非重叠监控视频中基于颜色空间分布模型的行人匹配方法

李 博[a], 李国辉[a,b], 涂 丹[a]
(国防科学技术大学 a. 信息系统与管理学院 系统工程系; b. 信息系统工程重点实验室, 长沙 410073)

摘 要: 针对非重叠视域监控系统中行人目标关联问题,提出了一种基于颜色空间分布模型的行人目标匹配方法。使用最邻近聚类方法对目标的像素进行聚类,得到目标的 MC 特征;建立目标的颜色空间分布模型,并进行特征转换;在建立的分布模型的基础上,计算目标相似度并进行匹配。实验结果表明,该算法能有效地去除目标检测与分割产生的边缘背景像素的影响,在对相同目标保持较高匹配率的情况下,能较好地处理由于目标颜色分布差异而造成的误匹配情况,同时对于摄像机视角的差异以及目标姿态的变化有较好的鲁棒性。

关键词: 行人匹配; 非重叠视域; 颜色空间分布模型; 主颜色谱直方图
中图分类号: TP391.4 **文献标志码**: A **文章编号**: 1001-3695(2012)02-0797-04
doi:10.3969/j.issn.1001-3695.2012.02.107

Human matching approach based on color spatial distribution model across non-overlapping camera views

LI Bo[a], LI Guo-hui[a,b], TU Dan[a]
(*a. Dept. of System Engineering, School of Information System & Management, b. Key Laboratory of Information System Engineering, National University of Defense Technology, Changsha* 410073, *China*)

Abstract: For the human matching requirements, this paper presented a special approach of human matching based on color spatial distribution model across non-overlapping camera views. First, it obtained the MC feature by clustering the target pixels through the nearest cluster method. Secondly, built the color spatial distribution model and transformed the MC feature. Finally, calculated the target pair similarity based on the built color spatial distribution model and got the matching result. Experimental results show that the presented approach can effectively eliminate the influence of the edge pixels which come with the detection and the segmentation. While it maintained a high matching rate to the same target in different cameras, it is capable of coping with the wrong matching by the difference of the color distribution. The results also show that the proposed approach is robust to the difference of the visual angle and the change of the object gesture.

Key words: human matching; non-overlapping camera views; color spatial distribution model; MCSHR

0 引言

近年来,视频监控系统已经在各个行业得到了广泛的应用。由于安全的需要,监控系统的规模变得越来越庞大,在监控系统上进行目标跟踪与行为分析的智能监控系统研究已经成为近年来的一个热点问题。每个摄像机有一个视域(field of view),由于大多数监控系统中摄像机的视域是非重叠的(non-overlapping views),因此研究在非重叠监控系统中目标的跟踪问题变得越来越重要。这其中最重要的步骤就是非重叠监控视频中的行人匹配技术。其主要目的在于对一个区域监控系统中的多个监控视频里的目标实现计算机的自动搜索,达到行人目标的实时跟踪。

目标匹配的核心问题是采用目标的特征对目标外观模型(object appearance model,OMA)进行描述,在现有的目标匹配研究中多采用颜色特征进行匹配。文献[1]介绍了一个智能监控系统中目标匹配的基本框架,包括目标提取、特征计算、特征数据库、特征匹配等部分。Javed 等人[2,3]根据目标的累积直方图建立了不同监控视频间的亮度转换函数(brightness transfer function,BTF),在此基础上对不同视频间的颜色分量进行转换,利用颜色转换结果进行目标相似度的计算,并使用 PCA 方法对高维空间的计算进行降维。Cheng 和 Piccardi 等人[4,5]建立了一种主颜色谱直方图方法(major colour spectrum histogram representation,MCSHR)来建立人物外观模型,并给出了针对这种模型的相似度计算方法,同时他们还进一步利用了目标局部直方图均衡化方法来消除光照强度的影响。Jeong 等人[6]提出了一种基于 UV 色度特征的方法。以上三种方法的相同之处在于利用了目标的颜色统计信息,从统计特征上对目标进行匹配,而较少利用颜色的空间信息,因此对于颜色统计信息相似、分布有差异的目标误匹配率较高。此外还有一些研究将 SIFT 特征也用于目标匹配的问题中[7]。虽然 SIFT 是一种鲁棒性好的尺度不变性描述方法,但是由于不同监控摄像机

