SUGISAKA MASANORI

Journal of Robotics, Networking and Artificial Life
Vol. 9(2); September (2022), pp. 159–170
ISSN (Online):2352-6386; ISSN (Print):2405-9021
https://alife-robotics.org/jrnal.html

Research Article

Investigation of Real-time Emotional Data Collection of Human Gaits using Smart Glasses

Nitchan Jianwattanapaisarn, Kaoru Sumi
School of Systems Information Science, Future University Hakodate, 116-2 Kamedanakano-cho, Hakodate, Hokkaido 041-8655, Japan

ARTICLE INFO

Article History

Received 12 May 2022
Accepted 30 June 2022

Keywords

Emotion recognition
Emotion induction
Motion capturing
Gait analysis
Smart Glasses
Non-straight walking

ABSTRACT

Emotion recognition is useful in many applications. Most methods available nowadays depend on facial features which are difficult to obtain from standard security cameras. Unlike traditional biometrics, gaits can be obtained noninvasively from afar. A novel method to perform gait data collection with real-time emotion induction was proposed. OptiTrack was used to capture 49 participants walking circularly while watching emotion-induced videos on Microsoft HoloLens 2. This is the first study that use real-time emotion induction technique in non-straight walking path..

1. Introduction

Emotion recognition and analysis have recently become very popular due to their useful applications in several scenarios such as improving the quality of human-robot interaction, evaluating customer satisfaction, detecting suspicious behaviors for crime prevention, assessing student engagement during online classes, and so forth. It can be used to prevent conflicts among groups of people by detecting when one or more people in a stressful conversation become upset. Appropriate personnel can respond and separate the person from the conversation. Emotional recognition and analysis can also be used in commercials, such as determining whether an advertisement elicits a positive emotion in people passing by. Many of these applications would improve the quality of life in modern days.

Because of the popularity of human emotion analysis research, there is a specific research field called *Affective Computing* [1] that aims to teach computers to understand and generate human-like affects that can be applied to a variety of applications. Many affective computing applications have been proposed in recent years. Some are extremely useful in current situation, particularly for educational purpose. As most students are unable to attend on-site classes due to the coronavirus pandemic. Students who would like to practice their programming skills at home cannot access to face-to-face guidance from lecturers and thus cannot improve their skills as effectively as it should be in normal situation. By incorporating affective computing techniques into the online exercise program, the program can evaluate how the students perform in each assignment. If students do not perform well in any assignment, the cause might be they are feeling exhausted or unmotivated; the program can advise students to focus on their weak points in that lesson, and the program can interact with students using animated agents to make their emotions getting better so they can continue studying effectively and happily [2].

Corresponding author's E-mail: g3219002@fun.ac.jp, kaoru.sumi@acm.org www.fun.ac.jp

Emotion prediction using human observers is a time-consuming task and is not accurate enough to be used in reality. Many techniques for automatic emotion prediction have been proposed. Nevertheless, most of publicly available methods presently use facial expressions as input for analysis and prediction. Facial features perform very well in some situations but have some limitations, such as being difficult to obtain in crowded conditions and by normal security cameras. Furthermore, some subjects do not show intense emotions on their faces. Due to these limitations, facial feature-based emotional detection techniques work accurately in certain situations, such as when the subjects are facing forward and close to the camera. If face images cannot be clearly captured, other features are required to make emotion recognition and analysis applications more practical and ready for real-world uses.

Human gaits and postures are the forms in which the human body moves and poses while walking or performing other activities. Gaits and postures can be observed from afar without the need for high-resolution photos or videos. Recording of gait and posture features does not tamper with the subjects' daily lives. These data can even be gathered without the subjects' awareness. Gait and posture recognition have been used successfully in a wide range of applications with high accuracies including human identification [3,4] human re-identification [5], human age estimation and gender recognition [6,7]. As a result, as clearly proven by many previous research, human gait and posture are appropriate features for detecting human emotions [8,9,10,11,12,13,14,15,16,17,18].

In this study, we proposed a method and environment for performing gait data collection using Microsoft HoloLens 2 to display the emotion-induced videos. Gait data were captured using OptiTrack motion capturing system. However, any other motion capture devices can be also used instead of OptiTrack, for instance, Vicon, Microsoft Kinect, Intel RealSense, or standard video cameras with pose-estimation software such as OpenPose.

2. Related Works

There are several studies about emotion recognition proposed during recent years due to the field popularity and usefulness. However, most of these works are based on facial expression. These studies provide accurate results for some applications but still have limitations in other real-world usages as mentioned in the introduction. There are fewer works which focused on emotion recognition using gaits and postures.

