

以颜色和形状直方图为线索的粒子滤波人脸跟踪

高建坡 王煜坚 杨浩 吴镇扬

(东南大学无线电工程系, 南京 210096)

摘要 跟踪器的设计和跟踪线索的选择与表达是人脸跟踪中的两大关键因素,针对一般人脸跟踪算法中常用简单椭圆来描述人脸形状线索时易受背景干扰的缺点,以及视频目标跟踪中动态模型和观测模型的非线性非高斯特点,提出了一种以颜色和形状直方图为线索的粒子滤波人脸跟踪算法,该算法在粒子滤波基本框架之下,引入了一种新的用直方图来描述人脸形状的方法,并对其进行了改进,用来作为人脸跟踪的形状线索。同时,为了减轻背景干扰,提出了一种经验有效边缘的检测方法。实验表明,该跟踪方法不仅能有效地处理人脸旋转、背景中的肤色干扰和部分遮掩问题,并且能够在由于大面积遮掩等原因而丢失目标的情况下,及时有效地重新捕获已丢失的目标。

关键词 人脸跟踪 粒子滤波 CONDENSATION 形状描述 肤色模型

中图分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2007)03-0466-08

Particle Filter Face Tracking Using Color and Shape Histogram as Clues

GAO Jian-po, WANG Yu-jian, YANG Hao, WU Zhen-yang

(Department of Radio Engineering, Southeast University, Nanjing 210096)

Abstract Face tracking problem comprises two key factors; one is the design of tracker; the other is the selection and representation of tracking cues. To relieve the shortcoming that the general face tracer is frequently disturbed by background when a simple ellipse model is used to describe the facial shape, and in accordance with the fact that the dynamic model and observation model of video object tracking problem are always non-Gaussian and non-linear, a robust particle filter face tracking algorithm which uses shape and color represented in the form of histogram as clues is proposed. The tracking algorithm is based on particle filter, in which a new shape representation method using histogram is introduced and modified to be as one of tracking clues. At the same time, an experiential edge detection algorithm is proposed in order to alleviate background disturbance. Experimental results show that the proposed algorithm is robust to the rotation of face, background skin color distraction and partial occlusion. In addition, the proposed tracker can recover the lost target due to the severe occlusion.

Keywords face tracking, particle filter, CONDENSATION, shape representation, skin color model

1 引言

图像序列中人脸的跟踪,对于视频监控、人机交互、人脸识别和视频检索等领域都具有十分重大的意义。人脸跟踪的根本任务是从连续的图像序列中跟踪捕获人脸的运动轨迹和大小变化。由于人脸跟

踪实质上就是对人脸的状态(位置、大小等)进行动态的估计,因此,人脸跟踪可以看成是在给定观测量的情况下求解系统隐含状态量的估计问题。大家知道,Kalman 滤波是解决动态估计问题的一个有力工具,但是 Kalman 滤波只有在线性系统和噪声为高斯的情况下才是最优的,然而大多数视频跟踪场合,这种苛刻的条件是难以满足的,虽然扩展 Kalman 滤波

基金项目:国家自然科学基金资助项目(60672094)

收稿日期:2005-10-27;改回日期:2006-01-09

第一作者简介:高建坡(1975~),男,1998 获山东理工大学学士学位,2004 年获南京理工大学硕士学位,现为东南大学博士研究生。主要研究方向为图像处理、计算机视觉。E-mail: jianpogao@163.com

可以利用 Taylor 展开式对非线性系统进行线性近似,然而这种近似必然会带来信息的损失,而且扩展 Kalman 滤波仍然没有摆脱高斯噪声的限制。近年来,随着计算机性能的提高及成本的下降,一种新的基于贝叶斯原理的序贯蒙特卡罗粒子滤波^[1]方法逐渐受到研究人员的青睐,粒子滤波彻底摆脱了对系统为线性及噪声为高斯特性的要求,因为它是一种自然的方式处理非高斯多模态概率分布和其他一些非线性观测过程。粒子滤波在不同的场合有不同的称呼,比如蒙特卡罗滤波器、CONDENSATION (conditional density propagation)^[2-4]、自举滤波器等等。粒子滤波一经 Isard 引入到计算机视觉领域^[2]就迅速成为了一个重要的研究方向。

