

学校编码: 10384
学号: 31520110154098

分类号_____密级_____
UDC_____

厦门大学

博 士 学 位 论 文

基于多信息融合的人体行为识别方法研究

Research on Multi-information Fusion based

Human Action Recognition

唐 超

指导教师姓名: 周昌乐 教授

专业名称: 人工智能基础

论文提交日期: 2014 年 月

论文答辩时间: 2014 年 月

学位授予日期: 2014 年 月

答辩委员会主席: _____

评 阅 人: _____

2014 年 月

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

摘要

基于视觉的人体运动分析是计算机视觉研究领域一个非常活跃的研究方向。它在人机交互、基于内容的视频检索、智能视频监控和机器人等领域有着广泛的应用前景。人体运动分析技术包括：人体运动目标检测和分割、目标跟踪、行为识别和场景的语义理解。人体行为识别属于人的运动分析的高级处理部分。人体行为识别可以简单地看成一个时变数据的分类问题，包括行为特征提取和分类算法的设计两部分。基于视觉的人体行为识别的研究融合了图像处理、计算机视觉、模式识别和人工智能等学科的研究成果，是一个交叉的研究领域。

本文通过对人体行为识别的实际应用需求和已有研究方法的分析，总结了人体行为识别方法中存在的几个关键性问题：(1)快速高效的运动人体目标检测方法；(2)较强的有判别能力的人体行为特征提取方法；(3)有效的人体行为识别模型的构建和(4)快速高效的行为分类模型设计。基于这些问题，我们从人体运动目标检测、人体行为特征表示、分类模型的构建和分类算法的设计四个角度出发并在已有工作的基础上提出了新的方法，本文的主要贡献是：

1. 提出了一种鲁棒性强，抗噪性能优的人体运动目标检测方法(C-CLG 方法)。该方法将 Canny 边缘检测技术与基于局部与全局光流算法 (Combined Local and Global Optical Flow Method, CLG) 相结合，通过 Canny 算子检测出物体的边缘信息，用 CLG 光流法得到物体的光流场信息，然后将边缘信息和光流场信息进行融合，并采用数学形态学方法对融合结果进行处理，从而得到运动目标。C-CLG 方法比传统的运动目标检测方法有较强鲁棒性和抗噪性能，可以适应不同场景下的运动人体的检测。

2. 提出了一种简单高效的多信息融合的混合行为特征表示方法和相应的识别算法。该混合行为特征具有简单、鲁棒和判别能力强的特点，它融合了基于曲率函数的傅立叶描述子特征(Curvature Function-Based Fourier Descriptors, CFD)、人体区域特征(Shape-Based Region Descriptors, SRD)和人体骨骼重要运动关节特征(Joints-based Polar Coordinates Descriptors, JPCD)，同时利用泛化能力较强的随机森林模型进行快速分类。相比其它监督行为识别分类算法，肯有简单、快速，高效的特点。

3. 提出了一种基于半监督学习框架的行为识别模型(Multi-classification Co-training Model, MC-CO)。其主要特点是:(1)可以有效利用少量有标记样本和大量未标记样本学习得到泛化能力较强的分类模型,满足了一些实际的应用要求。(2)不同于标准的协同训练算法,本文提出的基于熵度量的学习器选择算法(Classifiers Selection Algorithm based on Diversity Measure of Entropy, CSADME)不需要两个独立冗余视图的要求,这更符合实际的行为识别问题。(3)基于标记近邻和未标记近邻的置信度计算公式来评估未标记样本的置信度,增加了基学习器之间的差异性,从而可以更好的提高学习器相互学习能力。

4. 提出了一种新的基于选择性集成的旋转森林行为识别算法(Based on Selected Ensemble Rotation Forest Model, SERF)。该模型将集成学习中的选择性集成策略融入到基分类器选择中,增强了旋转森林中基分类器之间的差异性。分类器融合策略采用基于置信度的投票方法。

综上所述,本文针对人体行为识别中存在问题与缺陷,有针对性提出了几种解决方法。经过大量实验证明文中提出的方法是有效的,且具有良好的应用潜力。

关键词: 人体行为识别 多信息融合 行为特征表示 半监督学习 分类器设计

Abstract

Visual-based human action recognition is currently one of the most active research topics in computer vision with many important applications, including human-computer interfaces, content-based video indexing, video surveillance, and robotics, among others.

