

一种鲁棒的多特征融合目标跟踪新算法

王欢 王江涛 任明武 杨静宇

(南京理工大学计算机科学与技术学院, 南京 210094)

摘要 仅利用单一的目标特征进行跟踪是大多数跟踪算法鲁棒性不高的主要原因。提出了一种新的多特征融合目标跟踪算法,该算法将目标的颜色、纹理、边缘、运动特征统一使用直方图模型进行描述,以降低算法受目标形变和部分遮挡的影响,在 Auxiliary 粒子滤波框架内将所有特征观测进行概率融合,以突出状态后验分布中目标真实状态对应的峰值,从而有效避免了复杂背景的干扰,并给出了一种有效的融合系数计算方法,使融合结果更加准确可靠。实验结果表明,该算法能同时处理刚性与非刚性目标的跟踪,较单一特征的跟踪算法具有明显的优势,对复杂背景下的跟踪具有较高的鲁棒性。与现有多特征融合算法的比较也证明了本文算法的有效性。

关键词 目标跟踪 多线索跟踪 概率融合 Auxiliary 粒子滤波

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2009)03-0489-10

A New Robust Object Tracking Algorithm by Fusing Multi-features

WANG Huan, WANG Jiang-tao, REN Ming-wu, YANG Jing-yu

(School of Computer Science, Nanjing University of Science & Technology, Nanjing 210094)

Abstract Object tracking using single feature often leads to a poor robustness. In this paper, a new object tracking algorithm based on multiple features fusion is presented. to alleviate the affection of object deformation and partial occlusion, it analyzes and describes the color, texture, edge and motion feature of the object using a consistent histogram model, in order to conquer the distractions in the complex background, these features are rationally fused in the framework of Auxiliary Particle filter to obtain a more satisfying approximation of the posterior distribution of the object states. A new method to estimate the fusion coefficient is also proposed to improve the fusion result. Experiment results show that our algorithm can efficiently cope with both rigid and non-rigid objects, outperforms single feature based object tracking algorithms, and has a high robustness in complex background. The comparisons with other multi-cue tracking algorithm also show the validity of the proposed algorithm.

Keywords object tracking, multi-cue tracking, probability fusion, Auxiliary particle filter

1 引言

视觉目标跟踪是计算机视觉中一个重要的研究领域,它不仅在军事上有着举足轻重的作用,而且在人机交互、视频压缩和通信、道路交通管制、监控安防等领域也有着广泛的应用。视觉目标跟踪就是在连续的图像序列中精确地定位目标的位置、速度、姿

态等状态。几十年来,新的跟踪方法层出不穷,它们大致可以分为基于特征的方法,基于区域的方法,基于模型的方法,和基于学习的方法等。近年来, Mean-Shift 跟踪算法^[1]和粒子滤波跟踪算法^[2]受到了广泛重视, Mean-Shift 跟踪算法利用 Mean-Shift 算法对密度函数梯度进行估计,能快速搜索和匹配目标,具有较高的实时性,但只能处理呈单峰分布的状态概率密度函数,且难以跟踪快速运动目标;粒子滤

基金项目:国家自然科学基金重点项目(60632050);国家自然科学基金项目(60472060)

收稿日期:2007-07-27; 改回日期:2007-10-23

第一作者简介:王欢(1982~),男,南京理工大学计算机专业博士研究生。主要研究方向为模式识别、图像处理、计算机视觉。

E-mail: wanghuan_ywzq@tom.com

波算法利用一组带权值的粒子估计和逼近目标状态的后验分布,以此来预测目标的当前状态,它能有效处理非线性,非高斯的跟踪问题,以及处理密度函数呈多峰分布的状态预测问题,但跟踪精度和速度依赖于粒子数目的选取,很多学者围绕两种算法的不足展开了大量的研究,并取得了一定的成果。然而,Mean-Shift 跟踪算法和粒子滤波跟踪算法往往仅采用单一的颜色特征,众所周知,单一的特征无法持续稳健地胜任复杂多变环境下的目标跟踪问题,鲁棒性不高,而利用多个特征进行跟踪是提高跟踪算法鲁棒性的一种有效途径。

