

基于因果分析的群体行为识别

王 军, 夏利民

(中南大学信息科学与工程学院 长沙 410075)

【摘要】该文提出一种基于因果分析的群体行为识别方法,利用Grange因果检验分析个体行为间的因果关系,在此基础上,结合个体间的因果关系、空间位置关系和视觉注意力范围,利用基于主集的聚类法检测行为群体。为了有效地表示群体行为,用方向梯度直方图和光流直方图描述个体行为,用因果特征描述个体间的交互行为。采用稀疏表示进行群体行为识别,在公共数据库BEHAVE和collective activity上对该方法进行验证,并与其他方法进行对比试验,结果表明了该方法的有效性。

关 键 词 因果分析; 主集; 群体行为识别; 群体检测; 稀疏表示

中图分类号 TP391 **文献标志码** A **doi:**10.3969/j.issn.1001-0548.2018.02.015

Group Activity Recognition in Crowd via Causality Analysis

WANG Jun and XIA Li-min

(School of Information Science and Engineering, Central South University ChangSha 410075)

Abstract A novel method for group activity recognition in crowd is proposed using causality analysis. The Granger Causality Test is used to analyze the causality between individual actions. On this basis, we adopt a dominant set based clustering algorithm to detect interacting groups in crowded scenes using causality, spatial and directional relationships among people. To effectively represent group activity, low level visual features and causality features are used. The low level visual features, which included histograms of oriented gradients (HOG) and histograms of optical flow (HOF), are applied to describe the properties of individual activity, and the causality features obtained by causality analysis are introduced to depict the interaction information of people. Sparse representation is employed to recognize group activities in crowd. Experiments are performed on the BEHAVE and collective activity databases to test and evaluate the proposed method. The experiments results show that the proposed method is more effective than other state-of-the-art methods.

Key words causality analysis; dominant set; group activity recognition; group detection; sparse representation

行为识别是计算机视觉领域的一个重要研究方向,在人机交互、智能视频监控等方面有着广阔的应用前景。近年来人体行为识别取得了显著进展,但主要工作还是集中在单人行为识别和两人交互行为识别上,群体行为识别仍是一项挑战性的工作。

目前,大多数群体行为识别方法主要分为基于特征的方法和基于轨迹的方法。在基于特征的群体行为识别方面,文献[1]用位置、形状、运动以及时序等特征表示群体行为。文献[2]采用局部视觉信息和群体行为链模型识别群体行为。文献[3]利用局部表面信息作为群体行为描述子。文献[4]利用运动信息和外观信息作为行为特征,采用核密度估计法建立群内个体间的交互关系。基于轨迹的群体行为识

别^[5-7]是通过分析个体轨迹间的相互影响来识别群体行为。文献[5]利用Granger因果检验分析两个行为轨迹间的相互作用,并以此识别双人间的交互行为。文献[6]将该方法推广到群体行为识别。文献[7]用高斯过程表示人体运动轨迹,设计个体、双人、群体行为描述子进行群体行为识别。

然而,这些方法假设场景中只有一种群体行为,但实际环境中,可能同时存在个体行为、两人的交互行为以及群体行为,如果将整个行为视频看成是一个群体行为进行识别,将会影响群体行为的识别率。为了解决这一问题,文献[8-13]首先将行为群体从人群中检测出来,然后识别群体行为。文献[10]利用运动轨迹信息和外观信息定义了3种群体行为

收稿日期: 2017-06-21; 修回日期: 2017-11-15

基金项目: 国家自然科学基金(51678075); 国家863项目(2009AA11Z205)

作者简介: 王军(1971-),男,博士生,主要从事图像处理与模式识别方面的研究。

描述子, 分别描述子群内、子群间的内在交互行为, 采用SVM进行群体行为识别, 但该方法很难检测出子群。文献[11-12]利用群交互区域检测场景中有意义的群体, 用群体交互能量特征、吸引和排斥特征来描述GIZ中的群体行为, 采用SVM进行群体行为识别。文献[13]利用图聚类算法检测人群中的群体, 设计了一种表示运动和交互作用的描述子, 采用SVM识别群体行为。但这些方法仅仅依靠个体间的空间关系来检测行为群体, 这使得群体检测率不是很高, 从而影响群体行为识别; 同时, 只是用视觉特征描述群体行为, 而这些特征不能有效描述群体行为, 因此也限制了群体行为识别率。

