

视频图像理解的一般性框架研究^{*}

梁英宏, 王知行, 曹晓叶, 许晓伟

(华南理工大学 计算机科学与工程学院, 广州 510640)

摘要: 视频图像理解侧重于对视频序列进行解释, 既涉及到图像的空间特性, 也涉及到视频序列的时间特性, 是目前计算机视觉领域的一个研究热点。回顾了视频图像理解方法的研究现状, 提出视频图像理解的一般性框架, 包括层次结构、涉及的技术领域和应用的系统结构, 并以一个实际应用作为示例解释该框架的层次结构。
关键词: 视频图像理解; 视频图像分割; 目标识别; 行为理解; 场景理解
中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2008)07-2203-05

Research on common framework of video image understanding

LIANG Ying-hong WANG Zhi-zhan CAO Xiao-ye XU Xiao-wei

(School of Computer Science & Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Video image understanding technology places emphasis on interpreting video sequences including spatial and temporal features of video sequences. It is a hot spot in the field of computer vision. This paper first reviewed the current status of research on video image understanding methods and then proposed a common framework for understanding video sequences including the layered structure, the correlative fields and the common system structure. Finally discussed a real application for interpreting the layered structure of this framework.
Key words: video image understanding; video image segmentation; object recognition; behavior understanding; scene understanding

引言

随着计算机硬件和软件技术的进步, 越来越多的人工操作转向计算机自动处理。在计算机视觉领域, 人们试图使计算机代替人的视觉感官, 对视觉信息进行加工处理, 实现自动理解。在人们所接收到的信息中, 视觉信息占到绝大多数, 而图像在人的感知中扮演着非常重要的作用, 人们随时随处都要接触图像^[1]。图像理解就是通过计算机方法对数字化的图像进行解释, 是计算机视觉领域的重要内容。

视频序列与一般的图像区别在于: 视频序列不是单独的一幅图像, 它是由一个图像序列组成, 也称做视频帧。视频图像理解过程往往不是针对单独一幅图像来进行的, 而是要综合考虑该图像所处的图像序列特征, 既涉及到图像的空间特性, 也涉及到图像序列的时间特性。静态图像理解主要是在空间层面上, 即在单幅图像内进行理解, 不可避免地要利用一些先验知识; 而视频图像理解总是对一系列时间关联的图像进行处理, 从中找到感兴趣的区域, 既有空间上也有时间上的操作, 具备一定的自学习功能, 如在视频中进行背景学习^[2]。在实际应用中, 对视频图像处理有很强的实时性要求, 所以在一般情况下, 视频图像的理解更偏重于时间特性。

目前, 越来越多的应用从对静态图像的处理转到对视频图像序列的处理。视频图像处理一般包含两个主要步骤: 视频图像捕捉和视频图像理解。前者主要涉及图像的压缩和传输; 后

者主要对视频图像进行分割和目标识别, 获取有用信息。视频图像理解的研究在最近十年得到人们的广泛关注, 其应用范围相当广泛, 最典型的例子是利用计算机和视频捕捉设备对场景进行监控^[3-4]。在自动监控过程中, 首先要克服视频图像的畸变以及场景和视频设备本身所带来的噪声, 区分前景和背景; 然后要识别前景, 判断其所属类别 (如是否是人类); 最后理解其行为。这三个过程对应了视频图像理解的三个层次。

视频图像理解的潜在应用包括:

①智能视频监控。目前, 越来越多的场合安装了监控设备。这些设备主要用于人控监视或视频录像, 通过将计算机与视频监控设备相结合, 实现对场景的自动监控。以地铁站为例, 通过统计入口人数, 可以细致了解某个地铁站的人流高峰低峰时间, 从而为发车间隔和人员管理作出指导; 通过估计地铁站内候车的人群密度, 可以为人员疏导作出指示; 通过监视月台上的安全指示线, 可以自动发出报警。

