

DOI: 10.11992/tis.201605033

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20160808.0831.022.html>

## 粒化的 Mean Shift 行人跟踪算法

刘翠君, 赵才荣, 苗夺谦, 王学宽

(同济大学 电子与信息工程学院, 上海 201804)

**摘要:** Mean Shift 行人跟踪采用颜色特征直方图作为跟踪特征, 存在易受背景颜色干扰等问题。基于此, 在传统的 Mean Shift 行人跟踪算法中引入粒计算的思想, 提出粒化的 Mean Shift 行人跟踪算法, 对图像目标区域作粒层分块来提取块颜色特征信息, 并在颜色特征表示上作不同粒度的粒化, 最后在 Mean Shift 迭代框架下实现行人跟踪。该方法相比传统的跟踪方法具有计算复杂度更低、稳健性更好的优点。在 PETS2009 和 CAVIAR 数据库做的实验表明, 这种方法跟踪正确率更高, 在颜色干扰下稳健性更好, 能够实时有效地跟踪行人。

**关键词:** 信息粒; 粒计算; Mean Shift; 特征提取; 行人跟踪

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1673-4785(2016)04-0433-09

中文引用格式: 刘翠君, 赵才荣, 苗夺谦, 等. 粒化的 Mean Shift 行人跟踪算法[J]. 智能系统学报, 2016, 11(4): 433-441.

英文引用格式: LIU Cuijun, ZHAO Cairong, MIAO Duoqian, et al. Granular mean shift pedestrian tracking algorithm[J]. CAAI Transactions on Intelligent Systems, 2016, 11(4): 433-441.

## Granular mean shift pedestrian tracking algorithm

LIU Cuijun, ZHAO Cairong, MIAO Duoqian, WANG Xuekuan

(College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China)

**Abstract:** Mean shift pedestrian tracking that uses a color histogram as its tracking feature has drawbacks, e.g., performance can easily be affected by the introduction of a background color. To solve this problem, the idea of granular computing was introduced into the traditional mean shift pedestrian tracking algorithm, and a new granular mean shift pedestrian tracking algorithm, based on granular computing, is presented. The algorithm blocks the image's target area with specific granularity to extract color features, then adopts different color channels of granulation on the feature, and finally realizes target tracking under the framework of the mean shift iteration. Compared with other traditional methods the algorithm displays lower computational complexity and is more robust. Experimental results on PETS2009 and CAVIAR databases show that the algorithm achieves a higher tracking accuracy, better robustness and efficiency under color interference, and can track the target pedestrian in real time.

**Keywords:** information granules; granular computing; mean shift; feature extraction; pedestrian tracking

行人跟踪涉及行人检测和行人跟踪两部分。行人检测属于运动目标检测, 目的是从图像序列中将行人目标从背景图像中提取出来; 行人跟踪则是在视频图像序列中检测定位出行人, 包括目标建模、目

标匹配和目标定位。行人跟踪算法分属两大类: 确定性跟踪和随机跟踪。确定性跟踪以 Mean Shift (MS) 跟踪为主线, 是以相似度量作为代价函数的优化问题, 有一系列基于 MS 的跟踪算法<sup>[2-9]</sup>; 随机跟踪将视觉跟踪转化为贝叶斯理论框架下的状态估计问题, 目前典型的随机跟踪算法有 Kalman 滤波跟踪<sup>[10]</sup>和粒子滤波跟踪<sup>[11]</sup>。以 MS 为代表的行人跟踪算法因其简单稳健的优势而广受研究者青睐。文

收稿日期: 2016-05-30. 网络出版日期: 2016-08-08.

基金项目: 国家自然科学基金项目(61273304); 上海市中医药三年行动计划重点项目(ZY3-CCCX-3-6002)

通信作者: 苗夺谦. E-mail: dqmiao@tongji.edu.cn.

献[1]对 MS 跟踪算法的经典文献和研究历程做了详细综述。2003 年 D.Comaniciu 等<sup>[2]</sup>将 MS 应用于视觉跟踪领域,首次提出 MS 跟踪算法,开辟了视觉跟踪领域的新方向,自此有大量关于 MS 的改进算法被提出。为提高 MS 跟踪算法的适应能力,在目标模板更新策略上,文献[3]提出基于线性加权的模板更新;文献[4]提出基于滤波的模板更新方法;在目标尺度及方向估计上,文献[5]提出基于特征点匹配的尺度与方向估计方法,消除由于尺度变化导致的定位误差;文献[6]提出基于权重图的尺度与方向估计方法,能更准确地反应目标尺度信息;在抗遮挡跟踪上,文献[7]在 MS 跟踪算法的基础上引入 Kalman 滤波器来辅助预测目标位置,增强了稳健性;文献[8]提出目标分块的方法,通过子块的跟踪定位来获得目标的整体位置;在跟踪快速目标上,MS 迭代容易陷入局部最优值,使得无法处理快速运动的目标,文献[9]引入模拟退火算法,提出退火 MS 来逐步平滑代价函数,用于全局模式的搜索,跳出局部最优值。

MS 跟踪向来以计算简单、调节参数少为优势,这些改进算法尽管从不同方面改进了 MS 跟踪算法的缺陷,但是也一定程度上增加了计算复杂度和调节参数,无法很好地保证跟踪实时性。MS 跟踪存在的挑战问题有许多,本文旨在基本保留 MS 的跟踪框架上解决 MS 跟踪的基本问题:颜色直方图带来的干扰问题和计算复杂度问题。考虑到颜色直方图对于目标形状、姿态、旋转等变化及遮挡都具有一定的鲁棒性,但是由于摄像机的视频图像质量普遍较差,分辨率普遍较低,易受背景颜色干扰,使得采集到的目标颜色特征存在较大的噪声冗余信息,然而粒计算方法恰巧能有效地处理视频图像大数据冗余信息。

