Available at www.ijsred.com

RESEARCH PAPER OPEN ACCESS

Face Recognition Attendance System and Emotion Recognition Using Kinect Motion Capture Data of Human Gaits

Dr. Dinesh D. Patil*, Dr. Priti Subramanium**, Prof. Abhijeet Ingale***, Miss.Vrunda Mahajan****

**** (Computer Science and Engineering Department, Ssgbcoet, Bhusawal.
Email: mahajanvrunda671@gmail.com)

** (Computer Science and Engineering Department, Ssgbcoet, and Bhusawal.
Email: pritikanna559@gmail.com)

--************************----------------------------------

Abstract:

The face is a key part of the human body that uniquely identifies a person. Using facial characteristics as biometrics, a facial recognition system can be implemented. The most challenging task in any organization is attendance marking. In the traditional attendance system, students are called by teachers and their presence or absence is marked accordingly. However, these traditional techniques are time-consuming and tedious. In this research paper, an open CV based face recognition approach was proposed. The training database is created by training the system with the faces of the authorized person. The cropped images are then saved as a database with the appropriate labels. Automatic emotion recognition has great value in many applications, but more portable, non-intrusive and low-cost technologies need to be developed to fully realize the application value of recognition. Human gaits could reflect the emotional state of the walker and could be a source of information for emotion recognition. This work proposed a new method of emotional state recognition through human walking using Microsoft Kinect, a low-cost, portable camera-based sensor. The walking of 59 participants in neutral condition, induced anger and induced happiness were recorded by two Kinect cameras, and the original data were processed by joint selection, coordinate system transformation, Gaussian sliding window filtering, differential operation and data segmentation. Features of gait patterns were extracted from the 3-dimensional coordinates of the 14 joints of the main body using Fourier transform and principal component analysis (PCA). NaiveBayes, RandomForests, LibSVM, and Sequential Minimal Optimization (SMO) classifiers were trained and evaluated, and the recognition accuracy of anger and happiness from the neutral state reached 80.5% and 75.4%, respectively. Although the results of distinguishing between anger and happiness states were not ideal in the current study, it demonstrated the feasibility of automatic recognition of emotional states from walking, while the features meet the application requirements.

Keywords — **Face Recognition, Face Detection, OpenCV, Camera, Attendance, Biometrics, Emotion Recognition, Affective computing, Walking, Machine Learning, Kinect.**

--************************----------------------------------

I. INTRODUCTION

Build a facial recognition system using Python. Face recognition is one step further than face detection. In face detection, we only detect the position of the human face in the image, but in face recognition, we create a system that can identify people. “Face recognition is a broad challenge to verify or identify people in images or videos. The big tech giants are still working on creating a faster and more accurate face recognition model.” Emotion is a mental experience with high intensity and high hedonic content (pleasure/displeasure) (Cabanac, 2002) that profoundly affects our daily behavior by regulating an individual's motivation (Lang, Bradley & Cuthbert, 1998), social interaction (Lopes et al., 2005) and cognitive processes (Forgas, 1995). Recognizing the emotions of others and responding adaptively to them is the foundation of effective social interaction (Salovey & Mayer, 1990), and because users tend to view computers as social agents (Pantic & Rothkrantz, 2003), they expect their affective state to be perceived and taken into account. Consideration when interacting with computers. Given the importance of emotional intelligence for successful human interaction, the ability of a computer to automatically recognize and respond appropriately to a user's emotional feedback has been recognized as a key aspect of natural, effective, persuasive, and trustworthy human-computer interaction (Cowie et al. ., 2001; Pantic & Rothkrantz , 2003; Hudlička, 2003). The possible applications of such an emotion-sensitive system are numerous, including automatic customer service (Fragopanagos & Taylor, 2005), interactive games (Barakova & Lourens, 2010) and smart homes (Silva, Morikawa & Petra, 2012), etc. automatic emotion recognition is very challenging task, the development of this technology would be of great value.

As a common use of multiple ways of recognizing emotional states in human interaction, various cues such as facial expressions (e.g., Kenji, 1991), gestures (e.g., Glowinski et al., 2008), physiological signals are used in affective computing. (E.g. Picard, Vyzas & Healey, 2001), linguistic information (e.g. Alm, Roth & Sproat, 2005) and acoustic features (e.g. Dellaert, Polzin & Waibel, 1996). In addition, walking is another modality with great potential. As the most common daily behavior that can be easily observed, psychologists have found that body movement and walking style reflect a walker's emotional states. Human observers were able to identify different emotions from walking, such as the amount of arm swing, stride length, and heaviness (Montepare, Goldstein & Clausen, 1987). Even when gait information was minimized by using point light displays, which meant representing body movement with only a small number of illuminated points, observers could still judge the category and intensity of emotion (Atkinson et al., 2004). The attribution of gait features and other body language to the recognition of specific affective states has been reviewed in (Kleinsmith & Bianchi-Berthouze, 2013).

In recent years, gait information has already been used in affective computing. Janssen et al. (2008) reported emotion recognition using artificial neural networks in human walking using kinetic data collected by a force platform and kinematic data captured by a motion capture system. Using a tag-based motion tracking system, researchers have developed computational methods to recognize emotions from walking in both an inter-individual (comparable to recognizing the affective state of an unknown pedestrian) and a person-dependent state (comparable to recognizing the affective state of a pedestrian). Familiar pedestrian) (Karg et al., 2009a; Karg et al., 2009b; Karg, Kuhnlenz & Buss, 2010). These walking information recording technologies already made it possible to automatically recognize the emotional state of the pedestrian, but due to the high costs of a trained person, technical equipment and maintenance (Loreen, Markus & Bernhard (2013)), the applications of these non-portable systems were severely limited.