We found a survey that investigated many studies about gait analysis [19]. They found that gait analysis can be used not only for identification of subjects but also for subjects' current emotion prediction. They discovered that human walking in different emotions have different characteristics. By using these information, automatic emotion recognition can be performed. There are several advantages of using gaits compared with traditional biometrics such as facial features, speech features, physiological features etc. Gaits can be observed from far away without subjects' awareness. Gaits are difficult to be imitated. Gaits can be obtained without subjects' awareness. Because of these advantages, gaits are very effective expressions that can be used for automatic emotion recognition. Many types of devices can be used for gait data recording. For example, force plate can be used for recording velocity and pressure [8]. Infrared light barrier system also performs well for recording velocity data [8,20]. Motion capturing system e.g., Vicon can capture the coordinate data accurately using the markers attached on the body [9,10,11,12,21,22,23]. Microsoft Kinect is also another efficient tool that can capture human skeleton by processing the depth image with color image to predict the position of body joints [3,4,5,6,13,14,24]. Accelerometer sensor on wearable devices such as smartphones or smart watches can also record the movement data for gait analysis [15,16,18]. After gait data collection, there are several preprocessing steps that can be used. For instance, low pass Butterworth filter [23,25,26], sliding window Gaussian filtering [13,14]. Data transformation from time domain to others such as Discrete Fourier Transform [13,14,24] or Discrete Wavelet Transform are also widely used [27,28,29]. Gait features are categorized into Spatiotemporal Features such as stride length, velocity, step width, step length, and Kinematic Features such as coordinates data, joint angles, angular range of motion etc. Dimension reductions are also used such as Principal Component Analysis [30,31,32,33,34]. Finally, emotion recognition phase can be performed using many popular techniques, e.g., Multilayer Perceptrons [8], Naive Bayes [12,13,14], Nearest Neighbors [12,35], Support Vector Machine [12, 13,14,15,18], Decision Tree [15,18,35] etc. There are some interesting results from many studies they surveyed. For happiness, the subject step faster [17], strides are longer [36], arm movement increases [36], joint angle amplitude increases [9]. For sadness, the arm swing decreases [17], torso shape and limb shape are contracted [22], joint angles amplitudes reduce [9].

There are many gait analysis studies proposed in recent decades. Several applications can be achieved by analyzing human gaits. The following are some examples; human identification or re-identification [3,4, 5], gender prediction [6,7], emotion prediction [8,19], mental illness prediction [20,21] etc. There are several aforementioned methods to collect gait data such as using force plate, light barrier, motion capturing system, video camera, accelerometer, and so forth. We focus only the methods that extract 3-dimensional coordinates, binary silhouette, and body parts angles as gait features since these gait features are sensitive to walking pattern. Most studies proposed nowadays used straight walking path in their experiments to achieve high quality gait data [8,9, 10,11,13,14,18,20,21,24,25,26,33]. However, fewer studies used free-style walking path that the subjects can choose any walking pattern as they wanted instead of straight walking [3,4,5,6]. By developing methods for free-style walking data, opportunities for the proposed methods to be implemented in a real-world scenario which humans are walking without awareness of being observed in public spaces are increased because obtaining adequate straight walking data in a noisy environment is more difficult than obtaining free-style walking data.

As we used Microsoft HoloLens 2 smart glasses to show the videos to each participant while he or she is walking, another issue needs to be concerned is that whether the gaits could be interfered by watching contents on smart glasses or not. Risk of adverse effect such as slips and trips are matter. There are some studies about this issue. For example, [37] investigated the performance of gaits while using head-worn display while walking. They conducted experiments using 12 participants to test that the subjects can walk effectively in multiple conditions while using head-worn display or not. They assessed the minimum foot clearance, required coefficient of friction, foot placement location around the obstacle, walking speed and obstacle crossing speed. They found that using head-worn display did not affect with level walking performance compared to using a paper list and baseline walking which used nothing. In obstacle crossing tasks, more conservative and more cautious strategy were selected when the subjects used the head-worn display, and the obstacle crossing speed decreased by 3% compared to the baseline. However, location of foot placement around the obstacle was not affected when using head-worn display. [38] and [39] also conducted experiments to examine the adverse impacts on human gaits when the subjects use head-worn display e.g., smart glasses and walk at the same time. In their experiments, 20 participants including 10 males and 10 females walked on treadmill in 4 different conditions at their preferred walking speed. They asked the subjects to perform one single-task walking and three dual-task walking using different equipment to display the information to the subjects while walking. In dual-task walking, participants performed attention-demanded tasks on different display types. Attention-demanded tasks include Stroop test, categorizing, and arithmetic. Display types used in their experiments are paper-based, smartphones and smart glasses. When the subjects used paper-based display and smartphone, they used the head-down posture whereas they used the head-up posture for single task walking and for smart glasses display. Motion capturing was done using Vicon motion capture system with 7 cameras. They found that using smart glasses to perform tasks while walking can have more impacts on gait performance such as stability when compared with using other types of display. Also, they found that smart glasses can disturb the control of gait variables. However, the participants were more unstable when using smartphone and paper-based system compared to when using smart glasses.

According to these related studies, we decided to use Microsoft HoloLens 2 for displaying emotional videos to our participants while they are walking in the recording area even though there are possibilities that some adverse effects can be occurred such as walking stability when using smart glasses during walking, we coped with this issue by asking the participants to take one rehearsal walk in the walking area without wearing HoloLens 2 to make them familiar with the walking space and another rehearsal walk while wearing HoloLens 2 that displayed nothing to make them familiar with walking while wearing smart glasses at the same time.

For walking pattern, straight walking should result in cleaner gait data but has more limitations when implementing in real-world scenarios. On the other hand, walking freely without any path guidance could be too difficult for the subjects. Since they have to concentrate on the videos content showing on HoloLens 2 while walking, if they also need to determine the walking path at the same time, they cannot focus well on videos content and their gaits can be interfered. Therefore, we decided to use the lax circular walking path for our experiments. By walking circularly in clockwise or counter-clockwise direction without marking the path line on the floor, we can have both straight walking and non-straight walking data in one walking trial.

