Demo: LIVES: Learning through Interactive Video and Emotion-aware System

Min Chen
EPIC Lab
Huazhong Univ. of Science and Technology
Wuhan, China
minchen@ieee.org

Yixue Hao
EPIC Lab
Huazhong Univ. of Science and Technology
Wuhan, China
yixue.epic@gmail.com

Yong Li
Department of Electronic Engineering
Tsinghua University
Beijing, China
liyong07@tsinghua.edu.cn

Di Wu
School of Info. Sci. & Tech.
Sun Yat-Sen University
Guangzhou, China
wudi27@mail.sysu.edu.cn

Dijiang Huang
Department of CSE
Arizona State University
Arizona, USA
Dijiang.Huang@asu.edu

ABSTRACT

In order to improve the accuracy and efficiency of emotion recognition, we design a novel system called Learning through Interactive Video and Emotion-aware System (LIVES). LIVES includes data collection, emotion recognition, and result validation, as well as emotion feedback. We adopt transfer learning to label and validate moods in LIVES, while the emotion can be classified into six types of mood in a reasonable accuracy. Through transfer learning, the time-consuming and labor-intensive processing cost on data collection and labeling can also be greatly reduced. In our prototype system, LIVES is used to enhance an emotion-aware robot's intelligence provided by cloud. LIVES-based emotion recognition is executed in the remote cloud while corresponding result is sent to the robot for emotion feedback. The experimental results demonstrate LIVES significantly improves the accuracy and effectiveness of emotion classification.

Keywords

Affective interaction, transfer learning, sentiment anlysis

1. INTRODUCTION

Nowadays, there are lots of works on emotion recognition. Previous work on emotion recognition can be categorized as follows: (1) body signal-based recognition [4]; (2) audio-visual data based recognition, including method through face video processing [5], and (3) recognition based on the usage pattern of smartphone [7]. However, it is challenging to fuse large-scale heterogeneous emotion data in multi-dimensional spaces for analyzing user's emotion while proactively activating surrounding hardware resources to perform personalized actions to comfort user's emotion. LIVES based solution is the first to tackle above challenges through multi-space emotion data collection, emotion recognition, validation and feedback. Especially, LIVES explores transfer learning on top of robotics and cloud-terminal integration technologies.

MobiHoc'15, June 22–25, 2015, Hangzhou, China.
ACM 978-1-4503-3489-1/15/06.
http://dx.doi.org/10.1145/2746285.2764928 .

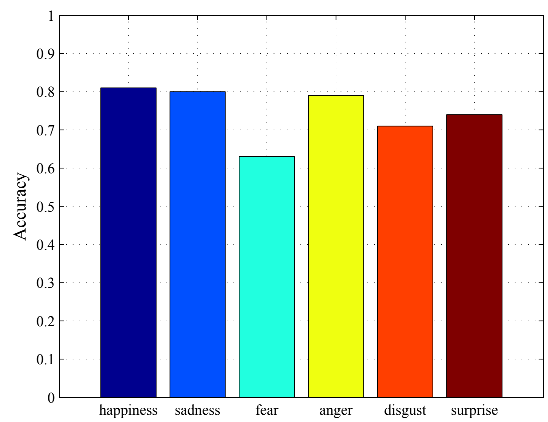
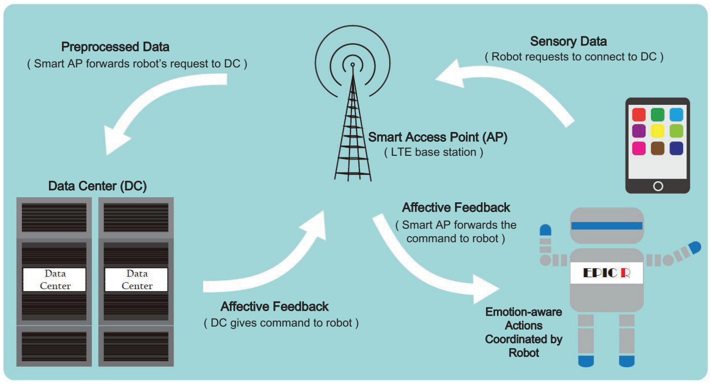


