

DOI:10.11992/201507005

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20160105.1532.010.html>

压缩感知 W-HOG 特征的运动手势跟踪

史东承,倪康

(长春工业大学 计算机科学与工程学院,吉林 长春 130012)

摘要:运用单一特征或较简单的特征进行手势跟踪容易产生跟踪偏差。为提高手势跟踪的精确性,将压缩感知运动目标跟踪运用到手势跟踪中,并提取运动区域的 HOG 特征替换原来广义的类 Harr 特征进行压缩感知目标跟踪。为减少不同图像块 HOG 特征等权值串联对手势跟踪所产生的累积跟踪误差,需计算 HOG 特征权值,并将特征权值与 HOG 特征进行有效融合,形成 W-HOG 压缩特征。实验数据统计结果表明本文所改进的算法在精确度方面较 CT 算法提高近 16%,较 HOG-CT 算法提高 6% 左右。且在复杂背景下能够较精确地检测出运动手势,提高了手势跟踪在光照变化、背景存在类肤色物体等情况下的鲁棒性,减少手势跟踪漂移的发生。

关键词:压缩感知;Harr 特征;HOG 特征;手势跟踪;跟踪漂移

中图分类号:TP391.9 **文献标志码:**A **文章编号:**1673-4785(2016)01-0124-05

中文引用格式:史东承,倪康.压缩感知 W-HOG 特征的运动手势跟踪[J].智能系统学报,2016,11(1):124-128.

英文引用格式:SHI Dongcheng, NI Kang. Motion gesture tracking based on Compressed sensing W-HOG features[J]. CAAI Transactions on Intelligent Systems, 2016, 11(1): 124-128.

Motion gesture tracking based on compressed sensing W-HOG features

SHI Dongcheng, NI Kang

(College of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China)

Abstract: The use of a single or simple feature for gesture tracking always induces tracking errors. To improve the accuracy of hand tracking, this study uses compressed-sensing motion tracking to track a hand and extracts HOG features in the movement area instead of original generalized class Harr features to track the target. At the same time, to reduce the accumulation of errors of gesture tracking generated by weight series, such as HOG features of different image blocks, the study calculates the HOG feature weight and effectively integrates the weight with HOG features to form W-HOG compression characteristics. The statistical experimental results show that the improved algorithm provided increased accuracy by approximately 16% compared with CT algorithm and approximately 6% compared with HOG-CT algorithm. Moreover, the algorithm can accurately detect the moving gesture in a complex background, improve the tracking robustness of gesture tracking in circumstances of illumination changes and background objects with a color similar to the skin, and reduce the occurrence of gesture tracking drifting.

Keywords: compressed sensing; Harr features; HOG features; gesture tracking; tracking drifting

在运动目标跟踪方面,实时性和鲁棒性是衡量跟踪效果必不可少的 2 个因素。传统方法在实时性和鲁棒性存在或大或小的缺陷。Zhang Kaihua 等 2012 年在 ECCV 会议上提出了基于压缩感知的目标跟踪算法(compress tracking)^[1],该算法在实时性和鲁棒性有了很大的提高,但复杂背景下效果欠佳。文献[2]在压缩感知目标跟踪的基础上结合了运动估计算法,较好地抑制了跟踪漂移,但在较复杂背景

下算法效果不容乐观。文献[3]和文献[4]都引用了特征加权的思想,但都是基于压缩感知跟踪原始特征,故在复杂背景下鲁棒性较差。文献[5]和文献[6]分别将光流法和 SIFT 特征点融入到压缩感知目标跟踪中,在复杂背景下,这 2 种算法均能够较准确进行目标跟踪,但是由于计算量的增加,实时性差。文献[7]提出了基于 HOG 特征的压缩感知目标跟踪,但由于特征误差的累积也容易产生跟踪漂移。本文摒弃了经典 HOG 特征描述中对 HOG 特征的等权值串联的理念,弥补了 HOG 特征的在特征描述上的不足。将 HOG 特征和 HOG 特征权值融合形

成 W-HOG 压缩特征,以此进行运动手势跟踪和对分类器进行更新,以提高算法在复杂背景下的鲁棒性。

1 经典压缩感知目标跟踪简介

Donoho 等于 2006 年首次提出了压缩感知理论,该理论指出,通过一个满足约束等距性条件的稀疏测量矩阵对原图像特征空间做投影,会把保留高维特征的信息压缩到低维特征子空间。文献[1]中的 Zhang Kaihua 等将压缩感知理论运用到运动目标跟踪中去(compressive tracking),此后,该跟踪算法在视频跟踪领域得到广泛应用。

