

基于多线索概率分布图像融合的目标跟踪

李子龙, 鲍 蓉

(徐州工程学院 信电工程学院, 江苏 徐州 221008)

摘 要: 为了克服多线索目标跟踪中固定线索权值的不足, 提出采用颜色、运动历史、视觉显著性等多种概率分布图像来描述目标的观测信息, 通过目标在前一帧的位置确定中心区域和周边区域, 并计算不同概率分布图像基于这两个区域的直方图, 即中心直方图和周边直方图, 每个概率分布图像的置信度由其中中心直方图和周边直方图的差异度来描述, 根据不同线索的置信度, 在线调节当前帧各种概率分布图像在目标位置判断中所占的权重, 实现对目标的多线索融合跟踪, 实验结果表明, 比常用的固定线索权值的融合算法效果更好。

关键词: 目标跟踪; 多线索融合; 概率分布图像; 置信度

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2012)03-0022-04

A Novel Visual Object Tracking Algorithm Based on Multiple Clues Probability Distribution Image Fusion

LI Zi-long, BAO Rong

(School of Information and Electrical Engineering, Xuzhou Institute of Technology, Xuzhou 221008, China)

Abstract: In order to overcome shortcoming of fixed clues weight, Different features are used to generate a set of likelihood maps for each pixel indicating the probability of that pixel belonging to foreground object or scene background. The confidence score of each likelihood map is computed based on distinction of histograms of likelihood values on the object versus values from the surrounding background region, measured from the likelihood map of the previous frame. The evidence combination framework dynamically updates the weights such that, in the fused likelihood map, discriminative foreground/background information is preserved while ambiguous information is suppressed. Experimental comparisons demonstrate the proposed method outperform the classical fixed clues Weight fusion technique.

Key words: visual object tracking; multiple clues fusion; probability distribution image; confidence score

1 引言

对视频序列中的目标进行视觉跟踪是计算机视觉的一个重要研究方向。由于目标环境的多变性, 单一固定特征缺乏足够的鲁棒性对目标特征进行自适应的表达。为了提高目标特征表达的性能, 基于多线索融合的目标跟踪方法开始引起研究者的注意。

多线索融合的目标跟踪方法的优点是各个线索(特征)间可以实现信息互补。但是这种方法的难点是如何设计一个好的融合准则来实现特征信息间的最优融合。

最常用的多线索融合准则就是在融合各线索信息的时候, 对各种线索采用固定权值^[1-2]。这种融合方法的不足是: 如果跟踪期间有一个特征信息变化明显, 而其权值保持不变, 则融合得到的信息会不可靠, 利用这种信息跟踪视频目标会导致跟踪效果不理想, 甚至跟踪失败。

文献[3]利用颜色和形状的融合信息描述目标的观测信息, 提高了观测信息描述的鲁棒性。文献[4]采用颜色、运动和声音三种线索对粒子滤波的框架下目标可视跟踪进行了研究。Wu 等^[5]将多线索融合问题看为一个基于因式分解图模型的概率推理

问题.文献[6]将目标的灰度特征和梯度特征结合起来描述目标模板,进而采用 mean-shift 算法来稳定跟踪目标.文献[7]提出了通过学习得到的高层目标性状、颜色统计信息与底层的图像边缘信息以水平集能量函数的形式嵌入到粒子滤波框架下,对于非刚体目标跟踪有一定效果.本文提出了一种基于中心直方图和周边直方图对比的多线索概率分布图像融合的视频目标跟踪算法.

2 目标的多线索概率分布图像

在概率分布图像中,每个像素点的亮度表示该像素点属于目标的概率.如果像素点的亮度为 255(最大值),则表示该像素点属于目标物体的概率为 100%;如果像素点的亮度为 0(最小值),则表示该像素点属于目标物体的概率为 0.像素点的亮度越高,其属于目标的概率越大;反之,亮度越低,其属于目标的概率越小.通过特定的阈值对概率分布图像进行二值化,可以将目标物体从场景中分离出来.