Figure 1: Accuracy of mood recognition

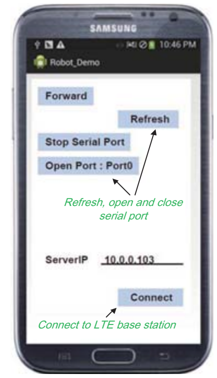
2. LIVES FRAMEWORK

2.1 LIVES Description

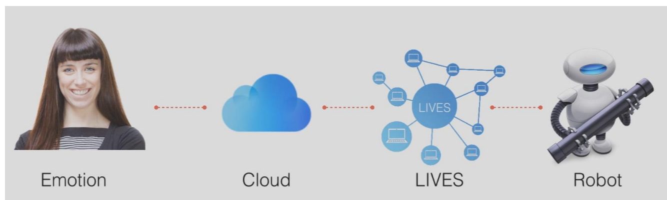
In LIVES, user's physical, cyber and social domain's data are first collected by wearable device, mobile devices and network devices. Physical data include physiological data, activity level, location, environmental and face/interactive video. Cyber data consist of call, SMS, email, application usage, WiFi and network control. Social network data include SNS content or image post, repost and comment. Then, the data are preprocessed. After preprocessing, we extract the data feature. For the emotion space, the data are labeled and validated by transfer learning based on previously labeled results. Now we introduce the label and validate phase. In the label phase, the source domain data x_s input come from ten users' personal information in terms of physical data, cyber and social network data and moods. The target domain data x_t input data come from other personal information. Our goal is to estimate the mood c probability $p(c|x_t)$, since the x_s and x_t may be in a different feature space, So we first need to find a translator to link the two feature spaces [3]. since the f_s and f_t are features and are conditionally independent x_s, where $x_s \in \chi_s$, so we can calculate $p(f_s, f_t)$ as



(a)



(b)



(c)

Figure 2: LIVES based Testbed: (a) System architecture; (b) Control messages shown in LTE cell phone; (c) Test scenario.

follows: $p(f_t, f_s) = \int_{\chi_s} p(f_t, x_s)p(f_s|x_s)dx_s$, Now we link the unlabel feature f_t and x_s through Jensen-Shannon divergence. In the validation phase, We overhear users' input of their own emotion when they use some applications such as Moodagent. But one principal problem is that, the mood we overhear is not the same as mood label space C. Thus, we need to compare the similarity of the label space C and the user's input mood as $M = \{m_1, \ldots, m_n\}$ which is collected in each time slot as ground-truth label. Using a framework similar to transfer learning [6], we can find the similarity between mood c and m as follows: $\mathrm{sim}(c, m) = \mathrm{MMD}^2[D_c, D_m]$, The $\mathrm{MMD}^2[D_c, D_m]$ is the maximum mean discrepancy [1], we validate the label which the most similar in C, Through transfer learning the time-consuming and labor-intensive processing cost can be reduced extensively. After a certain time of labeling and validation through transfer learning, the training sets are established. By the utilization of Hidden Markov model, the emotion data are classified into six mood. These mood classification results are forwarded to EPIC robot which is developed in Embedded and Pervasive Computing (EPIC) lab. During emotion feedback stage, EPIC robot system carries out hardware resource cognition then performs a series of actions to comfort user's emotion using Markov decision process. For example, a personalized video will be shown in an available screen such as a projector. If the user is not satisfied with the emotion feedback actions, video contents and display method can be fine tuned through continuous learning until user's emotions are cared more accurately and efficiently. This is called feedback data through the interaction between user and system. In order to improve the accuracy of emotion recognition, we use physical data, cyber data, social network data and feedback data in next time.

2.2 Demonstration

As shown in Fig. 2(a), we implement LIVES in AIWAC testbed [2], which consists of robot, smart access point (AP), and data center (DC). The testbed provides a version of software available to run in Linux in the DC, which generates command queue according the analyzed mood and these commands are transmitted from DC to the robot via TCP protocol. Fig. 2(b) illustrates control messages shown in LTE cell phone. In Fig. 2(c), we use LIVES to predicate the user's mood compared with the mood given by users, we can obtain the accuracy as shown in Fig. 1, where the recognition of user's moods is achieved with acceptable accuracy which is higher than Moodscope [7]. When we find people in a unpleasant mood, we use robot to do some interactive actions to care user's mood.

Acknowledgment

Many thanks to Yujun Ma, Yin Zhang, Hang Ruan, Wenye Wang, Shan Zhao, Chuanbei Wu for making tremendous system-building efforts on the testbed, where LIVES experiments are run.

3. REFERENCES

[1] K. Borgwardt, A. Gretton, and et al. Integrating stuctured biological data by kernel maximum mean discrepancy. In *Proc. of ISMB'06*, 2006.

[2] M. Chen, Y. Zhang, Y. Li, and et al. Aiwac: affective interaction through wearable computing and cloud technology. *IEEE Wireless Communications*, 22(1):20–27, Feb. 2015.

[3] W. Dai, Y. Chen, G.-R. Xue, Q. Yang, and Y. Yu. Translated learning: Transfer learning across different feature spaces. In *Proc. of ACM NIPS*. ACM, 2008.

[4] A. Gluhak, M. Presser, L. Zhu, and et al. Towards mood based mobile services and applications. In *Proc. of ACM EuroSSC*, pages 159–174. ACM, 2007.

[5] J. Hernandez, M. Hoque, W. Drevo, and R. Picard. Mood meter: Counting smiles in the wild. In *Proc. of Int'l Conf. on Ubiquitous Computing (UbiComp)*, pages 301–310. ACM, Sept. 2012.

[6] D. H. Hu, V. W. Zheng, and Q. Yang. Cross-domain activity recognition via transfer learning. *Pervasive and Mobile Computing*, 7(3):344–358, 2011.

[7] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proc. of ACM MobiSys'13*, pages 84–89. ACM, June 2013.