Identifying Emotion from Natural Walking

Liqing Cui†‡, Shun Li†‡, Wan Zhang†‡, Zhan Zhang†, Tingshao Zhu†∗
† Institute of Psychology, Chinese Academy of Sciences
‡ The 6th Research Institute of China Electronics Corporation (National Computer System Engineering Research Institute of China)

Abstract
To provide a means of identification of human emotion in walking, this paper analyzes the capability of walking activity to reveal a person’s affective states. We obtain pure wrist and ankle accelerometer data, because of redundant information existing in high dimension data, then we set different w (moving average filter window size) and utilize principal component analysis (PCA) to reduce dimension, respectively compare classification accuracy for wrist and ankle with respect to different w value with specific models. In fact, emotion identification from ankle has better performance than wrist. Best accuracy of anger is achieved for Decision Tree with 74% when w is 5, best accuracy of happiness is 85%, and the identification ratio of anger-happy is 85.34%. It is concluded that gait of walking is capable to reveal and identify the emotional state of person when walking.

Keywords: Sensors mining, emotion identification, cellphone, accelerometer sensor

1. Introduction
Nonverbal communication plays a major role in the future. To simplify human-machine interaction and to detect human’s health state, nonverbal signals of humans are observed delivering additional cues for a person’s physiological and psychological state and intentions. Within this research area, affective
computing faces the challenges of automatically identifying a person’s affect state. Generally, detection and identification of emotion is based on facial expressions, linguistic and acoustic features in speech. Psychological studies on visual analysis of body movement show that human movement differs from other movements because it is the only visual stimulus we have experience of both perceiving and producing\cite{2}.\cite{15}, here the modality we provide to identify human’s emotion from walking. In walking, we only record person’s wrist and ankle accelerometer data by cellphone devices’ sensors.

Nowadays, mobile devices have already become indispensable communication tools in daily life, and more and more latest generation of smart cellphones incorporates many diverse and powerful sensors. Common sensors are GPS sensors, light sensors, direction sensors, temperature sensors, proximity sensors, pressure sensors, acceleration sensors, gravity sensors. Some of them are not only complementarily help users manage their devices intelligently, but offer new opportunities for data mining and mining applications with advantages of small size, substantial computing power and high precision.

In this paper we describe and implement a system that uses phone-based acceleration sensors to perform human emotion identification. To collect accelerometer data, we conduct an experiment that participants tying respectively cellphone to wrist and ankle perform daily walking in a fixed area. Due to that we record the accelerometer and gravity data of cellphone of wrist and ankle, this mean we can obtain pure accelerometer data of wrist and ankle joint. because of redundant information in high dimension including time-domain, frequency-domain, power, distribution features, we utilize Principal Component Analysis (PCA) to reduce redundant information. On the one hand, we compare classification performance of Decision Tree, Random Forest, Simple Logistic, Naive Bayes, Multilayerperception and Random Tree, on the other hand, we compare the effect on performance of different data preprocessing.

This work is significant because the learning model permits us to obtain useful knowledge about human affection in a certain gait of millions of people to some extend -just by one person’s wrist and ankle data. Our work has a wide range of
applications, including generating a daily or weekly or monthly emotion profile to detect person’s affection state change during the latest period. If it works, we can replace cellphone with smart bracelet, these data sent and received by wireless devices can be used as medical information for doctors. In addition, they can also used for personal health by offering a benignant feedback that suggests having some exercise or entertainment.

To summarize, our research has two main contributions.

- We obtain pure accelerometer data from wrist and ankle in daily walking, to reveal the association between one’s walking activity and her or his current emotion.
- Data preprocess, especially in eliminating burr and noise, play a significant role apart from algorithm in improving performance of human emotion identification.

The outline of this paper is organized as follows. Section 2 summarizes related work about identification of emotion in walking. Description of data base and data preprocessing and feature extraction is presented in Section 3. Section 4 describes the results of the trained models and the performance of trained models for experiment, in section 5 we discuss our methods and summary our work. Finally, the paper ends with a conclusion in Section 6.

2. Related Work

There exist various models to categorize emotions in psychology. Ekman’s basic emotions, which are anger, disgust, fear, sadness and surprise, and the dimensional pleasure-arousal-dominance (PAD) model are widely used in automatic emotion recognition [8][13]. The PAD model spans a 3-dimensional space with the independent and bipolar axes pleasure, arousal and dominance. An affective state is described as a point within this state space.