收稿日期: 2011-05-29; **修回日期**: 2011-07-11

作者简介:李博(1987-),男,湖南邵阳人,硕士,主要研究方向为图像与视频智能处理、计算机视觉(libo_nudt@163.com);李国辉(1963-),男,湖南永州人,教授,博士,主要研究方向为图像与视频信息智能处理、数据挖掘;涂丹(1971-),男,湖南常德人,副教授,博士,主要研究方向为图像与视频智能处理、计算机视觉.

在视角上的差别以及目标在运动时姿态上的变化,使得其匹配效果并不理想。

本文基于传统的 MCSHR 匹配方法,提出了一种基于颜色空间分布模型的自适应主颜色相似度计算方法。其主要优点包括:a)提取目标主颜色(major color,MC)特征并根据建立的空间分布模型进行特征转换,能有效地处理目标颜色统计特征相似、分布存在差异的目标的误匹配情况;b)对提取的 MC 特征进行空间分布处理,以降低目标检测与分割产生的边缘背景像素对目标匹配造成的影响;c)相似度计算方法对于目标姿态、摄像机视角等变化有较好的鲁棒性。

1　基于颜色空间分布模型的目标匹配方法

传统的非重叠监控视频中目标匹配方法主要利用了目标的颜色统计信息,其他的利用分布信息的方法如 SIFT 等则对于目标姿态的变化比较敏感。为了处理由于颜色分布差异造成的行人目标误匹配问题,本文对大量监控视频中的行人目标进行了分析,发现使用颜色聚类得到的 MC 特征,在基于人体区域划分的基础上,具有一些明显的分布特征:a)与目标匹配相关的少数几种 MC 类的中心都位于目标的中心位置附近;b)由于检测和分割产生的干扰像素 MC 类的中心都位于目标的边缘;c)目标上身的分布特征较稳定并且对于目标有更好的识别性;d)在不同视角的监控视频中,目标垂直轴线上的分布差异与目标匹配更为相关。

基于以上事实,本文考虑以下的前提条件:a)目标为直立的行人目标;b)目标外观较完整,没有遮挡等情况;c)目标在外观上没有显著的变化。在此前提下将目标分为头部、上身、下身三个部分,在目标的上身主颜色聚类的基础上进行融合颜色分布信息的 MC 特征转换,在目标的下身则进行基于分布模型的 MC 权重归一化处理,这样既能去除边缘背景像素的影响,又能较好地建立目标识别所需的颜色分布特征。目标匹配的流程为:首先进行目标检测以及预处理,提取主颜色特征,建立颜色空间分布模型,进行颜色匹配特征处理,最后进行相似度计算得到目标匹配结果。

1.1　提取区域 MC 特征

在 MCSHR 方法中,Piccardi 等人[4,5]使用了 K-means 聚类方法对检测到的目标像素进行聚类。为了降低目标边缘背景像素对匹配的影响以及得到能够使用分布模型进行处理的 MC 特征,本文采用最近邻聚类方法,其优点在于对聚类结果的类别数目没有限制,通过设定合适的类间距离阈值,可以获得最能表现目标外观模型的 MC 特征。

提取 MC 特征前需要进行目标检测,本文的目标检测使用文献[8]的减背景方法。根据文献[9]的方法将目标分为头部、上身和下身部分。对后两个部分在 RGB 颜色空间中进行目标像素聚类,使用 RGB 颜色空间的欧式距离:

$$d(C_1,C_2)=|C_1-C_2|=\sqrt{(R_1-R_2)^2+(G_1-G_2)^2+(B_1-B_2)^2} \tag{1}$$

其中,C_1 和 C_2 分别代表两个图像像素点在 RGB 颜色空间中的位置,R_i、G_i、B_i 分别代表该像素点的 RGB 值。假设目标 A 经过聚类得到 M 个 MC 类,表示为

