Activity Recognition Exploiting Classifier Level Fusion of Acceleration and Physiological Signals

Haodong Guo
College of Computer Science
Zhejiang University
Hangzhou, 310027, China
hdguo@zju.edu.cn

Gencai Chen
College of Computer Science
Zhejiang University
Hangzhou, 310027, China
chengc@zju.edu.cn

Ling Chen
College of Computer Science
Zhejiang University
Hangzhou, 310027, China
lingchen@zju.edu.cn

Yanbin Shen
College of Computer Science
Zhejiang University
Hangzhou, 310027, China
tcsyb@zju.edu.cn

UbiComp'14 Adjunct, September 13-17, 2014, Seattle, WA, USA
ACM 978-1-4503-3047-3/14/09.
http://dx.doi.org/10.1145/2638728.2638777

Abstract
We investigate how to effectively combine physiological signals with acceleration signals to conduct activity recognition task. Firstly, features are extracted from acceleration and physiological signals, including heart rate variability (HRV). Secondly, classifier level fusion is utilized to combine the models built by acceleration and physiological features separately. Experiment results show that activity recognition task can benefit from HRV features, and classifier level fusion has its superiority over feature level fusion.

Author Keywords
Activity recognition; Wearable computing; Fusion

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

Introduction
Physiological signals have tight relationships with activities, e.g. heart rate would increase while crawling up the stairs. Therefore, employing physiological signals is expected to improve the performance of activity recognition systems. Centinela [1] is a representative activity recognition system based on acceleration and physiological signals. Transient

features [1] are extracted from physiological signals to distinguish certain activities similar in acceleration features. However, the features are strongly dependent on the slope of the line that best fits the physiological signal and only three states of physiological changes (i.e., increasing, decreasing, and constant) are considered. Moreover, Centinela conventionally combines different modalities in feature level, and has the feature compatibility problem [2].

To counter the above problems, we conduct activity recognition exploiting classifier level fusion of acceleration and physiological signals. The contributions of this work are as follows:

1. Use heart rate variability (HRV) features in activity recognition. Own to the relationships of HRV features and human activities, it is expected that using HRV features would improve the accuracy of activity recognition.
2. Exploit stacking framework to realize classifier level fusion of acceleration and physiological signals, which can address the compatible problem of feature level fusion.
3. Evaluate our approach on real-life naturalized setting dataset and perform extensive comparison with other methods.

Methodology

Our research problem can be formulated as, given a collection of raw sensor signals $TS=\{\{ts_1, \ldots, ts_m\}, \{ts_{m+1}, \ldots, ts_{m+n}\}\}=TS_a \cup TS_p$, where TS_a is the set of acceleration signals and TS_p is the set of physiological signals, how to build proper models on diffierent signals, and then fuze the multiple models to achieve better performance of activity recognition. As depicted in Figure 1, sensor data is transformed into features and then multiple models are built on these features. Stacking, an ensemble learning method, is modified to combine acceleration and physiological modalities.

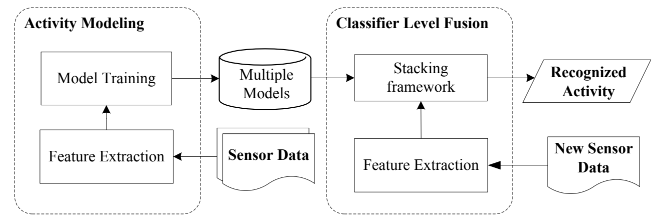

Figure 1: The architecture of the proposed method

Data Collection

We have developed an Android-based data acquisition tool, and the screenshot is shown in Figure 2. We use smart phone and BioHarness™ BT chest sensor strap[1] as our data collection devices. The strap provides 3D acceleration and various physiological signals, e.g. heart rate, heart rate RR intervals (i.e., duration of two R wave), respiration rate, breath amplitude, skin temperature, posture, electrocardiogram amplitude. Additionally, equipped with Bluetooth, the strap can easily communicate with a smart phone. By employing the tool, we establish our dataset, which include the acceleration and physiological data of 26 participants (all undergraduate or postgraduate students, 20 males and 6 females) on five activities (i.e. running, walking, sitting, ascending, and descending).