本文在粒子滤波器 CONDENSATION^[2-4]的框架之下,引入了一种新的用直方图的形式来描述人脸形状的方法,并对其进行了改进而作为人脸跟踪的形状线索,在此基础上,提出了一种以直方图描述的颜色和形状为线索的粒子滤波人脸跟踪方法;实验表明,该跟踪方法不仅能有效地处理人脸旋转、背景中的肤色干扰和部分遮掩问题,并且能够在由于大面积遮掩等原因而丢失目标的情况下,及时有效地重新捕获已丢失的目标。

2 跟踪线索

众所周知,图像中包含的信息量十分巨大,在图像序列中,可用于人脸目标跟踪的信息线索很多,如颜色、运动、形状、纹理等。这些信息线索各有特点,对不同的环境有着不同的鲁棒性,为了提高跟踪的可靠性,往往需要结合多种信息来共同完成复杂的跟踪任务^[5]。一个跟踪器的性能在很大程度上取决于跟踪线索的选择和表达,本文算法选择了颜色和形状两种特征线索用于人脸跟踪。

2.1 颜色

利用颜色信息进行图像分析和理解已有相当长的历史。在彩色空间中,一般而言,由于不同的物体往往聚集在各自的色度区域内,因此,颜色信息可以在图像中用于检测特定目标。经研究表明,由于人脸肤色稳定地分布在颜色空间中相对集中的范围,并且用肤色来定位人脸具有姿态不变性和简单易行的特点,因此人的肤色信息在人脸检测和跟踪问题中常作为一个有效的特征被广泛应用^[6-8]。建立一个基于肤色的人脸检测和跟踪系统,将面临着以下

3个问题:(1)采用哪种颜色空间;(2)如何建模人脸的肤色;(3)如何利用肤色模型来进行人脸检测或跟踪。

HSV 颜色空间是根据人感知颜色的方式建立的,因此比较符合人的视觉感受,该颜色空间明确区分了亮度和色度之间的特性。研究表明,HSV 空间 3 个分量的相关性从人的感觉上是很小的,从而使本空间在肤色分割中应用非常广泛^[6,7]。

本文就是在 HSV 空间,采用直方图建模的方法来描述颜色特征。为了减少光照的影响,通常情况下,对亮度 V 分量划分得粗略一些,本文对 H 、 S 、 V 3 个分量采用 $8 \times 8 \times 4$ 划分,即直方图的段数 M 总共为 256。某一目标区域颜色直方图 $H_o^c = \{h_o^c(u)\}_{u=1, \dots, M}$ 的构建方法为

$$h_o^c(u) = k \sum_{i=1}^N \delta[f(x_i) - u] \quad (1)$$

式中, k 为归一化常数, N 为目标区域像素总个数, u 为直方图段数索引值, $f(x_i)$ 是表示像素点 x_i 所在直方图段数的指示函数, δ 为 Kronecker delta 函数。

目标区域颜色直方图 H_o^c (上角 c 代表 color,下角 o 代表 object) 和人脸模型颜色直方图 $H_{face}^c = \{h_{face}^c(u)\}_{u=1, \dots, M}$ 之间的相似性采用以下 Bhattacharyya 距离来度量:

$$d_c = \sqrt{1 - \rho[H_o^c, H_{face}^c]} \quad (2)$$

式中, $\rho[H_o^c, H_{face}^c] = \sum_{u=1}^M \sqrt{h_o^c(u) h_{face}^c(u)}$,即为离散 Bhattacharyya 系数。

2.2 形状

Berkeley 加利福尼亚大学的 Belongie 等提出了形状上下文 (shape context) 的概念,具体请参见文献[9]。形状描述中一特征点的形状上下文指的是在某一指定极坐标系中形状上的其他特征点相对于该点的矢量长度及转角的统计直方图。本文借鉴文献[10]的思想,并对其进行改进而引入形状上下文来描述人脸形状,图 1 给出了一人脸图像建立形状上下文直方图的方法示例,从该图可以看出对于一人脸目标区域,建立形状上下文直方图的方法是首先对该区域进行边缘检测,并对边缘图像进行特征点抽取,为了抽样方便,本文采用了均匀抽样的方法;然后对抽取的每一特征点按文献[9]中的方法建立统计直方图,图 1(e) 给出了图 1(d) 中点 P 的形状上下文直方图示例;最后把所有特征点的形状上下文直方图拉直合并在一起,便可得人脸形状的