Early studies based on black displays showing only the joints of the body, then observers recognize its gender or judge it is a familiar person or not in
walking\[7\]. In Montepare’s psychological study, He expounded that observers can also identify emotions from variations in walking gaits\[11\]. Relatively, the emotions discrimination of sadness and anger are easier than pride for observers. Pollick quantified expressive arm movements in terms of velocity and acceleration suggesting which aspects of movement are important in recognizing emotions\[12\]. Crane and Gross illustrated that body’s reaction of felt and recognized emotions not only depended on gesticulatory behavior, but associated with emotion-specific changes in gait kinematics. In their study, they extracted some activity features, including velocity, cadence, head orientation, shoulder and elbow range of motion, as significant parameters to identify emotions\[6\].

In fact, emotion changes rapidly even in a short walking, and body movement is also complex, these factors effect the correct identification accuracy. Janssen investigated the recognition for four emotional states by means of artificial neural nets. The recognition results reaches 83.7% in average based on kinematic data for person-dependent recognition. But inter-individual recognition remains around chance level. There is a 79% correct classification of gait patterns including calming, excitatory or no music. Karg applied different methods such as Principal Component Analysis(PCA), Kernal PCA(KPCA) and Linear Discriminant Analysis(LDA) into kinematic parameters of person-dependent recognition and inter-individual recognition to compare results and improve accuracy rate. LDA in combination with Naive Bayes leads to an accuracy of 91% for person-dependent recognition of four discrete affective states based on observation of barely a single stride\[1\]. In \[9\], a combination of two consecutive PCA and Fourier Transformation is used for data reduction. Best accuracy is achieved for Naive Bayes with 72% for the four emotions sad, neutral, happy and angry during natural walking.

A general overview of analytical techniques for clinical and biomechanical gaits analysis is given in\[4\]\[5\]. It mainly refers to classification of clinical disorders, though the methods for feature extraction can also be taken for psychological gaits analysis. Dimension reduction techniques such as KPCA improves recognition of age in walking\[14\]. The performance comparison of Principal
Component Analysis (PCA) and KPCA is showed in [3]. Martinez and Kak showed that PCA can outperform when size of training sets are small to feature dimension [10].

Focus of our work is to extract relevant time domain, frequency domain, power, distribution features from kinematic data set to identify human emotion. Here, we assume the arousal from different video themes is easier to identify and thus expressed in walking. We utilize pure accelerometer data from wrist and ankle respectively to identify human emotion, and compare emotion identification accuracy of different moving filter window size under emotions of happiness or anger.

3. Methods

The proposed emotion identification method based on three-axis acceleration sensor and gravity sensor embedded in mobile phone comprises the following three steps: 1) collecting and pre-processing the sensor data from subjects, 2) extracting features, and 3) training classifiers. And in the last step, we train kinds of classifier models to compare and analyze performance.

3.1. Subjects and Data Base

To investigate identification of emotion in gaits patterns, a data base has been collected at xueyuan2 classroom 350, University of Chinese Academy of Sciences. We randomly recruited fifty-nine graduates to participate this experiment, including 31 male and 28 female. Our project employed two SAMSUNG GALAXY S2 and one SAMSUNG Tab as platform, and they were Android operation system, because the Android operating system is free, open-source, easy for us to program to access accelerometer and gravity sensor value in cellphone and develop an APP on SAMSUNG Tab to record the start and end of activity time. The cellphone had a 5 Hz sample frequency, i.e., sensor recorded one piece of data per 200ms.

Participant tied respectively one cellphones to wrist and another to ankle, then
was asked current emotion, rating own anger score, and standing at specified
starting line, then performed daily walking in a fixed rectangle-shaped area for
two minutes, then continued walking for one minute after watching infuriating
video. At three hours later (if interval is too small, it influences participant’s
next emotion arousal after video), the one participated the second round of
the experiment, was asked current emotion, rating own happiness score, then
tied one cellphones to wrist and another to ankle as same as the first round,
walked for two minutes in above area, then continued to walk for one minute
after watching prepared funny video. The start and end of each walking time
was marked in SAMSUNG Tab APP by hand by our host. Then we used the
time series data in Tab to cut and aggregate database file data in cellphone into
samples.

