Emotion recognition based on customized smart bracelet with built-in accelerometer

Zhan Zhang[1], Yufei Song[2], Liqing Cui[2], Xiaoqian Liu[2] and Tingshao Zhu[2]

[1] School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, China
[2] Institute of Psychology, Chinese Academy of Sciences, Beijing, China

ABSTRACT

Background: Recently, emotion recognition has become a hot topic in human-computer interaction. If computers could understand human emotions, they could interact better with their users. This paper proposes a novel method to recognize human emotions (neutral, happy, and angry) using a smart bracelet with built-in accelerometer.

Methods: In this study, a total of 123 participants were instructed to wear a customized smart bracelet with built-in accelerometer that can track and record their movements. Firstly, participants walked two minutes as normal, which served as walking behaviors in a neutral emotion condition. Participants then watched emotional film clips to elicit emotions (happy and angry). The time interval between watching two clips was more than four hours. After watching film clips, they walked for one minute, which served as walking behaviors in a happy or angry emotion condition. We collected raw data from the bracelet and extracted a few features from raw data. Based on these features, we built classification models for classifying three types of emotions (neutral, happy, and angry).

Results and Discussion: For two-category classification, the classification accuracy can reach 91.3% (neutral vs. angry), 88.5% (neutral vs. happy), and 88.5% (happy vs. angry), respectively; while, for the differentiation among three types of emotions (neutral, happy, and angry), the accuracy can reach 81.2%.

Conclusions: Using wearable devices, we found it is possible to recognize human emotions (neutral, happy, and angry) with fair accuracy. Results of this study may be useful to improve the performance of human-computer interaction.

INTRODUCTION

The recognition of emotions plays an important role in human-computer interaction (*Chen et al., 2015*). If a computer could recognize emotions, it would be able to interact better with human being (*Beale & Peter, 2008*; *Brave & Nass, 2003*; *Peter, Crane & Beale, 2006*). In psychology, emotion is defined as a complex state that consists of a subjective experience (how we experience emotion), a physiological response (how our bodies react to emotion), and an expressive response (how we behave in response to emotion) (*James, 1884*; *Smith & Lazarus, 1990*). It suggests that the observable aspects of

emotion (physiological and expressive components) could be used as indicators of emotional state, such as facial expressions, speech, physiological parameters, gestures, and body movements (*Peter & Beale, 2008*; *Zhang et al., 2013*).

Evidence exists that human emotions are expressed in walking (*Karg, Kühnlenz & Buss, 2010*; *Montepare et al., 1999*) to some extent. *Michalak et al. (2009)* found that both sadness and depression can be recognized through walking styles. *Crane & Gross (2007)* found that emotion is associated with body movements, including velocity, cadence, head orientation, and shoulder and elbow range of motion. *Montepare et al. (1999)* reported that human emotions can be identified through gait information. They found that among emotions (neutral, happy, sad, and angry), sad and angry can be identified more easily from gait information. Recently, a few studies have attempted to build computational models for human identification using gait (*Han & Bhanu, 2006*; *Kale et al., 2004*). However, computational methods of gait-based emotion detection have not yet been fully established. More importantly, human emotions change over time, only if we can access gait information ubiquitously, otherwise we are not able to implement real-time recognition of emotions. Traditional methods (e.g. self-report, interviewing, and observation) may lead to delayed reporting, which fail to meet the requirement.

The emergence of wearable smart devices shed light on this direction. Wearable smart devices are accessories incorporating computer and advanced electronic technologies, incorporated with various types of sensors. Taking the accelerometer sensor as an example, it can serve as a tool for tracking and recording a user's daily movements. *Lee & Mase (2002)* found that it is possible to recognize human activities through the use of wearable sensors. *Bao & Intille (2004)* placed wearable biaxial accelerometers at different body positions to collect acceleration data for recognizing human activities. *He & Jin (2008)* recognized human activities through the use of triaxial accelerometers. *Sun et al. (2010)* and *Kwapisz, Weiss & Moore (2011)* recognized physical activities with the help of a smartphone with built-in sensors. These studies motivate us to collect acceleration data as a good representation of human activities, which may be beneficial to identify human emotions in real-time.

