

基于新颖相似度的视觉跟踪算法

夏 瑜, 吴小俊

(江南大学 信息工程学院, 江苏 无锡 214122)

[摘要] 模板匹配是视觉跟踪领域中的重要环节. 以粒子滤波为跟踪框架, 提出了一种新颖的模板匹配的统计特征相似性指标, 引进了加权因子有效抑制目标边界噪声干扰和匹配区域背景成分的影响, 同时突出了目标特征的权重. 由于完全不同模板图像统计特征也会存在交集, 导致相似性偏差. 所以融合基于 HSV 颜色空间的相似度能够修正偏差, 改善匹配函数峰值特性, 使得搜索目标得到全局最优解, 最终实现鲁棒跟踪. 实验结果表明, 模板匹配具有良好的峰值特性, 算法在跟踪目标存在变形、噪声、遮挡时也可以达到比较理想的跟踪效果.

[关键词] 粒子滤波, 直方图, 相似性度量, HSV, Bhattacharyya 系数

[中图分类号] TP 391. 41 **[文献标识码]** A **[文章编号]** 1672-1292(2008) 04-0068-05

Visual Tracking Algorithm Based on a Novel Similarity Function

Xia Yu, Wu Xiaojun

(School of Information Engineering, Jiangnan University, Wuxi 214122, China)

Abstract Template matching is an important link of visual tracking. A novel template statistical feature matching similarity criterion is proposed with a tracking framework using particle filter. Weighted factor is introduced to effectively reduce the influence of border noise and background feature, and a great emphasis is given to the importance of target feature. Similarity bias may be got in the two completely different templates because of the intersection of statistical features. The proposed method can correct the bias and improve the peak modality of matching function by fusing similarity based on HSV color system, so it can obtain the global optimal solution and robust tracking. Experimental results show that template matching has an excellent peak distribution, and the proposed tracking algorithm exhibits good precision and robustness in the presence of noise, deformation and occlusion.

Key words particle filter, histogram, similarity measure, HSV, Bhattacharyya coefficient

视觉跟踪在机器人、安全监控、人机交互等领域有着广泛的应用^[1], 近年来吸引了众多研究者的注意, 是当前计算机视觉领域中的研究热点之一. 目前存在的视觉跟踪方法主要分为两类: 数据驱动方法和模型驱动方法^[2, 3]. 其中, 数据驱动方法是利用目标的运动信息对目标进行跟踪的方法, 如光流法和特征点法^[4, 5], 这种方法的缺点是计算量太大. 模型驱动方法主要是通过模板匹配的方式来实现, 根据其匹配特征对象不同, 分为基于目标边界特征方法^[6]和基于目标区域特征方法^[7, 8]. 其中, 基于目标边界方法主要利用图像中的边缘信息进行图像匹配; 而基于目标区域特征方法通过对目标整个区域颜色、纹理等特征信息的分析来实现图像匹配.

基于模板匹配的方法得到了广泛的应用, 然而基于目标边界特征方法有用信息获取方法复杂, 且容易受到背景的干扰, 匹配效果不太理想. 由于颜色、纹理具有旋转不变性、缩放不变性等性质且所得到的信息比边缘更加充分, 所以基于目标区域特征方法匹配效果要优于基于目标边界的匹配方法. 但是, 因为我们所获取的匹配图像往往是背景和目标特征常常相互交织在一起的, 直接使用模板生成直方图, 将会受到背景成分的影响, 而且边界信息容易受噪声的干扰, 这样会引起目标本身特征的削弱, 使得跟踪精度下降, 甚

收稿日期: 2008-06-18

基金项目: 2006 年教育部新世纪优秀人才计划项目 (NCET-06-0487)、国家自然科学基金 (6047206Q 60572034) 和江苏省自然科学基金 (BK2006081) 资助项目.

通讯联系人: 夏 瑜, 硕士研究生, 讲师, 研究方向: 计算机视觉和目标跟踪. E-mail: xiyu@changshu.net

至会错误匹配而导致跟踪失败.

