International Journal of Child-Computer Interaction 32 (2022) 100392

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

journal homepage: www.elsevier.com/locate/ijcci

Research paper

Supporting collective physical activities by interactive floor projection in a special-needs school setting

Mika Oki [a,*], Shuichi Akizuki [b], Baptiste Bourreau [a], Issey Takahashi [a], Yoshimitsu Aoki [c], Junichi Yamamoto [d], Kenji Suzuki [a]

[a] Artificial Intelligence Laboratory, University of Tsukuba, Ibaraki, 305-8573, Japan
[b] School of Engineering, Chukyo University, Aichi, 466-8666, Japan
[c] Faculty of Science and Technology, Keio University, Kanagawa, 223-8522, Japan
[d] Faculty of Letters, Keio University, Tokyo, 108-8345, Japan

ARTICLE INFO

Article history:
Received 21 October 2020
Received in revised form 15 June 2021
Accepted 20 August 2021
Available online 10 September 2021

Keywords:
Collective physical activity
Robust human tracking
Interactive floor projection
Behavior change
Neurodevelopmental disorders

ABSTRACT

This paper presents an algorithm to provide floor projection feedback according to the local distance and density of individuals. It is realized by a large-space floor projection system with a feedback function based on human tracking with laser ranging image sensors. The purpose is to support the cognition of spatial–temporal structures of groups of adolescents with neurodevelopmental disorders (NDs) that are conducting organized physical activity (PA). Observation and evaluation of behavioral changes in adolescents with NDs, when they were active with or without the floor projection based on the proposed algorithm, were conducted to validate its effectiveness. We observed that the proposed algorithm can be implemented in different organized PAs. It had the effect to help individuals in a behavior to keep a close distance to each other as a group rather than to keep the same distance apart from each other while walking.

1. Introduction

Neurodevelopmental disorders (NDs) are an impairment regarded as “a group of heterogeneous conditions characterized by delay or disturbance in the acquisition of skills in a variety of developmental domains, including motor, social, language, and cognition” (Thapar, Cooper, & Rutter, 2017). Examples of NDs include a wide range of disabilities such as intellectual disability (ID), autism spectrum disorder (ASD), developmental coordination disorder (DCD), and communication disorders (American Psychiatric Association, 2013).

Physical activity (PA) participation is beneficial for children with disabilities for therapeutic reasons as well as general physical and social development (Johnson, 2009; Pontifex, Fine, da Cruz, Parks, & Smith, 2014). A systematic review suggests that PA offers a safe and alternative form of treatment for children aged 21 and under (Tayla, April, Kirsten, & Jeanette, 2017). In addition to the physical benefits, enhanced social identity, including perceptions of competence and similarity to peers, enhanced self-worth, and strengthening of social interaction and bonding have been reported (Arim, Findlay, & Kohen, 2012; Taub & Greer, 2000).

In this paper, we use the term organized PA in reference to sports, lessons, instructions, and collective activities that are conducted under specific rules and predetermined tasks with a coach or instructor (Arim et al., 2012). School teachers and therapists spend much time and effort on teaching organized PA based on the curriculum; these are collective activities performed among a group of people (Choi & Savarese, 2014). Fig. 1 shows an example. To acquire these skills, practice levels are defined so as to make the elements of learning clearer (Phillips, Hannon, & Molina, 2015). Lessons are conducted for achieving the most basic skills to perform collective PAs with others, such as walking together, gathering, keeping a certain distance from each other, waiting in line, and so on.

In collective PAs, we consider that spatial–temporal coordination of understanding the task assigned either verbally or by demonstration, the ability of spatial awareness (perception of space and surrounding persons, prediction of action and interaction of other people, choosing the most appropriate interpersonal distance (Gabbard, 2012; Klatzky, 1998; Kozlowski & Bryant, 1977)), as well as the action determination and physical movement of the body accordingly, are needed.

However, existing reports on each of these elements state that people with ND have difficulties with them. For example, people with ASD have been characterized as visual learners, in that they process visual information more effectively than auditory information (Greenspan & Wieder, 1997; Quill, 1995; Tissot & Evans, 2003). People with ND have difficulty in reconstructing,

* Corresponding author.
E-mail address: m.oki@ai.iit.tsukuba.ac.jp (M. Oki).

https://doi.org/10.1016/j.ijcci.2021.100392

Fig. 1. One example of organized collective PA (Delivery Service Game (DSG)). Details are shown in Figs. 6 and 7.

understanding, and taking actions according to verbal words. In addition, imitation is reported to be challenging for them. They may be limited in terms of motor abilities and social skills, thus impacting their ability to participate in PA (Arim et al., 2012; Chester & Calhoun, 2012; Donnellan, Hill, & Leary, 2012). It is known that the development of spatial awareness is difficult for children with DCD, ASD, and other conditions (Bryson, Wainwright-Sharp, & Smith, 1990; Coulter, 2009).

