Improving Human Action Recognition Using Fusion of Depth Camera and Inertial Sensors

摘要：本文提出了一种融合方法，该方法基于两个不同的模态传感器（包括深度摄像头和惯性人体传感器）来改善人类动作识别。从深度相机提供的深度图像和惯性人体传感器提供的加速度计信号中提取出计算有效的动作特征。这些功能包括深度运动图和统计信号属性。对于动作识别，使用协作表示分类器来检查特征级融合和决策级融合。在特征级融合中，将从两个不同模态传感器生成的特征在分类之前合并，而在决策级融合中，DempsterShafer理论用于组合来自两个分类器的分类结果，每个分类器对应一个传感器。引入的融合框架是使用伯克利多模式人类行为数据库进行评估的。结果表明，由于来自这些传感器的数据的互补性，根据每个传感器单独使用时的情况所采取的措施，引入的融合方法可将识别率提高2％至23％。

01: A kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities

摘要：这项研究评估了在公立学校中使用基于Kinect的系统修复两名运动障碍年轻人的可能性。这项研究是根据ABAB序列进行的，其中A代表基线，B代表干预阶段。数据显示，两名参与者大大提高了他们进行身体康复的动机，从而在干预阶段改善了运动表现。讨论结果的实际和发展意义。

02: Home-based senior fitness test measurement system using collaborative inertial and depth sensors,

摘要：本文提出了一种通过协作使用惯性传感器和深度相机的家庭高级健身测试（SFT）测量系统。深度摄像头用于监控对象的正确姿势以进行体能测试，以及与正确姿势的任何偏差，而惯性传感器用于测量对象在指定的持续时间内进行的体能测试动作的次数。健身协议。结果表明，这种协作方法可在现实条件下提供SFT测量结果时获得很高的成功率。

03: Depth video-based human activity recognition system using translation and scaling invariant features for life logging at smart home

摘要：基于视频的人类活动识别系统对智能家居和医疗保健服务等各种应用具有潜在的贡献。在这项工作中，我们提出了一种利用深度轮廓的R变换的新颖的基于深度视频的翻译和缩放不变人类活动识别（HAR）系统。为了在室内环境中执行HAR，不变的HAR方法对于在摄像机视图中的任意位置自由执行活动而不会出现人体轮廓的平移和缩放问题至关重要。我们通过对深度轮廓进行R变换来获得这样的不变特征。此外，在R变换深度轮廓中，将在深度值中反映的人体的形状信息编码为特征。在R变换中，首先通过每个深度轮廓的Radon变换来计算2D特征图，然后通过R变换计算1D特征轮廓以获取平移和缩放不变特征。然后，我们应用主成分分析（PCA）进行降维，并应用线性判别分析（LDA）使特征更加突出，紧凑和健壮。最后，隐马尔可夫模型（HMM）用于训练和识别不同的人类活动。我们提出的系统显示出优于常规方法的识别率，六种典型人类活动的平均识别率高达93.16％，而基于PC和IC的常规深度轮廓仅达到74.83％和86.33％，而基于二进制轮廓的R转化方法分别达到67.08％。我们的实验结果表明，该方法在识别人类日常活动中具有鲁棒性，可靠性和高效性。

04: A medication adherence monitoring system for pill bottles based on a wearable inertial sensor,

摘要：本文提出了一种基于可穿戴惯性传感器的药瓶药物依从性监测系统。使用摄像头辅助训练阶段创建与旋盖和手到嘴这两个动作相对应的信号模板。然后，通过在信号模板与可穿戴惯性传感器获取的信号之间实时执行移动窗口动态时间扭曲来识别药丸摄入行为。进行的实验结果表明，已开发的医学依从性监控系统可以高度准确地识别药丸摄入的行为。

