

一种分块表示的彩色目标跟踪算法

孙中森^{1,2}, 孙俊喜^{1,3}, 宋建中¹

(1.中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2.中国科学院研究生院, 北京 100039;

3.长春理工大学, 吉林 长春 130022)

摘要: 一种基于分块加权彩色直方图特征的目标表示方法。将图像分为部分叠加的块, 分别对每块计算加权量化彩色直方图, 构成直方图组。将直方图组用于基于 mean shift 的彩色目标跟踪系统, 利用 Kalman 滤波估计目标状态。

关键词: 目标跟踪 直方图组 mean shift 算法

目标跟踪是计算机视觉领域的一个分支, 它在视频监控、压缩、传输领域以及高技术武器装备方面都有重要的意义。一个鲁棒性好的跟踪器应该包括两部分: 一是对目标特征的可靠表示, 二是对背景复杂以及遮挡具有鲁棒性的精确估计^[1]。对于具有鲁棒性的跟踪估计过程已有很多研究, 包括基于概率的^{[2][3]}以及非概率的^[4]。目标的表示也有很多方法, 有灰度特征、颜色特征、纹理特征、边缘特征等, 其中的彩色直方图由于容易计算、信息量多而被广泛采用。但是直方图只包括图像的全局信息, 丢失了像素的空间信息, 不同的目标图像可能具有相同的直方图^[5], 这样在目标跟踪中由于背景中与目标相似颜色的出现而产生错误的匹配。在文献[6]、[7]中采用的单一加权距离彩色直方图中, 虽然考虑了位置信息, 但却过分强调中心点的颜色, 忽略了其他颜色。

本文采用将椭圆目标分割为具有五点支撑的子部分, 分别求取 RGB 空间的彩色加权直方图, 构成直方图组作为目标特征, 应用于基于 mean shift 的跟踪系统。

1 目标表示与相似度测量

1.1 目标模型表示

直方图是图像像素颜色的函数, 描述的是图像中具有该种颜色在图像中出现的频率^[5]。颜色直方图除以图像面积即可得到归一化直方图, 即图像的颜色概率密度函数(PDF), 表示目标图像的颜色概率分布。所谓加权直方图, 就是针对图像中不同区域的像素, 增加不同的权值, 再统计其出现的频率。如果增加的权值是到某一固定点的距离函数, 则直方图就具有了一定的位置信息。文献[6]、[7]的使用权值就是到目标中心距离的函数, 其包络为 Epanechnikov 核:

$$K_E = \begin{cases} \frac{1}{2} C_d^{-1} (d+2) (1 - \|x\|^2), & \|x\| \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

权值随着距离的增加递减。

增加权值以后, 目标中心位置的颜色得到加强, 而其他部分的颜色特征则削弱了。为了弥补此不足, 采用了将感兴趣区的目标区用一个外框椭圆, 内部包括五个局部区域的图形表示方法, 如图 1 所示。

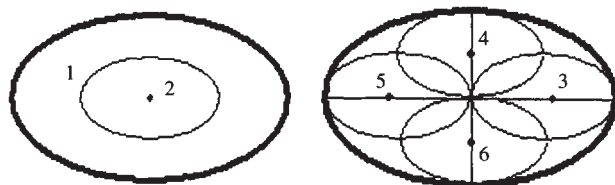


图 1 图形表示

第 1 部分表示整个目标区域, 反映目标的整体特征; 第 2 部分表示与第 1 部分同心的长短轴减半的区域, 反映目标的中心特征; 第 3 到第 6 部分区域大小与第 2 区域相同, 表示目标的右、上、左、下局部特征。分别求取以上 6 个区域的彩色加权直方图, 构成直方图组。

