Towards Modeling of Interpersonal Proximity Using Head-Mounted Camera for Children with ASD

Airi Tsuji[1](✉), Satoru Sekine[2], Soichiro Matsuda[3], Junichi Yamamoto[4], and Kenji Suzuki[3]

[1] Tokyo University of Agriculture and Technology, 2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
atsuji@go.tuat.ac.jp

[2] The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
ssekine@keio.jp

[3] University of Tsukuba, Tennodai, Tsukuba, Ibaraki 305-8573, Japan
matsuda@ai.iit.tsukuba.ac.jp, kenji@ieee.org

[4] Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
yamamotj@flet.keio.ac.jp

Abstract. This study suggests a novel method to clarify interpersonal proximity from the area occupied by a target person in the visual field. We use the size of the area as an index of interpersonal proximity to support therapeutic activities for the developmental support for children with ASD. We investigate the relationship between physical and interpersonal distance and its proximity as a sense of distance. Head-mounted camera and Mask-RCNN used as the measurement. The pilot experiment shows the possibility of measurement and experimentation in actual therapy showed a difference between interpersonal distance and interpersonal proximity.

Keywords: Chidren with ASD · Human recognition · Interpersonal proximity · Head-mounted camera · Machine learning

1 Introduction

Interpersonal proximity (a sense of distance) is an element of nonverbal communication and indicates aan dequate physical and social distance from communication targets. The concept of interpersonal distance as the physical distance was proposed by Hall [1] which is measured via the stop-distance and observation methods. It is widely used as one of evaluation of communications in psychology and sociology. To maintain a reasonable degree of distance from others requires spatial awareness that is regarded as the link between understanding proximity to others (interpersonal proximity) and social contexts. The concept is related to a sense of direction and spatial cognition [2]. However, it is known that the physical distance between two parsons is related to interpersonal proximity, but

K. Miesenberger et al. (Eds.): ICCHP 2020, LNCS 12377, pp. 104–111, 2020.
https://doi.org/10.1007/978-3-030-58805-2_13

these two factors are not the same. The interpersonal proximity is considered even in verbal communication and does not necessarily depend only on physical distance.

Children with autism spectrum disorder (ASD) experience difficulties in establishing social communication and interpersonal interaction with restricted interests and repetitive behaviors, which is defined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition published by the American Psychiatric Association [3]. Moreover, it has become clear that social communication skills can be improved by providing appropriate comprehensive developmental support for children with ASD [4].

In order to support the development of communication skills, it is important to assess the communication in terms of behavior and interaction in a quantitative manner. Previous work has been studied about physical and interpersonal distance that is measured as the target and used as an evaluation index. There are several therapeutic activities to help children to control the interpersonal distance to the appropriate degree in order to support the development of children with ASD. As an attempt to quantify the interpersonal proximity has been reported in the research on the interpersonal distance on VR space, which is originated from Bailenson [5]. It clarified that the interpersonal proximity is highly dependent upon visual perception. We consider that it is helpful for therapists or caregivers to observe to change the interpersonal proximity during sessions. We then try to explore a computational model to estimate the interpersonal proximity based on the measured interpersonal distance between two persons.

The main contribution in this study is to find a certain correlation between the image size of the target person in the image sequence obtained from first-person view and the physical interpersonal distance obtained by the motion capture system. The current study aimed to quantify interpersonal proximity to support the development of children with ASD. To this end, we aimed to identify a method via which to clarify interpersonal proximity from the area occupied by a target in the visual field of a first-person viewing camera.

2 Method

2.1 Interpersonal Proximity

Interpersonal distance refers to the physical distance between two persons and should be established to build interpersonal relationships. Hall divided personal space into close, individual, social, and public distance [1]. As the intimacy between the two persons increases, this distance decreases. Reducing distance, rather than intimacy, could cause others to experience strong discomfort. Children with ASD are often closer to or further away from each other relative to typically developing children. However, it is difficult to adjust their interpersonal distance or create appropriate distance from them [6,7].