本文提出一种新的基于因果分析的群体行为识别方法。利用Grange因果检验对个体行为间关系进行分析, 给出了个体间因果关系, 在此基础上, 首先, 根据个体间的空间位置关系、视觉注意力方向以及因果关系, 利用基于主集的聚类算法检测人群中的行为群体; 然后, 结合低层视觉特征和因果特征描述群体行为, 由于不同群体行为个体数目不同, 从而对应的轨迹数不同, 导致整个特征长度不同。为了解决该问题, 采用词包法将特征转化为行为视觉词表示; 最后, 采用稀疏表示进行群体行为识别。

1 Grange因果检验

文献[14]提出了Grange因果检验, 揭示不同经济因素之间的因果关系和反馈关系。Granger认为如果加入另一个时间序列 X_2 , 当前时间序列 X_1 的自回归预测误差的方差减小, 那么 X_2 对 X_1 有因果影响。

设有两个联合平稳随机过程分别为 $X_i(t)$ 和 $X_j(t)$, 其对应的自回归模型分别表示为:

$$X_i(t) = \sum_{k=1}^n a_{1k} X_i(t-k) + \varepsilon_1 \quad (1)$$

$$\text{var}(\varepsilon_1) = \Sigma_1$$

$$X_j(t) = \sum_{k=1}^n d_{1k} X_j(t-k) + \eta_1 \quad (2)$$

$$\text{var}(\eta_1) = \Gamma_1$$

式中, ε_1 、 η_1 分别为模型噪声; Σ_1 和 Γ_1 为方差。

$X_i(t)$ 和 $X_j(t)$ 的联合自回归模型可写为:

$$X_i(t) = \sum_{k=1}^n a_{2k} X_i(t-k) + \sum_{k=1}^n b_{2k} X_j(t-k) + \varepsilon_2 \quad (3)$$

$$\text{var}(\varepsilon_2) = \Sigma_2$$

$$X_j(t) = \sum_{k=1}^n c_{2k} X_j(t-k) + \sum_{k=1}^n d_{2k} X_i(t-k) + \eta_2 \quad (4)$$

$$\text{var}(\eta_2) = \Gamma_2$$

根据Grange因果检验, 可以得到下列结论:

1) 因果关系: 如果 $\Sigma_2 < \Sigma_1$, 则 X_2 是 X_1 的Granger因果关系。

2) 反馈: 如果 X_2 是 X_1 的Granger因果关系, 而 X_1 是 X_2 的Granger因果关系, 则 X_1 与 X_2 具有反馈。

3) 因果率 F_c :

$$F_{cij} = \ln \left(\frac{\Sigma_1}{\Sigma_2} \right) \quad (5)$$

因果率 F_c 测量因果关系的强度, 因果影响越强, F_{cij} 也越大; 若 X_2 对 X_1 没有因果影响, 则 $F_{cij} = 0$ 。

4) 反馈率 F_{fij} 表示反馈强度:

$$F_{fij} = \ln \left(\frac{\Gamma_1}{\Gamma_2} \right) \quad (6)$$

本文利用Grange因果检验分析群体行为。对于一对个体, 根据其运动轨迹, 由最小二乘法可估计运动模型式(1)~式(4)中的系数及方差, 再利用式(5)~(6)求得他们之间的因果律和反馈率, 根据这些因果关系进行群体检测和群体行为表示。

2 行为群体检测

由于复杂场景中可能存在独立的个体行为、双人交互行为和群体行为, 为了有效地识别群体行为, 首先必须检测出有意义的群体, 为此利用因果关系和基于主集的聚类方法进行行为群体检测。

2.1 主集

主集^[15]是最大团到无向加权图的推广。考虑一个无向加权图 $G=(V, E)$, $V=\{1, 2, \dots, n\}$ 是所有顶点的集合, $E \in V \times V$ 是所有边的集合, 设 $S \subseteq V$ 是顶点的一个非空子集, $i \in S$, 则 i 关于 S 的平均加权度为:

$$\text{awdeg}_S(i) = \frac{1}{|S|} \sum_{j \in S} a_{ij} \quad (7)$$

式中, $|S|$ 表示 S 的基数; a_{ij} 为节点 j 和 i 之间的相似度。如果节点 $j \notin S$, 则节点 j 和 i 的相似性为:

$$\phi_S(i, j) = a_{ij} - \text{awdeg}_S(i) \quad (8)$$

i 关于 S 的权值为:

$$w_S(i) = \begin{cases} 1 & |S|=1 \\ \sum_{j \in S \setminus \{i\}} \phi_{S \setminus \{i\}}(j, i) w_{S \setminus \{i\}}(j) & \text{其他} \end{cases} \quad (9)$$