In this paper, we propose to recognize human emotions (neutral, happy, and angry) using wearable smart devices. We aim to build classification models for differentiating different emotions, using human gait data collected from a smart device with built-in accelerometer.

METHODS

The procedure of our work consists of four steps: (1) Data collection, (2) Data preprocessing, (3) Feature extraction and selection, and (4) Model training. Methods and procedures of this study have been approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences, H15010.

Data collection

In this study, a total of 123 healthy postgraduate students (45 women and 78 men) were recruited from the University of Chinese Academy of Sciences (UCAS). All participants

were required to wear a customized smart bracelet with built-in accelerometer that can track their movements. Specifically, the bracelet records movement behavior from 3-dimentional acceleration data (X axis: longitude; Y axis: latitude; Z axis: elevation).

Before the experiment, we tested the sensitivity of smart bracelet in different walking style. The absolute value of smooth walking data (X axis) (see Fig. 1 red lines) were less than 6 (m/s^2), and many absolute value of race walking data (X axis) (see Fig. 2 red lines) were greater than 6 (m/s^2). Results indicated that the bracelet is effective for differentiating smooth walking from race.

After validation, we run the experiment as follows. To avoid any environmental noise, we selected a quiet room to conduct the experiment. In the room, the floor was covered by a rectangular red carpet (length: 5 meters; width: 1 meter). All 123 participants were required to wear the smart bracelet on their right wrist and ankle, then walked back and forth on the carpet. In this study, the bracelet tracked a participant's movements continuously (five times per second), which generated time series data for further analysis. Specifically, the experiment consists of three parts.

Firstly, participants were instructed to walk for two minutes, which served as walk in neutral emotion.

Secondly, all participants were required to rate their angry emotion by a 10-point Likert Scale (1 = Not angry to 10 = Extremely angry). They then watched a 2-min film clip to elicit angry (*Yan et al., 2014*). This clip tells a short story about a girl who was run over by a vehicle. As she lay bleeding on the road, at least two persons skirted around her boy and just ignored. Eventually, the girl died. After watching the clip, participants were required to walk for one minute, i.e., walk in angry as expected. After walking, they were asked to rate their emotion (angry) for the second time.

Thirdly, to minimize the effects of angry emotion, we conducted third part more than four hours after the second part. Participants were instructed to walk for two minutes, as neutral walking behaviors. They were required to rate their happy emotion by a 10-point Likert Scale (1 = A little happy to 10 = Extremely happy), and then watched a 1.5-min film clip to elicit the emotion of happy (*Yan et al., 2014*). This film clip is a funny cartoon. After watching the clip, participants were required to walk for another one minute, i.e., walk in happy, and they rated their happy emotion after then.

The mean of self-reported angry and happy emotions was shown in Fig. 3. Results indicated that the angry and happy videos had evoked the desired emotional response.

To avoid cold-start effect of walk in neutral, we only examined a participant's movements within the last minute. Therefore, for each one emotion condition, the time series has the same length (one minute).

Data preprocessing

The data preprocessing is done as follows.

Firstly, as unexpected walking vibrations might cause noise in data collection, we used Moving Average Filter to address this issue.