多特征跟踪,国外又称多线索跟踪,即利用多个具有互补性的目标特征同时对目标状态进行观测,并在一个合理框架内进行融合,从而可靠地得到目标的精确状态。采用多特征跟踪的出发点非常直观,通常背景是复杂多变的,且难以预测,因此不可避免会包含一些虚假物体,它们与目标具有相同或相似的特征,而一旦这些干扰物体进入目标的搜索区域,使用单一特征的跟踪算法常会陷入“盲然”,无法分清真实目标。此外,目标自身的特征也会不断地变化,比如目标颜色会随着环境光照的变化而改变,这时,单一特征的跟踪算法对目标模式更新的要求会非常高,既要灵敏地觉察出目标模型是否需要及时更新,又要防止可能的干扰或遮挡而产生的过更新。而如果同时使用多个特征,这些问题的影响便会得到有效地处理,当一个特征无法准确地分辨目标时,其他特征可以从另一个角度来衡量一个候选区域是否是真实目标,也许没有一个特征能唯一的确定目标,但利用多个特征联合表决,在绝大部分情况下能有效地区分出目标和背景,且当目标的一个特征发生明显变化时,互补的特征未必会发生显著的变化。

在多特征跟踪的研究中,文献[3]采用颜色和梯度特征跟踪特定的人脸目标,取得了很好的跟踪效果。文献[4]研究了可见光与红外图像特征的融合跟踪问题。文献[5]研究了视觉特征、语音特征和运动特征的融合跟踪问题。文献[6]采用颜色、纹理、边缘方向特征,并在粒子滤波框架内进行概率融合,给出了一种通用性较好的多特征跟踪方法。文献[7]使用角点,边缘,纹理和区域灰度特征实现了对刚性目标的跟踪。从这些研究成果中可知,多特征融合的跟踪算法较单一特征的目标跟踪算法相比,在抗复杂背景干扰、目标外观变化和光照变化等

方面具有明显的优势。然而,多特征跟踪算法仍依赖于单个特征的合理选取,其性能的发挥受单个特征鉴别能力的影响较大,此外,融合方法的有效性会直接影响其跟踪的性能。本文给出了一种多特征融合的目标跟踪方法,通过合理提取目标的颜色、纹理、边缘、运动 4 个有效特征,并将这些特征统一使用简单的直方图进行建模,以提高算法对目标形变、部分遮挡的影响,在 Auxiliary 粒子滤波框架内将所有特征观测进行概率融合,以突出后验分布中目标真实状态对应的峰值,从而有效避免了复杂背景的干扰,并给出了一种有效的融合系数计算方法,使融合结果更加可靠。与文献[7]相比,本文算法能同时处理刚体和非刚体目标的跟踪,与文献[6]相比,本文算法中单个特征的鉴别性较高,且融入了运动特征,进一步减轻了复杂背景的干扰,融合策略更为合理,实现更为简单。

2 Auxiliary 粒子滤波跟踪框架

从贝叶斯理论出发,粒子滤波是用来预测系统状态在每一时刻的后验概率密度函数的。它的有效性主要表现在适用于处理非高斯,非线性的滤波问题。其处理过程一般分为预测和更新两个步骤。设 $p(x_{k-1} | z_{1:k-1})$ 代表 $k-1$ 时刻的后验概率密度函数,其中 x_{k-1} 表示 $k-1$ 时刻的状态, $z_{1:k-1}$ 表示直到 $k-1$ 时刻为止系统所有的观测值。则在预测阶段,通过将 $k-1$ 时刻的后验分布与状态转换概率分布 $p(x_k | x_{k-1})$ 积分,就得到了 k 时刻的先验分布:

$$p(x_k | z_{1:k-1}) = \int p(x_k | x_{k-1}) p(x_{k-1} | z_{0:k-1}) dx_{k-1} \quad (1)$$

在更新阶段,当获得 k 时刻的观测 $p(z_k | x_k)$ 后,结合 Markov 原理,利用贝叶斯公式就得到了 k 时刻的后验分布,

$$p(x_k | z_k) = \frac{p(z_k | x_k) p(x_k | z_{k-1})}{\int p(z_k | x_k) p(x_k | z_{k-1}) dx_k} \quad (2)$$

通过不断地递归迭代就能得到各个时刻目标状态的后验分布。

由于很难从积分中解析地得到方程的解,因此,粒子滤波利用一组带有权值的粒子来逼近后验概率,其过程是:通过从 $k-1$ 时刻后验密度函数 $p(x_{k-1} | z_{k-1})$ 中采样得到一组粒子集合 $\{s_{k-1}^i, w_{k-1}^i\}$, s_{k-1}^i 为粒子状态, w_{k-1}^i 为其对应的权重。利用状态转

移模型 $p(x_k = s_k^i | x_{k-1} = s_{k-1}^i)$ 和 k 时刻获得的量测 z_k , 迭代地获得每个时刻各个粒子的权重:

$$w_k^i \propto w_{k-1}^i \frac{p(z_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{k-1}^i, z_k)} \quad (3)$$

