DOI: 10.13245/j.hust.15S1046

基于局部特征融合的目标再识别方法

袁文博[1] 陈尔奎[2] 曹志强[1] 王天柱[1] 谭 民[1]
(1 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190;
2 山东科技大学电气与自动化工程学院, 山东 青岛 266590)

摘要 提出一种基于局部特征融合的目标再识别方法.通过视觉传感网采集视频图像,针对场景内的目标,首先结合颜色和形体特征,对目标进行区域分割,进一步设计特征描述子,对关键区域的颜色和形体信息进行提取,基于此判断目标间的相似度,实现目标再识别.实验表明了所提方法的有效性.

关键词 目标识别; 区域分割; 局部特征; 再识别; 视觉传感网
中图分类号 TP24 **文献标志码** A **文章编号** 1671-4512(2015)S1-0192-04

Object re-identification method based on fusion of local features

Yuan Wenbo[1] *Chen Erkui*[2] *Cao Zhiqiang*[1] *Wang Tianzhu*[1] *Tan Min*[1]
(1 State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; 2 College of Electrical and Automation, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract An object re-identification method based on fusion of local features was proposed. Video image was obtained through visual sensor network and objects in scenes were concerned. Firstly, object was divided into multiple regions based on color feature and shape feature. Then, feature descriptors were designed to extract color and shape information of key regions. Finally, object re-identification was implemented by evaluating the similarity among the objects. The effectiveness of the proposed method was verified by experiments.

Key words objcect recognition; region segmentation; local feature; re-identification; visual sensor network

在无重叠视域的视觉传感器网络中,一个已识别的目标,在条件(包括监控场景、光照条件、目标位姿等)变化后再次将其识别的过程称为目标再识别.目标再识别技术在智能监控、多目标跟踪等领域有着重要研究意义,近年来得到了国内外学者的广泛关注与研究.

目前,目标再识别问题的研究方法主要分为两类,第一类方法是对目标的颜色、形态等外观特征直接进行统计和使用.文献[1]通过计算人体不同部分的空间协方差矩阵,获得局部空间之间的相关度,从而区分出不同的目标.直方图加缩影(histogram plus epitome, HPE)方法[2]通过融合人体局部和全局的颜色信息,对颜色特征有更好的描述.此外,先分割得到感兴趣区域,在感兴趣区域内提取如 SURF 和 SIFT 等特征点来完成再识别[3-4],是当前较热门的做法之一. Cai 将位置信息融入到 HSV 颜色直方图,提出了全局颜色背景(global color context, GCC)方法[5].文献[6]先将局部描述子转化为 Fisher 向量,然后再提取全局特征,进行目标再识别.第二类方法通过训练,找出有最大区分度的特征分量,再进行建模实现目标再识别. Zhao 等利用非监督学习方式来寻

收稿日期 2015-06-30.
作者简介 袁文博(1988-),男,博士研究生; 曹志强(通信作者),副研究员, E-mail: zhiqiang.cao@ia.ac.cn.
基金项目 国家自然科学基金资助项目(61273352, 61175111, 60805038); 山东省自然科学基金资助项目(ZR2011FM008).

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

找目标的突出特征进行匹配[7]. 文献[8]采用视觉单词树直方图和全局颜色直方图构建所跟踪目标的人体外观模型，并使用 SVM 增量学习进行在线训练. Gray [9]等提出 ELF 模型，通过分析局部特征的整体情况，来对不变视角下的行人进行再识别，其使用 Adaboost 算法进行学习. Schwartz [10]融合大量特征，并利用偏最小二乘方法(partial least squares，PLS)对获取到的高维特征降维，保证了较高的识别率. 本研究对目标进行区域分割，针对关键区域提取颜色和形体特征向量并进行融合，有效保证了目标再识别的效果.

1　基于局部特征融合的目标再识别

本文提出的目标再识别方法，涉及到的主要步骤包括目标的区域分割、特征描述子设计、测度选取及特征融合.

1.1　目标的区域分割

前景图像被提取出来之后，应对目标特征稳定、信息丰富的区域加以关注，这样提取的特征才更具鲁棒性，所以须要对目标进行区域分割. 合适的区域分割，并从中挑选出关键区域，可以帮助特征描述子更好的描述目标.

目标的区域分割由颜色分割算子和形体分割算子共同完成.

