Article

CEPP: Perceiving the Emotional State of the User Based on Body Posture

Suk Kyu Lee [1], Mungyu Bae [2], Woonghee Lee [2] and Hwangnam Kim [2,*]

[1] LG Electronics Inc., Seoul 06772, Korea; skjl25@gmail.com
[2] School of Electrical Engineering, Korea University, Seoul 02841, Korea; nardyen@korea.ac.kr (M.B.);
    tgorevenge@korea.ac.kr (W.L.)
* Correspondence: hnkim@korea.ac.kr; Tel.: +82-2-3290-4821

Abstract: Much research has been conducted in the area of face and gesture recognition in order to classify one's emotional state. Surprisingly, utilizing computerized algorithms which recognize emotional conditions based on body postures has not yet been systematically developed. In this paper, we propose a novel method, Computerized Emotion Perception based on Posture (*CEPP*), to determine the emotional state of the user. This method extracts features from body postures and estimates the emotional state by computing a similarity distance. With the proposed algorithm, we will provide new insights into automatically recognizing one's emotional state.

Keywords: emotion recognition; body posture; feature extraction

1. Introduction

The recent proliferation of mobile devices has enriched daily human life. With the advancement of technology, daily interactions between humans and mobile devices are not solely limited to communication, but are expanded to entertainment, learning, energy management, and many other areas. Furthermore, the user interface is being transformed, and recently, many approaches to enhance the interaction between the user and the mobile device have been introduced. However, one factor that is not considered a great deal is emotion. Emotion is very critical component of being a person. It can influence habits of consumption, relationships with others, and even goal-directed activity [1]. Traditionally, human–computer interaction (HCI) viewed that users needed to discard their emotional selves to work efficiently and rationally with the computer [2]. Thus, in the past the emotional aspect was not considered a great deal in interactions with computers. Nonetheless, with the past studies in the field of psychology, it is now impossible to think that an individual can engage in interactive activity with computers without considering his or her emotional system [3]. Henceforth, emotion plays a critical role in all computer-related activity. It is contemplated as an important component for the design process of devising novel services and applications for computers and mobile devices [4]. With emotion, human interactions with computers and mobile devices can evolve as the in human-centric and human-driven forms. To achieve human-centric and human-driven interaction, it is crucial that the devices perceive the user's emotional state and interact with the users accordingly.

Research related to emotion recognition has been widely conducted in the past decades, and has primarily focused on characterizing face and gestures to identify the user's emotional state [5,6]. Nonetheless, research related to recognizing emotion based on human body posture has not been extensively conducted. Even if many express their feelings using a specific posture, to some extent, body posture is considered to describe emotional intensity.

However, recent studies have indicated that body posture and movement present specific information about one's emotional state [7–9]. Furthermore, research using computational resources to automatically determine the emotional condition based on body posture has not been widely conducted.

In this paper, we propose the Computerized Emotion Perception based on Posture (*CEPP*) algorithm to identify one's emotional label. It is designed to extract the body posture's features and compute similarity distance based on dynamic time warping (DTW). To distinguish emotional condition of the user, we generated ground truth images, which depict fundamental emotional states regardless of one's sexuality, culture, and character. Moreover, we created multiple video sequences by having the user pose in random body postures based on specific sequence of emotional states. Based on these videos, we recognize the user's emotional labels to evaluate the effectiveness of *CEPP*. We believe that our novel algorithm based on body posture can provide new insights to determine emotional states automatically.

The remaining paper is organized as follows. We briefly go over the past and current research in Section 2. We describe the design of *CEPP* and then discuss how it is implemented in Section 3. In Section 4, we present the evaluation results of *CEPP*. Finally, we conclude the paper in Section 5.

2. Related Work

In this section, we briefly explain the background information and related work that are used in the proposed algorithm.

In terms of recognizing the emotion in an automatic fashion, Camurri et al. proposed four layers to recognize emotions from dance movements [10]. The authors relied on the body movement to recognize the emotion instead of focusing on the overall body posture as a whole. Additionally, Caridakis et al. proposed a method which utilized a Bayesian classifier to recognize emotion based on body gestures and speech [11]. In this literature, the authors collected multimodal data and then trained the data with a Bayesian classifier to recognize the emotion.

Much of the related research [12] has focused on recognizing emotion based on the user's body movements, gestures, and speech, with authors utilizing a Bayesian classifier or support vector machine (SVM) to train the data. Instead of using a single image of the user, these approaches require extra images to analyze the user's body movement. Furthermore, the machine learning algorithms consume extensive computational resources. We differentiate our algorithm from them by providing a fast emotional condition recognition algorithm (less than 1 second). Moreover, we focus on recognizing the user's entire body posture and provide novel methods for extracting a feature and analyzing the similarity distance.

Another related study was performed by Shibata et al. [13]. In this paper, the authors recognized body emotion from pressure sensors and accelerometers. The main idea of this work is the proposal of a scheme which can recognize emotional state while the user is sitting down, which is in contrast to our study.

One of the areas of application of identifying the emotional state of the user based on body posture is in clinical research. Loi et al. investigated on the problem of recognizing the emotional state of clinically depressed patients based on the body language [14]. Thus, the applications of recognizing emotion based on body posture can be utilized in clinical research to better understand patients.

Lastly, dynamic time warping (DTW) is an algorithm for measuring the similarity between two temporal time sequences. It is widely used in the area of speech and signature recognition, and in some partial shape matching [15]. However, only few works applied DTW to classify user emotional state. One related work was conducted by Castellano et al. [16]. The authors analyzed the speed of user's body movement and amplitude to specify emotional condition. Although this work may look very similar to our proposed algorithm, it is different in terms of feature extraction. The authors measured the quantity of motion based on the user's movement in terms of velocity and acceleration of hand movement. In our proposed algorithm, we utilize the user's overall body posture as a feature. Besides, the computation of similarity distance is an approach different to that used in traditional DTW.