3.2. Data Preprocessing

We have acquired two groups of sensor data, one is for wrist, and another
is for ankle. Each group also includes accelerometer data sets SensorLa and
gravity data sets SensorGra. According to Tab’s time record, we cut every
participant’s walking data before and after video. Then after SensorLa sets
subtracts SensorGra sets at same time, what we get is pure accelerometer data.
Because of noise and burrs existing in data, we need to do some preprocess
to pure accelerometer data, moving average filter is suitable for time domain
signal. Moving average filter expression shows as below:

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

w is adjustable for size of average process once, we set w 3 and 5 respectively
in our data procession. As shown in Figure 1, a raw data of X-axis from ankle.

The \( w \) value has significantly influence on smooth performance. The below
figures show the ankle wave signal with respect to \( w \).

When \( w = 3 \), the undulating signal has become smoother than raw data
shown in Figure 1. If \( w = 5 \), the signal become more smoother than that when
\( w \) is 3, as shown in Figure 2.

But if \( w \) value is too great, it will eliminate existing minor changes in wave
Figure 1: one raw accelerometer data of X-axis from ankle

Figure 2: one raw accelerometer data of X-axis from ankle is processed by moving average filter when w is 3

Figure 3: one raw accelerometer data of X-axis from ankle is processed by moving average filter when w is 5
signal. Though it makes wave smoother, perhaps we lose key undulatory information of signal. Here we respectively set w 3 and 5 to analyze results.

Due to the sampling frequency is 5Hz, i.e., the APP can access accelerometer data five times per second and write it into database one by one. A few minutes can accumulate hundreds of pieces of records, these records are too much to deal directly and extract features hastily. Generally, slice sliding window is common means used to cut data into sheets or slices, slice sliding window size we choose is 128, and the coverage ratio is 50%, like Figure. 4 shown.

3.3. Feature Extraction

Participant has a great difference in walking with own complex behavioral traits, different speed, and every participant’s records is not same, here we extract time domain, frequency domain, power and distribution features from each activity data sheet, including each axis’s standard deviation, means, kurtosis, skewness, the front 32 amplitude coefficients of FFT (the first value is current component, i.e., the axis’s means), the means and standard deviation of power spectral density (PSD), and the correlation coefficient of every two axes, total 38 dimensions one axis. So one activity data sheet produces a 1 * (38 * 3) feature matrix. Similarly, all participants’ activity sheets produce feature matrix, then we aggregate all these feature matrices into one matrix.
Different features show activity’s properties, means represents signal direct-current component in activity data sheet, belonging to low frequency property. Standard deviation shows the stability of acceleration signal, and represents degree of concentration of the motion data. Because of window width is 128, the front 32 amplitude coefficients of FFT denote activity property in low-frequency domain. And PSD feature explains acceleration signal from the point of energy.

All of the above features almost have big numerical difference, even some important features with small values are ignored in classification, and seriously affect the results. To avoid it, we need do Z-score normalization to features sets matrix.

In general, high dimension of feature vector increases computational complexity, and there exists much redundant information. In order to reduce the dimension of feature vectors, and to obtain the best description of the different behavior and the best classification characteristics, dimension reduction is an essential step.

4. Results

For train sets we get from wrist and ankle with different \( w (w=3 \text{ and } w=5) \) in the two rounds of experiment, the first round, we labeled each sample with “natural” or “anger” after PCA, then we trained models in Weka. Similarly, for train sets got in the second round, we labeled each with “natural” or “happy”, then trained models.

4.1. Anger Emotion Identification

In first round of experiment, we obtained pure accelerometer data from wrist and ankle. After a series of procession, we utilized diverse kinds of classification algorithms to train models in Weka with default parameters and standard 10-fold cross validation.

When \( w \) is 3, the results we obtained from wrist and ankle is shown in Table. [1]

From the above results, the video participant watched did have arouse her or his homologous emotion corresponding to own emotion score change and have a
Table 1: The classification accuracy in different models when $w=3$

| Joint | DT    | RF   | SL    | NB    | MLP   | RT    |
|-------|-------|------|-------|-------|-------|-------|
| Wrist | 56.25%| 62.21%| 53.60%| 51.42%| 63.92%| 58.81%|
| Ankle | 71.31%| 64.49%| 54.26%| 57.39%| 59.38%| 59.94%|

DT: Decision Tree.
RF: Random Forest.
SL: Simple Logistic.
NB: Naive Bayes.
MLP: Multilayerperception.
RT: Random Tree.