粒计算方法由 Zadeh<sup>[12]</sup>和 Lin<sup>[13]</sup>首次提出,是人工智能领域中的一种新理念和新方法,它覆盖了所有和粒度相关的理论、方法和技术,主要用于对不确定、不精确、不完整信息的处理,对大规模海量数据的挖掘和对复杂问题的求解<sup>[14]</sup>。粒计算的实质是通过选择合适的粒度,来寻找一种较好的、近似的解决方案,避免复杂的计算,从而降低问题求解的难度。文献[15]归纳了粒计算方法在大数据处理中的 3 种基本模式并讨论了粒计算方法应用于大数据的可行性及优势。考虑到摄像机视频图像中的海量数据挖掘和行人跟踪问题的复杂计算,试图通过粒计算方法找到降低计算代价并有效改进跟踪效果的方法。

本文提出粒化的 MS 行人跟踪算法,基本保留 MS 跟踪算法框架,引入粒计算方法来优化颜色特征表示,在颜色特征的采集和处理上都采用粒计算方法表示,得到粒化后的核函数加权颜色直方图作为颜色特征表示,通过 MS 迭代实现行人跟踪,一定程度上削弱了图像噪声干扰,不仅降低了计算代价,而且提高了跟踪实时性。

## 1 颜色特征粒化

将大量的复杂特征及信息按其各自的特征及性能划分为若干个较简单的块,每一块看作是一个信息粒,这些信息粒具有相似的特征及性能,这个划分过程就是信息粒化的过程。

**定义 1** 用一个三元组来描述信息粒  $G = (IG, EG, FG)$ , 其中 IG 称为信息粒  $G$  的内涵, EG 称为信息粒  $G$  的外延; FG 称为内涵和外延之间的转换函数。

信息粒的内涵 IG 是信息粒在特定环境下的表现知识,表示在一个特定的任务下信息粒中所有元素的一般性特征、规则、共同性等,可以定义一个语言,由这个语言的公式来表示;信息粒的外延 EG 是满足这个公式的所有对象的集合;信息粒的转换函数 FG 是 EG 到 IG 的泛化转换函数。对于信息粒  $G = (IG, EG, FG)$ ,  $X$  为对于信息粒  $G$  所有可能外延 EG 的对象的集合,即外延的论域;  $Y$  为对于信息粒  $G$  所有可能内涵 IG 的集合,即内涵的论域。下面给出转换函数 FG 的形式化定义。

**定义 2** 从  $X$  到  $Y$  的信息粒的泛化转换函数是笛卡尔积  $X \times Y$  的一个子集 FG, 对于每个属于  $X$  的 EG 都存在唯一的  $Y$  中的元素 IG 使得  $(EG, IG)$  属于 FG。

传统的 MS 行人跟踪算法采集行人目标区域的颜色特征信息,粒化的 MS 行人跟踪算法引入粒计算的思想对颜色特征信息进行粒化,包含两个信息粒化过程:图像粒化和颜色通道粒化。

### 1.1 图像粒化

行人目标区域存在背景颜色干扰,为了减弱背景颜色干扰和降低计算代价,将图像行人目标区域进行不同粒度的粒层分块,并采集每一粒块的  $R$ 、 $G$ 、 $B$  颜色均值,得到图像特征信息粒格,简称图像粒,该过程为图像粒化。用三元组  $G' = (IG', EG', FG')$  表示图像粒,内涵  $IG'$  是由外延  $EG'$  集合中的每个对象通过转换关系  $FG'$  计算出的  $R$ 、 $G$ 、 $B$  颜色均值集合。外延  $EG'$  的论域  $X' = \{\{M \times N\} \mid M, N \in \mathbb{N}\}$ , 其中  $\{M \times N\}$  表示大小为  $M \times N$  的图像粒块,  $L_i^{R_i} \in X', i = 1, 2, \dots, n_{R_i}$ , 表示图像粒度为  $R_i$  时的第

$i$  个图像粒块,其中  $R_l = \{1, 2, \dots\}$  表示图像粒层分块的粒度。相同粒度下的图像粒块大小一致,令原始图像块大小为  $m \times n$ ,图像粒度为  $R_l$  时的图像粒块  $L_i^{R_l}$  的大小为  $m^{R_l} \times n^{R_l}$ ,  $n_{R_l}$  表示粒度为  $R_l$  时的图像粒块总数,有

$$n_{R_l} = \lfloor m \times n / m^{R_l} \times n^{R_l} \rfloor \quad (1)$$

内涵  $IG^l$  的论域  $Y^l = \{(C_R, C_G, C_B) \mid C_R, C_G, C_B \in (0, 255)\}$ , 其中  $(C_R, C_G, C_B)$  为  $R, G, B$  三通道的颜色特征向量,  $C_i^{R_l} \in Y^l, i = 1, 2, \dots, n_{R_l}$  表示图像粒度为  $R_l$  时第  $i$  个图像粒块的颜色特征均值。转换函数  $FG^l: X^l \rightarrow Y^l$  定义为

$$f_l(\lfloor M \times N \rfloor) = \frac{\sum_{[M \times N]} (C_R, C_G, C_B)}{M \times N} \quad (2)$$

转换函数对外延集合中的每个图像粒块求  $R, G, B$  三通道的颜色特征均值,得到图像粒的内涵集合,外延集合中的粒块对象和内涵集合中颜色特征向量对象一一对应,内涵集合中的所有颜色特征向量对象构成该图像目标区域的颜色特征矩阵。