式中, $q(x_k^i | x_{k-1}^i, z_k)$ 是重要性密度函数, 当取 $q(x_k^i | x_{k-1}^i, z_k) = p(x_k^i | x_{k-1}^i)$ 时, 则通常将这种粒子滤波方法称为采样重要性重抽样算法, 此时:

$$w_k^i \propto w_{k-1}^i p(z_k | x_k^i) \quad (4)$$

将权重归一化后, 就得到了逼近的后验分布, 利用该分布就可以得到 k 时刻各种与目标状态相关的函数 $f(\cdot)$ 的估计值:

$$E(f(x_k)) = \sum_{i=1}^N w_k^i f(x_k^i) \quad (5)$$

然而, 取 $q(x_k^i | x_{k-1}^i, z_k) = p(x_k^i | x_{k-1}^i)$ 虽然易于实现, 但并没有考虑最近的观测 z_k , 因此, 当观测似然分布位于先验分布的尾部, 或较先验分布的峰值更加狭窄陡峭时, 性能便会恶化。

Auxiliary 粒子滤波^[8]通过在滤波过程中利用最新的观测来构造重要性密度函数。它引入一个变量 μ_k^i , 该变量由动态模型 $p(x_k^i | x_{k-1}^i)$ 获得, 利用该变量来辅助从后验分布中取样。与采样重要性重采样 (SIR) 算法相比, Auxiliary 粒子滤波充分考虑了当前观测, 克服了 SIR 中当观测似然分布位于先验分布的尾部, 或具有尖锐高峰时滤波性能不佳的问题, 且只需要较少的粒子数就能有效地逼近目标状态后验分布, 本文的跟踪算法就建立在 Auxiliary 粒子滤波的基础上。

下面给出粒子的状态转移模型, 由于难以获得任意目标的运动特性, 同时为了提高跟踪算法对遮挡的处理性能, 本文采用简单的随机运动模型:

$$x_k = x_{k-1} + C_k v_k \quad (6)$$

式中, x 为状态, k 为时刻, v_k 是均值为 0 的高斯白噪声, C_k 对应于粒子的传播半径, 它的大小正比于前 t 个时刻目标的平均状态变化。

$$C_k = \eta \frac{1}{l} \sum_{n=k-t}^{k-1} |x_n - x_{n-1}| \quad (7)$$

η 为一个比例常数, 本文实验中取 $l=5, \eta=2$ 。

3 特征提取与融合

本节分别给出目标颜色、纹理、边缘、运动特征的提取方法, 以及目标模型与候选区域相似度的计算方法, 并由此给出粒子的观测模型和多特征融合

算法。

3.1 颜色特征

颜色是目标最主要的特征之一, 很多文献专门研究了目标的颜色表示模型, 如文献[9]提出的 Spatiogram 模型, 文献[10]提出的空间颜色混合高斯 (SMOG) 模型等。相比之下, 颜色直方图模型由于计算简便, 处理速度快, 对目标尺度、旋转、部分遮挡不敏感等优点而被广泛使用。本文也利用目标的颜色直方图作为目标的第一个跟踪特征, 设一个以 y 为中心的区域, 区域中每个像素的坐标为 $x_i, i=1, \dots, N, N$ 为区域内的像素个数, 则 m 项的颜色直方图定义为

$$\hat{p}_u(y) = A \sum_{i=1}^m K\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta(b(x_i) - u) \quad (8)$$

式中,

$$A = \frac{1}{\sum_{i=1}^m K(\|x_i^*\|^2)} \quad (9)$$

是归一化常数, K 是核函数, 本文取 Gaussian 核 $K(x) = \exp\left(-\frac{1}{2}x^2\right)$, 它用来给靠近目标中心的像素赋予较高的权重, 因此, 能有效避免目标外围像素易受遮挡和干扰等不稳定问题, x_i^* 为目标图像中每个像素的归一化坐标, δ 为 Kronecker delta 函数, $b(x_i)$ 表示 x_i 处像素的颜色值, u 为直方图的颜色索引, 范围是 $1 \sim m, h$ 是核带宽, 决定着权重的分布, 其值越大, 核函数越平坦, 反之则越陡。

3.2 纹理特征

纹理是对于图像各像元灰度的空间分布的一种描述。目标纹理的提取方法有很多, 例如 Gabor 滤波器, Steerable pyramid 等。但需要卷积运算, 时间复杂度较高。本文使用灰度共生矩阵进行纹理提取, 因为其不仅简单, 快速, 而且实验中发现, 与使用 Gabor 滤波器或 Steerable pyramid 得到纹理图像后再建立纹理直方图的方法^[6]相比, 利用纹理矩阵建立纹理直方图对目标刻画更细致。共生矩阵的提取方法见文献[11]。本文分别提取 $0^\circ, 45^\circ, 90^\circ$ 和 135° 4 个共生矩阵, 灰度图像均量化成 32 个等级, 并分别将其展开成 1 维直方图, 再将它们串联起来构成完整的纹理直方图, 其大小为 $4 \times 32 \times 32$ 项。