S 的整体相似性定义为:

$$W(S) = \sum_{i \in S} w_S(i) \quad (10)$$

满足下列条件的非空子集 $S \subseteq V$ 称为主集:

$$w_S(i) > 0 \quad \forall i \in S \quad (11)$$

$$w_{S \cup \{i\}}(i) < 0 \quad \forall i \notin S \quad (12)$$

这表明主集等同类。求解下列优化问题可得到主集:

$$\begin{aligned} & \text{Maximize } f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \\ & \text{subject to } \mathbf{x} \in \Delta_n \end{aligned} \quad (13)$$

式中, $\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : x_i \geq 0, i \in V \text{ and } \mathbf{1}^T \mathbf{x} = 1\}$ 。

如果 \mathbf{x} 是式(13)的局部极值解, 那么 \mathbf{x} 的支撑集 $\sigma(\mathbf{x}) = \{i \in V : x_i > 0\}$ 是节点的一个主集。

2.2 基于主集图聚类的行为群体检测

类具有较高的内部一致性, 并且类内元素与内外元素具有较高的可区分性。同样, 对于相互作用的群体, 同一群体内的成员之间具有很强的相互作用, 而群内成员与群外成员之间相互作用比较小。基于这种比较, 将行为群体检测看成在图中搜索最大相互作用节点的主集, 于是行为群体检测就转化为基于主集的图聚类问题。

将场景中的人看成无向图 $G=(V, E, A)$, 图的顶点代表人, 而 $A=(a_{ij})$ 代表场景中人与人之间的交互作用程度。很显然, 当两人间的因果关系越强, 他们之间的交互作用也就越强; 当两个人越接近, 他们之间交互作用的可能性也越大; 另外, 两个人处在对方的视觉注意力范围时, 他们交互作用的可能性也越大。基于这些直觉, 利用个体间的因果关系、空间位置关系以及视觉注意力方向关系定义个体间的交互程度。假设人的运动轨迹、空间位置、头部姿态已经确定。则根据个体的运动轨迹, 利用Grange因果分析可得到个体间的因果律 F_{cij} 和反馈率 F_{fij} ; 由空间位置可得到他们之间的空间距离 d_{ij} ; 给定第 i 人的头部姿态, 可确定其视觉注意力范围 α_i , 如第 j 个人出现在视觉注意力范围 α_i 内, 则记为 $j \in \alpha_i$ 。于是, 第 i 和第 j 个人之间的交互程度 a_{ij} 定义为:

$$a_{ij} = \begin{cases} 0 & i \notin \alpha_j, j \notin \alpha_i \\ \frac{1}{2d_{ij}}(F_{cij} + F_{fij}) & (i \in \alpha_j, j \notin \alpha_i) \text{ 或 } (i \notin \alpha_j, j \in \alpha_i) \\ \frac{1}{d_{ij}}(F_{cij} + F_{fij}) & i \in \alpha_j, j \in \alpha_i \end{cases} \quad (14)$$

根据式(14), 可以确定矩阵 A , 然后根据文献[16]的方法求解式(13)的局部优化解和它的支撑集, 该支撑集即为相互作用的群体。

3 群体行为识别

3.1 群体行为表示

对于检测到的群体, 其行为用群内个体的行为

和个体间的交互作用来表示。利用低层视觉特征描述个体行为, 因果关系描述个体间的交互行为。

1) 个体行为的低层特征表示

采用局部外观特征HOG和局部运动特征HOF表示个体行为。首先, 利用多目标跟踪算法^[17]得到每个人的运动轨迹。为了避免跟踪过程中轨迹的漂移, 只跟踪一段轨迹而不是整段轨迹, 本文设跟踪轨迹的长度 $L=50$ 。其次, 沿着每个人的轨迹, 建立一个包围人体的、大小为 $N_x \times N_y \times L$ 的立方体, 并将该立方体划分为 $n_x \times n_y \times n_L$ 的时-空网格, 本文取 $n_x=n_y=2$, $n_L=3$; 然后, 在网格上计算4-bin HOG和5-bin HOF。最后将所有网格的特征连起来得到HOG描述子和HOF描述子, HOG、HOF的维数分别为48 ($2 \times 2 \times 3 \times 4$)和60 ($2 \times 2 \times 3 \times 5$)。连接HOG和HOF, 得到个体行为的低层特征表示 F_1 , 其维数为108。

