

# 基于梯度特征与彩色特征相融合的 mean shift 跟踪方法

李平生, 李小霞

(西南科技大学 信息工程学院, 四川 绵阳 621010)

**摘要:** 针对 mean shift 跟踪方法中存在的光照变化不稳定问题, 提出了基于梯度特征与彩色特征相融合的 mean shift 跟踪方法。首先分别提取目标的梯度特征和彩色特征, 利用多尺度的相似度计算方法进行特征的匹配, 然后通过最大化相似度对目标进行跟踪。通过物体和人体等运动目标的跟踪, 验证了改进的跟踪算法在光照变化情况下的鲁棒性优于原有的算法, 显著降低了跟踪位置误差。

**关键词:** 目标跟踪; mean shift; 梯度图像; 多尺度相似度量

中图分类号: TP391.41

文献标识码: A

文章编号: 1674-7720(2011)03-0035-04

## Mean shift tracking algorithm based on gradient feature and color feature fusion

Li Pingsheng, Li Xiaoxia

(School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China)

**Abstract:** A novel mean shift algorithm based on object tracking gradient feature and color image fusion is proposed for the illumination unstable problem in traditional mean shift method. Firstly, gradient features and color features of the target are extracted separately, the features are matched using multi-scale similarity calculation method. Then, the target can be tracked by maximizing the similarity. Experiments on the tracking of moving targets such as object and human demonstrate that the proposed algorithm has more robust than the original algorithm under the situation of illumination changes and reduces the tracking position error obviously.

**Key words:** object tracking; mean shift; gradient image; multi-scale similarity

目标跟踪是计算机视觉领域的热门研究课题, 在军事制导、视觉导航、安全监控、视频编码等方面有着广泛的应用。基于 mean shift 的目标跟踪方法由于其计算速度快、模型简单和较好的跟踪性能, 近年来一直受到广大研究者的重视。这种方法采用加权颜色直方图的形式对目标进行特征描述, 在每帧图像中通过求解目标模型和候选模型相似度的最大值来确定目标的位置。虽然 mean shift 跟踪方法采用的模型比较简单, 对目标的形变、旋转有一定的稳定性, 但是这种方法只利用了目标的颜色信息, 当场景中光照发生变化时, 目标的颜色分布也会随之变化, 从而导致跟踪过程的不稳定性。同时, mean shift 跟踪方法中采用单层直方图的目标特征描述, 在光照变化时, 会带来较大的误差。

针对 mean shift 跟踪方法跟踪过程中的光照不稳定的问题, Hager<sup>[2]</sup>提出了建立精确的模型来对光照变化建模; Freedman<sup>[3]</sup>结合了光流信息, 采用 Graph-Cut 算法提

出了一种光照鲁棒的跟踪方法; 针对目标直方图丢失了目标的空间信息的缺点, Yang<sup>[4]</sup>等人提出了一种新的目标表示方法及其对应的 mean shift 跟踪算法, 以反映目标的空间信息; 李培华<sup>[5]</sup>提出了将图像的位置信息加入到颜色直方图中; 针对原始算法只采用一个特征的缺点, 王永忠<sup>[6]</sup>等人提出将颜色信息和纹理信息进行自适应融合的跟踪方法; 针对跟踪过程中单层相似度存在匹配不准确的问题, Rubner<sup>[7]</sup>提出了一种 cross-bin 直方图度量, 考虑了不同位置的特征量化级(bin)之间的相关性; Granman<sup>[8]</sup>提出了一种用于物体类别识别的匹配方法 (Pyramid Match), 将特征空间划分为多层次的 Pyramid 直方图, 在不同尺度下计算 bin 之间的相似度, 融合了多尺度的 bin 之间的关系。虽然这些方法都在一定程度上改进了原有的 mean shift 算法, 但是都存在计算量大的问题, 不能满足实时性的要求。