05: Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions

摘要：体育活动对人们的福祉具有积极影响，也可能减少慢性疾病的发生。具有可穿戴传感器的活动识别可以向用户提供有关他/她关于体育活动和运动的生活方式的反馈，从而促进更加积极的生活方式。到目前为止，活动识别主要是在有监督的实验室环境中进行的。这项研究的目的是检验与无监督环境相比，无监督环境下受试者的日常活动和体育运动如何被识别。通过使用混合分类器，将包含先验知识的树结构和人工神经网络相结合，并通过使用三个参考分类器，可以识别活动。从12名受试者中收集了68小时的活动数据，其中对活动进行了21小时的监督，对未进行47小时的监督。活动是根据臀部和手腕上的3-D加速度计的信号特征以及GPS信息进行识别的。这些活动包括躺下，坐着，站着，步行，跑步，骑健身车骑自行车，用划船机划船，踢足球，越野行走以及骑普通自行车骑车。使用监督和非监督数据进行活动识别的总准确性为89％，仅比使用监督数据进行活动识别的准确性低1％。但是，当仅使用监督数据进行训练而仅使用非监督数据进行验证时，准确性降低了17％单位，这强调了在活动识别系统开发中实验室外数据的必要性。结果支持在现实生活中认识到更广泛的领域和更复杂的活动的愿景。

06: Action recognition based on a bag of 3D points,

摘要：本文提出了一种从深度图序列中识别人类动作的方法。具体来说，我们使用动作图来显式地模拟动作的动力学，并使用一袋3D点来表征与动作图中的节点相对应的一组显着姿势。此外，我们提出了一种简单但有效的基于投影的采样方案，以从深度图中采样3D点包。实验结果表明，仅从深度图中采样大约1％的3D点即可达到90％以上的识别精度。与基于2D轮廓的识别相比，识别错误减少了一半。此外，我们通过仿真演示了袋状姿势模型处理遮挡的潜力。

07: Recognizing actions using depth motion maps-based histograms of oriented gradients,

摘要：在本文中，我们提出了一种有效的方法，该方法可以从深度图序列中识别人的动作，从而为动作识别提供更多的身体形状和动作信息。在我们的方法中，我们将深度图投影到三个正交平面上，并通过整个视频序列累积全局活动以生成深度运动图（DMM）。然后从DMM计算定向梯度直方图（HOG）作为动作视频的表示。Microsoft Research（MSR）Action3D数据集上的识别结果表明，尽管我们的表示形式更加紧凑，但我们的方法明显优于最新方法。此外，我们调查了在我们的框架中需要多少帧才能识别MSR Action3D数据集上的动作。我们观察到，短的30-35帧子序列足以获得与在整个视频序列上运行的结果相当的结果。

08: Robust 3D action recognition with random occupancy patterns,

摘要：我们从深度相机捕获的深度序列中研究动作识别问题，其中噪音和遮挡是常见问题，因为它们是用单个商用相机捕获的。为了解决这些问题，我们提取了称为随机占用模式（ROP）特征的半局部特征，该特征采用了一种新颖的采样方案，可以有效地探索极大的采样空间。我们还利用稀疏编码方法对这些特征进行鲁棒性编码。所提出的方法不需要仔细的参数调整。由于使用了高维积分图像，因此其训练非常快，并且对遮挡具有鲁棒性。我们的技术在商品深度相机捕获的两个数据集上进行了评估：动作数据集和手势数据集。我们的分类结果优于在两个数据集上通过最新技术方法获得的分类结果。

09: HON4D: Histogram of oriented 4D normals for activity recognition from depth sequences,

摘要：我们提出了一种新的描述符，用于从深度传感器获取的视频中识别活动。以前的描述符大多独立地计算形状和运动特征，因此，它们通常无法捕获像素级的复杂关节形状运动线索。相反，我们使用直方图描述深度序列，该直方图捕获时间，深度和空间坐标的4D空间中表面法线方向的分布。为了构建直方图，我们创建了4D投影仪，该投影仪对4D空间进行量化并表示4D法线的可能方向。我们使用常规的polychoron的顶点初始化投影仪。因此，我们使用可区分的密度度量来优化投影仪，以便在4D法线更加密集和可区分的方向上引入其他投影仪。通过广泛的实验，我们证明了我们的描述符可以更好地捕获深度序列中的关节形状运动线索，因此在所有相关基准上均优于最新技术。

10: Spatio-temporal depth cuboid similarity feature for activity recognition using depth camera,