2) 交互行为的因果特征表示

因果特征描述个体间的交互行为, 包括两个部分: 相互作用强度和如何相互作用。用因果率 F_{cij} 和反馈率 F_{fij} 表示两个个体之间的相互作用强度, 而个体间如何相互作用按下列方法确定。

设两个个体运动轨迹 X_i 和 X_j 的联合自回归模型为式(2)和(4), 如果将式(2)看成输入、输出分别为 X_j 和 X_i 的数字滤波器模型, 则对应的系统函数为:

$$H_{ij}(z) = \frac{X_i(z)}{X_j(z)} = \frac{\sum_{k=1}^n b_{2k} z^{-k}}{1 - \sum_{j=1}^n a_{2k} z^{-k}} \quad (15)$$

式中, $X_i(z)$ 和 $X_j(z)$ 分别为 X_i 和 X_j 的Z变换。利用对应的频率特性来表示个体间如何相互作用^[5], 即取频率特性在 $0, \pi/4, \pi/2, 3\pi/4, \pi$ 处的幅值, 以及在 $\pi/4, \pi/2, 3\pi/4$ 处的相角(因为在 0 和 π 处相角为常数):

$$\mathbf{F}_{ij} = \left[|H_{ij}(e^{j0})|, |H_{ij}(e^{j\pi/4})|, \dots, |H_{ij}(e^{j\pi})|, \angle H_{ij}(e^{j\pi/4}), \dots, \angle H_{ij}(e^{j\pi}) \right] \quad (16)$$

\mathbf{F}_{ij} 反映 X_j 如何影响 X_i 。同样, 分别以 X_i 、 X_j 为滤波器的输入和输出, 得到特征向量 \mathbf{F}_{ji} , 反映 X_i 如何影响 X_j 。

因果率和反馈率表示一对个体之间的相互作用的强度, 而频率特性 \mathbf{F}_{ij} 、 \mathbf{F}_{ji} 反映一对个体如何相互作用, 因此这些特征是互补的, 将其合在一起表示一对个体的交互行为。另外, 个体间的相对距离和相对速度通常也用于描述交互行为。因此, 用 \mathbf{F}_{ij} 、 \mathbf{F}_{ji} 、 d_{ij} 、 Δv_{ij} 表示一对个体的交互行为:

$$\mathbf{F}_2(X_i, X_j) = (\mathbf{F}_{ij}, \mathbf{F}_{ji}, F_{cij}, F_{fij}, d_{ij}, \Delta v_{ij}) \quad (17)$$

式中, F_2 是一个20维的因果特征向量。

3) 群体行为表示

如上所述, 每个个体行为用一条轨迹的108维低层特征来描述, 一对个体的交互行为用一对轨迹的20维因果特征来描述, 而整个群体行为由所有个体的行为特征及所有行为对特征联合表示。但由于不同的群体行为包含个体的数目不同, 相应的轨迹条数、轨迹对数不同, 导致不同的群体行为特征维数不同。为此, 利用词包法(bag-of-words, BOW)来表示群体行为。首先根据低层特征和因果特征, 用均值聚类法分别建立个体行为的字典 BOW_{ind} 和交互行为字典 BOW_{int} 。对于每个群体行为, 根据最近邻原则, 把它的每条轨迹/轨迹对归入每一类, 可得到各个视觉词在样本中出现的频率, 这些频率就构成了样本的视觉词表示。本文两个字典大小分别为160、40, 因此, 群体行为用一个200维的视觉词向量 y 表示。

3.2 基于稀疏表示的群体行为识别

利用稀疏表示识别群体行为。首先, 对待测试群体行为样本进行编码; 然后, 计算在每一类上的重建误差; 最后, 将待测试群体行为样本分类到重建误差最小的类中。

设 C 有类群体行为, 每个行为用上述特征向量表示, $D=[D_1, D_2, \dots, D_C]$ 表示稀疏字典, 其中 D_i 是由 K 个第 i 类行为构成的子字典, 本文利用 K -SVD算法来确定稀疏字典^[18]。对于测试样本 y 可表示为:

$$y = D\alpha \quad (18)$$

式中, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_C]^T$ 为稀疏编码向量, 通过求解下列优化问题得到:

$$\hat{\alpha} = \arg \min_{\alpha} \left\{ \|y - D\alpha\|_2^2 + \lambda \|\alpha\|_1 \right\} \quad (19)$$