为了提高跟踪算法的鲁棒性, 本文提出的方法首先

融合了目标的梯度特征。该特征计算简单,描述了目标的边缘特性,对光照的变化具有较强的鲁棒性。其次,在目标匹配的过程中,采用多尺度直方图特征进行目标的匹配,这种匹配方法避免了由于光照变化引起的单尺度直方图匹配的误差。

### 1 梯度图像

梯度特征描述了图像的边缘、角点等局部区域变化的信息,对于光照的变化具有较强的鲁棒性,广泛运用在目标特征描述、图像匹配和目标检测中。本文采用了梯度特征的幅度信息,首先用 Sobel 算子分别计算图像在  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$  和  $135^\circ$  的幅值  $I_1$ 、 $I_2$ 、 $I_3$ 、 $I_4$ ,然后合并各个方向的梯度幅值图像,其计算如下:

$$I=(I_1+I_2+I_3+I_4)/4 \quad (1)$$

其中  $I$  为合并后的图像。

Sobel 算子模板如图 1 所示。

-1	-2	-1	-2	-1	0	-1	0	1	0	1	2
0	0	0	-1	0	1	-2	0	2	-1	0	1
1	2	1	0	1	2	-1	0	1	-2	-1	0
(a) $0^\circ$	(b) $45^\circ$	(c) $90^\circ$	(d) $135^\circ$								

图 1 Sobel 算子模板

### 2 目标模型

设  $\{x_i\}_{i=1 \dots n}$  为目标区域的像素点集,区域中心位于  $x_0$  点。用凸核函数对各点进行加权。定义函数  $b(x)$ ,  $b(x_i)$  是像素  $x_i$  在量化的特征空间  $\{t=1 \dots T\}$  的量化级。则量化特征  $t$  在目标模型中出现的概率可表示为:

$$q_t=C_q \sum_{i=1}^n k\left(\left\|\frac{x_0-x_i}{h}\right\|^2\right) \delta[b(x_i)-t] \quad (2)$$

其中,常数  $C_q$  是归一化因子,使得  $\sum_{t=1}^T q_t=1$ ,  $\delta$  是 Kronecker 函数; $h$  为窗宽。

在得到了目标模型后的每一帧中,假定候选目标的位置在  $y$ , 可以通过计算  $y$  的特征分布和目标分布的相似度判断  $y$  是否是真实位置。设候选目标的像素点集为  $\{y_i\}_{i=1 \dots n}$ , 以  $y$  为中心,使用相同的核函数对候选目标的各个像素进行加权,则特征  $t$  在候选目标中出现的概率为:

$$p_t(y)=C_q \sum_{i=1}^n k\left(\left\|\frac{y-y_i}{h}\right\|^2\right) \delta[b(x_i)-t] \quad (3)$$

### 3 分层相似度量

原始的 mean shift 采用单层的 Bhattacharyya 系数来表示相似度:

$$\rho[p(y)q]=\sum_{i=1}^T \sqrt{p_i(y)q_i} \quad (4)$$

但是由于光线的变化,目标的特征分布,特别是目标的颜色分布也会发生变化。因此,这种单层的相似度量就会出现比较大的误差,进而出现跟踪的失败。为了

解决这个问题,本文采用了多层的相似度量。首先要得到多尺度的颜色特征分布  $p_i^{\text{color}}$  和  $q_i^{\text{color}}$  的多尺度描述。第  $l$  层( $l=1 \dots L$ )将特征空间划分为  $2^l$  个 bin, 每层可以分别得到直方图:  $q_1, q_2, q_3 \dots q_L$  和  $p_1(y), p_2(y), p_3(y) \dots p_L(y)$ , 而  $q_i^{\text{color}}$  和  $p_i^{\text{color}}(y)$  之间的相似度由不同尺度下的 Bhattacharyya 系数加权计算得到:

$$\rho^{\text{color}}(y)=\sum_{i=1}^L \beta_i \rho[p_i(y)q_i]=\sum_{i=1}^L \frac{1}{2^{L-l}} \sum_{i=1}^{2^l} \sqrt{p_i(y)q_i} \quad (5)$$