该算法低维特征提取公式为

$$V = PX \quad (1)$$

式中: $X \in R^{n \times 1}$ 为高维的原始特征, $P \in R^{k \times n}$ 为满足约束等距性条件的随机测量矩阵, $V \in R^{k \times 1}$ 为压缩后的低维特征。文献[1]中提出的 X 为各种不同尺度的矩形滤波器与原图像卷积结果的组合,实际为计算一系列矩形内所有像素的灰度差。样本的实际特征提取公式为

$$v_i = \sum_{k=1}^{NR} \frac{p_{i,k} R_{i,k}}{\sqrt{NR}} \quad (2)$$

式中:NR 为图像块的数目,其数值一般在 2~4 之间随机取, $p_{i,k}$ 在 1 和 -1 之间等概率随机取, R 为候选区域随机选取的图像块。

其后的目标跟踪就是通过迭代运算寻找与目标特征最相似的候选区域,文献[7]中首先在后续帧中遍历目标位置临近区域,然后在候选区域选用贝叶斯分类器计算其分类分数,以得分数值最大的候选区域作为目标位置,判别公式^[8]如下:

$$H^k(v) = \sum_{i=1}^k \log\left(\frac{p(v_i | y=1)}{p(v_i | y=0)}\right) \quad (3)$$

式中: $p(v_i | y=1)$ 和 $p(v_i | y=0)$ 满足标准正态分布,后文的分类器更新中将会使用到标准正态分布的均值和方差。我们选取最大的 $H(v)$ 作为跟踪目标,最后根据文献[1]进行相关系数的更新。

2 压缩感知 W-HOG 的运动手势跟踪

HOG 特征是用梯度和方向来表现局部目标的表象和形状,其优点是具有旋转和平移不变性、局部光照不变性。文献[7]提出的基于 HOG 特征的压缩感知目标跟踪,虽然使用 HOG 特征代替原来的灰度特征,但是由于每个 HOG 图像块特征对整体手势跟踪效果的影响存在差异,若等权值相加每个图像块的 HOG 特征,必会对手势跟踪造成较大影响。故

在背景信息中含有噪声点、背景灰度变化较大以及背景中含有类肤色信息时,容易发生跟踪偏离和漂移。本文使用 HOG 特征和特征权重代替原来的灰度特征,从而避免了等权值相加对进行运动手势跟踪所造成的累计误差。

2.1 W-HOG 压缩特征空间的生成

采用 W-HOG 特征代替原来的灰度特征后,原始的高维特征为不同尺度的矩形滤波器与原图像卷积后的结果提取 HOG 特征。为了增加算法对复杂背景的鲁棒性,需要对 HOG 特征和特征权值进行特征融合,形成 W-HOG 特征,该特征为压缩后的低维特征。

2.1.1 梯度方向直方图的构建与特征计算

对于候选区域随机选择不同大小的 2~4 个矩形图像块 $R_{i,k}$,统计每个图像块的梯度方向直方图 $H_{i,k}$ 。将梯度方向分成 bin 块,块内的每个像素点向该方向进行投票加权,运用积分直方图^[9]快速计算该方向上的梯度幅值,继而得到子块的梯度方向直方图,总的梯度方向直方图特征为子块特征的加权和。最终特征^[10]为

$$v_i^k = \sum_{b=1}^{NR} \frac{p_{i,k} \cdot h_{i,k}^b}{\sqrt{NR}}, b = 1, 2, \dots, \text{bin} \quad (4)$$

式中: $h_{i,k}^b$ 为第 b 个像素块的梯度方向直方图。低维梯度直方图特征^[11]形成示意图如图 1 所示。

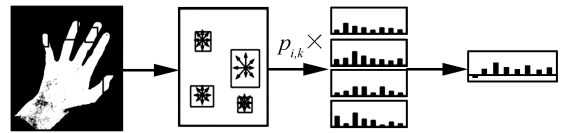


图 1 低维 HOG 特征生成示意图

Fig.1 Schematic diagram of low dimensional HOG feature generating

2.1.2 HOG 特征权值的计算

手势图像中不同图像块中所包含的特征信息量是不同的,信息熵反映了图像中所带信息量的大小,信息熵越大,该图像块所包含信息量就越大,则对手势跟踪的影响也就越大。本文对压缩后的矩形块进行信息熵计算来构造 HOG 加权系数。首先把该手势图像转换成灰度图像,则第 i 个子块信息熵的计算公式如下^[12]:

$$E_i = - \sum_{k=1}^n p_i^k \log_2 p_i^k \quad (5)$$

式中: n 为像素级数,本文取 256; p_i^k 为第 k 级像素点出现的概率。该像素块的 HOG 加权系数计算公式如下:

$$w_i = \frac{E_i}{\sum_{i=1}^m E_i} \quad (6)$$

式中: w_i 为该像素块的加权系数, E_i 为第 i 个子块信息熵^[13]。计算得到的 HOG 特征权值与相应的梯度方向直方图相乘即得到 W-HOG 融合特征。融合后的低维 W-HOG 特征可表示为

$$v_i^k = \sum_{k=1}^{NR} w_i \frac{p_{i,k} \cdot h_{i,k}^b}{\sqrt{NR}}, \quad b = 1, 2, \dots, \text{bin} \quad (7)$$

2.2 算法步骤

1) 初始化分类器

a) 视频序列的第一帧运用坐标信息选择运动手势图像。

b) 正负样本的采集: 在 $0 < \|I(x) - I\| < \alpha$ 范围内取正样本; 在 $\beta < \|I(x) - I\| < \gamma$ 范围内取负样本, 并计算 $0 < \|I(x) - I\| < \gamma$ 范围内的积分直方图并保存。

c) 利用积分直方图, 快速计算运动手势目标和背景的梯度直方图^[14]。

$$\begin{cases} \mu_i^1 = \lambda \mu_i^1 + (1 - \lambda) \mu^1 \\ \sigma_i^1 = \sqrt{\lambda (\sigma_i^1)^2 + (1 - \lambda) (\sigma^1)^2 + \lambda(1 - \lambda) (\mu_i^1 - \mu^1)^2} \end{cases} \quad (8)$$

继而按照式(7)计算融合低维 W-HOG 特征, 按

照式(8)更新分类器的均值和方差。

2) 运动手势目标跟踪

读取下一图像帧, 在 $0 < \|I(x) - I\| < \gamma$ 周围采集手势样本, 并根据算式(7)计算候选区域样本的低维 W-HOG 特征, 继而代入算式(3), 取其最大值作为运动手势目标区域。

3) 分类器的更新

在所确定的运动手势目标周围采集正负样本^[15], 并计算正负样本 HOG 特征, 根据算式(8)分别更新分类器。

4) 跳转到步骤 2) 进行下一帧的处理。

3 实验结果和分析

本实验背景选择为复杂静态背景, 背景中含有多个类肤色物体的存在, 自行拍摄的运动手势视频约为 9 秒, 经拆帧视频拆分为 260 帧, 另外本实验还对该视频分别进行 CT 跟踪以及基于 HOG 特征的 CT 跟踪进行对比实验。实验中, 更新率取值为 0.6、 $\alpha = 4$ 、 $\beta = 8$ 、 $\gamma = 20$ 。实验环境为 Intel Core i5 CPU 3.0GHZ, 安装内存: 2.0GB, Window7 操作系统, 使用 MATLAB 2014b 进行仿真实验。经仿真实验后, 视频序列第 95 帧仿真结果如图 2 所示, 视频序列第 197 帧仿真结果如图 3 所示。



图 2 第 95 帧仿真结果

Fig.2 Simulation results of Video sequence frame95



图 3 第 197 帧仿真结果

Fig.3 Simulation results of Video sequence frame197

经对比实验可以看出,基于 W-HOG 的压缩感知目标跟踪算法在背景复杂、背景中含有类肤色物体等情况下能够准确地检测出运动手势。表 1 为 3 种跟踪算法运动手势跟踪准确率的统计结果,其中总帧数为 260。

表 1 准确性统计结果
Table 1 Results statistics of accuracy

算法	正确跟踪的帧数	正确率/%
CT	197	76
HOG-CT	223	86
W-HOG-CT	239	92

从表格数据统计可以看出:较其他 2 种手势跟踪算法,基于 W-HOG 的压缩感知运动手势跟踪算法手势跟踪正确率较高。特征对每个图像块的 HOG 特征进行加权处理,摒弃了经典 HOG 特征描述中对 HOG 特征的等权值串联的理念,弥补了 HOG 特征在特征描述上的不足。为验证该算法的实时性,本文对该算法的运行时间进行统计,统计结果如表 2 所示,其中总帧数数为 260。