Filters are functions that convert one time series into another. By choosing appropriate filter, certain patterns in the original time series can be eliminated in the new one.
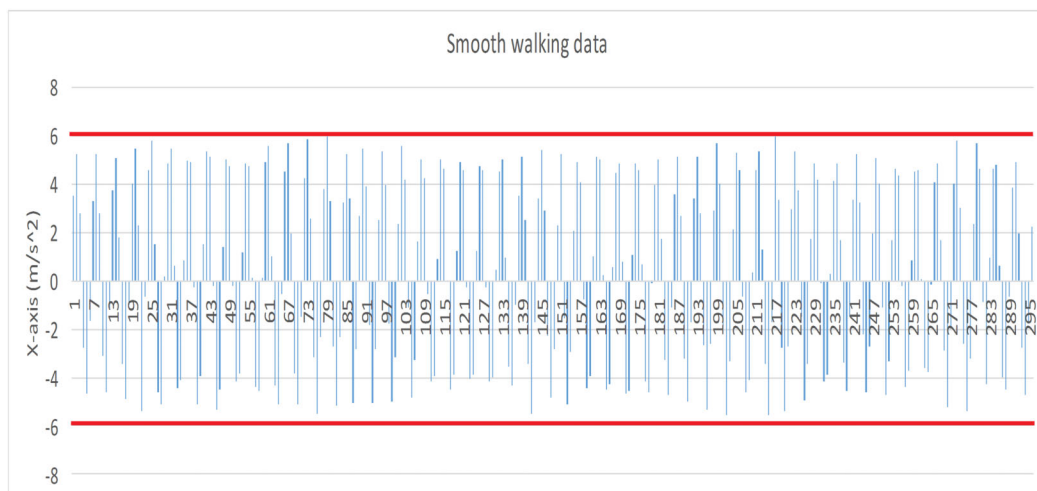
**Figure 1 Data record produced by smooth walking (X axis).**
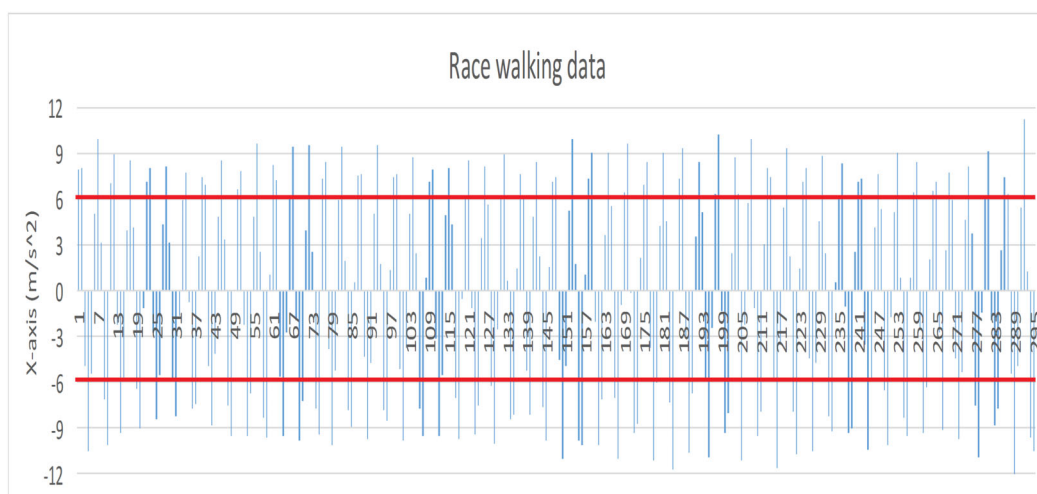
Figure 2 Data record produced by race walking (X axis).

Moving Average Filter is a kind of low-pass filter, which removes high frequency components and yields an estimate of the slow-moving trend. It operates by averaging a number of points from the input signal to generate a new point in the output signal. The convert process is defined as:

$$Output\ [i] = \frac{1}{w} \sum_{j=0}^{w-1} Input\ [i+j] \tag{1}$$

Input refers to the input signal (original time series data recorded by the bracelet), *Output* refers to the output signal (new time series data), and w refers to the number of points used to generate a new point. In this study, we set $w \in \{3, 5\}$ (*Cui, Li & Zhu, 2015*; *Ma, 2013*).
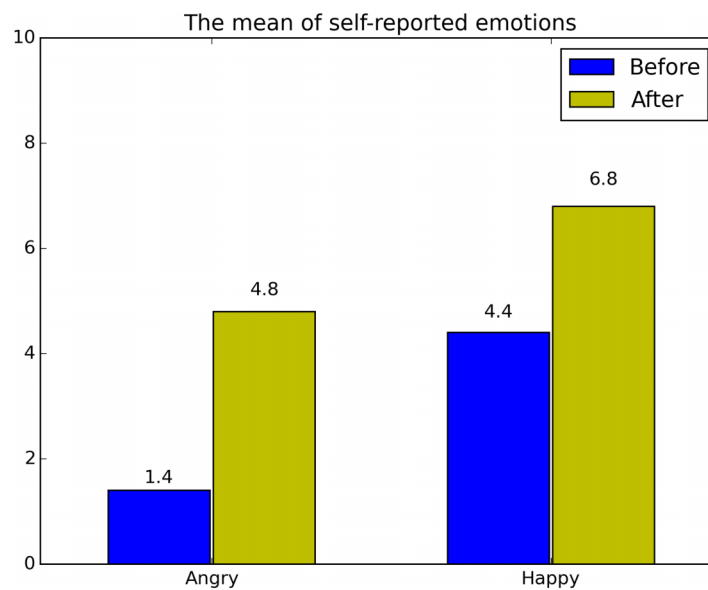
Figure 3 The mean of self-reported emotions (angry and happy).

