

# 多摄像机视域内的目标活动分析

蒋建国, 顾占冰, 胡珍珍, 齐美彬

(合肥工业大学计算机与信息学院, 安徽合肥 230009)

**摘 要:** 本文提出了一种能够分析多摄像机非重叠视域中运动目标行为之间时空相关性的方法. 该方法基于特征空间中目标活动模式的相似性和活动空间的关联性, 将摄像机网络中每个视域分为多个有意义的语义活动区域; 利用交叉核典型相关分析(XKCCA)分析语义活动区域之间的时空相关性, 得到摄像机网络的拓扑关系, 该拓扑关系能够反映目标在跨摄像机的语义区域之间运动的时空信息; 将这些信息有效地融入到跨摄像机的目标再确认过程中, 有利于排除虚假目标, 提高跨摄像机目标再确认的准确度. 与现有的方法相比, 本文方法不依赖于个体目标的跟踪, 实验结果表明本方法在复杂、拥挤、低帧频和低分辨率的多摄像机视频监控网络中能够有效地理解和分析视频内容, 更准确的实现跨摄像机目标再确认.

**关键词:** 语义区域划分; 核典型相关性分析; 多摄像机视域内容理解

**中图分类号:** TN911Y.23      **文献标识码:** A      **文章编号:** 0372-2112 (2014)02-0306-06

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2014.02.015

## Activity Analysis Cross Muti-Camera

JIANG Jian-guo, GU Zhan-bing, HU Zhen-zhen, QI Mei-bin

(Institute of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China)

**Abstract:** This paper proposes an approach to analyze the temporal and spatial correlations between objective activities from multiple non-overlapping camera network. Based on the similarity of moving models and the relationship of moving space, each vision field of camera network is segmented into semantic active regions automatically. Then a Cross Kernel Canonical Correlation Analysis is implemented to explore the correlations between these active regions and the topology of the camera network. This topology can reflect the temporal and spatial information of objectives cross multi-cameras and improve the accuracy of objects re-identification by removing false objects. Compared with existing methods, our approach does not depend on the individual tracking and is efficient in complex and crowded scene. The experiment results show that our approach performs effectively and efficiently in multi camera surveillance network.

**Key words:** camera view decomposition; KCCA; multiple camera content understanding

## 1 引言

对大型、复杂拥挤的公共场所进行视频监控往往由多个摄像机组成, 诸如商场、学校、机场、地铁站、火车站等. 如何更有效地利用多个摄像机进行监控和自动分析, 以减小监控人员的工作量是当前视频分析和智能监控的研究热点. 在监控网络中, 对视频语意的理解、摄像机拓扑关系的估计和跨摄像机目标的相关性分析对于全局监控和理解尤其重要. 国内外的许多研究机构和科研小组在这方面做出了重要的贡献.

E E Zelmiler<sup>[1]</sup>等用基于非监督的学习方法, 对目标的运动轨迹进行相关性分析, 对全局活动进行聚类建立

全局模型来进行异常行为监测. X Wang<sup>[2]</sup>等人通过分析行人在不同摄像机视域内相同活动模式区域的轨迹进行分析来建立运动轨迹网络, 从而建立运动模型进行视频内容分析处理. 此外, Gray D 等<sup>[3]</sup>以及 Prosser B 等<sup>[4]</sup>在视频分析中采用 CBTF (Cumulative Brightness Transfer Function) 相似函数学习等不同的方法得到较好的效果. 刘亚楠等<sup>[10-12]</sup>针对视频理解的工作注重语意的理解和分析, 有些工作注重目标在摄像机网络中移动的时间和空间拓扑关系, 王涛等<sup>[8]</sup>运用 ViBe 检测和跟踪目标建立摄像机网络的拓扑模型, 也有王亮等<sup>[9]</sup>运用步态来进行目标的检测与识别.