$$\mathrm{MC}(A)=\{C_{A_1},C_{A_2},\cdots,C_{A_i},\cdots,C_{A_M}\} \tag{2}$$

其中,$C_{A_i}(i=1,2,\cdots,M)$是目标的 MC 类,每个类的像素频率可以表示为

$$p(A)=\{p(A_1),p(A_2),\cdots,p(A_i),\cdots,p(A_M)\} \tag{3}$$

1.2　MC 空间分布模型

为了对目标 MC 类的空间分布进行估计,建立以目标区块中心为原点的坐标系,如图 1 所示。根据聚类的结果,可以得到 MC 类的空间分布。

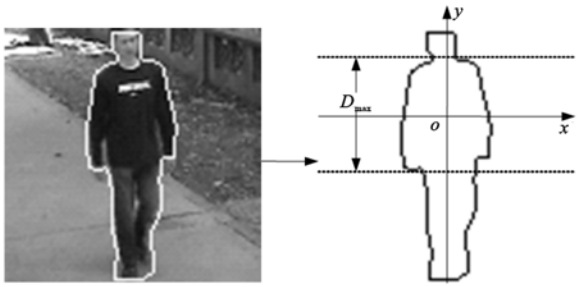

图1　目标颜色特征空间坐标系

聚类区域的高度定义为 D_{max},假设目标 A 的 M 个 MC 类的空间位置可以表示为

$$O(A)=\{O_{A_1},O_{A_2},\cdots,O_{A_M}\}=\{(x_{A_1},y_{A_1}),(x_{A_2},y_{A_2}),\cdots,(x_{A_M},y_{A_M})\} \tag{4}$$

定义坐标系中心到边缘的最长距离为 L_{max},计算每个中心点到坐标原点 $O(x_{A_0},y_{A_0})$的距离 L_{A_i},计算式为

$$L_{A_i}=\sqrt{(x_{A_i}-x_{A_0})^2+(y_{A_i}-y_{A_0})^2} \tag{5}$$

定义每个颜色类的位置权重 l_i:

$$l_i=\frac{(L_{max}-L_{A_i})}{L_{max}} \tag{6}$$

将其归一化得到每个颜色类的相对权重系数:

$$V_{A_i}=\frac{l_i}{\sum_{i=1}^{M}l_i} \tag{7}$$

将主颜色 $p(A_i)$与空间分布相对权重系数 V_{A_i}相乘之后进行归一化处理,可以得到融合颜色类空间分布位置权重的主颜色特征 $p_{sp}(A_i)$:

$$p_{sp}(A_i)=\frac{p(A_i)\times V_{A_i}}{\sum_{i=1}^{M}p(A_i)\times V_{A_i}} \tag{8}$$

使用上文所述的颜色空间模型处理后的主颜色特征 $p_{sp}(A_i)$融合了目标的颜色空间分布信息,位于目标边缘的像素类对应权重相应地降低,因此能够较好地处理目标检测与分割所产生的边缘像素的干扰,同时能提高与目标匹配相关的 MC 特征在相似度计算时的匹配贡献率。

1.3　基于空间分布模型的相似度计算

在文献[3]中,通过计算累积主颜色的像素比例来计算目标相似度,这种方法对于统计信息相似度较高的目标有较好的匹配效果,但不能有效识别颜色的差异分布。为此本文在 1.2 节建立的分布模型基础上,使用一个转换函数对目标的 MC 特征进行转换。考虑到在通常的监控视频中由于摄像机在视角、高度上的差异以及目标姿态变化等因素,使得目标的 MC 特征在 x 轴上的分布差异较大,而在 y 轴上的分布却具有较好的鲁棒性,本文在 1.2 节中建立的空间分布模型上,提出了一种带惩罚因子的自适应权重转换方法。

定义两个目标中两个相似 MC 类的 y 轴矢量距离 $D_{(A_i,B_j)}$为

$$D_{(A_i,B_j)}=S_{A_i}(y_i)-S_{B_j}(y_j)=y_{A_i}-y_{B_j} \tag{9}$$

其中,$S_{A_i}(y_i)$为 C_{A_i}沿 y 轴的矢量,$S_{B_j}(y_j)$为 C_{B_j}沿 y 轴的矢量,定义 $D_{(A_i,B_j)}$的归一化距离为

$$d_{(A_i,B_j)}=\frac{D_{(A_i,B_j)}}{(D_{\max}^{[A]}+D_{\max}^{[B]})/2} \tag{10}$$