利用OMP法(orthogonal matching pursuit)求解上述优化问题^[19]。

然后计算相对第 i 类的重建误差:

$$R_i = \|y - D_i \hat{\alpha}_i\|_2 \quad (20)$$

最后, 群体行为样本 y 根据下式分类(识别):

$$\text{label}(y) = \arg \min_i \{R_i\} \quad (21)$$

4 实验与结果

为了验证提出方法的有效性, 本文在公共数据库BEHAVE和collective activity(CAD)上进行了实验, 并与文献[10, 12-13]的方法进行对比。所有实验都在工作站(2.8 GHz CPU, 32GB RAM)上进行。在实验中, 首先根据个体运动轨迹, 由最小二乘法估计运动模型的系数及方差, 利用Grange因果检验求得

个体间因果律和反馈率; 其次, 根据因果律、反馈率、个体间的空间距离及视觉注意力范围, 利用基于主集的聚类法找出行为群体; 然后沿着每条轨迹提取HOG和HOF特征, 以及两条轨迹间的因果特征, 并采用 K -均值聚类法得到160个的低层特征视觉词和40个因果特征视觉词表示群体行为; 最后, 利用利用稀疏表示识别群体行为。

4.1 BEHAVE数据库

BEHAVE数据库包括4段视频, 帧速为25 fps, 分辨率为640×480, 记录了10类群体行为, 包括InGroup, Approach, Meet, WalkTogether, Split, Ignore, Chase, Fight, RunTogether, Following。本文只考虑Approach(A), Split(B), WalkTogether(W), RunTogether(R), Fighting(F)和InGroup(I)这6种群体行为, 其他群体行为不考虑。

图1为该数据库的一些群体行为的视频帧。图1a为多人追赶; 图1b为多人打架。实验包含470个群体行为样本, 用220个群体行为学习稀疏字典, 250个群体行为用于测试实验。表1为几种方法的识别率, 可看出本文方法的识别率最高, 平均识别率达到了94.7%。表2为几种方法的平均识别时间, 本文的方法识别时间为0.86 s, 略小于其他方法, 表明本文方法识别速度快。

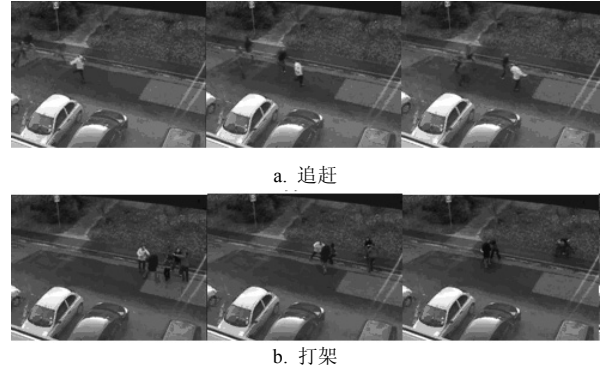


图1 BEHAVE数据库

表1 在BEHAVE数据库上的识别率

行为	识别率/%			
	文献[10]	文献[12]	文献[13]	本文方法
Approach	83.3	83.3	85.4	88.5
Split	76.9	91.4	92.3	94.6
WalkTogether	96.0	91.6	96.0	96.2
RunTogether	91.3	89.2	90.7	92.5
Fighting	88.7	88.9	89.4	90.6
InGroup	96.2	93.4	94.5	97.0
平均	88.5	90.6	91.2	94.7

表2 在BEHAVE数据库上的平均识别时间

方法	文献[10]	文献[12]	文献[13]	本文方法
识别时间/s	0.88	0.92	0.91	0.86

BEHAVE数据库中部分视频同时含有个体行为、双人交互行为和群体行为,为了验证文中提出的群体检测对群体行为识别的影响,对有、无行为群体检测进行了对比实验。表3和表4为两种情况下,在BEHAVE数据库上的群体行为识别的混淆矩阵。结果表明有群体检测明显提高群体行为的识别率。