②人机交互。利用计算机和视频设备对人的行为进行理解, 如表情、姿态、手势等。传统的人机交互依赖于人的手动输入, 计算机无法理解人的行为; 基于视频图像的人机交互模式能够智能化地将人的行为映射成为控制命令, 从而减轻人与计算机对话的复杂程度。

③视频压缩。利用视频图像理解的视频压缩也可以称为基于视频内容理解的视频压缩方式, 这也是未来视频图像压缩领域的一个重要方向。目前的一个简单应用就是采用运动触

收稿日期: 2007-04-17; 修回日期: 2007-07-11 基金项目: 国家科技部科技型中小企业技术创新基金资助项目 (02C26214400224); 广东省科技计划资助项目 (2002A1020104)

作者简介: 梁英宏 (1978-), 男, 安徽蚌埠人, 博士研究生, 主要研究方向为图像处理、嵌入式开发 (Mhwir@yeah.net); 王知行 (1946-), 男, 浙江杭州人, 教授, 博导, 主要研究方向为计算机辅助设计、图像处理、模式识别; 曹晓叶 (1976-), 女, 讲师, 主要研究方向为计算机图形图像; 许晓伟 (1977-), 男, 博士研究生, 主要研究方向为计算机图形图像。

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

发录像的硬盘录像监控设备,通过运动检测(motion detection)来启动视频监控设备录像,从而大大减少视频存储空间。

d)运动提取和运动分析。提取视频图像中的运动目标,并跟踪其运动路径,分析其运动姿态。通过将运动物体从场景中抽离,可以将该运动物体映射到其他场景中,从而实现场景替换。此外运动抽取和分析还可以用于三维建模,在游戏、电影等产业有着广阔的发展前景。

e)机器人视觉。机器人通过单目或双目的图像传感器捕获场景信息,通过分析找到最优路径,避免与障碍物发生碰撞。

视频图像理解的应用范围远不止于此,然而在研发基于视频图像的应用时,往往只针对个案进行处理,难以形成一个通用的框架和技术路线。

回顾

有关视频图像理解方面的研究已经有数十年的历史,最早的研究是从运动检测和跟踪开始。由于在单幅图像中寻找运动物体存在一定的困难,转向视频序列的分析。早期由于计算机硬件的限制导致对视频序列的处理在速度上存在瓶颈,研究难以应用于实际,而计算机技术的进步使得实时视频处理成为可能,于是越来越多的研究偏向实际应用。视频监控是视频图像理解最重要的研究方向,其研究主要包括运动检测、运动跟踪和运动分析等。人机交互是视频图像理解的另一个重要研究方向,包括表情识别和动作识别等。通过对实际应用的总结,类似于图像理解的层次结构^[5,6],将视频图像理解分为三个主要层次,分别为视频图像分割、目标识别和场景/行为理解。图 1 描述了视频图像理解层次结构。

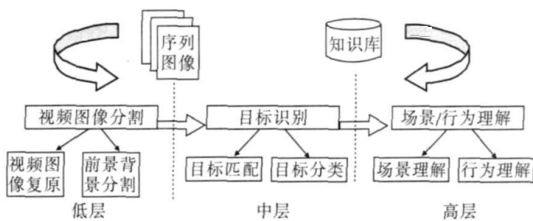


图 1 视频图像理解的层次结构

视频图像分割

视频图像分割主要是在一段图像序列中获取感兴趣的目标,也就是将前景与背景区分开来。视频图像分割效果的好坏,极大地影响了后续目标识别的精确度。一段时间以来,国内外学者对视频图像的分割方法作了大量的研究,研究的重点包括:如何在视频图像中精确地获得前景目标的轮廓;如何对背景进行更新,以适应环境光照和景物变化的影响;如何消除光线变化和背景噪声运动的影响。