图1为粒化的MS行人跟踪算法的图像粒化过程示意图。

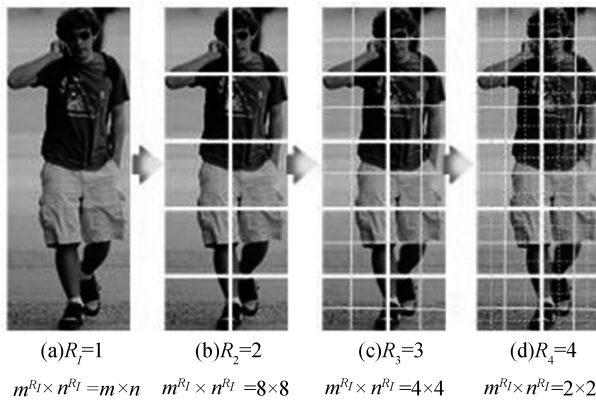


图1 图像粒化过程

Fig.1 Image granulation

$R_l$  为图像粒度,  $m^{R_l} \times n^{R_l}$  为图像粒度为  $R_l$  时的图像粒块的大小。当图像粒度  $R_l = 1$  时,图像粒块为行人目标区域  $m \times n$ ,图像粒的外延集合中只有一个对象,即  $EG^l = \{L_i^{R_l=1}, i = 1\}$ ,图像粒的内涵集合也只有一个图像粒块  $L_i^{R_l=1}$  的颜色特征均值,即  $IG^l = \{C_i^{R_l=1}, i = 1\}$ ;当图像粒度  $R_l = 2$  时,考虑到行人目标区域  $m \times n$  尺度较小且形状倾向于矩形,因此将  $m \times n$  的目标区域首次粒化成  $8 \times 8$  的图像粒块,更新外延集合为  $EG^l = \{L_i^{R_l=2}, i = 1, 2, \dots, n_2\}$  和内涵集合为  $IG^l = \{C_i^{R_l=2}, i = 1, 2, \dots, n_2\}$ , 其中  $n_2 = \lfloor m \times n / 8 \times 8 \rfloor$ ;当图像粒度  $R_l = 3$  时,对  $n_2$  个  $8 \times 8$  大小的图像粒块再粒化成  $n_3$  个  $4 \times 4$  大小的粒

块,同时更新外延集合和内涵集合;依此类推,当图像粒度  $R_l = 5$  时,粒化为  $n_5$  个  $1 \times 1$  大小的粒块,已经达到像素点级,粒化过程结束。实验表明图像粒度  $R_l = 4$  时,实验效果较好。

## 1.2 颜色通道粒化

行人跟踪的视频图像往往分辨率较低,存在一定程度的颜色噪声干扰。颜色级数越多,表示的颜色信息就越多,但是相应的颜色干扰信息也越多,且运算代价也越大,通过颜色特征计算相似性度量参数(巴氏系数)的速度也越慢,这对行人跟踪需要高实时性的要求具有很大挑战,因此不利于实时的行人跟踪。粒计算的思想能有效地对这种不确定的模糊信息进行简化处理,因此在颜色特征表示上引入粒计算的思想:行人目标区域经过图像粒化和转换函数处理后得到行人目标区域的颜色特征矩阵,对颜色特征矩阵做颜色通道粒化处理得到最终粒化后的颜色特征矩阵。

分别将  $R, G, B$  三通道颜色区间进行不同粒度的量化,每通道的颜色区间都均匀二分,得到一系列均匀的子颜色区间,称为颜色区间信息粒格,简称颜色粒,该过程为颜色通道粒化。用三元组  $G^c = (IG^c, EG^c, FG^c)$  来表示颜色粒,内涵  $IG^c$  是由外延  $EG^c$  集合中的每个颜色区间通过转换关系  $FG^c$  计算出的  $R, G, B$  颜色量值集合。外延  $EG^c$  的论域为  $X^c = \{[a, b], 0 \leq a, b \leq 255, a, b \in \mathbb{N}\}$ ,  $L_i^{R_c} \in X^c, i = 1, 2, \dots, n_{R_c}$  表示颜色粒度为  $R_c$  时第  $i$  个颜色粒区间,  $n_{R_c} = 2^{R_c-1}$  表示颜色粒度为  $R_c$  时的颜色粒总数,则颜色粒度为  $R_c$  时的每一个颜色粒区间长度为

$$\text{length}(L_i^{R_c}) = 256 / n_{R_c} = 2^8 / 2^{R_c-1} = 2^{9-R_c} \quad (3)$$

颜色粒度为  $R_c$  时第  $i$  个颜色粒区间  $L_i^{R_c}$  可表示为  $[\text{length}(L_i^{R_c}) \times (i-1), \text{length}(L_i^{R_c}) \times i]$ , 内涵  $IG^c$  的论域  $Y^c \in [0, 255]$ , 转换函数  $FG^c$  可定义为

$$f_c(L_i^{R_c}) = i \quad (4)$$

令  $v$  表示颜色粒区间  $L_i^{R_c}$  的一个颜色值,即  $v \in [\text{length}(L_i^{R_c}) \times (i-1), \text{length}(L_i^{R_c}) \times i]$ , 则转换函数可变换为

$$f_c(v) = v / \text{length}(L_i^{R_c}) = v \cdot 2^{R_c-9} \quad (5)$$

至此可直接通过转换函数式(5)计算出颜色特征值在任一粒度下粒化后的颜色特征值,较大程度降低计算量。

考虑到采集的原始颜色特征为  $R, G, B$  颜色值,有  $R$  通道、 $G$  通道和  $B$  通道。每个颜色通道对颜色敏感度不一样,如若对每个颜色通道采用一样的粒度来量化,则不仅没有降低噪声干扰,而且模糊了行人目标的颜色特征。因此对每一个颜色通道采用相应的

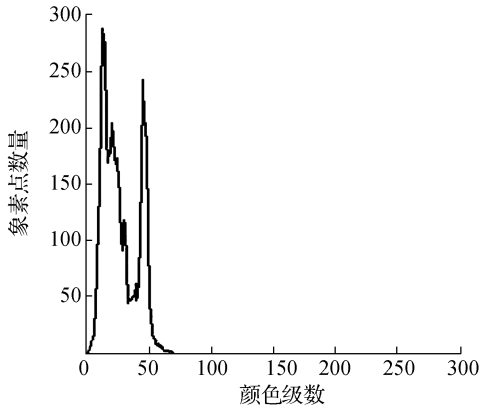
粒度来量化,经验得出  $G$  通道最敏感,  $R$  通道次之,  $B$  通道敏感度最低。因此  $G$  通道要表现更多的颜色特征信息,颜色粒度则赋予最大;而敏感度相对最弱的  $B$  通道表现最少的颜色特征信息,颜色粒度则赋予最小。令  $R, G, B$  通道的量化粒度分别表示为  $R_C^R, R_C^G, R_C^B$ , 经验得出当  $R_C^R=5, R_C^G=6, R_C^B=3$  时表现的颜色特征信息最为有效,行人跟踪效果较好。