其中,  $\beta_i$  随着每层 bin 数的增加而增大,从而给精细层更大的权值,以保证精细特征在整个特征空间中具有较大的分量。这种匹配方法考虑了多尺度之间的关系,在光照变化情况下具有更好的鲁棒性。

### 4 改进的 mean shift 算法

本文中采用了梯度图像和彩色图像相融合的 mean shift 跟踪方法,在得到目标模板和候选模板的相似度后,目标的位置由最大化相似度得到:

$$y=\max((1-\alpha)\rho^{\text{color}}[p(y)q]+\alpha\rho^{\text{grad}}[p(y)q]) \quad (6)$$

其中  $\alpha$  表示了梯度特征所占的比重。

为了求取  $y$  的最大值,把式(6)在初始位置  $p_u(y_0)$  处用 Taylor 公式展开得到:

$$\begin{aligned} \rho^{\text{color}}+\alpha\rho^{\text{grad}} &\approx \frac{1}{2}[(1-\alpha)\sum_{i=1}^L \frac{1}{2^{L-l}} \sum_{i=1}^{2^l} \sqrt{p_i^{\text{color}}(y_0)q_i^{\text{color}}} + \\ &\alpha \sum_{i=1}^T \sqrt{p_i^{\text{grad}}(y_0)q_i^{\text{grad}}] + \frac{C}{2}[(1-\alpha)\sum_{i=1}^L \frac{1}{2^{L-l}} \sum_{i=1}^{2^l} p_i^{\text{color}}(y) \\ &\sqrt{\frac{q_i^{\text{color}}}{p_i^{\text{color}}(y_0)}} + \alpha \sum_{i=1}^T p_i^{\text{grad}}(y) \sqrt{\frac{q_i^{\text{grad}}}{p_i^{\text{grad}}(y_0)}}] \end{aligned} \quad (7)$$

其中  $y_0$  为上一帧的真实位置。

式(7)右边第一项是一个常数,要想求取式(7)的最大值,等价于求取右边第二项的最大值,通过对  $y$  求偏导,利用 mean shift 理论,可以得到目标的新位置:

$$y_1=\frac{\sum_{i=1}^n [((1-\alpha)x_i\omega_i^{\text{color}}+\alpha x_i\omega_i^{\text{grad}}) \cdot g\left(\left\|\frac{y_0-x_i}{h}\right\|\right)]}{\sum_{i=1}^n [((1-\alpha)\omega_i^{\text{color}}+\alpha\omega_i^{\text{grad}}) \cdot g\left(\left\|\frac{y_0-x_i}{h}\right\|\right)]} \quad (8)$$

其中:

$$\omega_i^{\text{color}}=\sum_{l=1}^L \frac{1}{2^{L-l}} \sum_{i=1}^{2^l} \sqrt{\frac{q_i^{\text{color}}}{p_i^{\text{color}}(y_0)}} \delta(b(x_i)-t) \quad (9)$$

$$\omega_i^{\text{grad}}=\sum_{i=1}^T \sqrt{\frac{q_i^{\text{grad}}}{p_i^{\text{grad}}(y_0)}} \delta(b(x_i)-t) \quad (10)$$

$$g(x)=-k'(x) \quad (11)$$

在本方法中用到的核函数  $k(x)$  是 Epanechnikov 核:

$$k(x)=\begin{cases} 1-r^2, & -1 \leq r \leq 1 \\ 0, & \text{其他} \end{cases} \quad (12)$$

在计算过程中不断地迭代式(8),直到  $\|y_1-y_0\|$  小于设定的阈值,就得到了目标的真实位置。

整个算法的步骤如下:

《微型机与应用》2011年第30卷第3期

(1)初始化:给出目标模型  $\{q_i\}_{i=1, \dots, 2', l=\dots, L}^{color}$  和  $\{q_i\}_{i=1, \dots, T}^{grad}$  以及目标上一帧的位置  $y_0$ 。