Feature Extraction

Similar to [1], all signals are divided into fixed size time windows (i.e., 6s, 12s, and 24s) with 50% overlap to

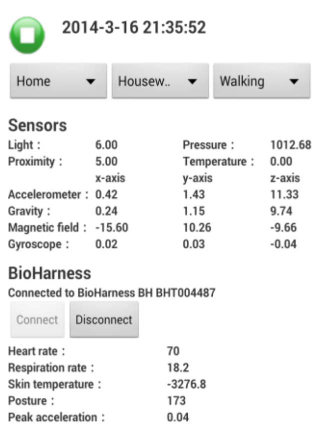

Figure 2. Our Android-based data acquisition tool.

[1] http://www.zephyr-technology.com/bioharness-bt.html

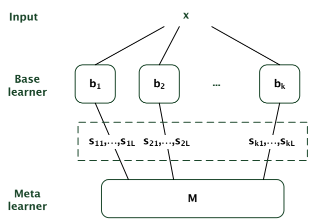

Figure 3. The standard score-based stacking algorithm.

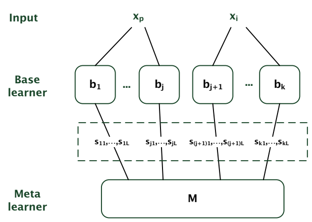

Figure 4. Our modified score-based stacking algorithm.

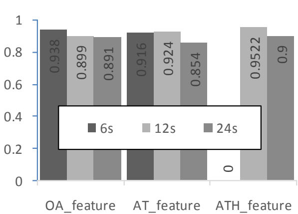

Figure 5. Classification accuracy under different time window sizes using J48 classifier (6s is too small to extract HRV features, OA, AT, ATH denote datasets D_{oa}, D_{at},D_{ath},respectively).

overcome the issue of detecting activity transitions. For each time window, eight statistical features are extracted from acceleration signal, and nine structural features and two transient features [1] are extracted from each physiological signal. Moreover, we also extract time and frequency domain features of HRV, including mean and standard deviation of heart rate, absolute very low-frequency (VLF, 0.0033-0.04 Hz), absolute low-frequency (LF, 0.04-0.15 Hz), absolute high-frequency (HF, 0.15-0.4Hz), and ratio of LF to HF. HRV describes the variations between consecutive heartbeats. Stress, certain cardiac diseases, and other physiological states can affect the HRV [3]. Therefore, we attempt to exploit HRV in activity recognition.

Multiple Modalities Combination

Considering the compatibility issues arising from different time shifts, window length configurations, and sampling frequencies, fusion acceleration and physiological modalities in feature level would weaken the utility of physiological features. While fusing in classifier level, the problem no longer exists, because the outputs from different classifiers have uniform expressions, e.g. predictive scores. We modify stacking [4], an ensemble learning method, to combine the predictions of several learning algorithms. Stacking gives us a combination framework that can be easily utilized in our work.

Let X and Y denote the space of inputs and the set of class labels, respectively, assuming $Y = \{l_1, l_2, \ldots, l_L\}$, where L is the number of class labels. Given training data set $TS = \{(x_1, y_1), \ldots,(x_m, y_m)\}$, where $x_i \in X$ and $y_i \in Y$ (i = 1, …, m), suppose that there are K base machine learning classifiers: $b1$, $b2$, …, b_K with b_i: $x \rightarrow Y$, where $\forall\ i$, $b_i = \arg\max_{l \in Y} s_{i,l}(x)$, $s_{i,l}(x)$ is the predictive score returned by the classifier b_i when the input x is labeled with l. And the score $s_{i,l}(x)$ is usually the conditional probability $p_i(l|x)$ in statistical machine learning methods. Score-based stacking (see Figure 3) tries to find a meta-learner M: $S \rightarrow$Y, where $S = \{s_{11}(x), s_{12}(x), \ldots, s_{1L}(x), \ldots, s_{i1}(x), \ldots, s_{ij}(x), \ldots, s_{KL}(x)\}$, $s_{ij}(x)$ is the predictive scores returned by the ith base classifier for the class label j for $1 \leqslant i \leqslant K$ and $1 \leqslant j \leqslant L$. To combine physiological modality with acceleration modality properly, the base classifiers are built on their data separately and then fuse their scores in a meta level classifier. Figure 4 presents our modified score-based stacking algorithm, which can ensure the independence of different modalities in classifier level.