计算经过颜色通道粒化后的 3 个颜色通道的直方图虽然已经降低了一定的计算代价和存储代价,但是为了更进一步简化计算量和降低存储代价,将颜色粒化后的三通道颜色值整合成一维颜色值作为该颜色粒的颜色特征值,由于颜色的权重信息已经在颜色粒化的时候通过颜色粒度表现,因此整体的转换函数  $f(v)$  直接体现为各颜色通道的转换函数的叠加,且保证  $f(v) \geq 1$ , 表达为

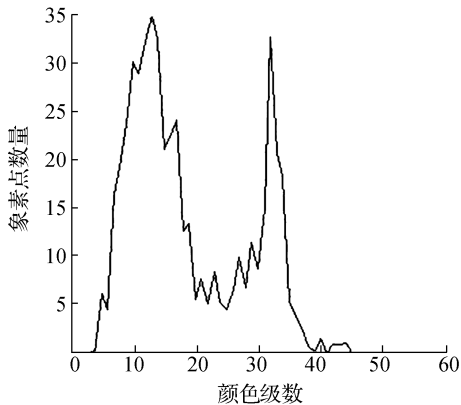
$$f_C(v) = f_C(v_R) + f_C(v_G) + f_C(v_B) = v_R \cdot 2^{R_C^R-9} + v_G \cdot 2^{R_C^G-9} + v_B \cdot 2^{R_C^B-9} + 1 \quad (6)$$

式中:  $v = (v_R, v_G, v_B)$  表示  $R, G, B$  颜色特征值向量,  $m_{RC} = n_{R_C^R} + n_{R_C^G} + n_{R_C^B} + 1$  表示最终得到的一维颜色特征值的级数。

未经颜色特征粒化(a)和经过颜色特征粒化后的颜色直方图(b)如图 2 所示。



(a) 颜色特征直方图



(b) 粒化的颜色特征直方图

图2 颜色直方图

Fig.2 Color histogram

比较两图,可看到粒化后的颜色直方图在颜色级数上降低了大约 5 倍,曲线与横轴所围成的面积是采集的目标区域的像素点数量,可见粒化后的像素点数量大约降低了 4 倍,大大降低了计算量。这一点将在后续的实验中得到进一步的验证。

## 2 粒化的 Mean Shift 行人跟踪算法

粒化的 Mean Shift 行人跟踪算法(GMS)是利用图像粒化及颜色通道粒化后的核函数空间加权的颜色直方图作为目标模型,采用巴氏系数作为相似度量,通过自适应步长的 MS 迭代来实时跟踪行人目标的位置。

### 2.1 算法描述

设  $\{L_i^{R_l}, i = 1, 2, \dots, n_{R_l}\}$  表示中心位置为  $y$  的行人目标区域经过图像粒度  $R_l$  粒化后的图像粒,  $x_i^*$  表示每个图像粒  $L_i^{R_l}$  的中心位置,则经过图像粒化后的加权核函数为

$$K(X) = \begin{cases} 1 - X, & X \leq 1 \\ 0, & X > 1 \end{cases} \quad X = \left\| \frac{y - x_i^*}{h} \right\|^2 \quad (7)$$

式中:  $h$  表示行人目标区域核半径。 $\{x_{i,j}, j = 1, 2, \dots, m^{R_l} \times n^{R_l}\}$  表示图像粒  $L_i^{R_l}$  的像素位置,  $b_f(x_{i,j}) = [b_f^R(x_{i,j}) \ b_f^G(x_{i,j}) \ b_f^B(x_{i,j})]$  表示像素  $x_{i,j}$  的  $R, G, B$  颜色特征值向量,其中  $b_f^R(x_{i,j}), b_f^G(x_{i,j}), b_f^B(x_{i,j})$  分别表示  $R, G, B$  三颜色通道的特征值,则每个图像粒  $L_i^{R_l}$  的颜色特征值向量通过转换函数式(2)得到

$$f_l(L_i^{R_l}) = \frac{\sum_{j=1}^{m^{R_l} \times n^{R_l}} b_f(x_j)}{m^{R_l} \times n^{R_l}} \quad (8)$$

通过式(6)和式(8),求得行人目标区域经过图像粒化和颜色通道粒化后的参考颜色直方图  $q = \{q_u\}_{u=1,2,\dots,m_{RC}}$  为

$$q_u = C \sum_{i=1}^{n_{R_l}} K\left(\left\| \frac{y - x_i^*}{h} \right\|^2\right) \delta[f_C(f_l(L_i^{R_l})) - u] \quad (9)$$

式中:  $\delta$  为 Kronecker 函数;  $C$  为归一化常数,使得  $\sum_{u=1}^{m_{RC}} q_u = 1$ , 因此有

$$C = \frac{1}{\sum_{i=1}^{n_{R_l}} K\left(\left\| \frac{y_0 - x_i^*}{h} \right\|^2\right)} \quad (10)$$

同理,设在跟踪过程中行人目标候选区域的颜色直方图经过图像粒化和颜色通道粒化后表示为  $p(y_0) = \{p_u(y_0)\}_{u=1,2,\dots,m_{RC}}$ , 其中  $y_1$  为目标候选区域的中心位置,同样有

$$p_u(\mathbf{y}_0) = C \sum_{i=1}^{n_{R_l}} K\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i^*}{\mathbf{h}}\right\|^2\right) \delta[f_c(f_l(L_i^{R_l})) - u]$$