针对上述问题, 本文跟踪算法则是以粒子滤波^[9]理论为基础, 利用颜色信息建立新颖的直方图, 引入加权因子, 有效抑制区域边界受噪声的干扰和减小背景的影响, 并融合符合人眼感知特性、接近于人类视觉系统的 HSV 颜色空间^[10]所计算的相似度来提高匹配精度. 其实质是通过色彩信息所建立的新颖直方图来提高匹配精度, 改善模板匹配的峰值特性, 从而提高粒子采样率. 利用本文算法来实现运动目标跟踪, 具有很高的跟踪准确性和鲁棒性, 且对复杂环境也有很强的目标识别能力, 适用的场合更普遍.

1 粒子滤波

粒子滤波又称蒙特卡罗仿真, 是一种以贝叶斯推理为核心的滤波技术. 它用一组加权的随机采样来近似表示所求的后验概率密度函数. 当随机采样的粒子数足够多时, 粒子滤波近似于最优贝叶斯估计.

通过建立状态模型和观测模型, 把运动目标的跟踪抽象为状态向量的估计. 状态向量用来描述目标的位置、速度、加速度和轮廓等信息, 在视频中对目标的跟踪就相当于在状态空间中对相应的状态向量的跟踪.

状态模型可以表示为:

$$X_k = f_k(X_{k-1}, V_{k-1}), \quad (1)$$

其中: f_k 为 $R^n \times R^n \rightarrow R^n$, 是系统传递方程; X_k 表示所跟踪的目标在时刻 k 的状态向量; $V_{k-1} \in R^n$ 是系统噪声. 利用系统传递方程可以递归地求出各个离散时间上的状态向量.

观测模型可以表示为:

$$Z_k = h_k(X_k, W_k), \quad (2)$$

其中: h_k 为 $R^n \times R^n \rightarrow R^n$, 是观测方程; Z_k 是观测向量; W_k 是观测噪声. 使用最近的观测值对系统传递状态向量进行修正.

在粒子滤波中, 概率分布被近似为粒子及其权重定义的离散样本值 $\{X_k^{(m)}, W_k^{(m)}\}_{m=1}^M$ 其中 $X_k^{(m)}$ 为 k 时刻第 m 个粒子, $w_k^{(m)}$ 为对应的 k 时刻第 m 个粒子的权重, M 为粒子数. 于是 k 时刻的后验概率密度可以加权近似为:

$$p(X_k | Z_{1:k}) = \sum_{m=1}^M w_k^{(m)} \delta(X_k - X_k^{(m)}), \quad (3)$$

其中, $\delta(\cdot)$ 是 Dirac delta 函数. 而 Particle Filter 算法的具体实现步骤如下:

① 初始化: 在 $k=0$ 时刻, 根据概率分布函数 $p(X_0)$ 抽样出 N 个粒子, 假定抽样出的每个粒子用 $\{X_{0:k}^i, w_k^i\}_{i=1}^N$ 表示, 其中 $\{X_{0:k}^i, i=1, \dots, N\}$ 为粒子, $\{w_k^i, i=1, \dots, N\}$ 为权重.

② 根据上面 (1) 式进行目标状态更新.

③ 重要性权重计算. 由 $X_k^i \sim q(X_k | X_{0:k-1}^i, Z_{0:k})$ 采样, 重要性权重计算为下式:

$$w_k^i = w_{k-1}^i \frac{p(Z_k | X_k^i) p(X_k^i | X_{k-1}^i)}{q(X_k^i | X_{0:k-1}^i, Z_{0:k})} \quad (4)$$

④ 归一化重要性权重:

$$\bar{w}_k^i = w_k^i / \sum_{i=1}^N w_k^i \quad (5)$$

⑤ 重采样: 从 X_k^i 集合根据重要性权重 \bar{w}_k^i 重新采样得到新的 N 个粒子的集合 \tilde{X}_k^i , 并重新分配粒子权重: $w_k^i = \bar{w}_k^i = 1/N$.

⑥ 目标状态估计:

$$\tilde{X} = \sum_{i=1}^N w_k^i \tilde{X}_k^i \quad (6)$$

设定 $k \rightarrow k+1$, 返回步骤 ②.