Microsoft Kinect is a low-cost, portable, camera-based sensor system with an official software development kit (SDK) (Gaukrodger et al., 2013; Stone et al., 2015; Clark et al., 2013). As a tag-free motion sensing system, Kinect could continuously monitor three-dimensional body movement patterns and is a practical option for developing a low-cost, widely available motion recognition system for people's daily walks. Kinect's validity has been demonstrated in gesture and motion recognition studies. Kondori et al. (2011) identified head position using Kinect, Fern'ndez-Baena, Susin & Lligadas (2012) found it worked well for tracking simple stepping movements, and Auvinet et al. (2015) successfully detected walking cycles on a treadmill using Kinect. In a report by Weber et al. (2012) the accuracy and sensitivity of Kinect-derived kinematic measures such as reach distance, joint angles, and spatiotemporal parameters of gait were estimated and found to be comparable to gold standard marker-based motion sensing systems such as Vicon. On the other hand, the use of Kinect in the medical field has also been reported recently. Lange et al. (2011) used Kinect as a game-based rehabilitation tool for balance training. Yeung et al. (2014) found that the Kinect was valid in assessing body sway in clinical settings. Kinect also performed well in measuring some clinically relevant movements of people with Parkinson's disease (Galna et al., 2014).

Since the emotional states of pedestrians could be reflected in their walking (Montepare, Goldstein & Clausen, 1987; Atkinson et al., 2004), Kinect was found to be an inexpensive, portable, but valid tool for recording human body movements (Auvinet et al. ., 2015; Weber et al., 2012), using Kinect to recognize emotions from gait could be a feasible practice. Given the great value of automatic emotion recognition (Cowie et al., 2001; Silva, Morikawa & Petra, 2012), this practice is worth trying. By automating the recording and analysis of body expressions, especially by applying machine learning methods, researchers have been able to use more and more low-level properties of configurations directly described by 3D coordinate values in emotion recognition (De Silva & Bianchi-Berthouze, 2004; Kleinsmith & Bianchi-Berthouze, 2007). The low-level data-based features extracted from the original 3D coordinates could not provide an intuitive high-level description of the gait pattern under certain affective states, but can be used to train computational models to effectively recognize emotions. We hypothesize that the emotional states of pedestrians (such as happiness and anger) could be reflected in their walking information recorded by Kinect in the form of the coordinates of the body's main joints, and the states could be recognized using machine learning methods. We conducted an experiment to test this hypothesis and try to develop a computational method to recognize emotions from Kinect walking recordings.

II. FACE RECOGNITION

During the 1990s, holistic approaches were used for face recognition. Handcrafted local descriptors have become popular. Local approaches to feature learning were followed in the early 1920s and then in the late 21st century. The following algorithms are currently widely used and implemented in OpenCV: Eigenfaces (1991), Local Binary Patterns Histograms (LBPH) (1996), Fisherfaces (1997), Scale Invariant Feature Transform (SIFT) (1999), Speed Up Robust Features (SURF) (2006). Each method uses a different approach to extract image information and compare it to the input image. Fischer-faces and Eigenfaces have almost similar approaches, as well as SURF and SIFT. LBPH is a simple but very effective method, but it is slow compared to modern face recognizers. These algorithms are not faster compared to modern facial recognition algorithms. Traditional algorithms cannot be trained just by taking a single image of a person. Some of the widely used Deep Learning based face recognition systems are as follows: DeepFace, DeepID series of systems, VGGFace, FaceNet. Face recognition generally takes pictures of the face and finds important points such as the corner of the mouth, eyebrows, eyes, nose, lips, etc. The coordinates of these points are called facial feature points, there are 66 such points. In this way, different technique of finding feature points gives different results.

A. Steps to build the face recognition system

We need to install 2 libraries to implement face recognition. Dlib is a modern C++ toolkit containing a machine learning algorithm and tools for building complex software in C++ to solve real-world problems. The face_recognition library, created and maintained by Adam Geitgey, includes the dlib face recognition feature. OpenCV for some image processing. Now that you've downloaded all the important libraries, let's import them and build the system. After importing the libraries, the image needs to be loaded. Face_recognition library loads images in BGR form, so to print the image you should convert it to RGB using OpenCV.

B. Find face locations and draw bounding boxes

You need to draw a bounding box around the faces to show whether a human face has been detected or not.

C. Train an image for face recognition

This library is built to automatically find the face and only work with faces, so you don't have to crop the face from the images. In this phase, we will convert the image of the train into some encodings and save the encoding with the given name of the person for this image. For testing we load an image and convert it to encoding and now in training we match the encoding with the stored encoding, this matching is based on finding the maximum similarity. When you find the encoding matching the test image, you get the name associated with the train encoding.

Note: face_recognition.compare_faces returns true, false if the person is the same in both images.

D. Building a face recognition system

Import the necessary libraries. Define the path to the folder where the training image dataset will be stored.

Note: for training, we only need to place the training images in the path directory, and the image name must be in the format of person_name.jpg/jpeg.

For example: As you can see in my student_images path, I have 6 people. So our model can only recognize these 6 people. You can add more images to this directory to recognize more people.

- Now create a list to store the person_name and image array.
- Go through all the image files present in the path directory, read the images and append the images array to the images list and the file name to classNames.
- Create a function to encode all train images and store them in the encoded_face_train variable.
- Creation of a function that creates an Attendance.csv file to store attendance with time.

Note: here you need to create the Attendance.csv file manually and specify the path in the function with open ("filename.csv", 'r+') it will create the file and 'r+' mode is used to open the file for reading and writing. First we check if the participant's name is already available in the attendance.csv, we will not write the attendance again. If the participant name is not available in visitance.csv, we write the participant name with the time of the function call.

Read webcam for real-time recognition

- Resize the image by 1/4 for the recognition part only. The output frame will be the original size.
- Resizing improves frames per second.
- face_recognition.face_locations () is called on a resized image (imgS).