在视频序列图像中分割前景和背景,是一个动态的处理过程。前景和背景在时间上是不断变化的,有效的分割方法必须能够适应这种变化。视频序列图像的动态分割也可以看做是背景的自适应更新。总结国内外的研究成果,目前比较通用的方法可以分成三大类:

a)滤波器方法^[2,7-9]是一种在时间序列上的连续预测方法,已知直到时间 t 的滤波结果,可以根据新的证据获得 $t+1$ 时刻的滤波结果。这个过程通常被称为递归估计(recursive estimation)。在视频图像分割中,通过滤波器方法估计每个像素点的背景值,用于分割前景背景。滤波器方法通常速度比较

快。用于背景更新的常用滤波器方法包括低通滤波^[8]、IIR滤波器^[7]及卡尔曼滤波器方法^[2]。

b)参数统计方法^[8,10,11]的实质在于从已知的数据集合中拟合出一个概率密度函数,然后用该函数进行预测。这种方法假定数据集合的概率模型是可以已知概率密度函数描述的,也就是说给定数据集合 X 以及已知的概率函数 $f(\cdot|\theta)$,寻找一个优化的 θ 来估计 θ 。参数的估计一般是通过最大似然估计方法完成。采用参数统计方法分割视频图像前景/背景,就是对图像中的每一个像素建立一个概率密度函数,并随着时间的推移不断更新其参数,利用该函数计算下一个时间像素值的概率,推断其是否属于背景。

最经典的背景参数统计方法就是采用无监督聚类的混合高斯方法^[10,11]。采用单个高斯分布估计背景值最大的问题在于把背景的分布看得过于简单,也就是认为背景的所有像素都是单模态分布的。在实际视频场景中,这种理想的状况很少,背景会受到各种干扰的影响,实际的分布很可能是多模态的,如闪烁的显示器或电视机一般是双模态的。文献^[11]中提出,背景中像素的多模态分布可以看做是一种混合高斯分布,也就是背景中的每个像素都属于一个或多个高斯分布。采用混合高斯方法更新背景的方法就是利用多个高斯分布反映背景像素变化的多模态性,将其变化纳入可统计范围。该方法成功地解决了一些复杂场景下的背景更新问题,但也存在计算量大的实时应用问题。

参数统计方法^[12-14]假设概率密度的参数形式是已知的,并以此为条件进行学习。然而现实世界是纷繁复杂的,往往很多实际的模式识别应用难以用经典的密度函数来概括。上面所提到的无监督聚类的混合高斯方法虽然能够解决背景的多模分布问题,但参与混合高斯的数量毕竟有限,增加其数量会导致计算量和存储空间增大。Ahmed Elgammal 给出一种基于核函数的非参数统计方法来分割背景^[13,14],并将其发展到颜色空间领域,获得了很好的效果,解决了部分背景噪声的影响。

还有一些其他方法是基于背景最大可能性像素值的方法^[15-17],包括侯志强^[15]和王俊明等人^[16]提出的基于像素归类的背景重构算法提出的背景渐进构建办法^[16]等。这些方法都是基于一个简单的假设:背景像素值是过去一段时间观察中出现频率最高的值。这些方法计算简单,但都难以消除背景噪声的干扰,一般都要通过图像腐蚀消除较小的噪声区。由于并没有对背景建立参数统计模型,这些方法可以笼统地归入非参数统计方法。

目标识别

目标识别是对分割出的前景目标进行识别,理解其所属的类别。例如判断前景区域是否属于人体目标,或者判断多幅序列图像中的目标是否属于同一目标。类似于视频图像分割,目标识别与图像分类也存在着时空操作与空间操作的差别。

目标识别包含两个主要内容:目标匹配和目标分类。

1)目标匹配 包括两种:单幅视频图像中的目标与已知的目标进行匹配,以及在视频序列中对目标进行连续匹配,即目标跟踪。对于视频图像理解来说,更侧重于后者。目标跟踪是对有效的目标(已分类的目标)在序列图像中进行匹配,从而保持对该目标位置的掌握。

