的最新变化。本文采用递增 PCA(Incremental PCA, IPCA)方法构建更新特征空间, 保持目标模式的最新表达, 以适应目标本身的变化和周围环境的改变, 实现复杂环

境下的人脸特征点跟踪<sup>[6]</sup>。

本文后续部分将依次介绍 ASM 方法, 基于 IPCA 的 ASM 人脸跟踪方法和自适应粒子滤波人脸跟踪技术, 描述验证实验, 讨论改进 ASM 方法的实验效果。

## 2. ASM 方法

ASM 是一种基于统计模型的图像处理方法, 通过对具有一定代表性的同类目标图像进行统计建模, 以获得反映目标图像二维形状变化规律的形状统计模型。ASM 首先对一组标有特征点的图像进行形状建模和局部纹理建模, 然后在搜索过程中不断调节形状和姿态参数, 从而使形状达到最优化。

ASM 需要对一组标有特征点的图像进行形状建模和局部纹理建模, 该过程需要大量的人脸样本作为训练集。

### 1) 形状建模(Shape Model)

形状建模旨通过训练多幅图像序列, 获得平均形状、转换矩阵、搜索需要的初始形状和模型约束的相关参数。主要包括特征点标定、对齐操作和 PCA 构建形状模型三个步骤。

#### (a) 特征点标定

ASM 算法首先需要对训练集中的每幅人脸图像进行特征点标定, 刻画人脸形状。图像中的每个特征点都被标注唯一的固定序列编号, 训练集中每幅图像的特征点构成的形状用  $2n \times 1$  的向量表示为

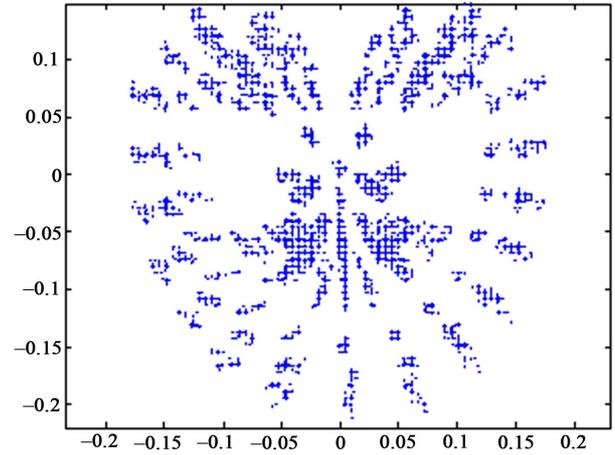
$\mathbf{x} = (x_1, \dots, x_n, y_1, \dots, y_n)^T$ , 其中,  $x_1, x_2, \dots, x_n$  表示有序特征点的  $x$  坐标,  $y_1, y_2, \dots, y_n$  表示有序特征点的  $y$  坐标,  $n$  表示训练集的大小。

#### (b) 对齐操作

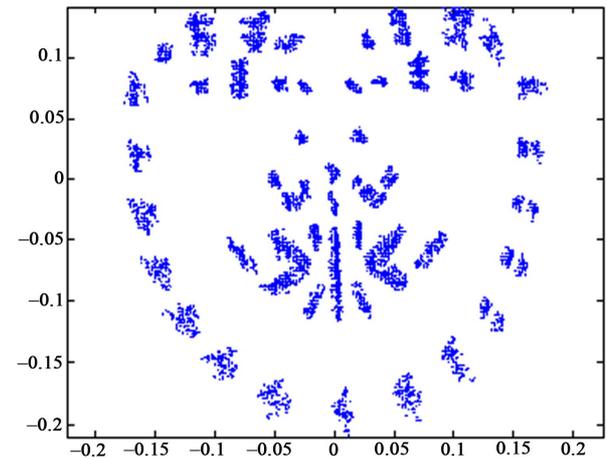
为了获得训练集中模型点坐标的统计特征, 使不同样本对应点之间具有可比性, 需要消除训练样本大小、位置等差异造成的影响, 即需要进行对齐操作。由于每个点分布向量具有旋转不变性, 一个形状可以通过旋转、缩放、平移等仿射变换对齐另一个形状, 使得两个形状之间的差异最小。两个形状向量  $\mathbf{X}_i$  和  $\mathbf{X}_j$  的对齐操作可通过平移、缩放、旋转实现, 最后求解变换后的两个形状向量之间的最小距离。