在单摄像机跟踪领域, 跟踪算法较为成熟, 如田

广等<sup>[14]</sup>利用特征变换和支持向量机对目标进行跟踪,郑锦等<sup>[18]</sup>通过背景概率方法提出较为鲁棒的背景重构法. 先前<sup>[15,16,19]</sup>的工作有些是在单独摄像机视域内检测和跟踪行人来建立转移时间关系或进行运动分类,有些是对目标在摄像机网络之间移动进行跟踪和检测来获得目标的运动模式. 然而,在大型、复杂、拥挤的场景内,由于目标之间相互遮挡严重,对于运动目标的检测、跟踪很多时候是失效的,不能满足分析的需求. C C Loy 等<sup>[5,13]</sup>通过对运动目标的统计特征建立模型,用交叉典型相关分析(Cross Kernel Canonical Correlation Analysis, Xcca)分析特征线性组合的相关性达到很好的效果. 但是,不同摄像机视域内,运动目标由于速度、方向、目的地不同,运动模式并非都是线性关系,这是 CCA 方法的局限性所在. 而 Kernel CCA 对于非线性组合的特征的分析具有很好的效果,能有效克服 CCA 的局限性,所以本文用 Kmet CCA 分析视频场景内非线性组合的运动目标特征. 本文方法不依赖于运动目标检测和跟踪,而是对不同视域场景内运动模式的非线性组合特征利用 Kernel CCA 进行分析,得到它们之间在时间和空间上的相关性信息,能够在低帧频、低分辨率的情况下实现对视频监控网络在语意、时间、空间上的理解和分析.

## 2 多摄像机视域语意区分割

在视频场景中,定义具有相似、相同运动模式的场景为一种语意区域,通常一个摄像机视域内有多个语意区域. 例如商店中的售货区、收银台和通道就是不同的语意区域. 将视频监控场景划分为多个具有不同语意的区域,对于视频分析具有重要意义. 对每个摄像机拍摄的视频进行处理时,本文采用文献<sup>[6]</sup>提出的背景建模方法建立背景模型. 每个固定时间间隔更新背景. 通过背景去除得到前景像素. 在实验中,摄像机分辨率为  $320 \times 240$ ,帧频为 5fps,共处理 6000 帧(20min).

将摄像机每个视域划分为若干个块 block(每个块大小  $10 \times 10$  像素). 一个运动目标覆盖多个块,在前景中,对每个块求出该块的质心  $p(i, j, t)$ .  $i, j$  分别表示该块在此摄像机视域内在第  $t$  帧上行和列的块索引值( $i \in [1, 24], j \in [1, 32], t \in [1, 6000]$ ,下同). 通过与上一帧在该块内的质心比较,来反映经过该块的目标的运动信息. 运动信息分为两类特征  $\langle f, s \rangle$ ,其中  $f$  是  $f(i, j, t)$  的简写,表示快速经过该块的物体运动速度.  $s$  是  $s(i, j, t)$  的简写,表示慢速经过该块的物体运动速度,也能表示在该块暂时停留、徘徊的运动目标信息:

$$\text{Mov}(i, j, t) = \|p(i, j, t) - p(i, j, t-1)\| \quad (1)$$

$$\begin{cases} f(i, j, t) = \text{Mov}(i, j, t), & \text{if } \text{Mov} > \alpha \\ s(i, j, t) = \text{Mov}(i, j, t), & \text{if } \text{Mov} \leq \alpha \end{cases} \quad (2)$$

式中, Mov 代表相邻两帧前景质心的运动速度,  $\alpha$  是区

分快速和慢速运动质心的经验值,这里取<sup>[3,5]</sup>.

视域内的语意区是由运动模式相似,空间距离较近的若干个块构成,构造一个近似矩阵  $A \in R^{N \times N}$ ,此处  $N$  是该摄像机视域中块(block)的总数,则:

$$A_{ij} = \begin{cases} \exp\left(-\frac{(\bar{r}_{ij}^f)^2}{2\sigma_i^f\sigma_j^f}\right) \exp\left(-\frac{(\bar{r}_{ij}^s)^2}{2\sigma_i^s\sigma_j^s}\right) \exp\left(-\frac{\|b_i - b_j\|^2}{2\sigma_b^2}\right), & \text{if } \|b_i - b_j\| \leq R \text{ and } i \neq j \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中  $\bar{r}_{ij}^f$  是指视域中第  $i$  块和第  $j$  块的动态特征  $f$  的相关性度量,而  $\bar{r}_{ij}^s$  是指视域中第  $i$  块和第  $j$  块的静态特征  $s$  相关性度量,  $\bar{r}_{ij}^f, \bar{r}_{ij}^s$  通过皮尔森相关系数求出,即  $\bar{r} = 1 - |r|$ .  $\bar{r}$  越接近 0,两个块之间的相关性就越大,  $\bar{r}$  接近 1 相关性就越小.  $R$  决定每个块与  $R$  邻域内的块计算相关性.  $[\sigma_i^f\sigma_j^f]$  以及  $[\sigma_i^s\sigma_j^s]$  分别是  $\bar{r}_{ij}^f$  和  $\bar{r}_{ij}^s$  的相关性尺度因子,是每个块与  $R$  邻域内所有块相关距离的均值.  $b_i$  和  $b_j$  是第  $i$  和  $j$  块的二维坐标值. 然后,归一化近似矩阵为  $\bar{A} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ ,其中  $D$  是对角矩阵,即  $D_{ii} = \sum_{j=1}^n A_{ij}$ . 本文采用文献<sup>[7]</sup>提出的聚类方法进行聚类,该方法根据归一化的近似矩阵  $\bar{A}$ ,以及初始设置的最大最小聚类数值自动选择最优聚类结果,本文初始聚类数值最小值和最大值分别取 3 和 15,对摄像机视域内的块(block)进行聚类,得到多个语意区域,完成语意区域的划分.

如图 1 所示,图中同一颜色区域为一个语意区域,用一个序号来标识,其中每个语意区域包含若干个块(block).



图1 摄像机语意区分割

由于后面要分析各个语意区域之间的相关性,本文通过对每个语意区域内各个块特征值的求和平均值作为该语意区的特征来表示该语意区域. 即:

$$\begin{cases} \hat{f}_n = \frac{1}{R_n} \sum_{b \in R_n^f} f \\ \hat{s}_n = \frac{1}{R_n} \sum_{b \in R_n^s} s, \quad n = 1, \dots, N \end{cases} \quad (4)$$

$\hat{f}_n, \hat{s}_n$  是指第  $n$  个区域的动态和静待特征.  $R_n$  是属于某一语意区的块的数目,  $N$  是该摄像机视域经过聚类后得到的语意区域总数. 在进行语意区相关性分析前, 通过去除低活动区域来使得语意区域相关性分析更有效和更快速. 低活动区域定义为: 如果某一个区域内含有一半以上的低活动块(如果某一块的  $f$  或  $s$  在对于时间  $t$  分量上的方差小于  $[2, 4]$  就定义其为低活动块).

### 3 多摄像机视域语意区 xKCCA 分析

对于视频监控网络内的任意一对语意区, 我们希望知道它们在时间和空间上的相关信息, 例如这对语意区是否相关以及相关性的大小. Kernel CCA 对于非线性组合的特征的分析具有很好的性能, 为了获得这些信息本文采用 Kernel CCA(即 KCCA)来分析视频区域之间的相关性, 并获得时间信息. 本文中的交叉核典型相关分析(Cross Kernel Canonical Correlation Analysis, xKCCA), 是指任意两个语意区在不同的时序间隔下计算此时对应的核典型相关系数值的一种方法. 通过改变时序间隔来得到两个语意区域在取最大相关系数值时对应的时序间隔值, 这个值就是相应语意区域之间最大概率的时间延迟.

本文采用的 KCCA 是 CCA 的非线性推广, 它借助核理论, 将待分析的变量映射到高维空间, 然后在这个高维空间中进行 CCA 的分析<sup>[17]</sup>. 假设待分析的两个语意区域的特征是  $X = \langle \hat{f}_{n-1}, \hat{s}_{n-1} \rangle$  和  $Y = \langle \hat{f}_n, \hat{s}_n \rangle$ , 并假设他们在时间间隔  $t \in [0, \tau]$ . 根据特征空间理论和核函数理论, 通过改变式(5)中  $t$  的取值  $[0, \tau]$ , 来达到在不同时间延迟  $t$  下,  $X$  和  $Y$  的 KCCA 的求解, 这个过程即是 xKCCA 分析的过程. 相关系数  $\rho_{x,y}$  定义为:

$$\rho_{x,y}(t) = \frac{\alpha_1^T \mathbf{K}_X \mathbf{K}_Y \beta_1}{\sqrt{\alpha_1^T \mathbf{K}_X \mathbf{K}_X \alpha_1} \sqrt{\beta_1^T \mathbf{K}_Y \mathbf{K}_Y \beta_1}}$$

$$T_{xy}^{\text{delay}} = \arg \max_{\tau} \frac{\sum_{\Gamma} \rho_{x,y}(t)}{\Gamma}$$

$$\hat{\rho}_{x,y} = \arg \max_{\tau} \frac{\sum_{\Gamma} (T_{xy}^{\text{delay}})}{\Gamma} \quad (5)$$

此处,  $\mathbf{K}_X, \mathbf{K}_Y$  分别是  $X$  和  $Y$  的核矩阵,  $\alpha_1$  和  $\beta_1$  分别是  $\langle \hat{f}_{n-1}, \hat{s}_{n-1} \rangle$  和  $\langle \hat{f}_n, \hat{s}_n \rangle$  核函数系数. 函数  $\Gamma = \min(\text{rank}(X), \text{rank}(Y))$ , 通过  $\Gamma$  函数可以在每个时刻得到唯一的相关系数值. 通过式(5)可以在每个时间延迟索引值  $t$  上取得唯一的一个相关性数值, 并能够求得  $X$  与  $Y$  在该时间索引值下的最大相关系数. 据此, 视频监控网络中的两个区域的相关性和时间延迟的大小已经得到, 同理可以求出视频监控网络的所有语意区的相关性和时间延迟, 构成语意区相关矩阵  $\mathbf{P}$  和时间延迟矩阵  $\mathbf{D}$  其中,  $\mathbf{P} = \{\mathbf{P}_{ij}\} \in R^{N \times N}$ ,  $\mathbf{P}_{ij} = \hat{\rho}_{x,y}$ ,  $\mathbf{D} = \{\mathbf{D}_{ij}\} \in R^{N \times N}$ ,

$\mathbf{D}_{ij} = T_{xy}^{\text{delay}}$ . 这里  $\mathbf{P}_{ij}$  反映任意两个语意区  $i$  和  $j$  的相关性的大小,  $\mathbf{P}_{ij}$  越接近于 1, 两个语意区相关性就越大. 而  $\mathbf{D}_{ij}$  用来衡量这两个语意区在时间延迟多久时两个语意区域达到最相关.

### 4 摄像机间的拓扑估计

摄像机间的拓扑关系用来分析视频监控网络中各个摄像机之间的时间和空间的逻辑关系, 有利于分析全局运动模式. 通过上面的工作, 已经得到摄像机网络所有语意区域之间的时间和空间相关性. 利用前文提到的矩阵  $\mathbf{P}$  和  $\mathbf{D}$  可以通过式(6)计算一个包含所有语意区的关联矩阵  $\Psi = \{\Psi_{ij}\} \in R^{N \times N}$  ( $N$  是指总的语意区数目), 用来量化任意两个语意区之间的关联性, 其中:

$$\Psi_{ij} = \frac{\hat{\rho}_{x,y}}{\sqrt{\hat{\rho}_{x,y}^2 + (1 - |\hat{\rho}_{x,y}|)^2}} \quad (6)$$

式中  $\hat{\rho}_{x,y}$  和  $T_{xy}^{\text{delay}}$  是经过归一化后的区域相关矩阵  $\mathbf{P}_{ij}$  和  $\mathbf{D}_{ij}$ .

