

# 基于全方位视觉的多目标跟踪技术<sup>①</sup>

严杭晨, 汤一平, 马宝庆, 田旭园, 吴立娟

(浙江工业大学 信息工程学院, 杭州 310023)

**摘要:** 为了快速准确地检测并跟踪多目标对象, 提出了一种基于全方位视觉的多目标对象跟踪方法. 首先采用全方位视觉传感器(ODVS)实时地采集现场 360° 全景视频图像; 接着融合运动历史图像算法(MHI)和运动能量算法(MEI)实现了快速高效的 MHoEI(Motion History or Energy Images)自动跟踪算法, 对多目标对象进行检测和跟踪; 最后, 本文采用面向对象技术融合目标对象进行匹配跟踪实验结果表明本文提出的方法能较好地跟踪多目标对象, 具有鲁棒性高、运算量小、便于硬件实现、高效等优点.

**关键词:** 多目标检测与跟踪; 数据关联; ODVS; 面向对象技术; MHoEI

## Multi-object Tracking Based on Omni-Directional Vision Sensor

YAN Hang-Chen, TANG Yi-Ping, MA Bao-Qing, TIAN Xu-Yuan, WU Li-Juan

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

**Abstract:** To rapidly and accurately detect and track multiple objects, this paper presents a method of multiple objects detection and tracking based on Omni-directional images. Firstly, getting 360-degree Omni-directional images by Omni-directional Vision Sensor in real time, which is an effective way to solve the problem of the multiple video source correlation. Secondly, the MHoEI algorithm that is created by combining Motion History Images algorithm with Motion Energy Images algorithm can efficiently detect and track multiple objects in this paper. Finally, the matching algorithm based on object oriented technology is proposed through fusing many properties of objects, which was used to identify different objects. The algorithm better solves the problem of data association of different properties. Experiment results show that the method in this paper can better track objects in complex background. The system has robustness, lightly computational load, high efficiency features.

**Key words:** multiple objects detection and tracking; data association; ODVS; object oriented technology; Motion History or Energy Images algorithm

随目标对象的跟踪是计算机视觉的一个重要研究方向, 实时而准确地从监控视频图像序列检测出目标前景并对检测出的目标前景进行稳定地跟踪是一些系统做下一步分析处理的重要前提, 如目标的分类与识别, 行为理解与描述等<sup>[1-5]</sup>. 在文献[6]中提出了一种 MHI(Motion History Image)的算法, 采用多帧视频图像进行分析, 可以实时获得运动目标区域及运动梯度方向. 文献[7]和文献[8]分别对 MHI 算法进行了改进, 实现了在一定条件下的多运动目标持续跟踪. 但是基于 MHI 方法的缺陷在于只能跟踪运动目标, 当目标静止时, 算

法会失去跟踪目标, 而当其再次运动时, 算法会误识别为新目标. 文献[9]提出了 CamShift 跟踪算法, 能够自动更新模型以适应目标在图像中的尺度变化, 能有效地跟踪目标对象. 文献[10]对现有的 CamShift 算法进行了改进, 在一定程度上消除了背景中相似颜色对目标的干扰, 可以更好地对目标进行跟踪. 基于 CamShift 的跟踪算法具有可以锁定静止目标的优势, 但跟踪目标时易受附近颜色干扰, 也难于获知目标的运动趋势, 且不含对各目标的图像分割, 属于半自动跟踪.

针对上述问题, 本文一种基于全方位视觉的多目

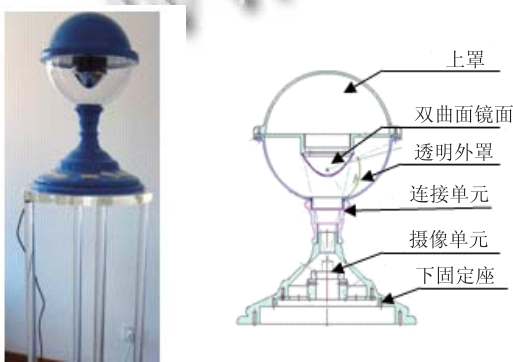
<sup>①</sup> 基金项目: 国家自然科学基金(61070134)