从图 1 中不难发现, 经过对齐操作以后, 特征点的分布更加集中。

#### (c) 用 PCA 建立形状模型



(a) 对齐操作前



(b) 对齐操作后

Figure 1. Alignment samples  
图 1. 样本对齐前后的比较示意

形状对齐操作以后, 对样本集进行主成分分析 (PCA), 可获取样本集在分布空间中的主方向。形状建模以后, 一个形状可以表示为:  $\mathbf{X} \approx \bar{\mathbf{X}} + \mathbf{P}\mathbf{b}$ 。其中,  $\mathbf{P}$  是通过 PCA 得到的前  $t$  维特征向量矩阵,  $\mathbf{b}$  是形变参数, 表示任意人脸在特征空间中的投影。为了约束人脸形状, 一般限制  $\mathbf{b}$  的范围为  $|b_i| \leq 3\sqrt{\lambda_i}$ 。

### 2) 局部纹理建模(Profile Model)

第一步, 获得每个特征点的局部纹理。比较训练集中的各图像的某个特定点周围的灰度分布。沿着特征点法线方向进行像素级灰度值采样。特征点的法线定义为: 当前点的前一个点和后一个点连线的中垂线。

第二步, 进行局部纹理建模。将第  $i$  个样本的第  $j$  个特征点记为  $p_{ij}$ , 采样的第  $k$  个灰度值记为  $g_{ijk}$ , 则  $p_{ij}$  的  $2k+1$  个灰度层可记为:

$$g_{ij} = (g_{ij1}, g_{ij2}, \dots, g_{ij(2k)}, g_{ij(2k+1)})^T.$$

### 3) 特征点搜索

通过形状建模和局部纹理建模以后, 利用建模获得的参数对待定位的图像进行目标搜索。

搜索过程中, 两个特征点之间的相似度的匹配函数可用马氏距离(Mahalanobis Distance)表示, 定义为:  $f_j(g_j) = (g_j - \bar{g})^T S^{-1} (g_j - \bar{g})$ 。其中,  $\bar{g}$  表示第  $j$  个特征点的灰度层差分均值,  $g_j$  表示第  $j$  个特征点的第  $i$  个灰度差分,  $S$  为第  $j$  标记点局部纹理的协方差矩阵。

通过迭代不断调整姿势和形状参数, 获得每个点的最佳匹配。不过, 这只能使每个点的匹配最优, 而非整体形状匹配最优, 即不能保证得到的人脸形状就是最终的人脸形状。因此, 需要对模型进行约束, 将当前搜索的形状投影到模型空间, 并约束模型中  $b$  值的范围, 以保证每个特征点在局部和整体上最优。

## 3. 基于 IPCA 的 ASM 人脸跟踪

### 3.1. 递增主成分分析 IPCA (Incremental PCA)

由于目标表观容易受形状、自身反射特性、姿态和光照等因素影响, 需要一个稳定的表观模型来适应这些因素的影响。可以提前采集目标在各种光照和复杂环境下的样本, 利用 PCA 求解样本的主元方向, 并进行降维。传统 PCA 本质上是在最小均方意义下寻找最代表原始数据的投影, 从而简洁有效地表示原始数据。IPCAS 利用最新获得的图像更新特征空间, 获得的特征空间能很好地描述图像之间的相似性或差异性<sup>[7]</sup>。

