

学校编码: 10384

分类号 _____ 密级 _____

学号: 31520091152827

UDC _____

基于局部时空共现特征的人体行为识别方法研究

陈延平

指导教师

李绍滋 教授

厦门大学

厦门大学

硕士学位论文

基于局部时空共现特征的人体行为识别方法研究

Research on Techniques in Human Action Recognition base on Local Spatial-Temporal Co-occurrence Features

陈延平

指导教师姓名: 李绍滋 教授

专业名称: 计算机技术

论文提交日期: 2012 年 5 月

论文答辩时间: 2012 年 月

学位授予日期: 2012 年 月

答辩委员会主席: _____

评 阅 人: _____

2012 年 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文(包括纸质版和电子版)，允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

()1.经厦门大学保密委员会审查核定的保密学位论文，于
年 月 日解密，解密后适用上述授权。

()2.不保密，适用上述授权。

(请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。)

声明人(签名)：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

人体行为识别是计算机视觉领域的研究热点和难点，在智能监控、高级人机交互、体育运动分析和基于语义的视频检索等方面有着重要的意义和广阔的应用前景。计算机理解人体动作和行为将成为未来高层计算机视觉技术发展的重要方向。

目前，人体行为识别的研究尚处于初级阶段，尽管在单一背景的简单动作识别上取得了一定的进展，但是由于人体行为的复杂性、特征提取的时间复杂度高等问题，真实场景下的人体行为识别仍然面临着许多困难，如实时性差、识别率未达到实际应用等。针对上述情况，本文围绕基于视频的人体行为识别展开研究，进行了以下三个方面的工作：

1. 查阅了大量文献，对现有的基于局部时空特征的人体行为识别方法的进展的进行详细概述，重点介绍了局部时空特征的检测算子和描述算子。为本文后续工作的展开提供理论支持。
2. 提出了一种基于 CUDA 的 HOG 特征算子的快速计算方法。针对 HOG 特征提取的时间复杂度高的问题，根据 CPU 下的 HOG 特征算子，研究了 GPU 下的 HOG 特征快速计算方法。在 GPU 下分别对 HOG 特征的梯度计算、直方图统计、归一化处理、线性 SVM 分类器检测部分，特征提取的五个关键步骤都分别设计了相应的并行模块。实验结果表明 GPU 下的 HOG 特征计算相较于 CPU 下的 HOG 特征计算平均提高了约 10 倍的速度。
3. 提出了一种基于表观码本同现的特征包模型算法。针对传统特征包模型中的缺点，即简单的用单个码本来表示每个聚类中心的做法，本文引入了表观码本同现的方法来增加码本之间的相关性。首先利用局部特征检测算子和特征描述算子计算出特征向量。接着利用 k-means(k-均值聚类)算法对所有的训练视频特征向量进行聚类，构建 k 个码本项。然后对 k 个码本计算共现码本矩阵和共现码本相关度矩阵，根据共现码本相关度矩阵得出相关度最大的 k 对共现码本。然后对训练视频和测试视频分别计算 k 个码本的原码本直方图和 k 对共现码本的直方图。最后合并原码本直方图和共现码本直方图组成一个新的码本直方图，这样每个视频特

征向量由原码本直方图和表观码本共现直方图组成。在包含 6 种动作的 KTH 数据库进行实验,实验结果表明,本文的方法提高了 2.4%的准确率。

总之,本文在总结现有人体行为识别方法的基础上,提出了一种基于 CUDA 的 HOG 特征算子的快速计算方法,并在此基础上提出了一种基于表观码本同现的特征包模型的改进算法,实验结果表明了本文提出的方法的有效性。

关键词: 人体行为识别; 表观码本同现; 特征包模型; HOG; CUDA; HOG/HOF

厦门大学博硕士论文摘要库

Abstract

As a hot topic in the field of computer vision, human behavior recognition could be widely used in many practical applications range from intelligent video surveillance, advanced man-machine interface, sports motion analysis, and video retrieval based on semantic. Computer understanding of human action and behavior will be an important direction of high-level computer vision technology development in the future.

Currently, research of human behavior recognition is still in its infancy. Although having a good performance in simple background and simple action, human behavior recognition in real scenes is still faced with great challenges because of time complexity of the feature extraction and complex human behaviors. In response to these problems, in this paper we focus on the research of the following three aspects on human behavior recognition in videos:

1. Summarize existing behavior recognition methods of local spatio-temporal feature representation, and focuses on the methods of local feature detector and feature descriptor. This paper presents the basic idea of common algorithms in details and compares their advantages and disadvantages.