收稿时间: 2013-03-12; 收到修改稿时间: 2013-04-07

标对象跟踪方法. 首先, 通过 ODVS 采集场景的全景图, 可以避免多视频源(需要多个摄像装置并通过拼接才能获取全景图像)所带来的视频数据关联问题; 接着, 采用一种计算快速和高效 MHoEI 算法来实现进入场景目标对象的分割与定位, 从而实现对多目标对象的检测和简单跟踪; 最后, 本文提出了一种面向对象的多目标对象跟踪算法, 将实体目标对象的空间数据、属性数据及其行为封装为对象, 实现了目标对象的状态和属性数据的关联.

### 1 全方位视觉传感器(ODVS)

ODVS 由上罩、透明外罩、双曲面镜面、摄像单元、连接单元、下固定座等单元组成, 本文采用自主研发的 ODVS 作为全景图像获取的装置, 用于解决多视频源数据关联问题, 其实物图和结构图分别如图 1(a)和图 1(b)所示.



(a) 实物图 (b) 结构图

图 1 ODVS

ODVS 通过 USB 接口与 PC 相连, 将采集的 360 度视频图像通过 USB 接口传输给 PC 机, PC 机对全景图像进行展开、背景更新、目标对象提取和跟踪等处理. 由于篇幅有限, 关于 ODVS 的详细设计, 请参阅文献[11].

ODVS 采集的 360° 实时全景视频图像如图 2(a)所示. 为了便于后续的动态图像理解、分析等高层图像处理, 将得到的全景图通过公式(4)和公式(5)描述的柱状展开处理, 得到如图 2(b)所示的全景柱状展开图.



(a) 全方位原图



(b) 柱状展开图

图 2 全方位原图及柱状展开图

### 2 多目标对象的检测算法

实时性要求, 必须选择计算量小的算法, 但是目标跟踪另一个重要性能是跟踪的准确性, 而准确性的保证往往是在复杂运算和处理的基础上的, 这就造成了两者的矛盾. 一个好的运动目标跟踪系统必须要兼顾两者, 要对两个性能指标加以平衡. 因此, 目标对象的有效跟踪关键点是要求算法的计算快速和高效. 运动历史图像算法(MHI)和运动能量算法(MEI)的计算是递归, 只要存储当前最新信息, 因此这两种算法均符合计算快速和高效的要求.

#### 2.1 MHI 和 MEI 算法

MHI 算法中得到的运动历史图像是一段间隔中的相邻图像帧进行帧间差分 and 灰度化处理之后得到的, 这种得到运动历史图像的算法对于始终处于运动状态的目标可以很好地得到运动轮廓模板并且计算量比较小, MHI 算法能用于创建一个运动梯度图像并通过 SOBEL 算子计算得到运动梯度的方向和大小, 同时利用该结果可进一步用于估算对象的运动方向, 由于 MHI 算法不需要背景建模只进行帧间差分处理就可以得到前景运动对象, 因而该算法对于提取前景运动对象具有很高的实时性, 计算方法也极其简单, 这些是 MHI 算法的长处; 但是其短处是, 当处于运动状态的对象目标暂时性停止运动时, 存在着运动历史图像就无法得到问题; 另外当目标对象的整体处于暂时静止状态, 部分处于运动状态情况下, MHI 算法只能检测出那些处于运动状态目标对象的部分, 比如人体处于静止状态而双手的位置不断在变化, 这时 MHI 算法只能检测出活动的手. MHI 的更新方式如公式(6)所示:

$$H_{\tau} = \begin{cases} \tau \cdots \cdots \text{if } D(x, y, t) = 1 \\ \max(0, H_{\tau}(x, y, t-1) - 1) \cdots \text{otherwise} \end{cases} \quad (6)$$

式中,  $D(x, y, t)$  为运动区域的二进制图像序列,  $\tau$  为持续时间,  $H_{\tau}(x, y, t)$  为说明运动历史的二进制图像序列,  $H_{\tau}$  为运动历史图像; 从公式(6)可以看出