3. Data Collection

In most previous studies in the field of emotional recognition and analysis, participants were asked to walk in a straight line after watching emotional movies or asked to walk in a straight line while thinking about personal experiences. There are some issues with these settings. In cases where participants were asked to walk after watching emotional videos, it is possible that some participants do not feel the intended emotions toward the end of the walk, or some participants do not have the intended emotions at all after watching the videos. These can lead to inaccurate relationship between gaits and emotions. In case where participants were asked to feel certain ways using their personal experiences, it is also possible that some participants cannot recall their feelings well enough to reflect on their body movements. These causes can lead to faulty information.

To eliminate those issues that lead to faulty information and inaccurate relationship between gaits and emotions, our experiments are designed so that participants are constantly incited with emotion-induced videos while walking. We used the latest smart glasses technology named Microsoft HoloLens 2 for showing videos to the subjects while they are walking. To the best of our knowledge, currently available studies have never used this kind of emotion induction method before. By using HoloLens 2 for showing videos, subjects can see the room environment and the videos at the same time. Because we show emotional videos to participants while walking, the results are closer to real-life situations when a subject seeing some events and feels some emotions because of those events. In other words, we attempted to simulate the real-time emotions of the participants by showing emotion-induced videos while they are walking. Moreover, intensity of induced emotions should be more consistent than conventional method that shows emotional videos to subjects before walking. Our proposed emotional gait data collection technique should be useful for advancement of emotion recognition research field.

3.1. *Equipment for Data Collection*

Currently, there are two main types of motion capturing equipment: *marker-less* and *marker-based* devices. Marker-less devices are more convenient to use in real-life situation because there is no need to attach any equipment to subject's body. Coordinates of body parts are calculated by image processing technology using depth data recorded by infrared camera together with RGB images from color camera. For marker-based type, several markers have to be attached to subject's body at the desired positions such as on the head, hand, elbow, and so forth. Marker-based device is more complex to setup because it requires several cameras to capture the infrared reflection from the markers attached to subject's body for reconstruction of markers coordinates in 3-dimensional space. However, body tracking accuracy of marker-less system is lower than marker-based type since marker-less system predicts the position of each body part while marker-based type uses the actual position obtained from several cameras.

In this study, OptiTrack which is a famous marker-based motion capturing system were used for our data collection. We used the baseline marker set with 37 markers which is the standard configuration for human skeleton tracking. With baseline marker set configuration, 37 markers were attached to each subject's body. The names of markers are listed in Table 1, and the position of markers are shown in Figure 1. Sample image of OptiTrack installation is shown in Figure 2.

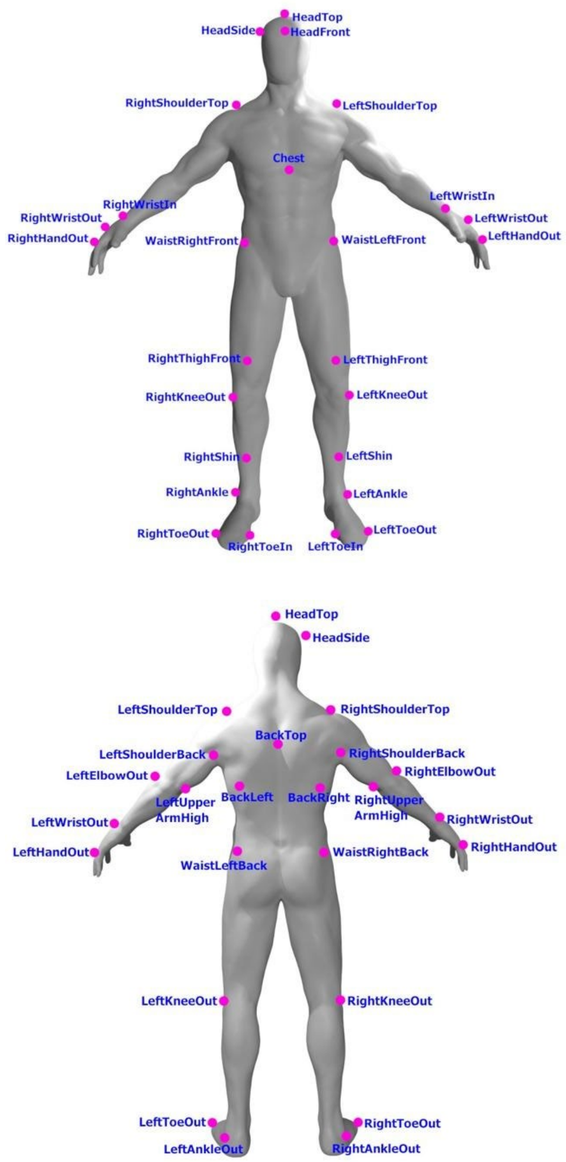

Fig. 1. Position of Front and Back Markers (Original Human Figure Source: dog012 on sketchfab.com *)

Fig. 2. Two OptiTrack Flex 3 cameras were installed on one camera stand at different height levels

3.2. *Recording Environment*

We marked a rectangle on the floor to use as the walking area that can be captured by OptiTrack motion tracking system using the black tape as shown in Figure 3. Fourteen OptiTrack motion capture cameras were