Secondly, we reduced data redundancy. Since walking can be considered as a set of repetitive behaviors, a long-term data collection may lead to data redundancy, which can cause inefficiency in computation.

As the smart bracelet can track a participant's movements continuously, we are able to get a time series composed of 300 ($5 \times 60 = 300$) sets of 3-dimentional acceleration data (x, y, z) within one minute. To cope with such data effectively, we used sliding window (*windowsize* = 128) to divide the original time series into segments with equal length. Each segment can be regarded as a new time series (a data sample). In order to ensure an overlap ratio of 50% (*Cui, Li & Zhu, 2015*; *Xue & Jin, 2011*), the sliding step was set as 64. For example, for each user, the original time series $\{(x_1, y_1, z_1), \ldots, (x_{300}, y_{300}, z_{300})\}$ can be divided into three segments, such as $\{(x_1, y_1, z_1), \ldots, (x_{128}, y_{128}, z_{128})\}$, $\{(x_{65}, y_{65}, z_{65}), \ldots, (x_{192}, y_{192}, z_{192})\}$, and $\{(x_{129}, y_{129}, z_{129}), \ldots, (x_{256}, y_{256}, z_{256})\}$. Each segment consists of 128 sets of 3-dimensional acceleration data. Because of the limited size (less than 128), the rest of data $\{(x_{193}, y_{193}, z_{193}), \ldots, (x_{300}, y_{300}, z_{300})\}$ would be removed.

After preprocessing, we can reduce data redundancy and increase our sample size by three times, which is beneficial to build classification models.

Feature extraction and selection

In this paper, three kinds of features are extracted: temporal domain, frequency domain and temporal-frequency domain features.

Temporal domain features

For any participant's data on each one of three axes, we calculated skewness (S^t), kurtosis (K^t), and standard deviation (σ^t), in which, t denotes the axis label ($t \in \{x, y, z\}$). In addition, we also estimated correlations between every two axes.

Specifically, skewness refers to the asymmetry of the probability distribution of data, which is defined as:

$$S^t = \frac{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_i^t - \bar{x}^t\right)^3}{\left(\frac{1}{n}\sum_{i=0}^{n-1}\left(x_i^t - \bar{x}^t\right)^2\right)^{\frac{3}{2}}} \tag{2}$$

where x_i^t refers to acceleration data on the t axis, n refers to the width of a segment, and $\bar{x}^t$ refers to a mean value of all data on t axis.

Kurtosis measures the flatness of the probability distribution, which is defined as:

$$K^t = \frac{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_i^t - \bar{x}^t\right)^4}{\left(\frac{1}{n}\sum_{i=0}^{n-1}\left(x_i^t - \bar{x}^t\right)^2\right)^2} \tag{3}$$

Standard deviation measures how spread out a distribution is, which is defined as:

$$\sigma^t = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}\left(x_i^t - \bar{x}^t\right)^2} \tag{4}$$

Correlation defines the degree to which two variables vary together, or a measure of the intensity of the association between two variables, which is defined as:

$$P_{t_1,t_2} = \frac{Cov(t_1, t_2)}{\sigma^{t_1}\sigma^{t_2}}$$
$$Cov(t_1, t_2) = E[t_1 - E(t_1)][t_2 - E(t_2)] \tag{5}$$

where P_{t_1,t_2} refers to the correlation coefficient between every two axes (t_1 axis and t_2 axis), $Cov(t_1, t_2)$ refers to the covariance between every two axes, and $E(t_1)$ refers to a mean value of all data on the t_1 axis.