根据语意区关联矩阵  $\Psi$  来计算摄像机网络关联矩阵  $\Phi = \{\Phi_{mn}\} \in R^{M \times M}$  ( $M$  是指总的摄像机数目). 现在假设摄像机网络中摄像机总数为  $M$ , 通过前文提出算法被划分的语意区数目为  $N$ , 其中第  $m$  个摄像机视域内被划分的语意区数目为  $L_m$ , 则  $N = \sum_{m=1}^M L_m$ . 第  $m$  个摄像机视域内的语意区分别是  $R_{m1}, R_{m2}, R_{m3}, \dots, R_{mL_m}$ , 第  $n$  个摄像机视域内的语意区分别为  $R_{n1}, R_{n2}, R_{n3}, \dots, R_{nL_n}$ . 首先根据语意区关联矩阵  $\Psi = \{\Psi_{ij}\}$  求得  $R_{m1}$  与  $R_{n1}, R_{n2}, R_{n3}, \dots, R_{nL_n}$  的相关性的最大值  $\Phi_{mn1}$ , 同理求得第  $m$  个摄像机内其他语意区与第  $n$  个摄像机所有语意区的相关性最大值, 得到  $\Phi_{mn1}, \Phi_{mn2}, \Phi_{mn3}, \dots, \Phi_{mnL_m}$  对其进行数值由大到小的排序后, 取前  $0.5 \cdot L_m$  个的平均值作为摄像机  $m$  与摄像机  $n$  的相关系数  $\Phi_{mn}$ . 得到  $\Phi = \{\Phi_{mn}\} \in R^{M \times M}$  后, 若  $\Phi_{mn} > \text{mean}(\Phi)$ , 则摄像机  $m$  和摄像机  $n$  是拓扑相关, 否则视此对摄像机不相关.

### 5 基于上下文目标再确认

当运动目标从一个摄像机消失后, 希望能够快速、有效地确定该目标再次出现在哪一个摄像机中, 确认其出现的区域. 目标再确认, 就是根据既定目标来在其他摄像机中迅速地找到这个目标. 然而在摄像机监控网络中, 目标的再确认主要有以下三个难点: 同一个目标在不同摄像机下表现的特征不同, 不同目标在不同摄像机下表现的特征可能相同以及目标消失和出现在时间关系上的不确定性.

通过在目标再确认过程中加入时间和空间相关信息, 可以减小搜索空间, 提高搜索的准确性. 很多有效

的特征会增加目标在确认的准确度如 SIFT、SURF、BOW 等,本文这里只以简单的颜色直方图匹配法来说明增加时空信息后,对于目标再确认的算法准确性的提高程度.具体来说,对于两个给定目标  $a$  和  $b$ ,他们的颜色直方图的 Bhattacharya 距离相似度分数为:

$$S_{bha} = \sum_{i=1}^{N_{bin}} \sqrt{H_a^i H_b^i} \quad (7)$$

每个颜色直方图总区间(bin)数为  $N_{bin} = 256$ . 实验中用彩色图像的总像素数目来归一化所有的区间. 然后分别计算 GRB 每个通道的 Bhattacharyya 距离分数,然后三个通道取平均值得到  $S_{bha}$ . 运动目标可能同时覆盖多个语义区域,计算目标覆盖各个语义区域的比率得到:

$$\mu = \{\mu_i | i = 1, \dots, N_r\} \\ \sum_{i=1}^{N_r} \mu_i = 1 \quad (8)$$

这些权重代表目标在不同的语义区域所占比例的大小,用这个比率利用文献[5]中提出的式(9)计算目标在不同语义区域之间的相关性和时间延迟:

$$\hat{\rho}^{a,b} = \sum_{i=1}^{N_r} \mu_i^a \left( \sum_{j=1}^{N_r} \mu_j^b \hat{\rho}_{x,y} \right) \\ \hat{T}^{a,b} = \sum_{i=1}^{N_r} \mu_i^a \left( \sum_{j=1}^{N_r} \mu_j^b \hat{T}_{x,y} \right) \quad (9)$$

式中  $\hat{\rho}_{x,y}$  和  $\hat{T}_{x,y}$  分别可以从  $P$  矩阵和  $D$  矩阵得到. 则对于两个目标之间总相似性量度为:

$$S_{final}^{a,b} = \begin{cases} S_{bha}^{a,b} \hat{\rho}^{a,b}, & \text{if } 0 < t_{gap}^{a,b} < \beta \hat{T}^{a,b} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

式中  $t_{gap}^{a,b}$  是指目标分别出现在两个语义区之间的时间差.  $\beta$  决定在多大的时间范围内对目标进行搜索和匹配.