3. Overview of CEPP

In this section, we explain in detail the proposed algorithm. Figure 1 represents the overview of *CEPP* algorithm. *CEPP* consists of two fundamental components: feature extraction and computation of similarity distance. Based on these two components, it recognizes the emotional state of the user by comparing the image of the user with the ground truth images. The procedure of the *CEPP* algorithm works as follows. The first two steps are high-level approaches which rely on image processing techniques. Firstly, we identify the user within the sight of the camera and create a silhouette of the user. Afterwards, we acquire the edge of the user from the user's silhouette. Based on the high-level approach for obtaining the silhouette and the edge of the user, we proceed to the low-level feature based approach by obtaining the center point of the user and compute the similarity distance. We determine the center point of the user and extract the features based on the user's body posture once we obtain the edge of the user. After the feature extraction is complete, we compare the user's body posture image with the ground truth image's feature by computing the Euclidean distance based on the center point to the detected edge of the body posture from the 0 to 360 degrees. When the comparison of Euclidean distance is done, we analyze the similarity distance to find the best match with the ground truth image and classify the emotional condition. Details on each design choice will be explained later in this section.
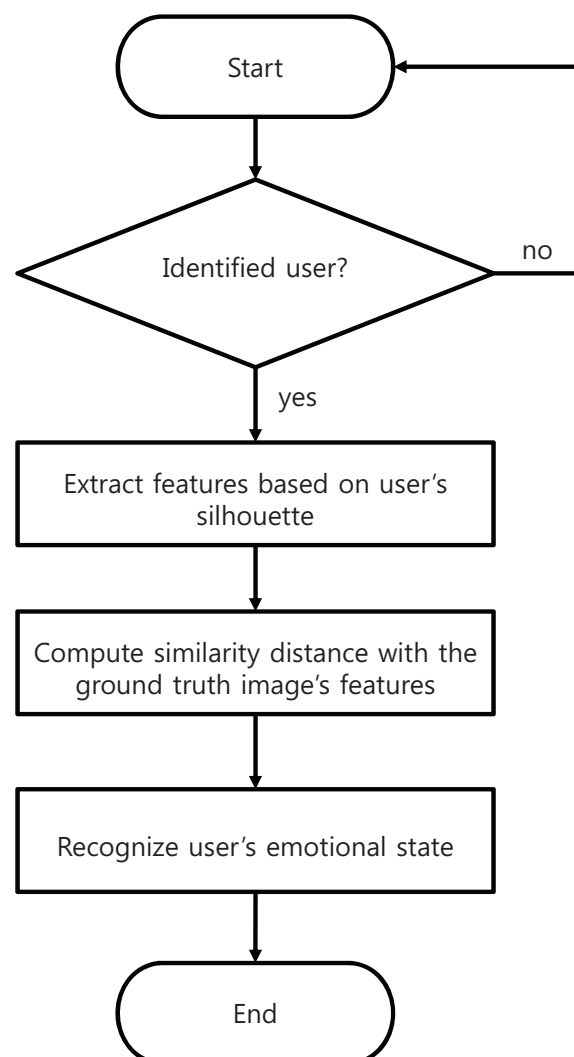
Figure 1. Operation flow of Computerized Emotion Perception based on Posture (*CEPP*).

3.1. Classifying Emotional State

It is crucial to analyze the relationship between body posture and emotional condition in order to identify the emotional labels of the user. Emotional states can be broadly categorized as either positive or negative emotions. Wallbott and Darwin classified emotion based on widely perceived body postures [17,18]. In these works, the authors investigated widely accepted body postures that depicted various emotional labels. Based on these analysis, we identified the common body postures of the emotional states of the user and created ground truth images, which illustrate the body postures that correspond to each emotional state of the user. The ground truth images were generated based on the descriptions of body posture based on the emotional state of the user, shown in Table 1. We created a set of ground truth images of each emotional state of the users. Further, we utilize it as a medium to identify the emotional state of the user. We selected six emotional states for this paper, and these are classified as either positive (elation joy, happiness, and interest) or negative emotions (boredom, disgust, and hot anger). Table 1 depicts our selected emotional labels based on body postures.

Table 1. Description of body postures based on emotional states.

Emotional State	Body Posture
Happiness	Shoulders up
Interest	Lateral hand and arm movement and arm stretched out frontal
Boredom	Raising the chin (moving the head backward), collapsed body posture, and head bent sideways
Disgust	Shoulders forward, head downward and upper body collapsed, and arms crossed in front of the chest
Hot anger	Lifting the shoulder, opening and closing hand, arms stretched out frontal, pointing, and shoulders squared

3.2. Recognizing the User's Emotional State

In this section, we explain how *CEPP* recognizes the emotional state of the user. The pseudo code for *CEPP* is illustrated in Algorithm 1. Firstly, the algorithm is started when the user is identified in the field of view of the camera. After the user is identified, the proposed algorithm obtains the silhouette of the user. The silhouette of the user is obtained by utilizing one the features from OpenNI library that can identify the user with a three-dimensional (3D) camera [19]. With this feature, the user can be classified as white and the background is illustrated as black based on the acquired silhouette of the user. Once the silhouette of the user is completely generated, we compute the center of mass of the user, where we proceed to the next equation:

$$\gamma_x = \frac{1}{N} \sum_{(x,y) \in S} x, \quad \gamma_y = \frac{1}{N} \sum_{(x,y) \in S} y, \tag{1}$$

where γ_x and γ_y represent the center of the user, N depicts the number of detected pixels within the domain S,

$$S = \{(x,y)|I_s(x,y) < 255\}, \tag{2}$$

and I_s is the user's silhouette image.

Note that the center of mass coordinate of x and y (γ_x and γ_y) in the above equation is the average of the pixels that have a value which is not 255 (white). The same procedure is applied to the ground truth images. While we are computing the center of mass of the user, we transform the silhouette

image of the user into an edge image of the user, I_e. Afterwards, we calculate the degree between the center point and the edge of the user's body posture:

$$\phi = \arctan \frac{I_{e(y)} - \gamma_y}{I_{e(x)} - \gamma_x} \times \frac{180}{\pi}, \tag{3}$$

where $I_{e(x)}$ and $I_{e(y)}$ are the x and y coordinates of the pixel that are detected in I_e and ϕ depicts the degree of the center point, γ, corresponding to $I_{e(x)}$ and $I_{e(y)}$. After we obtain ϕ for each corresponding edge pixel, we compute the Euclidean distance, ρ_ϕ, as the following equation:

$$\rho_\phi = \sqrt{(\gamma_x - I_{e(x)})^2 + (\gamma_y - I_{e(y)})^2}. \tag{4}$$

Algorithm 1 Pseudo code of *CEPP*.