表3 有行为群体检测混淆对比

行为	A	S	W	R	F	I
A	88.5	0.00	0.00	0.00	0.00	11.5
S	0.00	94.6	2.20	1.00	1.20	1.00
W	0.00	0.00	96.2	3.80	0.00	0.00
R	0.00	0.00	5.20	92.5	0.00	2.30
F	9.50	0.00	0.00	0.00	90.5	0.00
I	3.00	0.00	0.00	0.00	0.00	97.0

表4 无行为群体检测混淆对比

行为	A	S	W	R	F	I
A	82.5	0.00	6.50	0.00	0.00	11.0
S	0.00	90.6	3.40	2.50	2.25	1.25
W	1.30	0.00	91.8	4.20	0.00	2.70
R	0.00	0.00	6.50	88.3	0.00	5.20
F	8.70	1.60	0.00	0.00	89.7	0.00
I	4.00	0.00	2.40	3.60	0.00	90.0

4.2 Collective Activity数据库

CAD数据库包含5不同的群体行为:横过马路、等待、步行、谈话和排队。图2为CAD数据库的部分群体行为的视频帧。图2a横过马路;图2b多人步行。实验包含550个群体行为样本,用250个群体行为学习稀疏字典,以300个群体行为进行测试实验。表5为几种方法在CAD数据库上识别率,表中结果表明,本文方法的平均识别率达到了86.7%,明显高于其他方法。表6为几种方法的平均识别时间,本文方法识别时间为0.94 s,小于其他方法。



a. 过马路



b. 步行

图2 CAD数据库

表5 在CAD数据库上的识别率

行为	识别率/%			
	文献[10]	文献[12]	文献[13]	本文方法
Crossing	72.3	69.3	78.4	79.5
Walking	83.4	81.5	81.5	84.3
Waiting	67.6	71.6	77.8	78.4
Talking	85.7	89.2	100	100
Queueing	71.5	76.3	85.9	87.5
平均	76.2	79.8	84.2	86.7

表6 在CAD数据库上的平均识别时间

方法	文献[10]	文献[12]	文献[13]	本方法
识别时间/s	0.98	1.21	1.12	0.94

表7 有行为群体检测混淆对比

行为	Crossing	Waiting	Queueing	Walking	Talking
Crossing	79.5	5.80	1.00	13.7	0.00
Waiting	3.10	78.4	1.00	16.5	1.00
Queueing	0.00	0.00	87.5	12.5	0.00
Walking	12.0	2.00	1.00	84.3	0.00
Talking	0.00	0.00	0.00	0.00	100

表8 无行为群体检测混淆对比

行为	Crossing	Waiting	Queueing	Walking	Talking
Crossing	67.5	6.50	3.50	22.5	0.00
Waiting	5.40	64.6	3.40	24.1	2.50
Queueing	0.00	0.00	81.4	18.6	0.00
Walking	22.6	4.00	2.00	71.2	0.00
Talking	0.00	1.30	0.00	0.00	98.7

该数据库中的每段视频都包同时含个体行为、双人交互行为和群体行为,实验中,在CAD数据库上验证了文中提出的群体检测对群体行为识别的影响。表7和表8分别为有、无群体检测的混淆矩阵。结果表明加入了群体检测,群体行为的识别率明显得到提高。

从上述结果可看出,相比其他几种方法,文中提出的方法在各种情况下都能提高群体行为的识别率,并且识别速度最快。在BEHAVE数据库和collective activity数据库中,场景中都可能同时存在独立的个体行为、双人的交互行为和群体行为,虽然文献[10]利用子群信息可以识别群体行为,但很难检测子群,导致这种方法的群体识别率不是很高。文献[12]利用群体交互区检测有意义的群体,排除了没有参与群体行为的个体,所以识别率得到了提高。然而,群体的检测方法仅仅依靠个体之间的空间关系,这种关系并不能完全表明个体之间存在交互关系(如相互接近的人不一定有交互行为),这使得群体检测率不是很高,从而限制了群体行为识别率。文

献[13]利用视觉注意力和社会距离检测交互群体,检测率得到提高,利用运动信息和交互信息识别群体行为,因此能有效地识别群体行为。本文的方法同时利用因果关系、空间位置关系和视觉注意力范围,因此更能有效地从人群中检测行为群体,采用低层视觉特征和因果特征可有效表示群体行为,因此可提高群体行为识别率。