2.1 颜色概率分布图像

在对刚体和非刚体目标的视频跟踪中,颜色直方图被广泛用于表示目标的外观特性.对颜色直方图采用反向投影操作,即原始图像中的像素值被颜色直方图中相应像素的统计量所代替,然后将得到的结果归一量化后就得到了颜色概率分布图像.

2.2 运动历史概率分布图像

运动历史图像法是一种运动分割方法^[8].它利用了连续图像中目标轮廓在空间上的相关性,通过每帧图像对应的不同时刻将连续图像加权叠加形成运动历史图像,然后再进行分割以得到最终目标.对运动历史图像归一化到 $[0, 255]$,就可以得到运动历史概率分布图像.

2.3 视觉显著性概率分布图像

根据认知心理学和神经物理学对人类视觉过程的深入研究发现,人观察周围环境时,往往由于行为目的需求或局部景象线索将注意力有选择地集中在某个或某些景物上,从而选择一定的点或者区域作为景物的代表,进行视觉信息处理,这就是视觉显著性.

根据对视觉显著区域的检测,可以得到视觉显著性概率分布图像.视觉显著性概率分布图像反映了目标的“醒目性”,即一个目标在视觉上与其周围环境差异的大小.本文采用文献[9]的方法进行视觉显著性特征的提取.

3 多线索概率分布图像的融合

3.1 概率分布图像的置信度

视频目标跟踪的问题,可以看作是根据输入图像 I 的信息推断状态变量 x (例如目标位置等)的问题,也即为求解 $P(x|I)$ 的问题.采用多个概率分布图像融合的方法来解决这个问题.

对图像 I ,从不同的角度,以不同的方式定义 M 个概率分布图像,假设其分布为 $P(x|L_i), i=1, \dots, M$.根据贝叶斯准则,可以得到:

$$P(x|I) = \int P(x|L)P(L|I)dL \approx \sum_{i=1}^M w_i P(x|L_i) \quad (1)$$

其中, M 是概率分布图像的总个数, $w_i = P(L_i|I)$ 是对每个独立概率分布图像有效性的置信度. $P(x|I)$ 可以看作是概率分布图像集合的加权和.

3.2 最优置信度分布的求解

根据式(1)的描述,问题的关键在于如何确定每个概率分布图像的权重,即其有效性的置信度大小.简单来说,就是对每个概率分布图像 $P(x|L_i)$ 的信赖程度.

定义 $1 \times M$ 概率分布图像置信度矩阵 $W = [w_1 \ w_2 \ \dots \ w_M]$,注意到 $\sum_{i=1}^M w_i = 1$,因此有 $W^T 1 = 1$,其中 $1 = [1 \ 1 \ \dots \ 1]$ 是 $1 \times M$ 矩阵,元素都为 1.

采用平方损失函数,则当前图像的最优的置信度分布 W_{opt} 可以通过求对真实概率 $P_t(x|I)$ 的最小均方误差的方法来解决:

$$W_{opt} = \operatorname{argmin}_W E \left[\left(\sum_{i=1}^M w_i P(x|L_i) - P_t(x|I) \right)^2 \right] \quad (2)$$

定义 $M \times M$ 协方差矩阵:

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1M} \\ C_{21} & C_{22} & \dots & C_{2M} \\ \dots & \dots & \dots & \dots \\ C_{M1} & C_{M2} & \dots & C_{MM} \end{bmatrix}$$

其中,协方差矩阵 C 的第 i 行第 j 列元素 C_{ij} 的值为:

$$C_{ij} = E \left[(P(x|L_i) - P_t(x|I))(P(x|L_j) - P_t(x|I)) \right] \quad (3)$$

则式(2)可以简化为:

$$W_{opt} = \operatorname{argmin}_W W^T C W \quad (4)$$

对于上面这个极值问题,本文采用拉格朗日乘数法来求解,求出的最优的置信度分布 W_{opt} 的第 i 个分量 $w_{i,opt}$ 为:

$$w_{i,opt} = \frac{C_{ii}^{-1}}{\sum_{j=1}^M C_{jj}^{-1}} \quad (5)$$

由式(5)可以看出, C_{ii}^{-1} 越大,第 i 个线索在多线索融合中所占权重应该越大,置信度越高。

3.3 中心直方图和周边直方图

在实际应用中,无法获知当前图像帧的真实概率 $P_i(x|I)$,为了解决这个矛盾,采用中心直方图和周边直方图的分布对比来求得 $w_{i,opt}$ 的近似解。在概率分布图像上,用一个覆盖目标且略大于目标范围的矩形(或椭圆)区域作为中心区域,用中心区域外的一个“口”形(或环形)区域作为周边区域,中心区域内的像素用来描述目标的特性,周边区域内的像素用来描述背景环境的特性。

