深度与惯性传感器融合的实时人体动作识别系统

摘要：本文提出了一种实时运行的人类动作识别系统，该系统同时使用深度摄像头和基于以前开发的传感器融合方法的惯性传感器。计算有效的深度图像特征和惯性信号特征被馈送到两个计算有效的协作代表分类器中。然后执行决策级融合。通过考虑一组全面的人类行为，使用可公开获得的多模式人类行为识别数据集对开发的实时系统进行评估。所开发的实时系统的总体分类率显示为> 97％，这比单独使用每种传感方式时至少高出9％。离线和实时实验的结果都表明了该系统的有效性及其实时吞吐量。

01: 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。

02: Hierarchical spatio-temporal context modeling for action recognition

摘要：由于其在许多实际应用中的巨大潜力，在逼真的视频中识别动作的问题既具有挑战性，又具有吸收性。由于在受控环境下使用简化的动作数据库或在没有充分封装时空上下文的情况下专注于过度本地化的功能，大多数先前的研究受到了限制。在本文中，我们建议以分层的方式对时空上下文信息进行建模，其中以抽象的升序使用三个级别的上下文：1）点级上下文（SIFT平均描述符），2）轨迹内上下文（轨迹过渡描述符），以及3）轨迹间上下文（轨迹接近描述符）。为了获得后两个级别的高效紧凑表示，我们将时空上下文信息编码到马尔可夫过程的转换矩阵中，然后提取其平稳分布作为最终上下文描述符。在多通道非线性SVM的基础上，我们验证了拟议的逼真的动作（HOHA）和事件（LSCOM）识别数据库的分层框架，相对于最新结果分别实现了27％和66％的相对性能提升。我们进一步建议采用多核学习（MKL）技术来修剪内核，以加快算法评估的速度。

03: 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。

04: 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。

05: 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。

06: Action recognition from depth sequences using depth motion maps-based local binary patterns

摘要：本文提出了一种从深度视频序列中识别动作的高效计算方法。它使用来自三个投影视图（正面，侧面和顶部）的所谓的深度运动图（DMM）来捕获运动线索，并使用局部二进制模式（LBP）获得紧凑的特征表示。考虑了两种类型的融合，包括特征级融合和决策级融合。在特征级融合中，来自三个DMM的LBP特征在分类之前被合并，而在决策级融合中，使用软决策融合规则来组合分类结果。引入的方法在两个标准数据集上进行了评估，并与现有方法进行了比较。结果表明，其性能优于现有方法，并且能够实时处理深度视频序列。

07: 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。

08: Feature selection and activity recognition system using a single triaxial accelerometer

摘要：在各种应用中（例如医疗监测和康复）都需要活动识别。先前开发的利用三轴加速度计的活动识别系统已经提供了混合的结果，并且对象之间存在差异。本文提出了一种精确的活动识别系统，该系统利用人体穿戴的无线加速度计，将其用于患者监护的实际应用中。该算法利用来自单个腰部安装的三轴加速度计的数据将步态事件分类为六个日常活动和过渡事件。加速度计可以在腰围周围的任何位置佩戴，从而减少了用户培训。使用Relief-F和顺序向前浮动搜索（SFFS）从一系列先前发布的功能以及本文介绍的新功能中进行功能选择。选择对加速度计在腰部周围不敏感的相关且强大的功能。与Relief-F相比，SFFS选择了几乎一半的功能，并提供了比Relief-F更高的准确性。使用朴素贝叶斯和k最近邻（k-NN）进行活动分类，并比较结果。对七个主题的活动识别结果具有一人一人的错误估计，这两个分类器的整体准确性约为98％。每个活动的准确性也超过95％。

09: Distributed recognition of human actions using wearable motion sensor networks

摘要：我们提出了一种分布式识别框架，该框架使用称为分布式稀疏分类器（DSC）的低带宽可穿戴运动传感器网络对人类的连续行为进行分类。该算法使用一组训练运动序列作为先前示例对人类动作进行分类。它还能够拒绝不在训练类别中的孤立动作。分类在各个传感器节点和基站计算机上以分布式方式进行。我们将多个动作类的分布建模为混合子空间模型，每个动作类一个子空间。给定一个新的测试样本，我们将在所有训练示例中寻求样本的最线性表示。我们表明，表示中的主导系数仅对应于测试样本的动作类，因此其成员资格在稀疏表示中进行编码。提供了快速线性求解器，以通过ℓ1-最小化来计算此类表示。为了验证框架的准确性，构建了一个公共的可穿戴动作识别数据库，称为可穿戴动作识别数据库（WARD）。该数据库由13个动作类别的20位人类受试者组成。DSC在WARD数据库中最多使用五个运动传感器，从而实现了最先进的性能。我们进一步表明，使用主动传感器的较小子集，识别精度只会适度降低。它在不可靠的无线网络上验证了分布式识别框架的鲁棒性。它还展示了DSC节省传感器能量以进行通信，同时保留准确的全局分类的能力。