Adjusting and structuring the environment by providing additional visible instructions to support the cognitive abilities of children with ND is helpful (Olley, 2005; Sharmin et al., 2018). We have been studying an interactive floor projection system called FUTUREGYM featuring image projecting and capturing capabilities as implemented in the gymnasium of a special needs school (Oki, Bourreau, Takahashi, & Suzuki, 2019; Takahashi, Oki, Bourreau, Kitahara, & Suzuki, 2018a, 2018b, 2018c). With the help of FUTUREGYM, students with ND can perform PA in an efficient manner (enhancing their social interactions (Takahashi et al., 2018c), understanding of cleaning (Takahashi et al., 2018a), and cognition, as well as augmenting PA classes (Oki et al., 2019; Takahashi et al., 2018b)).

In this study, we propose an algorithm to provide floor projection feedback according to the local distance and density of individuals, in order to support the cognition of spatial–temporal structures of groups of individuals. This is supposed to lead to the success of a certain predetermined task in organized PA. By using these floor projection algorithms associated with different collective PAs, we consider it has a potential to support the behavior of students with ND in performing various tasks.

To validate the effectiveness of the proposed algorithm, we designed three different collective PAs, and performed observation and evaluation of behavioral changes in adolescents with ND when they were active with or without the floor projection feedback. All of them are conducted with the participation of four people at once. The predetermined task in the PAs is that the four people cooperate when walking in a certain spatial location, as shown in Fig. 2.

We hypothesize that the proposed projection algorithm works as a visual aid to change their behavior to accomplish the desired tasks.

Our contributions are three-fold:

1. Proposing a floor projection feedback algorithm in which the projection pattern changes according to the local distance and density of individuals.
2. Verifying that the proposed floor projection feedback can be effectively used as a visual aid for adolescents with ND in a behavior to keep a close distance to each other as a group while walking.
3. Introducing possible applications of the proposed algorithm and the potential of floor projection for the acquisition of skills taught in special needs education.

2. Background and related work

2.1. VR/AR/MR for individuals with ND

Interventions for children and people with ND are performed based on the concept of sensory integration, stimulating the most basic perceptual mechanisms and promoting perceptual learning (Fahle & Poggio, 2002), and the concept of embodied cognition (Wilson, 2002). As a result, attempts have been made to functionalize and structure the environment.

Research focused on the application of virtual reality (VR), augmented reality (AR), and mixed reality (MR) for use by individuals with ND is increasing exponentially globally. This is a reasonable method because visual assistance is effective for children and individuals with ND who are reported to be visually dominant (Greenspan & Wieder, 1997; Quill, 1995; Tissot & Evans, 2003). This is also pointed out in Kientz, Hayes, Goodwin, Gelsomini, and Abowd (2020) as one of the several myriad putative affordances associated with VR technologies for autism. By implementing interactive floor projection, we aim to lower the cognitive burdens of reconstructing and understanding verbal words. It acts as a visual medium so that users in the environment can perceive the context that should be shared. This will lead to a mutual understanding of verbal words among people in the environment.

A VR environment using tablet games and floor-wall projection has been proposed as a method for interactively providing visual assistance. For example, studies aimed at supporting social interaction (Parés et al., 2005; Parsons, 2015; Parsons & Mitchell, 2002) and promoting collaborative behavior by providing interactions using the whole body by projecting onto the floor and walls were conducted, and the effects of the intervention have been reported (Cibrian, Ortega, Escobedo, & Tentori, 2015; Cibrian, Tentori, & Martinez-Garcia, 2016; Malinverni & Parés, 2018; Mora-guiard, Crowell, Parés, & Heaton, 2016a, 2016b). In addition, studies have been reported that use projections to teach procedures to achieve certain tasks that are difficult for children with ND (Mueller, Byrne, Andres, & Patibanda, 2018). In addition to the use of visual stimuli, therapy rooms that make use of multi-modal

