

# 基于均值移动与贪婪算法的多人脸跟踪

蔡灿辉 刘 强

(华侨大学信息科学与工程学院, 泉州 362021)

**摘要** 提出一个基于均值移动(Mean Shift)和贪婪算法的多人脸跟踪器。首先建立多个均值移动目标跟踪器以进行多人脸跟踪。结合卡尔曼滤波逐个检测目标并从视频帧中清除已跟踪到的人脸,以解决当多个目标相邻或相互遮挡时相应的跟踪窗口会收敛于最大目标、导致其他目标丢失的难题。引入辅助窗口并根据其纹理信息确定粘连目标的对应。实验结果表明,该多人脸跟踪算法可实现稳健的实时多人脸跟踪。

**关键词** 人脸跟踪 多目标跟踪 均值移动 贪婪算法

**中图法分类号:** TP301.6 **文献标识码:** A **文章编号:** 1006-8961(2009)12-2611-04

## Multiple Face Tracking Using Mean Shift and Greedy Algorithm

CAI Can-hui, LIU Qiang

(School of Information Science and Technology, Huaqiao University, Quanzhou 362021)

**Abstract** A novel multiple face tracking algorithm is proposed in this paper. Multiple Mean Shift trackers are first built to enable multiple face tracking. To overcome the weakness of Mean Shift tracking, which is prone to converge to the local maximum target if tracked objects are adjacent or partially occluded, a greedy tracking method is used to pursue the targets one by one, during which a Kalman filter is first employed to locate the initial position, and then the tracked object is removed from the scene to guarantee no other Mean Shift tracker iterates the same target. An accessory window featured with local texture distribution is introduced to correspond to candidate windows and targets. Experimental results have indicated the proposed algorithm can track multiple faces robustly in real time.

**Keywords** Face tracking, multiple target tracking, Mean Shift, greedy algorithm

## 1 引言

多人脸跟踪技术是计算机视觉及相关领域中的具有挑战性的课题,在视频处理、安全监控、人机交互等方面具有重要的意义。均值移动(Mean Shift)<sup>[1]</sup>算法具有计算复杂度低,可用于不同亮度下对不同大小、形状的物体进行跟踪的优点,得到了广泛的应用。但该算法在整个跟踪过程中,目标模型与跟踪窗口的大小保持不变,且容易收敛到局部极值,导致跟踪目标丢失。针对上述问题,不少专家学者提出了改进算法<sup>[2-3]</sup>,但多个窗口容易收敛到局

部最大目标的问题仍没有解决。为此,本文提出了结合均值移动和贪婪算法的多人脸跟踪方案,有效地解决了当两个人脸相邻时均值移动算法容易收敛到面积较大的人脸(局部最大值),导致面积较小的人脸丢失的难题;并通过引入卡尔曼滤波器和辅助窗口,解决了不同目标的个体对应问题。

## 2 改进的均值移动算法

### 2.1 均值移动(Mean Shift)算法介绍

均值移动算法是一种基于核密度估计的无参快速模式匹配算法<sup>[1]</sup>,其基本思想就是不断地把跟踪

**基金项目:**国家自然科学基金项目(60772164);福建省自然科学基金项目(A0710009);福建省科技计划项目(2005H034)

**收稿日期:**2009-09-04;**改回日期:**2009-09-11

**第一作者简介:**蔡灿辉(1954~),男,教授。2003年于天津大学获博士学位。主要研究方向为图像处理。E-mail:chcai@hqu.edu.cn

窗口中心向样本均值移动使之收敛。用于跟踪时,通常采用目标模型和候选区域所对应的颜色直方图之间的相似性(即 Bhattacharyya 系数)作为收敛准则,把跟踪问题转化为模式匹配问题。均值移动算法可简单地描述为:

(1) 计算跟踪窗口内样本的均值,调整跟踪窗口的中心到样本均值位置。

(2) 重复(1),直到跟踪窗口中心和质心“会聚”。

均值移动目标匹配算法不需要进行全局搜索,计算简单,实时性好。但由于参数固定,缺乏必要的模型更新方法,目标外观变化将导致模型不能很好地刻画当前的目标特征。另外,当目标运动速度较快或多个目标粘连及存在遮挡等干扰时,会丢失目标造成跟踪失败。