Frequency domain features

In addition to temporal domain features, we also converted the acceleration data from its original domain (temporal domain) to the frequency domain. For any participant's data on each axis, we calculated the mean value and the standard deviation of Power Spectral Density (PSD), which are frequency domain features. Specifically, PSD measures one signal's power intensity in the frequency domain. The mean value of PSD represents the average power per unit of bandwidth, and the standard deviation of PSD represents the degree of dispersion in terms of power.

Temporal-frequency domain features

After extracting temporal and frequency domain features, we combine these two kinds of features as a third kind of features, using Fast Fourier Transform (FFT) which converts the sampled function from its original domain (time domain) to the frequency domain.

The conversion is defined as:

$$X_k^t = \sum_{j=0}^{n-1} x_j^t e^{-i2\pi k \frac{j}{n}} \qquad k = 0, \cdots, 31 \tag{6}$$

For any user's data on each one axis, we run FFT analysis. In this study, we selected the first 32 amplitude coefficients as temporal-frequency domain features.

Feature selection

For each time series (i.e., each segment), we extract 38 features on each axis, to acquire a total of 114 features ($38 \times 3 = 114$) (See Table 1). To improve the performance of model training, we run Principle Component Analysis (PCA) for feature selection, and we selected the features with accumulative contribution rate over 95%.

We calculated the Load Matrix, acquiring the factor loading between the new features and the original features. We chose the factor loading which is equal or greater than 0.71 to determine the original features mainly contributing to the new features. We found the temporal domain features (standard deviation) and temporal-frequency domain features (coefficient of FFT) are highest contribution to the new features.

Model training

For differentiating three emotions (neutral, happy, and angry), we trained classification models using WEKA, and four different algorithms (decision tree, support vector machine, random forest, and random tree) were used to build classification models, respectively.

Decision Tree is a method commonly used in classification learning. In this study, parameters of the Decision Tree model were defined as: J48 -C 0.25 -M 2. The parameter C serves to set confidence threshold for pruning, and the parameter M serves to set minimum number of instances per leaf.

Support Vector Machine (SVM) is an algorithm, which constructs hyperplanes in a high-dimensional space for classification. In this study, parameters of the SVM model were defined as: LibSVM -S 0 -K 2. The value of parameter S ($S = 0$) indicates that $C - SVC$ type of SVM is used, and the value of parameter K ($K = 2$) indicates that the Radial Basis Function (RBF) is selected as the kernel of SVM.

Random Forest is a meta-estimator that fits a number of decision tree classifiers on various sub-samples of the dataset. In this study, the parameter of the Random Forest model was defined as: RandomForest -I 10. The value of parameter I indicates that how many decision trees are constructed.

Random Tree is a method for constructing a tree or tree-map. In this approach, a tree or tree-map is formed by a stochastic process. In this study, the parameter of the Random Tree model was defined as: RandomTree -M 1.0. The parameter M serves to set minimum number of instances per leaf.

In this study, we applied 10-fold cross validation on training models. The performance of classification models were evaluated by examining the emotion recognition rate (Q), which is defined as:

Table 1 All features extracted.

Features		X	Y	Z
Temporal	Skewness	1	1	1
	Kurtosis	1	1	1
	Standard deviation	1	1	1
	Correlation coefficient	1	1	1
Frequency	Mean of PSD	1	1	1
	Standard deviation of PSD	1	1	1
Temporal-frequency	FFT	32	32	32
Total		38	38	38

$$Q = \frac{The\ number\ of\ samples\ which\ are\ classified\ correctly}{The\ total\ number\ of\ samples} \qquad (7)$$

RESULTS

Two-category classification

For two-category classification, we examined the performance of models in differentiating every two emotions. We eliminated the noise from dataset using a Moving Average Filter, and selected features using PCA. During the noise removal process, w is set as 3 or 5, which allows us to examine the classification performance according to w.

Differentiation between neutral and angry

For the differentiation between neutral and angry, the number of features after PCA was shown in Table 2. The classification performance is shown in Table 3.