The objective of human action analysis is to detect and recognize human actions from videos so that the computer system is able to understand human behaviors and make further semantic description of the scene.

Computer systems understand human actions from the scene involving two major steps: human motion detection and human action recognition. There are challenges in both of these two research areas. Generally speaking, the main challenge of human motion detection from video is to detect humans with different moving speeds in complex background clusters, and under different illumination changes. For human action recognition, there have been a number of action classifiers proposed, but the crucial factors are how to give effective and efficient representations of high dimensional human actions for categorization or recognition, how to employ unlabeled video data to train and enhance the performance of the action recognition system, and how to categorize the disorderly and unsystematic video data in human action aspect quickly and efficiently.

In this thesis, new algorithms are proposed from multi-information fusion based approach to solve these problems in human action detection and recognition.

1. A new Canny-Combined Local and Global Optical Flow Method (C-CLG) is proposed to employ optical flow and Canny operator for human motion detection via direct analysis from videos. The Canny operator is used to compute the contour of the objects in the video. Meanwhile, the Optical Flow method is adopted to establish the motion vector field. After using the morphologic operators to process the results, the moving objects are obtained. The C-CLG is robust for object detection under illumination changes.

2. This paper presents a simple but efficient action recognition algorithm using mixed visual features (CFD+SRD+JPCD, CSJ). The mixed features fuse three action descriptors, namely curvature function-based Fourier descriptors (CFD), shape parameters-based regional descriptors (SRD) and joints-based polar coordinates descriptors (JPCD). The frame-based human action classifier is developed using random forests algorithm. The CSJ can be efficiently extracted from human silhouette and human skeleton data, and used for training the human action classifier.

3. This thesis develops a semi-supervised algorithm for human action recognition, as labeled data are costly to obtain whereas unlabeled data are abundantly available. To tackle this problem, we propose a novel multi-class semi-supervised learning algorithm in which the variant co-training algorithm namely MC-CO is adopted to leverage the information from unlabeled data. Also, a classifiers selection algorithm based on diversity measure of entropy (CSADME) is shown to solve the classifiers selection issue. CSADME is adopted to select two base classifiers through using entropy diversity measure instead of requiring two views of the data, and thus it is applicable to action recognition problems with no natural attribute partitions.

4. This paper proposes a Based on Selected Ensemble Rotation Forest (SERF) action classification model in which selective ensemble strategy is used to pick the base learner, to increase difference between the classifiers in the Rotation Forest model.

The proposed methods have been extensive evaluated using Kinect human action classification database. Comparison between the proposed methods and existing state-of-the-art methods are also reported in this thesis.

Key words: Human action recognition; Multi-information fusion; Action Representation; Semi-supervised learning; Classifier Designing

目 录

摘 要.....	I
Abstract.....	III
目 录	V
Table of Contents	IX
第一章 绪论	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 动作特征描述	3
1.2.2 行为识别方法	6
1.3 本研究的主要贡献.....	9
1.4 本文的组织结构.....	10
第二章 融合边缘信息与运动信息的运动目标检测方法.....	11
2.1 相关工作介绍.....	11
2.2 微分光流法.....	13
2.4 基于 C-CLG 的运动目标检测方法	15
2.3 实验与分析.....	17
2.3.1 数据集介绍	18
2.3.2 室外场景实验	19
2.3.3 室内场景实验	20
2.3.4 与其它方法的比较实验	21
2.4 本章小结.....	22
第三章 基于混合特征的人体行为识别方法	23
3.1 方法框架.....	23