本文借助卡尔曼滤波器<sup>[4]</sup>对跟踪目标进行预测,采用类似文献[2]的方法连续自适应地调整跟踪窗口的大小,取得了良好的跟踪效果。

## 2.2 卡尔曼滤波器的参数设定

卡尔曼滤波器是一种常用的状态估计快速算法,能够较准确地对目标的位置、速度做出预测。在跟踪过程中,由于相邻两帧图像之间的时间间隔较短,目标运动状态变化较小,所以可假设目标在单位时间间隔内是匀速运动,即假设系统为线性状态模型。令  $k$  时刻目标窗口中心位置和运动速度分别为  $(x_{ck}, y_{ck})$  和  $(v_{xc_k}, v_{yc_k})$ , 观测向量为  $\mathbf{z}_k = [xz_k, yz_k]$ , 则  $k$  时刻的系统状态可表示为  $\mathbf{x}_k = [x_{ck}, y_{ck}, v_{xc_k}, v_{yc_k}]$ 。令  $\Delta T$  为  $k$  时刻和  $k-1$  时刻的时间差,则状态转移矩阵  $\mathbf{A}$  和观测矩阵  $\mathbf{H}_0$  可分别表达为

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{H}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

## 2.3 跟踪窗口的自适应调整

基于人脸的特征,可把跟踪窗口定为椭圆形。令椭圆的长、短轴分别为  $2h_x$  和  $2h_y$ , 根据文献[3]对核密度公式的推导,可得:

$$h_x = \sqrt{\frac{2 \sum_{i=1}^{n_k} w_i (x_{i,x} - y_x)^2}{\sum_{i=1}^{n_k} w_i}} \quad (1)$$

$$h_y = \sqrt{\frac{2 \sum_{i=1}^{n_k} w_i (x_{i,y} - y_y)^2}{\sum_{i=1}^{n_k} w_i}} \quad (2)$$

式中,  $y_x, y_y$  是该跟踪窗口的中心,  $x_{i,x}, x_{i,y}$  是第  $i$  个样本点的坐标,  $w_i$  是权重系数。引入权重系数的目的是降低边缘像素对目标的影响,使目标被部分遮挡时,仍有较好的跟踪效果。

## 2.4 目标模型的建立

研究表明,肤色的差异更多表现在亮度而不是色度<sup>[5]</sup>。为了减少肤色受照明强度的影响,本文选用 HSV 空间的 H 分量建立肤色概率分布模型。H 分量和 RGB 各分量的映射关系为

$$H = \arccos \frac{2R - G - B}{2 \sqrt{(R - G)^2 - (R - B)(G - B)}} \quad (3)$$

通过采集肤色样本,把肤色样本中的每个像素从 RGB 空间转换到 HSV 空间,然后计算肤色样本 H 分量的统计直方图,并对该直方图进行归一化处理,就可以得到肤色在 H 空间的概率分布  $f(H)$ , 该分布就是跟踪目标的颜色分布。

为了利用目标的空间信息,本文采用反向投影技术<sup>[2]</sup>把目标模型和候选区域的概率分布映射到观测图像。所谓的反向投影就是对观测图像中的像素进行  $H \rightarrow f(H)$  映射,把观测图像中的每个像素的 H 分量值用肤色直方图中对应的  $f(H)$  代替,所得到图像称为观测图像的颜色概率分布图像 (probabilistic distribution image, PDI)。这时, PDI 中每一点的值就代表该像素在肤色统计直方图下的出现概率。由于肤色统计直方图中非肤色成分很少,因此,非肤色区域经过反向投影后数值也很小,接近黑色。在本文的改进均值移动算法中,采用目标 PDI 与候选区域 PDI 间的相似性作为收敛准则。

## 2.5 改进的均值移动算法描述

综上所述,可得改进后的均值移动算法如下:

- (1) 建立目标颜色概率模型,初始化搜索窗口;
- (2) 把每一帧输入图像转化为概率分布图像 (PDI);
- (3) 采用均值移动算法调整搜索窗口的中心使之和样本质心“会聚”。
- (4) 根据系统状态  $\mathbf{x}_k$ , 用卡尔曼滤波器预测下一时刻目标窗口的中心位置,用公式(1)和公式(2)