2. Propose a method of fast HOG feature computation based on CUDA. In order to solve the problem of time complexity of the HOG feature extraction, base on the HOG feature calculation under the CPU, the GPU's algorithm contains: gradient computation , gradient of histogram, normalization and the linear SVM classifier detection, feature extraction on the GPU is designed to parallel modules, the final test results have more than 10 times the speed faster.

3. Propose a method of bag of feature model base on co-occurrence. For the shortcomings of the bag of feature model, that is with a single codebook express each cluster center, the introduction of the co-occurrence codebook with the present method to increase the codebook relation. Use k-means algorithm to cluster the vectors, result of constructing k codebooks. Then calculate the matrix of co-occurrence and relevance of co-occurrence matrix, base on the k codebooks. Using

the relevance of co-occurrence matrix, we can obtain the maximum relevance of the k pair codebooks. Then we calculate the histogram of k codebooks and the histogram of k pair codebooks. Then merge the two histogram to construct a new histogram. Finally, each video can be represented of the above two histograms.

In summary, this paper research base on existing methods of human action recognition. It proposes fast HOG feature computation based on CUDA, and an improved bag of feature model base on co-occurrence. Experiment results show that our methods and model are effective.

Key Words: Human Behavior Recognition; Co-occurrence; Bag of Feature; HOG; CUDA; HOG/HOF

目 录

摘 要	I
Abstract.....	III
第一章 绪论	1
1.1 选题背景及研究意义	1
1.2 国内外研究现状及存在的问题	2
1.2.1 国内外研究现状.....	2
1.2.2 存在的问题.....	8
1.3 本文主要研究工作及创新点	9
1.4 本文组织结构	10
第二章 基于局部时空特征的动作识别方法及现状分析	13
2.1 局部特征表示	13
2.1.1 时空兴趣点的检测.....	13
2.1.2 局部描述符.....	14
2.1.3 基于网格的局部表示.....	15
2.1.4 局部描述符之间的关联.....	16
2.2 局部时空特征	17
2.2.1 局部时空特征检测.....	17
2.2.2 局部特征描述算子.....	21
第三章 基于 CUDA 的快速 HOG 特征计算	27
3.1 梯度方向直方图（HOG）	27
3.2 CUDA 软件开发环境.....	29
3.2.1 GPU 概述.....	29
3.2.2 CUDA 编程模型	30
3.2.3 CUDA 存储模型	31
3.2.4 CUDA 编程接口	33
3.3 基于 CUDA 下的 HOG 特征算法实现	34

3.3.1 梯度的计算.....	35
3.3.2 梯度直方图的计算.....	36
3.3.3 归一化处理.....	37
3.3.4 线性 SVM 分类器.....	37
3.3.5 提取图像的 HOG 特征.....	38
3.4 实验结果与分析	38
3.5 本章小结	39
第四章 基于表观码本同现的特征包模型	41
4.1 特征包模型 (BOF)	42
4.2 表观码本同现 (Co-occurrence)	44
4.3 基于表观码本同现的特征包模型改进算法.....	47
4.4 实验结果与分析.....	48
4.5 本章小结.....	52
第五章 总结与展望	53
5.1 回顾与总结	53
5.2 未来研究工作展望	54
参考文献.....	55
致 谢.....	61
附录 攻读硕士学位期间发表的论文	62

Table of Contents

Chapter1 Introduction	1
1.1 Background and Significant.....	1
1.2 Present Research and Problems	2
1.2.1 State-Of-Art	2
1.2.2 Problems	2
1.3 Common database.....	8
1.3 Main Research Contents	10
1.4 Outline.....	11
Chapter2 Human Behavior Recognition Methods	13
2.1 Local Feature Presentation	13
2.1.1 Spatial-Temporal Interest Point Detector.....	13
2.1.2 Local Descriptor.....	14
2.1.3 Local Representation base on Grid.....	15
2.1.4 Local Descriptor Co-occurrence	16
2.3 Local Spatial-Temporal Feature.....	17
2.3.1 Local Spatial-Temporal Feature Detector.....	17
2.3.2 Local Spatial-Temporal Feature Descriptor.....	21
Chapter3 HOG Feature Coputation base on CUDA.....	27
3.1 Histograms of Oriented Gradients	27
3.2 CUDA Development Environment	29
3.2.1 GPU Outline	29
3.2.2 CUDA Program Model.....	30
3.2.3 CUDA Memory Model.....	32
3.2.4 CUDA Program Interface.....	33
3.3 Implementation Of HOG Feature Coputation base on CUDA	34
3.3.1 Gradient Computation.....	35
3.3.2Histogram Of Gradient Computation.....	36
3.3.3 Normalization.....	37