- face_recognition.distance() returns the distance array of the test image with all images present in our train directory.
- The index of the minimum distance of the surfaces will be the corresponding surface.
- After finding a matching name, we call the markAttendance function.
- Draw a bounding box using cv2.rectangle().
- We insert the matching name into the output frame using cv2.putText()

III. METHODS FOR EMOTION RECOGNITION TASKS

A. *Experiment design*

Fifty-nine graduate students of the University of the Chinese Academy of Sciences, including 32 females and 27 males with a mean age of 24.4 years (SD = 1.6), participated in this study. Good health was required based on self-report, and individuals were excluded if they reported any injury or disability affecting walking. The experiment was conducted in a bright, quiet room with a 6 m*1 m walkway marked with adhesive tape on the floor in the middle of the room (Fig. 1). Two Kinect cameras were placed oppositely at the two ends of the sidewalk to record walking information (Fig. 2). After informed consent, participants performed the first round of the experiment to produce movements in the neutral and angry conditions. Starting at one end of the trail, participants first walked back and forth along the trail for 2 min (neutral condition) while Kinect cameras recorded their body movements. Participants were then asked to indicate their current emotional state of anger on a scale from 1 (no anger) to 10 (very angry). Then, participants watched an approximately 3-minute video clip of an irritating social event, which was selected from a database of Chinese emotional film clips and used to elicit audience anger (Yan et al., 2014), on a computer in the same room. To ensure that the emotion evoked by the video would last while walking, participants began walking back and forth on the footpath immediately after watching the video (angry condition). When this 1-minute anger walk ended, they were asked to indicate their current emotional state of anger and their state when the video ended on a ten-point scale. Figure 3 shows the whole process of this first round of the experiment.

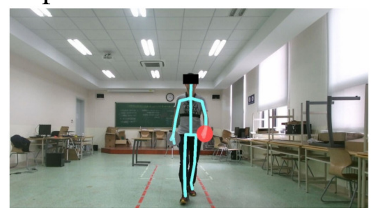

Fig. 1. The experiment scene.

The second round of the experiment was conducted according to the same procedures, with the video being a funny movie clip (Yan et al., 2014) and the scale measuring the emotional state of happiness. There was an interval of at least 3 hours between the two rounds of experiments (participants left and then returned) to avoid possible interference between the two induced emotional states. Each participant completed two rounds of the experiment and left a 1-min recording of anger-primed walking (angry condition), a 1-min recording of happiness-primed walking (happy condition), and two 2-min pre-emotion walking recordings. Baseline of each wheel (neutral state). Each time before starting to walk on the trail, the participants were instructed to walk naturally as in everyday life. The entire protocol was approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences (Approval number: H15010).

B. *Gaits data collection*

With Kinect cameras placed at the two ends of the path, the participants' walking information was recorded as video with a frame rate of 30 Hz. Each image contains 3-dimensional information about 25 body joints, including the head, shoulders, elbows, wrists, hands, spine (shoulder, middle, and base), hips, knees, ankles, and feet, as shown in Fig. 4. With the help of the official Microsoft SDK Beta2 for Kinect and customized software (Microsoft Visual Studio 2012) were exported and further processed the 3-dimensional coordinates of 25 joints with the camera position as the starting point. Gait data recorded by the two Kinect cameras were processed independently.

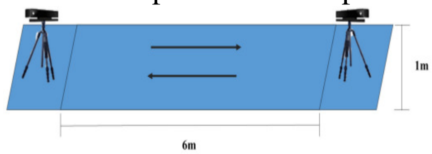

Fig. 2. The schematic of the experiment environment.

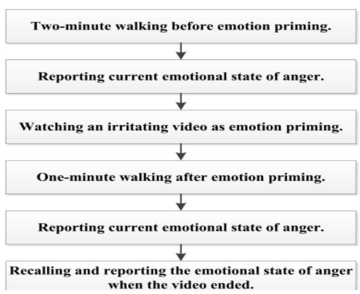

Fig. 3. The procedures of the first round experiment.

C. Data Processing

1) ***Preprocessing-Joint selection:*** As shown by psychological studies on the perception of biological motions (e.g., Giese & Poggio, 2003), a few points representing the trunk and limbs were sufficient to provide information for accurate recognition. Referring to Troje's virtual marker model (Troje, 2002), which has been used in many psychological studies, and on the principle of simplification, we selected 14 joints to analyse walking patterns, including the spine, neck, shoulders, wrists, elbows, hips. , knees and ankles. The base joint of the spine was also used to represent the position of the subject on the footpath relative to the Kinect and to transform the coordinate system. After selecting a joint, one frame contains the position of 3 dimensions out of 14 joints, which can provide a vector of 42 dimensions. Each participant left 4 uninterrupted gait recordings (before anger exposure, after anger training, before happiness exposure, and after happiness training), assuming each recording consisted of T frames, then the data of one recording can be described by the matrix T ∗42 .

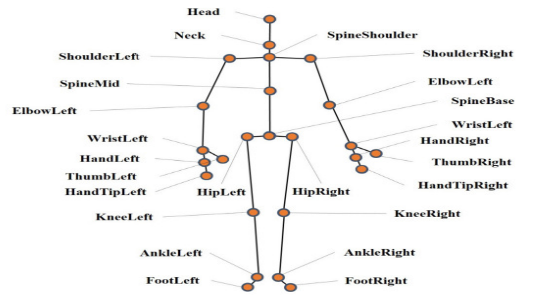

Fig. 4. Stick figure and location of body joint centers recorded by Kinect.

2) ***Coordinate system transformation:*** Different objects may have different positions relative to the Kinect camera when they walked along the path, and using 3D coordinates directly with the camera position as the starting point may cause many errors in the gait pattern analysis. To solve this problem, we changed the coordinate system to instead use the spine joint position in each frame as the starting point.

