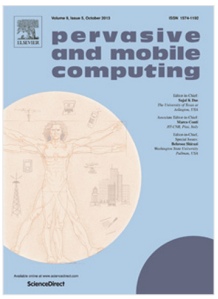
Accepted Manuscript



Learning multi-level features for sensor-based human action recognition

Yan Xu, Zhengyang Shen, Xin Zhang, Yifan Gao, Shujian Deng,
Yipei Wang, Yubo Fan, Eric I-Chao Chang

PII: S1574-1192(17)30114-1
DOI: http://dx.doi.org/10.1016/j.pmcj.2017.07.001
Reference: PMCJ 867

To appear in: *Pervasive and Mobile Computing*

Received date : 28 February 2017
Revised date : 25 May 2017
Accepted date : 5 July 2017

Please cite this article as: Y. Xu, Z. Shen, X. Zhang, Y. Gao, S. Deng, Y. Wang, Y. Fan, E. Chang, Learning multi-level features for sensor-based human action recognition, *Pervasive and Mobile Computing* (2017), http://dx.doi.org/10.1016/j.pmcj.2017.07.001

Learning Multi-level Features For Sensor-based Human Action Recognition

Yan Xu[a,b], Zhengyang Shen[a], Xin Zhang[a], Yifan Gao[c], Shujian Deng[a], Yipei Wang[a], Yubo Fan[a,*], Eric I-Chao Chang[b,*]

[a] *State Key Laboratory of Software Development Environment and Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education and Research Institute of Beihang University in Shenzhen, Beihang University, Beijing 100191, China*
[b] *Microsoft Research Asia, Beijing 100080, China*
[c] *School of Electronic and Information Engineering, Beihang University, Beijing 100191, China*

Abstract

This paper proposes a multi-level feature learning framework for human action recognition using a single body-worn inertial sensor. The framework consists of three phases, respectively designed to analyze signal-based (low-level), components (mid-level) and semantic (high-level) information. Low-level features capture the time and frequency domain property while mid-level representations learn the composition of the action. The Max-margin Latent Pattern Learning (MLPL) method is proposed to learn high-level semantic descriptions of latent action patterns as the output of our framework. The proposed method achieves the state-of-the-art performances, 88.7%, 98.8% and 72.6% (weighted F_1 score) respectively, on Skoda, WISDM and OPP datasets.

Keywords: Multi-level, Human action recognition, Latent pattern, High-level, Semantic

1. Introduction

In recent years, sensor-based human action recognition (HAR) plays a key role in the area of ubiquitous computing due to its wide application in daily life [1–6]. In most cases, utilization of raw data without special process is impractical since raw recordings confound noise and meaningless components. Therefore, the process of feature extraction and feature learning is supposed to be conducted [7], which is commonly based on the sliding window

*Corresponding author
Email addresses: xuyan04@gmail.com (Yan Xu), hbgtjxzbbx@gmail.com (Zhengyang Shen), xinzhang0376@gmail.com (Xin Zhang), goghivan0017@gmail.com (Yifan Gao), cxwldsj@gmail.com (Shujian Deng), edithwang525@gmail.com (Yipei Wang), yubofan@buaa.edu.cn (Yubo Fan), echang@microsoft.com (Eric I-Chao Chang)