计算跟踪窗口的大小。

改进的均值移动算法在跟踪单目标时不受跟踪目标的形状、大小和光照等环境因素的影响,具有较强的抗噪声干扰能力。

### 3 多人脸跟踪算法的实现

#### 3.1 构建多人脸跟踪器

实现多人脸跟踪的第一步就是为每一个目标(人脸)构建一个改进的均值移动跟踪器。考虑到人脸的特征,把每个跟踪窗口定为椭圆形。并对各目标窗口编号,存入初始窗口队列。

#### 3.2 贪婪跟踪算法

均值移动是一种局部最优化搜索算法。当两个具有相似颜色分布的物体相邻时,则两个窗口都会跟踪面积较大的物体,导致目标丢失。为了正确跟踪多个相邻或部分遮盖的人脸,本文把贪婪算法引入多人脸跟踪,提出了如下贪婪跟踪策略:(1)利用改进的均值移动算法找到最大跟踪窗口存入窗口队列;(2)对相应的区域置零,消除已搜索到目标;(3)继续运行改进的均值移动程序寻找下一个目标。重复该算法步骤直到找到该局部区域的所有目标。图 1 给出了该贪婪跟踪算法的部分实验结果。其中上半部分给出了具有两张人脸的视频序列的 5 帧图像的 PDI,下半部分是经过一次均值移动运算并清除已跟踪的人脸后的 PDI。为了指示两张人脸的位置,图中分别用红、黄两种颜色填充这两张人脸在 PDI 图中的对应位置。其中,中间帧较小的人脸完全被较大的人脸遮住了。从图中可以看出,除了完全遮挡的情况外,面积较小的人脸也可以跟踪得到。换言之,采用贪婪跟踪算法,可成功地跟踪相邻和部分遮盖的目标。

#### 3.3 跟踪窗口匹配

完成了目标搜索后,下一个问题是如何解决当前帧的各个目标(窗口)与前一帧各目标之间的对应关系,即目标鉴别问题。对于两张距离较远的人脸,根据其位置信息与运动信息就可以轻易地解决这个问题。但对于窗口相邻或相互遮挡的目标而言,首先搜索到的是面积较大的人脸,而对同一人脸来说,相对于镜头的距离不同,面积会发生变化,因此,必须增加额外的辅助信息才能解决窗口对应问题。本文采用衣服的纹理作为附加辅助信息。在初始化跟踪窗口后,在跟踪窗口下方

选择一个  $30 \times 30$  像素的辅助窗口,将该辅助窗口的局部纹理模型(local binary patterns, LBP)<sup>[6]</sup> 作为一个辅助信息。当没有发生遮挡时,辅助窗口的纹理信息与主窗口同步更新。实验结果表明,在辅助窗口没有被遮挡时,可结合卡尔曼滤波器的系统状态和辅助窗口的 LBP 信息确定窗口间的对应关系。当辅助窗口被遮挡时,也可根据上述系统状态估计其对应关系,等遮挡消失后再确认其对应关系。必须指出,当窗口间出现遮挡时,较大的(强势)窗口对应的辅助窗口一般不会被遮挡,这时强势目标窗口可辨识(对应关系可确定),则另一窗口的对应关系也完全可以确定。这样就很好地解决了个体对应问题。限于篇幅,建立 LBP 的方法详见文献[6]。

#### 3.4 目标更新

为了动态刷新人脸进入或退出场景的信息,本系统每 10 帧启动一次 Adaboost 人脸检测<sup>[7]</sup> 程序检测场景中的人脸,更新目标窗口。当边界附近的目标连续 10 帧不出现,则认为该目标已退出监控区域,删除该目标窗口。

#### 3.5 算法描述

本文提出的多人脸跟踪算法可归纳如下:

- (1)初始化已知目标、待定目标和临时窗口队列,清零已知目标、待定目标和临时窗口数。
- (2)对当前帧,计算色度信息,产生色度帧。若已知目标、待定目标数均为零,转步骤(7)进行目标更新。
- (3)从窗口队列中选择一个窗口,启动改进的均值移动算法,找到对应的目标。计算该目标辅助窗口的纹理信息。将该窗口及其辅助信息存入临时窗口队列。
- (4)在色度帧上清除该目标。
- (5)转步骤(3)重复上述过程,直到所有已知目标、待定目标均处理完毕。
- (6)根据卡尔曼滤波器的系统状态和辅助窗口的 LBP 纹理对临时窗口队列的所有目标进行匹配,把已匹配对象存入已知目标队列,并更新辅助窗口的纹理信息,否则存入待定目标队列,不更新辅助窗口的纹理信息。
- (7)如果该帧号为 10 的整数倍,启动改进的 Adaboost 程序搜索新人脸。对新增人脸进行编号,建立目标窗口。计算目标的质心、面积、PDI 和辅助窗口的纹理信息,存入目标队列。