摘要：本地时空兴趣点（STIP）以及RGB视频所产生的功能已被证明在活动识别方面很成功，可以处理杂乱的背景和部分遮挡。在本文中，我们在深度视频中提出了它的对应部分，并展示了其在活动识别方面的功效。我们提出了一种从深度视频中提取STIP（称为DSTIP）的过滤方法，该方法可以有效地抑制噪声测量。此外，我们建立了一个新颖的深度长方体相似度特征（DCSF），用于描述DSTIP周围具有自适应支撑尺寸的局部3D深度长方体。我们使用公共MSRAction3D，MSRDailyActivity3D数据集和我们自己的数据集在活动识别应用程序上测试此功能。实验评估表明，该方法优于深度视频方面的最新活动识别算法，并且该框架比现有方法具有更广泛的适用性。我们还将与其他功能进行详细比较，并对参数选择进行分析，以作为应用指南。

11: Human daily activity recognition with sparse representation using wearable sensors

摘要：使用移动个人感应技术进行的人类日常活动识别在普及医疗保健领域中发挥着核心作用。一个主要的挑战在于人体运动固有的复杂性以及人们进行某种活动时的各种风格。为了解决这个问题，在本文中，我们提出了一种新的人类活动识别框架，该框架基于最近开发的使用可穿戴惯性传感器的压缩感知和稀疏表示理论。我们的方法将人类活动信号表示为训练集中所有活动类别的活动信号的稀疏线性组合。活动信号的类别成员资格通过解决l1最小化问题来确定。我们通过识别14位受试者执行的9种最常见的人类日常活动，实验性地验证了基于稀疏表示的方法的有效性。我们的方法实现了96.1％的最大识别率，这比基于最近邻居，朴素贝叶斯和支持向量机的传统方法高了6.7％。此外，我们证明，通过使用随机投影，在我们的框架内寻找“最佳特征”以实现最佳活动识别性能的任务并不重要。

12: A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer

摘要：通过可穿戴式传感器进行的体育活动识别可以提供有关个人功能能力和生活方式程度的有价值的信息。在本文中，我们提出了一种基于加速度传感器的人类活动识别方法。我们提出的识别方法使用分层方案。在较低级别上，通过统计信号特征和人工神经网络（ANN）识别活动所属的状态，即静态，过渡或动态。高层识别使用加速度信号的自回归（AR）建模，因此，将导出的AR系数与信号幅值区域和倾斜角结合在一起，以形成增强特征向量。通过线性判别分析和ANN对所得的特征向量进行进一步处理，以识别特定的人类活动。我们提出的活动识别方法仅使用连接到受试者胸部的单个三轴加速度计即可识别三种状态和15种活动，平均准确度为97.9％。

13: A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation,

摘要：背景技术集成电路，无线通信和生理感应的最新技术进步允许使用微型，轻便，超低功耗的智能监控设备。这些设备中的许多设备都可以集成到无线人体局域网（WBAN）中，这是一种用于健康监视的新技术。方法使用现成的无线传感器，我们设计了WBAN原型，该原型具有标准的ZigBee兼容无线电和通用生理，动力学和环境传感器集。结果我们引入了多层远程医疗系统，并描述了我们如何针对计算机辅助的身体康复应用和动态监测来优化我们的原型WBAN实现。该系统对传感器数据进行实时分析，为用户提供指导和反馈，并可以根据用户的状态，活动水平和环境条件生成警告。此外，所有记录的信息都可以通过Internet传输到医疗服务器，并无缝集成到用户的电子病历和研究数据库中。

14: Wearable sensors for reliable fall detection,

摘要：意外跌倒是老年人口严重受伤的常见原因。通过引入小型，无创的传感器节点以及无线网络，常春藤项目旨在为老年人提供更独立生活的途径。在家庭环境中，使用穿戴在腰上的小型设备和固定的尘埃网络，我们可以检测到跌倒的发生和受害者的位置。低成本，低功耗MEMS加速度计用于检测跌落，而RF信号强度用于定位人员

15: Monitoring intake gestures using sensor fusion (Microsoft Kinect and inertial sensors) for smart home tele-rehab setting

摘要：智能家居技术可帮助中风后患者独立完成日常生活（ADL）的活动，同时节省他们的时间，金钱和精力。否则，要求患者前往康复诊所进行正式护理。为了实现这一目标，我们提出了一种在家庭环境中发现特定ADL饮食习惯的方法。我们将惯性传感器和Microsoft Kinect传感器融合在一起，以监视患者的进气姿势，包括精细切割，装载食物以及将食物操纵到嘴里。对于身体的两侧，我们测量（i）手腕，肘部和肩膀的位置；（ii）肘关节和肩关节的角位移；（iii）受试者保持的勺子/叉子/杯子的加速度。Kinect的使用可以区分健康的身体侧面和瘫痪的身体侧面，这是远程康复中的常见问题。该系统已在健康受试者上成功测试；因为中风患者在较短的范围内显示出较慢的运动，所以该系统至少可以为他们提供同样好的服务。