1: **procedure** *CEPP*
2:     *init*:
3:         Compute a set of ground truth silhouettes $I_s(g)$
4:         Generate a set of ground truth edges $I_e(g)$
5:         Create a set of ground truth center of point $\gamma(I_s(g))$
6:     *loop*:
7:         **if** user u is identified **then**
8:             Create silhouette of the user $I_s(u)_i$
9:             Generate edge of the user $I_e(u)_i$
10:             Compute center of mass $\gamma(I_s(u)_i)$
11:             **if** γ is computed **then**
12:                 Extract feature $f_u(\gamma(I_s(u)_i), I_e(u)_i)$
13:                 Extract feature $f_g(\gamma(I_s(g)_i), I_e(g)_i)$
14:                 Compute similarity distance $\zeta(f_u, f_g)$
15:             **end if**
16:             Recognize emotional state based on ζ
17:         **end if**
18:         **goto** *loop*.
19: **end procedure**

Then, we sort a set of computed ρ_ϕ based on ϕ from the lowest degree to the highest degree with the matched Euclidean distance. Moreover, we transform the sorted set of ρ_ϕ as the feature f_g or f_u when the computed image is one of ground truth images (the user image).

Once we obtain the features of the ground truth and the user, we proceed to compute similarity distance. Note that f_g is the feature of the ground truth image, and f_u is the feature of the user. f_g and f_u are represented as the following:

$$f_g = \{f_{g(1)}, f_{g(2)}, \cdots, f_{g(i)}, \cdots, f_{g(m)}\}, \tag{5}$$
$$f_u = \{f_{u(1)}, f_{u(2)}, \cdots, f_{u(j)}, \cdots, f_{u(n)}\},$$

where m is the length of the ground truth image's feature, n is the length of user image's feature, i and j are the indexes within the features f_g and f_u, respectively, the component $f_{g(i)}$ of f_g is the ρ_ϕ for the

ground truth image, and $f_{u(j)}$ of f_u is the ρ_ϕ for the user's image. Afterwards, we define the warping matrix D using dynamic programming, where the warping matrix D is as follows:

$$D(i,j) = d(i,j) + \min \begin{cases} D(i-1,j) \\ D(i-1,j-1), \\ D(i,j-1) \end{cases} \tag{6}$$

where $d(i,j)$ is $d(i,j) = |f_{g(i)} - f_{u(j)}|$. Based on this information, we can obtain the similarity distance, ζ_k, which is $\zeta_k = D(m,n)$, where k is the index of the similarity distance for the designated emotional state. If the similarity distance is smaller, the feature of the ground truth image and the user is similar. Thus, once computing of ζ_k for each emotional state is complete, *CEPP* algorithm compares all of the ζ_k of each emotional state and selects the emotional label with the lowest ζ_k,

$$\zeta = \min\{\zeta_1, \zeta_2, \zeta_3, ..., \zeta_l\}, \tag{7}$$

where l denotes the total number of emotional states, and ζ indicates the similarity distance of the recognized emotion by finding the lowest ζ_k. By finding the lowest ζ_k, we can identify the emotional state of the user with the corresponding index of ζ_k.

4. Performance Evaluation

In this section, we present the performance evaluation of our proposed algorithm. We conducted our experiments in terms of evaluating the accuracy of emotion recognition and computation time to operate *CEPP*. Furthermore, we have conducted a user survey with the participants and the viewers of the experimental videos.

4.1. Experimental Configuration

The entire algorithm was built on *C++* in *Ubuntu 14.04*. We tested our proposed algorithm within a hardware platform consisting of an *Intel Core i5* CPU M450, 2.40 GHz, 4.00 GB RAM, *GeForce* 310 M. We utilized *Kinect* for Microsoft Xbox 360 and *OpenNI SDK* was applied as the primary source to acquire the body postures of the user based on the emotional states that we specified in Section 3.1. The user generated the body postures by following and characterizing the emotional states from Table 1 with the sequence of the emotional feeling as described in Table 2.

Table 2. Configuration of the experiment.

Video Scene	Emotional States
All of the emotions	(1) Happiness, (2) Disgust, (3) Sadness, (4) Fearfulness, (5) Hot anger

4.1.1. Details about the Participants

In order to conduct the study, we recruited six subjects to participate in our experiment. The details on the subjects are as follows. Firstly, two out of six subjects were female and rest of the subjects were male. The subjects were between 25 and 34 years of age. Further, the heights of the subjects were between 160 cm and 180 cm. Lastly, the subjects were of Asian ethnicity. None of the subjects had any prior experience with this type of system.

As for the viewers, we recruited four viewers to watch the experimental videos that were obtained with respect to the actions of the subjects. All four viewers were males between 25 and 30 years of age. Each viewer watched the video and answered the user survey questions to verify whether the subjects' body posture presents emotional feeling.

4.1.2. Details on the Experimental Setup

In terms of the experimental setup for this experiment, we set up one camera on top of the shelf so that the camera could capture the entire body of the subject. We focused on obtaining the image of the user from the knee to the head and acquired all of the hand or leg movements through the camera. One of the reasons that we focused on acquiring the image of the entire body of the user was to make sure that we could capture all of the body parts of the user starting from the head to the knee. If the camera can capture only the upper part of the body, then that upper body part cannot be considered as representative of the whole-body posture of the subject. Thus, we set up the camera on the shelf and made sure that the camera looked down at the subject to acquire the entire body posture.