3) ***Sliding window gauss filtering:*** The original walking recordings contained noises and burrs and needed to be smoothed out. We apply sliding window Gaussian filtering to each column of the matrix of each record. The window length is 5 and the convolution kernel c = [1, 4, 6, 4, 1]/16, which is a frequently used lower Gaussian filter (Gwosdek et al., 2012).

4) ***Differential operation:*** Because the change in joint position between each frame reflects the dynamic part of walking more than the joint position itself, we applied a differential operation to the matrix of each recording to obtain the 3-dimensional position changes of the 14 joints between each frame.

5) ***Data segmentation:*** Because the joint coordinates obtained by the Kinect were not accurate when the participant rotated. We ditched the pivot frames and split one board into multiple straight-walk sections. Anterior segments recorded movements when participants faced the camera, and posterior segments recorded movements when participants returned to the camera. To ensure that each segment covered at least one step, we kept only segments containing at least 40 frames.

D. Feature extraction

It looks quite different for the front and back of the same walk, so we extracted features from the front and back segments separately. Because human walking is periodic and each segment in our study covered at least one step, we performed a Fourier transform and extracted 42 main frequencies and 42 corresponding phases on each segment. The averages of the anterior segments and the averages of the posterior segments of one recording were then calculated separately, obtaining 84 features from the anterior segments and 84 features from the posterior elements. This means that we extracted a total of 168 functions from one board.

Since the value of different features varied greatly, in case some important features with small values might be ignored during model training, all features were first processed by Z-score normalization. To improve computational efficiency and reduce redundant information, Principal Component Analysis (PCA) was then used for feature selection, as it was found that PCA can perform much better than other techniques on training sets with small size (Martinez & Kak, 2001). . The selected features were used in model training.

E. Model training

In this process, three computational models were created to distinguish anger from neutral (before anger activation), happiness from neutral (before happiness activation), and anger from happiness. Because different classifiers may result in different classification accuracies for the same data, we trained and evaluated several commonly effective classifiers, including NaiveBayes, Random Forests, LibSVM, and SMO, to select a better model. These four classification methods were used with 10-fold cross-validation. A more detailed description of the training process can be seen in the report by Li et al. (2014). In this process, three computational models were created to distinguish anger from neutral (before anger activation), happiness from neutral (before happiness activation), and anger from happiness. Because different classifiers may result in different classification accuracies for the same data, we trained and evaluated several commonly effective classifiers, including NaiveBayes, Random Forests, LibSVM, and SMO, to select a better model. These four classification methods were used with 10-fold cross-validation. A more detailed description of the training process can be seen in the report by Li et al. (2014).

TABLE I
SELF-REPORT EMOTIONAL STATES BEFORE AND AFTER EMOTION PRIMING.

	Before priming(BP)	**After priming I**	**After priming II**
		(API:before walking)	**(API:after walking)**
Round 1: anger priming	1.44(0.93)	6.46(1.99)	5.08(1.98)
Round 2: happiness priming	3.88(2.49)	6.61(2.22)	5.63(2.24)

Notes. The average of participants' self-ratings was shown in the table with the standard deviation in the parenthesis.

TABLE II

THE ACCURACY OF RECOGNIZING ANGRY AND NEUTRAL

	NaiveBayes	**RandomForests**	**LibSVM**	**SMO**
KINECT1	80.5085	52.5424	72.0339	52.5424
KINECT2	75.4237	-	71.1864	-

Notes. Table entries are accuracies expressed as a percentage. Values below chance level (50%) are not presented.

TABLE III

THE ACCURACY OF RECOGNIZING HAPPY AND NEUTRAL

	NaiveBayes	**RandomForests**	**LibSVM**	**SMO**
KINECT1	79.6610	51.6949	77.9661	-
KINECT2	61.8644	51.6949	52.5414	-

Notes. Table entries are accuracies expressed as a percentage. Values below chance level (50%) are not presented

IV. RESULTS

An interface for the Intelligent Attendance System was created. Using the interface, images of individual people are recorded and stored in a training data set. At the same time, their information is stored in the database. Finally, images of individual persons are tracked and recognized.

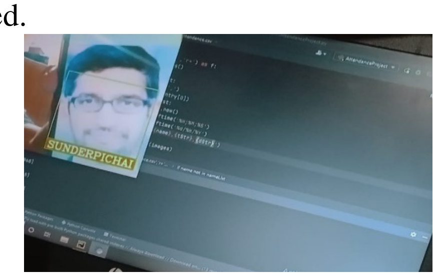

Fig. 5.Webcam capturing the image.

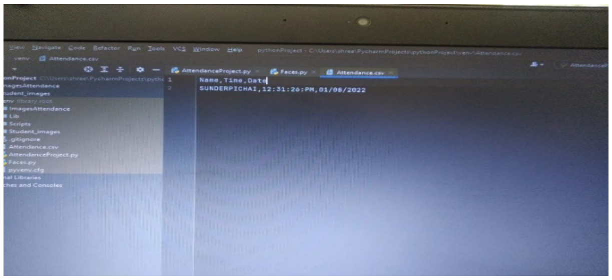

Fig. 6.Attendance showing mark in csv file.

A. *Self-reports of emotional states*

In the current study, emotional states on 10-point scales were used to estimate the effect of emotional priming. As shown in Table 1, for both anger and happiness priming, AP (post-priming) I and APII emotional state ratings were higher than BP (pre-priming). A paired samples t test (by SPSS 15.0) showed that: for anger activation, pre-activation anger ratings were significantly lower than API ($t [58] = 18.98$, $p < 0.001$) and APII ($t [58] = 14.52$, $p < 0.001$); for happiness priming, prepriming ratings of happiness were also significantly lower than API ($t [58] = 10.31$, $p < .001$) and APII ($t [58] = 7.99$, $p < .001$). These results indicated that both anger and happiness priming successfully induced changes in emotional state along the corresponding dimension. In the first round of the experiment, participants generally experienced more anger when walking after the video than before the video, and the same happened for happiness in the second round of the experiment.