为减小目标尺度的影响, 分别在行、列方向选择不同的尺度 $h_{0,j} = \{h_{x,j}, h_{y,j}\}$, $j=1 \dots 6$, 将每个椭圆归一化为单位圆。设彩色直方图量化等级为 m , 表示目标密度函数为:

$$\hat{q}_j = \{\hat{q}_{j,u}\}_{u=1 \dots m} = C \sum_{i=1}^n k(\|x_{j,i}^* - u\|) \delta(b(x_{j,i}^*) - u), j=1 \dots 6 \quad (2)$$

其中: $x_{j,i}^*$ 为第 j 个区域的像素, $\|x_{j,i}^*\|^2$ 为其到区域中心的距离, C 为满足 $\sum_{u=1}^m \hat{q}_{j,u} = 1$ 的归一化常数。候选目标的密度函数为:

$$\begin{aligned} \hat{p}_j(s) &= \{\hat{p}_{j,u}(s)\}_{u=1 \dots m} \\ &= C_{j,h} \sum_{i=1}^n k\left(\left\|\frac{s - x_{j,i}}{h_j}\right\|^2\right) \delta(b(x_{j,i}) - u) \end{aligned} \quad (3)$$

$C_{j,h}$ 为归一化常数, $h_j = \{h_{j,c}, h_{j,l}\}$ 为窗口半径。采用式

本刊邮箱: eta@ncse.com.cn

(2)、式(3)表示的目标特征模板计算相似性函数具有核轮廓函数 $k(x)$ 的特性^[7]。

1.2 相似度表示

通常采用 Bhattacharyya 系数表示两个概率分布的相关性,其离散估计表达式为:

$$\hat{\rho}(s) = \sum_{u=1}^m \sqrt{\hat{p}_u(s) \hat{q}_u} \quad (4)$$

分别计算各部分的概率分别相关度,整体的相关度表示如下:

$$\hat{\rho}(s) = \sum_{j=1}^N k_j \hat{\rho}_j(s) = \sum_{j=1}^N k_j \times \sum_{u=1}^m \sqrt{\hat{p}_u(s) \hat{q}_u} \quad (5)$$

其中: N 为块个数,此处 $N=6$; k_j 为各块相关度的加权值,满足下式:

$$\sum_{j=1}^N k_j = 1, k_j > 0 \quad (6)$$

设预测目标位置为 s_0 , 则将 $\hat{\rho}(s)$ 在 s_0 处泰勒展开,近似为:

$$\begin{aligned} \hat{\rho}(s) &= \frac{1}{2} \sum_{j=1}^N \sum_{u=1}^m \sqrt{\hat{p}_{j,u}(\hat{s}_{0,j}) \hat{q}_{j,u}} \\ &+ \sum_{j=1}^N \frac{C_{j,h}}{2} \sum_{i=1}^n w_{j,i} k\left(\left\|\frac{s - x_{j,i}}{h_j}\right\|^2\right) \end{aligned} \quad (7)$$

$$\text{其中: } w_{j,i} = \sum_{u=1}^m \sqrt{\frac{\hat{q}_{j,u}}{p_{j,u}(\hat{s}_{0,j})}} \delta(b(x_{j,i}) - u) \quad (8)$$

式(7)的极大值点就是搜索目标的当前位置,由于第一项为定值,即为核密度函数 $K(x)$ 表示的加权密度估计的最大值位置。

2 mean shift 跟踪算法

Mean Shift 是一种基于核函数密度估计的图像特征分析方法。将图像分成若干子集,每个图像子集的感兴趣特征参数构成目标特征空间。最重要的特征对应特征空间的密集区域。描述这一特征空间常采用概率密度估计的方法,即将特征空间看成特征参数的经验概率密度分布。特征空间的密集区对应于这一概率密度分布的局部极大值,即分布的众数。Mean shift 就是在不估计概率密度分布的情况下,采用梯度优化算法寻找到这一密度函数的众数的一种搜索算法。当 mean shift 用于视频跟踪定位时,设定目标区域的某些特征(如颜色、纹理)为感兴趣的特征参数,将后续图像序列分别映射到这一特征空间,并用核概率密度分布描述,它实际反映的是目标与候选目标在特征空间的相似性。这一密度分布的众数就是与目标最相似的特征,它对应的位置即为目标当前位置。