2 基于新颖统计直方图相似度计算

在基于粒子滤波器的目标跟踪方法中, 目标模板匹配是影响目标跟踪性能的重要因素. 颜色特征由于

具有旋转不变性、缩放不变性等性质而成为目标跟踪算法中的一种主要特征, 本文算法采用其建立的新颖直方图信息充分、抗干扰性更强. 在得到可靠直方图后, 采用 Bhattacharyya 系数计算目标模板直方图与待测粒子直方图之间的相似度, 所得到的匹配曲面更加尖锐, 匹配准确性高. 而新颖直方图的构建步骤如下:

- 步骤 1 在初始化帧图像中选取目标模板, 采用颜色空间特征作为直方图特征
- 步骤 2 分别计算目标模板和候选目标的每一个像素所属直方图的柱形条, 对于每一个像素的 RGB 值为 (r, g, b) , 而 r, g, b 的范围分别为 $[0 \sim 255]$, 如果直方图的柱形条数为 m , 这样每一个像素所属的 2D 直方图的柱形条可以通过下式来计算:

$$b_{in_index} = \left\lceil \alpha \frac{r}{256} + \beta \frac{g}{256} + \gamma \frac{b}{256} \right\rceil \times m, \tag{7}$$

其中, α, β, γ 为色彩值调节系数, 可以根据视频图像进行调节来提高跟踪直方图的差异性, 从而能更加清晰地识别目标.

- 步骤 3 计算生成直方图中每个像素的权重, 并添加到相关的直方图柱形条. 其方法为: 假如目标模板或候选目标的边界到中心的距离为 l 这个距离不一定所有像素全部统一, 可以采用分块方法得到局部距离来提高精度; 假设每一个像素到中心的距离为 d , 则权重可以如下进行计算:

如果 $d \leq l$ 则 $av_i = \frac{l^2 - d^2}{l^2} = 1 - \left(\frac{d}{l}\right)^2$.

如果 $d > l$ 则 $av_i = 0$ 表示该特征目标模板不具有或影响为零.

计算各个像素权重的和 av_{sum} , 然后进行权重归一化: $av_i = av_i / av_{sum}$.

- 步骤 4 构造新颖直方图, 根据 b_{in_index} 来添加相应的归一化权重到指定柱形条.
- 目标模板及候选目标区域直方图表示的是图像在颜色空间的概率密度分布, 度量候选目标区域与目标模板的颜色概率密度分布的相似性采用 Bhattacharyya 系数. 假如目标模板色彩直方图特征向量为 $P = \{P^{(u)}\}_{u=1 \dots m}$, 候选目标区域色彩直方图特征向量为 $Q = \{Q^{(u)}\}_{u=1 \dots m}$, 则 Bhattacharyya 系数可以按下式计算:

$$\rho(P, Q) = \sum_{u=1}^m \sqrt{P(u)Q(u)}. \tag{8}$$

3 基于 HSV 颜色空间相似度计算

Smith 提出的适用的基本主观颜色模型 HSV 颜色空间, 其中 Hue 表示颜色的主色调, 用 0 到 360° 的角度表示, Saturation 表示饱和度, Value 表示颜色的亮度. HSV 颜色空间由两个重要特点: 首先亮度分量与色调分量是分开的, V 分量与图像的彩色信息无关. 其次, 色调 H 及饱和度 S 分量的概念相互独立并与人感受色彩的方式紧密相连. 这些特性使得 HSV 颜色模型非常适合基于人的视觉系统对彩色感知特性进行处理分析的图像算法.

在 HSV 颜色空间中计算两个像素 p_i 和 p_j 的颜色欧氏距离 (Euclidean distance), 如下式:

$$D_{ij} = \sqrt{(h_i - h_j)^2 + (s_i - s_j)^2 + (v_i - v_j)^2}. \tag{9}$$

于是, 两个像素颜色相似度为:

$$s'(p_i, p_j) = \begin{cases} e^{-kd}, & d < T, \\ 0 & d \geq T. \end{cases} \tag{10}$$

其中, d 为像素 p_i 和 p_j 的颜色欧氏距离 (Euclidean distance), 而 k 为常数, T 为阈值. 显然, 容易得出两个相同颜色的像素的颜色相似度 $s'(p_i, p_j) = 1$ 像素的颜色相似度随着欧氏距离的增大而单调递减. 于是目标模板及候选目标区域的相似度可以如下计算:

$$s(A, B) = \frac{1}{N} \sum_{(p_i, p_j) \in A, B} s'(p_i, p_j), \tag{11}$$

其中 A, B 分别代表目标模板及候选目标区域, N 为像素总数. 这样计算出来的匹配图像相似度更加合理, 克服了与人主观感受不一致的缺陷.