TABLE IIV

THE ACCURACY OF RECOGNIZING ANGRY AND HAPPY

	NaiveBayes	RandomForests	LibSVM	SMO
KINECT1	52.5424	55.0847	-	51.6949
KINECT2	-	51.6949	-	50.8475

Notes. Table entries are accuracies expressed as a percentage. Values below chance level (50%) are not presented.

B. *The recognition of primed emotional states and neutral state*

The accuracy of the classifier was the proportion of correctly classified cases in the test set. Table 2 shows the accuracy of each classifier in recognizing angry and neutral. The results from the data captured by the two Kinect cameras were presented separately. With the NaiveBayes and LibSVM classifiers, the computational model was able to recognize anger with relatively high accuracy, especially for NaiveBayes. Table 3 shows the accuracy of each classifier in recognizing happy and neutral. NaiveBayes and LibSVM, especially the former, also performed better than the other two classifiers. The accuracy can be up to 80% for both angry state recognition and happy state recognition. There were consistent differences between KINECT1 and KINECT2 results: accuracy using data recorded by KINECT1 was better than KINECT2.

C. *The recognition of angry and happy*

While the recognition accuracy of the prime states and the neutral state was quite high, the performance of the same classifiers in distinguishing between the angry and happy states was not ideal. As Table 4 shows, the highest that came from random forests was only around 55%. In general, the recognition accuracy of these computational models in the current study did not appear to be well above chance. The results from the KINECT1 data also seemed slightly better than the results from KINECT2.

V. DISCUSSION

Our results partially supported the hypothesis stated at the end of the Introduction: the emotional states of pedestrians (such as happiness and anger) could be reflected in their gaits recorded by Kinect in the form of coordinates of the main joints of the body, and the states could be recognized using machine learning methods. Participants' self-reports indicated that emotion priming in the current study successfully achieved the expected effect: participants walked in an angrier emotional state than before anger activation and walked in a happier emotional state than before happiness activation. In fact, the emotional state during walking before the activation in the study could be considered neutral, representing the "normal state" of the participant, since we did not exert any influence. Thus, the present results could be seen as differentiating the angry/happy condition from the neutral condition based only on the gait data recorded by the Kinect.

These results also demonstrated the feasibility of recognizing emotional states through walking with the help of low-cost portable Kinect cameras. Although some gait characteristics, such as arm swing rate, stride length, and gait speed, have been found to reflect a walker's emotional state (Montepare, Goldstein, & Clausen, 1987), it is not surprising that accurate judgments of emotional state are made automatically based on these isolated the indicators could be difficult for a computer and even for human observers if the intensity of the target emotion was minute. The recognition method in our study did not depend on several certain emotionally relevant indices of body movements, but used continuous dynamic walking information in the form of 3D coordinates. Machine learning made full use of these low-level features, and the high recognition accuracy of some common classifiers implied the validity of the method. In fact, the low-level features used in this study were not exclusive to walking, and there was great potential for using this method in emotion recognition using other types of body movements.

However, in the current study, we did not distinguish very well between the two primary emotional states, anger and happiness. Basically, there were two possible reasons. First, experimentally induced anger and happiness may be relatively mild and appear similar when reflected in walking. In previous studies, participants were often required to recall a past situation or imagine a situation associated with a certain affect while walking (Roether et al., 2009; Gross, Crane & Fredrickson, 2012), and the emotional states induced may be relatively stronger than in our study. As highly arousing emotions, both anger and happiness share some similar kinematic characteristics (Gross, Crane & Fredrickson, 2012), which also makes it difficult to distinguish each other using computational models. Second, it was also possible that the difference between anger and happiness was already manifested in walking, but the feature extraction model training methods were not sensitive enough to take advantage of it. Considering evidence from other reports of a difference in gait characteristics between angry and happy states (Montepare, Goldstein & Clausen, 1987; Roether et al., 2009; Gross, Crane & Fredrickson, 2012), other optimized computational methods may be valid to distinguish these minute but distinct emotional states.

Although the current method can only recognize induced emotions and neutral states, it has the potential to bring benefits in application. The automatically recognized state of emotional arousal could, together with other information, be a valuable resource for decision-making. For example, if emotional arousal were detected in a person diagnosed with major depressive disorder, it would likely be an indicator of a negative mood episode. Another example could be the practice of security personnel in detecting hostility or fraud. In this situation, with some background knowledge, automatic detection of an unusual state of emotional arousal from the target would be a meaningful cue for the observer to alleviate information overload and reduce erroneous judgments, even if the emotion type of that arousal is not precisely determined. Compared to asking participants to act with certain emotions while walking (Westermann, Stahl & Hesse, 1996), the walking of the participants in our study was more common in everyday life, which increased the applicability of our method. In addition, one Kinect camera was able to track up to six individual body movements, making it easier to monitor and analyze the effects on emotional social interactions between individuals. Since emotional features of mental disorders have often appeared in the interpersonal interaction setting (Premkumar et al., 2013; Premkumar et al., 2015), using Kinect data to recognize affect may have even more practical value than other methods.

This study still had some limitations. There was a consistent difference in data quality between the two Kinect cameras, possibly indicating that we did not perfectly control the illumination intensity of our experimental environment. To make the experimental conditions as close as possible to the natural state, we used the natural emotional state before activation as a baseline without any manipulation, so that the baseline was only "approximately" neutral and could vary between participants. Despite these limitations, this study demonstrates a feasible method for automatic recognition of emotional states from walking using portable, low-cost Kinect cameras with great potential for practical value. It would be worthwhile for future studies to improve this method to increase the effectiveness of distinguishing certain types of affect. In our study, the extraction of low-level features from gait was indiscriminate. Indeed, it has been found that different aspects of gait patterns may relate to certain dimensions of emotion (such as level of arousal or valence) to varying degrees (Pollick et al., 2001). So, a possible strategy in future study could be to add some features designed based on the characteristics of the target emotion, to better utilize the features in model training.