Results showed that, for differentiating neutral from angry, the overall classification accuracy was over 61.5%. The performance of models using data collected from wrist were better than those using data collected from ankle. When we examined data collected from wrist-worn accelerometers with $w = 5$, we got the best classification accuracy ($Q = 91.3\%$) using the LibSVM algorithm.

Differentiation between neutral and happy

For the differentiation between neutral and happy, the number of features after PCA was shown in Table 4. The classification performance can be shown in Table 5.

For differentiating neutral from happy, the overall classification accuracy was over 61.0%. When we examined data collected from wrist-worn accelerometers with $w = 3$, the best classification accuracy is ($Q = 88.5\%$) using LibSVM.

Differentiation between happy and angry

For the differentiation between happy and angry, the number of features after PCA was shown in Table 6. The classification performance can be shown in Table 7.

Table 2 Number of features after PCA for differentiating neutral and angry.

	$w = 3$	$w = 5$
Wrist	56	49
Ankle	59	53

Table 3 Classification performance for differentiating neutral and angry.

		LibSVM	DecisionTree	RandomForest	RandomTree
$w = 3$	Wrist	86.0%	76.2%	71.1%	65.7%
	Ankle	72.5%	64.2%	63.8%	64.3%
$w = 5$	Wrist	91.3%	83.8%	82.8%	69.8%
	Ankle	71.3%	61.5%	62.3%	61.9%

Table 4 Number of features after PCA for differentiating neutral and happy.

	$w = 3$	$w = 5$
Wrist	61	55
Ankle	61	54

Table 5 Classification performance for differentiating neutral and happy.

		LibSVM	DecisionTree	RandomForest	RandomTree
$w = 3$	Arist	88.5%	77.8%	64.9%	61.0%
	Ankle	80.9%	73.1%	65.7%	63.0%
$w = 5$	Wrist	78.2%	67.7%	63.1%	62.3%
	Ankle	71.7%	62.4%	61.5%	62.6%

Table 6 Number of feature after PCA for differentiating neutral and happy.

	$w = 3$	$w = 5$
Wrist	56	49
Ankle	59	53

Table 7 The classification accuracy for differentiating happy and angry.

		LibSVM	DecisionTree	RandomForest	RandomTree
$w = 3$	wrist	88.5%	83.3%	73.8%	68.1%
	Ankle	79.1%	70.1%	65.3%	60.6%
$w = 5$	Wrist	82.5%	72.9%	66.3%	63.1%
	Ankle	71.1%	60.4%	62.0%	60.9%

Zhang et al. (2016), *PeerJ*, DOI 10.7717/peerj.2258

9/14

Table 8 Number of features after PCA for differentiating neutral, happy, and angry.

	$w = 3$	$w = 5$
Wrist	56	49
Ankle	59	53

Table 9 Classification performance for differentiating neutral, happy, and angry.

		LibSVM	DecisionTree	RandomForest	RandomTree
$w = 3$	Wrist	79.6%	65.8%	59.0%	52.4%
	Ankle	68.6%	60.1%	53.2%	49.3%
$w = 5$	Wrist	81.2%	70.6%	66.2%	56.6%
	Ankle	62.3%	49.6%	52.4%	47.8%

For differentiating happy from angry, the overall classification accuracy was over 60.4%. When we examined data collected from wrist-worn accelerometers with $w = 3$, the best classification accuracy is ($Q = 88.5\%$) on LibSVM.

Three-category classification

Besides conducting two-category classification, we also examined the classification performance for differentiating three emotions (neutral, happy, and angry). For the three-category classification, the number of feature after PCA was shown in Table 8. The classification performance can be shown in Table 9.

Results showed that, the overall performance of three-category classification is worse than two-category classification. When we examined data collected from wrist-worn accelerometers with $w = 5$, the best classification accuracy is ($Q = 81.2\%$) using LibSVM.

DISCUSSION

In this paper, we used wearable devices (smart bracelets) to conduct a gait analysis on 123 participants to recognize their emotions. This study demonstrates that human emotions (neutral, happy, and angry) are expressed in walking to some degree, and more importantly, a real-time recognition of human emotions (neutral, happy, and angry) can be realized by using wearable smart devices.