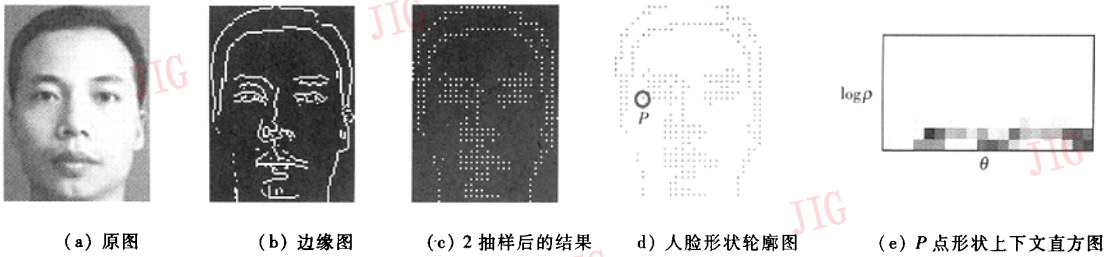


图 1 人脸形状上下文直方图示例

Fig. 1 The example of facial shape context

总体上下文直方图。上述用直方图描述形状的方法,由于统计的是点与点之间在某极坐标系中的相对距离和转角,显然具有平移不变性。根据所有距离的中值将点间距离归一化,这样可使直方图形状描述子保持一定的尺度不变性。同时,可以有多种方法使描述子做到旋转不变,本文采用了以所得形状特征点的特征主轴建立极坐标系的方法来完成旋转不变操作。另外,当存在部分人脸遮掩的情况时,由于那些没有被遮掩的特征点间的相对矢量长度及转角是不受影响的,所以这种以统计为工具的描述子还能在一定程度上处理遮掩效应。由此可见,上述用直方图描述形状的方法,可以胜任人脸跟踪的线索。

考虑到实际的人脸跟踪过程,在对某一目标区域进行边缘检测时,由于背景干扰,可能会形成一些不利于人脸形状分析的虚假人脸边缘特征点,因此为了有效避免背景,特别是复杂背景的干扰,本文提出了经验有效边缘的概念,所谓经验有效边缘就是指对人脸形状分析有用的边缘。不言而喻,由于这些边缘一定分布在人脸上,而且在人脸上的边缘像素点周围一般至少有一肤色像素点,因此,经验有效边缘像素点应该是那些梯度值超过一定阈值,且该像素及其四邻域至少有一点为肤色像素的点,由此可引入以下经验有效边缘的判断准则:

$$P(i, j) = (G(i, j) > T_{\text{gradient}}) \text{ 且 } (\bigcup_{(x,y) \in R_4} (10 < \tilde{V}(x,y) < 50)) \quad (3)$$

式中, $G(i, j)$ 为像素点 (i, j) 处的梯度值, T_{gradient} 为梯度阈值, R^4 为像素点 (i, j) 的四邻域, $\tilde{V}(x, y)$ 为 $Y\tilde{U}\tilde{V}$ 颜色空间^[8]中的分量 \tilde{V} , 统计实验表明,人脸肤色的 \tilde{V} 分量范围主要集中在 $10 < \tilde{V} < 50$ 。图 2 给出了利用本文提出的经验有效边缘检测和常规边缘检测的对比,显然,经验有效边缘检测大大抑制了非肤色背景的干扰,进而减少了形状上下文直方图的统计计算量。

对于包括 I 个边缘特征点的人脸模型,其总的形状上下文直方图可以表示成

$$H^S_{\text{face}} = \{h_i^{\text{face}}(j), i = 1, 2, \dots, I; j = 1, 2, \dots, K\} \quad (4)$$

式中, K 表示每个特征点对应形状上下文直方图的段数。同样,假设目标区域边缘特征点数为 J , 则目标区域总的形状上下文直方图可以表示成:

$$H^S_0 = \{h_i^0(j), i = 1, 2, \dots, J; j = 1, 2, \dots, K\} \quad (5)$$

对于 H^S_0 (上角 S 代表 shape) 中的任一 $h_i^0(\cdot)$, 总可以在 H^S_{face} 中找到一 $h_i^{\text{face}}(\cdot)$ 使得它们的直方图分布最接近, 可称之为匹配特征点。由此可定义 $h_i^0(\cdot)$ 和 $h_i^{\text{face}}(\cdot)$ 间的匹配指标为如下形式的 χ^2 距离:

$$C_i = \frac{1}{2} \sum_{j=1}^K \frac{[h_i^0(j) - h_i^{\text{face}}(j)]^2}{h_i^0(j) + h_i^{\text{face}}(j)} \quad (6)$$