Furthermore, we collected and analyzed the results with our proposed algorithm with total of 2300 images. For each subject, we acquired at least 450 images and executed our proposed algorithm for evaluation. Some of the representative images of the participants based on their emotional feeling are illustrated in Figure 2. The silhouette of each subject is depicted in Figure 3. We analyzed the data by comparing with the subject's ground truth with the experimental video. In addition, we utilized the ground truth of different subjects to analyze the emotional state of the user.
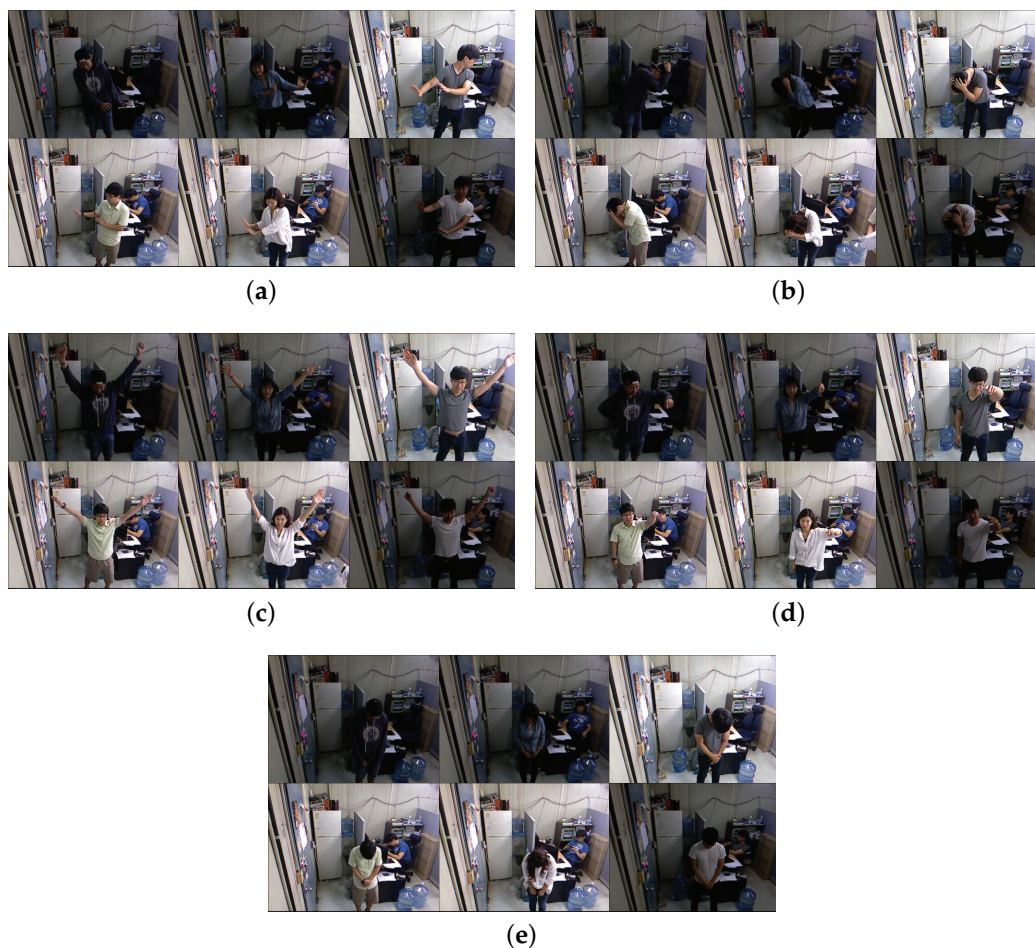

Figure 2. Visual results of all of the emotions. (**a**) Disgust; (**b**) Fearfulness; (**c**) Happiness; (**d**) Hot anger; (**e**) Sadness.
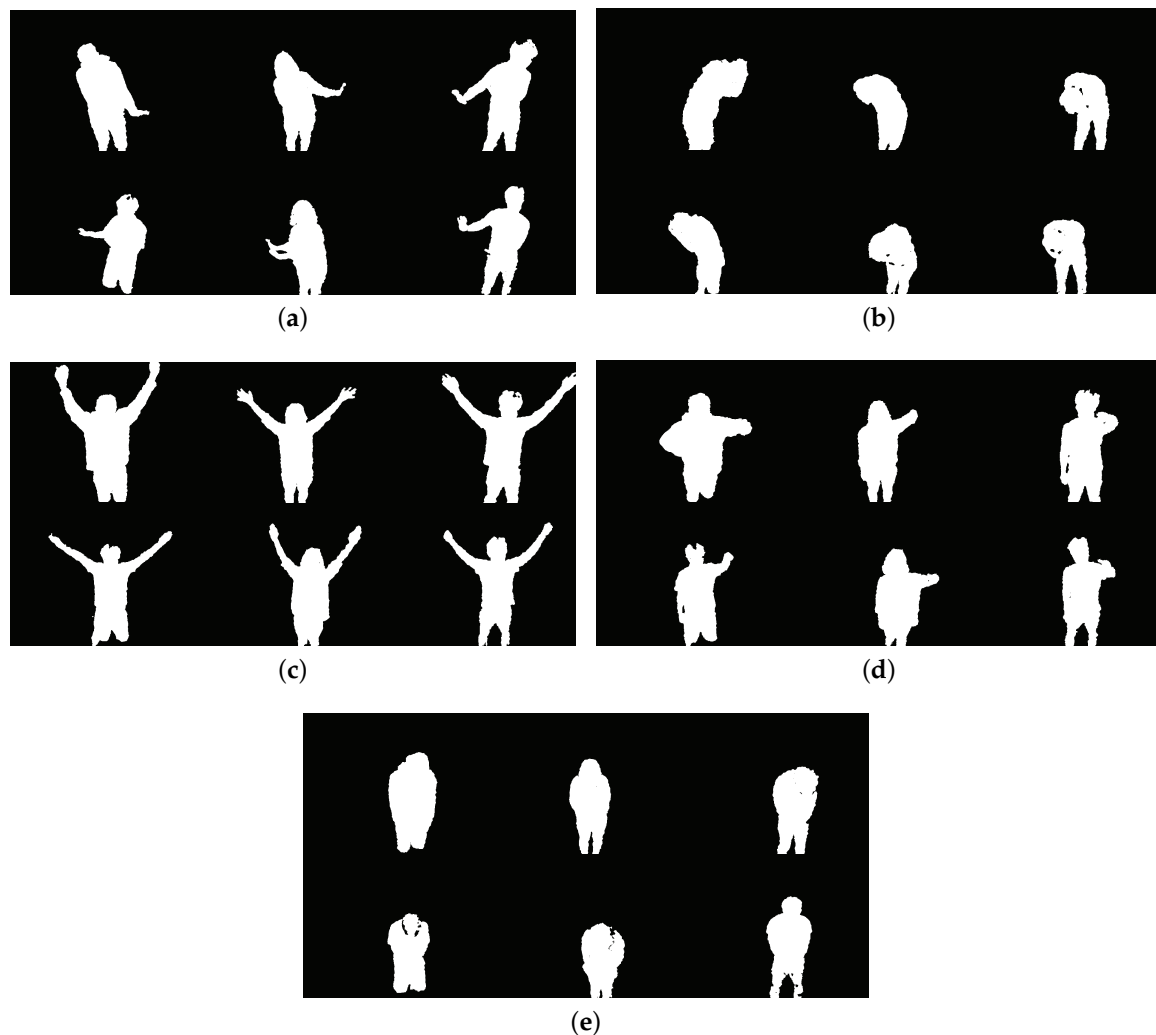
Figure 3. Visual results of all of the emotions in terms of binary image. (**a**) Disgust; (**b**) Fearfulness; (**c**) Happiness; (**d**) Hot anger; (**e**) Sadness.