5 结 束 语

本文提出了一种新的基于因果分析的群体行为识别方法。首先,利用Grange因果检验分析个体间因果关系,其次,结合个体间因果关系、空间关系和视觉注意力范围,采用基于主集的聚类算法进行群体检测;然后,利用低层视觉特征和因果特征描述群体行为,其中,HOG和HOF用于表示个体行为,因果特征用于表示群体内个体间的交互行为,提高了特征的区分能力;最后,利用稀疏表示识别群体行为。在公共行为数据库BEHAVE和collective activity上与几种方法进行了对比实验,结果表明提出的方法比其他方法具有更高的识别率和识别速度。

参 考 文 献

- [1] KANEKO T, SHIMOSAKA M, ODASHIMA S. A fully connected model for consistent collective activity recognition in videos[J]. Pattern Recognition Letters, 2014, 43(1): 109-118.
- [2] AMER M R, TODOROVIC S. A chains model for localizing participantsof group activities in videos[C]//IEEE International Conference on Computer Vision. Barcelona: IEEE, 2011: 786-793.
- [3] ZHU G, YAN S, HAN T, et al. Generative group activity analysis with quaternion descriptor[J]. Adv Multimed Model, 2011, 6524: 1-11.
- [4] STEPHENS K, BORS A G. Group activity recognition on outdoor scenes[C]//IEEE International Conference on Advanced Video and Signal Based Surveillance. Colorado Springs: IEEE, 2016: 59-65.
- [5] ZHOU Y, HUANG T S, NI B, et al. Recognizing pair-activities by causality analysis[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(5): 1-20.
- [6] NI B, YAN S, KASSIM A. Recognizing human group activities with localized causalities[C]//IEEE Conference on Computer Vision and Pattern Recognition. Miam: IEEE, 2009: 1470-1477.
- [7] CHENG Z, QIN L, HUANG Q. Group activity recognition by gaussian processes estimation[C]//IEEE International Conference on Pattern Recognition. Istanbul: IEEE, 2010.
- [8] YIN Y, YANG G, XU J. Small group human activity recognition[C]//IEEE International Conference on Image Processing. Lake Buena Vista: IEEE, 2012: 2709-2712.
- [9] ZHANG C, YANG X, LIN W, et al. Recognizing human group behaviors with multi-group causalities[C]//International Conference on Web Intelligence and Intelligent Agent Technology Workshops. Macau: IEEE, 2012: 44-48.
- [10] JU J, YANG C, SCHERER S. Recognition of human group activity for video analytics[C]//Pacific-Rim Conference on Multimedia. Gwangju: IEEE, 2015.
- [11] KIM Y J, CHO N G, LEE S W. Group activity recognition with group interaction zone[C]//International Conference on Pattern Recognition. Stockholm: IEEE, 2014: 3517-3521.
- [12] CHO N G, KIM Y J, PARK U. Group activity recognition with group interaction zone based on relative distance between human objects[J]. International Journal Pattern Recognition and Artificial Intelligence, 2015, 29(5): 1-10.
- [13] TRAN K N, YAN Xu, KAKADIARIS I A, et al. A hybrid approach for individual and group activity analysis in crowded scene[J]. Communications in Computer and Information Science, 2016, 598: 189-204.
- [14] GRANGER C. Investigating causal relations by econometric models and cross-spectral methods[J]. Econometrica, 1969, 37(3): 424-438.
- [15] PAVAN M, PELILLO M. Dominant sets and pairwise clustering[J]. IEEE Transactions on PAMI, 2007, 29(1): 167-172.
- [16] BULO S R, BOMZE I M. Infection and immunization: a new class of evolutionary game dynamics[J]. GAMES and Economic Behavior, 2011, 71(1): 193-211.
- [17] MILAN A, SCHINDLER K, ROTH S. Multi-target tracking by discrete-continuous energy minimization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2054-2068.
- [18] ZHANG Y S, GUO Y. 2D OMP algorithm for space-time parameters estimation of moving targets[J]. Electronics Letters, 2015, 51(22): 1809-1820.
- [19] MUKHERJEE S, BASU R, SEELAMANTULA C S. 1-K-SVD: a robust dictionary learning algorithm with simultaneous update[J]. Signal Processing, 2016, 123: 42-52.

编辑 叶 芳