已知 n 点特征空间的采样数据为 $x_i, i=1, 2, \dots, n, x_i \in R^d$, 核函数 $K(x)$ 的窗口半径 h , 在点 x 处的核函数密度估计值为:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right) \quad (9)$$

核函数 $K(x)$ 有多种不同的形式,常用的有 Epanechnikov 核以及正态核。正态核是最优的核,但其计算量大。Epanechnikov 核是估计密度与真实密度间平均全局误差最小化的核^[4], 表达式如下:

$$K_E = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \|x\|^2), & \|x\| \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

c_d 为常数。这一密度函数的众数点位置即为 $f(x)$ 的极大值点,而这样的点位于 $\nabla f(x)=0$ 处^{[5][6]}。为此引入核函数轮廓函数 $k(x)$ 。假定 $k(x)$ 在 $x \in [0, 1]$ 上除少数点外导数存在,其微分表示为 $g(x)$, 则 $g(x) = -k'(x)$, 则 K 核密度估计的导数表示为:

$$\begin{aligned} \hat{\nabla} f_{hc}(x) &= \frac{2}{nh^{d+2}} \times \left[\sum_{i=1}^n g\left(\left\|\frac{x - x_i}{h}\right\|^2\right) \right] \\ &\times \left[\frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x \right] \end{aligned} \quad (11)$$

令 $\hat{\nabla} f_{hc}(x)=0$, 则有:

$$M_{h,G} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x = 0 \quad (12)$$

$M_{h,G}$ 为 Mean Shift 向量, 令:

$$\hat{f}_{hc}(x) = \frac{C}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{x - x_i}{h}\right\|^2\right) \quad (13)$$

$$\text{则, } M_{h,G} = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(x)}{\hat{f}_{h,G}(x)} \quad (14)$$

Mean shift 递推过程就是计算 mean shift 的向量 $M_{h,G}$, 并根据 $M_{h,G}$ 移动 K 核中心的过程。

设 K 核中心位置的递推序列为 $\{s_j\}_{j=1, 2, \dots}$, 则递推序列如下:

$$s_{j+1} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{s_j - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{s_j - x_i}{h}\right\|^2\right)} \quad (15)$$

3 实验结果

用 Kalman 滤波预测目标的实际运动状态^[8], 选择文献[9]中的状态表达式以及状态方程。实验中,选取 $\rho=0.5, \sigma=0.1, \sigma_r=3, \Delta t=0.04, \lambda=1.2, \xi=0.7$ 。不考虑目标尺度变化,即 $h=h_0$, 采用 VS-PETS 中的序列图像。图 2 为序列图像分块直方图组的部分跟踪结果,图中的松树为遮挡区。第 48 帧目标(骑自行车的人)行使途中要进入遮挡区,第 128 帧目标被挡在遮挡物后面,由 kalman 预测目标在当前帧位置,第 170 帧目标穿过遮挡区被重新捕获,继续跟踪。图中的白色线条为目标运动的轨迹。



图2 序列图像分块直方图组的部分跟踪算法

本文利用加距离权彩色直方图作为目标特征,增加了算法对遮挡的鲁棒性;采用遮挡因子 occ 表示目标遮挡的程度,能够准确指示目标实际遮挡的程度。在目标遮挡后重新出现时能够快速捕获到目标,继续跟踪。

但是,当目标较长时间遮挡,对目标位置估计误差增大时,靠增加目标搜索的范围来实现重新捕获,将会大大增加计算量。

参考文献

- [1] POLAT E, YEASIN M, SHARMA R. A 2D/3D model-based object tracking framework[J]. Pattern Recognition, 2003, 36: 2127-2141.
- [2] CORMICK J M, BLAKE A. A probabilistic exclusion princi-

ple for tracking multiple objects.in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Corfu, Greece, 1999.