4.2. Evaluation with Respect to Accuracy

As the first step to evaluate *CEPP*, we conducted an experiment where we evaluated the accuracy of recognizing the emotional state of the user using multiple scenes from Table 2. Figure 4a illustrates the results, where the *x*-axis represents the index of the video frame and the *y*-axis depicts the similarity distance. Figure 4a indicates the result whereby the user demonstrated all of the emotions. For scenes that describe the positive and negative emotions, results are shown in Figure 4b,c. Lastly, Figure 4d shows the result of the scene, which depicts two specific emotional conditions. The emotion with the least similarity distance is indicated as the current emotional label of the user. Based on Figure 4a, we were able to verify that all of the emotional labels were recognized correctly in terms of the sequence from Table 2 by our proposed algorithm. We could clearly see that all of the emotional states were recognized properly based on the sequence of characterizing the emotional conditions. Moreover, similar results for the positive and the negative emotions could be observed. For the two specific emotional states, even though the two emotional states were categorized into either positive or negative emotions, *CEPP* was able to recognize the emotional labels correctly.

In addition, we evaluated the results from Figure 4 in terms of the frequency of appearance of each emotional state that was being recognized and its accuracy within the experimental videos. Based on the results from the figure, we could identify that the number of recognized emotional

states was evenly distributed within the video as we designed in Table 2. For example, emotions recognized within the positive emotion video scene were evenly recognized, with values of 34.83%, 33.48%, and 31.69%. Similar results were observed with negative emotions, where the values were 29.34%, 37.15%, and 33.51%.

As the next step for evaluating *CEPP*, we analyzed the results from Figure 4 in terms of percentages of accurate recognition of emotional state. The summary of the results is listed in Table 3. Moreover, it is depicted in Figures 5 and 6. For all of the video scenes, the accuracies in terms of correct recognition of emotional states for: (1) all of the emotions; (2) positive emotions; (3) negative emotions; and (4) specific emotions were 87.3%, 96.9%, 95.8%, and 97.0%, respectively. For specific one emotional state such as happiness, elated joy, hot anger, or boredom, in the video scenes (all of the emotions, positive emotions, negative emotions, and specific emotions), we were able to see that the percentage of accuracy was 100%. Overall, we were able to verify that the accuracy of *CEPP* maintained above 85% for one entire video sequence. For one specific emotional state such as boredom in the video sequence of all of the emotions, the accuracy of recognizing that particular emotional state was below 80%.

However, for most of the emotional states, the percentage of accuracy was above 80% and for some emotional conditions, it reached to 100%.
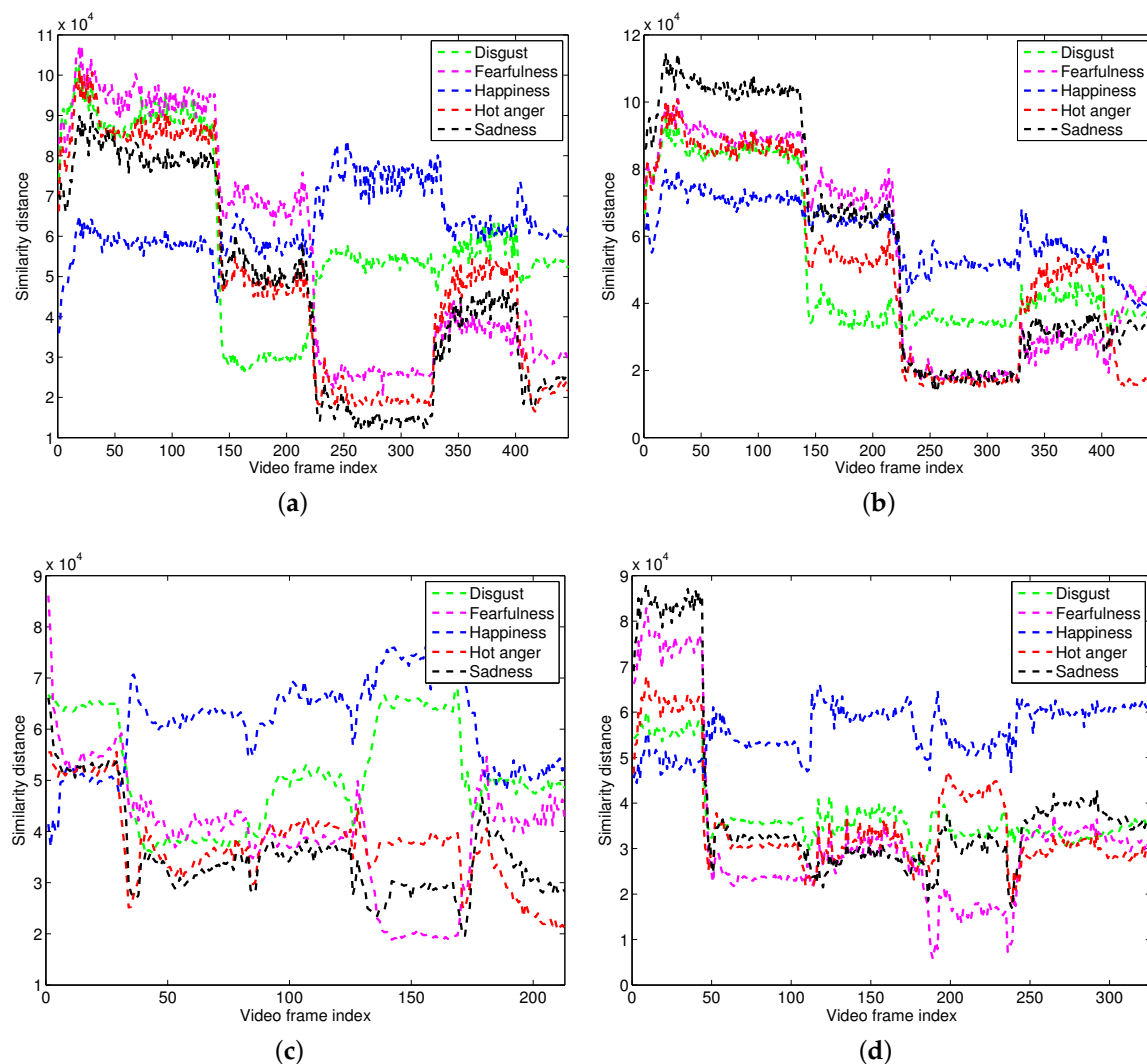

