·综述·

智能监控系统下的老年人异常行为研究进展*

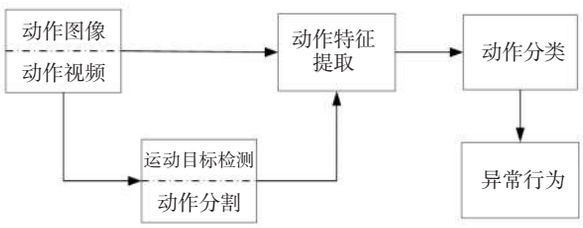
杨 鹏[1,2] 李潇婧[2] 张慧君[2] 王喜太[3]

智能监控系统下的老年人异常行为研究是近年的热点，是计算机视觉领域针对独居老年人的智能服务与及时监护技术，它融合了图像处理和模式识别等相关技术。对老年人异常行为的研究旨在当独自生活的老人因跌倒、突发疾病而处于异常行为状态时，计算机系统能及时发现并报警，让老人得到及时救助。

目前对老年人家庭监护的实现方法一般是利用摄像头、传感器等数据采集设备对老年人日常生活活动进行监测。利用传感器的家庭监护技术需要给老人佩戴穿戴式传感器或是在老人房间内安装各种监测传感器，前者老人容易忘记穿戴，后者依赖于精确的活动模型，随着季节、年龄的变化，活动模型需要及时修正，计算复杂度较高。而摄像头位置固定且无需佩戴，不易受外界因素干扰，近些年基于计算机视觉的老年人家庭监护的研究逐年增多。

基于计算机视觉的老年人家庭监护技术从采集的图像序列中检测动作数据并符号化动作信息，继而提取动作特征从而实现行为分类并判断是否有异常行为（图1）。大多数情况采用单个摄像头，有带红外功能摄像头[1]，全方位摄像头[2]和深度摄像头[3—5]，此外还有通过多个摄像头[6—7]进行多角度视觉采集。

图1 典型老年人异常行为识别流程图



1 老年人动作特征提取

动作特征提取通过从图像序列中抽取部分特征信息对动作进行表征，主要分为全局特征方法和局部特征方法，全局特征方法往往利用背景减除或人体检测技术对人体进行分割，并借助跟踪技术得到人体的剪影、光流、姿态和深度等特征对动作进行描述。局部特征方法通过局部区域的形状和运动信息提取时空兴趣点，避免了运动分割或者人体检测等预处理方法造成的错误[8]。见表1。

表1 老年人动作特征提取方法

类别	形式	代表例子
全局特征	剪影特征	近似椭圆、最小边界矩形、运动历史/能量图
	光流特征	光流
	姿态特征	3D姿态、骨架姿态
	深度特征	三维位置
局部特征	时空兴趣点	HOG、HOF描述符、SIFT描述符

1.1 全局特征

1.1.1 剪影特征：利用人体轮廓等信息来估计人体形态，利用形态序列来描述人体运动。通常采用背景减除法或差分图像法来检测运动区域，通过分析剪影区域的近似椭圆或最小边界矩形来分析人体形状变化。

Nguyen等[9]用背景减法提取人体目标，用近似椭圆及其外接矩形来标定目标对象，利用形状的长宽比特征和角度特征来判断站立、坐和跌倒状态。Vishwakarma等[10]通过在关键帧中分析轮廓方向的变化来识别老人的异常行为。Tang等[11]由最小边界矩形的长宽比来判断老人的姿态（如坐、躺、站立、蹲等）。孙媛媛等[12]采用头、左手、左脚、右脚、右手5个顶点对人体进行五边形拟合，采用面积有效比、周长比、中心比、宽高比来判断老人的行为。

Xiang等[2]利用运动历史或能量图像（MHoEI）对全方位摄像头中的目标进行跟踪，用最小边界矩形对目标进行标定，通过宽高比来判断姿态。Wu等[13]通过提取运动人体轮廓的运动能量图像（motion energy images, MEI），之后通过Hu矩算法提取人体的运动特征。李一波等[14]对4种动作（行走、跌倒、躺下、坐下）能量图边缘提取傅里叶描述子，然后结合动态能量描述子对跌倒这种异常行为进行识别。