Fig. 3. Rectangle Walking Area marked with Black Tape on the Floor

Table 1. List of OptiTrack Baseline Markers

HeadTop	
HeadFront	
HeadSide	
BackTop	
Chest	
Back	Left
	Right
WaistFront	Left
	Right
WaistBack	Left
	Right
ShoulderBack	Left
	Right
ShoulderTop	Left
	Right
ElbowOut	Left
	Right
UpperArmHigh	Left
	Right
WristOut	Left
	Right
WristIn	Left
	Right
HandOut	Left
	Right
ThightFront	Left
	Right
KneeOut	Left
	Right
Shin	Left
	Right
AnkleOut	Left
	Right
ToeOut	Left
	Right
ToeIn	Left
	Right

installed on 7 camera stands. That is, two cameras on each stand at different heights as shown in Figure 2. Seven camera stands were placed around the walking

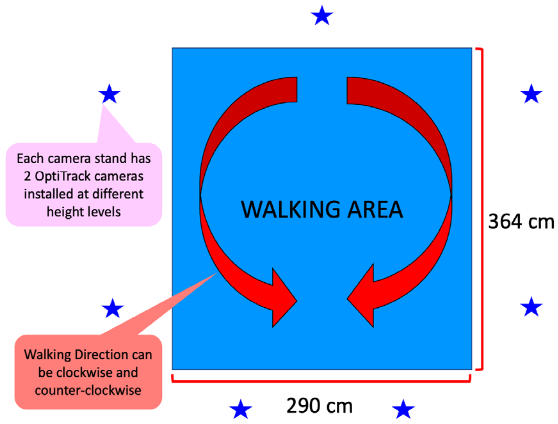

Fig. 4. Size of Walking Area and Position of Recording Equipment (OptiTrack)

area as illustrated in Figure 4. The size of the walking area is 2.9 meters by 3.64 meters.

3.3. *Materials for Data Collection*

We selected 3 videos as stimuli for emotion induction. These videos were shown to the subjects using HoloLens 2 while each subject is walking in the recording area.

- ***Neutral Video:*** The nature landscape video from YouTube named *Spectacular drone shots of Iowa corn fields* uploaded by the YouTube user named *The American Bazaar* *
- ***Negative Video:*** An emotional movie selected from LIRIS-ACCEDE database named *Parafundit* by *Riccardo Melato*
- ***Positive Video:*** An emotional movie selected from LIRIS-ACCEDE database named *Tears of steel* by *Ian Hubert* and *Ton Roosendaal*

Neutral video was selected from nature landscape videos on YouTube that does not induce any emotion. Positive videos and negative videos will be used for induction of happy emotion and sad emotion respectively. Negative and positive videos were selected from the public annotated movies database LIRIS-ACCEDE † published by [40].

* https://www.youtube.com/watch?v=4R9HpESkor8, Last Accessed: May 12, 2022

† https://liris-accede.ec-lyon.fr/, Last Accessed: April 6, 2022

The LIRIS-ACCEDE database contains many creative commons movies and their emotional annotations. There are several movie collections with different properties published in LIRIS-ACCEDE database such as Discrete LIRIS-ACCEDE, Continuous LIRIS-ACCEDE, MediaEval 2015 Affective Impact of Movies, and so forth. In our study, we used the Continuous LIRIS-ACCEDE collection that contains 30 movies and emotion annotations in Valence-Arousal dimension. Most movies contain both positive and negative valence in the same movie. We carefully selected one movie that has only positive valence for entire movie and one movie that has only negative valence for entire movie to make our walking trial to contain only one emotion in each trial. Sample plots of valence annotations from the movies in Continuous LIRIS-ACCEDE collection are shown in Figure 5, and plots of our selected movies i.e., negative movie and positive movie are illustrated in Figure 6.

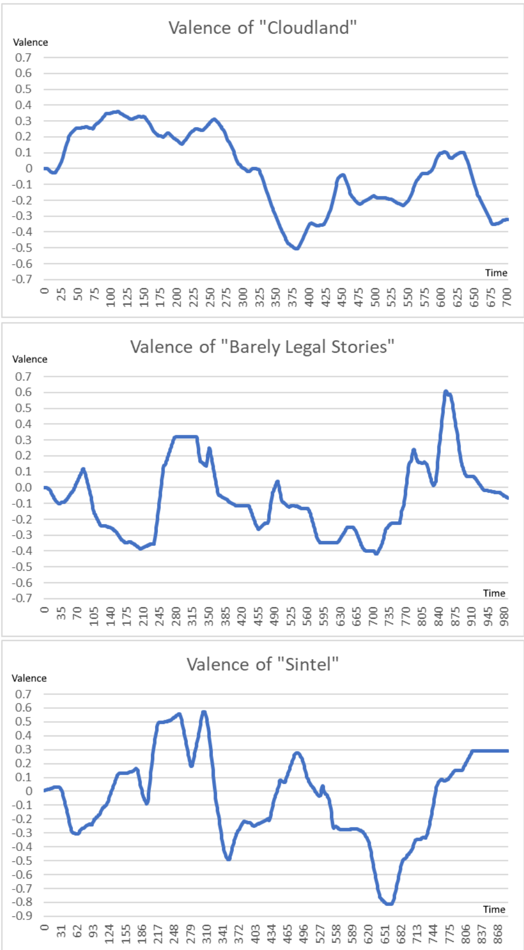

Fig. 5. Valence Annotation of 3 Sample Movies

We also concerned about the length of each video since all participants will walk and watch the videos at the same time, so all videos we used are not exceed 15 minutes length. Length of neutral video, negative video, and positive videos are 5:04, 13:10, and 12:14 minutes respectively. Audio of negative and positive videos contain music, sound effects, and conversations in English. Subjects can hear the sound from HoloLens 2 build-in stereo speakers when they walk. Neutral Video does not have any sound to ensure that it will not induce any emotion.