We found that human gait can be used to differentiate different emotions, which is consistent with previous research (*Montepare, Goldstein & Clausen, 1987*; *Michalak et al., 2009*). Results indicated that, for both two-category and three-category classification, the highest classification accuracy can reach over 88%. As humans can recognize emotions from signals (e.g. face and voice) with an accuracy of 70% − 98% (*Takahashi, 2005*; *Bos, 2006*), it suggests that classification models on gait work fairly well. In addition, SVM classifiers outperform other classifiers (e.g. Decision Tree, Random Forest, and Random Tree).

The models perform differently across different emotions. For neutral vs. non-neutral classification, it was easier to identify neutral from angry than from happy, which is consistent with other research on emotion recognition (*Montepare, Goldstein & Clausen, 1987*). We also found that, when the number of categories (i.e. number of emotion types) increases, the

classification accuracy decreases. Probably it is partly due to an increasing complexity of the classification task. More specifically, it might be challenging to recognize various types of emotions within only one minute. If we can extend the period of observation, the classification accuracy would be increased.

The placement of wearable smart devices at different body positions may influence the performance as well. Results indicated that, in general, data collected from wrist might be much more useful to recognize human emotions.

Our method is based on cut-edge technology (*Montepare, Goldstein & Clausen, 1987*; *Michalak et al., 2009*; *Sun et al., 2010*). The customized smart bracelet is convenient to wear for a long time, and the procedure of acceleration data collection is non-intrusive and ecological. If we embed bluetooth in the bracelet to upload real-time walking behavioral data, we are able to monitor emotion timely, which could be used on different circumstances, such as monitoring emotion of mental disorders, children's psychological research, etc.

It is important to note the limitations of this study. In this study, we conducted the gait analysis on a total of 123 participants, and the sample size is a bit limited. Collecting data from a larger number of participants might further validate the modeling performance. As all participants are Chinese, we do not know whether this method still works in other countries. In addition, we are not sure whether there exist cultural differences in the relationship between emotion and walking patterns. Because of the limited length of observation (1 min), we cannot compare the classification accuracy among observation periods with different lengths. Therefore, we cannot figure out the optimal observation time window for recognizing human emotions. As we took each segment as one sample, and the samples were classified independently from each other. We just looked into the classification result of each sample. In the future work, for each participant, we may use the voting method based on classification result of the segments to predict his/her emotion.

This study provides an innovative method for recognizing human emotions (neutral, happy, and angry) in an efficient manner. Through the use of wearable smart devices, we can recognize human emotions (neutral, happy, and angry) in real-time, which could be beneficial to other research including human-computer interaction.

CONCLUSION

This paper built classification models for differentiating various emotions (happy, neutral, and angry), using human gait data collected from wearable smart devices. Results indicate that it is efficient to recognize human emotions (happy, neutral and angry) by using a wearable smart bracelet. This method can be helpful to realize an automatic recognition of human emotions (neutral, happy, and angry) in human-computer interaction.

ADDITIONAL INFORMATION AND DECLARATIONS

REFERENCES

Bao L, Intille SS. 2004. Activity recognition from user-annotated acceleration data. In: *Pervasive Computing*. Berlin Heidelberg: Springer, 1–17.

Beale R, Peter C. 2008. The role of affect and emotion in HCI. In: *Affect and Emotion in Human-Computer Interaction*. Berlin Heidelberg: Springer, 1–11.

Bos DO. 2006. EEG-based emotion recognition, the influence of visual and auditory stimuli. Internal report. Enschede: Department of Computer Science, University of Twente, 1–17.

Zhang et al. (2016), *PeerJ*, DOI 10.7717/peerj.2258

12/14

Brave S, Nass C. 2003. Emotion in human–computer interaction. *The Human-Computer Interaction Handbook*. Hillsdale: L. Erlbaum Associates Inc., 81–96.

Chen J, Hu B, Moore P, Zhang X, Ma X. 2015. Electroencephalogram-based emotion assessment system using ontology and data mining techniques. *Applied Soft Computing* **30**:663–674 DOI 10.1016/j.asoc.2015.01.007.