16: Fuzzy inference-based reliable fall detection using kinect and accelerometer,

摘要：跌倒是老年人死亡和发病的主要原因。然而，流行的方法仅利用加速度计或将加速度计和陀螺仪两者都用于将跌落与日常生活活动分开。这使得区分真正的跌倒和类似跌倒的活动并不容易。现有的基于CCD相机的解决方案需要安装，相机校准的时间，并且通常并不便宜。在本文中，我们展示了如何实现可靠的跌倒检测。该检测是通过使用低成本Kinect的模糊推理系统以及由加速度计和陀螺仪组成的设备完成的。实验结果表明该系统具有较高的检测精度和有效性。

17: Indoor activity recognition by combining one-vs.-all neural network classifiers exploiting wearable and depth sensors,

摘要：活动识别最近引起了很多兴趣，并且似乎是一种帮助老年人寻求独立生活的有前途的方法。已经存在几种基于可穿戴式传感器或基于相机的人类活动检测方法，但是很少有方法将这两种方式结合在一起。本文提出了一种通过结合可穿戴和深度传感器来增强室内人类活动识别的鲁棒性的策略。为了利用这些传感器捕获的数据，我们使用了二元一对多神经网络分类器的集合。每个活动特定的模型都配置为最大化其性能。完整系统的性能可与需要整个数据集的惰性学习方法（k-NN）媲美。

18: Fusion of inertial and depth sensor data for robust hand gesture recognition

摘要：本文提出了在隐马尔可夫模型框架内融合惯性和视觉深度传感器数据以进行手势识别的首次尝试。本文介绍的数据融合方法是通用的，因为它可以用于识别各种身体运动。结果表明，与每个传感器单独使用的情况相比，来自视觉深度和惯性传感器的数据融合以互补的方式起作用，从而导致更可靠的识别结果。在Microsoft MSR数据集中获得的单个手势的识别率表明，我们的融合方法可在实时和现实条件下提供改进的识别。

19: Berkeley MHAD: A comprehensive multimodal human action database,

摘要：多年来，已经提出了许多方法来分析图像，视频以及最近来自深度数据的人体姿势和运动信息。但是，大多数方法都是在对每个应用程序都过于特定，仅限于特定模式，更重要的是在未知条件下捕获的数据集上进行评估的。为了解决这些问题，我们引入了伯克利多模式人类动作数据库（MHAD），该数据库由来自光学运动捕获系统的时间同步和几何校准的数据，来自多个视图的多基线立体摄像机，深度传感器，加速度计和麦克风组成。这个受控的多模式数据集为研究人员提供了一个包容性的测试平台，可以在各种研究领域的已知捕获条件下跨多种模式开发和基准化新算法。为了演示将MHAD用于动作识别的可能性，我们比较了使用流行的词袋算法（独立于每个模态）的结果与使用多核学习的模态的各种组合的结果。我们的比较结果表明，对人体运动的多模式分析比单模式分析产生更好的动作识别率。

20: Robust face recognition via sparse representation

摘要：我们考虑这样的问题：从正面视图自动识别具有变化表情和照明以及遮挡和伪装的人脸。我们将识别问题视为多个线性回归模型之间的分类之一，并认为稀疏信号表示的新理论为解决此问题提供了关键。基于通过C1最小化计算的稀疏表示，我们提出了一种用于（基于图像的）目标识别的通用分类算法。这个新框架为人脸识别中的两个关键问题提供了新见解：特征提取和遮挡的鲁棒性。对于特征提取，我们表明，如果正确利用了识别问题中的稀疏性，则特征选择不再是关键。但是，关键是特征的数量是否足够大以及稀疏表示是否已正确计算。只要特征空间的尺寸超过稀疏表示理论所预测的阈值，非常规特征（如下采样图像和随机投影）的性能就与常规特征（如本征面和拉普拉斯脸）一样好。通过利用这些错误通常相对于标准（像素）基础稀疏的事实，该框架可以统一处理由于遮挡和损坏引起的错误。稀疏表示的理论有助于预测识别算法可以处理多少遮挡，以及如何选择训练图像以最大程度地增强遮挡的鲁棒性。我们在可公开获取的数据库上进行了广泛的实验，以验证所提出算法的有效性并证实上述主张。