(11)

基于巴氏系数,求得目标区域的参考颜色直方图  $q$  和目标候选区域的颜色直方图  $p(\mathbf{y}_0)$  的相似度为

$$\rho[p(\mathbf{y}_0), q] = \sum_{u=1}^{m_{R_C}} \sqrt{p_u(\mathbf{y}_0) q_u}$$

(12)

至此,粒化的 MS 行人跟踪算法的 MS 跟踪迭代步骤可以总结为算法 1。

**算法 1** 粒化的 Mean Shift 行人跟踪算法 (GMS)

**输入** 行人跟踪视频图像序列。

**初始化** 图像粒度  $R_l = 3$  和颜色通道粒度  $R_C^R = 5, R_C^G = 6, R_C^B = 3$ ; 给定行人目标初始中心位置  $\mathbf{y}_0$  及行人目标区域核半径  $h$ , 由式 (9) 得到行人目标区域经过粒化后的参考颜色直方图  $q$ 。

**跟踪过程**

1) 以  $\mathbf{y}_0$  为行人目标初始中心位置, 分别由式 (11) 和式 (12) 计算  $\{p_u(\mathbf{y}_0)\}_{u=1,2,\dots,m_{R_C}}$  和

$$\rho[p(\mathbf{y}_0), q] = \sum_{u=1}^{m_{R_C}} \sqrt{p_u(\mathbf{y}_0) q_u}。$$

2) 计算权值  $\{w_i\}_{i=1,2,\dots,n_{R_l}}$ 。

$$w_i = \sum_{u=1}^{m_{R_C}} \delta[f_c(f_l(L_i^{R_l})) - u] \sqrt{\frac{q_u}{p_u(\mathbf{y}_0)}}$$

3) 计算行人目标新位置  $\mathbf{y}_1$ :

$$\mathbf{y}_1 = \frac{\sum_{i=1}^{n_{R_l}} x_i^* w_i g\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i^*}{\mathbf{h}}\right\|^2\right)}{\sum_{i=1}^{n_{R_l}} g\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i^*}{\mathbf{h}}\right\|^2\right)}$$

4) 计算新位置的颜色直方图  $\{p_u(\mathbf{y}_1)\}_{u=1,2,\dots,m_{R_C}}$ 。

5) 由式 (12) 计算新位置和行人目标的相似度

$$\rho[p(\mathbf{y}_1), q] = \sum_{u=1}^{m_{R_C}} \sqrt{p_u(\mathbf{y}_1) q_u}。$$

6) 如果  $\rho[p(\mathbf{y}_1), q] < \rho[p(\mathbf{y}_0), q]$ , 则令  $\mathbf{y}_1 \leftarrow \frac{1}{2}(\mathbf{y}_0 + \mathbf{y}_1)$ 。

7) 若  $\|\mathbf{y}_1 - \mathbf{y}_0\| < \varepsilon$ , 则本轮迭代结束; 否则  $\mathbf{y}_0 \leftarrow \mathbf{y}_1$  返回 1) 继续迭代更新行人目标位置。

**输出** 当前帧行人目标区域实时最优位置。

2.2 算法计算复杂度

GMS 算法的跟踪迭代框架基本维持 MS 算法的跟踪框架, 因此 GMS 算法和 MS 算法的计算代价主要区别在每次迭代的计算代价, 表现在颜色直方图  $p$ 、新位置  $\mathbf{y}_1$  和相似度  $\rho$  的计算代价上, MS 算法中表示为  $C_H$ 、 $C_y$ 、 $C_\rho$ , GMS 算法中表示为  $C'_H$ 、 $C'_y$ 、 $C'_\rho$ , 如表 1。

表 1 MS 和 GMS 的运算代价

Table 1 The calculation cost of MS and GCMS

代价	MS	GMS
直方图	$C_H = RC_w +$	$C'_H = R'C_w +$
	$R(C_p + 3C_m + C_a)$	$R'(C_p + 3C_r + C_a)$
新位置	$C_y = u(C_w + C_d) +$	$C'_y = u'(C_w +$
	$R(3C_m + C_a)$	$C_d) + R'(3C_m + C_a)$
相似度	$C_\rho =$	$C'_\rho =$
	$u(C_s + C_m + C_a)$	$u'(C_s + C_m + C_a)$

表中  $C_w$ 、 $C_p$ 、 $C_m$ 、 $C_a$ 、 $C_r$ 、 $C_d$ 、 $C_s$  分别表示核函数加权、获取像素值、浮点数乘法、加法、移位运算、除法运算和开方的计算代价;  $R$ 、 $u$  分别表示 MS 算法中的行人目标区域大小及颜色直方图的级数;  $R'$ 、 $u'$  分别表示 GMS 算法中的行人目标区域大小及颜色直方图的级数; 满足近似关系  $R/R' = 4$ ,  $u/u' \approx 5$ ,  $C_m$  和  $C_r$  的关系无法准确比较, 但是  $C_m > C_r$  一定成立。至此可近似估计出 MS 和 GMS 的每次迭代的计算代价关系, 为式 (13)。理论计算得出, 仅考虑各运算代价时, GMS 算法比 MS 算法的计算代价至少快 4 倍, 是个可观的结果。在实际实验中, 由于图像分块函数消耗、硬件等原因会使得真正计算效率低于 4 倍。

$$\frac{C_H + C_y + C_\rho}{C'_H + C'_y + C'_\rho} \approx 4$$

(13)

3 实验结果与分析

在 CPU2.6 GHz、内存 4 GB 的 PC 机、MATLAB (R2013b) 的环境下进行实验, 实现本文的改进算法 GMS 跟踪算法, 并实现传统 MS 跟踪算法、Kalman 滤波跟踪算法和粒子滤波跟踪算法与之比较。