MHI 算法属于一种时间向后看算法。

此外, MHI 是一个标量值图像, 敏感于运动的方向, 灰度的变化体现出行为运动的方向; 最近运动的像素被设置为当前时间, 亮度值最大, 而运动较久的像素则被清除; 故 MHI 的生成很好地把运动的空间特征和时间信息都体现在图像中, 它既在空间上很好地描述了运动发生的位置和区域, 同时在时间上反映了运动是如何发生的。

MEI 算法中得到的运动图像是一段间隔中的相邻图像帧进行帧间差分图像求和处理之后得到的, 我们将其称为二进制累积运动能量图像; 这里定义  $I(x, y, t)$  图像序列,  $D(x, y, t)$  运动区域的二进制图像序列, 对于许多应用程序中差分图像能生成  $D(x, y, t)$ , 因此 MEI 算法的  $E_{\tau}(x, y, t)$  可以用公式(7)进行计算:

$$E_{\tau}(x, y, t) = \bigcup_{i=0}^{\tau-1} D(x, y, t-i) \quad (7)$$

式中,  $D(x, y, t)$  为运动区域的二进制图像序列,  $\tau$  为持续时间,  $E_{\tau}(x, y, t)$  为  $t$  时刻像素点  $(x, y)$  的灰度值, MEI 算法也是属于一种时间向后看算法; MEI 在空间上很好地记录和保持了目标对象运动发生的位置和区域。

MEI 和 MHI 算法分别具有两个不同向量图像编码的动作属性; 采用这两种跟踪算法的好处是, 由于它们的计算是递归的, 只有对于最新信息需要存储, 使得计算快速和高效; 如何融合这两种算法的长处是实现目标对象跟踪的实时性需求的关键。

对于运动目标对象跟踪, 在某些情况下 MEI 和 MHI 匹配的方法不一定是相同的, 这是两个不同的匹配标准; 因此, 需要给出匹配标准之间的区别, 即要区分在什么地方发生着运动和运动是如何进行的; MEI 算法主要解决目标对象在什么地方发生着或者发生过运动的问题, 而 MHI 算法主要解决目标对象是如何进行运动的问题。

## 2.2 MHoEI 跟踪算法

对于目标对象跟踪, 目标对象存在着运动和相对静止两种状态, 对于移动中的目标对象, 利用 MHI 算法并选择一个合适的持续时间是能较好的检测出来的, 但是处于暂时性停止运动状态时用 MHI 算法提取的目标对象随着时间推移将逐步消失; 而用累积运动能量图像 MEI 来描述目标对象处于暂时性停止运动状态时不会出现随着时间推移将逐步消失问题; 因此分别

采用 MHI 算法和 MEI 算法是难以实现对时而运动时而静止的目标对象跟踪的。

为了解决这个问题, 本文提出了一种计算快速和高效的跟踪算法, MHoEI(Motion History or Energy Images)算法, 用公式(8)表示:

$$H_{\tau} = \begin{cases} \tau \cdots \cdots & \text{if } D(x, y, t) = 1 \\ \max(0, H_{\tau}(x, y, t-1)) \cdots & \text{if } S \leq \delta \\ \max(0, H_{\tau}(x, y, t-1) - 1) \cdots & \text{otherwise} \end{cases} \quad (8)$$

式中,  $S$  为目标对象的运动速度,  $\tau$  为持续时间, 持续时间  $\tau$  需要根据目标对象运动速度  $S$  进行动态调整, 一般来说, 目标对象运动速度越快设定的持续时间  $\tau$  值越小, 反之目标对象运动速度越慢设定的持续时间  $\tau$  值越大; 当目标对象运动速度小于等于一个阈值  $\delta$  时对目标对象的运动像素点灰度值不作减 1 运算, 这样能使得处于静止状态的目标对象不会出现随着时间推移将逐步消失的问题. 这里目标对象的运动速度  $S$  可以根据感兴趣区域(ROI)的质心移动速度来得到. 同时, 我们也可以通过某个目标对象的 ROI 来计算得到该目标对象的各种形状状态属性和运动状态属性, 依托于这些属性信息可以很好地将运用于多目标的跟踪来区分多目标对象之间的差异, 实现目标对象的状态和属性数据的关联。