Crane E, Gross M. 2007. Motion capture and emotion: affect detection in whole body movement. In: *Affective Computing and Intelligent Interaction*. Berlin Heidelberg: Springer, 95–101.

Cui L, Li S, Zhu T. 2015. Emotion detection from natural walking. *PeerJ PrePrints* **3**:e1794 DOI 10.7287/peerj.preprints.1384v3.

Han J, Bhanu B. 2006. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28(2)**:316–322 DOI 10.1109/TPAMI.2006.38.

He Z-Y, Jin L-W. 2008. Activity recognition from acceleration data using AR model representation and SVM. In: *Proceedings of the 7th International Conference on Machine Learning and Cybernetics*. Vol. 4. Piscataway: IEEE, 2245–2250.

James W. 1884. II—What is an emotion? *Mind* **os-IX(34):**188–205 DOI 10.1093/mind/os-IX.34.188.

Kale A, Sundaresan A, Rajagopalan A, Cuntoor NP, Roy-Chowdhury AK, Krüger V, Chellappa R. 2004. Identification of humans using gait. *IEEE Transactions on Image Processing* **13(9):**1163–1173 DOI 10.1109/TIP.2004.832865.

Karg M, Kühnlenz K, Buss M. 2010. Recognition of affect based on gait patterns. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **40(4):**1050–1061 DOI 10.1109/TSMCB.2010.2044040.

Kwapisz JR, Weiss GM, Moore SA. 2011. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter* **12(2):**74–82 DOI 10.1145/1964897.

Lee S-W, Mase K. 2002. Activity and location recognition using wearable sensors. *IEEE Pervasive Computing* **1(3):**24–32 DOI 10.1109/MPRV.2002.1037719.

Ma L. 2013. Sensor-based activities recognition on mobile platform. Master's Thesis, Harbin Institute of Technology.

Michalak J, Troje NF, Fischer J, Vollmar P, Heidenreich T, Schulte D. 2009. Embodiment of sadness and depression-gait patterns associated with dysphoric mood. *Psychosomatic Medicine* **71(5):**580–587 DOI 10.1097/PSY.0b013e3181a2515c.

Montepare J, Koff E, Zaitchik D, Albert M. 1999. The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior* **23(2):**133–152 DOI 10.1023/A:1021435526134.

Montepare JM, Goldstein SB, Clausen A. 1987. The identification of emotions from gait information. *Journal of Nonverbal Behavior* **11(1):**33–42 DOI 10.1007/BF00999605.

Peter C, Beale R. 2008. *Affect and Emotion in Human-Computer Interaction: From Theory to Applications*. Vol. 4868. Berlin Heidelberg: Springer Science & Business Media.

Peter C, Crane E, Beale R. 2006. The role of emotion in human-computer interaction. In: *Proceedings of the 21st British HCI Group annual conference on people and computers: HCI...but not as we know it - Volume 2*. Swinton: British Computer Society, 211–212.

Smith CA, Lazarus RS. 1990. Emotion and adaptation. In: Pervin LA, ed. *Handbook of Personality: Theory and Research*. New York: Guilford Press, 609–637.

Sun L, Zhang D, Li B, Guo B, Li S. 2010. Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In: *Ubiquitous Intelligence and Computing*. Berlin Heidelberg: Springer, 548–562.

Takahashi K. 2005. Remarks on emotion recognition from bio-potential signals. In: *Second international conference on autonomous Robots and Agents*186191.

Xue Y, Jin L. 2011. Discrimination between upstairs and downstairs based on accelerometer. *IEICE Transactions on Information and Systems* **94(6):**1173–1177 DOI 10.1587/transinf.E94.D.1173.

Yan W-J, Li X, Wang S-J. 2014. Zhao G, Liu Y-J, Chen Y-H, Fu X. 2014. CASME II: an improved spontaneous micro-expression database and the baseline evaluation. *PLoS ONE* **9(1):**e86041 DOI 10.1371/journal.pone.0086041.

Zhang X, Hu B, Chen J, Moore P. 2013. Ontology-based context modeling for emotion recognition in an intelligent web. *World Wide Web* **16(4):**497–513 DOI 10.1007/s11280-012-0181-5.