3.2 行为特征提取	24
3.2.1 基于曲率函数的傅立叶描述子	25
3.2.2 人体区域统计量特征	29
3.2.3 基于关节点极坐标特征	31
3.2.4 特征融合	34
3.3 基于单帧的随机森林行为识别算法	34
3.3.1 随机森林	35
3.3.2 识别算法	36
3.4 实验结果与分析	37
3.4.1 Kinect 视觉系统.....	37
3.4.2 数据集	38
3.4.3 实验分析	40
3.5 本章小结	43
第四章 基于半监督学习的行为识别方法	45
4.1 相关工作介绍	46
4.1.1 半监督学习	46
4.1.2 基于视图的半监督学习算法	49
4.2 系统框架	51
4.3 基于熵的差异性度量学习器选择算法	53
4.3.1 学习器差异性度量方法	53
4.3.2 基于熵的差异性度量算法	55
4.4 多分类的协同训练算法 (MC-CO)	57
4.5 实验结果与分析	60
4.5.1 实验设置	60
4.5.2 方法评估	61
4.5.3 与其余方法的对比	67
4.6 本章小结	69
第五章 基于选择性集成的旋转森林人体行为识别方法	71
5.1 方法框架	71
5.2 特征提取	72

5.2.1 质心边界距	72
5.2.2 光流特征	73
5.2.3 混合特征	75
5.3 识别算法.....	75
5.3.1 旋转森林	75
5.3.2 基于不一致度量的选择性集成算法	78
5.3.3 多分类器的融合	79
5.4 实验分析.....	81
5.5 本章小结.....	85
第六章 总结与展望	87
6.1 研究工作总结.....	87
6.2 未来工作展望.....	88
参 考 文 献	89
攻读博士学位期间发表录用的学术论文.....	105
致 谢	107

厦门大学博硕士学位论文摘要库

Table of Contents

Abstract (in Chinese)	I
Abstract (in English)	III
Table of Contents (in Chinese)	V
Table of Contents (in English)	IX
Chapter I Introduction	1
1.1 Background and Meaning	1
1.2 States-of-the-arts	2
1.2.1 Human action representation	3
1.2.2 Models and classification methods	6
1.3 Contribution	9
1.4 Outline	10
Chapter II Fusion motion and edge information for moving target detection	11
2.1 Related works	11
2.2 Differential optical flow technology	13
2.2 C-CLG-based moving target detection method	15
2.3 Experiments and Analysis	17
2.3.1 Data sets	18
2.3.1 The outdoor experiments	19
2.3.2 The indoor experiments	20
2.3.3 Comparison experiments	21
2.4 Conclusion	22
Chapter III Hybrid features for human action recognition	23
3.1 System overview	23

3.2 Feature extraction	24
3.2.1 Curvature function-based Fourier descriptors	25
3.2.2 Shape-based region descriptors	29
3.2.3 Joints-based polar coordinates descriptors	31
3.2.4 Features fusion	34
3.3 Frame-based action recognition.....	34
3.3.1 Random forest	35
3.3.3 Recognition algorithm.....	36
3.4 Experiments and analysis	37
3.4.1 Kinect sensor	37
3.4.2 Data sets	38
3.4.3 Experimental analysis.....	40
3.5 Conclusion.....	43
 Chapter IV Action recognition based on semi-supervised learning	
.....	45
4.1 Related works	46
4.1.1 Semi-supervised learning	46
4.1.2 Views-based SSL algorithms.....	49
4.2 System overview	51
4.3 Classifiers selection algorithm based on entropy measure	53
4.3.1 Diversity measures of classifiers	53
4.3.2 Classifiers selection algorithm based on diversity measure of entropy	55
4.4 Multi-classification co-training algorithm (MC-CO)	57
4.5 Experiment ans analysis	60
4.5.1 Experimental settings	60
4.5.2 Performance evaluation	61
4.5.3 Comparative results	67
4.6 Conclusion.....	69
 Chapter V Human action recognition based on selected ensemble	

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.

厦门大学博硕士学位论文摘要库