设概率分布图像在中心区域的直方图为 $H_{bj}(b)$,周边区域的直方图为 $H_{bg}(b)$,其中 b 的取值范围是从 1 到 m , m 是直方图的级数。为了进行比较,需要分别对 $H_{bj}(b)$ 和 $H_{bg}(b)$ 进行归一化。记归一化的中心直方图和周边直方图分别为 $\{p(b)\}$ 和 $\{q(b)\}$,其中 $b \in \{1, \dots, m\}$ 。

中心直方图和周边直方图的对数概率比定义为:

$$L(b) = \log \frac{\max(p(b), \delta)}{\max(q(b), \delta)} \quad (6)$$

其中, δ 是一个很小的数(设其为 0.001)。引入 δ 的目的是为了避免出现被 0 除和对 0 取对数的情况。一般来说,对于任意一个 b , $p(b)$ 越大(属于目标的概率大), $q(b)$ 越小(属于背景的概率小),则对数概率比 $L(b)$ 越大。给定归一化直方图 $\{h(b)\}$,定义对数概率比 $\{L(b)\}$ 相对于 $\{h(b)\}$ 的方差为:

$$\begin{aligned} \text{var}(L; h) &= E[L^2(b)] - (E[L(b)])^2 \\ &= \sum_b h(b) L^2(b) - \left[\sum_b h(b) L(b) \right]^2 \end{aligned} \quad (7)$$

3.4 置信度的近似方法

本文用概率分布图像相对于中心直方图和周边直方图的方差来近似 C_{ii}^{-1} :

$$C_{ii}^{-1} \approx \frac{\text{var}(L; (p+q)/2)}{\text{var}(L; p) + \text{var}(L; q)} \quad (8)$$

其中, $\text{var}(L; p)$ 和 $\text{var}(L; q)$ 分别是概率分布图像的目标概率方差和背景概率方差, $\text{var}(L; (p+q)/2)$ 是概率分布图像的混合概率方差。如果目标概率方差和背景概率方差较低(表示目标像素和背景像素分别聚集在一起),混合概率方差较高(表示目

标像素和背景像素可以很好的区分开来),则这个概率分布图像具有很好的置信度,应该赋予在多线索融合中赋予较大的权重。

根据式(8),最优的置信度分布 W_{opt} 的第 i 个分量 $w_{i,opt}$ 可以近似为:

$$w_{i,opt} \approx \frac{VR(L_i; p, q)}{\sum_{j=1}^M VR(L_j; p, q)} \quad (9)$$

其中

$$VR(L_i; p, q) = \frac{\text{var}(L_i; (p+q)/2)}{\text{var}(L_i; p) + \text{var}(L_i; q)}$$

4 实验结果及分析

4.1 置信度的计算

图 1 是不同视频的置信度对比。图中上面一行图像是 ATON 视频库的 Highway II 视频的一帧,目标是一辆车。下面一行取自于 PETS2001 视频库中一段视频,目标是一个行人。

			
	0.416714	0.267047	0.316239
			
原视频	颜色 置信度	运动历史 置信度	视觉显著性 置信度

图 1 不同视频不同线索的置信度对比

在对车的跟踪中,由于车的颜色和背景路面的颜色差别较大,因此颜色置信度较高(0.416714)。而由于目标(汽车)运动速度较快,所以运动历史概率分布图像上很多目标过去的位置像素点仍然高亮,导致运动历史置信度较低(0.267047)。在对人的跟踪中,由于目标(行人)运动速度较慢,因此运动历史概率分布图像较准确,运动历史置信度较高(0.417349)。而目标的视觉显著性较差,因此视觉显著性置信度较低(0.244562)。