The concept of interpersonal distance also exists in VR space involving the manipulation of interpersonal distance and behavior change [5]. The manipulation of interpersonal distance, using augmented reality in the real world, reduces

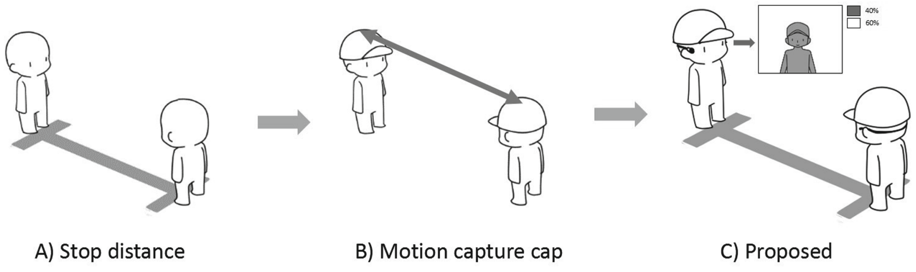

Fig. 1. Previous and proposed measurement methods

discomfort or pressure [8]. It is noted that interpersonal distance is perceived and recognized mainly via vision, because visual interpersonal distance manipulation, using a head-mounted display, affects recognition and behavior in the actual interpersonal distance.

However, it is not simply considered to be another representation of interpersonal distance but is evaluated subjectively based on more types of information rather than physical distance. The term of interpersonal proximity is also used in the remote exchange of text information. It may be possible to obtain different interpersonal proximity to two persons maintaining the same interpersonal distance. Interpersonal proximity can be acquired even in communication in a virtual space, using only visual information; therefore, it is possible to clarify interpersonal proximity using visual information. We assume that an index of interpersonal proximity can be described in relation to the peer within the field of view from the participant.

2.2 Proposed System

We proposed the use of the area around the person in a first-person view as a method for estimating interpersonal proximity between two people. The images from the first-person viewpoint is acquired from the head-mounted camera. Figure 1 shows an overview of previous methods and the proposed method compared to conventional methods.

The conventional method to measure the interpersonal distance in a quantitative manner is the stop-distance method shown in Fig. 1A [9]. This method requires two persons to approach one another until they feel discomfort or unpleasantness; the distance between their toe positions at the final point is then measured. Although this method enabled quantitative measurement of interpersonal distance, it was difficult to measure changes in interpersonal distance dynamically during activity.

Then, a method using the motion capture system [10], allows us to measure the interpersonal distance in several activities, as illustrated in Fig. 1B. It measures the distance between the heads of two persons. However, this was limited to physical distance measurement and did not allow consideration of the influence of the field of view.

On the other hand, in the proposed method as illustrated in Fig. 1C, the head-mounted camera is used to acquire the first-person view field and the area occupied by the target person. We consider that we can estimate the interpersonal proximity that reflects actual feelings most closely.

2.3 Measuring Human Area

Mask-RCNN [11] is an object-recognition method using deep learning. The proposed system used this method to calculate the area occupied by the target person in the image sequence from a first-person view. The outline of the process is shown in Fig. 2. In Fig. 2, X means the position of each person, D means the distance between two persons and S means the area occupied by the target person in each frame. The area around each object each the image is calculated by processing the first-person view images taken with the head-mounted camera, using Mask-RCNN. Learning models use Common Objects in Context Dataset [12]. The only area occupied by the target person (the largest human tagged by the mask) is used as the index of interpersonal proximity. In addition, interpersonal distance measurement in motion capture was performed, to compare it with the existing method.

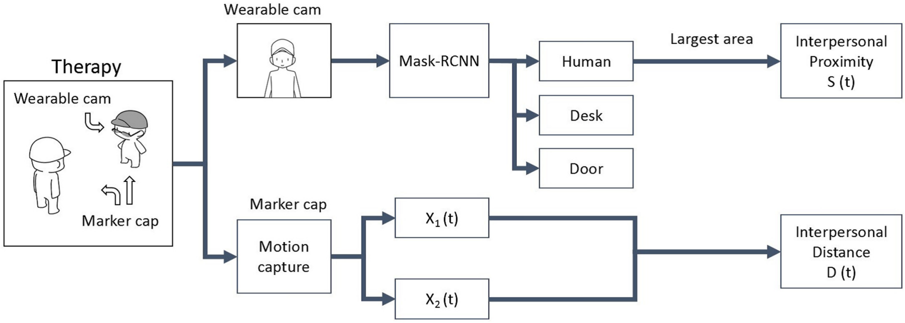

Fig. 2. Overview of system

3 Experiments

An experiment was conducted to verify the proposed method. In the experiment, the eye gaze and first-person images are acquired by using Tobii glass (Tobii Technology K.K.) [13] that is a glasses-type, head-mounted, first-person viewpoint camera with a small burden on the experiment participants. The motion capture marker cap used in a previous study was attached and interpersonal distance measurement was conducted at the same time. We will show the relationship between interpersonal distance and interpersonal proximity was examined using the proposed method. The experiment was conducted with two participants wearing the Tobii glass and marker caps. Participants were asked to perform the stop-distance method as a classical measurement of interpersonal distance. One participant was standing or sitting, and the other moved to approach him until discomfort was felt by one participant. The approaching participant stopped upon feeling discomfort or when the standing/sitting person raised a hand to indicate discomfort.