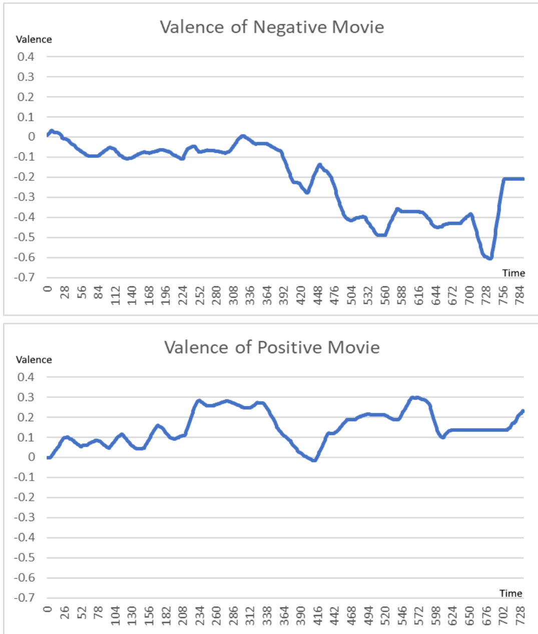

Fig. 6. Valence Annotation of Selected Negative Movie (Parafundit) and Positive Movie (Tears of steel)

3.4. *Methods for Data Collection*

Each participant was asked to answer the health questionnaire and signed the consent form in the beginning of the experiment. Questions in the health questionnaire are listed below.

1. Do you have any neurological or mental disorders?
2. Do you have a severe level of anxiety or depression?
3. Do you have hearing impairment that cannot be corrected?
4. Do you have any permanent disability or body injury that affects walking posture?
5. Do you currently feel sick now? (e.g., fever, headache, stomachache etc.)

6. If you have any problem with your health condition, please describe.

After each subject was confirmed to be healthy i.e., they could walk, watch, and listen normally, we instructed each participant to walk in circular pattern inside the marked rectangle space. Participants could choose the direction they want to walk between *clockwise* or *counter-clockwise*. During each walking trial, subjects could switch directions whenever they want. We asked each subject to walk in the recording area for 3 minutes without wearing HoloLens 2 to establish a subject's natural way of walking. The purpose of the first walking trial *(Rehearsal Walk)* is to make the subject feel familiar with the walking space. Next, each subject performed the second Rehearsal Walk by walking again for another 3 minutes and wearing HoloLens 2 that did not show any content to make them feel familiar with walking while wearing HoloLens 2. As we found in related studies, if the participants never have experiences using smart glasses while walking before, gait performance can be unstable [37,38,39]. We attempted to cope with this problem by asking the subjects to take the rehearsal walks with and without HoloLens 2 before performing actual recording.

Then, we showed *Neutral Video* on HoloLens 2 and asked each subject to walk and watch the video at the same time to capture *Neutral Walk*. Each subject started walking when the video started and stopped walking when the video ended. *Positive Walk* and *Negative Walk* were done using *Positive Video* and *Negative Video* from LIRIS-ACCEDE database as described in Section 3.3.

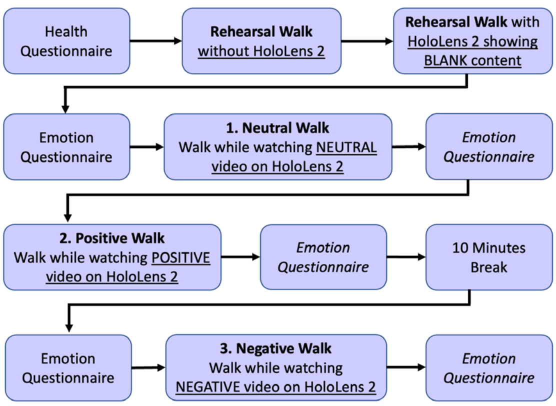

Fig. 7. Data Collection Process

These videos were shown on HoloLens 2 while each subject was walking using the same procedure with Neutral Walk. Additionally, after finishing Positive Walk, we asked the subject to go take 10 minutes break to reset their emotion back to normal before starting Negative Walk. The order of Negative Walk and Positive Walk was swapped for the next subject. That is, swapping between Neutral, Positive, Negative and Neutral, Negative, Positive. The entire process for data collection is shown in Figure 7. Note that we also asked each subject to answer the self-reported emotion questionnaire before and after walking for each video. The questions are as follows.

- Please choose your current feeling: Happy, Sad, Neither (Not Sad and Not Happy)
- How intense of your feeling: 1 (Very Little) to 5 (Very Much)

Fig. 8. Sample of Walking Subject while Watching Video on HoloLens 2

Sample images of a subject walking in circular pattern in the recording area while watching an emotion-induced video on HoloLens 2 is shown in Figure 8, and a photo of a subject wearing the OptiTrack Motion Capture Suit with 37 markers and HoloLens 2 is shown in Figure 9.