4.2 与其他融合方法的对比

常用的多线索融合方法中,都会事先约定每个线索对跟踪的贡献程度,而不是向本文算法一样自适应的更新不同线索的置信度。为了说明本文算法的效果,我们特将本文算法和固定置信度的多线索融合跟踪算法(简单起见,每个线索平等对待,赋予

相同的权重)对比。

图2显示了对PETS2001视频序列的一段视频使用固定置信度的多线索融合跟踪算法和本文算法多线索跟踪的效果对比。其中上面一行显示的是固定置信度的多线索融合跟踪算法对图像的第250、300、340帧的跟踪结果,下面一行是本文算法对图像的第250、300、340帧的跟踪结果。根据图像可以看出,固定置信度的多线索融合导致了对目标位置判断的误差,而本文算法则能较好的跟踪到目标。

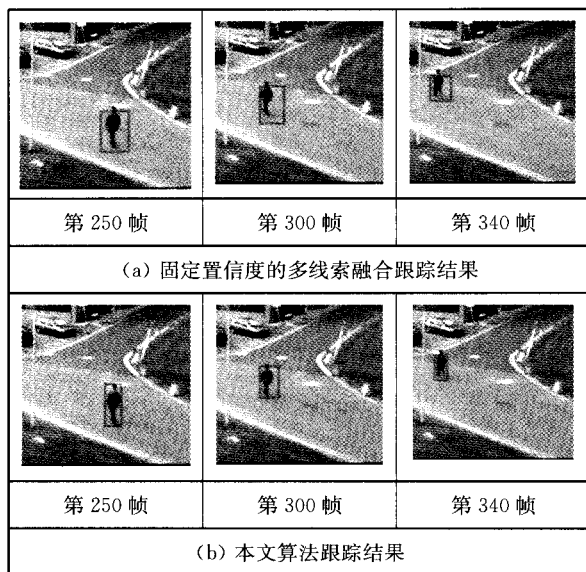


图2 本文算法与单线索跟踪算法的对比

5 结束语

本文对视频目标跟踪中的多线索融合问题进行了研究,提出了一种基于中心直方图和周边直方图对比的多线索概率分布图像融合的视频目标跟踪算法。实验结果表明,本文算法比起常用的固定线索权值的融合算法,效果更好。根据实际应用的需要,可以采用更多的概率分布图像来描述目标的观测信息。对实际目标选择最适合的几种概率分布图像,既可保证较好的跟踪效果,又可避免计算复杂度。

参考文献:

- [1] 冀治航,陈家新,黎蔚,等. 基于空间直方图的 CamShift 目标跟踪算法[J]. 微电子学与计算机, 2009, 26(7): 194-197.
- [2] 王楠,刘威,陈雪,等. 基于多线索混合的交通标志检测与跟踪[J]. 控制与决策, 2010, 25(8): 1277-1280.
- [3] 李安平. 复杂环境下的视频目标跟踪算法研究[D]. 上海:上海交通大学, 2006.
- [4] Perez P, Vermaak J, Blake A. Data fusion for tracking with particles[J]. the IEEE, 2004, 92(3): 495-513.
- [5] Wu Y, Huang T S. Robust visual tracking by integrating multiple cues based on co-inference learning [J]. International Journal of Computer Vision, 2004, 58(1): 55-71.
- [6] 陈爱华, 朱明, 王艳华, 等. 融合梯度特征的灰度目标跟踪[J]. 微电子学与计算机, 2009, 26(2): 69-71.
- [7] 赵晓林, 胡峰, 孙立国, 等. 基于多线索融合的非刚体目标精确跟踪[J]. 清华大学学报, 2010, 50(10): 1703-1707.
- [8] Bradski G R, Davis J W. Motion segmentation and pose recognition with motion history gradients [J]. Machine Vision and Applications, 2002, 13(7): 174-184.
- [9] Hou X, Zhang L. Saliency detection: a spectral residual approach[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA, 2007: 1-8.

作者简介:



李子龙 男, (1979—), 硕士, 讲师。研究方向为图像处理、模式识别与目标跟踪。