采用 4 组视频图像序列进行实验, 图像序列选自 PETS2009 和 CAVIAR 视频图像库, 涵盖了行人跟踪的主要场景及挑战场景, 包括简单背景、复杂背景、目标遮挡、形变及光照变化等。为检测本文方法的性能, 将与传统 MS 跟踪算法、Kalman 滤波跟踪算法和粒子滤波跟踪算法进行比较分析。

3.1 与传统 MS 跟踪的比较

图 3 选取的视频图像是 S2\_L1 项目中 View\_005 中的部分序列, 序列中红衣女子从视频窗口右端水平穿过走向视频窗口左端, 期间活动窗口只有红衣女子一人, 跟踪图像背景较简单。其中图 3(a) 是 MS 的跟踪结果, 图 3(b) 是 GMS 的跟踪结果。从跟踪结果可发现该序列 MS 和 GMS 的跟踪效果都较好, 在整个行人跟踪过程中都有正确跟踪到目标行人, 但是观察表 2 发现该序列 MS 算法的平均迭

代次数接近 5,而 GMS 算法的平均迭代次数 3.5 左右,跟踪实时性提高了 30%。可见在跟踪背景比较简单,行人目标颜色特征较为明显时,虽然 MS 算法和 GMS 算法的跟踪准确率相差无几,但是 GMS 算法由于引入了粒计算,降低了计算复杂度,因此一定程度上提高了跟踪实时性。



(a) MS 跟踪结果



(b) GMS 跟踪结果

图 3 Image I 跟踪结果比较

Fig.3 Comparison of tracking results of image I

图 4 选取的视频图像是 S2\_L1 项目中 View\_005 中的部分序列,序列中蓝衣女子从视频窗口右端水平穿过走向视频窗口左端,走到视频窗口中央时与两名黑衣男子发生遮挡随后分离。由于有行人遮挡及背景颜色干扰,跟踪图像背景略复杂。其中图 4(a)是 MS 的跟踪结果,图 4(b)是 GMS 的跟踪结果。比较图 4 的跟踪结果,发现目标行人未发生碰撞前,MS 和 GMS 都有正确跟踪到行人,在第 144 帧目标行人与两个行人发生碰撞后(即遮挡后分离),在第 147 帧及第 148 帧 MS 目标跟踪丢失,发现跟踪到干扰行人的蓝色牛仔褲上,与目标行人的颜色特征相似度较高,导致误判,直至最后目标彻底丢失。而在整个行人跟踪过程中,GMS 跟踪目标未丢失,跟踪正确率较高,仅在第 144 帧行人发生遮挡时跟踪略有偏差,遮挡分离后,跟踪回复正常。可见当跟踪场景中存在与行人目标的颜色特征相似的候选目标与行人目标碰撞时,MS 算法跟踪失效,GMS 算法的粒化后颜色直方图一定程度上降低了背景颜

色特征的干扰,使得跟踪正确保持。该序列 MS 的平均迭代次数接近 6,GMS 的平均迭代次数约为 4.5,跟踪实时性提高了 25%,但是相比序列一的 GMS 平均迭代次数,每帧图片平均增加了一次迭代,这是因为 Image II 跟踪场景相比 Image I 更为复杂。



(a) MS 跟踪结果



(b) GMS 跟踪结果

图 4 Image II 跟踪结果比较

Fig.4 Comparison of tracking results of image II

图 5 选取的视频图像是 S2\_L1 项目中 View\_001 中的部分序列,序列中红衣女子从视频窗口右下端斜向上走向视频窗口左上端,走到视频窗口中央时与两名黑衣男子发生遮挡随后分离,紧接着又被电线杆遮挡后又分离。跟踪图像背景略复杂,发生多次遮挡,包括两行人的动态遮挡和电线杆的静态遮挡,而且背景局部区域存在光照变化干扰。其中图 5(a)是 MS 的跟踪结果,图 5(b)是 GMS 的跟踪结果。MS 跟踪在第 467 帧时,行人跟踪丢失,跟踪结束;GMS 在第 467 帧时,行人跟踪略有偏差,但是仍继续正确跟踪。整个跟踪过程中,行人目标的颜色特征与背景颜色特征和遮挡物体的颜色特征能很好地区分,GMS 在目标行人发生遮挡时,能适当减弱背景及遮挡物体的颜色干扰,跟踪效果很好,而 MS 算法更多地受背景及遮挡物体的颜色干扰导致多次遮挡之后行人彻底丢失。该序列 MS 的平均迭代次数约为 5.5,GMS 的平均迭代次数约为 4.5,跟踪实时性仅提高了 18%,可见 GMS 算法在跟踪场景复杂的情况下,跟踪效果还有待提升。

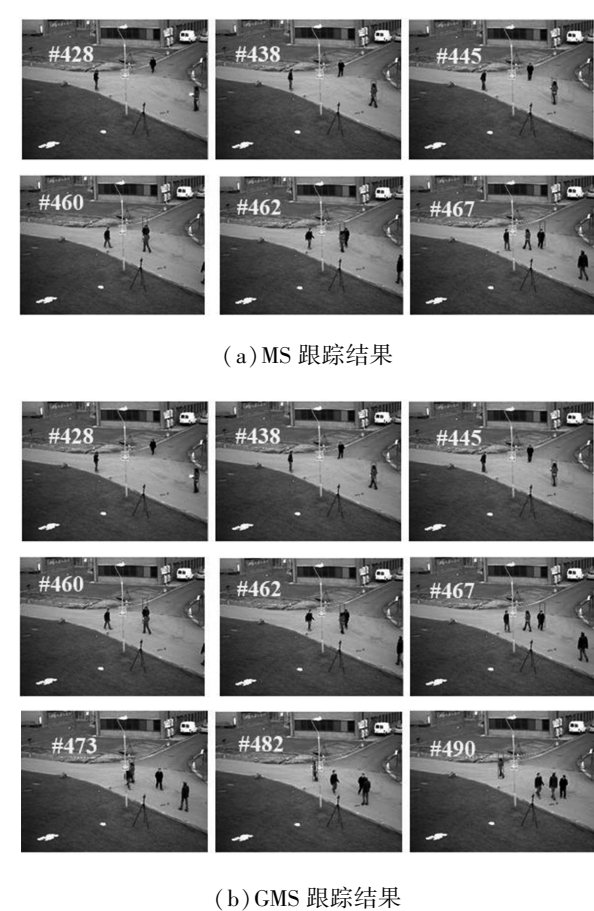


图 5 Image III 跟踪结果比较

Fig.5 Comparison of tracking results of image III

表 2 MS 和 GMS 的平均迭代次数(次/帧)