2)目标分类 是在分割出的前景中提取有效的特征或特征向量,将目标分成不同的类别(如将前景中的人体与刚性物

体区分开), 属于无监督聚类; 或判断目标是否与已知的类别相同或相似, 属于有监督聚类, 需要利用先验知识。

总结有关目标匹配的研究成果, 在方法上主要分为基于匹配的方法和基于运动预测的方法。前者主要是利用前景目标的特性在不同图像中匹配目标; 后者主要利用物体的运动特性, 预测该目标在下一帧中的位置。

基于匹配的方法主要包括特征匹配方法^[18~22]和轮廓匹配方法^[23~24]。前者利用前景目标的一些特征点进行匹配, 这些特征点在目标运动中应当具有不变性, 如点^[18]、线^[19]、区域^[20, 21]、角点^[22]等; 后者利用前景目标的轮廓, 即边缘曲线来匹配。基于匹配的方法较简单, 但是存在的问题也很多, 主要是难以确保目标特征不会因为遮挡或物体的变形而变化, 甚至无法提取。

基于运动预测的方法^[25~28]主要是采用滤波器对目标在下一个时刻的运动状态进行估算, 利用估算出的运动状态信息, 如位置, 在一定范围内搜索目标, 进行匹配。这些方法主要包括卡尔曼滤波方法及其改进方法^[25]、序列蒙特卡罗方法^[26]、粒子滤波器方法^[27, 28]等。运动预测方法在精度上还存在一定问题, 特别是当多目标跟踪时, 目标难以辨别更会影响匹配的准确度。

有关目标匹配(跟踪)方法的详细介绍可以参阅文献^[29]。文献^[29]同时提出目标跟踪的三大难点: 鲁棒性、准确性及快速性问题。对于视频图像理解的应用来说, 往往对鲁棒性和准确性的要求不尽相同, 但都有速度上的要求, 即算法的实时性, 这就导致了关于目标匹配的研究成果难以应用到实际。在一些应用中, 通过多传感器融合(如可见光传感器和红外传感器融合)可以大大提高算法的鲁棒性和准确性, 而通过双目视频测距或三维建模也可以解决一些目标相互遮挡的问题。算法的实时性问题影响到实际应用, 是研究者应该着重解决的难点。

目标分类主要涉及模式识别一些学习算法, 实质是对运动物体的特征集合进行训练, 获得该目标本身或其所属类别的分类器, 用于目标聚类。神经网络、支持向量机等知识经常用于目标分类。

场景和行为理解

场景和行为理解处于视频图像理解的最上层。目前国内外的研究还处于初级阶段, 还有很多问题无法解决。行为理解最典型的一个例子就是人的动作理解。大量人体在场景中的运动倾向就构成了场景理解。

Haritaoglu等人^[30]建立了一个实时视频监控系统, 用于在视频序列中发现和跟踪人体目标, 并分析其运动, 理解其行为。所谓的 W4 即 who what where when。该系统可以发现(理解)人们是否随身携带物品, 以及互相交换物品、在场景中遗留物品这样的简单行为。

场景和行为的理解本身非常复杂, 涉及面很广。正确的理解必须有知识库作为支撑。理解通常会涉及到两种知识聚类过程: 有监督聚类和无监督聚类。有监督聚类事先将动作符号与行为含义对应, 通常是一个有监督学习过程, 训练阶段大量的事例样本被输入训练器, 获得动作的分类器; 训练完成的样本被用于动作与其含义的匹配。无监督聚类无须了解动作本身的含义, 通过自聚类, 将场景中的目标动作分成多个类别。

与目标分类类似, 各种模式识别的方法都可以用于场景和

行为理解, 隐马尔可夫模型是最常用的训练方法^[31, 32]。

视频图像理解的一般性框架

视频图像理解内容的层次结构

视频图像理解由于包含在时间序列的动态处理过程, 比对静态图像的理解更能反映人类视觉系统的工作模式。观察图 1 所示的视频图像理解的层次结构, 低层的视频图像分割接收原始的视频图像数据, 对图像进行复原操作, 然后将前景与背景分离, 分割出的前景用于下一步操作; 中层的目标识别是对前景信息的进一步分析, 理解前景目标的类别, 或者能够在不同的序列图像中找到该目标; 最高层的场景/行为理解是对目标在场景中的行为辨别, 或者对场景的状态作判断。图 2 显示了各个层次理解内容的差别。