A. Appendix. The preprocessing and feature extraction of data

1) ***The description of one frame data:***One frame gait data, after selecting 14 joints, can be expressed by a 42 dimension vector:

$$\mathrm{j_t} = [\mathrm{x_1}, \mathrm{y_1}, \mathrm{z_1}, \mathrm{x_2}, \mathrm{y_2}, \mathrm{z_2}, \dots, \mathrm{x_{14}}, \mathrm{y_{14}}, \mathrm{z_{14}}]$$
(1)

2) ***The description of the data of one record:***One record could be expressed by the $T * 42$ matrix, withj_1,j_2,...,j_t,...,j_T as 42 dimension vectors describing each frame:

$$\mathrm{J} = [\mathrm{j_1}, \mathrm{j_2}, \dots, \mathrm{j_t}, \dots, \mathrm{j_T}]^{\mathrm{T}}$$
(2)

3) ***The process of coordinate system transformation:***In one frame data j_t , the first three columns were the 3-dimension coordinates of the spinebase joint, so the coordinate transformation was conducted by:

$$x_i^t = x_i^t - x_1^t$$
$$y_i^t = y_i^t - y_1^t$$
$$z_i^t = z_i^t - z_1^t$$
$(1 \le t \le T, 2 \le i \le 14)$(3)

4) ***The process of sliding window gauss filtering:***The process of filtering was conducted as follow:

$$x_i^t = [x_i^t, x_i^{t+1}, x_i^{t+2}, x_i^{t+3}, x_i^{t+4}] \cdot c$$
$$y_i^t = [y_i^t, y_i^{t+1}, y_i^{t+2}, y_i^{t+3}, y_i^{t+4}] \cdot c$$
$$z_i^t = [z_i^t, z_i^{t+1}, z_i^{t+2}, z_i^{t+3}, z_i^{t+4}] \cdot c$$
$(1 \le t \le T - 4, 1 \le i \le 14)$(4)

5) ***The process of differential operation:***To obtain the changes of 3-dimension position of 14 joints between each frame, the differential operation is conducted by:

$$\mathrm{j_{t-1}} = \mathrm{j_t} - \mathrm{j_{t-1}} (2 \le \mathrm{t} \le \mathrm{T})$$(5)

6) ***The process of data segmentation:***One record was divided into several front segments and back segments. If one record J contained n front segments and m back segments, it could be described as a series of matrices as follow. Each matrix had 42 columns, but may have different counts of rows, because each segment may contain different number of frames:

$$\begin{cases} \mathrm{Front_i}, 1 \le \mathrm{i} \le \mathrm{n} \\ \mathrm{Back_j}, 1 \le \mathrm{j} \le \mathrm{m}. \end{cases}$$(6)

7) ***The process of feature extraction-:***We ran Fourier transformation to acquire features of each $Front_i$and $Back_i$, obtaining the main frequencies $f_1^i, f_2^i, \dots, f_{42}^i$ and the corresponding phases $\varphi_1^i, \varphi_2^i, \dots, \varphi_{42}^i$ for each segment. The averages of n front segments and m back segments of a single record were calculated separately, obtaining$Feature_{front}$ and $Feature_{back}$ of each record:

$$\mathrm{Feature_{front}} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}}^{\mathrm{n}} [\mathrm{f_1^i}, \mathrm{f_2^i}, \dots, \mathrm{f_{42}^i}, \varphi_1^{\mathrm{i}}, \varphi_2^{\mathrm{i}}, \dots, \varphi_{42}^{\mathrm{i}}]$$
(7)

$$\mathrm{Feature_{back}} = \frac{1}{\mathrm{m}} \sum_{\mathrm{i}}^{\mathrm{m}} [\mathrm{f_1^i}, \mathrm{f_2^i}, \dots, \mathrm{f_{42}^i}, \varphi_1^{\mathrm{i}}, \varphi_2^{\mathrm{i}}, \dots, \varphi_{42}^{\mathrm{i}}]$$
(8)

$\mathrm{Feature_{front}}$ & $\mathrm{Feature_{back}}$ of each record contained 84 features respectively, and the complete feature matrix Feature of each record was the combination of $\mathrm{Feature_{front}}$and $\mathrm{Feature_{back}}$, containing 168 features:

$$\mathrm{Feature} = [\mathrm{Feature_{front}}, \mathrm{Feature_{back}}]$$
(9)

VI. CONCLUSIONS

This paper presents the most productive OpenCV face recognition method available for attendance system. The system was implemented using the LBPH algorithm. Therefore, LBPH is the most authentic and competent face recognition algorithm found in OpenCV for adequate identification and attendance marking using proxy avoidance. In this post, we discussed how to create a face recognition system using the face_recognition library and created an attendance system. You can further design the GUI using Tkinter or Pyqt for the facial recognition attendance system. Various experiments have also been conducted to study emotion recognition using kinect motion capture data of human walking.

ACKNOWLEDGMENT

Support was provided by the National Basic Research Program of China (2014CB744600), the Key Research Program of the Chinese Academy of Sciences (CAS) (KJZD-EWL04), the Strategic Priority Research Program of CAS (XDA06030800) and the Scientific Foundation of the Institute of Psychology of the Academy of Sciences of the Czech Republic (Y4CX143005). The investors had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors disclosed the following grant information: National Basic Research Program of China: 2014CB744600. Chinese Academy of Sciences (CAC) Key Research Program: KJZD-EWL04. CAS Strategic Priority Research Program: XDA06030800. Scientific Foundation of the Institute of Psychology, CAS: Y4CX143005. The authors declare that there are no competing interests. The following information regarding ethical approvals (ie, approval authority and reference numbers, if any) was provided: Institute of Psychology, Chinese Academy of Sciences: H15010. The following information was provided regarding data availability: Raw data was provided as a supplemental dataset. Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.2364#supplemental-information.