21: Sparse representation for computer vision and pattern recognition

摘要：稀疏信号表示技术开始对计算机视觉产生重大影响，通常对非传统应用程序具有重大影响，其目标不仅是获得观察到的信号的紧凑高保真表示形式，而且还提取语义信息。词典的选择在弥合这一差距方面起着关键作用：由训练样本本身组成或从中学习的非常规词典为获得最新结果并将语义含义附加到稀疏信号表示中提供了关键。理解此类非常规词典的良好性能反过来又需要新的算法和分析技术。这篇综述文章重点介绍了稀疏信号表示与计算机视觉之间的相互作用如何丰富这两个领域的几个代表性示例，并提出了许多悬而未决的问题需要进一步研究。

22: Sparse representation or collaborative representation: Which helps face recognition?

摘要：作为一种最近提出的技术，基于稀疏表示的分类（SRC）已被广泛用于面部识别（FR）。SRC首先将测试样本编码为所有训练样本的稀疏线性组合，然后通过评估哪个类别导致最小表示误差来对测试样本进行分类。尽管在SRC和许多相关工作中都非常强调稀疏性的重要性，但在大多数文献中都忽略了在SRC中使用协作表示（CR）。但是，真的是l1范数稀疏性提高了FR精度吗？本文致力于对SRC的工作机制进行分析，指出SRC对于人脸分类的强大作用在于CR，而不是l1-norm稀疏性。因此，我们提出了一种非常简单但效率更高的人脸分类方案，即基于CR的正则化最小二乘分类（CRC\_RLS）。广泛的实验清楚地表明，CRC\_RLS具有非常有竞争力的分类结果，而其复杂度却大大低于SRC。

23: Solutions of Ill-Posed Problems

摘要：摘要：预言我们先前使用非成像诊断技术所证明的那样，微波技术为院前中风的检测提供了可能性。此处的重点是基于图像的诊断，其中图像重建的技术和计算复杂性是临床实现中面临的挑战。这里，我们调查中风事件之前获得有关患者大脑解剖结构的信息是否可用于促进基础图像的中风诊断。通过从磁共振图像中分割患者的头部组织获得先验信息。专家手动分级目前是黄金标准，但是它既费力又主观。因此，需要一种自动的方法。此处介绍了使用合成磁共振成像（MRI）数据和来自四个健康的真实数据对几种以下方法的评估。分割是在完整的3D MRI数据上进行的，而电磁评估是使用2D切片进行的。根据以下方法对这些方法进行了评估：i）关于地面真相的所有组织的组织分类精度； ii）通过头传播的模拟电磁波的精度； iii）泄漏图像重建的精度。分割精度是根据与真实情况的重叠程度（骰子重叠）来减小的。电磁仿真精度是根据信号偏差最后，通过Dice评分，介电特性的相对误差以及真实和重建的脑出血之间的视觉比较来测量图像重建的精度。结果表明，根据MRI数据对组织进行准确的切分（切分分数= 0.97）可以对取向的大脑内出血进行准确的图像重建（相对误差= 0.24）。他们还建议，准确的自动分割人工分割的替代方法，并且可以使用微波成像系统快速诊断中风患者的脑出血。

24: Compressed-sensing recovery of images and video using multihypothesis predictions,

摘要：考虑了多假设预测驱动的静止图像和视频序列的压缩感知重构。具体地，对于静止图像，从初始非预测重建内的空间周围的块做出针对图像块绘制的多个预测。对于视频，根据一个或多个先前重建的参考帧生成当前帧的多假设预测。在每种情况下，预测均用于在压缩感知随机投影的域中生成残差。通常，该残差比原始信号更可压缩，从而改善了重建质量。为了适当地加权假设预测，提出了将Tikhonov正则化为不适定最小二乘优化的方法。实验结果表明，所提出的重建方案优于不采用多假设预测的替代策略。

25: Reconstruction of hyperspectral imagery from random projections using multihypothesis prediction