为了根据分布信息对目标的相似度计算进行相应的处理,根据前文的假设建立一个转换函数,将目标的归一化距离使用式(11)所示的转换函数进行转换。

$$T(d_{(A_i,B_j)})=2(|d_{(A_i,B_j)}|-1)^2-1 \quad 0\leqslant|d_{(A_i,B_j)}|\leqslant 1 \tag{11}$$

将转换函数应用到 1.2 节中得到的 MC 特征 $p_{sp}(A_i)$ 中,可以得到最终用于计算相似度的 MC 相似度特征:

$$p_{\text{norm}}(A_i)=\frac{T(D_{(A_i,B_j)})\times p_{sp}(A_i)}{\sum\limits_{i=1,2,\cdots,M}p_{sp}(A_i)} \tag{12}$$

同样地,可以得到目标 B 中 B_j 的 MC 特征:

$$p_{\text{norm}}^{[A_i]}(B_j)=\frac{T(d_{(A_i,b_j)})\times p_{sp}^{[A_i]}(B_j)}{\sum\limits_{j=1,2,\cdots,N}p_{sp}(B_j)} \tag{13}$$

根据文献[4]中 MCSHR 的相似度值计算方法,目标 A 中每个 MC 特征相似度值的计算为

$$\text{sim}(C_{A_i},C_{B_j|A_i})=\min\{p_{\text{norm}}(A_i),p_{\text{norm}}^{[A_i]}(B_j)\} \tag{14}$$

目标 A 到 B 的相似度使用 MC 特征相似度累加得到:

$$\text{sim}(A,B)=\sum_{i=1}^{M}\text{sim}(C_{A_i},C_{B_j|A_i}) \tag{15}$$

最后根据设定的相似度阈值 σ,可以得到不同监控视频中目标匹配的结果,具体实验的结果及分析将在下一部分中进行详细论述。

2　实验结果及分析

本文所提出的基于颜色空间分布模型的目标匹配算法在一个园区的非重叠多摄像头监控系统下进行了测试实验。实验的视频采集了六个不同场景的监控视频,如图 2 所示。该监控系统中所装的摄像机分辨率为 720 × 576 像素,帧率为 25 fps。测试算法运行于 Intel 2.40 GHz CPU 和 4.00 GB 内存的 64 位 PC 机上。

图2　非重叠视域监控视频

由于目标的外观特征受到光照、噪声等外部条件以及相对于摄像机的位置、姿态等因素的影响,其在外观上有一定的差异。实验采集了不同视域监控视频中的多个运动目标,部分实验图像如图 3 所示。实验中使用两个目标图像组成目标对进行计算,得到目标匹配的结果。

2.1　不同监控视频中相同目标的匹配

同一目标在不同监控视频中外观上的差异主要是受到外界光照条件等因素影响造成的。算法对于目标在外观差异上的容忍度是匹配的一个关键。为了测试同一目标在不同监控视频中的匹配结果,实验使用了六个不同场景中 20 组目标图像进行匹配,测试结果如表 1 所示。从表 1 中可以看出,对于相同的目标,由于 MC 特征相似,目标匹配的结果为 95.00%。而相似度数值上的差异,主要是由于外界环境光照条件影响造成的颜色差异引起的,通过设定合理的相似度阈值,可以使得数值差异在算法的容忍度范围之内。

图3　非重叠监控视频中的行人目标

表 1　不同监控视频中相同目标的匹配结果

目标编号	摄像机编号	上身相似度	结果	下身相似度	结果	目标匹配结果
1	1 2	0.9516	1(Yes)	0.9095	1(Yes)	1(Yes)
2	1 3	0.6023	1(Yes)	0.7589	1(Yes)	1(Yes)
3	2 3	0.6183	1(Yes)	0.6523	1(Yes)	1(Yes)
…	…	…	…	…	…	…
实验数目:20　正确匹配:19　错误匹配:1　匹配率:95.00%						