REFERENCES

[1] Alm CO, Roth D, Sproat R. 2005. Emotion from text: machine learning for textual emotion prediction. In: Proceedings of the Conference on Human Language Technology and Empirical Methods of Natural Language Processing. Association for Computational Linguistics, 579–586.

[2] Atkinson AP, Dittrich WH, Gemmell AJ, Young AW. 2004. Emotion perception from dynamic and static body displays in spot and fully illuminated displays. Perception 33(6):717–746 DOI 10.1068/p5096.

[3] Auvinet E, Multon F, Aubin C-E, Meunier J, Raison M. 2015. Detection of gait cycles during treadmill walking using Kinect. Gait & Posture 41(2):722–725 DOI 10.1016/j.gaitpost.2014.08.006.

[4] Barakova EI, Lourens T. 2010. Expression and interpretation of emotional movements in social games with robots. Personal and Ubiquitous Computing 14(5):457–467 DOI 10.1007/s00779-009-0263-2.

[5] Cabanac M. 2002. What is emotion? Behavioral Processes 60(2):69–83 DOI 10.1016/S0376-6357(02)00078-5.

[6] Clark RA, Pua Y-H, Bryant AL, Hunt MA. 2013. Validity of the Microsoft Kinect for providing trunk lateral tilt feedback during gait retraining. Gait & Posture 38(4):1064–1066 DOI 10.1016/j.gaitpost.2013.03.029.

[7] Cowie R, Douglas-Cowie E, Tsapatsoulis N, Votsis G, Kollias S, Fellenz W, Taylor JG. 2001. Emotion recognition in human-computer interaction. IEEE Signal Processing Magazine 18(1):32–80 DOI 10.1109/79.911197.

[8] Dellaert F, Polzin T, Waibel A. 1996. Emotion recognition in speech. In: Spoken Language, 1996. ICSLP 96. Proceedings. Fourth International Conference on St. 3. 1970–1973.

[9] De Silva PR, Bianchi-Berthouze N. 2004. Modeling human affective postures: an information-theoretic characterization of posture features. Computer Animation and Virtual Worlds 15(3–4):269–276 DOI 10.1002/cav.29.

[10] Fern'ndez-Baena A, Susin A, Lligadas X. 2012. Biomechanical validation of upper and lower body joint movements of Kinect motion capture data for rehabilitation treatment. In: Intelligent networking and collaborative systems (INCoS), 2012 4th International Conference on, 656–661.

[11] Forgas JP. 1995. Mood and judgment: the affect infusion model (AIM). Psychological Bulletin 117(1):39–66 DOI 10.1037/0033-2909.117.1.39.

[12] Fragopanagos N, Taylor JG. 2005. Emotion recognition in human-computer interaction. Neural Networks 18(4):389–405 DOI 10.1016/j.neunet.2005.03.006.

[13] Galna B, Barry G, Jackson D, Mhiripiri D, Olivier P, Rochester L. 2014. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. Gait & Posture 39(4):1062–1068 DOI 10.1016/j.gaitpost.2014.01.008.

[14] Gaukrodger S, Peruzzi A, Paolini G, Cereatti A, Cupit S, Hausdorff J, Mirelman A, Della CU. 2013. Gait tracking for clinical virtual reality applications: a low-cost solution. Gait and Posture 37 (Suppl. 1): S31 DOI 10.1016/j.gaitpost.2012.12.062.

[15] Giese MA, Poggio T. 2003. Neural mechanisms for biological motion recognition. Nature Reviews Neuroscience 4(3):179–192 DOI 10.1038/nrn1057.

[16] Glowinski D, Camurri A, Volpe G, Dael N, Scherer K. 2008. A technique for automatic emotion recognition by body gesture analysis. In: Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on, 1–6.

[17] Gross MM, Crane EA, Fredrickson BL. 2012. Effort shape and kinematic evaluation of bodily emotion expression during walking. Human Movement Science 31(1):202–221 DOI 10.1016/j.humov.2011.05.001.

[18] Gwosdek P, Grewenig S, Bruhn A, Weickert J. 2012. Theoretical foundations of Gaussian convolution using extended box filtering. In: Bruckstein A, Ter Haar Romeny B, Bronstein A, Bronstein M, eds Scale and variational methods in computer vision. Lecture Notes in Computer Science, Vol. 6667. Berlin Heidelberg: Springer 447–458.

[19] Hudlicka E. 2003. To feel or not to feel: the role of affect in human-computer interaction. International Journal of Human–Computer Studies 59(1):1–32 DOI 10.1016/S1071-5819(03)00047-8.
[20] Janssen D, Schöllhorn WI, Lubienetzki J, Fölling K, Kokenge H, Davids K. 2008. Emotion recognition in gait patterns using artificial neural networks. Journal of Nonverbal Behavior 32(2):79–92 DOI 10.1007/s10919-007-0045-3.

[20] Janssen D, Schöllhorn WI, Lubienetzki J, Fölling K, Kokenge H, Davids K. 2008. Emotion recognition in gait patterns using artificial neural networks. Journal of Nonverbal Behavior 32(2):79–92 DOI 10.1007/s10919-007-0045-3.

[21] Karg M, Jenke R, Kühnlenz K, Buss M. 2009a. A dual PCA approach for interindividual emotion recognition during natural walking. In: MLDM Posters, 51–61.

[22] Karg M, Jenke R, Seiberl W, Kuuhlenz K, Schwirtz A, Buss M. 2009b. Comparison of PCA, KPCA and LDA for feature extraction

for affect recognition in gait kinematics. In: Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on, 1-6.

[23] Karg M, Kuhnlenz 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. Kenji M. 1991. Facial expression recognition from optic flow. IEICE Transactions on Information and Systems 74(10):3474–3483.