3.3 边缘特征

对目标的边缘特征, 本文使用边缘方向直方图

来描述。对目标区域图像 I 的每个像素 $I(x, y)$ 计算其梯度幅值 $g(x, y)$ 和梯度方向 $\theta(x, y)$:

$$g(x, y) = \sqrt{\left(\frac{\partial I(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I(x, y)}{\partial y}\right)^2}$$

$$\theta(x, y) = \tan^{-1}\left(\frac{\partial I(x, y)}{\partial y} / \frac{\partial I(x, y)}{\partial x}\right) \quad (10)$$

将边缘方向角 $0 \sim 360$ 度量化成 36 级的角度直方图, 统计并累计对应角度的像素梯度幅值, 再进行归一化就得到了区域的边缘方向直方图。

3.4 运动特征

运动特征在剔除背景干扰方面非常有效, 尤其是在摄像机静止的情况下。通常, 运动特征的提取是利用背景差或帧间差得到差分图像, 再对其进行运动分割, 得到运动掩模图像来指导跟踪, 然而, 带来的问题是分割的阈值较难选取, 阈值取大了容易得到残缺的目标或根本得不到目标, 阈值取小了, 虽然能得到比较完整的目标但却又带入了较多的背景干扰, 而如果直接使用差分图像就可以避免分割的问题。此外, 考虑到最好能将所有特征统一使用直方图描述, 因此, 本文采用文献[5]提出的差分直方图, 它是基于如下的实际考虑: 在相邻两帧的差分图像中, 分析背景区域所对应的差分图像直方图, 可发现几乎所有的像素灰度值均集中在 0 附近, 而其他灰度级上基本没有像素, 分布极不均匀, 而运动目标区域的差分直方图在各个灰度级上都含有部分像素, 只是分布不太均匀而已, 但随着运动剧烈程度的加大, 分布越均匀。因此, 设目标模型直方图为均匀分布, 使用候选区域的差分直方图与均匀分布进行比较, 得到的相似度一定程度上便反映了目标的运动剧烈程度, 并以此作为衡量目标运动特征的判据。由于本文使用的是帧间差, 运动目标边缘轮廓的运动最为突出, 因此, 应对候选区域长宽均放大 5 个像素以较好地目标边缘轮廓包含进来。

3.5 直方图相似性度量

在第 1 帧中, 跟踪算法对目标图像分别按上述方法提取颜色、纹理、边缘方向和运动差分直方图, 设 $\mathbf{q}, \mathbf{p}(y)$ 分别为目标模型和以 \mathbf{y} 为中心候选区域的某一特征直方图, 本文通过 Bhattacharyya 系数来衡量它们的相似度:

$$\rho_i(\mathbf{p}_y, \mathbf{q}) = \sum_{u=1}^m \sqrt{\mathbf{p}_u(y) \mathbf{q}_u} \quad (11)$$

因此, 每个粒子都获得了 4 个特征的相似度观测 ρ_i , $i \in \{\text{color, texture, edge, motion}\}$, 下面将给出本文

Auxiliary 粒子滤波的观测模型和特征融合策略。

3.6 特征融合策略

在 Auxiliary 粒子滤波算法中, 对于每个特征而言, 各个粒子相对于该特征的观测概率都离散的刻画了目标状态的似然密度函数, 每个特征都可以按照自身的似然密度, 并结合先验概率密度, 对目标状态的后验密度函数给出一个估计。然而, 对于不同的跟踪场景而言, 每个特征的鉴别能力是不同的, 也即其观测似然函数中对应于目标真实状态处的峰值的尖锐程度不同, 因此, 对后验概率分布预测的准确性也不同。为了利用有限的粒子较好的刻画目标的观测似然以获得目标的后验概率准确估计, 一种有效的方法是, 设各个特征的相似度观测是相互独立的, 利用它们的联合相似度来构建观测似然, 即第 i 个粒子的联合相似度为

$$\rho^i = \prod_l \tilde{\rho}_l^i \quad (12)$$

其中

$$\tilde{\rho}_l^i = \frac{\rho_l^i - \rho_l^{\min}}{\rho_l^{\max} - \rho_l^{\min}} \quad (13)$$

l 是特征索引, ρ_l^{\min} 和 ρ_l^{\max} 为所有粒子在第 l 个特征上的相似度的最小值和最大值。则第 i 个粒子的观测概率为

$$p(z|x^i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1-\rho^i}{2\sigma^2}\right) \quad (14)$$