注:其中颜色距离阈值取 50,相似度阈值 σ = 0.5

2.2　不同监控视频中不同目标的匹配

不同监控视频中目标的误匹配是 MCSHR 算法存在的主要问题,也是本文要解决的问题之一。实验使用了六个场景下 20 组不同目标的图像进行匹配,结果如表 2 所示。从表 2 中可以看出,由于目标的上身颜色信息更丰富且具有更好的鲁棒性,经过空间分布模型处理之后的目标上身 MC 特征相比下身区域对于目标的识别更具有辨识性,对于不同目标的匹配能够实现较好的区分。

表 2　不同监控视频中不同目标的匹配结果

目标编号	摄像机编号	上身相似度	结果	下身相似度	结果	目标匹配结果
1 2	1 2	0.1893	0(No)	0.7841	1(Yes)	0(No)
2 3	1 3	0.1843	0(No)	0.4120	0(No)	0(No)
2 3	2 3	0.1923	0(No)	0.6523	1(Yes)	0(No)
…	…	…	…	…	…	…
实验数目:20　正确匹配:17　错误匹配:3　匹配率:85.00%						

注:其中颜色距离阈值取 50,相似度阈值取 σ = 0.5

2.3　与 MCSHR 方法的比较

传统的方法都是利用目标颜色的统计信息,对于颜色分布有差异的目标有较高的误检测,而且对于不同目标的匹配结果也不理想。为了与 MCSHR 方法进行比较,分别取了六个场景中共 20 组目标图像进行了实验,且只使用经过分布模型处理

过的上身的 MC 特征。图 4 所示是相同目标在不同场景中的相似度值曲线。由图中可以看出,对于相同的目标,MCSHR 方法计算出的相似度值较稳定,而经过分布模型处理的目标相似度值则扩大了目标的差异,通过选取合适的阈值,可以使得差异控制在可容忍的范围之内,同时更好地表现了目标之间的差异。

在不同目标的匹配方面,同样选取三个场景中的共 20 组目标图像进行实验,结果如图 5 所示。从图 5 中可以看出,MCSHR 特征在计算不同目标时,相似度值整体偏高,而经过分布模型处理之后的目标组,则由于扩大了不同目标组之间的差异,使得其相似度值更能表现不同目标之间的区别。

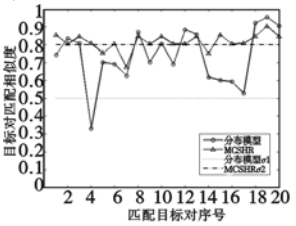

图4　不同监控视频中相同目标的匹配相似度值曲线

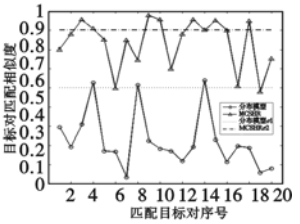

图5　不同监控视频中不同目标的匹配相似度值曲线

从以上三个部分的实验结果看来,本文所提出的颜色空间分布模型不仅对于颜色分布有差异的目标能进行较好的识别,同时还能去除检测的边缘效应,减少噪声干扰。与传统的方法相比,在保持了相同目标在不同场景的匹配准确率的同时,在不同目标的准确率上有了较大的提高。

3　结束语

本文提出了一种适用于非重叠视域监控系统中行人目标的匹配方法。首先,使用最近邻法对目标各个区域的像素进行聚类,得到能表征目标主要外观的 MC 特征,不仅能有效去除噪声的影响,而且对于光照等引起的目标外观变化因素有一定的容忍度。其次,通过建立的颜色空间分布模型对得到的 MC 特征进行转换,一方面能够突出主要表现目标的 MC 特征,同时能够降低由于检测和分割产生的目标边缘背景像素对 MC 特征的影响。最后,基于建立的颜色空间分布模型计算匹配目标对的相似度,在保持相同目标在不同场景中的匹配准确率的同时,降低了不同目标之间的误匹配。

目标匹配使用的外观特征主要是颜色的 MC 特征,这不可避免地会受到环境光照的影响。本算法通过颜色聚类和设定合理的相似度阈值,在一定程度上降低了光照的影响,但仍然有一些目标的特征变化较大,因此如何对目标环境光照条件进行估计并对目标进行颜色转换,仍然是一个值得研究的问题。

参考文献:

[1] REDDING N J, OHMER J F, KELLY J, *et al*. Cross-matching via feature matching for camera handover with non-overlapping fields of view[J]. Digital Image Computing: Techniques and Applications, 2008(38):343-350.