3.1 Result

Figure 3 shows the plots that represent interpersonal distance measured via motion capture and the area obtained by the proposed method. Blue lines represent the closest position. Table 1 shows the measurement results obtained via the stop-distance method. The interpersonal distance measured via motion capture and the frame obtained by the proposed method were acquired every second. Figure 4 shows the correlation between interpersonal distance and interpersonal proximity for the standing/sitting participants. Both correlation coefficients were higher than .50.

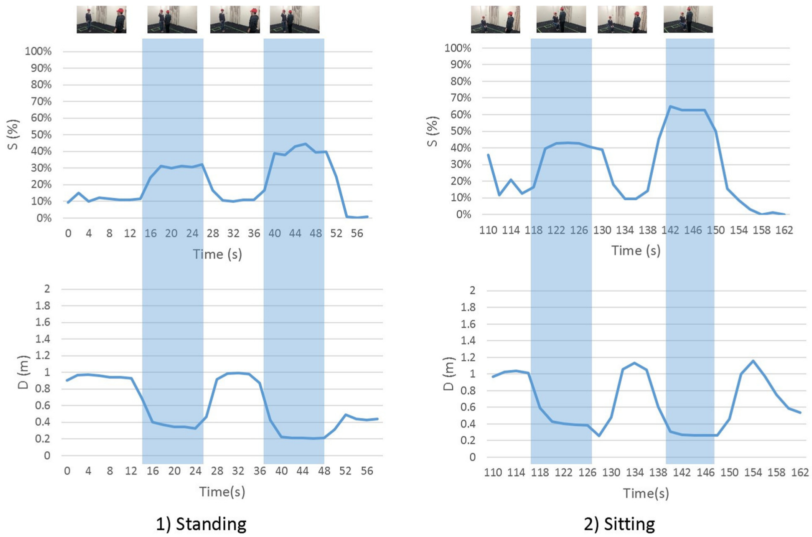

Fig. 3. Result of preliminary experiment

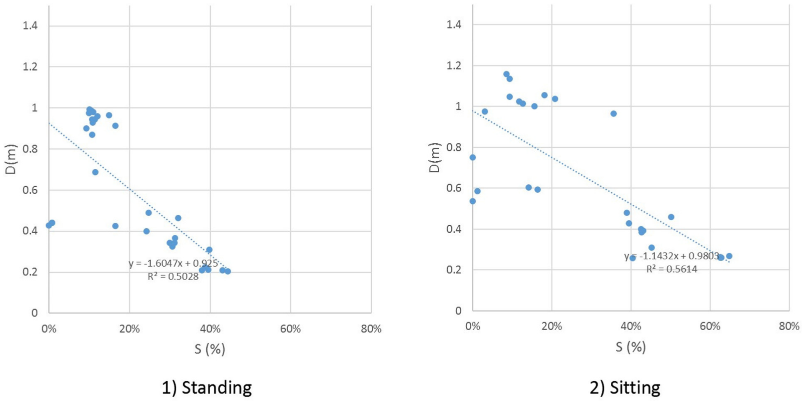

Fig. 4. Correlation between interpersonal distance and interpersonal proximity in preliminary experiments

Table 1. Result of each method

Standing/Sitting	Standing		Sitting	
Limit decision	Approaching	Non-approaching	Approaching	Non-approaching
Stop distance	43 cm	25 cm	40 cm	15 cm
Motion capture	35.72 cm	26.02 cm	37.05 cm	27.00 cm
Proposed	29.8%	40.59%	41.18%	57.99%

4 User Study with Children with ASD

We conducted an experiment for actual therapy for children with ASD. Participants in such therapy are children diagnosed with ASD. The chronological age was 5 years, 1 month, while the developmental age was 3 years, 4 months. The developmental age was calculated via the Kyoto Scale of Psychological Development [14], which is used in Japan as the standardized diagnosis scale. This experiment was approved by the Institutional Review Board of Anonymous Institute, and participants joined the experiment after providing informed consent.

4.1 Result

Figure 5 shows the results for interpersonal distance and interpersonal proximity. Figure 6 shows the correlation between interpersonal distance and interpersonal proximity. The correlation coefficient was .33.