DOI:10.3969/j.issn.1001-1242.2019.01.025
*基金项目：国家科技支撑计划课题（2015BAI06B03）
1 河北工业大学控制科学与工程学院，天津，300130；2 河北工业大学智能康复装置与检测技术教育部工程研究中心；3 国家康复辅具研究中心
作者简介：杨鹏，男，教授；收稿日期：2016-08-22

由于剪影不受颜色和纹理等特征的影响,且能对老人的生活隐私进行保护,目前被广泛采用。但是在复杂场景和有遮挡的情况下,无法获得准确的人体外观。

1.1.2 光流特征:不考虑人体的形状信息,而是直接从图像序列中提取出人体运动信息来表征运动状态。

Pal等[15]使用光流法创建运动图,通过简单的运动特征建模,使用神经网络将老人的日常活动划分为三类:高活动水平、低活动水平和无活动水平,依据活动水平来判断老人的健康状况。运动特征较适合中远距离和能见度低的情况,但是光流等方法计算较复杂且抗噪性能差。

1.1.3 姿态特征:人体动作最直接的定义是由一组关节位置的运动变化而组成的姿态序列,姿态特征比底层的表观和运动特征包含更高层次的语义,因此有助于后续的分类识别。

Meinel等[3]利用一个立体视觉传感器提供RGB-D数据,进行3D人体检测、目标识别和姿态估计最终形成人体行为识别系统。该系统的目的是检测老年痴呆患者的日常活动,使用收集的数据帮助患者(如提醒用药、喝水)并为医疗护理人员提供病情变化信息。Yao等[4]使用Kinect提取三维骨架数据,在得到的三维关节点的基础上,使用关节向量计算姿态特征,系统可以检测老人的六种行为(跌倒,喝水/吃饭,走,跑,坐和站)。

姿态特征的获取依赖于人体姿势估计的准确性,但是人体姿势估计本身也是一个还没有完全解决的问题。

1.1.4 深度特征:能描述人体运动过程中的三维位置信息,深度特征通常由深度相机获得,也可通过几个二维相机的组合获取。

Bian等[5]使用深度相机,采用一种改进的随机决策树(random decision tree, RDT)算法来提取三维人体关节。通过分析人体关键节点跟踪三维轨迹,根据头部的运动轨迹变化和位置来判断是否跌倒。Martinez等[16]提出了一种床匹配图(bed aligned map, BAM)描述符,从深度相机获得单人床校准表示,对高层次的场景理解用基于BAM的特征表示,对床上的老人进行不间断监测。

随着微软Kinect的发展,为提取姿态特征和深度特征提供方便,三维骨架模型还可适应遮挡等情况,但是Kinect的有效距离有局限性。

1.2 局部特征

局部特征通常采用时空兴趣点等信息得到特征描述算子,其不依赖底层的人体分割和跟踪,且对噪声和遮挡问题不敏感。但是它需要提取足够数量的稳定兴趣点,因此预处理阶段较重要。

Avgerinakis等[17]通过采集时空兴趣点的运动,基于梯度直方图(histogram of oriented gradients, HOG)和光流直方图(histograms of oriented optical flow, HOF)产生运动边界活动区(MBAA)的多尺度密集网格采样兴趣点和行为描述符,以获得外观和运动信息。此系统通过监测轻度老年痴呆症患者的日常生活,来分析他们的病情变化。Iqbal等[18]使用尺度不变特征转换(scale-invariant feature transform, SIFT)描述符提取运动(弯腰,跌倒,举手,站立)边缘和边界特征,SIFT对尺度、旋转以及一定视角和光照变化等图像变化都具有不变性,并且SIFT具有很强的可区分性。