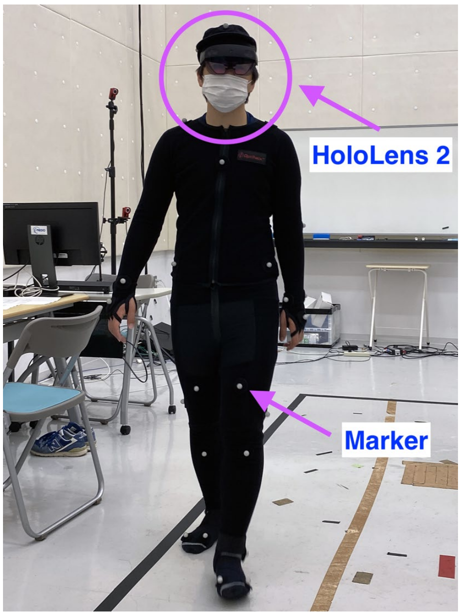

Fig. 9. A Subject Wearing HoloLens 2 and OptiTrack Motion Capture Suit with 37 Markers

4. Results and Discussion

In this study, an emotional gait dataset has been collected. Each subject walked in non-straight walking path and watched 3 videos including Neutral Video, Negative Video and Positive Video during walking. In summary, we have total 147 walking trials in this dataset.

An overview of our collected dataset is as follow.

- *Number of participants:* 49 subjects (41 male, 8 female)
- *Average age:* 19.69 years
- *Standard deviation of age:* 1.40 years
- *Average height:* 168.49 centimeters
- *Standard deviation of height:* 6.34 centimeters
- *Average weight:* 58.88 kilograms
- *Standard deviation of weight:* 10.84 kilograms

As we swapped the order of videos shown to the subjects, the number of subjects watching negative video before positive video, and subjects watching positive video before negative video are almost balance as listed.

- Neutral Video → Negative Video → Positive Video: ***24 subjects***
- Neutral Video → Positive Video → Negative Video: ***25 subjects***

According to the answers from self-reported emotion questionnaire after finished walking and watching each video, we have **44 sad** walking trials, **44 happy** walking trial, and **59 neither** walking trials. Table 2 shows the numbers of subjects who felt *happy, sad,* and *neither* from the self-reported emotion questionnaire for each video stimulus.

We analyzed each video stimulus including negative video, neutral video and positive video with its resulting reported emotion including happy, neither and sad. The comparison between expected emotion and reported emotion is illustrated in Figure 10. According to Figure 10 and Table 2, the resulting emotion from each stimulus are as follows.

- ***Negative Video:*** 13 subjects feel happy, 17 subjects feel neither, 19 subjects feel sad
- ***Neutral Video:*** 19 subjects feel happy, 28 subjects feel neither, 2 subjects feel sad
- ***Positive Video:*** 12 subjects feel happy, 14 subjects feel neither, 23 subjects feel sad

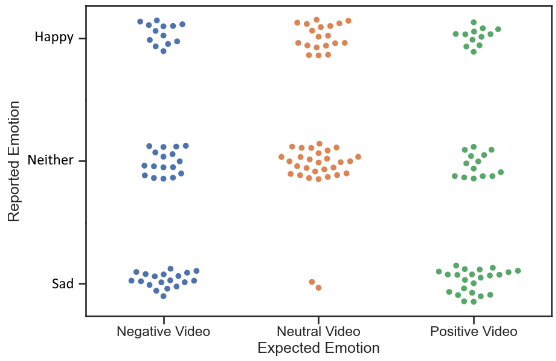

Fig. 10. Comparison of Subjects' Feelings between Reported Emotion and Expected Emotion

From these results, not all subjects feel sad after watching negative video, and not all subjects feel happy after watching positive video. For neutral video, it should not elicit any emotion so the reported emotion after watching neutral video might come from other factors such as trying HoloLens 2 for the first time can make some subjects feel happy.

We can imply from these results that, reported emotion of each subject is not always similar to the annotated emotion from video stimulus we showed. This could happen because of many factors. For example, it is possible that some subjects have different sensitivity for feeling sad. In other word, seeing some stories can make some subjects feel very sad while some other subjects can feel little sad, neither sad nor happy, or even very happy because different people have different emotion perception. This reason also applied to positive video stimulus, even though the video was annotated as positive emotion, some subjects feel happy but other subjects can also feel the opposite emotion as there could be some components or some stories in the positive movie that trigger them to feel sad instead of happy as we expected them to feel. Last but not least, individual preferences can also have effects on reported emotion after watching video stimulus. For instance, some subjects who do not like action movies or animation movies can feel sad after watching positive stimulus if that stimulus is the kind of movies they do not like. Also, the music soundtracks or conversations in the movie can affect with the subjects' feeling.

By performing real-time method for emotion induction using smart glasses i.e., Microsoft HoloLens 2 to display emotion induction videos to the participants while they are walking, we found that this type of emotion induction method is very interesting since it can simulate the situation when the subjects saw some stories happened so their gaits and postures are changed unconsciously. As we asked all subjects to answer the self-reported questionnaire after finished walking, we can ensure that this way of emotion induction can elicit subjects' emotion more consistent than conventional method that shows the video stimuli to the participants before walking.

Especially, since the reported emotion which is the actual emotion that the subjects feel could be different from the emotion we expected the participants to feel i.e., annotated emotion of the videos, it is very important that we must ask the subject how they feel after watching our video stimuli and compare their reported emotion with our expected emotion. If the reported emotion is different from the expected emotion, we should use the reported emotion to label that walking trial instead.