在某个实验场景中, 我们分别用 MhoEI 和 MHI 算法对前景目标进行提取, 如下图 3 所示:

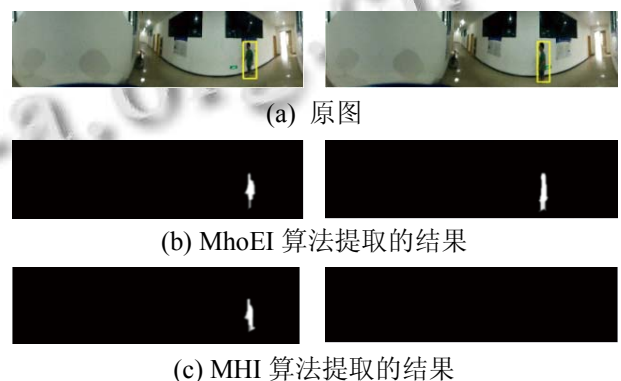


图 3 MHoEI(b)和 MHI(c)前景目标提取图

图 3 中(a)代表原图, (b)为 MhoEI 算法提取的结果, (c)为 MHI 算法提取的结果. 左列图像为目标在运动的状态, 右列图像为目标在静止的状态. 对比两个算法提取的效果图中, 我们可以看出 MhoEI 能较好的克服 MHI 无法提取静止前景目标的局限性, 而且 MhoEI 的提取前景目标轮廓精度相对较高。

### 3 面向对象的多目标对象跟踪算法

本文的研究目的是对进入全景视频场景中的多目标对象进行持续的跟踪,并获取其运动轨迹,并对各前景目标对象进行姿态、动作和行为的识别.当前景目标对象进入全景监控视频场景中时,由于其运动本质, MHoEI 的目标对象提取和跟踪算法会将其检测出来,并加以分割编号并实施跟踪.但是当多目标对象在场景中发生遮挡、分离和合并等情况时, MHoEI 跟踪算法会出现跟踪丢失问题,因此需要利用目标对象的属性进行匹配跟踪.从描述一个对象的角度来说,可以用大小、颜色、纹理、形状、空间位置和运动方向等方面的属性来描述其状态. MHoEI 的目标对象提取和跟踪算法能较好地获得前景目标对象的大小、形状、位置、运动方向、颜色概率密度等属性.

本文采用彩色直方图作为颜色特征模型和多目标对象出现分裂情况时第一匹配依据,具体做法是在多个对象合并前,将各个对象的色度信息保存在 **Object** 中;在多目标对象出现分裂情况时分别读取存储在 **Object** 中的相关目标对象的颜色直方图  $P_i = \{p_i^{(\mu)}\}_{\mu=1, \dots, m}$  数据;同时对分裂情况时的相关目标对象的 ROI 做颜色直方图  $Q_j = \{q_j^{(\mu)}\}_{\mu=1, \dots, m}$ , 将巴特查理亚距离作为两个彩色直方图相似性的度量,离散巴特查理亚系数的计算方法如公式(9)所示:

$$\rho[P_i, Q_j] = \sum_{\mu=1}^m \sqrt{p_i^{(\mu)} q_j^{(\mu)}} \quad (9)$$

式中,  $\rho[P_i, Q_j] \in [0, 1]$  为巴特查理亚系数,巴特查理亚距离用公式(10)进行计算

$$d = \min\{d_{i,j}\} = \min\{\sqrt{1 - \rho[P_i, Q_j]}\} \quad (10)$$

式中,  $d$  为巴特查理亚距离,如果该值小于某一规定的阈值 *Threshold* 就表示匹配成功.

接着采用 **Hu** 不变矩作为目标对象形状特征和多目标对象出现分裂情况时第二匹配依据,具体做法是在多个对象合并前,将各个对象的 **Hu** 不变矩特征的 7 个向量信息保存在 **object** 中;在多目标对象出现分裂情况时分别读取存储在 **object** 的相关目标对象中(合并前各目标对象)的 **Hu** 不变矩特征向量信息;同时对分裂情况时的相关目标对象做 **Hu** 不变矩特征向量运算,将欧氏距离作为两个 **Hu** 不变矩特征相似性的度量.