式中, σ 为均方差。

为了说明这一策略的有效性, 图 1 给出了两组相似度平面图。分别为原图、颜色、纹理、边缘、运动, 联合特征的相似度平面图。例如, 对于颜色特征而言, 以该区域的每一个点为中心, 在与目标模板同样尺寸的矩形框内提取颜色直方图, 并计算其与颜色模型直方图的相似度。将相似度归一化后以 2 维图像进行显示, 其他特征类似。在相似度平面图中, 越亮的点表示相似度越大, 越可能是目标点, 反之, 越可能是背景点。每组的最后一幅图为经过式(12)和式(14)计算后得到的 $p(z|x^i)$ 。从图中可以看出, 单一直方图特征的概率平面或呈多峰分布, 或者峰值较为平坦, 目标处对应的峰值不够突出, 不能有效地鉴别目标。而在联合概率观测上, 真实目标被有效地突显出来。

联合概率观测克服了单一特征观测不可靠的缺点, 但如果一个特征观测非常不准确, 势必会影响到联合观测的准确性, 为了避免这一情况, 可以给

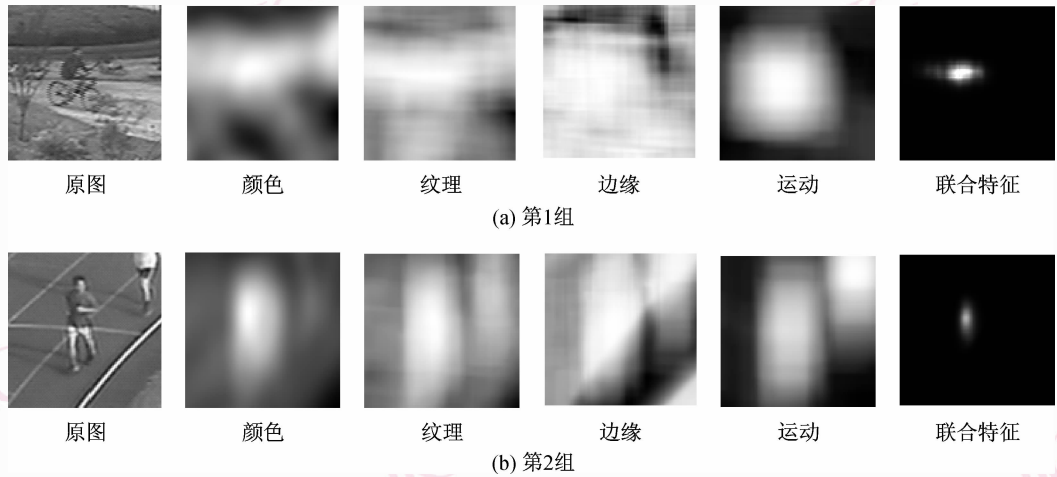


图1 各特征独立相似度和联合相似度平面示意图
Fig. 1 The similarity plane graph of each single feature and fusing all features

式(12)中每个独立的相似度 $\tilde{\rho}_i^i$ 赋予一个权重,如果某个特征观测的准确性较高,则给予一个较高的权重,反之,则给予较小的权重以抑制它在联合观测中所起的作用,即

$$\rho^i = \prod_l (\tilde{\rho}_l^i)^{\varepsilon_l} \quad (15)$$

式中, ε_l 表示第 l 个特征的权重。为了有效地识别每个特征的鉴别能力,以获得准确的 ε_l ,文献[6]给出了一种依据每个特征对应的所有粒子观测的最大值来确定加权系数的方法。这一依据并不可靠,比如,当背景中混杂边缘比较丰富时,边缘方向直方图的鉴别性不高,但每个粒子相对于边缘特征的观测都很大,即该特征的观测最大值总是高于其他特征。这时,边缘观测将一直被赋予较大的权重,反而影响融合的性能。

本文给出了一种新的融合权系数计算方法,该方法以每个特征较优粒子的空间分布为依据获得每个特征的鉴别能力。由于粒子滤波跟踪方法是通过所有粒子加权来获得各时刻目标状态的,因此,权重较大的粒子(本文称为较优粒子)对目标状态的计算作用较大。对每个特征 l ,将所有粒子在该特征上的相似度系数进行排序,并选取前 10% 个相似度最大的粒子 $\{x_l^i\}_{i=1}^m$, $m = 10\% \cdot N$ 作为较优粒子。分别计算每个特征 l 中较优粒子的空间位置的均值 $\bar{\mu}_l$ 和方差 σ_l^2 ,其中,

$$\bar{\mu}_l = \frac{1}{m} \sum_{i=1}^m x_l^i \quad (16)$$

$$\sigma_l^2 = \frac{1}{m-1} \sum_{i=1}^m (x_l^i - \bar{\mu}_l)(x_l^i - \bar{\mu}_l)^T \quad (17)$$