图 2 边缘检测结果对比

Fig. 2 Comparison of the edge detection results

本文将目标区域和人脸模型总的形状相似性度量定义为所有特征点的匹配指标之和,即

$$d_s = \sum_{i=1}^J C_i \quad (7)$$

3 粒子滤波

粒子滤波是一种为处理后验概率密度 $p(X_t | Z_t)$ 和观测过程概率密度 $p(Z_t | X_t)$ 为非高斯的问题而发展起来的方法。粒子滤波,又称作蒙特卡罗滤波,它是通过非参数化的蒙特卡罗模拟方法来实现递推贝叶斯滤波,而贝叶斯滤波的基本思想是由先验概率结合观测量来得到后验概率,贝叶斯规则的数学表达式为

$$p(X | Z) = k p(Z | X) p(X) \quad (8)$$

式中, $p(X)$ 为先验概率, $p(Z | X)$ 为观测更新部分, k 为归一化常数。

对于动态演化过程,假设被跟踪量用状态矢量 X 表示, $Z_t = \{z_1, z_2, \dots, z_t\}$ 表示到目前时刻 t 为止的所有观测量集合,假定 z_1, z_2, \dots, z_t 相互独立,同时又独立于状态动态过程,那么对于一阶连续马尔可夫过程, t 时刻的条件概率密度函数可由以下两个过程得到:

预测:

$$p(X_t | Z_{t-1}) = \int_{X_{t-1}} p(X_t | X_{t-1}) p(X_{t-1} | Z_{t-1}) \quad (9)$$

更新:

$$p(X_t | Z_t) = k p(z_t | X_t) p(X_t | Z_{t-1}) \quad (10)$$

然而,对于一般的非线性、非高斯的动态系统,通常很难得到式(9)和式(10)所示的贝叶斯估计的解析解,而粒子滤波却巧妙地解决了这个问题。粒子滤波的核心思想是用一组有权采样集 $S = \{(X^{(n)}, w^{(n)}) | n = 1, 2, \dots, N\}$ (其中 w 为权重)来表示概率分布,且其表示精度随着粒子点数的增加可以达到足够高。

用于视频序列目标跟踪的粒子滤波形式 CONDENSATION^[2,3] 是因子取样的动态化推广,其系统状态随时间的演变采用随时间演变的采样粒子来模拟。这种方法状态的跟踪过程没有明确的显式解析定义和闭式的状态计算,它是以一种更自然的方式来处理非高斯非线性过程。具体而言,CONDENSATION 方法主要包括如下两个模型和 3 个过程:

(1) 两个模型

动态模型:用于指导采样集的传播。

观测模型:用于指导采样集权重的观测更新。

(2) 3 个过程

选择(重取样):用于完成对 $t-1$ 时刻粒子的选择,可有效减轻退化现象。

预测:根据动态模型来完成采样集的传播。

观测:用于完成对采样集的权重进行观测更新。

图 3 和图 4 分别给出了 CONDENSATION 方法框图及人脸跟踪实现示意图。

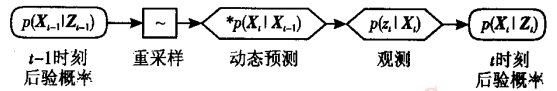


图 3 CONDENSATION 方法框图

Fig. 3 Block diagram of CONDENSATION algorithm

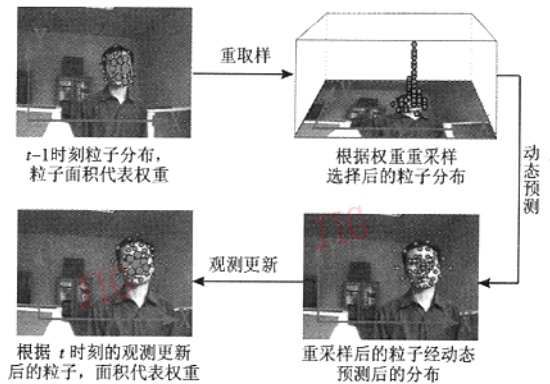


图 4 CONDENSATION 方法人脸跟踪示意图

Fig. 4 Sketch map of face tracking using CONDENSATION

4 人脸跟踪实现

4.1 动态模型

定义 $X = (x, y, \sigma, \Delta x, \Delta y, \Delta \sigma)$ 为系统状态变量。其中, x, y 表示人脸的中心位置, σ 表示人脸的大小(把人脸近似看成 1:1.2 的椭圆), $\Delta x, \Delta y, \Delta \sigma$ 分别表示 x, y 和 σ 的帧间动态变化。本文中, $t-1$ 时刻第 n 个粒子状态 $X_{t-1}^{(n)}$ 到 t 时刻第 n 个粒子状态 $X_t^{(n)}$ 的动态模型是采用一阶 AR (auto regression) 模型的形式,即 $X_t^{(n)}$ 可由 $X_{t-1}^{(n)}$ 经过下式确定性漂移和随机扩散得到:

$$X_t^{(n)} = AX_{t-1}^{(n)} + BN_{t-1}^{(n)} \quad (11)$$

式中 A 表示确定性部分, B 为随机部分, $N_{t-1}^{(n)}$ 表示 $t-1$ 时刻的噪声矩阵。这些参数的确定可以根据经

验设定,也可以通过图像序列的学习得到。

4.2 观测模型

由于本文人脸跟踪算法中采用了两种跟踪线索,因此对于每一滤波粒子的观测测量都由颜色观测测量和形状观测测量两部分组成,其中,

目标区域颜色观测测量为

$$w_c(z_c | X_t) \propto \exp(-\lambda_c \cdot d_c) \quad (12)$$

式中 d_c (下角标 c 代表 color) 表示 t 时刻采样粒子状态 X_t 下的目标区域颜色直方图和人脸模型颜色直方图间的 Bhattacharyya 距离。

目标区域的形状观测测量为

$$w_s(z_s | X_t) \propto \exp(-\lambda_s \cdot d_s) \quad (13)$$

式中, d_s (下角标 S 代表 shape) 表示 t 时刻采样粒子状态 X_t 下的目标区域形状直方图和人脸模型形状直方图间的形状相似性度量。

假设每一线索所提供的观测测量在统计上独立,则结合 t 时刻状态 X_t 下的颜色观测测量和形状观测测量,就可以给出状态 X_t 粒子的权重值,即

$$w = w_c(z_c | X_t)w_s(z_s | X_t) \quad (14)$$

4.3 部分采样粒子的初始化

为了有效地处理目标遮掩和其他一些特殊情况,本文借鉴了文献[4]的做法,即使动态模型中包含一定比例粒子的重新初始化,也就是独立于任何观测值,根据一个先验概率分布 $p(X_t)$ 产生一定数量的采样点。这将有助于应付由于遮掩等原因而引起的跟踪失败,以便重新获得跟踪目标。本文粒子动态演化始化过程可以表示为

$$\hat{p}(X_t | Z_{t-1}) \propto (1 - q)p(X_t | Z_{t-1}) + qp(X_t) \quad (15)$$

式中 q 为初始化比例调整因子。具体实现时,本文对其中一部分采样点按均匀分布随机产生,而且对已经漂移到该图像帧以外的采样点也按同样方法初始化,对剩下的采样点则采用动态模型演化得到。

4.4 本文算法总结

对于本文中的一次帧间人脸跟踪过程,也就是通过对 $t-1$ 时刻的粒子状态集合 $S_{t-1} = \{(X_{t-1}^{(n)}, w_{t-1}^{(n)}) | n = 1, 2, \dots, N\}$ 到 t 时刻粒子状态集合 $S_t = \{(X_t^{(n)}, w_t^{(n)}) | n = 1, 2, \dots, N\}$ 的演化来得到 t 时刻人脸状态的方法如图 5 所示。