(2)在给定的初始位置  $y_0$  处提取特征,计算  $\{p_i(y_0)\}_{i=1, \dots, 2', l=\dots, L}^{color}$  和  $\{p_i(y_0)\}_{i=1, \dots, T}^{grad}$ , 得到相似度:

$$\rho(y_0) = (1-\alpha)\rho^{color}(y_0) + \alpha\rho^{grad}(y_0) = (1-\alpha) \sum_{l=1}^L \frac{1}{2^{L-l}} \sum_{i=1}^{2'} \sqrt{p_i^{color}(y)q_i^{color}} + \alpha \sum_{i=1}^T \sqrt{p_i^{grad}(y)q_i^{grad}} \quad (13)$$

(3)根据式(9)、式(10)分别求得  $\omega_i^{color}$  和  $\omega_i^{grad}$ 。

(4)根据式(8)得到下个位置  $y_1$ 。

(5)计算  $\{p_i(y_1)\}_{i=1, \dots, 2', l=\dots, L}^{color}$  和  $\{p_i(y_1)\}_{i=1, \dots, T}^{grad}$ , 得到新的位置的相似度  $\rho(y_1)$ 。

(6)如果  $\rho[p(y_1)q] < \rho[p(y_0)q]$ , 则  $y_1 = (y_0 + y_1) / 2$ , 直到  $\rho[p(y_1)q] > \rho[p(y_0)q]$ 。

(7)若  $\|y_1 - y_0\| < \varepsilon$ , 则迭代停止; 否则  $y_0 = y_1$  转第(3)步。

### 5 试验结果

仿真环境为: 联想台式机电脑, 其 CPU 为 Intel(R) Pentium(R)D2.80 GHz, 内存 1 GB; 操作系统为 Windows XP Professional, SP2, 软件采用 VC6.0 及 OPENCV1.0。

实验用到的视频序列是在实验室环境下采集到的物体运动的视频和人体运动的视频, 序列图像的尺寸为 640×480, 在图像序列一中, 人手拿的一支笔在视野中做无规则的运动, 如图 2 所示。视频序列二是人体在视野中做无规则的运动, 在跟踪的过程中阳光从窗户照进来, 目标在靠近窗户和远离窗户的时候, 其外观特征会跟随光照的变化而变化, 如图 3 所示。实验由人工手动捕捉目标。

实验比较了原始的 mean shift 算法和改进的 mean shift 算法, 两种算法都用到图像的 RGB 色彩空间。其中原始的 mean shift 算法采用 16×16×16 的直方图, 改进的 mean shift 算法则采用 3 尺度的颜色模型, 每个尺度的直方图的 bin 分别为 16×16×16、8×8×8、4×4×4。

梯度图像采用 16×16×16 的直方图。在通过大量的实验得出的取值范围在 0.3~0.6 之间最好, 本文中  $\alpha$  取 0.5。



图 2 物体效果比较



图 3 人体效果比较

从图 2、3 可以看出, 原始的 mean shift 跟踪方法由于受到光照的影响, 当目标从光线弱的区域进入到光线较强的区域时, 目标的颜色特征发生了变化, 导致了目标的跟踪不稳定, 以至于最后失去了目标。改进的 mean shift 跟踪方法由于融合了目标的光照不变特征描述, 在光照变化时, 还是能够匹配到目标, 因此在整个过程中始终能够跟踪到目标。而且在相似度的计算上采用多尺度的颜色特征描述, 在一定程度上增加了色彩特征匹配的鲁棒性。

分别计算两种算法中每一帧的跟踪结果与参考位置的误差, 图 4 为两种算法在实验中的位置误差比较的结果。可以看出, 在光线变化缓慢的区域两种算法的误差接近, 但是当光线变化显著时, 原始算法的误差就会大大地增加, 以至于最后失去目标, 而改进的 mean shift 跟踪方法仍能够跟踪目标。