颜色分割算子定义为

$$C(j,\tau)=\frac{1}{\chi}\sum_{v=1}^{\tau}\sum_{i}d(p_{i,j-v},p_{i,j+v}),\qquad(1)$$

式中：$p_{i,j-v}\in I$；$p_{i,j+v}\in I$. I 为前景图像；$p_{i,j}$ 为全局坐标为(i,j)的像素点；$d(p_{i,j-v},\ p_{i,j+v})$为两个像素点在 RGB 空间的欧氏距离；χ 为累加的总次数；τ 为固定值，决定了计算颜色分割算子所用到的图像范围；$C(j,\tau)$表征了以第 j 行为分界线，上下各. τ 行区域间颜色特征差异情况. 在此基础上，利用颜色分割算子计算各区域间的颜色分割线 c，其中，需要分割的区域数目根据目标确定.

形体分割算子定义为

$$B(j,\varepsilon)=\frac{1}{\varepsilon}\sum_{v=1}^{\varepsilon}\max\left(\frac{L_{j-v}}{L_{j+v}},\frac{L_{j+v}}{L_{j-v}}\right),\qquad(2)$$

式中：L_j 为前景图像第 j 行的长度；ε 为固定值，决定了计算形体分割算子所用到的图像范围；$B(j,\varepsilon)$表征了以第 j 行为分界线，上下各 ε 行区域间形体特征差异情况.

与颜色分割算子类似，利用形体分割算子求取各区域间的形体分割线 b. 将颜色分割线和形体分割线融合，得到最终的分割结果 f.

1.2　特征描述子设计

完成目标的区域分割后，需要设计特征描述子，来提取目标关键区域的特征信息. 特征描述子包括加权的 HSV 颜色直方图和形体信息描述子两部分.

通常，目标再识别假设目标在一定时间内外观特征不会改变，因此颜色是进行再识别最常用和有效的特征. 但传统颜色直方图内并没有包含空间位置信息，一定程度上降低其特征描述能力，因此本文将像素的空间位置转化为权重，引入颜色直方图内，设计了加权的 HSV 颜色直方图(使用色调和饱和度两个参数).

R^k 表示第 k 个区域的加权的 HSV 颜色直方图，其各维可以表示为

$$R_{h,s}^{k}=\sum_{j=f_k}^{f_{k+1}}\sum_{i}\delta_{i,j}^{h,s}\,\eta_1 Q_{i,j},\qquad(3)$$

式中：$h\in 1,2,\cdots,H$，H 为色调的量化级数；$s\in 1,2,\cdots,S$，S 为饱和度的量化级数，若像素 $p_{i,j}$ 的色调在 h 内，同时饱和度在 s 内，则 $\delta_{i,j}^{h,s}$ 值为 1，否则为 0；$Q_{i,j}$ 为像素 $p_{i,j}$ 的权重，服从以区域重心 G 为期望值的二维正态分布；η_1 为 $Q_{i,j}$ 的归一化常数. 靠近重心位置的像素点，在直方图统计中所占权重更大，空间位置信息融合到颜色信息里面，有更好的鲁棒性. R^k 的维数固定不变，其大小由 H 和 S 量化级数决定.

颜色信息对目标位姿变化具有一定的不变性，但是遇到较为复杂的情况，如目标与背景颜色相似、不同目标颜色接近或相同时，仅靠颜色信息很难准确完成目标再识别. 考虑到目标间形体特征大多存在差别，因此设计出一种形体信息描述子.

形体信息描述子实现过程如下：

步骤 1　根据第 k 个区域的分割结果 f_k 和 f_{k+1}，获取该区域的前景图像 M.

步骤 2　计算 M 的区域重心 G.

步骤 3　计算 M 左右边界的切线 L_l、R_l.

步骤 4　依据重心坐标 G，将 M 划分为四个子部分 $M_1\sim M_4$，如图 1 所示.