[24] Kleinsmith A, Bianchi-Berthouze N. 2007. Recognizing affective dimensions from posture. In: International Conference on Affective Computing and Intelligent Interaction. Berlin Heidelberg: Springer, 48–58.

[25] Kleinsmith A, Bianchi-Berthouze N. 2013. Perception and recognition of affective body expression: a survey. IEEE Transactions on Affective Computing 4(1):15–33 DOI 10.1109/T-AFFC.2012.16.

[26] Kondori F, Yousefi S, Li H, Sonning S, Sonning S. 2011. 3D Head Position Estimation Using Kinect. In: Wireless Communications and Signal Processing (WCSP), 2011 International Conference on, 1-4.

[26] Kondori F, Yousefi S, Li H, Sonning S, Sonning S. 2011. 3D Head Position Estimation Using Kinect. In: Wireless Communications and Signal Processing (WCSP), 2011 International Conference on, 1-4

[27] Lang PJ, Bradley MM, Cuthbert BN. 1998. Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. Biological Psychiatry 44(12):1248–1263 DOI 10.1016/S0006-3223(98)00275-3.

[28] Lange B, Chang C-Y, Suma E, Newman B, Rizzo A, Bolas M. 2011. Development and evaluation of a low-cost gaming tool for balance rehabilitation using the Microsoft Kinect sensor. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual IEEE International Conference, 1831–1834.

[28] Lange B, Chang C-Y, Suma E, Newman B, Rizzo A, Bolas M. 2011. Development and evaluation of a low-cost gaming tool for balance rehabilitation using the Microsoft Kinect sensor. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual IEEE International Conference, 1831–1834.

[29] Li L, Li A, Hao B, Guan Z, Zhu T. 2014. Personality prediction of active users based on microblogging behavior. PLoS ONE 9(1):e84997 DOI 10.1371/journal.pone.0084997.

[30] Loreen P, Markus W, Bernhard J. 2013. Field study of inexpensive marker-free motion analysis for rehabilitation and sports medicine. Gait and Posture 38 (Suppl. 1):S94–S95 DOI 10.1016/j.gaitpost.2013.07.195.

[31] Lopes PN, Salovey P, Côté S, Beers M, Petty RE. 2005. Emotion regulation skills and the quality of social interaction. Emotion 5(1):113–118 DOI 10.1037/1528-3542.5.1.113.

[32] Martinez AM, Kak A. 2001. PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(2):228–233 DOI 10.1109/34.908974.

[33] Montepare J, Goldstein S, Clausen A. 1987. Emotion identification from gait information. Journal of Nonverbal Behavior 11(1):33–42 DOI 10.1007/BF00999605.

[34] Pantic M, Rothkrantz LJ. 2003. towards affect-sensitive multimodal human–computer interaction. Proceedings of IEEE 91(9):1370–1390 DOI 10.1109/JPROC.2003.817122.

[35] Picard RW, Vyzas E, Healey J. 2001. Toward machine emotional intelligence: an analysis of the affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(10):1175–1191 DOI 10.1109/34.954607.

[36] Pollick FE, Paterson HM, Bruderlin A, Sanford AJ. 2001. Perception of affect from arm movements. Cognition 82(2):B51–B61 DOI 10.1016/S0010-0277(01)00147-0.

[37] Premkumar P, Onwumere J, Albert J, Kessel D, Kumari V, Kuipers E, Carretié L. 2015. The relationship between schizotypy and early attention to rejection interactions: the influence of neuroticism. The World Journal of Biological Psychiatry 16(8):587–601 DOI 10.3109/15622975.2015.1073855.

[38] Premkumar P, Williams SC, Lythgoe D, Andrew C, Kuipers E, Kumari V. 2013. Neural processing of criticism and positive comments from relatives in individuals with schizotypal personality traits. The World Journal of Biological Psychiatry 14(1):57–70 DOI 10.3109/15622975.2011.604101.

[39] Roether CL, Omlor L, Christensen A, Giese MA. 2009. Critical features for emotion perception from walking. Journal of Vision 9(6):15–15 DOI 10.1167/9.6.15.

[40] Salovey P, Mayer JD. 1990. Emotional Intelligence. Imagination, Cognition and Personality 9(3):185–211 DOI 10.2190/DUGG-P24E-52WK-6CDG.

[41] Silva LCD, Morikawa C, Petra IM. 2012. State-of-the-art smart homes. Engineering Applications of Artificial Intelligence 25(7):1313–1321 DOI 10.1016/j.engappai.2012.05.002.

[42] Stone E, Skubic M, Rantz M, Abbott C, Miller S. 2015. Average household walking speed: exploring a new metric for assessing mobility and fall risk in the elderly. Gait and Posture 41(1):57–62 DOI 10.1016/j.gaitpost.2014.08.019.

[43] Three NFs. 2002. Decomposing biological movement: a framework for the analysis and synthesis of human gait patterns. Journal of Vision 2(5):2–2 DOI 10.1167/2.5.2.

[44] Weber I, Koch J, Meskemper J, Friedl K, Heinrich K, Hartmann U. 2012. Is MS Kinect suitable for motion analysis? Biomedical Engineering/Biomedizinische Technik 57(SI-1 Track-F) DOI 10.1515/bmt-2012-4452.

[45] Westermann R, Stahl G, Hesse F. 1996. Relative effectiveness and validity of mood induction procedures: an analysis. European Journal of Social Psychology 26:557–580 DOI 10.1002/ (SICI) 1099-0992(199607)26:43.0.CO; 2-4.

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

[47] Yeung L, Cheng KC, Fong C, Lee WC, Tong K-Y. 2014. Evaluation of the Microsoft Kinect as a clinical tool for body sway assessment. Gait & Posture 40(4):532–538 DOI 10.1016/j.gaitpost.2014.06.012.