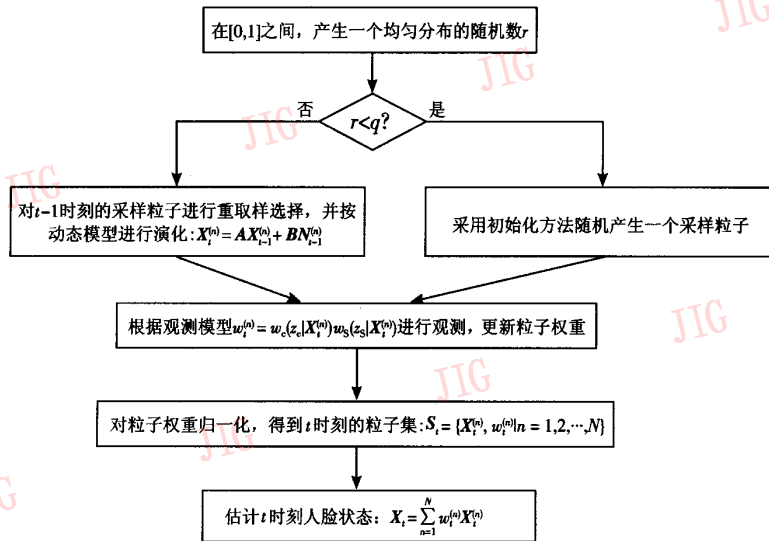


图 5 本文人脸跟踪算法框图

Fig. 5 Block diagram of the face tracking algorithm in this paper

5 跟踪实验结果与分析

为了验证本文方法的跟踪效果,进行了大量的人脸跟踪实验,实验数据来自于斯坦福大学视觉实验室(Stanford Vision Laboratory)所提供的专门为检

验人脸跟踪算法所设计的人脸序列^[11]以及笔者自己在实验室里所拍制的人脸序列。实验中,本文选择了3个有代表性的序列(斯坦福大学的 seq_mb、seq_sb 序列和一自拍序列)来进行跟踪结果的分析比较。

基于粒子滤波的目标跟踪已经被证明有着一些

其他跟踪方法不可比拟的优点,而当今不同的基于粒子滤波框架下的人脸跟踪方法的一个主要差异就是如何构建观测环节,也就是选择何种图像信息来进行观测度量。考虑到本文的主要工作在于为粒子滤波框架下的人脸跟踪引入一种新的用直方图来描述人脸形状的方法,并根据实际跟踪环境的复杂背景干扰问题,为此提出了经验有效边缘的概念,以减少背景干扰。为了充分体现本文算法的优越性,下面给出了本文算法同现行的一些典型的基于粒子滤波框架下的人脸跟踪方法的分析比较。

基于颜色的粒子滤波人脸跟踪是一种非常普遍的现行人脸跟踪算法^[6],该方法以单一的颜色为线索来进行人脸跟踪。其优点是速度快,但是肤色受环境的影响大,且易受肤色背景的干扰。从图6(a)所给出的基于颜色粒子滤波人脸跟踪的部分结果可以明显看出,该跟踪算法容易受到肤色背景干扰,如seq_mb序列中与肤色相近的门、seq_sb序列中的手掌和自拍序列中的书橱等,这些干扰使得部分跟踪区域偏离了目标。

目标的轮廓形状是粒子滤波跟踪中常用的一种信息^[3,5],其对目标的轮廓描述方式虽然多样,但对目标轮廓形状的观测度量的出发点却都是基于图像梯度信息的。如今用椭圆边界来近似人脸是一种常用的方法^[5,12],而结合颜色和椭圆边界形状进行粒子滤波人脸跟踪则是另一种常用的现行人脸跟踪算法。图6(b)给出了基于颜色和椭圆边界的粒子滤波人脸跟踪的部分结果,从跟踪结果中可以看出,椭圆边界形状信息的加入,虽然在一定程度上减少了肤色背景的干扰,但是由于对目标椭圆边界形状的观测度量是简单基于图像梯度信息的,因此其容易受到背景,特别是复杂背景、人脸器官等大图像梯度区域以及遮挡现象的影响,使得部分跟踪结果不够准确。

而本文的人脸跟踪算法采用的跟踪线索为颜色直方图和形状直方图,由于这两种线索相互补充,因此显然可有效地减少肤色背景干扰。大家知道,颜色直方图本身所固有的优点之一就是目标一定范围内的旋转、尺度变化及遮挡具有较强的鲁棒性。对于形状直方图线索而言,当存在部分人脸遮挡时,那些没有被遮挡的特征点间的相对矢量长度及转角是不受影响的,因而此时用直方图的方式去统计描述特征点间的相对矢量长度及转角,这在很大程度上能有效处理遮挡效应。另外,笔者对目标区域和