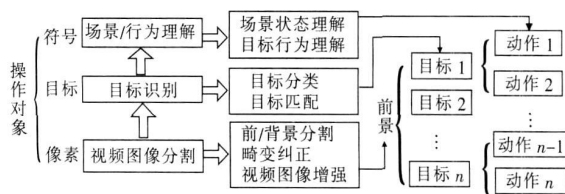


图 2 视频图像理解各个层次的内容差别

视频图像分割是在像素级别进行操作, 实际上是去除图像序列中无意义的像素, 将有意义像素聚类成集合表达, 即目标, 前景可能会对应多个目标。目标识别操作的对象是目标, 包括目标匹配或目标分类。目标分类可以是利用先验知识进行分类, 也包括在时间序列上的聚类; 场景/行为理解操作对象是符号。符号反映了目标的动作(或者说是目标的变化), 往往利用先验知识对动作进行识别, 确定其含义。还有一种情况是对场景中的动作进行聚类, 发现与众不同的行为。视频图像理解的各个层次具体内容包括:

a) 视频图像增强。对像素矩阵进行操作, 去除电子噪声。该操作往往在前/背景分割中采用统计学方法来解决。

b) 畸变纠正。对像素矩阵进行操作。畸变是指由于摄像机拍摄三维场景时的透视效应导致视频图像出现的透视畸变现象。在一些应用中, 畸变会导致目标识别或分类处理的误差, 可以采用相机标定的方法加以解决。

c) 前/背景分割。对像素矩阵进行操作, 取得有意义的前景信息。该操作是视频图像分割的主要目的, 也是视频图像分割的主要研究内容。

d) 目标匹配。对前景目标进行操作, 提取目标特征与已知的目标特征进行比较, 或者在序列图像中连续进行匹配(目标跟踪)。

e) 目标分类。对前景目标进行操作, 确定其类别。

f) 目标行为理解。对符号进行操作, 确定目标动作符号的含义, 需利用先验知识。

g) 场景状态理解。对符号进行操作, 将符号集合聚类, 解释这些符号集合的含义, 或者发现背离其他集合的孤立符号。

值得注意的是, 实际应用中, 以上步骤有时并不一定需要全部加以处理。例如在某些应用中, 并不需要非常精确的视频图像分割结果, 所以可以省去复原和畸变纠正操作。

视频图像理解的层次结构对于研究其方法, 以及应用于实

际项目都可以起到指导性的作用。

视频图像理解的技术领域

视频图像理解所涉及的技术范围很广，特别在解决实际问题时，还会涉及到其他相关领域的知识。视频图像理解侧重于在时间上动态地进行目标识别以及行为理解，其操作包含像素级别、目标级别及符号级别。每个级别的操作都涉及到相关技术领域的知识。视频图像理解与相关技术领域的关联如图 3 所示。

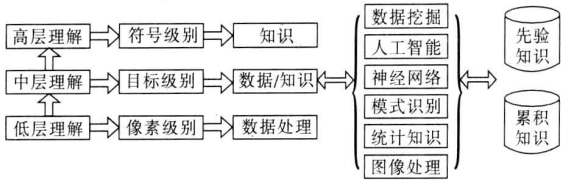


图 3 视频图像理解与相关技术领域的关联

图 3 中先验知识是经验知识或者训练所得的知识；累计知识是通过理解形成的知识。

低层理解是像素级操作，属于数据处理，主要应用到图像处理以及统计学方面的知识，所生成的目标表达式由于没有归类和解释，仍然属于数据；中层理解是目标级操作，处理的类型除了低层理解产生的数据外，还包括一些先验知识，中层理解可以产生新的知识；高层理解是符号级操作，是对知识进行操作，包括知识的搜索、匹配、分类等，可以积累产生新的知识。