表 2 实时性统计结果
Table 2 Results statistics of the real time

算法	运行时间/s	运行速率/(frame · s) ⁻¹
CT	13.70	19
HOG-CT	23.70	11
W-HOG-CT	25.90	10

从表格数据统计可以看出:基于 W-HOG 的压缩感知运动手势跟踪运行时间较其他两种算法运行时间稍长,造成该现象的主要原因是 W-HOG 特征的计算增加了对图像块特征权值的计算。虽然 HOG 特征的计算过程中增加了积分直方图以加快特征的计算,但较经典 HOG 特征,计算复杂程度较高。从手势跟踪对算法实时性的要求来看,该速率已经能够满足基本要求,不会造成跟踪延迟等问题的发生。

4 结束语

本文将压缩感知运动目标跟踪运用到了运动手势跟踪中,并且将 HOG 特征替代原来的类 Harr 特征,同时运用 HOG 特征权值进行特征融合,确定运动手势的位置,最后进行分类器的更新,完成了运动手势跟踪。该算法摒弃了经典 HOG 特征描述中对 HOG 特征的等权值串联的理念,有效减少了 HOG 特征等权值相加对手势跟踪所造成的累积误差,并

且能够在较复杂环境下准确地进行运动手势跟踪,本算法还运用积分直方图减少运算量,但是手势跟踪的实时性有待提高。在尺度方面,该算法不能够有效地满足运动手势跟踪在尺度自适应方面的要求,手势发生形变时,手势跟踪产生一定的误差。因此,在实时性的提高和尺度自适应方面是以后研究工作的重点。

参考文献:

[1]ZHANG Kaihua, ZHANG Lei, YANG M H. Real-time compressive tracking [C]//Proceedings of the 12th European Conference on Computer Vision. Florence, Italy, 2012: 866-879.

[2]WU Jiayun, CHEN Daquan, YI Rui. Real-time compressive tracking with motion estimation [C]//Proceedings of IEEE International Conference on Robotics and Biomimetics (RO-BIO). Shenzhen, China, 2013: 2374-2379.

[3]ZHU Jianzhang, MA Yue, ZHENG Chen, et al. Adaptive weighted real-time compressive tracking [J]. IET computer vision, 2014, 8(6): 740-752.

[4]王松林, 项欣光. 基于压缩感知的多特征加权目标跟踪算法 [J]. 计算机应用研究, 2014, 31(3): 929-932.

WANG Songlin, XIANG Xinguang. Real-time tracking using multi-feature weighting based on compressive sensing [J]. Application research of computers, 2014, 31(3): 929-932.

[5]刘丹, 窦勇. 基于光流和压缩感知的目标跟踪 [J]. 计算机技术与自动化, 2013, 32(4): 84-87.

LIU Dan, DOU Yong. Object tracking based on optical flow and compressive sensing [J]. Computer technology and automation, 2013, 32(4): 84-87.

[6]钟权, 周进, 崔雄文. 融合 SIFT 特征的压缩跟踪算法 [J]. 光电工程, 2015, 42(2): 66-72.

ZHONG Quan, ZHOU Jin, CUI Xiongwen. Compressive tracking algorithm based on SIFT [J]. Opto-Electronic Engineering, 2015, 42(2): 66-72.

[7]孙晓燕, 常发亮. 梯度特征稀疏表示目标跟踪 [J]. 光学精密工程, 2013, 21(12): 3191-3197.

SUN Xiaoyan, CHANG Faliang. Object tracking based on sparse representation of gradient feature [J]. Optics and precision engineering, 2013, 21(12): 3191-3197.

[8]LIU Jin, HAN Chongzhao, HAN Feng. A novel compressed sensing based track before detect algorithm for tracking multiple targets [C]//Proceedings of the 16th International Conference on Information Fusion (FUSION). Istanbul, the Republic of Turkey, 2013: 1514-1519.

[9]LI Hanxi, SHEN Chunhua, SHI Qinfeng. Real-time visual tracking using compressive sensing [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, USA, 2011: 1305-1312.

[10]张雷, 王延杰, 何舒文. 基于相位一致性的实时压缩跟踪方法 [J]. 光子学报, 2014, 43(8): 124-131.