## 6 实验结果与分析

本文选取室外两个强关联的摄像机 Cam1、Cam2 和室内两个弱关联摄像机 Cam3、Cam4 进行分析,室内的两个摄像机与室外的两个摄像机关联程度较低. 每个摄像机的分辨率为  $320 \times 240$ ,从每个摄像机记录的视频时间长度为 20 分钟的视频中,等间隔抽取 6000 帧.

### 6.1 语义区域分割

语义区域分割的结果如图 2 所示, Cam1 中的 1 区域是门口,行人从 1 区域出门,经 6、10 区通过 3 或 5 到达 Cam2 的 13 或 16 区. 从 Cam2 中可以看到,路口被准确的分开,图中 20 区是有学生在进行测量活动,被准确的分割出来. Cam3 和 Cam4 有行人经过的地面和门口区域被准确标示出. 同时在一些低活动区域如 Cam3 和 Cam4 的墙面被划分为低活动区域不进行标号和分析,同时显示出算法在室内和室外视域分割的鲁棒性.

### 6.2 摄像机网络拓扑估计

摄像机网络拓扑分析能够得到各个摄像机之间的

拓扑结构和摄像机之间的关联性. 图 3 分别是用 xCCA 和 xKCCA 得到的语义区的关联矩阵.

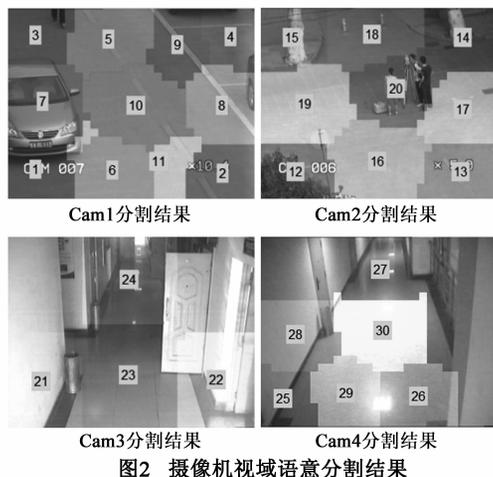


图2 摄像机视域语义分割结果

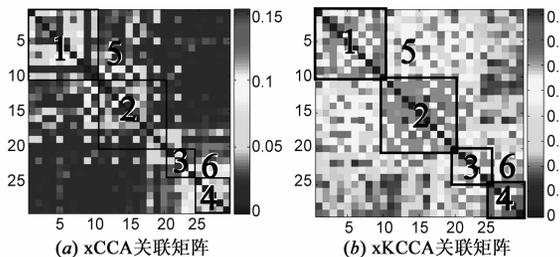


图3 xCCA和xKCCA关联矩阵对比

图 3(a)和(b)分别是 xCCA 和 xKCCA 所得到的关联矩阵. 横纵坐标都是摄像机网络中语义区域的编号. 摄像机 Cam1、Cam2、Cam3、Cam4 中含有的高活动语义区数目分别是 10、10、4、6 个. 上图中标号 1、2、3、4 分别包含了各个摄像机内部语义区之间的相关性,从实验结果来看本文算法中单摄像机视域内语义区域之间的关联性更加显著,更符合实际情况. 特别的,图中标号 5 区域内的数值表示 Cam1 和 Cam2 摄像机之间相应语义区之间的相关性,图 3(b)中区域之间相关性对比更加明显,相关系数大小区分更明显. 同样,图中标号 6 区域的数值表示摄像机 Cam3 和 Cam4 中语义区域之间的相关性.

### 6.3 基于上下文的目标再确认

本实验选用摄像机 Cam1 和摄像机 Cam2 来说明算法的有效性和准确性. 如前文所述,跨摄像机跟踪的难点在于:同一个目标在不同摄像机下表现的特征不同,不同目标在不同摄像机下表现的特征也可能相同,这对于目标匹配的影响较大. 把摄像机间时空关系加入到目标匹配算法中会使得目标匹配更精准. 具体来说,分别计算目标的颜色直方图,然后和待匹配目标用式(10)计算匹配分数(Bhattacharyya 距离分数)并降序排序,认为匹配分数高的为匹配到的目标. 实验中,1 号摄