5. Conclusion

In summary, we proposed a novel emotional gait data collection method that uses the different technique to induce human emotion while walking. Unlike conventional emotion induction method which shows the emotional video stimuli to the subjects before walking on a conventional display such as a television or a computer monitor, our method utilized the latest smart glasses technology named Microsoft HoloLens 2 to display the emotion induction videos to the participants while they are walking. By showing the stimuli on HoloLens 2, subjects can walk and watch the videos at the same time. This makes emotion induction more realistic as the real-time emotion perception is simulated in our proposed method. Additionally, we used the non-straight walking path for the subjects to walk while watching the video stimuli. Non-straight walking path could result in noisier gait data than straight walking path, but it gives us more real walking posture as in the real-world scenarios. Moreover, annotated emotion of the stimuli could be much different from actual emotion reported by the subjects, it is highly recommended to always ask the subjects to report their actual feelings after walking. Finally, in this study, even we used OptiTrack motion capturing system to capture the gait data in 3-dimensional coordinates, it is not mandatory to use the marker-based system for gait data recording. Marker-less system such as Microsoft Kinect or standard video camera with pose-estimation software can be also used to record human movement data for gait analysis.

Acknowledgements

We would like to express our gratitude to all volunteers who participated in our experiments. Also, we would like to thank all members of Kaoru Sumi Laboratory in Future University Hakodate who helped us a lot in our experiments such as setting up the venue and recording equipment, translation of written instructions and questions for Japanese participants, and interpretation between Japanese and English when communicating with participants who preferred to talk in their native language.

References

1. Rosalind W Picard. *Affective computing*. MIT press, 2000.
2. Thomas James Tiam-Lee and Kaoru Sumi. Analysis and prediction of student emotions while doing programming exercises. In *International conference on intelligent tutoring systems*, pages 24–33. Springer, 2019.
3. Nirattaya Khamsemanan, Cholwich Nattee, and Nitchan Jianwattanapaisarn. Human identification from freestyle

walks using posture-based gait feature. *IEEE Transactions on Information Forensics and Security*, 13(1):119–128, 2017.

4. Piya Limcharoen, Nirattaya Khamsemanan, and Cholwich Nattee. View-independent gait recognition using joint replacement coordinates (jrcs) and convolutional neural network. *IEEE Transactions on Information Forensics and Security*, 15:3430–3442, 2020.
5. Piya Limcharoen, Nirattaya Khamsemanan, and Cholwich Nattee. Gait recognition and re-identification based on regional lstm for 2-second walks. *IEEE Access*, 9:112057–112068, 2021.
6. Kotcharat Kitchat, Nirattaya Khamsemanan, and Cholwich Nattee. Gender classification from gait silhouette using observation angle-based geis. In 2019 *IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, pages 485–490. IEEE, 2019.
7. Ebenezer RHP Isaac, Susan Elias, Srinivasan Rajagopalan, and KS Easwarakumar. Multiview gait-based gender classification through pose-based voting. *Pattern Recognition Letters*, 126:41–50, 2019.
8. Daniel Janssen, Wolfgang I Schöllhorn, Jessica Lubienetzki, Karina Fölling, Henrike Kokenge, and Keith Davids. Recognition of emotions in gait patterns by means of artificial neural nets. *Journal of Nonverbal Behavior*, 32(2):79–92, 2008.
9. Claire L Roether, Lars Omlor, Andrea Christensen, and Martin A Giese. Critical features for the perception of emotion from gait. *Journal of vision*, 9(6):15–15, 2009.
10. Avi Barliya, Lars Omlor, Martin A Giese, Alain Berthoz, and Tamar Flash. Expression of emotion in the kinematics of locomotion. *Experimental brain research*, 225(2):159–176, 2013.
11. Gentiane Venture, Hideki Kadone, Tianxiang Zhang, Julie Grèzes, Alain Berthoz, and Halim Hicheur. Recognizing emotions conveyed by human gait. *International Journal of Social Robotics*, 6(4):621–632, 2014.
12. Michelle Karg, Kolja K¨uhnlenz, and Martin Buss. Recognition of affect based on gait patterns. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(4):1050–1061, 2010.
13. Baobin Li, Changye Zhu, Shun Li, and Tingshao Zhu. Identifying emotions from non-contact gaits information based on microsoft kinects. *IEEE Transactions on Affective Computing*, 9(4):585–591, 2016.
14. Shun Li, Liqing Cui, Changye Zhu, Baobin Li, Nan Zhao, and Tingshao Zhu. Emotion recognition using kinect motion capture data of human gaits. *PeerJ*, 4:e2364, 2016.
15. Zhan Zhang, Yufei Song, Liqing Cui, Xiaoqian Liu, and Tingshao Zhu. Emotion recognition based on customized smart bracelet with built-in accelerometer. *PeerJ*, 4:e2258, 2016.
16. Juan Carlos Quiroz, Elena Geangu, and Min Hooi Yong. Emotion recognition using smart watch sensor data: Mixed-design study. *JMIR mental health*, 5(3):e10153, 2018.
17. 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.
18. Mangtik Chiu, Jiayu Shu, and Pan Hui. Emotion recognition through gait on mobile devices. In 2018 *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 800–805. IEEE, 2018.
19. Shihao Xu, Jing Fang, Xiping Hu, Edith Ngai, Yi Guo, Victor Leung, Jun Cheng, and Bin Hu. Emotion recognition from gait analyses: Current research and future directions. *arXiv preprint arXiv:2003.11461*, 2020.
20. Matthias R Lemke, Thomas Wendorff, Brigitt Mieth, Katharina Buhl, and Martin Linnemann. Spatiotemporal gait patterns during over ground locomotion in major depression compared with healthy controls. *Journal of psychiatric research*, 34(4-5):277–283, 2000.
21. Johannes Michalak, Nikolaus F Troje, Julia Fischer, Patrick Vollmar, Thomas Heidenreich, and Dietmar Schulte. Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosomatic medicine*, 71(5):580–587, 2009.
22. M Melissa Gross, Elizabeth A Crane, and Barbara L Fredrickson. Effort-shape and kinematic assessment of bodily expression of emotion during gait. *Human movement science*, 31(1):202–221, 2012.
23. Matthieu Destephe, Takayuki Maruyama, Massimiliano Zecca, Kenji Hashimoto, and Atsuo Takanishi. The influences of emotional intensity for happiness and sadness on walking. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7452–7455. IEEE, 2013.
24. Bingli Sun, Zhan Zhang, Xingyun Liu, Bin Hu, and Tingshao Zhu. Self-esteem recognition based on gait pattern using kinect. *Gait & posture*, 58:428–432, 2017.
25. Gu Eon Kang and M Melissa Gross. Emotional influences on sit-to-walk in healthy young adults. *Human movement science*, 40:341–351, 2015.
26. Gu Eon Kang and M Melissa Gross. The effect of emotion on movement smoothness during gait in healthy young adults. *Journal of biomechanics*, 49(16):4022–4027, 2016.
27. MN Nyan, FEH Tay, KHW Seah, and YY Sitoh. Classification of gait patterns in the time–frequency domain. *Journal of biomechanics*, 39(14):2647–2656, 2006.
28. Adham R Ismail and Shihab S Asfour. Discrete wavelet transform: a tool in smoothing kinematic data. *Journal of biomechanics*, 32(3):317–321, 1999.
29. E Baratin, L Sugavaneswaran, K Umapathy, C Ioana, and S Krishnan. Wavelet-based characterization of gait signal for neurological abnormalities. *Gait & posture*, 41(2):634–639, 2015.
30. ME Wootten, MP Kadaba, and GVB Cochran. Dynamic electromyography. i. numerical representation using