基于视频图像理解的应用系统结构（图 4）

视频图像理解是随着计算机视频监控技术的进步而发展起来的，早期无论是视频采集设备还是计算机硬件都难以满足实时视频应用的需求，所以视频图像理解还停留在静态图像理解的层次上。随着技术的进步，产生越来越多的实时视频应用，视频图像理解也随之发展起来。研发一个基于视频图像理解的应用，要在视频图像理解的一般步骤的基础上，对具体问题具体对待，往往不是一定要完成全部三层的理解操作，但其系统的结构都有一定的共性。

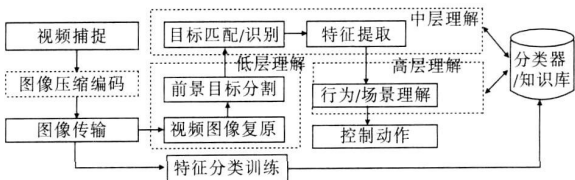


图 4 基于视频图像理解的应用系统结构

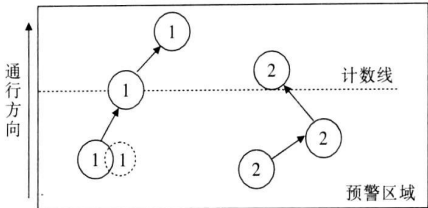
视频图像理解方法的应用实例——行人计数

本文考虑一个基于视频图像理解的行人计数问题。行人计数主要解决一些通行口的过往行人计数，如超市、地铁站出入口、公共汽车的上下门等。解决此类问题的关键是区分出多个前景目标，解决遮挡问题，进行精确计数。本文采用基于运动分析的方式统计人数。为简化问题，只考虑单向通行的情况，摄像机垂直拍摄。图 5 显示了在场景中设置一个虚拟计数线。

具体的处理步骤如下：

a) 视频图像分割。取出运动前景，由于场景相对静止，可以采用低通滤波方法更新背景；然后通过背景差取得每帧的前

景区域；考虑到行人阴影会导致后面的目标误检，还需要进行阴影去除。 $D=B+I+S$ 其中： D 表示视频图像、 B 表示背景、 I 表示前景、 S 表示阴影。



(圆形实线区域表示目标,圆形虚线区域表示目标阴影)

图 5 通行场景示意

(a)背景差获得前景区域 $I=D-B$ 更新 B

(b)采用数学形态学方法去除噪声；

(c)采用基于颜色空间转换的阴影去除方法^[33]去除阴影，最终 $I=I-S$

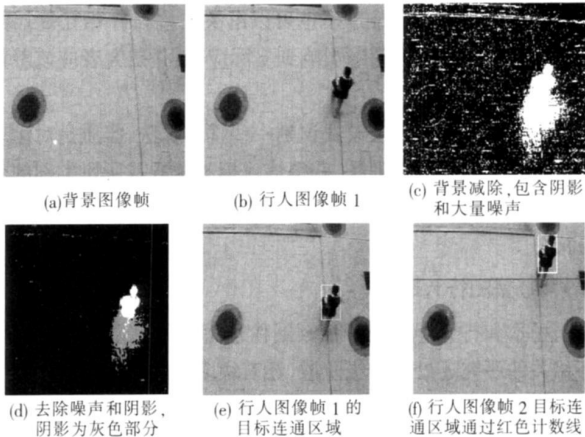
b)目标识别。分割前景为单个目标，提取目标特征以便目标匹配，这里可以提取目标的位置坐标、目标的外接矩形面积以及长宽比作为特征。