对于多目标跟踪算法的具体实现步骤是:首先是设计一个用于描述前景目标对象的类,对每个目标对

象以 16 位整数进行自动编号,每出现一个新的目标对象其编号就自动加 1,在程序中以该编号为对象参考名创建一个对象 (**Object**),令 **Object**{size, color, texture, shape, location, direction, old} 为对象参数结构,具体实现算法如下:

**Step1:** 用 MHoEI 算法对输入全景展开视频图像进行更新,对目标区域进行分割,得到 ROI(感兴趣区域)的个数  $i$ ;

**Step2:** 判断当前帧的 ROI 个数  $i$  与上一帧中已存在的目标对象 ROI 的个数  $j$  进行比较;

**STEP3:** 对于两个以上目标对象存在的情况下,如果的个数  $i$  等于已存在的目标对象的个数  $j$  表明在场景中的目标对象数没有发生变化同时也没有出现合并或者分离的情况下,接着转入 **Step4** 配对.如果当前 ROI 的个数  $i$  大于已存在的目标对象的个数  $j$  表明在场景中有新的目标对象进入或者出现了分离的情况,接着转入 **Step5** 处理步骤;如果的个数  $i$  小于已存在的目标对象的个数  $j$  表明在场景中已有的目标对象离开或者出现了合并的情况,接着转入 **Step6** 处理步骤;

**Step4:** 基于颜色特征模型和 **Hu** 矩匹配;接着 **Step7** 处理;

**Step5:** 首先采用 **Step4** 处理步骤进行配对计算,然后采用空间范围变化很小准则进行判断场景中有新的目标对象进入或者出现了分离的情况,接着 **Step7** 处理;

**Step6:** 首先采用 **Step4** 处理步骤进行配对计算,然后采用空间范围变化很小准则进行判断场景中有离开或者出现了合并的情况,接着 **Step7** 处理;

**Step7:** 返回 **Step1**,对下一帧进行处理.

在上述算法中采用了颜色、形状和姿态属性进行匹配计算,多目标对象的持续跟踪的鲁棒性有了很大的提升.如果加入运动方向和纹理特征等进行匹配计算,多目标对象的持续跟踪的鲁棒性会更高的提升,跟踪的出错率会降低,但这些检测也会消耗一些计算资源,这是下一步研究的工作.

### 4 实验研究

本文采用的硬件由奔腾双核 2.7G CPU、3GB 内存配置的 PC 机在 Windows XP 操作系统环境进行实验验证. ODVS 通过 USB 与 PC 相连,全景视频图像分辨率为  $640 \times 480$  像素.全景展开图图像分辨率为  $640 \times 180$  像素.

实验环境为 80 平方米大小的展示室, 将 ODVS 置于展示室的中间, 安置高度约为 1.8 米, 以获得整个空间内全景视频图像, 实验图如下, 图 4 为三人(分别为目标对象 A、目标对象 B 和目标对象 C)同时在展示室内的目标对象检测与跟踪实验结果. 各目标对象分别从运动到暂时静止、再从暂时静止到运动、变化各种人体姿态以及不停地来回走动, 但是实验中多目标对象在场景中没有发生遮挡、分离和合并等情况. 从图 8 所示的跟踪结果来看, 无论是暂时的停止走动(目标对象 A)和姿态的变化(目标对象 B), 还是改变前后行走方向(目标对象 C), 目标对象 A、目标对象 B 和目标对象 C 均能被有效地跟踪, 跟踪的 ROI 也比较准确, 这为人体姿态识别和行为识别奠定了良好的基础.