Table 2 The average iterations of MS and GMS

图像	MS	GMS
Image I	4.978 7	3.659 6
Image II	5.958 3	4.541 7
Image III	5.583 3	4.625 0
Image IV	5.238 1	3.904 8

图 6 选取的视频图像是 CAVIAR 数据库 FRONT VIEW 中的部分序列,序列中目标行人从视频窗口右端水平穿过走向视频窗口左端,走过视频窗口中央时与多名行人发生遮挡随后分离。整个视频图像质量较差,行人特征极其不明显,光照变化很大,加之有行人遮挡及背景颜色干扰,跟踪图像背景更为复杂。其中图 6(a)是 MS 的跟踪结果,图 6(b)是 GMS 的跟踪结果。观察(a)、(b)的跟踪结果,在第 493 帧发现 MS 目标跟踪片刻丢失,GMS 目标跟踪仅仅略有偏差,这是因为 MS 需要获取更多的颜色信息,相对应的对光照变化会更加敏感;第 549、564、573 帧捕捉了目标行人碰撞过程,由于行人特征实在不明显使得 MS 跟踪失效,而 GMS 能适当地容忍颜色特征干扰,

使得碰撞过程行人跟踪正常。可见当跟踪视频质量较差时 GMS 比 MS 更具有优势。

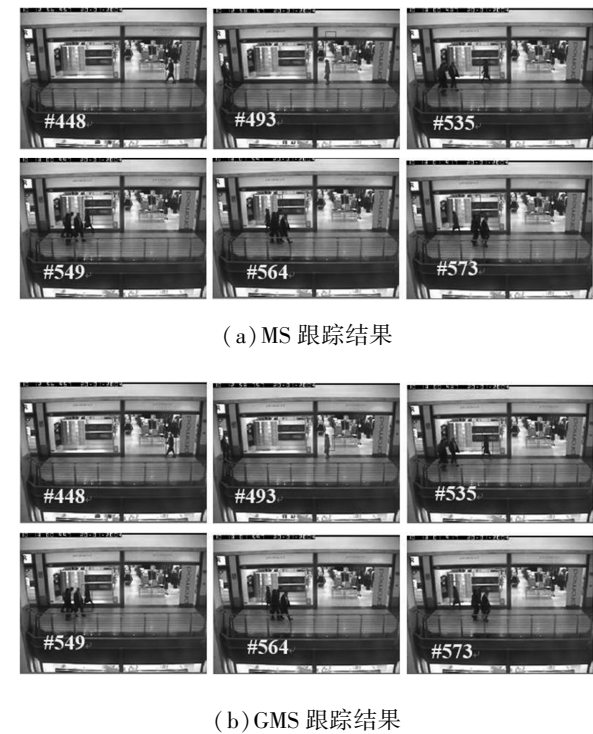


图 6 Image IV 跟踪结果比较

Fig.6 Comparison of tracking results of image IV

3.2 与其他类跟踪方法的比较

观察 Image I 红衣女子的跟踪结果(图 7),发现针对简单背景的行人跟踪,Kalman 滤波和粒子滤波总体上都能较好地跟踪到目标行人,不会丢失。该序列光照基本无变化,目标行人运动基本匀速,两个算法在第 477 帧跟踪位置都有偏差,但是随后跟踪正常,这是因为 Kalman 滤波预测初始参数的原因,随着跟踪状态的变化,参数会随时调整,这是不可避免的;粒子滤波的粒子采样初始是无经验无知识的,粒子覆盖整个跟踪图像,伴随着不断的重采样,粒子会渐渐地集中在行人目标上,跟踪位置也不断精确。

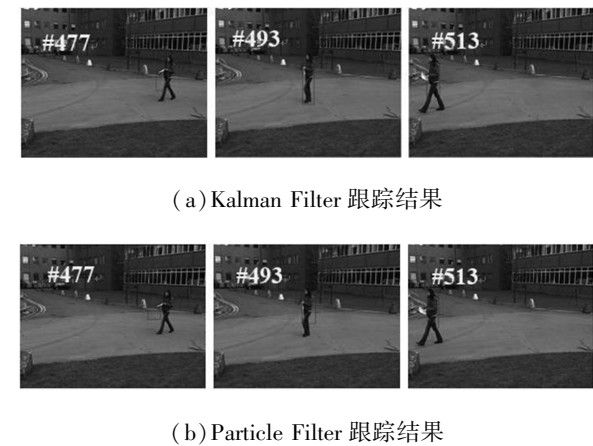


图 7 Image I 跟踪结果比较

Fig.7 Comparison of Tracking Results of Image I

观察 Image III 红衣女子的跟踪结果(图 8),发现复杂背景下且存在行人互遮挡、光照变化等挑战因素时,Kalman 滤波和粒子滤波跟踪都失效,行人目标极易丢失,这是因为传统 Kalman 滤波和粒子滤波跟踪都是利用状态估计来预测行人目标的位置,而光照变化等因素极易干扰粒子采样,可见 MS 跟踪采取颜色直方图作为跟踪特征具有一定的优势。