3.3.4 Linear SVM Classifier.....	37
3.3.5 HOG Feature Extraction	38
3.4 Experiment Result and Analysis	38
3.5 Conclusion	39
Chapter4 Bag Of Feature Model Base on Co-occurrence	41
4.1 Bag Of Feature	42
4.2 Co-occurrence	44
4.3 Bag Of Feature Model Base on Co-occurrence Algorithm	48
4.4 Experiment Result	49
4.5 Conclusion	52
Chapter5 Conclusion and Prospect.....	53
5.1 Conclusion	53
5.2 Propection of the Future work	54
References	55
Acknowledgements	61
Appendix Published Papers	62

第一章 绪论

1.1 选题背景及研究意义

计算正渗透和影响人们生活的各个方面, 凭借其本身的超大的存储能力与运算速度深受人们的欢迎。但是计算机对人的依赖性还是很强, 为此人们想让计算机具有自主学习、理解和分析的能力, 计算机视觉由此产生。人体行为识别是近年来计算机视觉研究的前沿方向和热点, 其主要目标如下: 通过从视频中提取图像序列的特征并且给出相应动作的分类标签。

目前, 人体行为识别技术已经应用于很多领域, 下面将从生物特征识别、智能监控、感知接口、虚拟现实和基于内容的视频分析五个方面介绍其应用。

1. 生物特征识别

生物识别技术研究了基于身体或行为特征的人体识别方法和算法。传统的方法基于指纹、面部或虹膜等物理特征识别, 也可称为为生理生物识别技术。然而这些方法需要收集人体的生物特征。最近几年, 行为的生物识别技术研究越来越多, 所用的特征需要跟传统的物理特征一样识别效果好。这种方法的优点是, 获取人体特征变得没有必要, 由于它可以不进行中断或干扰人体活动。观察行为意味着长期的观察主体, 动作识别方法自然可以应用到这个地方。目前, 最有前景的行为生物识别特征则是人类的步态研究。

2. 智能监控

传统的视频监控依赖网络视频系统, 操作者需要时刻监控摄像机的视野范围内的异常活动。随着近期摄像头需求数量的增长和大范围的部署, 人工操作的效率和准确性已经很难控制。因此, 许多安全机构正在寻求这些任务基于视觉解决方案可以取代或协助人类操作者。在摄像头的视野范围自动识别异常的行为问题已经引起视觉研究人员的注意。这跟一个相关的应用问题很相似, 即通过从长视频学习行为模式, 从而在一个庞大的数据库搜索感兴趣的行为。

3. 感知接口

如何让计算机和人类之间更好的互动, 仍然是设计人机交互接口的挑战之

一。视觉是最主要沟通方式之一。有效地利用这种方式，如姿态和行为，能帮助建立可以更好地与人类互动的计算机。同样，基于视觉方法可以使用户的姿态与智能环境进行互动。然而这种技术还不成熟，还不能通过“图灵测试”。

4. 虚拟现实

游戏和动画行业依赖于虚拟逼真的人体和人体运动的技术。人体运动的虚拟在游戏行业被广泛的使用，因为游戏为了产生尽量多的动作可以在一些质量上做一些妥协。另一方面，电影行业传统上更多依靠动画师提供高品质的动画。然而，这一趋势正在迅速改变。随着计算机算法和硬件的改善，更现实的运动合成变得有可能。这项技术的一个相关应用是通过模拟环境中来学习。这方面的例子包括虚拟出一个现实环境中的危险情景来训练战士，消防战士或者其他救援人员。

5. 基于内容的视频分析

视频已成为我们的日常生活一部分。随着视频分享网站的视频数量快速的增长，这使得很有必要的开发高效的视频检索和存储策略，以改善用户体验。这需要从原始视频学习识别的模式，并且基于其内容概括视频。基于内容的视频检索已经和基于内容的图像检索一样受到高度的关注。概括和检索视频内容（如体育影片）最可行的商业应用之一。

由于人体行为识别应用广泛，使得很多学者和研究机构开始关注这一研究领域。然而人体行为识别仍然处于探索阶段，因此具有潜在的巨大经济利益和社会价值。目前人体识别的研究都还是建立在简单场景下，并没有考虑光照、遮挡、视角、尺度变换和复杂背景等诸多因素的影响，这使人体行为识别很难应用到真实环境中。因此对于行为识别技术进行深入研究具有重要的理论和现实意义，将会更好改善人们的生产生活。

1.2 国内外研究现状及存在的问题

1.2.1 国内外研究现状

根据最新的研究工作，本文把人体行为识别的方法分为 3 类：人体模型法，全局方法，局部特征法。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库