Figure 4. Cross evaluation of different ground truths. (**a**) Configuration 1; (**b**) Configuration 2; (**c**) Configuration 3; (**d**) Configuration 4.

Conclusively, with the *CEPP* algorithm, we were able to identify the emotional state of the user based on body posture. However, it is necessary that the body area be classified and the similarity distance be computed separately to increase the overall accuracy of the algorithm. For example, the hand area needs to be identified and then compared, and the head area needs to be classified to see if the head was leaning backwards or not. Based on the similarity distance from each body region, we can analyze the similarity distance for each region separately and add these similarity distances into one whole result for recognizing the emotional state of the user. With this approach, we believe that the accuracy of *CEPP* can be enhanced.

Figure 5. Overall accuracy.

Table 3. The number of recognized emotional states and their proportions within the video.

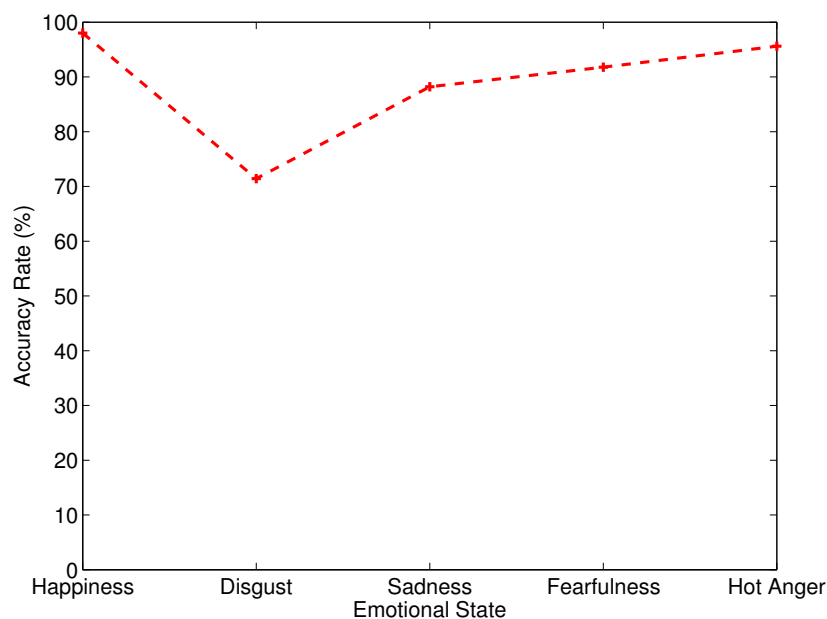| Emotions	Recognized Emotional State						
	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Overall
Happiness	100%	100%	100%	100%	100%	100%	100%
Disgust	60.32%	92.68%	95.45%	95.08%	92.86%	80.70%	86.18%
Sadness	100%	100%	100%	48.57%	84.62%	100%	88.86%
Fearfulness	89.83%	77.78%	92.59%	86.79%	86%	89.87%	87.14%
Hot anger	100%	100%	100%	100%	100%	100%	100%

Figure 6. Accuracy rate of the overal experiment.