ZHANG Lei, WANG Yanjie, HE Shuwen. Real-time com-

- pressive tracking method based on phase congruency[J]. Acta photonica sinica, 2014, 43(8): 124-131.
- [11] 张洁玉, 武小川. 加权局部二值模式的人脸特征提取[J]. 中国图象图形学报, 2014, 19(12): 1794-1801.
- ZHANG Jieyu, WU Xiaochuan. Feature extraction of faces based on a weighted local binary pattern[J]. Journal of image and graphics, 2014, 19(12): 1794-1801.
- [12] 朱秋平, 颜佳, 张虎, 等. 基于压缩感知的多特征实时跟踪[J]. 光学精密工程, 2013, 21(2): 437-444.
- ZHU Qiuping, YAN Jia, ZHANG Hu, et al. Real-time tracking using multiple features based on compressive sensing[J]. Optics and precision engineering, 2013, 21(2): 437-444.
- [13] 王梅, 屠大维, 周许超. SIFT 特征匹配和差分相乘融合的运动目标检测[J]. 光学精密工程, 2011, 19(4): 892-899.
- WANG Mei, TU Dawei, ZHOU Xuchao. Moving object detection by combining SIFT and differential multiplication[J]. Optics and precision engineering, 2011, 19(4): 892-899.
- [14] HASAN H S, KAREEM S A. Human computer interaction for vision based hand gesture recognition: a survey[C]// Proceedings of International Conference on Advanced Computer Science Applications and Technologies. Kuala Lumpur, Malaysia, 2012: 55-60.

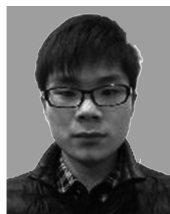
- [15] 宋海东, 陈二虎, 边敏. 基于 HOG-LBP 特征的自适应实时手掌检测算法[J]. 计算机与数字工程, 2013, 41(11): 1826-1828.

SONG Donghai, CHEN Erhu, BIAN Min. HOG-LBP feature adaptive real-time palm detecting algorithm[J]. Computer and digital engineering, 2013, 41(11): 1826-1828.

作者简介:



史东承,男,1959 年生,教授。主要研究方向为图像处理与机器视觉。主持完成信息产业部重点项目 1 项,吉林省省级课题 6 项。曾获中国通信学会科学技术二等奖 1 项、吉林省自然科学学术成果二等奖 1 项。发表学术论文 50 余篇。



倪康,男,1991 年生,硕士研究生,主要研究方向为图像处理与机器视觉,发表学术论文 3 篇。

全国第 15 届可拓学年会征文启事

全国第 15 届可拓学年会定于 2016 年 8 月 10-12 日在浙江省宁波市浙江大学宁波理工学院召开。会议将交流近年来可拓学的理论和应用研究成果,讨论可拓学发展方向、学科建设和人才培养等问题。为此诚挚邀请国内外相关领域的专家学者、管理人员和工程技术人员参加这次会议。

本次会议征文包括理论研究和应用研究成果两部分。

1. 理论研究

1) 基元理论;2) 可拓数学;3) 可拓逻辑;4) 可拓学的哲学基础;4) 可拓创新方法;5) 可拓学与其他学科的交叉融合。具体形式为学术论文。

2. 应用研究

可拓论和可拓创新方法与各专业领域相结合产生的可拓工程理论、方法与应用研究的论文,如(但不限于此):人工智能、计算机、自动化、控制、检测、设计、管理、经济等。

征文要求:

1) 文笔精练、资料可靠,字数不超过 7 千字,并附中英文摘要和关键词;

2) 论文应是新的研究成果,严格按照科技论文规范撰写,且未在国内外学术会议或刊物上发表过。

此次会议由中国人工智能学会主办,由中国人工智能学会可拓学专业委员会、浙江大学宁波理工学院和《智能系统学报》编辑部联合承办。会议征文中被大会录用的论文,需要作者参加会议交流,再由可拓学学术委员会评审出优秀论文,推荐到《智能系统学报》、《数学的实践与认识》等核心期刊和《广东工业大学学报》(科技核心期刊)“可拓论坛”正式发表。欢迎相关领域的研究人员、高校师生、企业家、工程技术人员、以及一切爱好和有志于可拓学理论研究与应用研究的朋友踊跃投稿。

本届大会的征文采用 email 投稿的方式,投稿 email 地址:extenics@vip.163.com

重要日期:

征文截止日期:2016 年 5 月 30 日

录用通知日期:2016 年 6 月 30 日

会议联系地址:广东工业大学可拓学与创新方法研究所

Email:extenics@vip.163.com

电话:020-39322973 传真:020-39322019

地址:广州大学城外环西路 100 号广东工业大学可拓学与创新方法研究所(邮编:510006)

联系人:李剑明老师,汤龙博士