principal component analysis. *Journal of Orthopaedic Research*, 8(2):247–258, 1990.
31. R Shiavi and P Griffin. Representing and clustering electromyographic gait patterns with multivariate techniques. *Medical and Biological Engineering and Computing*, 19(5):605–611, 1981.
32. Sandra J Olney, Malcolm P Griffin, and Ian D McBride. Multivariate examination of data from gait analysis of persons with stroke. *Physical Therapy*, 78(8):814–828, 1998.
33. Heydar Sadeghi, Paul Allard, and Morris Duhaime. Functional gait asymmetry in able-bodied subjects. *Human Movement Science*, 16(2-3):243–258, 1997.
34. Kevin J Deluzio, Urs P Wyss, Benny Zee, Patrick A Costigan, and Charles Serbie. Principal component models of knee kinematics and kinetics: normal vs. pathological gait patterns. *Human Movement Science*, 16(2-3):201–217, 1997.
35. Ferdous Ahmed, Brandon Sieu, and Marina L Gavrilova. Score and rank-level fusion for emotion recognition using genetic algorithm. In *2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, pages 46–53. IEEE, 2018.
36. Shaun Halovic and Christian Kroos. Not all is noticed: Kinematic cues of emotion-specific gait. *Human movement science*, 57:478–488, 2018.
37. Sunwook Kim, Maury A Nussbaum, and Sophia Ulman. Impacts of using a head-worn display on gait performance during level walking and obstacle crossing. *Journal of Electromyography and Kinesiology*, 39:142–148, 2018.
38. Alireza Sedighi, Sophia M Ulman, and Maury A Nussbaum. Information presentation through a head-worn display ("smart glasses") has a smaller influence on the temporal structure of gait variability during dual-task gait compared to handheld displays (paper-based system and smartphone). *PLoS one*, 13(4):e0195106, 2018.
39. Alireza Sedighi, Elaheh Rashedi, and Maury A Nussbaum. A head-worn display ("smart glasses") has adverse impacts on the dynamics of lateral position control during gait. *Gait & posture*, 81:126–130, 2020.
40. Yoann Baveye, Emmanuel Dellandr´ea, Christel Chamaret, and Liming Chen. Deep learning vs. kernel methods: Performance for emotion prediction in videos. In *2015 international conference on affective computing and intelligent interaction (acii)*, pages 77–83. IEEE, 2015.

Authors Introduction

Mr. Nitchan Jianwattanapaisarn

Nitchan is currently a doctoral student at Future University Hakodate, Japan. He received his M.Eng. and B.Sc. from Sirindhorn International Institute of Technology, Thammasat University, Thailand. He previously worked as an internship student at Interaction Science Laboratories, Advanced Telecommunications Research Institute International (ATR), Japan. His research interests include affective computing, human-computer interaction, and applications of machine learning and artificial intelligence.

Prof. Kaoru Sumi

Kaoru Sumi is a professor at Future University Hakodate, Japan. She received her Ph.D. in engineering from the University of Tokyo. She previously worked at ATR MI&C Research Laboratories, Communications Research Laboratory (CRL), Osaka University, National Institute of Information and Communications Technology (NICT), and Hitotsubashi University. She is currently working on affective computing, human agent interaction, persuasive technology, interactive digital storytelling, and serious games.