2 老年人异常行为识别

目前对老年人异常行为识别的方法大多都是基于统计模型的方法,将人体序列的排列组合情况及各个时间点的动作进行相关的训练和分类。常用的方法有模板匹配法、距离分类法、状态空间法和状态统计法,见表2。

表2 老年人异常行为识别方法

类别	具体方法	特点
模板匹配法	形状模板、时序模板	计算量低,但受时间跨度影响较大
距离分类法	KNN、NN	根据样本与训练类别的最小距离进行分类
状态空间法	HMM	对时空微小变化鲁棒性好,但模型训练复杂
	SVM、BC	有监督,适合线性不可分情况
状态统计法	K-means、BOW	无监督,聚类,不需要输入标签,对异常数据敏感
	深度学习	分层学习,训练集越大精度越高

2.1 模板匹配法

模板匹配法是最简单直接的行为分类方法,常用的有基于剪影的形状模板或基于运动历史或能量图的时序模板,通过计算待测动作特征数据与样本模板的相似度来判断是否属于样本动作。文献[9—11]都是模板匹配法。模板匹配法需要一个稳定的大容量特征数据库作为基础,其计算量低,但时间跨度的变化对其影响较大。

2.2 距离分类法

K近邻分类器(K-nearest neighbors, KNN)和最近邻分类器(nearest neighbor, NN)都是根据样本与训练类别的最小距离进行分类。NN将样本与距离它最近的训练类别判定为同类,KNN依据样本中的K个近邻中的多数属于哪一类,样本则归为哪一类。Vishwakarma等[10]将提取最小边界矩形的长宽比特征通过KNN分类器检测跌倒或其他异常活动。李一波等[14]提出一种基于NN的分层识别方法,先根据四种行为(行走、坐下、躺下、跌倒)对应动作能量图的傅里叶描述子,用NN进行第一层分类,将跌倒和正常躺下归为一类,再把动态能量描述子带入NN进行第二层识别,将正常躺下和跌倒分类,从而识别异常行为。

2.3 状态空间法

状态空间法又称为基于概率网络的方法，这种方法对时间和空间尺度上的微小运动变化的鲁棒性较好，可以避免行为时间间隔建模，但模型训练复杂。常用的模型有隐马尔可夫模型（hidden Markov model, HMM）。Iqbal等[18]利用SIFT描述符提取运动边缘和边界特征得到特征向量，特征向量与从数据集中得到的HMM模型进行比较，从而识别异常行为。Farzad等[19]根据构建的星状骨架模型，提取姿态特征，根据各姿态序列训练HMM，最后将HMM作为分类器对5种行为（从床上跌落、从椅子上跌落、跌倒、弯腰、坐下）进行分类识别。

2.4 状态统计法

不同于状态空间法依赖于状态间的概率关系，状态统计法在全局状态下处理动作事件。典型的技术有支持向量机（support vector machine, SVM）、贝叶斯分类器（Bayes classifier, BC）、K均值聚类（K-means），词袋（bag of word, BOW）模型和深度学习等方法。

SVM、BC属于有监督的学习算法。SVM通过求解最优分类面进行分类，对线性不可分情况也适用，可以将样本映射到高维空间进行区分。BC是通过某对象的先验概率，利用贝叶斯公式计算出其后验概率，选择具有最大后验概率的类作为该对象所属的类。Wang等[20]使用SVM作为跌倒检测分类器。采用从人体轮廓提取的胡矩不变量和人体姿态信息来训练SVM。Wu等[13]根据由MEI提取的Hu矩特征和整个动作序列作为特征描述，然后使用BC进行分类和识别。