(a)检查前景中的每一个连通区域，提取目标特征；

(b)在下一帧中进行匹配。

c)场景/行为理解。在这个例子中，场景和行为理解内容相当简单，如果跟踪的目标位置越过计数警戒线的位置，则计数，即该目标的行为被理解成走过通行区域。在单位时间内越过警戒线目标的数量被认为是行人流量（客流量），可以以此为依据判断场景的拥塞状况。

图 6 演示了该应用算法中的主要步骤。



(a)~(d)属于视频图像分割步骤;(e)~(f)属于目标识别步骤;(f)中行人通过计数线则进行计数,为行为和场景理解的简单实例

图 6 基于视频图像理解的行人计数应用的算法演示

结束语

视频图像理解有着广泛的应用前景，也是目前的一个研究热点。由于其涉及较多技术领域，在解决具体问题时，往往没有一个通用的思路和方法参考。本文对视频图像理解方法作了整体的介绍，提出了视频图像理解的一般性框架，并对目前该领域的研究成果进行了阐述。对该领域的研究也来源于实际项目，笔者发现基于视频图像理解的应用大部分内容的解决方法都是可以放在一般性框架之内的。该领域中很多问题还有待解决。由于长期以来很多基于视频图像理解的应用都视具体问题采取不同的技术手段完成，难以形成一个统一的平

台和一般性的处理过程。总结目前工作的基础, 希望在下一步的工作中完善这个框架, 并对其中的一些难点进行研究。

参考文献:

- [1] 李弼程, 彭天强, 彭波, 等. 智能图像处理技术 [M]. 北京: 电子工业出版社, 2004.
- [2] RIDDER C MUNKELT Q K RCHNER H. Adaptive background estimation and foreground detection using Kalman-filter [C] // Proc of International Conference on Recent Advances in Mechatronics 1995 193-199.
- [3] 方帅. 计算机智能视频监控系统关键技术研究 [D]. 长春: 东北大学, 2005.
- [4] 郑江滨. 智能监视方法研究 [D]. 西安: 西北工业大学, 2002.
- [5] 章毓晋. 中国图像工程及当前的几个研究热点 [J]. 计算机辅助设计与图形学报, 2002 14(6): 490-500.
- [6] 张会章, 张利霞, 郭雷. 图像理解 [J]. 计算机工程, 2003 29(11): 23-24 71.
- [7] 皮文凯, 刘宏, 查红彬. 基于自适应背景模型的全方位视觉人体运动检测 [J]. 北京大学学报: 自然科学版, 2004 40(3): 458-464.
- [8] 王典. 基于混合高斯的背景建模与阴影抑制算法研究 [D]. 西安: 西北工业大学, 2006.
- [9] KARMANN K YON B A. Moving object recognition using an adaptive background memory [C] // Proc of Time-Varying Image Processing and Moving Object Recognition Amsterdam: Elsevier Science 1990.
- [10] STAUFFER C GRMSON W E L. Adaptive background mixture models for real-time tracking [C] // Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1999.
- [11] STAUFFER C GRMSON W E L. Learning patterns of activity using real-time tracking [J]. IEEE Trans on Pattern Analysis and Machine Intelligence 2000 22(8): 747-757.
- [12] MITTAL A PARAGIOS N. Motion-based background subtraction using adaptive kernel density estimation [C] // Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2004.
- [13] ELGAMMAL A DURA SWAMI R HARWOOD D et al. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance [C] // Proc of IEEE 2002.
- [14] ELGAMMAL A HARWOOD D DAVIS L. Non-parametric model for background subtraction [C] // Proc of European Conference on Computer Vision 2000.
- [15] 侯志强, 韩崇昭. 基于像素灰度归类的背景重构算法 [J]. 软件学报, 2005 16(9): 1569-1576.
- [16] 王俊明, 陈世旺. 渐进背景影像的建构 [J]. 国立台湾师范大学学报: 数理与科技类, 2002 47(2): 43-54.
- [17] GIL J M ENEZ P MALDONADO-BASC ON S GIL-PITA R et al. Background pixel classification for motion detection in video image sequences [J]. Computational Methods in Neural Modeling 2003 2686 718-725.
- [18] TESSANAYAGAN P SUTER D. Object tracking in image sequences using point features [J]. Pattern Recognition 2005 38(1): 105-113.
- [19] LIU Yun-qai HUANG T S. Determining straight line correspondences from intensity images [J]. Pattern Recognition 1991 24(6): 489-504.
- [20] MEYER F BOUTHEMY P. Region-based tracking using affine motion models in long image sequences [J]. CVGIP: Image Understanding 1994 60(2): 119-140.
- [21] BASCLE B DERICHE R. Region tracking through image sequences [C] // Proc of IEEE International Conference on Computer Vision 1995.
- [22] MIIRA P MURTHY C PAL S. Unsupervised feature selection using feature similarity [J]. IEEE Trans on Pattern Analysis and Machine Intelligence 2002 24(3): 301-312.
- [23] JAIN A K YU Zhong LAKSHMANAN S. Object matching using deformable templates [J]. IEEE Trans on Pattern Analysis and Machine Intelligence 1996 18(3): 267-278.
- [24] TERZOPOULOS D SZELESK IR. Tracking with Kalman snakes [M]. Cambridge: MIT Press 1992.
- [25] LI Pei-hua ZHANG Tian-wen MA Bo. Unscented Kalman filter for visual curve tracking [J]. Image and Vision Computing 2004 22(2): 157-164.
- [26] ARULAMPALAM M MASKELL S GORDON N et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking [J]. IEEE Trans on Signal Processing 2002 50(2): 174-188.
- [27] HUE C CADER J PEREZ P. Tracking multiple objects with particle filtering [J]. IEEE Trans on Aerospace and Electronic Systems 2002 38(3): 791-812.
- [28] 江宝安, 卢焕章. 粒子滤波器及其在目标跟踪中的应用 [J]. 雷达科学与技术, 2003 1(2): 170-178.
- [29] 侯志强, 韩崇昭. 视觉跟踪技术综述 [J]. 自动化学报, 2006 32(4): 603-617.
- [30] HARTIAGOGLU I HARWOOD D DAVIS L W4. real-time surveillance of people and their activities [J]. IEEE Trans on Pattern Analysis and Machine Intelligence 2000 22(8): 809-830.
- [31] WILSON A BOBICK A. Recognition and interpretation of parametric gesture [C] // Proc of International Conference on Computer Vision 1998.
- [32] YAMATO J OHYA J EH II K. Recognition human action in time sequential images using hidden Markov model [C] // Proc of IEEE Conference on Computer Vision and Pattern Recognition 1992.
- [33] SPAGNOLO R D O RAZIO T LEO M et al. Advances in background updating and shadow removing for motion detection algorithms [C] // Proc of CAIP 2005.
- [8] PAPPS C P KARACAPILD EN N J. A comparative assessment of measures of similarity of fuzzy values [J]. Fuzzy Sets and Systems 1993 56(2): 171-174.
- [9] 刘华文. 模糊模式识别的基础—相似度量 [J]. 模式识别与人工智能, 2004 17(2): 141-145.
- [10] SEINESM BABUSKA R. Fuzzy relation classifier trained by fuzzy cluster [J]. IEEE Trans on Systems Man and Cybernetics Part B: Cybernetics 1999 29(5): 619-625.
- [11] DUDAR Q HART P E STORK D G. 模式分类 [M]. 李宏东, 姚天翔, 等译. 2版. 北京: 机械工业出版社, 2005: 157-160.
- [12] GYSTAFSON D E KESSEL W C. Fuzzy clustering with a fuzzy covariance matrix [C] // Proc of IEEE Conference on Decision and Control and California [J], 1979 761-766.
- [13] 李士勇. 工程模糊数学及应用 [M]. 哈尔滨: 哈尔滨工业大学出版社, 2004 25-136.

(上接第 2202 页)