4.3. Evaluation with Respect to Computation Time

As the last part of this experiment, we acquired the computation time for extracting features and computed the similarity distance, ζ, for all of the scenes. Figure 7 represents the result of computational time for conducting feature extraction and similarity distance, where the *x*-axis illustrates the index of the video frame and the *y*-axis depicts the computation time. Within the figure, we acquired the computation time for computing the overall similarity distance for one input image and the ground truth image. Moreover, we obtained the average time for computing one similarity distance. The same procedure is applied to extract the features. As we can see in the figure, recognizing one emotional state for one image cost less than 1 second. For the video scenes of: (1) all of the emotions; (2) positive emotions; (3) negative emotions; and (4) specific emotions, the computation time for recognizing the emotional condition were 827.08 ms, 911.36 ms, 749.61 ms, and 807.76 ms, respectively. For computing single similarity distance the average computation times were 137.84 ms, 151.9 ms, 124.94 ms, and 134.63 ms for four video scenes. This refers to computing the similarity distance between the input image and only one ground truth image. The computation time for feature extraction did not change much compared to computation of the similarity distance. As we can see in the figure, feature extraction was not heavily affected, even though the user posed in different body postures. However, computing similarity distance was different. Different body posture affected the computation time. If the user modifies his body posture, the computation time for similarity distance is affected. We believe that different body posture changed the features of the user's body and affected the computation time. The summary of the computation time of the proposed algorithm is listed in Table 4.
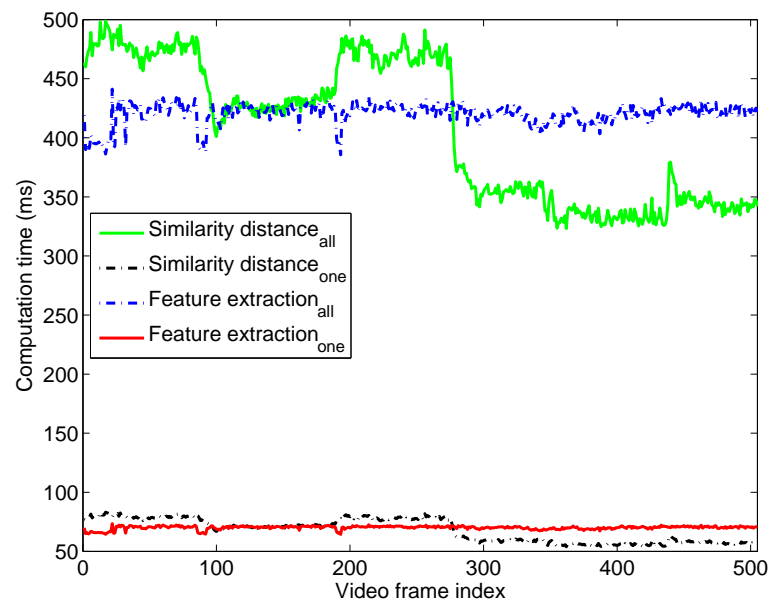
Figure 7. Computation time for recognizing the emotional states.

Conclusively, we could verify that extracting the feature within *CEPP* cost more than computing the similarity distance. Moreover, for analyzing one video frame with all of the ground truth images, including feature extraction and computing similarity distance, it took less than 1 s. Overall, *CEPP* is light and requires less computational resources for its operation. We observed that similar result is shown if *CEPP* is implemented in the mobile device, since the hardware specification for our testing machine and the mobile device is similar.

All in all, we can verify that our proposed algorithm's computational time is less than 1 second. However, compared with the other algorithms that were proposed to identify the emotional state of the user, our algorithm is more robust. One of the motives for this reasoning is that the other algorithms are based on the machine learning algorithm such as the Bayesian classifier and the support vector machine. Machine learning algorithms require extra computational time and yet more time for training of the data. Nevertheless, our algorithm does not need such computational complexity to recognize emotions.

Table 4. Computation time for recognizing emotional state.

Category	Computation Time
Computing all similarity distances	406.58 ms
Computing one similarity distance	67.76 ms
Extracting all features	420.50 ms
Extracting one feature	70.08 ms

4.4. User Study

We have also conducted an user survey to see whether the body posture can present certain emotions. We gave each subject and the viewer of the experimental video four statements to assess. We primarily focused on whether the body posture can represent the emotional state of the person in terms of the subject and the viewers. Thus, we asked the participants of the user study to give a score between the number of 1 (strongly disagree) to 5 (strongly agree). The statements for the users are as the follows.

(*S*.1) "The body posture within the experiment is the most representative body posture to express one's emotional state".

(S.2)   "There is a body posture that I used to express my emotional feeling".

(S.3)   "The body postures that I expressed within the experiment are body postures that I use within my daily life".

(S.4)   "The body posture that I formulated within the experiment is accurate in terms of expressing my emotional feeling"

For the viewers, we provided four similar statements as compared to those that were given to the subjects. Moreover, for statements (V.1) and (V.2), the viewers were given the same statements ((S.1) and (S.2)). For the (V.3) and (V.4), the statements are listed as follows.

(V.3)   "The body postures that the subjects expressed within the experiment are body postures that I use within my daily life".

(V.4)   "The body postures that the subjects formulated within the experiment are accurate in terms of expressing my emotional feeling."

Lastly, for statements (S.3) and (V.3) for the subjects and the viewers, the participants of the user study were asked to give a written comment based on the score that they have given.

The result of the user study is described in Figure 8. The average score of the subjects for statement (S.1) was 3.5, for (S.2) it was 4, for (S.3) it was 3.33, and for (S.4) it was 3.33. For the case of the viewers, the average scores for each statement were as follows: for (V.1) the score was 4.75, for (V.2) it was 4.5, for (V.3) it was 3.75, and for (V.4) it was 4.25. In terms of the comments from questions (S.3), many of the subjects commented that

"The motions chosen in the experiment seem to be somewhat exaggerated. My actions are not so extravagant".

However, the expression itself may still represent the emotional state. Similar comments came out from the viewers of the video. In summary, we found out that body posture may reflect one's emotional state. However, it is noted that different ethnicity and the culture may reflect their emotional feelings with different body postures.
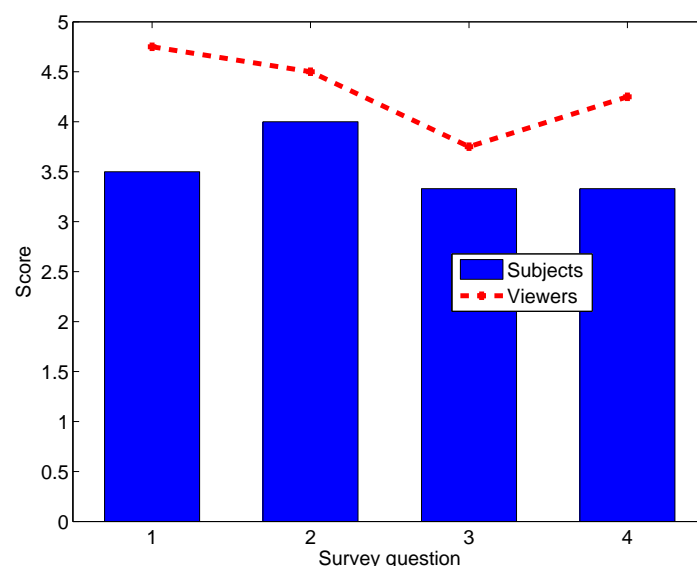

Figure 8. User survey score from the subject and the viewer.