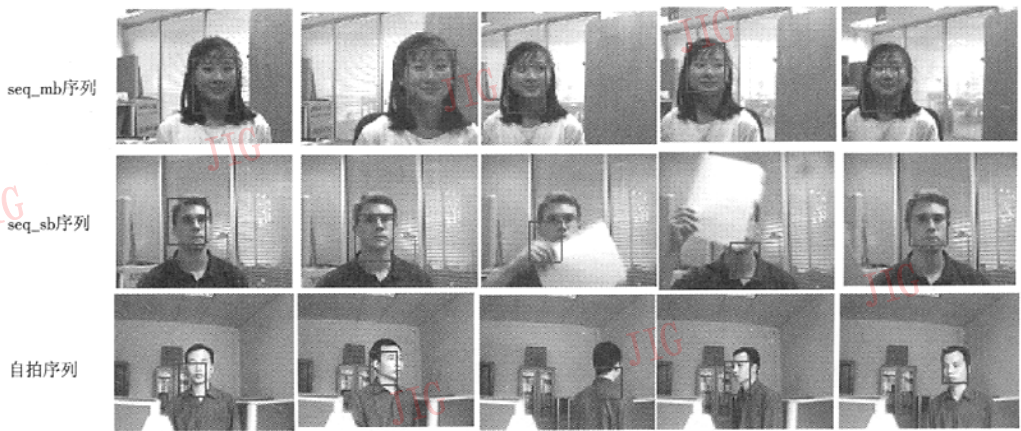
人脸模型总的形状相似性是采用文中式(7)定义的所有特征点匹配指标之和的形式来进行度量,用这种相似性度量来描述形状相似性隐含着一种最优匹配选择,这也将一定程度上可应付遮掩效应。当人脸目标由于严重遮掩丢失时,由于本文的跟踪控制器是基于CONDENSATION框架的,而在CONDENSATION方法中的粒子状态演化过程中,对权重较小的粒子并没有立即去除,而是进行了部分传播并给予了再次观测度量的机会,另外,由于在每次的粒子传播过程中,均按文中式(15)又进行了部分粒子的随机初始化,这样就能在很大程度上保证当人脸目标被遮掩后,被遮掩区域仍会有部分粒子覆盖(也即存在部分粒子的状态和被遮掩的人脸状态接近),这样当遮掩过后,这些粒子的权重就迅速变大,并使得它们在决定人脸状态中起主要作用,因此跟踪器又很快重新恢复对目标的跟踪。由以上分析可以看出,本文的跟踪方法对跟踪过程中的遮掩现象具有较强的鲁棒性。通过类似的分析不难看出,本文跟踪方法对有人脸转动的情况,同样具有较强的鲁棒性。

图6(c)所给出的跟踪结果也证明了以上对本文算法的性能分析,从seq_mb序列可以看出,背景中虽有明显的肤色干扰,但是由于本文在跟踪的线索中有对形状信息的判断,所以跟踪器可以有效地避免背景中的肤色干扰,即跟踪结果没有发生向肤色干扰区域的移动。对于seq_sb序列,跟踪的序列图像中有人脸被短暂遮掩的现象发生,从对seq_sb序列的跟踪结果可以看出,本文跟踪算法能有效应付小面积的轻度遮掩,对于由大面积的遮掩而导致的跟踪失败的情况,等遮掩过后,跟踪器能很快恢复对人脸的准确跟踪。对于自拍序列,图像序列中包含人脸转动,从对自拍序列的跟踪结果可以看出,本文的跟踪算法能有效地处理人脸较大角度的转动,虽然部分帧由于人脸转动角度过大而使跟踪目标丢失,但当人脸转回时,跟踪器又能很快恢复正常工作。

通过大量的试验表明,对于帧间15%内的人脸尺度变化,30°内的人脸旋转,20%左右内的人脸遮掩情况下,本文算法都能进行较好的跟踪。基于粒子滤波框架下的人脸跟踪,其计算复杂度主要集中在观测相似度计算上,对于本文中的跟踪线索,颜色相似度计算时比较简单,而基于直方图的形状相似度计算虽相对较为复杂,但是由于本文在对人脸形