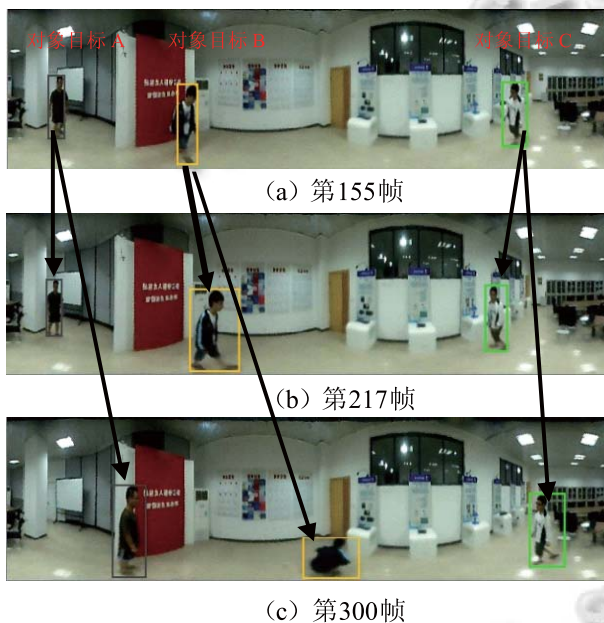


图 4 简单的多目标跟踪实验

图 5 为多目标对象发生合并分离的跟踪结果. 实验中安排将目标对象 A 与目标对象 B 首先进行合并, 如图 5(b)所示; 接着进行分离, 如图 5(c)所示.

在图 5(b)所示的合并情况下, 两个目标对象合并成一个新的目标对象, 在该新的目标对象中暂时保存了图 5(a)中红色跟踪框目标对象 B 和蓝色跟踪框目标对象 A 两个对象目标以及它们的一些状态属性; 当这两个对象发生分离时, 从暂时保存的状态属性中对分离后的对象进行颜色概率密度分布匹配, 区分分离后的两个目标对象中哪个是目标对象 A、哪个是目标对象 B; 图 5(c)所示的实验结果表明, 目标对象 A 仍然保

持着蓝色跟踪框, 目标对象 B 仍然保持着红色跟踪框.

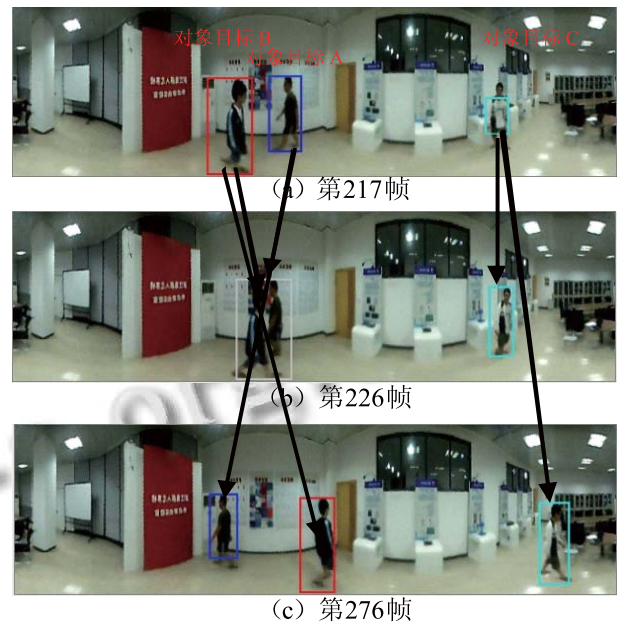


图 5 发生合并分离时的多目标跟踪实验结果

本实验中, MHoEI 算法表达式(8)中的速度  $\delta$  设置为 2 Pixel 效果较好, 从实验可以看出, 多个目标能够通过本文提出的 MHoEI 算法被快速和高效地检测并跟踪, 并且能够通过面向对象的多目标对象跟踪算法识别同一目标对象, 具有较好的适应性. 在三人的情况下整个多目标跟踪算法平均处理速度为帧数为 7 帧/秒, 基本达到了实时处理的速度. 本文提出的多目标跟踪技术有运算量小、跟踪范围有广、跟踪参数容易设置、鲁棒性高、跟踪策略融合容易等特点.