4 相似度融合

上面第 2、3 两节所取得的相似度, 融合后可以避免 Bhattacharyya 系数的统计偏差, 且更能反映采样粒子与目标之间的逼近度. 融合相似度后目标模板和候选目标区域的 Bhattacharyya 距离为:

$$d_B = \sqrt{1 - \mu \times \rho(P, Q) + \nu \times s(A, B)}, \quad (12)$$

其中 μ 、 ν 为调节系数, 可以根据跟踪序列图像的实际情况进行调节, $\rho(P, Q)$ 为新颖直方图统计特征相似度, $s(A, B)$ 为基于 HSV 颜色空间相似度.

5 仿真实验分析

为了验证本文算法的有效性, 选用相关视频序列计算相似度并实现目标视觉跟踪, 在 Pentium Dual E2140 CPU, 1 G 内存的 PC 机上, 用 Matlab 7.0 编程实现并进行相应的分析.

第一组实验是搜索区域中候选目标与目标模板相似度计算. 如图 1, 两组视频序列的左边图像为视频序列的局部图像, 大框表示搜索区域, 小框表示所跟踪的目标; 右边图像为根据本文算法所计算的相似度分布图. 相似度分布图峰值特性反映了算法对目标和背景的分辨和识别能力, 分布曲面越尖锐说明目标识别精度越高. 视频序列 1 (图 1(a)) 中运动中的乒乓球与背景有一定色差, 但因为运动速度快, 目标边缘与背景存在融合现象, 然后利用本文算法所得的相似度分布曲面相当尖锐, 目标识别效果佳. 视频序列 2 (图 1(b)) 中行人的头部除头发外皮肤存在色彩融合现象, 且与颈部和手臂的色彩纹理相同, 而采用本文算法, 有效地突出了目标本身特征, 生成曲面峰值特性明显, 能够精确锁定目标.