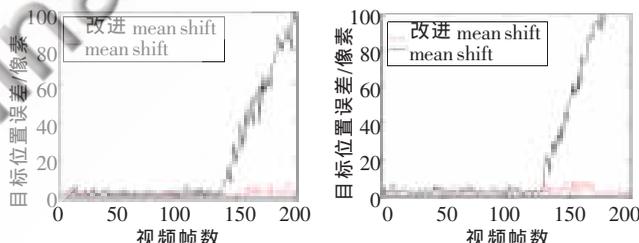


图 4 跟踪误差比较

为了比较改进的与原始的跟踪算法, 实验采用了 Li<sup>[9]</sup> 等人提出的指标  $S(A, B) = \frac{A \cap B}{A \cup B}$  评价相似度(A 为检测出的目标区域, B 为真实的区域),  $S(A, B)$  的值在 0~1 之间, 当 A、B 完全相同, 相似度则为 1; A、B 差别越大, 相似度越接近 0。由于序列较长, 且原始算法中出现跟踪失败的情况, 所以从每个视频中抽取相同的光照变化区域的 50 帧进行比较, 其结果如表 1 所示。

表 1 运动目标跟踪性能比较  $\sum S(A, B)$

序列名称	原始跟踪方法	本文跟踪方法
物体	0.5033	0.7126
人体	0.4832	0.7935
平均结果	0.4933	0.7531

从表 1 可以看出,本文提出的算法在跟踪过程中平均相似度比原始算法提高了 0.252 8,这说明了本文的算法在跟踪过程中比原始算法更加精确。

由于 mean shift 算法本身存在速度的限制,在跟踪过程中,目标的运动速度不能过快。通过实验得出在前后两帧中,目标的运动速度不能超过目标区域长度的一半。

本文将图像梯度特征引入到 mean shift 跟踪算法中,同时,该方法采用了多尺度的相似度度量,物体和人体的跟踪实验验证了改进的跟踪算法在光线的变化时具有鲁棒性,显著降低了跟踪位置误差,相似度提高了 0.252 8。mean shift 跟踪算法本质上是梯度下降跟踪方法,对目标的运动速度有较大的限制,在以后的研究中将尝试采用粒子滤波的方法来解决这个问题。

#### 参考文献

- [1] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25:564-577.
- [2] HAGER G D, BELHUMEUR P N. Efficient region tracking with parametric models of geometry and illumination[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(10):1025-1039.
- [3] FREEDMAN D, TUREK M. Illumination-invariant tracking via graph cuts [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego: IEEE Press, 2005:10-17.

- [4] YANG C J, DURAISWAMI R, DAVIS L S. Efficient mean-shift tracking via a new similarity measure [C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005:176-183.
- [5] 李培华. 一种改进的 Mean Shift 跟踪算法[J]. 自动化学报, 2007, 33(4): 347-354.
- [6] 王永忠, 梁彦, 赵春晖, 等. 基于多特征自适应融合的核跟踪方法[J]. 自动化学报, 2008, 34(4): 393-399.
- [7] RUBNER Y, TOMASI C, GUIBAS L J. The earth mover's distance as a metric for image retrieval [J]. International Journal of Computer Vision, 2000, 40(2):99-121.
- [8] GRAUMAN K, DARRELL T. The pyramid match kernel: discriminative classification with sets of image features [C]. IEEE International Conference on Computer Vision. Beijing: IEEE Press, 2005:1458-1465.
- [9] LI L Y, HUANG W M, GU I Y H, et al. Statistical modeling of complex backgrounds for foreground object detection[J]. IEEE Transactions on Image Processing, 2004, 13(11):1459-1472.

(收稿日期: 2010-09-07)

#### 作者简介:

李平生,男,1983年生,硕士研究生,主要研究方向:目标跟踪。

李小霞,女,1976年生,副教授,硕士研究生导师,主要研究方向:模式识别,生物特征识别。