经过光照补偿和去噪等图像预处理后, 得到二维灰度图<sup>[8]</sup>。可以用一个  $d \times n$  的矩阵  $A = \{I_1, \dots, I_n\}$  表示经预处理的训练图集, 存储  $1 \sim n$  时刻的图像向量, 其中  $I_i$  是一个  $d$  维向量, 代表第  $i$  帧图像。对矩阵  $A$  进行 SVD 分解:  $A = U \Sigma V^T$ , 得到  $A$  的特征值。当一个  $d \times m$  维的新观测图像矩阵到达时, 可以将其表示为  $[A E] = [U \tilde{E}] \begin{pmatrix} \Sigma & U^T E \\ 0 & \tilde{E}^T E \end{pmatrix} \begin{pmatrix} V^T & 0 \\ 0 & I \end{pmatrix}$ , 记为  $k+m$  维的方阵  $R$ , 其中  $k$  为  $\Sigma$  的特征值个数。  $R$  的 SVD 分解为  $p_{ijR} = \tilde{U} \tilde{\Sigma} \tilde{V}^T$ , 它可在初始化时得到, 而不必随目标图像的更新而更新。于是,  $[A E]$  的 SVD 分解可表示为

$$[A E] = ([U \tilde{E}] \tilde{U}) \tilde{\Sigma} \begin{pmatrix} \tilde{V}^T & V^T \\ 0 & I \end{pmatrix}.$$

采用 SKL 方法更新特征基, 首先对  $A$  进行 SVD 分解得到  $U$  和  $\Sigma$ , 然后按下述方法实现

$$A' = [AE] = U' \Sigma' V'.$$

1) 获得旧数据  $A = U \Sigma V^T$  和新到达的数据  $E$ ;

2) 对  $[U \Sigma E]$  进行 QR 分解得到  $\tilde{E}$  和  $R$ ,  $[U \Sigma E] = [U \tilde{E}] R$ ;

3) 对  $R$  进行 SVD 分解  $R = \tilde{U} \tilde{\Sigma} \tilde{V}^T$ ;

4) 若第三步分解后基向量的个数少于其非零特征值的个数, 则丢弃  $\tilde{U}$  和  $\tilde{\Sigma}$  中多余的特征向量和特征值;

5) 使得  $U' = [U \tilde{E}] \tilde{U}, \Sigma' = \tilde{\Sigma}$ 。

由上述算法可知,  $A = \{I_1 I_2 \dots I_n\}$  和  $B = \{I_1 I_2 \dots I_n\}$  分别表示原始图像和新到达的图像, 它们的组合记为  $C = [A B]$ , 分别记  $A$  和  $B$  的样本均值和散射矩阵为  $I_A I_B I_C$  和  $S_A S_B S_C$ , 给定  $(A - I_A)$  的 SVD 分解  $U$  和  $\Sigma$ , 通过 IPCA 方法计算  $(C - I_C)$  的 SVD 分解, 得到相应的  $U'$  和  $\Sigma'$ 。随着新样本数据的到达, 更新训练数据的样本均值和散射矩阵。

### 3.2. 粒子滤波跟踪

粒子滤波通过寻找一组在状态空间中传播的随机样本, 对概率密度函数  $P(X_k | Z_{1:k})$  进行逼近,  $X_k$  为系统状态,  $Z_{1:k}$  为系统的观测值, 用样本均值代替积分运算, 从而获得状态最小方差估计, 这些样本称为“粒子”<sup>[9]</sup>。数学描述为: 对于平稳的随机过程, 若  $k-1$  时刻系统的后验概率密度为  $P(X_k | Z_{1:k-1})$ , 依据一定原则选取  $n$  个随机样本点,  $k$  时刻获得测量信息后, 经过状态和时间更新过程,  $n$  个粒子的后验概率密度可近似为  $p(X_k | Z_k)$ 。随着粒子数目的增加, 粒子的概率密度函数逐渐逼近状态的概率密度函数, 粒子滤波估计即达到了最优贝叶斯估计的效果。

视频流相邻两帧图像内的目标跟踪可以看成目标在图像中位置的仿射变换。目标在  $t$  时刻的仿射变换表示为  $X_t = (x, y, s, \theta)$ , 其中,  $x, y$  分别表示目标区域中心在  $x$  和  $y$  方向的坐标,  $s$  表示比例因子,  $\theta$  表示目标区域的旋转角度。对于给定的一系列观测图像  $I_t = \{I_1 I_2 \dots I_t\}$ , 根据贝叶斯理论, 有

$$p(X_t | I_t) \propto p(I_t | X_t) \int p(X_t | X_{t-1}) p(X_{t-1} | I_{t-1}) dX_{t-1}.$$