摘要：考虑从光谱随机投影重建高光谱图像。具体而言，针对感兴趣的像素矢量绘制的多个预测是根据初始非预测重建中空间相邻的像素矢量进行的。提出了一种根据频段之间的相关系数对频段进行划分和合并的两阶段假设生成程序，以对假设进行微调。所得的预测用于在投影域中生成残差。通常，该残差比原始像素矢量更可压缩，从而改善了重建质量。为了适当地加权假设预测，提出了对病态最小二乘优化的距离加权Tikhonov正则化。实验结果表明，所提出的重建方案明显优于不采用多假设预测的替代策略。

26: Single-image super-resolution using multihypothesis prediction,

摘要：考虑了由多假设预测驱动的单图像超分辨率。所提出的策略利用了单个图像内图像块之间存在的自相似性。具体而言，低分辨率图像的每个面片都表示为空间周围的假设面片的线性组合。使用Tikhonov正则化计算该表示的系数，然后将其用于生成高分辨率图像。实验结果表明，与双三次插值法相比，该算法可提供更高质量的超分辨率，而无需像竞争性单图像技术那样在大量图像训练集上进行训练。

27: Study of dempster-shafer theory for image segmentation applications

摘要：本文使用Dempster-Shafer的证据理论对图像分割和融合提出了一种新的知识模型。所提出方法的显着方面在两个层面上。首先，Dempster-Shafer理论中的质量分布是直接从图像直方图中得出的。其次，融合过程并非像先前报告的大多数作品中那样从单一识别帧开始，而是从首先定义与要融合的两个图像相关联的两个独立识别帧开始，然后合并它们形成了新的识别框架。详细讨论了用于定义组合框架中的质量分布的策略。在图像分割的背景下说明了所提出的融合方法。所得结果表明了该方法的鲁棒性。

28: Connectionist-based Dempster–Shafer evidential reasoning for data fusion

摘要：Dempster-Shafer证据理论（DSET）是处理不确定性和不精确性的流行范例。其相应的证据推理框架在理论上具有吸引力。但是，有一些突出的问题阻碍了它在现实生活中的应用。在这方面，两个突出的问题是：1）基本概率分配（质量）问题和2）信息源之间的依赖性问题。本文试图通过在模式分类应用中利用神经网络来解决这些问题。首先，实现了以均方误差作为成本函数的多层感知器神经网络，以针对每个信息源计算所有类别的后验概率。其次，从贝叶斯决策观点出发，开发了一种证据结构构建方案，用于将估计的后验概率与相应的焦点元素一起转移到一组质量。第三，设计和分析了Dempster-Shafer证据推理的网络实现，并将其扩展到基于DSET的神经网络，称为DSETNN，以操纵证据结构。为了解决源之间的依赖性问题，DSETNN通过有监督的学习过程进行了调整，以实现最佳性能。为了证明该方法的有效性，我们将其应用于三个基准模式分类问题。实验表明，DSETNN优于DSET，并且在分类准确性和学习收敛速度方面提供了令人鼓舞的结果。

29: Real time human action recognition based on depth motion maps

摘要：本文提出了一种使用深度运动图（DMM）的人体动作识别方法。深度视频序列中的每个深度帧都投影到三个正交的笛卡尔平面上。在每个投影视图下，两个连续投影图之间的绝对差通过形成DMM的整个深度视频序列进行累积。然后将具有距离加权的Tikhonov矩阵的l 2正规化协作表示分类器用于动作识别。所开发的方法显示出计算效率高，可以实时运行。应用于Microsoft Research Action3D数据集的识别结果表明我们的方法优于现有方法。

30: Learning patterns of activity using real-time tracking

摘要：我们的目标是开发一个视觉监视系统，该系统可以被动地观察站点中的移动物体并从这些观察中学习活动模式。对于扩展站点，系统将需要多个摄像机。因此，系统的关键要素是运动跟踪，摄像机协调，活动分类和事件检测。在本文中，我们将重点放在运动跟踪上，并说明如何使用观察到的运动来学习站点中的活动模式。运动分割基于自适应背景减法，该方法将每个像素建模为高斯混合，并使用在线逼近来更新模型。然后评估高斯分布，以确定最有可能是由后台过程导致的。这样就产生了一个稳定，实时的室外跟踪器，该跟踪器可以可靠地处理照明变化，杂波造成的重复运动以及长期的场景变化。跟踪系统不知道其跟踪的任何对象的身份时，对于整个跟踪序列，身份都保持不变。我们的系统通过累积序列中表示的联合共现来利用此信息。这些联合共现统计然后用于创建表示的分层二叉树分类。此方法对于分类序列以及站点中活动的单个实例很有用。