以及所有特征较优粒子的空间位置总体均值

$$\bar{\mu} = \frac{1}{4} \sum_l \bar{\mu}_l \quad (18)$$

由此,权系数的计算基于如下依据,其一,某个特征的较优粒子在空间上分布越集中,即较优粒子的空间方差越小,表明其后验分布越接近单峰分布,鉴别性越高;其二,多个特征的联合最优粒子的总体均值位置与目标真实位置一般较接近,因此,每个特征较优粒子的空间均值应当离总体均值位置越近越好。因此,第 l 个特征的鉴别能力可以通过下式计算:

$$\varepsilon_l = \frac{1}{|\bar{\mu}_l - \bar{\mu}| \cdot \sigma_l^2} \quad (19)$$

将其归一化 $\varepsilon_l = \frac{\varepsilon_l}{\sum_l \varepsilon_l}$,就得到了各个特征的融合权系数。

与文献[6]相比,本文的融合策略同样也将每个特征视为一个独立的观测,但通过对所有粒子在每个特征上相似度的简单拉伸代替了观测概率的计算,从而有效避免了文献[6]算法中需要对每个特征观测概率的方差进行估计。虽然最终观测中也需要给定 σ 的值,但实验发现,对 σ 指定一个固定值已经能充分满足实际跟踪的需要。本文取 $\sigma = 0.2$ 。此外,本文的融合系数的计算充分利用了较优粒子的空间分布,由于每个特征均是对相同空间分布的粒子进行衡量的,因此使鉴别性的度量更具有可比性,而通过方差和均值差进行鉴别性的度量比相似度大小的度量更具合理性。

到此, 本文的跟踪算法可以总结如下:

(1) 初始化, $k=0$, 对目标模板图像提取颜色、纹理、边缘方向模型直方图, (由于运动模型直方图设为均匀分布, 因此无需提取); 设先验分布为高斯分布, 以目标在第 1 帧中的状态 $(x, y, s)^T$ 为均值向量, 该状态各成分分别表示目标中心在图像中的 X 坐标, Y 坐标, 和目标的尺寸 (相对于模板大小的比例系数), 以

$$\begin{bmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}$$

为协方差矩阵, 从先验

分布中抽取 N 个初始粒子 $\{x_0^i\}_{i=1}^N$, 权值均设为 $\{w_0^i\}_{i=1}^N = 1/N$;

(2) $k \leftarrow k+1$ For $i=1, \dots, N$

计算辅助变量 $\mu_k^i \sim p(x_k | x_{k-1}^i)$;

计算 $w_k^i = w_{k-1}^i \cdot p(z_k | \mu_k^i)$;

其中观测概率 $p(z_k | \mu_k^i)$ 由式 (14) 和式 (15) 给出, 下同。

(3) 计算新的粒子索引集合: 从 $w_k^i, i=1, \dots, N$ 中重新采样 N 次得到新的粒子标号 $\{i'\}_{j=1}^N$;

(4) 获得 k 时刻粒子和权重:

$\tilde{x}_k^j \sim p(x_k | x_{k-1}^{i'})$, $j=1, 2, \dots, N$, 计算粒子的权重 $\tilde{w}_k^j = \frac{p(z_k | \tilde{x}_k^j)}{p(z_k | \mu_k^{i'})}$; 归一化各粒子的权重, 使 $\sum_{j=1}^N \tilde{w}_k^j = 1$;

(5) 评估粒子匮乏程度: 若 $1 / \sum_{j=1}^N (\tilde{w}_k^j)^2 < N/2$, 则从 $\{\tilde{x}_k^j\}_{j=1}^N$ 中依粒子权重重新采样 N 次得到 $\{\tilde{x}_k^{i''}\}_{j=1}^N$, 并重新赋粒子权值 $\tilde{w}_k^{i''} = 1/N$, 否则不做处理;

(6) 输出结果: $E(f(x_k)) = \sum_{j=1}^N \tilde{w}_k^j f(\tilde{x}_k^j)$;

(7) 转步骤 2。

4 实验结果与分析

本节通过多个序列图像的测试来验证本文算法的有效性。所有算法均在 AMD2.0 微处理器, 512 M 内存的 PC 机上使用 VC++6.0 编程实现, 实验中, 粒子数目设为 100。