4.5. Remark

In this section, we will address the acceptability and the limitation of the proposed system. We believe that there were limitations throughout the process of conducting the experiment. First of

all, we believe that the number of participants in terms of the gender was not evenly distributed. Also, this may be applied same for the viewers. Throughout the experiment, we tried to recruit diverse range of participants for our study. Nonetheless, with limited resources, it was difficult to achieve the ideal number of participants. It is true that the number of participants may be very small. However, we are planning to add novel features and then recruit more participants to acquire more data in order to derive significant results. Overall, in future work we will try to overcome this issue and diversify the characteristics of the participants and viewers for our user study.

In addition, there were technical limitations in terms of identifying the emotional state of the users with our proposed system. Light conditions did not affect our proposed system in terms of accuracy.

However, when the user is sitting, we believe that the accuracy of the proposed system might be degraded. One of the reasons behind this reasoning is that if the user is sitting on top of the chair, the part of the chair might be included within the silhouette of the user. Thus, the accuracy of the algorithm may be degraded, since our algorithm is based on assuming that the user is standing up. In future work, we do believe that it would be interesting to overcome this challenge by identifying the user while he/she is sitting on a chair. Moreover, in terms of multi-users, we are planning to expand our research by integrating the feature of identifying multiple number of users to see if our proposed algorithm can recognize the emotions of various users.

5. Conclusions

In this paper, we propose the Computerized Emotion Perception based on Posture (*CEPP*) algorithm to determine the emotional state of the user based on body posture. Many papers have dealt with gestures and the face to identify emotional conditions. However, in this paper, we designed a novel method to extract the features and analyzed the similarity distance to classify how user body posture can reflect the user's emotional state.

As for the future work, we will identify more body postures to diversify the emotional states. Furthermore, we will classify the body posture based on the body's region and compute separate similarity distance to improve the overall accuracy of recognizing emotion. Lastly, we will implement *CEPP* into a mobile device. Based on user's daily interaction with the mobile device, we believe that it can be utilized as a novel tool for interacting multimedia contents.

Author Contributions: Suk Kyu Lee designed the system and the experiments. Mungyu Bae conducted and assisted with the experiments. Moreover, Suk Kyu Lee and Mungyu Bae analyzed the data and wrote the paper. Woonghee Lee verified the system design and the algorithm. Hwangnam Kim governed the overall procedure of this research.

References

1. Decisions, O.T.D. The Power of Emotional Appeals in Advertising. *J. Advert. Res.* **2010**, 169–180, doi:10.2501/S0021849910091336.
2. Brave, S.; Nass, C. Emotion in human-computer interaction. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*; Sears, A., Jacko, J.A., Eds.; Taylor & Francis: Milton Park, UK, 2002; pp. 81–96.
3. Gross, J.J. Emotion and emotion regulation. *Handb. Personal. Theory Res.* **1999**, *2*, 525–552.
4. Isomursu, M.; Tähti, M.; Väinämö, S.; Kuutti, K. Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *Int. J. Hum. Comput. Stud.* **2007**, *65*, 404–418.

5. Li, Y. Hand gesture recognition using Kinect. In Proceedings of the 2012 IEEE 3rd International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 22–24 June 2012; pp. 196–199.

6. Saneiro, M.; Santos, O.C.; Salmeron-Majadas, S.; Boticario, J.G. Towards Emotion Detection in Educational Scenarios from Facial Expressions and Body Movements through Multimodal Approaches. *Sci. World J.* **2014**, *2014*, doi:10.1155/2014/484873.

7. Libero, L.E.; Stevens, C.E.; Kana, R.K. Attribution of emotions to body postures: An independent component analysis study of functional connectivity in autism. *Hum. Brain Mapp.* **2014**, 35, 5204–5218.

8. Geiser, M.; Walla, P. Objective measures of emotion during virtual walks through urban environments. *Appl. Sci.* **2011**, *1*, 1–11.

9. Dael, N.; Mortillaro, M.; Scherer, K.R. Emotion expression in body action and posture. *Emotion* **2012**, *12*, 1085–1101.

10. Camurri, A.; Lagerlöf, I.; Volpe, G. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *Int. J. Hum. Comput. Stud.* **2003**, *59*, 213–225.

11. Caridakis, G.; Castellano, G.; Kessous, L.; Raouzaiou, A.; Malatesta, L.; Asteriadis, S.; Karpouzis, K. Multimodal emotion recognition from expressive faces, body gestures and speech. In *Artificial Intelligence and Innovations 2007: From Theory to Applications*; Boukis, C., Pnevmatikakis, A., Polymenakos, L., Eds.; Springer: Berlin, Germany, 2007; pp. 375–388.

12. Jia, X.; Liu, S.; Powers, D.; Cardiff, B. A Multi-Layer Fusion-Based Facial Expression Recognition Approach with Optimal Weighted AUs. *Appl. Sci.* **2017**, *7*, doi:10.3390/app7020112.

13. Shibata, T.; Kijima, Y. Emotion recognition modeling of sitting postures by using pressure sensors and accelerometers. In Proceedings of the IEEE 21st International Conference on Pattern Recognition (ICPR), Tsukuba, Japan, 11–15 November 2012; pp. 1124–1127.

14. Loi, F.; Vaidya, J.G.; Paradiso, S. Recognition of emotion from body language among patients with unipolar depression. *Psychiatry Res.* **2013**, *209*, 40–49.

15. Sankoff, D.; Kruskal, J.B. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*; Addison-Wesley: Reading, MA, USA, 1983; Volume 1.

16. Castellano, G.; Villalba, S.D.; Camurri, A. Recognising human emotions from body movement and gesture dynamics. In *Affective Computing and Intelligent Interaction*; D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C., Eds.; Springer: Berlin, Germany, 2007; pp. 71–82.

17. Wallbott, H.G. Bodily expression of emotion. *Eur. J. Soc. Psychol.* **1998**, *28*, 879–896.

18. Darwin, C. *The Expression of the Emotions in Man and Animals*; Oxford University Press: Oxford, UK, 1998.

19. OpenNI Consortium. OpenNI, the Standard Framework for 3D Sensing. Available online: http://openni.ru (accessed on 15 August 2017).