- [3] RASMUSSEN C, HAGER G D. Probabilistic data association methods for tracking complex visual objects. IEEE Trans. Pattern Anal. Mach. Intell, 2001; 23(6): 560-576.
- [4] HARITAOGLU I, HARWOOD D, DAVIS L. W4: real-time surveillance of people and their activities. IEEE Trans. Pattern Anal. Mach. Intell, 2000, 22(8): 809-830.
- [5] 朱志刚. 数字图像处理[M]. 北京: 电子工业出版社, 2002.
- [6] NUMMIARO K, ESTHER K M, GOOL L V. Color feature for tracking non-rigid objects[J]. 自动化学报, 2003, 29(3).
- [7] COMANECI D, RAMESH V, MEER P. Kernel-based object tracking[J]. Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5).
- [8] WELCH G, BISHOP G. An Introduction to the kalman filter[M]. UNC- Chapel Hill, TR 95-041, April 5, 2004.
- [9] GREWAL M S, ANDREWS A P. Kalman filtering: theory and practice using matlab[M], Second Edition, 2001.

(收稿日期: 2006-09-20)

(上接第56页)

距离的增大而增加;而 FRR 随着汉明距离的增大而减小。FAR 和 FRR 曲线的交叉点,即等错误率(ERR)为 1.85%,这说明本文的算法是可行的和有效的。

在识别模式下,将虹膜图像分为二组,从每一类的两个阶段采集的虹膜图像中各选一幅组成 GROUP1,共包含 216 幅虹膜图像,作为标准虹膜模板;GROUP2 包含剩下的 540 幅图像,作为测试的虹膜图像。识别时取阈值为 0.39,将测试图像与每一类的两个模板的汉明距离中较小的距离作为与该类的距离,其识别率可达 98.22%。如果应用系统对安全性的要求较高,可以适当地下调识别阈值。

本文提出了一种基于 Log-Gabor 子带图像分割的虹膜识别算法,该算法克服了传统方法中实现旋转不变时存在的缺点,并且具有平移和放缩不变性。通过对包含 756 幅图像组成的虹膜库进行实验,取得了良好的效果。

参考文献

- [1] JAIN A. Biometrics personal identification in networked society[M]. Kluwer Academic Publishers, 1999.
- [2] DAUGMAN J. High confidential visual recognition by test of statistical independence[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(11): 1148-1161.
- [3] WILDES R P. Automated iris recognition: an emerging biometric technology[J]. Proceeding of the IEEE, 1997, 85(9):

1348-1363.

- [4] LIM S, LEE K, BYEON O, et al. Efficient iris recognition through improvement of feature vector and classifier[J]. ETRI Journal, 2001, 23(2): 61-70.
- [5] 王蕴红, 朱勇, 谭铁牛. 基于虹膜识别的身份鉴别[J]. 自动化学报, 2002, 28(1): 1-10.
- [6] 王林, 练秋生. 虹膜定位算法研究[J]. 电视技术, 2006, (7).
- [7] DAUGMAN J. Biometric personal identification system based on iris analysis[P]. US Patent 5291560, 1994.
- [8] Li Ma, TieNiu Tan. Efficient iris recognition by characterizing key local variations[J]. IEEE Transactions on Image Processing, 2004, 13(13): 739-750.
- [9] GABOR D. Theory of communication[J]. Journal of Institute for Electrical Engineering, 1946, (93): 429-457.
- [10] DAUGMAN J. Uncertainly relation for resolution in space, spatial frequency and orientation optimized by two dimension visual cortical filters[J]. Journal of the Optical Society of America, 1985, (2): 1160-1169.
- [11] FIELD D J. Relations between the statistics of natural images and the response properties of cortical cells[J]. Journal of the Optical Society of America A, 1987, 4(12): 2379-2394.
- [12] 张德馨. 基于虹膜识别的身份鉴别方法研究[D]. 硕士学位论文. 北京: 中国科学院自动化研究所, 2003.

(收稿日期: 2006-11-21)