(a) Kalman Filter 跟踪结果



(b) Particle Filter 跟踪结果

图 8 Image III 跟踪结果比较

Fig.8 Comparison of Tracking Results of Image III

3.3 跟踪效率分析

表 3 展示了不同方法在不同图像序列中的平均跟踪效率(s/帧),发现 GMS 的跟踪效率均比 MS 的跟踪效率高,说明 GMS 跟踪算法在实际跟踪场景下确实降低了运算代价,提高了跟踪实时性。同时发现相比简单场景下,Kalman 滤波和粒子滤波的跟踪效率更胜 GMS。观察序列三各方法的跟踪效率发现 Kalman 滤波和粒子滤波结果很小,这是因为 Kalman 滤波和粒子滤波跟踪目标已经彻底丢失,所以该序列数据失去参考价值。

表 3 不同跟踪方法的平均跟踪效率比较

Table 3 Comparison of the average tracking efficiency of different tracking methods

图像	MS	Kalman Filter	Particle Filter	GMS
Image I	1.011 5	0.657 0	0.397 2	0.776 8
Image II	0.498 2	0.538 7	0.354 8	0.326 5
Image III	0.976 6	0.684 1	0.400 4	0.799 0
Image IV	0.373 8	0.312 3	0.359 9	0.343 0

4 结论

在跟踪场景中若存在与行人目标颜色特征相似的干扰物体时,特别是碰撞时或被遮挡时,传统的 MS 行人跟踪算法会使得跟踪目标丢失。因此,本文提出了粒化的 MS 行人跟踪算法,总结如下:

1)采用粒计算的思想对行人目标区域进行粒

化分块,并且对分块后的块颜色特征进行颜色通道粒化,归一化后得到粒化后的颜色直方图作为跟踪过程的行人目标颜色特征表示;

2)通过巴氏系数衡量行人目标区域和候选区域的相似度,最后通过 MS 迭代框架来实时跟踪目标行人的位置;

3)算法保持了传统 MS 跟踪算法的基本框架,因此也延续了传统 MS 跟踪算法的优势,且使得算法的计算复杂度更低,跟踪更快;

4)通过实验分析得出,本算法减弱了跟踪场景的背景颜色干扰,同时提高了跟踪实时性。

参考文献:

[1]顾幸方,茅耀斌,李秋洁. 基于 Mean Shift 的视觉目标跟踪算法综述[J]. 计算机科学, 2012, 39(12): 16-24.  
GU Xingfang, MAO Yaobin, LI Qiujie. Survey on visual tracking algorithms based on mean shift[J]. Computer science, 2012, 39(12): 16-24.

[2]COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE transactions on pattern analysis and machine intelligence, 2003, 25(5): 564-577.

[3]KHAN Z H, GU Yuhua, BACKHOUSE A G. Robust visual object tracking using multi-mode anisotropic mean shift and particle filters[J]. IEEE transactions on circuits and systems for video technology, 2011, 21(1): 74-87.

[4]PENG Ningsong, YANG Jie, LIU Zhi. Mean shift blob tracking with kernel histogram filtering and hypothesis testing[J]. Pattern recognition letters, 2005, 26(5): 605-614.

[5]CHEN Aihua, ZHU Ming, WANG Yanhua, et al. Mean shift tracking combining SIFT[C]//Proceedings of the 9th International Conference on Signal Processing. Beijing: IEEE, 2008: 1532-1535.

[6]NING Jifeng, ZHANG Lei, ZHANG D, et al. Scale and orientation adaptive mean shift tracking[J]. IET computer vision, 2012, 6(1): 52-61.

[7]CHEN Kangli, GE Wancheng. Pedestrian tracking algorithm based on kalman filter and partial mean-shift tracking[C]//Proceedings of the 2nd International Conference on Systems and Informatics (ICSAI). Shanghai: IEEE, 2014: 230-235.

[8]FANG Jiangxiong, YANG Jie, LIU Huaxiang. Efficient and robust fragments-based multiple kernels tracking[J]. AEU-international journal of electronics and communications, 2011, 65(11): 915-923.

[9]SHEN Chunhua, BROOKS M J, VAN DEN HENGEL A. Fast global kernel density mode seeking: applications to lo-

calization and tracking[J]. IEEE transactions on image processing, 2007, 16(5): 1457-1469.

[10] CHANG S. Stochastic peak tracking and the Kalman filter [J]. IEEE transactions on automatic control, 1968, 13(6): 750.

[11] DOUCET A, GORDON N. Efficient particle filters for tracking manoeuvring targets in clutter[C]//Proceedings of IEE Colloquium on Target Tracking: Algorithms and Applications. London: IEEE, 1999: 4/1-4/5.

[12] ZADEH L A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic [J]. Fuzzy sets and systems, 1997, 90(2): 111-127.

[13] LIN T Y. Granular computing: fuzzy logic and rough sets [M]//ZADEH L A, KACPRZYK J. Computing with Words in Information/Intelligent Systems 1. Berlin Heidelberg: Physica-Verlag HD, 1999: 183-200.

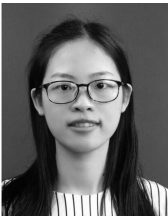
[14] 苗夺谦, 王国胤, 刘清, 等. 粒计算: 过去、现在与展望 [M]. 北京: 科学出版社, 2007.

MIAO Duoqian, WANG Guoyin, LIU Qing, et al. Granular computing: past, present and future[M]. Beijing: Science Press, 2007.

[15] 徐计, 王国胤, 于洪. 基于粒计算的大数据处理[J]. 计算机学报, 2015, 38(8): 1497-1517.

XU Ji, WANG Guoyin, YU Hong. Review of big data processing based on granular computing[J]. Chinese journal of computers, 2015, 38(8): 1497-1517.

作者简介:



刘翠君,女,1993 年生,硕士研究生,主要研究方向为计算机视觉、粒计算等。



赵才荣,男,1981 年生,副研究员,博士,主要研究方向为人脸识别、计算机视觉等。



苗夺谦,男,1964 年生,教授、博士生导师,博士,主要研究方向为粒计算、粗糙集、中文信息处理等。

第 34 届中国控制会议

The 34th Chinese Control Conferencel