图 2 首先给出了文献 [6] 算法和本文算法在行人经过光照区序列上的跟踪结果, 图 2(a) 为文献 [6] 算法的跟踪结果, 图 2(b) 为本文算法的跟踪结果。其中, 蓝色矩形框表示算法输出的目标位置, 绿色点表示各个粒子的位置 (下同)。在该序列中, 由

于光照的突变, 当行人经过光照区时, 颜色、纹理特征均丧失了跟踪性能, 虽然边缘特征受光照影响较小, 但由于光照区中存在物体阴影的强边缘干扰, 使边缘特征也失去效用, 因此, 文献 [6] 算法中所有的特征均不可靠, 导致算法跟踪失败。而本文算法虽然也受到了同样的影响, 但还引入了运动特征, 运动特征能有效地将运动的行人与静止的背景区分开, 因此, 使得本文的跟踪算法得以在整个序列中均能可靠地跟踪目标。

在图 3 中, 跟踪的目标——骑车人也存在于复杂的背景环境中, 首先, 受摄像机远距离架设的影响, 颜色信息受到较大衰减, 与背景颜色混为一体, 而且背景中存在电线杆, 树木等强干扰物体, 此外, 目标还受树木遮挡的影响, 使用文献 [6] 算法几乎不能正常跟踪, 而本文的算法依然能有效地跟踪整个序列, 这里仅给出本文的跟踪结果, 图 (3) 上方大图为序列的第 1 帧图像, 在该图像上叠加了目标在整个序列上的跟踪轨迹, 其中 6 个顺序的蓝色框分别对应于其右方的 6 幅跟踪结果图。

图 4 给出了本文算法在动态背景下人脸序列中的跟踪结果。该序列中包含了遮挡、相似颜色背景干扰, 尺度变化, 目标平面内外旋转, 以及同时存在多种变化的影响等。使用本文算法进行测试, 发现颜色与边缘特征起着主要作用, 大多数情况下, 颜色特征能很好地跟踪面部和头发颜色, 当出现背景干扰时, 纹理特征起到了较好的补充; 当头部发生平面外旋转时, 颜色、纹理特征发生骤变, 而边缘特征能有效捕捉未发生明显变化的头部轮廓; 遮挡发生时, 多个特征由于遮挡的位置和程度的不同均受到了不同程度影响, 但经过融合后, 依然有效地保持稳定的跟踪效果。图 5 给出了本文算法在多人场景中的跟踪结果, 其中目标经历了多次部分遮挡。可以看出, 虽然多个特征均使用了忽略空间位置信息的直方图模型, 但由于多特征融合的作用, 在目标发生部分遮挡时并没有发生较大的空间定位误差。

最后, 在 Runner 序列和 car 序列上对各特征单独跟踪和所有特征融合跟踪的跟踪误差进行了定量的比较, 每帧中目标的精确位置由手工获得, 在 Auxiliary 粒子滤波算法的观测模型 (式 14) 中, ρ 分别取 ρ_{color} 、 ρ_{texture} 、 ρ_{edge} 、 ρ_{motion} 和 ρ_{fusion} 进行跟踪实验, 并比较各自与目标精确位置的均方误差。从图像跟踪结果 (如图 6、图 7 所示) 和跟踪误差曲线 (如图 8 所示) 可以看出, 融合后的跟踪精度具有明显提高, 就

图 2 本文算法与文献[6]的跟踪算法结果对比

Fig. 2 The comparison of tracking results with the algorithm in reference [6] and our algorithm

图 3 本文算法在复杂背景干扰和遮挡情况下的部分跟踪结果

Fig. 3 Tracking results of face sequence with our algorithm

图 4 本文融合跟踪算法在人脸序列上的部分跟踪结果示意图

Fig. 4 Tracking results of face sequence with our algorithm

图 5 本文融合跟踪算法在多人序列中的部分跟踪结果示意图

Fig. 5 Tracking results of multiple pedestrian sequence with our algorithm

Runner 序列而言,由于跟踪目标的局部背景中存在另一个运动目标,使目标的运动特征、边缘特征受到干扰,但由于颜色和纹理特征能有效地鉴别目标,融合跟踪的性能并未随着运动和边缘特征的不可靠而受到较大的影响。与最可靠的颜色特征的跟踪误差非常接近。而在 Car 序列中,4 个特征单独跟踪时

均能有效地跟踪目标,然而这时融合跟踪的误差最小。充分说明了融合算法的有效性。由于粒子的产生具有一定的随机性,因此,这里的跟踪误差取的是 20 次跟踪实验的平均值。图 6、图 7 中目标框的颜色所对应的特征与图 8 所表示的特征颜色相互对应。