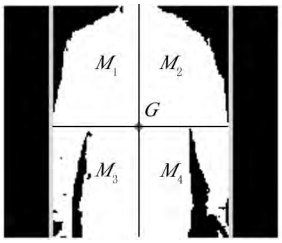


图 1　前景图像 M 划分示意图

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

步骤5 对 $M_1 \sim M_4$ 分别以 L_l 与 R_l 和 f_k 与 f_{k+1} 的交点为原点,统计各子部分的形体信息,具体方法的数学公式为

$$V_{m,a}^k = \sum_{p_{x,y} \in M_m} \delta_{x,y}^a (1 - \eta_2 T_{x,y}); \tag{4}$$

$$\delta_{x,y}^a = \begin{cases} 1 & ([2A\arctan(y/x)/\pi] \in [a-1, a)), \\ 0 & (\text{否则}), \end{cases} \tag{5}$$

式中:$m \in 1,2,3,4$;$a \in 1,2,\cdots,A$,A 为各子部分形体信息的向量维数;$p_{x,y}$ 为 M 内的像素点;x 和 y 为相对其所在的子部分的原点的坐标,可根据式(5)求取其对应的 $\delta_{x,y}^a$;$T_{x,y}$ 为像素 $p_{x,y}$ 的权重,服从以区域重心 G 为期望值的二维正态分布;η_2 为 $T_{x,y}$ 的归一化常数. 离重心位置越远的像素点,包含了目标更多的形体信息,在形体信息统计中权重也越大.

步骤6 将统计好的形体信息级联得到 V^k,即第 k 个区域的形体信息向量.

1.3 测度选取及特征融合

得到目标的加权的 HSV 颜色直方图和形体信息向量后,须要进行相似性度量,来完成目标再识别. 由于其提取特征的方式不同,因此使用的测度也不相同.

颜色直方图的测度要能表征两个区域颜色接近的程度,因此本文选择使用直方图的交. 具体方法为

$$r_c(I, \hat{I}) = \sum_{k=1}^{N} \omega_k \frac{\sum_h \sum_s \min(R_{h,s}^k, \hat{R}_{h,s}^k)}{\sum_h \sum_s R_{h,s}^k}, \tag{6}$$

式中:I 和 $\hat{I}$ 是两个待比较的前景图像;$r_c(I, \hat{I})$ 为两个前景图像的颜色特征的相似度;N 为划分的区域数目;. ω_k 为相应区域在颜色直方图测度中的权重.

形体信息向量由四个子部分的形体信息级联而成,形体信息向量的测度需要重点表征对应部分的相似程度,因此本文使用皮尔逊相关系数. 具体方法为

$$r_s(I, \hat{I}) = \sum_{k=1}^{N} \xi_k \rho(V^k, \hat{V}^k), \tag{7}$$

式中:$r_s(I, \hat{I})$ 为两个前景图像的形体特征的相似度;ξ_k 为相应区域在形体信息测度中的权重;$\rho(V^k, \hat{V}^k)$ 为 V^k 和 $\hat{V}^k$ 的皮尔逊相关系数.

采用加权组合的方式对两种特征的相似度结果进行融合,即

$$r(I, \hat{I}) = \lambda r_c(I, \hat{I}) + (1 - \lambda) r_s(I, \hat{I}), \tag{8}$$

式中:$r(I, \hat{I})$ 为目标间最终相似度结果;λ 为权值.

2 实验结果

基于视觉传感网采集的视频图像,组建数据库,其中共使用 4 个场景,有 21 个目标(包括 20 个人和 1 个机器人)共计 269 个样本.

实验中,相关参数取值如下:$\tau = 20$,$\varepsilon = 15$,$A = 30$,$\lambda = 0.6$. 对于加权的 HSV 颜色直方图,$H = 12$,$S = 8$,将色调均匀量化为 12 级,饱和度按照{0, 0.075, 0.15, 0.275, 0.4, 0.575, 0.75, 0.875, 1.0}非均匀量化为 8 级. 当目标为人时,依次划分为头、躯干、下肢 3 个区域,躯干和下肢区域为关键区域,而对于本实验中的机器人,自下向上依次划分为 3 个区域,选取中间和上面区域为关键区域,设定 $\omega_1 = 0$,$\omega_2 = 0.5$,$\omega_3 = 0.5$,$\xi_1 = 0$,$\xi_2 = 1$,$\xi_3 = 0$.

目标再识别领域的主流评价标准是累计匹配特征曲线(cumulative matching characteristic, CMC),本实验即采用此评价标准,并与仅使用颜色特征、仅使用形体特征两种方法进行对比. 实验具体过程如下,每轮实验开始时,随机选出某一目标的一对样本,分别作为目标样本和探测样本,同时从数据库中随机选取 149 个其他目标的样本,和探测样本共同组成容量为 150 的检测图库,对目标样本和检测图库中的样本求取相似度,并且记录探测样本的相似度排名,重复进行 100 轮实验.