像机内目标个数为 56 人, 2 号摄像机目标个数为 80 人, 其中有 42 人从 1 号摄像机进入 2 号摄像机中的行人. 图 4 中第一、第三行为颜色直方图匹配法 (CH) 进行匹配的结果, 第二、第四行为使用本文提出的方法 (CH + xKCCA) 匹配的结果, 左边第一列是 1 号摄像机中待检测目标 (实线框定), 对应右侧为待匹配目标. 图中第一、第三行中的 CH 算法匹配结果是第二列目标, 发生误匹配的原因是由于目标和准匹配目标的巴氏距离匹配分数高而被误判为最佳匹配目标, 实际目标应为对应虚线框定的目标. 第二、第四行的 CH + xKCCA 方法将时空信息融入后, 通过时间和空间的制约, 有效地将虚假目标排除, 达到准确匹配结果. 由第四行结果, 由于排除虚假目标, 使待匹配目标的数目减小, 加快匹配速度.

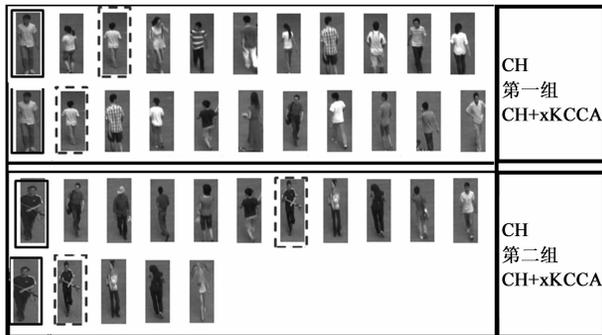


图4 目标再确认实验对比

## 7 结论

本文基于运动目标的运动信息, 将摄像机监控网络视域自动划分为多个语意区域, 这对于视频内容分析起重要作用, 采用 xKCCA 来分析不同语意区域在时间和空间上非线性组合特征的相关性, 得到了关联矩阵和相关时间矩阵, 并推导出摄像机网络的拓扑关系. 本文有效地解决了多摄像机网络的理解和分析的关键问题: 摄像机之间的时空关系分析、跨摄像机视域的运动目标再确认, 算法能够在低帧频和低分辨率情况下达到很好的效果.

## 参考文献

- [1] Zelniker E E, Gong S, Xiang T. Global abnormal behaviour detection using a network of CCTV cameras [A]. The Eighth International Workshop on Visual Surveillance [C]. Marseille: ECCV, 2008. 1-8.
- [2] Wang X, Tieu K, Grimson E L. Correspondence-free activity analysis and scene modeling in multiple camera views [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(1): 56-71.
- [3] Gray D, Tao H. Viewpoint invariant pedestrian recognition with

an ensemble of localized features [A]. Proceedings of the 10th European Conference on Computer Vision [C]. Berlin, Heidelberg: Springer-Verlag, 2008. 262-275.

- [4] Prosser B, Gong S, Xiang T. Multi-Camera matching using bi-directional cumulative brightness transfer functions [A]. Proceedings of British Machine Vision Conference [C]. Leeds: BMVA, 2008. 463.
- [5] Loy C C, Xiang T, Gong S. Incremental activity modeling in multiple disjoint cameras [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(9): 1799-1813.
- [6] Russell D, Gong S, Mary Q. Minimum cuts of a time-varying background [A]. Proceedings of British Machine Vision Conference [C]. Edinburgh: BMVA, 2006. 809-818.
- [7] Perona P, Zelnik-Manor L. Self-tuning spectral clustering [A]. Advances in Neural Information Processing Systems 17 [C]. USA: MIT Press, 2004. 1601-1608.
- [8] 王涛. 非重叠视域多摄像机目标匹配算法研究 [D]. 安徽合肥: 合肥工业大学, 2012, 4.  
Wang Tao. Research of the Object Matching Algorithm in Non-overlapping Multi-camera [D]. Hefei, Anhui: University of Technology of Hefei, 2012, 4. (in Chinese)
- [9] 王亮, 胡卫明, 谭铁牛. 基于步态的身份识别 [J]. 计算机学报, 2003, 26(3): 353-360.  
Wang Liang, Hu Weiming, Tan Tieniu. Gait-based human identification [J]. Chinese Journal of Computer, 2003, 26(3): 353-360. (in Chinese)
- [10] 刘亚楠, 吴飞, 庄越挺. 基于多模态子空间相关性传递的视频语义挖掘 [J]. 计算机研究与发展, 2009, 46(1): 1-8.  
Liu Yanan, Wu Fei, Zhuang Yueting. Video semantics mining using multi-modality correlation propagation [J]. Journal of Computer Research and Development, 2009, 46(1): 1-8. (in Chinese)
- [11] Liu Y, Wu F. Video semantic concept detection using multi-modality subspace correlation propagation [A]. Advances in Multimedia Modeling [C]. Berlin, Heidelberg: Springer-Verlag, 2006. 4351: 527-534.
- [12] 张鸿, 吴飞, 庄越挺. 跨媒体相关性推理与检索研究 [J]. 计算机研究与发展, 2008, 45(5): 869-876.  
Zhang Hong, Wu Fei, Zhuang Yueting. Cross-media correlation research and retrieval [J]. Journal of Computer Research and Development, 2008, 45(5): 869-876. (in Chinese)
- [13] Loy C C, Xiang T, Gong S. Multi-camera activity correlation analysis [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Miami, Florida, USA: IEEE, 2009. 1988-1995.
- [14] 田广, 戚飞虎. 移动摄像机环境下基于特征变换和 SVM 的分级行人检测算法 [J]. 电子学报, 2008, 36(5): 1024-1028.  
Tian Guang, Qi Feihu. Feature transformation and SVM based