记观测模型为  $P(X_t|I_t)$ ，通过观测  $I_t$  的状态，计算状态  $X_t$  的概率，建立动态时变系统

$P(X_t|X_{t-1}) = N(X_t; X_{t-1}, \Psi)$ ，其中， $\Psi$  代表仿射参数  $X_t$  的协方差矩阵。这样，就可以用表观模型表示被跟踪目标，用 PCA 的概率表示图像的观测模型。样本训练图集经过特征分解以后得到一个中心为  $\mu$  的子空间  $U$ ，待跟踪图像  $I_t$  中的目标概率  $P$  可以用样本到空间中的参考点的马氏距离  $d$  求得。 $d$  可以分解为空间  $I_t$  和  $U$  之间的距离  $d_t$  和样本点到  $U$  的正交子空间的距离  $d_w$  的乘积<sup>[10]</sup>：

$$P(I_t|X_t) = P_{d_t}(I_t|X_t)P_{d_w}(I_t|X_t) \quad (1)$$

由于空间产生的样本概率满足高斯分布，可以表示为：

$$P_{d_t}(I_t|X_t) = N(I_t; \mu, UU^T + \varepsilon I) \quad (2)$$

其中， $I$  为单位矩阵， $\mu$  为样本均值， $\varepsilon I$  表示观测过程中的高斯噪声。

子空间内样本的概率可以用样本均值的马氏距离建模：

$$P_{d_w}(I_t|X_t) = N(I_t; \mu, U \Sigma^{-2} U^T) \quad (3)$$

式(2)和(3)结合，可以计算出  $P(I_t|X_t)$ <sup>[10]</sup>。

采用粒子滤波方法可以跟踪局部遮挡、光照变化和多角度等人脸图像，并能在丢失后恢复目标跟踪。

## 4. 改进的 ASM 算法流程

### 4.1. 基于粒子滤波跟踪的 ASM 形状搜索过程

搜索目标的初始位置对 ASM 的影响非常大。如果搜索位置偏离了搜索目标或与搜索目标的大小偏差较大，搜索过程很容易陷入局部最小，导致整个搜索失败。目前，大多数 ASM 搜索算法都基于人工手动设置初始搜索位置，而非自动确定初始搜索位置。若利用人脸自动检测实现人脸特征定位，须严格限制背景、光照等条件，否则人脸检测失败会导致整个搜索过程的失败。

为了自动确定跟踪目标，本文方法自视频流的第一帧就利用人脸自动检测捕获跟踪目标。自动人脸检测能确定视频流中人脸的确切位置，获得人脸位置、大小和旋转等参数，并将人脸位置作为粒子滤波人脸

跟踪算法的初始化窗口。

利用 3.2 节描述的粒子滤波跟踪算法，实现复杂条件下的人脸特征点跟踪。具体算法描述如下：

(a) 如果是视频流的第一帧，则用自动人脸检测方法定位人脸位置；否则，用粒子滤波方法定位第一帧中特定目标的人脸位置；

(b) 估计当前人脸脸颊的特征点位置，将人脸框的宽度进行适当缩小和平移，再将新人脸框的左上角点记为左脸颊位置，右上角点即为右脸颊位置；

(c) 利用(b)得到的两个脸颊特征点的位置坐标与平均形状模型中脸颊两个特征点的坐标之间的对应关系，求得仿射变换矩阵的参数  $Ia, b, t_x, t_y$

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

(d) 利用仿射变换，将平均形状模型中的其它点也进行相应映射，获得 ASM 初始形状搜索位置，开始 ASM 搜索过程。自适应 ASM 跟踪流程如图 2 所示。

### 4.2. 改进后 ASM 的实验结果

实验针对复杂背景、人脸局部遮挡、强光照等情况，分别采用传统 ASM 和基于自适应滤波的 ASM 方法进行人脸跟踪，实验图像选用 640×480 像素。部分实验效果如图 3 所示。

采用传统 ASM 方法时，若第  $k$  帧人脸检测失败，将导致整个 ASM 搜索失败。由于视频流中连续几帧图像中的人脸位置变化不会很大，利用这一特性可大致估算第  $k$  帧的人脸位置。根据前几帧的人脸检测结果，取连续前几帧图像中人脸位置进行加权平均，分别赋予不同的权重，估算后一帧的人脸位置。