K-means等聚类算法都属于无监督的学习算法。词袋模型一般需要利用聚类算法先把相近的特征单词归为一类，再利用聚类结果组成词袋，最后通过类比统计计算出视觉对象的词袋模型中的词语序列。聚类算法不需要输入数据标签，但是对异常数据较敏感。Banerjee等[1]提出了一种扩展的模糊C均值聚类算法，基于泽尼克距（Zernike Moments）的GK（Gustafson Kessel）聚类算法。这种聚类算法中每个簇都有自己独特的协方差矩阵，在跌倒检测和跌倒风险评估方面有较好的效果。Martinez等[16]首先通过K-means聚类学习1000个词的码本模型，然后运用矢量量化（vector quantization, VQ）或局部约束线性编码（locality-constrained linear coding, LLC）和池化来获得每个序列的词袋表示，识别老人的异常行为。Jalal等[21]跟踪人的深度轮廓和特征产生人体骨架模型，通过23个节点的运动参数得到运动特征矢量，经K均值聚类和HMM模型训练，训练结束后系统生成生活日志。

深度学习是机器学习的新浪潮，其采用深层神经网络（deep neural network, DNN）模型，即包含多个隐藏层的人工神经网络，多层非线性结构使其具备强大的特征表达能力和对复杂任务建模的能力。深度学习方法也有监督学习与无监督学习之分，例如，卷积神经网络（convolutional neural network, CNN）属于监督学习、深度置信网（deep belief network, DBN）属于半监督学习，其具体步骤为无监督训练，有监督微调。Feng等[22]提出采用深度学习方法分析姿态的跌倒检测系统。首先提取人体剪影特征，经过后处理输入到DBN分类器。分类器首先由受限玻尔兹曼机（restricted Boltzmann machine, RBM）部分进行无监督的训练，直到所有的隐藏层初始化后，再进行有监督的训练。在对姿态进行训练分类后，可以确定老人在房间里的姿势。与传统方法比较，深度学习分类器具有较低的误检率和较高的检测率。

基于深度学习的分类方法有以下优点：通过分层学习，可以得到不同层次的不同因素间的相互影响；分层学习是从低水平到高水平层次结构的学习，符合自然的发展过程，更易于监控学习的内容，引导机制向更好的子空间发展，并且训练集越大精度越高。

3 存在问题和发展趋势

当前的老年人异常行为识别受到姿势、动作、衣着的个体差异和光照变化、遮挡、复杂背景的影响，给准确的特征提取和行为识别造成很大的困难。此外人体行为需要结合涉及的对象进行深层次的理解，然而在目前的研究中通过人工选择提取的动作特征，基本都属于人体行为的浅层特征。在近几年人体行为分析的研究中，有研究者通过深度学习建立特征的层次结构来获取更有效的数据表示[23]。例如，Le等[24]采用无监督的特征学习方式直接从视频中学习人体行为特征，采用层叠和卷积两种深度学习策略。Rajeswar等[25]采用3D CNN模型对监控视频中的人体动作进行识别。这两种基于深度学习的特征提取方法在人体动作识别中都得到了很好的效果。

随着智能监控系统的发展，视频资源越来越多，在视频大数据发展背景下，基于深度学习的老年人动作特征提取和异常行为分析的研究将有可能成为未来的发展趋势，将会从老年人日常活动中得出更准确的行为理解。

参考文献

[1] Banerjee T, Keller JM, Skubic M, et al. Day or night activity recognition from video using fuzzy clustering techniques [J]. IEEE Transactions on Fuzzy Systems,2014, 22(3): 483—493.

[2] Xiang Y, Tang YP, Ma BQ, et al. Remote safety monitoring for elderly persons based on omni-vision analysis[J]. PLoS One, 2015, 10(5):e0124068.

[3] Meinel L, Richter J, Schmidt R, et al. OPDEMIVA: An integrated assistance and information system for elderly with de-

mentia [C]// Consumer Electronics (ICCE'15). 2015: 76—77.

[4] Yao B, Hagras H, Alghazzawi D, et al. A Big Bang-Big Crunch Type-2 Fuzzy Logic System for Machine Vision-Based Event Detection and Summarization in Real-world Ambient Assisted Living [J]. IEEE Transactions on Fuzzy Systems, 2015,99:1—13.

[5] Bian ZP, Hou J, Chau LP, et al. Fall detection based on body part tracking using a depth camera[J]. IEEE J Biomed Health Inform, 2015, 19(2):430—439.