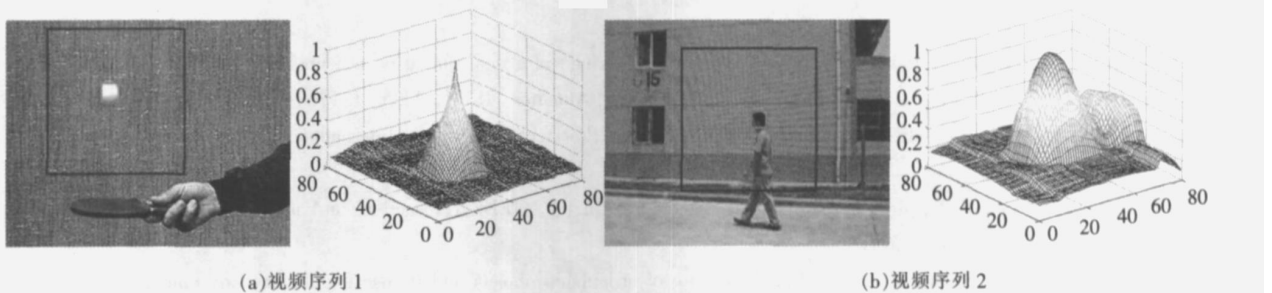


图 1 目标模型的相似度分布

Fig.1 Similarity distribution of target model

第二组实验是采用本文算法对视频序列中的运动人体进行跟踪, 在整个运动过程中, 运动中的人体的

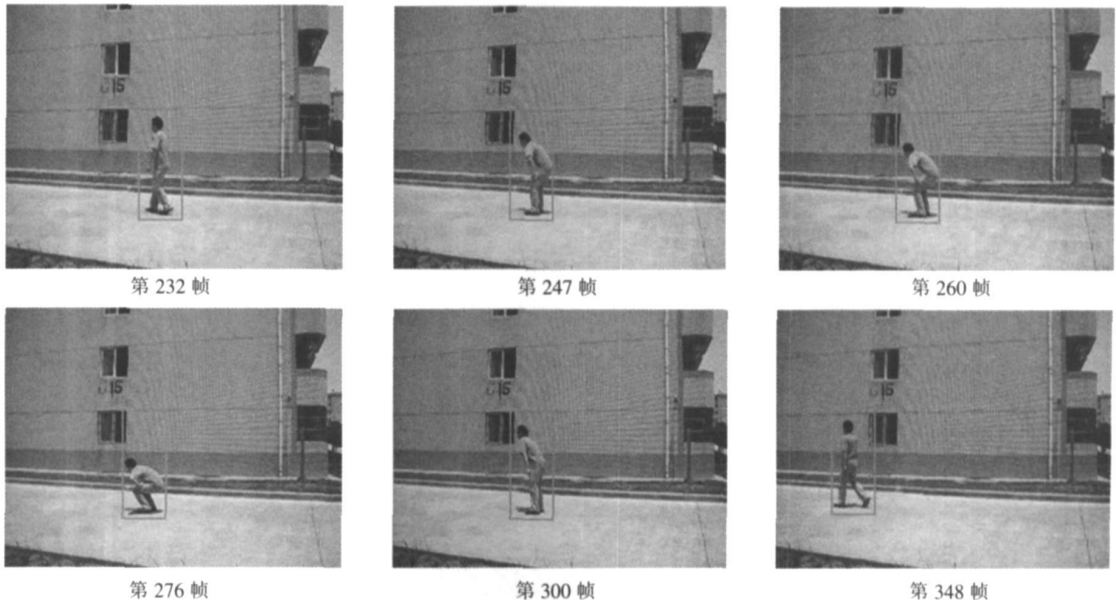


图 2 跟踪外形变化的行人

Fig.2 Tracking an appearance-changing passerby

姿态不断地发生改变,并且每次候选目标中所渗入的背景成分较多.如图 2 中视频图像第 (247)、(260)、(276)行人在做蹲下动作时,人体本身在候选模板中占的比例缩小了,而背景成分却增加了,如果直接生成普通直方图,背景因素的影响会使得一些原本属于目标的特征严重衰减.但采用本文算法构建新颖的直方图能够较好地描述目标对象,与实际目标对应性强,从而在搜索区域中生成辨析能力较强的相似度分布曲面,保证获得全局最优解,在跟踪过程中锁定跟踪目标.

6 结语

在基于贝叶斯理论及蒙特卡罗仿真的粒子滤波框架结构下,模板匹配精度决定了粒子采样的有效性.本文算法在粒子采样时通过新颖加权直方图构建来减少模板匹配中边缘噪声和交融背景的影响,并融合 HSV 颜色空间的相似度,在一定程度上提高了目标的辨别能力.该方法对采样粒子样本进行重新分配,使得粒子向目标状态的最大后验概率密度方向移动,提高了粒子的有效性,从而提高了视觉跟踪的可靠性和鲁棒性.两组仿真实验说明本文算法的目标识别能力和跟踪的准确度.

[参考文献] (References)

- [1] 王亮,胡卫明,谭铁牛.人运动的视觉分析综述[J].计算机学报,2002(3):225-237.
Wang Liang HuW eim ing TanT ien iu A survey ofvisual analysis of hum an motion [J]. Chinese Journal of Computers 2002 (3): 225-237. (in Chinese)
- [2] Lee K W, Ryu S W, Lee S J et al Motion based object tracking with mobile camera[J]. Electronics Letters 1998 34(3): 256-258
- [3] Yang C, Duraiswami R, Davis L. Efficient mean-shift tracking via a new similarity measure[C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition Los Alamitos IEEE, 2005 1: 176-183.
- [4] Barron J, Fleet D, Beauchemin S. Performance of optical flow techniques[J]. International Journal of Computer Vision 1994 12(1): 42-77.
- [5] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2003 25(5): 564-577.
- [6] Cheung G, Baker S, Kanade T. Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Madison IEEE, 2003 6(1): 77-84.
- [7] Nguyen H T, Smeyers A W M. Robust tracking using foreground-background texture discrimination[J]. Int J Comput Vis 2006 69(3): 277-293
- [8] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions Pattern Anal Mach Intel 2002 24(5): 603-619.
- [9] Gordon N, Salmond D. Novel approach to non-linear and non-Gaussian Bayesian state estimation[J]. Proceedings of Institute Electric Engineering 1993 140(2): 107-113.
- [10] 章毓晋. 图像工程[M]. 2版. 北京:清华大学出版社,2007 168-171.
Zhang Yujin. Image Engineering[M]. 2nd ed Beijing Tsinghua University Press 2007: 168-171. (in Chinese)

[责任编辑: 丁蓉]