Significant influence on her or his gait. In addition, both wrist and ankle have a relatively higher accuracy in Decision Tree than other models. Meanwhile, the identification accuracy from ankle is higher than wrist. A major reason is that the activity of hand is more complex than ankle when walking, causing much noise which is not easily separated from records.

In fact, when we set $w=5$, the results we obtained is of greatly significant difference. As shown in below Table 2.

Table 2: The classification accuracy in different models when $w=5$

| Joint | DT    | RF   | NB   | MLP   | RT    |
|-------|-------|------|------|-------|-------|
| Wrist | 54.99%| 59.54%| 50.71%| 58.97%| 52.99%|
| Ankle | 74.07%| 65.81%| 52.99%| –     | 62.68%|

- : invalid.

The results show that when $w=5$, Classification results of most of above models have a little higher accuracy than results when $w$ is set 3. And the accuracy of wrist is still lower as same as results when $w$ is 3.
4.2. Happiness Emotion Identification

In second round of experiment, the way we obtained pure accelerometer data from wrist and ankle is same with first round of experiment. After a series of procession, we utilized diverse kinds of classification algorithms to train models in Weka. The classification results is shown in Table 3 with \( w = 3 \).

| Joint | DT    | RF    | SL    | NB    | MLP   | RT    |
|-------|-------|-------|-------|-------|-------|-------|
| Wrist | 61.19% | 61.49% | –     | –     | 58.51% | 61.19% |
| Ankle | 74.93% | 67.46% | 57.61% | 51.94% | 61.19% | 62.39% |

From above results, we can induce that the funny video arouse participants’ emotion so that their gaits have a significant difference to be easily identified relatively between walking before video and walking after video. As same with Table 1, ankle has a better stability to identify emotion than wrist in all above models. The ankle accuracy reaches 74.93%. Similarly, \( w \) has a great influence on classification accuracy in second round experiment. As shown in Table 4.

| Joint | DT    | RF    | SL    | NB    | MLP   | RT    |
|-------|-------|-------|-------|-------|-------|-------|
| Wrist | 63.88% | 58.20% | 54.63% | –     | 51.94% | 62.69% |
| Ankle | 85.07% | 70.45% | 54.33% | 59.40% | 54.32% | 60.60% |

The table fully demonstrates that \( w \) do influence emotion identification to some extend, and Decision Tree model has best accuracy than other models we use, up to 85.07%.

4.3. Different Emotion Identification

We aggregated both data sets of the first round of experiment after video and data sets of the second round after video, and respectively tagged “anger” and “happy”. The classification accuracy is shown in Table 5 and Table 6.
Table 5: Different emotion classification accuracy in different models when w=3

| Joint | DT | RF | RT |
|-------|----|----|----|
| Wrist | 63.34% | – | – |
| Ankle | 74.49% | 63.64% | 56.60% |

Table 6: Different emotion classification accuracy in different models when w=5

| Joint | DT   | RF   | RT   |
|-------|------|------|------|
| Wrist | 63.05% | 54.25% | 54.25% |
| Ankle | 85.34% | 67.16% | 66.86% |

Form above two tables, we can induce that person’s gaits exist significant difference when his or her walking with anger emotion or happy emotion. And Decision Tree has better performance for ankle, reaching 85.34% to identify anger or happy emotion when w is 5 than decision tree model whose accuracy reaching 74.49% when w is 3.

5. Discussion

In the experiment, we utilized PCA to complete dimension reduction, then run train sets in different methods in Weka with default parameter values and standard 10-fold cross validation.

We extract activity’s 114 dimensional features from pure accelerometer data, then to identify person’s emotion in different models from participant’s gait. The above results presented in this paper are quite interesting and promising, indeed, there exists significant difference in walking within different emotion, and different w values (moving average filter window size) have an evident effect on identification accuracy, because that the w is greater, signal sequence is more smoother in time-domain, but if w is too great, it will eliminate tiny key wave changes so that this increases learning complexity. Otherwise, small w is without an evident performance of moving smooth filter. In summary, whether w is 3
or 5, ankle has a better stability to identify emotion than wrist with higher accuracy, reaching 74.31% in first round of experiment and 85.07% in second round of experiment. One of main reasons is that wrist has complex additional movement when human performs walking. Besides, two kinds of emotion(anger-happiness) is relatively easily to be identified, the accuracy reaches 85.34%.