hierarchical pedestrian detection with a monocular moving camera[J]. Acta Electronica Sinica, 2008, 36(5): 1024 - 1028. (in Chinese)

- [15] 耿玉亮, 须德. 一种鲁棒的摄像机运动分类算法[J]. 电子学报, 2006, 34(7): 1342 - 1346.

Geng Yuliang, Xu De. A Robust approach to camera motion classification[J]. Acta Electronica Sinica, 2006, 34(7): 1342 - 1346. (in Chinese)

- [16] Tieu K, Dalfy G, Grimson W E L. Inference of non-overlapping camera network topology by measuring statistical dependence[A]. Proceedings of Tenth IEEE International Conference of Computer Vision[C]. Beijing, China: IEEE, 2005, 2: 1842 - 1849.

- [17] 胡仿民. 基于改进核典型相关分析的人脸识别方法[J].

计算机工程与设计, 2009, 30(5): 1183 - 1185.

Hu Fangmin. Face recognition based on improved kernel CCA [J]. Computer Engineering and Design, 2009, 30(5): 1183 - 1185. (in Chinese)

- [18] 郑锦, 李波. 面向室外视频监控的背景重构算法[J]. 电子学报, 2009, 37(8): 1854 - 1859.

Zheng Jin, Li Bo. A background reconstruction algorithm for outdoor video surveillance[J]. Acta Electronica Sinica, 2009, 37(8): 1854 - 1859. (in Chinese)

- [19] Makris D, Ellis T, Black J. Bridging the gaps between cameras [A]. Proceedings of the 2004 IEEE Computer Society Conference of Computer Vision and Pattern Recognition [C]. Washington, DC: IEEE, 2004, Vol. 2: II-205-II-210.

## 作者简介



**蒋建国** 男, 1955年10月出生于安徽黄山. 教授, 博士生导师. 全国信息与电子学科研究生教育委员会、本科生教育委员会常务委员, 安徽省计算机学会副理事长. 研究方向为数字图像处理、智能信息系统.

E-mail: jjg@ah165.NET



**胡珍珍** 女, 1986年出生于安徽省淮北市. 2011年获得合肥工业大学硕士学位, 现为合肥工业大学在读博士, 主要研究方向是视频分析、姿态识别、计算视觉和多媒体计算.

E-mail: huzhen.ice@gmail.com



**顾占冰** 男, 1984年10月出生于黑龙江青冈县. 2008年于合肥工业大学获得学士学位, 读硕士研究生, 主要研究方向为目标跟踪、视频检索、内容分析.

E-mail: guzhanbing@gmail.com



**齐美彬** 男, 1969年出生. 博士, 现为合肥工业大学计算与信息学院教授. 主要研究方向为运动目标检测与跟踪、视频编码等.

E-mail: qimeibin@163.com