由图 3 可见，采用传统 ASM 进行人脸检测与跟踪，在复杂背景、遮挡和恶劣光照条件下均会失败。

实验结果表明，采用自适应粒子滤波的 ASM 方法跟踪复杂背景、局部遮挡、强光照等视频图像中的人脸，具有更高的精确性和鲁棒性。

## 5. 总结

本文提出基于自适应粒子滤波的 ASM 方法，能自动确定跟踪目标的初始位置，并能在复杂背景、局部遮挡和强光照等图像中检测跟踪人脸，较传统 ASM

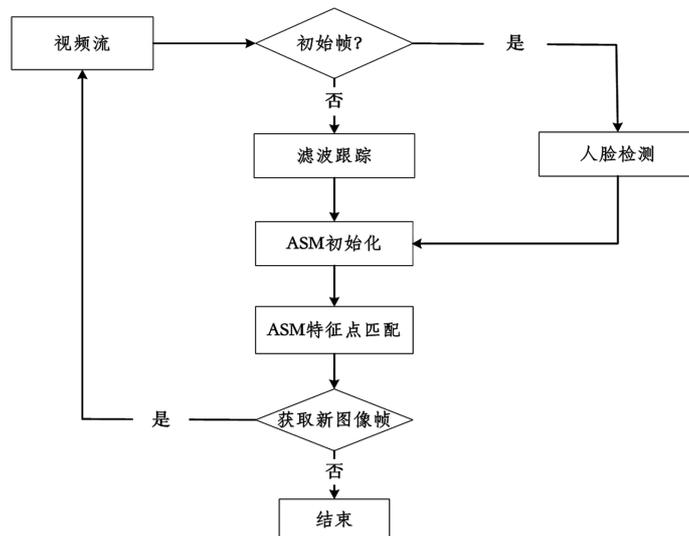


Figure 2. Flow chart of adaptive ASM for face tracking  
图 2. 自适应 ASM 算法跟踪流程图



Figure 3. Comparative experimental results  
图 3. 传统 ASM 与改进 ASM 人脸检测与跟踪对比

方法，改进的 ASM 方法具有更强的鲁棒性、精确性和实时性。

## 6. 致谢

感谢国家自然科学基金 60973070 对本项目的支持。

## 参考文献 (References)

[1] T. F. Cootes, C. J. Taylor, et al. Active shape models-their training and application. *Journal of Computer Vision and Image Understanding*, 1995, 61(1): 38-59.

[2] T. F. Cootes, G. J. Edwards and C. J. Taylor. Active appearance models. *Proceedings of European Conference on Computer Vision*, 1998, 2: 484-498.

[3] T. F. Cootes, C. J. Taylor. Statistical models for appearance for computer vision, 2004. <http://www.isbe.man.ac.uk/bim/refs.html>

[4] E. Hjelm, B. K. Low. Face detection: A survey. *Computer Vision and Image Understanding*, 2001, 83(3): 236-274.

[5] P. Viola, M. J. Jones. Robust real-time object detection. *International Journal of Computer Vision*, 2001, 57(2): 137-154.

[6] D. Ross, J. Lim and M. H. Yang. Adaptive probabilistic visual tracking with incremental subspace update. *Proceedings of European Conference on Computer Vision*, 2004, 2: 470-482.

[7] D. A. Ross, J. Lim, R.-S. Lin and M.-H. Yang. Incremental learning for robust visual tracking. *International Journal of Computer Vision*. 2008, 77(1-3): 125-141.

- [8] J. Shermina. Illumination invariant face recognition using discrete cosin transform and principal component analysis. Proceedings of International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT), Tamil Nadu, 23-24 March 2011: 826-830.
- [9] C. Shen, A. van Hengel and A. Dick. Probabilistic multiple cue integration for particle filter based tracking. International Conference on Digital Image Computing: Techniques and Applications, Sydney, 2003, 1: 399-408.
- [10] B. Moghaddam, A. Pentland. Probabilistic visual learning for object recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 696-710.