4.2 Discussion

From the result of the pilot experiments, a certain correlation between the interpersonal distance obtained from motion capture and the proposed method was

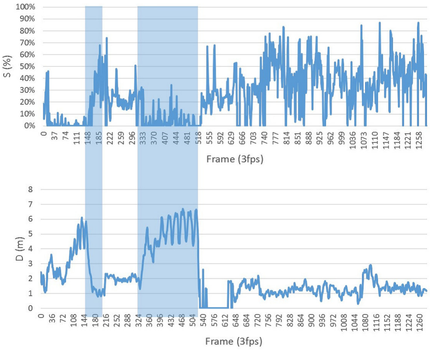

Fig. 5. Result of actual therapy

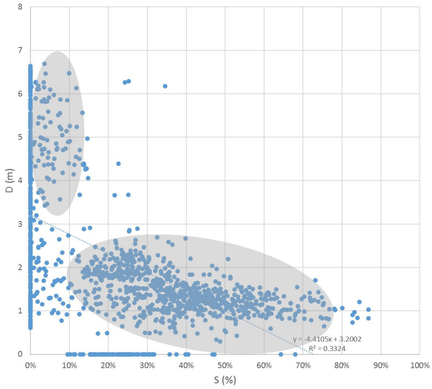

Fig. 6. Correration between interpersonal distance and interpersonal proximity of actual therapy

observed, as shown in Fig. 4. This result explains the target person looks larger as the distance is shorter when the target person appears within the field of view. We carefully checked if the head-mounted camera disturbed the therapeutic activities, but there was no critical problem during the experiment.

On the other hand, the results of the user study did not show a strong negative correlation compared to the pilot experiment, as shown in Fig. 6. We consider that there are several times when the participant was not looking at the therapist although they are actually at a close distance. However, except for that area which is illustrated the upper gray circle, a certain correlation also appears in the lower gray-circle area in Fig. 6. We assume that the participant may consider close proximity, although the target person does not appear in the field of view, the participant may consider close proximity if s/he knows well the target person stands behind or beside. However, its proximity may decrease compared with the target person standing in front of them. We may develop a time-varying algorithm that includes temporal characteristics of proximity. Although further analysis is needed, our findings provide insights between the interpersonal distance and interpersonal proximity.

5 Conclusion

In this study, we proposed a method via which to measure interpersonal proximity using an image sequence from a first-person view. From experimental results, we found high feasibility to use the system for therapeutic activities and also the possibility to estimate interpersonal proximity from the computer vision approach. In the future, we will conduct long-term therapy measurements to determine whether the proposed method can be used to measure changes in interpersonal proximity in therapy and verify the resultant effects.

References

1. Hall, E.T.: The hidden dimension (1966)
2. Kozlowski, L., Bryant, K.: Sense of direction, spatial orientation, and cognitive maps. J. Exp. Psychol.: Human Percept. Perform. **3**, 590–598 (1977)
3. American Psychiatric Association, et al.: Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub (2013)
4. Rogers, S.J., et al.: Effects of a brief early start Denver model (ESDM)-based parent intervention on toddlers at risk for Autism spectrum disorders: a randomized controlled trial. J. Am. Acad. Child Adolescent Psychiatry **51**(10), 1052–1065 (2012)
5. Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M.: Interpersonal distance in immersive virtual environments. Pers. Soc. Psychol. Bull. **29**(7), 819–833 (2003)
6. Rogers, A.L., Fine, H.J.: Personal distance in play therapy with an autistic and a symbiotic psychotic child. Psychotherapy: Theory Res. Pract. **14**(1), 41 (1977)
7. Kennedy, D.P., Adolphs, R.: Violations of personal space by individuals with Autism spectrum disorder. PLoS ONE **9**(8), e103369 (2014)
8. Maeda, M., Sakata, N.: Controlling the interpersonal distance using the virtual body size. IPSJ Interact. **2016**, 47–53 (2016). (in Japanese)
9. Gessaroli, E., Santelli, E., di Pellegrino, G., Frassinetti, F.: Personal space regulation in childhood autism spectrum disorders. PLoS ONE **8**(9), e74959 (2013)
10. Tsuji, A., Matsuda, S., Suzuki, K.: Interpersonal distance and face-to-face behavior during therapeutic activities for children with ASD. In: Miesenberger, K., Bühler, C., Penaz, P. (eds.) ICCHP 2016. LNCS, vol. 9759, pp. 367–374. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41267-2_52
11. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969 (2017)
12. COCO - Common Objects in Context. http://cocodataset.org/
13. Tobii pro glass. https://www.tobiipro.com/ja/product-listing/tobii-pro-glasses-2/
14. Ikuzawa, M., Matsushita, Y., Nakase, A.: Kyoto scale of psychological development 2001. Kyoto International Social Welfare Exchange Centre, Kyoto (2002)