As mentioned in [9], M. Karg and R. Jenke optimized two-fold PCA algorithm and improved its performance, classified train sets into four categories: angry, happy, neutral, and sad. Among results, The accuracy of angry reaches 69% and happy reaches 76% for Naive Bayes. In our work, performance of Naive Bayes on data sets does not match Decision Tree model. For our experiment, because we obtain person’s pure wrist and ankle accelerometer data, there existing no much noise information relatively, and we analyze all features which represents one person’s gait characteristics to certain extend, the results are more credible. However, further consideration and improvement is required. This includes what effect the size of slice sliding window has on identification, whether it can also improve model’s performance.

6. Conclusion

This experiment investigates identification of human emotion in person’s walking. For this purpose, we obtained pure wrist and ankle accelerometer data base. After a serial preprocess, we extract 114 features from each data sheet. The results of different trained models lead to the conclusion that ankle data is more capable to reveal human emotion than wrist, best accuracy reaching 85.34%, and preprocess is also key factor which can improve model performance. It is also concluded that different emotion is recognizable from human’s characteristics of walking, which reaches an accuracy of 85.07%.

Acknowledgments

The authors gratefully acknowledges the generous support from National High-tech R&D Program of China (2013AA01A606), National Basic Research
Program of China(2014CB744600), Key Research Program of Chinese Academy of Sciences(CAS)(KJZD-EWL04), and CAS Strategic Priority Research Program (XDA06030800).

References

[1] A comparison of PCA, KPCA and LDA for feature extraction to recognize affect in gait kinematics, Sept 2009.

[2] Christian Peter Russell Beale. Affect and emotion in human-computer interaction.

[3] L.J. Cao, K.S. Chua, W.K. Chong, H.P. Lee, and Q.M. Gu. A comparison of pca, {KPCA} and {ICA} for dimensionality reduction in support vector machine. Neurocomputing, 55(1C2):321 – 336, 2003. Support Vector Machines.

[4] Tom Chau. A review of analytical techniques for gait data. part 1: fuzzy, statistical and fractal methods. Gait & Posture, 13(1):49 – 66, 2001.

[5] Tom Chau. A review of analytical techniques for gait data. part 2: neural network and wavelet methods. Gait & Posture, 13(2):102 – 120, 2001.

[6] Elizabeth Crane and Melissa Gross. Motion capture and emotion: Affect detection in whole body movement. In AnaC.R. Paiva, Rui Prada, and RosalindW. Picard, editors, Affective Computing and Intelligent Interaction, volume 4738 of Lecture Notes in Computer Science, pages 95–101. Springer Berlin Heidelberg, 2007.

[7] JamesE. Cutting and LynnT. Kozlowski. Recognizing friends by their walk: Gait perception without familiarity cues. Bulletin of the Psychonomic Society, 9(5):353–356, 1977.
[8] Paul Ekman and Wallace V. Friesen. A new pan-cultural facial expression of emotion. *Motivation and Emotion*, 10(2):159–168, 1986.

[9] Michelle Karg, Robert Jenke, Kolja Kühnlenz, and Martin Buss. A two-fold pca-approach for inter-individual recognition of emotions in natural walking. In *MLDM Posters*, pages 51–61, 2009.

[10] Aleix M. Martinez and A. C. Kak. Pca versus lda. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(2):228–233, Feb 2001.

[11] Joann M. Montepare, Sabra B. Goldstein, and Annmarie Clausen. The identification of emotions from gait information. *Journal of Nonverbal Behavior*, 11(1):33–42, 1987.

[12] Frank E Pollick, Helena M Paterson, Armin Bruderlin, and Anthony J Sanford. Perceiving affect from arm movement. *Cognition*, 82(2):B51 – B61, 2001.

[13] James A Russell and Albert Mehrabian. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3):273 – 294, 1977.

[14] Jianning Wu, Jue Wang, and Li Liu. Feature extraction via {KPCA} for classification of gait patterns. *Human Movement Science*, 26(3):393 – 411, 2007.

[15] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, Jan 2009.