(a) 基于颜色的粒子滤波人脸跟踪结果



(b) 基于颜色和椭圆边界的粒子滤波人脸跟踪结果



(c) 本文算法人脸跟踪结果

图 6 人脸跟踪结果比较

Fig. 6 Comparison of face tracking results

状提取过程中采用了经验有效边缘的方法,从而使得检测到的虚假人脸形状边缘(即背景干扰边缘)

大大减少,这不仅在很大程度上节约了形状相似统计计算时间,并且,在提取形状特征点时,如适当减

少均匀取样精度,还可以进一步减少计算量。另外,由于本文中的形状直方图线索可对人脸形状进行较好的描述,也即粒子滤波中的形状线索观测精度较高,这样就可以用较少的状态粒子来完成对人脸的跟踪,以节省大量时间,再加上当今日益提高的计算能力,本文算法中形状直方图线索所带来的适量计算复杂度的增加,同其所带来的跟踪性能的提高相比,还是完全可以接受的。

6 结 论

跟踪控制器的设计和跟踪线索的选择与表达是人脸跟踪中的两大关键因素,针对视频跟踪场合中的非线性、非高斯的特点,本文选择了粒子滤波作为跟踪控制的基本框架。对于跟踪线索,本文引入了一种新的用直方图方式描述形状的方法,因为该形状描述对平移、尺度变化、旋转和部分遮掩均有较强的鲁棒性。根据实际跟踪环境,本文又对其进行了改进,并作为其中一个跟踪线索。在此基础上,文中提出了一种以直方图描述的颜色和形状为线索的粒子滤波人脸跟踪方法,实验表明,该跟踪方法能有效地处理人脸旋转和背景中的肤色干扰,并且能够在由于遮掩等原因而丢失目标的情况下,及时有效地重新捕获已丢失的目标。

致 谢 感谢斯坦福大学所提供的人脸序列^[11],同时感谢李中科博士的支持与帮助,本文同时凝结着他的部分工作^[10]。

参考文献 (References)

1 Sanjeev M, Smon M. A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking[J]. IEEE Transactions on

- Signal Processing, 2002, 50(2):174~188.
- 2 Isard M, Blake A. Visual tracking by stochastic propagation of conditional density[A]. In: Proceedings of 4th European Conference Computer Vision[C], Cambridge, England, 1996: 343~356.
- 3 Isard M, Blake A. CONDENSATION-conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1):5~28.
- 4 Isard M, Blake A. ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework[A]. In: Proceedings of European Conference Computer Vision[C], Freiburg, Germany, 1998, 1:893~908.
- 5 Shen C, Anton Van den Hengel, Dick A. Probabilistic multiple cue integration for particle filter based tracking[A]. In: Proceedings of Digital Image Computing Techniques and Applications[C], Sydney, Australia, 2003, 1:399~408.
- 6 Nummiaro K, Koller-Meier E, Van Gool L. An adaptive color-based particle filter[J]. Image and Vision Computing, 2003, 21(1):99~110.
- 7 Stern H, Efros B. Adaptive color space switching for face tracking in multi-colored lighting environments[A]. In: Proceedings of the International Conference on Automatic Face and Gesture Recognition[C], Washington DC, USA, 2002: 236~241.
- 8 Yao Hongxun, Gao Wen. Face detection and location based on skin chrominance and lip chrominance transform from color images[J]. Pattern Recognition, 2001, 34(8):1555~1564.
- 9 Belongie S, Malik J. Matching with shape contexts[A]. In: Proceedings of IEEE Workshop on Content-based Access of Image and Video-Libraries[C], Hilton Head Island, South Carolina, USA, 2000: 20~26.
- 10 LI Zhong-ke. Model-based and Image-based Video Representation[D]. Nanjing: Southeast University, 2004. [李中科. 模型基和图像基的视频描述[博士学位论文][D]. 南京: 东南大学, 2004.]
- 11 Head Tracking Sequence [DB/OL]. <http://robotics.stanford.edu/~birch/headtracker/seq/>
- 12 Birchfield S. An elliptical head tracker[A]. In: Proceedings of Conference Record of the 31st Asilomar Conference on Signals, Systems, and Computers[C], Pacific Grove, California, USA, 1997, 2: 1710~1714.