10: 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。

11: 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。

12: 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。

13: 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，通过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，Dempster-Shafer理论用于合并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％。

14: 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。

15: 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捉系统的时间同步和几何校准的数据，来自多个视角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。

16: 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l 1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。

17: Solutions of Ill-Posed Problems

摘要：正如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。本文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面临的挑战。本文中，我们调查中风事件之前获得的有关患者大脑解剖结构的信息是否可用于促进基于图像的中风诊断。通过从磁共振图像中分割患者的头部组织可以获得先验信息。专家手动细分目前是黄金标准，但是它既费力又主观。因此，需要一种全自动的方法。本文介绍了使用合成磁共振成像（MRI）数据和来自四个健康受试者的真实数据对几种此类方法的评估。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准确性； ii）通过头传播的模拟电磁波的准确性； iii）出血图像重建的准确性。分割精度是根据与真实情况的重叠程度（骰子得分）来衡量的。电磁仿真精度是根据信号偏差相对于基于地面真实情况的仿真进行测量的。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准确性。结果表明，根据MRI数据对组织进行准确的切分（切分评分= 0.97）可以对受试者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，准确的自动分割可用作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。

18: Local binary patterns and extreme learning machine for hyperspectral imagery classification

摘要：利用纹理信息对高空间分辨率的高光谱图像（HSI）进行分类具有极大的兴趣。提出了一种利用HSI丰富的纹理信息的分类范式。提出的框架采用局部二进制模式（LBP）来提取局部图像特征，例如边缘，拐角和斑点。将两级融合（即特征级融合和决策级融合）与全局Gabor特征和原始光谱特征一起应用于提取的LBP特征，其中特征级融合包括在模式分类过程之前将多个特征串联在一起，而决策级融合对每个单独分类流水线的概率输出执行，并且采用软决策融合规则来融合来自分类器集合的结果。此外，具有非常简单结构的高效极限学习机被用作分类器。在多个HSI数据集上的实验结果表明，提出的框架优于某些传统替代方案。

19: G3D: A gaming action dataset and real time action recognition evaluation framework

摘要：本文提出了一种用于评估实时动作识别方法性能的新型评估框架。评估框架将扩展基于时间的事件检测指标，以对多个不同的动作类别进行建模。提议的度量标准提供了游戏和其他类似应用程序的动作识别算法性能的更准确指示，因为它考虑了与时间和连续重复相关的限制。此外，提供了一个新的数据集G3D，用于在游戏中进行实时动作识别，其中包含同步的视频，深度和骨骼数据。我们的结果表明，需要专门针对游戏和其他类似实时应用程序设计的高级指标。

20: Inertial sensor-based touch and shake metaphor for expressive control of 3D virtual avatars

摘要：在本文中，我们提出了一种基于惯性传感器的触摸和摇动隐喻，用于在虚拟环境中对3D虚拟化身进行表达控制。直观的六自由度无线惯性运动传感器用作具有传感器融合算法的手势和运动控制输入设备。该算法使用户可以通过磁性，角速率和重力传感器在3D空间中跟踪用户的手部动作。实现了基于四元数的互补滤波器，以减少噪声和漂移。开发了一种基于动态时间扭曲的算法，可通过实时自动手势分割有效地识别动态手势。我们的方法能够识别手势并估计手势变化以进行连续交互。我们通过使用交互式灵活的手势映射界面演示了手势表达，该界面用于通过跟踪用户动态手势来创作和控制3D虚拟化身及其运动。这可以合成3D虚拟化身中的风格变化，并使用来自单个惯性运动传感器的手势序列生成运动数据库中不存在的运动。

21: UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor

摘要：人类动作识别具有广泛的应用，包括生物识别，监视和人机交互。用于人类动作识别的多模式传感器的使用正在稳步增加。但是，有限的公共数据集可以同时捕获深度相机和惯性传感器数据。本文介绍了一个名为UTD-MHAD的免费数据集，该数据集由四个时间同步的数据模态组成。这些模式包括RGB视频，深度视频，骨骼位置以及来自Kinect相机和可穿戴惯性传感器的惯性信号，可对27种人类动作进行全面设置。提供了实验结果，以显示该数据库如何用于研究涉及同时使用深度相机数据和惯性传感器数据的融合方法。该公共领域数据集有益于各种研究小组正在开展的多模式研究活动，以进行人类行为识别。