将实验结果进行统计,绘制成 CMC 曲线,结果如图 2 所示. 排名等级 n 代表在每轮实验中,选

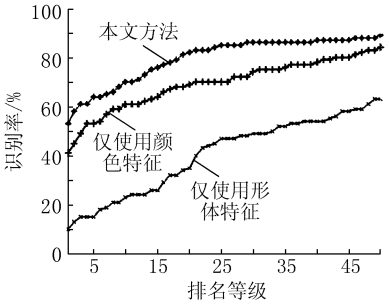
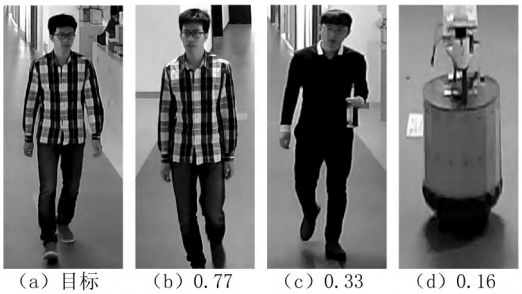


图 2 实验结果

取相似度最高的前 n 个结果,而如果其中包含探测样本,即认为本轮实验识别正确. 从实验结果可以看出:由于目标的形体特征具有易变性,因此仅使用形体特征的识别率低于仅使用颜色特征. 而使用本文方法,即对局部特征进行融合,识别率比仅使用颜色特征高出 10%左右,显示出更强的有效性.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

图3和图4为本文方法的效果展示，共两组，分别以人和机器人为目标，图(b)～(d)均给出了其与目标的相似度，可以看出：对于同一个目标，在监控场景、目标位姿等发生改变后，相似度仍保持在较高水平，而对不同的目标，相似度则一直较低.



(a) 目标　(b) 0.77　(c) 0.33　(d) 0.16

图3　本文方法的效果展示(人为目标)

(a) 目标　(b) 0.71　(c) 0.40　(d) 0.25

图4　本文方法的效果展示(机器人为目标)

本文面向视觉传感网监控的大尺度场景，提出一种基于局部特征融合的目标再识别方法. 实验结果表明，所提方法能够保证目标再识别的质量. 下一步拟将目标的区域分割进行优化，提升分割效果.

参　考　文　献

[1] Bak S, Corvee E. Person re-identification using spatial covariance regions of human body parts[C]//Proc of Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance. Piscataway: IEEE Computer Society, 2010: 435-440.

[2] Bazzani L, Cristani M. Multiple-shot person re-identification by chromatic and epitomic analyses[J]. Pattern Recognition Letters, 2012, 33(11): 898-903.

[3] Aziz K, Merad D, Fertil B. People re-identification across multiple non-overlapping cameras system by appearance classification and silhouette part segmentation[C]//Proc of 2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance. Piscataway: IEEE Computer Society, 2011: 303-308.

[4] Bäuml M, Stiefelhagen R. Evaluation of local features for person re-identification in image sequences[C]//Proc of 2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance. Piscataway: IEEE Computer Society, 2011: 291-296.

[5] Cai Y, Pietikäinen M. Person re-identification based on global color context[C]//Proc of Lecture Notes in Computer Science. Heidelberg: Springer, 2011: 205-215.

[6] Ma B, Su Y, Jurie F. Local descriptors encoded by Fisher Vectors for person re-identification[C]//Proc of Lecture Notes in Computer Science. Heidelberg: Springer, 2012: 413-422.

[7] Zhao R, Ouyang W, Wang X. Unsupervised salience learning for person re-identification[C]//Proc of Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2013: 3586-3593.

[8] 许允喜，蒋云良，陈方. 基于支持向量机增量学习和LPBoost的人体目标再识别算法[J]. 光子学报, 2011, 40(5): 758-763.

[9] Gray D, Tao H. Viewpoint invariant pedestrian recognition with an ensemble of localized features[C]//Proc of Lecture Notes in Computer Science. Heidelberg: Springer, 2008: 262-275.

[10] Schwartz W R, Davis L S. Learning discriminative appearance-based models using partial least squares [C]//Proc of Proceedings of SIBGRAPI 2009. Piscataway: IEEE Computer Society, 2009: 322-329.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net