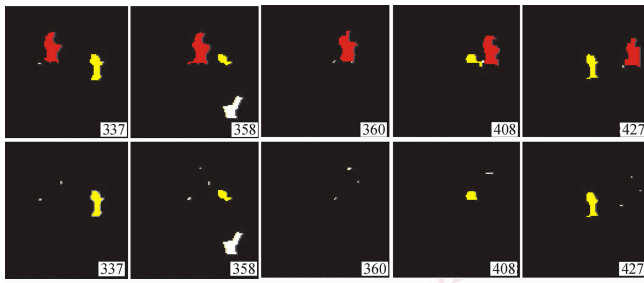


图 1 贪婪跟踪的仿真结果

Fig. 1 Simulating results of the greedy tracking strategy

## 4 仿真结果及讨论

在 Pentium IV 3.0 GHz, 512 MB RAM 的微机平台上用多个自己录制的多人脸序列对本文提出的人脸跟踪算法进行测试, 每个序列长度为 5 分钟。表 1 列出了部分实验数据, 图 2 给出部分实验结果。

表 1 多人脸跟踪仿真结果

Tab. 1 Multiple faces tracking result

序列	1	2	3
大小	352 × 288	352 × 288	352 × 288
人脸数	2	3	4
人脸被遮挡帧数	15	23	33
时间 (ms/帧)	26.16	27.36	29.83



图 2 对 3 人脸系列跟踪结果

Fig. 2 Tracking three faces by the proposed algorithm

从图 2 中可以看出, 当人脸进入场景后系统对目标自动更新, 分别建立红、兰、黄 3 个目标窗口。随后兰色目标窗口向其右方运动, 被黄色目标窗口遮挡, 目标丢失。遮挡过后系统又重新检测到 3 个

目标, 并对 3 个目标窗口进行正确匹配。接着红色目标窗口走到前方遮挡黄色目标窗口, 再沿原路线返回。利用卡尔曼滤波器的系统状态和辅助窗口的 LBP 纹理信息黄色目标窗口仍能匹配到自己的目标。可见, 本文的算法具有很强的鲁棒性。从表 1 可知, 本文提出的多人脸跟踪算法平均在每帧 CIF 图像中对多张人脸的跟踪需要运算时间不足 30 ms, 可以满足实时要求。

## 5 结 论

本文提出一个均值移动和贪婪算法相结合的多人脸跟踪算法。通过建立多个改进的均值移动跟踪器跟踪多人脸, 结合贪婪算法解决了多个均值移动跟踪器同时收敛到局部最大值导致目标丢失的难题。借助卡尔曼滤波器和辅助窗口成功地解决了目标个体辨识问题。引入 Adaboost 算法实现了目标的自动更新。实验结果证实了本文提出的实时多人脸跟踪算法的可靠性与有效性。

## 参考文献 (References)

- Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. IEEE Transactions Pattern Analysis Machine Intelligence, 2003, 25(5): 564-575.
- Bradski G R. Computer vision face tracking as a component of a perceptual user interface [A]. In: Proceedings of IEEE Workshop Applications of Computer Vision [C], Princeton, NJ, USA, 1998: 214-219.
- Jiang Zhuo-Lin, Li Shao-fa, Gao Dong-fa. An adaptive mean shift tracking method using multiscale images [A]. In: Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition [C], Beijing, China, 2007: 1060-1066.
- Kalman R E. A new approach to linear filtering and prediction problems [J]. Transactions of the ASME-Journal of Basic Engineering, 1960: 35-45.
- Yang J, Lu W, Waibel A. Skin color modeling and adaptation [A]. In: Proceedings of the 3rd Asian Conference on Computer Vision [C], Hong Kong, China, 1998: 687-994.
- Tan X, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions [A]. In: Proceedings of IEEE International Workshop on Analysis and Modeling of Faces and Gestures [C], Rio de Janeiro, Brazil, 2007: 168-182.
- Viola P, Jones M. Robust real time object detection [A]. In: Proceedings of the 8th IEEE International Conference on Computer Vision [C], Vancouver, British Columbia, 2001: 747.