## 5 总结

本文采用一种基于 ODVS 的多目标跟踪技术, 通过 ODVS 实时获取场景全景视频图像, 很好地规避了多视频源的视频数据关联问题, 是一种十分理想的整体性视觉的解决方案; 在多目标跟踪方法方面, 模拟人类视觉的特点, 通过确认信息对象, 提出了一种计算快速和高效的 MHoEI 跟踪算法, 算法简单、利用递归计算方法, 只有对于最新信息需要存储, 使得计算快速和高效; 当多目标对象发生遮挡、合并和分离情况时, 采用面向对象的多目标对象跟踪算法来识别不同的目标, 解决了目标对象的多种状态属性数据关联问题. 本文将被跟踪目标对象的空间数据、属性数据及其行为封装为软件对象, 这种面向对象技术的多目

标跟踪技术具有较好的可扩展能力,可为今后的多目标对象的行为识别奠定良好的基础。

### 参考文献

- 1 Stauffer C, Grimson W. Learning patterns of activity using real time tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 2000,22(8):747-767.
- 2 Haritaoglu I, Harwood D, Davis LS. W4: real-time surveillance of people and their activities, *IEEE Trans. Pattern Anal. Mach. Intell.* 2000,22(8):809-830.
- 3 Amer A, Dubois E, Mitiche A. Real-time system for high-level video representation: application to video surveillance: Conference on Visual Communication and Image Processing (VCIP). Proc. of the SPIE International Symposium on Electronic Imaging. Santa Clara, USA, January 2003, 5022: 530-541.
- 4 Collins R, Lipton A, Kanada T, et al. A system for video surveillance and monitoring: VSAM final report, Technical report CMU-RI-TR-00-12, Carnegie Mellon University, May 2000.
- 5 Lv F, Kang J, Nevatia R, Cohen I, Medioni G. Automatic tracking and labeling of human activities in a video sequence. Proc. of the 6th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS04), Prague, Czech Republic, May 2004.
- 6 Davis J, Bobick A. The representation and recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(3): 257-267.
- 7 孟繁锋, 屈桢深, 曾庆双, 等. isMHI 结合 CAMSHIFT 的多目标跟踪. *计算机应用*, 2009, 29(6): 1673-1676.
- 8 GARY RB, JAMES WD. Motion segmentation and pose recognition with motion history gradients. *Machine Vision and Application*, 2002, 13(7): 174-184.
- 9 Aggarwal J, Cai Q. Human motion analysis: a review. *IEEE Wkshp. Nonrigid and Articulated Motion Workshop*, 1997: 90-102.
- 10 王亮亮, 孙即祥, 谭志国. 基于人脸检测和 CamShift 算法的人脸跟踪系统. *微计算机应用*, 2008, 29(2): 14-17.
- 11 汤一平, 叶永杰, 朱艺华, 顾校凯. 智能全方位视觉传感器及其应用研究. *传感技术学报*, 2007, 20(6): 1316-1320.
- 12 彭州, 赵保军. 基于 Contourlet 变换和模糊理论的红外图像增强方法. *激光与红外*, 2011, 41(6): 0635-0640.
- 13 刘会明. 基于模糊自适应算法的胸片图像增强研究. *电子测试*, 2011, 9: 39-42.
- 14 李刚, 桂预风, 肖新平. 一种改进的基于模糊对比度的图像增强方法. *湖北工业大学学报*, 2008, 23(1): 76-78.
- 15 姜桃, 赵春江, 陈明, 等. 自适应图像模糊增强快速算法. *计算机工程*, 2011, 37(19): 0213-0214.
- 16 欧阳琰, 陈先桥, 初秀民. 路面破损图像的模糊增强方法研究. *计算机工程与应用*, 2010, 46(10): 184-186.

(上接第 194 页)

- 8 钱苏敏, 张琳絮, 张云, 等. 基于小波阈值改进的去噪算法研究. *国外电子测量技术*, 2012, 31(5): 49-51.
- 9 惠飞, 史昕, 赵祥模, 等. 基于改进阈值函数的小波去噪算法研究. *现代电子技术*, 2011, 34(12): 61-64.
- 10 孟志强, 王鼎顺, 周华安. 小波图像消噪中阈值与信噪比的单峰规律与阈值试探方法. *电子测量与仪器学报*, 2006, 20(5): 63-67.
- 11 Pal SK, King RA. On edge detection of X-ray images using fuzzy sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1983, 5(1): 69-77.