Experiments

To evaluate our approach, we construct five different feature sets: D_{oa} only contains the features of acceleration data; D_{ot} only includes transient features; D_{oh} only includes HRV features; D_{at} has both acceleration and transient features, which are combined in feature level; D_{ath} contains acceleration, transient features and HRV measures, which are also combined in feature level. Each dataset has three versions with different time window sizes. Figure 5 presents classification accuracy under different time window sizes. The results show that time window size 12s make the best performance. Therefore, we set time window size to 12s in the following experiments.

Analyzing the impact of physiological features

To quantify the benefit obtained from combining physiological signals, we compare activity classification accuracy of three datasets (i.e., D_{oa}, D_{at} and D_{ath}). In this experiment, different features are combined into a long feature by feature level fusion. We use J48

	OA	$AT_{feature}$	$AT_{a\text{-}stack}$	$ATH_{feature}$	$ATH_{a\text{-}stack}$
Walking	80.0	87.0	80.9	84.5	86.6
Running	100	92.3	96.0	100	96.1
Ascending	79.2	79.2	90.7	90.4	91.0
Descending	94.6	94.6	87.1	89.9	96.4
Sitting	95.7	100	98.4	98.2	100
Total	89.9	90.6	90.6	92.6	94.0

Table 1. Performance (%) of different level fusion on our datasets ("OA" denotes the dataset D_{oa}, "feature" means feature level fusion, "a-stack" means classifier level fusion).

classifier (an implementation of c4.5 algorithm in Weka[2]) and 10-fold cross validation. As depicted in Table 1, the results (white columns) show that combing physiological signals can significantly improve the classification accuracy for certain activities, especially for ascending, which tends to be misclassified as walking while only using acceleration data. And combining HRV features can improve the overall accuracy. It is also worthy to notice that to some activities, e.g. running, acceleration is sufficient to capture the activity pattern, while combination with physiological signals in feature level would weaken the classification performance.

Analyzing the impact of combination strategies

To further quantify the benefit obtained from combination strategy, we also compare activity classification accuracy under different combination strategies, i.e. feature level combination and classifier level combination. D_{at}, D_{ath} are used in feature level combination, and D_{oa}, D_{ot} and D_{oh} are used in classifier level combination. Our modified stacking method is used as classifier level combination strategy, and J48 and multinomial logistic regression are used as base classifier and meta classifier, respectively. 10-fold cross validation is also used in the experiment. In Table 1 (gray columns), it can be seen that to most activities, classification accuracy is improved by using classifier level fusion. The results suggest that combining acceleration and physiological signals in classifier level is more effective than in feature level, and our modified stacking method is an effective classifier level combination method.

[2] http://www.cs.waikato.ac.nz/ml/weka/.

Conclusion

In this paper, we explored the challenging issue of effectively combining acceleration and physiological signals to conduct activity recognition task. We employ HRV features in activity recognition and present a modified classifier level combination method based on the stacking framework. Experimental results show that the HRV features do have their positive effects on discriminating certain activities, and classifier level combination can achieve higher performance than feature level combination in activity recognition.

Acknowledgments

This work was funded by the Ministry of Industry and Information Technology of China (No. 2010ZX01042-002-003-001), the Natural Science Foundation of China (Nos. 60703040 and 61332017), the Science and Technology Department of Zhejiang Province (Nos. 2007C13019 and 2011C13042).

References

[1] Lara, Ó.D., Pérez, A.J., Labrador, M.A., Posada, J.D. Centinela: A human activity recognition system based on acceleration and vital sign data. *Pervasive and mobile computing* (2012), 8(5), 717-729.

[2] Li, M., Rozgic, V., Thatte, G., Lee, S., Emken, B.A., Annavaram, M., Narayanan, S. Multimodal physical activity recognition by fusing temporal and cepstral information. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2010), 18(4), 369-380.

[3] Berntson, G., Jr., J.B., Eckberg, D., Grossman, P., Kaufmann, P., Malik, M., Nagaraja, H., Porges, S., Saul, J., Stone, P., Molen, M.V.D. Heart rate variability: origins, methods, and interpretive caveats, *Psychophysiology* (1997), 34, 623-648.

[4] Wolpert, D. H. Stacked generalization. *Neural networks* (1992), 5(2), 241-259.