31: On space-time interest points

摘要：局部图像特征或兴趣点提供图像中图案的紧凑抽象表示。在本文中，我们将空间兴趣点的概念扩展到时空域，并显示所产生的特征如何经常反映出有趣的事件，这些事件可用于视频数据的紧凑表示以及时空事件的解释。为了检测时空事件，我们基于Harris和Förstner兴趣点算子的思想，并检测时空的局部结构，其中图像值在时空上都有明显的局部变化。我们通过在空间和时间尺度上最大化标准化的时空拉普拉斯算子来估计检测到的事件的时空范围。为了表示检测到的事件，我们然后计算局部，时空，尺度不变的N喷射，并根据其喷射描述符对每个事件进行分类。对于人体运动分析问题，我们说明了根据本地时空特征表示的视频表示如何允许在具有遮挡和动态混乱背景的场景中检测步行的人。

32: A k-nearest neighbor classification rule based on Dempster-Shafer theory

摘要：在本文中，从Dempster-Shafer理论的角度出发，解决了基于已记录数据集中最接近的邻居对未看到的模式进行分类的问题。要分类的样本的每个邻居都被视为支持有关该模式的类别成员身份的某些假设的证据项。支持程度定义为两个向量之间距离的函数。然后，根据Dempster的组合规则，将k个最近邻居的证据汇总起来。这种方法为诸如歧义性和距离拒绝之类的问题提供了全球性的处理，并且对有关培训模式的班级成员的知识不完善。使用几组模拟和现实数据，证明了该分类方案与投票和距离加权k-NN程序相比的有效性。

33: An evidence-theoretic

摘要：本文提出了一种用于优化证据理论k最近邻规则中的参数的学习程序，这是一种基于Dempster-Shafer信念函数理论的模式分类方法。在这种方法中，要分类的模式的每个邻居都被视为支持有关该模式的类成员的某些假设的证据。基于此证据，将基本信念质量分配给该类集合的每个子集。对于正在考虑的模式的每个k个最近邻均获得此类质量，并使用Dempster组合法则进行汇总。在许多情况下，通过实验发现，与使用相同信息的其他方法相比，该方法产生的错误率更低。但是，到目前为止尚未解决调整分类规则的参数的问题。作者通过最小化误差函数来从数据中确定最佳或接近最佳的参数值。实验证明了对原始方法的改进，从而大大提高了分类精度。

34: Learning realistic human actions from movies

摘要：本文旨在解决人们在多种逼真的视频环境中对自然动作的认识。由于一些问题，该挑战性但重要的主题在过去已被大多数人忽略，其中一个问题是缺少逼真的和带注释的视频数据集。我们的第一个贡献是解决这一局限性，并研究使用电影脚本自动注释视频中的人为行为。我们评估了从脚本中检索动作的替代方法，并展示了基于文本的分类器的优势。使用检索到的动作样本进行视觉学习，我们接下来转向视频中的动作分类问题。我们提出了一种新的视频分类方法，该方法基于并扩展了一些最新的思想，包括本地时空特征，时空金字塔和多通道非线性SVM。结果表明，该方法通过达到91.8％的准确度，可以改善标准KTH动作数据集的最新结果。鉴于自动标注中带有噪声标签的固有问题，我们特别研究并展示了我们的方法对训练集中标注错误的高度容忍度。最后，我们将该方法应用于电影中具有挑战性的动作类的学习和分类，并显示出令人鼓舞的结果。

35: Recognizing human actions: A local SVM approach

摘要：本地时空特征捕获视频中的本地事件，并且可以适应移动模式的大小，频率和速度。在本文中，我们演示了如何将这些功能用于识别复杂的运动模式。我们根据本地时空特征构造视频表示，并将此类表示与SVM分类方案进行集成以进行识别。为了进行评估，我们引入了一个新的视频数据库，其中包含2391个序列，这些序列由25个人在四种不同情况下执行的六个人类动作组成。提出的动作识别结果证明了该方法的合理性，并证明了与其他相关方法相比，该方法具有优势。