图 6 Runner 序列各特征单独跟踪与融合跟踪部分结果示意图

Fig. 6 Tracking results of Runner sequence with our algorithm using each single feature and feature fusion strategy

图 7 Car 序列各特征单独跟踪结果与融合跟踪部分结果示意图

Fig. 7 Tracking results of car sequence with our algorithm using each single feature and feature fusion strategy

为了进一步说明本文融合策略的有效性,且更有可比性,在本文算法的框架下,采用相同特征进行跟踪的基础上分别使用式(12)对应的不加权融合策略,式(15)对应的加权融合策略,以及文献[6]的融合策略分别在以上两个序列上进行对比试验,跟踪误差如图 9 所示,可见文献[6]的融合策略最差,不加权融合策略其次,加权融合策略精度最高。

最后,在算法实时性方面,本文算法的时间开销主要集中在各个特征的提取和相似度的计算上,而它们又直接与目标的大小和粒子数选取的多少有关。目标越大,粒子数取得越多,时间花费越大,因此,对于较大的目标可以通过对目标进行抽样,降低分辨率来减少计算时间。在未对程序做任何优化的情况下,经测试,取 100 个粒子,对 96×100 的目标

图8 各特征单独跟踪与融合跟踪结果误差曲线图

Fig. 8 The tracking error curve with our algorithm using each single feature and feature fusion

图9 不同融合策略跟踪结果误差曲线对比图

Fig. 9 The Comparison of tracking error with different fusion strategy

进行跟踪,本文算法能每秒钟处理4帧。如果对算法进行优化,使用数字信号处理(DSP)或现场可编程门阵列(FPGA)等硬件实现,完全能够达到实时处理的要求。

5 结论

本文在 Auxiliary 粒子滤波框架内,采用颜色、纹理、边缘、运动4个具有一定互补性的特征进行联合概率跟踪。4个特征均采用简单的直方图模型,充分利用了直方图对形变、部分遮挡、尺度不敏感的优点。实验发现,多特征进行融合避免了单一特征

的不稳定。对背景干扰(包括相似颜色、干扰边缘)、部分遮挡、光照变化、均具有较高的鲁棒性。同时,合理的融合策略保证了有利特征跟踪性能的充分发挥。下一步研究将集中在各个特征的在线更新上,以及将算法推广到多目标跟踪等一般情况。

参考文献 (References)

- 1 Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5):564-575.
- 2 Nummiaro K, Koller-Meier E, Van-Gool L. An adaptive color-based particle filter[J]. Image and Vision Computing, 2003, 21(1): 99-110.

- 3 Birchfield S, Elliptical head tracking using intensity gradients and color histograms [A]. In: Proceedings of the International Conference on Computer Vision and Pattern Recognition [C], Santa Barbara, CA, USA, 1998; 232-237.
- 4 Conaire C, Connor N. Thermo-visual feature fusion for object tracking using multiple spatio-gram trackers [A]. In: Proceedings of Conference on Machine Vision and Applications [C], New York, NY, USA, 2007; 483-494.
- 5 Perez P, Vermaak J, Blake A. Data fusion for visual tracking with particles [J]. Proceedings of the IEEE, 2004, 92 (3): 495-513.
- 6 Brasnett P, Mihayhova L, Bull D. Sequential monte carlo tracking by fusing multiple cues in video sequences [J]. Image Vision Computing, 2007, 25 (8): 1217-1227.
- 7 Serby D, Koller-Meier E, Van-Gool L. Probabilistic object tracking using multiple features [A]. In: Proceedings of 17th International Conference on Pattern Recognition [C], Cambridge, UK, 2004; 184-187.
- 8 Pitt M, Shephard N. Filtering via simulation: auxiliary particle filters [J]. Journal of the American Statistical Association, 1999, 94 (446): 590-599.
- 9 Birchfield S, Sriram R. Spatiograms versus histograms for region-based tracking [A]. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Patten Recognition [C], San Diego, CA, USA, 2005; 1158-1163.
- 10 Hanzhi W, Suter D, Konrad S. Effective appearance model and similarity measure for particle filter and visual tracking [A]. In: Proceedings of the International Conference on ECCV [C], Graz, Austria, 2006; 606-618.
- 11 Zhang Hong-lin. Digital Image Pattern Recognition Technique and Project Practice [M]. Beijing: Posts & Telecommunications Press, 2003, 452-455. [张宏林. 数字图像模式识别技术及工程实践 [M]. 北京: 人民邮电出版社, 2003; 452-455.]