[6] Chaaraoui AA, Padilla-López JR, Ferrández-Pastor FJ, et al. A vision-based system for intelligent monitoring: human behaviour analysis and privacy by context[J]. Sensors (Basel), 2014, 14(5):8895—8925.

[7] Hung DH, Saito H, Hsu GS. Detecting Fall incidents of the elderly based on human-ground contact areas [C]// IAPR Asian Conference on Pattern Recognition (ACPR).2013: 516—521.

[8] 李瑞峰,王亮亮,王珂.人体动作行为识别研究综述[J].模式识别与人工智能,2014,(1):35—48.

[9] Nguyen VD, Le MT, Do AD, et al. An efficient camera-based surveillance for fall detection of elderly people [C]// Industrial Electronics and Applications (ICIEA'14).2014: 994—997.

[10] Vishwakarma DK, Kapoor R, Maheshwari R, et al. Recognition of abnormal human activity using the changes in orientation of silhouette in key frames [C]// Computing for Sustainable Global Development (INDIACom). 2015: 336—341.

[11] Tang YP, Ma BQ, Yan HC. Intelligent video surveillance system for elderly people living alone based on ODVS [J]. Advances in Internet of Things, 2013, 3: 44—52.

[12] 孙媛媛,王春才,李哲.一种直方图均衡化的五边形拟合老年人异常行为检测算法的研究[J].长春理工大学学报(自然科学版),2014,(2):128—132.

[13] Wu DM , Wang L, Wang J, et al. Solitary oldies abnormal action recognition based on MEI[C]//Signal Processing, Communications and Computing (ICSPCC), 2015 IEEE International Conference on. IEEE, 2015: 1—4.

[14] 李一波,丁裕林.基于智能监控的独居老人室内异常行为检测[J].计算机应用与软件,2014,31(2):188—190.

[15] Pal S, Feng T, Abhayaratne C. Real-time recognition of activity levels for ambient assisted living [C]// Consumer Electronics - Berlin (ICCE-Berlin). 2015: 485—488.

[16] Martinez M, Rybok L, Stiefelhagen R. Action recognition in bed using BAMs for assisted living and elderly care [C]// Machine Vision Applications (MVA). 2015:329—332.

[17] Avgerinakis K, Briassouli A, Kompatsiaris I. Recognition of activities of daily living for smart home environments [C]// Intelligent Environments (IE).2013: 173—180.

[18] Iqbal JLM, Lavanya J, Arun S. Abnormal human activity recognition using scale invariant feature transform[J]. International Journal of Current Engineering and Technology, 2015, 5 (6):3748—3751.

[19] Farzad A, Asli RN. Recognition and classification of human behavior in intelligent surveillance systems using hidden markov model[J]. I.J. Image, Graphics and Signal Processing, 2015,12:31—38.

[20] Wang RD, Zhang YL, Dong LP, et al. Fall detection algorithm for the elderly based on human characteristic matrix and SVM [C]// Control, Automation and Systems (ICCAS).2015: 1190—1195.

[21] Jalal A, Kamal S. Real-time life logging via a depth silhouette-based human activity recognition system for smart home services [C]// Advanced Video and Signal Based Surveillance (AVSS). 2014: 74—80.

[22] Feng PM, Yu M, Naqvi SM, et al. Deep learning for posture analysis in fall detection [C]// Digital Signal Processing (DSP). 2014: 12—17.

[23] 雷庆,陈锻生,李绍滋.复杂场景下的人体行为识别研究新进展[J].计算机科学,2014,(12):1—7.

[24] Le QV, Zou WY, Yeung SY, et al. Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis [C]// Computer Vision and Pattern Recognition (CVPR). 2011:3361—3368.

[25] Rajeswar MS, Sankar AR, Balasubramaniam VN, et al. Scaling up the training of deep CNNs for human action recognition [C]// Parallel and Distributed Processing Symposium Workshop (IPDPSW). 2015: 1172—1177.