[2] JAVED O, SHAFIQUE K, RASHEED Z, *et al*. Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views[J]. Computer Vision and Image Understanding,2008(109):146-162.

[3] JAVED O, SHAFIQUE K, SHAH M. Appearance modeling for tracking in multiple non-overlapping cameras[C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005:26-34.

[4] MADDEN C, CHENG E D, PICCARDI M. Tracking people across disjoint camera views by an illumination-tolerant appearance representation[J]. Machine Vision and Applications,2007,18(3):233-247.

[5] CHENG E D, MADDEN C, PICCARDI M. Mitigating the effects of variable illumination for tracking across disjoint camera views [C]// Proc of IEEE International Conference on Video and Signal Based Surveillance. 2006:32-38.

[6] JEONG K, JAYNES C. Object matching in disjoint cameras using a color transfer approach [J]. Machine Vision and Applications, 2008(19):443-455.

[7] TEIXEIRA L F, CORTE-REAL L. Video object matching across multiple independent views using local descriptors and adaptive learning [J]. Pattern Recognition Letters,2009,30(2):157-167.

[8] CUCCHIARA R, PICCARDI M, PRATI A. Detecting moving objects, ghosts, and shadows in video streams[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2003,25(10): 1337-1342.

[9] SHASHUA A, GDALYAHU Y, HAYUN G. Pedestrian detection for driving assistance systems: single-frame classification and system level performance[C]//Proc of Intelligent Vehicles Symposium. 2004:1-6.

(上接第 796 页)

[3] 刘海宾, 何希勤, 刘向东. 基于分水岭和区域合并的图像分割算法[J]. 计算机应用研究,2007,24(9):307-308.

[4] ZHANG Y J. Advances in video and image segmentation[M]. New York: IRM Press, 2006:1-15.

[5] CAREVIC D, CAELLI T. Region based coding of color image using Karhunen-Loeve transform[J]. Graphics Models and Image Processing,1997,59(1):27-38.

[6] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms[M]. New York: Plenum Press,1981.

[7] COMANICIU D, MEER P. Mean-Shift: a robust approach toward feature space analysis[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,2002,24(5):603-619.

[8] FUKUNAGA K,HOSTETLER L D. The estimation of the gradient of a density function with applications in pattern recognition[J]. IEEE Trans on Information Theory,1975,21(1):32-40.

[9] COMANICIU D, MEER P. Mean-Shift analysis and applications [C]//Proc of IEEE International Conference on Computer Vision. 1999:1197-1203.

[10] COMANICIU D, RAMESH V, MEER P. The variable bandwidth Mean-Shift and data-driven scale selection[C]//Proc of IEEE International Conference on Computer Vision. 2001:438-445.

[11] TURLACH B. Bandwidth selection in kernel density estimation: a review[EB/OL]. (1993). http:// citeseer. ist. psu. edu/214125.

[12] 王永明,王贵锦. 图像局部不变性特征与描述[M]. 北京: 国防工业出版社, 2010.

[13] KOENDERINK J J. The structure of images[J]. Biological Cybernetics,1984,50(5):363-370.

[14] LINDEBERG T. Scale-space theory in computer vision[M]. [S. l.]: Kluwer Academic Publishers, 1994.

[15] FLORACK L M J, HAAR R B M, KOENDERINK J J, *et al*. Scale and the differential structure of images[J]. Image and Vision Computing,1992,10(6):376-388.

[16] SONKA M, HLAVAC V, BOYLE R. Image processing, analysis, and machine vision[M]. [S. l.]:Nelson Engineering,2003.