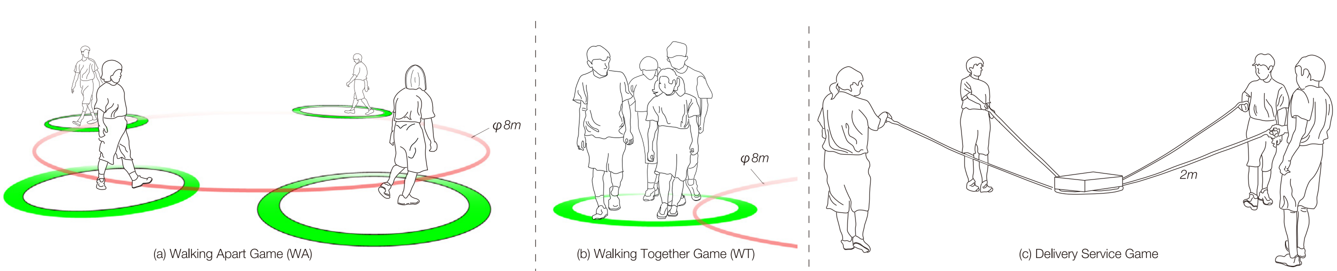

Fig. 2. Concept of (a) Walking Apart Game (WA), (b) Walking Together Game (WT), and (c) Delivery Service Game (DSG). Details are given in Section 5.

stimulation in a complex manner have been proposed (Garzotto & Gelsomini, 2018; Gelsomini, Leonardi, & Garzotto, 2020; Lancioni, Cuvo, & O'Reilly, 2002; Ringland et al., 2014). Through these exploratory studies, a certain effect on the support of cognition, communication, interpersonal interaction, emotion, and motion has been demonstrated. However, its effectiveness in intervention has not yet been sufficiently verified, and further research is required (Kientz et al., 2020) (Khowaja et al., 2020); we believe further investigation is needed in a school setting.

In addition to children and people with ND, AR space is expected to be applied in various situations. For example, a study on the effect of revealing the social presence of people using an interactive floor projection installed in a public space (Monastero & McGookin, 2018) has been reported. Graf et al. (2019) introduced an interactive floor projection system featuring peripersonal circle for inclusive exergames among wheel-chair and non-wheel-chair users. However, few works focused on proposing a specific model to support each element required in the accomplishment of collective PAs.

2.2. Object tracking methods for floor projection

The problem of determining an adequate illumination condition for conducting PAs in a floor projection environment is difficult because the illumination condition should cover the following two requirements, which exhibit a trade-off relation: The first one is ensuring visibility of the field. When conducting PAs among multiple people in the gymnasium, it is necessary to secure a minimum amount of light in order to provide an environment in which other people can be seen while ensuring safety. Therefore, the amount of light is adjusted by opening the dark curtain of the gymnasium, turning on the electricity, or using a highly bright color (e.g., white) for the content to be projected. The second one is ensuring darkness of the field so that the floor projection is visible enough.

In the literature, a method to capture the position of a moving projection target by distance measurement using light with a wavelength band that is different from the image, such as infrared (IR) light (Cibrian et al., 2015, 2016; Zhou, Xiao, Tang, Wei, & Chen, 2016), has been used. In addition, a new sensor dedicated to tracking the target (Hoberman, Pares, & Pares, 1999; Mora-guiard et al., 2016a) was employed. Many methods increase the robustness against illumination fluctuations by dividing the wavelength band of light handled in image projection and measurement (Cibrian et al., 2015, 2016; Siegl, Lange, Stamminger, Bauer, & Thies, 2017; Zhou et al., 2016). Visible light is used to project the image, while a distance sensor consisting of an IR projector and IR camera typified by Kinect is used to measure the object. Because controlling the lighting conditions in the gymnasium is challenging, we need to measure the user position using a distance sensor that is robust to lighting fluctuations, as in these approaches.

There are methods for real-time position tracking of humans and objects in an image. Recently, several machine-learning-based approaches using feature extraction by convolutional neural networks (CNNs), such as Faster R-CNN (Ren, He, Girshick, & Sun, 2015) and YOLOv3 (Redmon & Farhadi, 2018), have been adopted for the detection of the target object (Bewley, Ge, Ott, Ramos, & Upcroft, 2016). Although these methods convert the appearance of a person in an image into features, the environment in this study involves a wide field of view in addition to large illumination variations, leading to changes in the appearance of the target depending on the position in the image. Therefore, in this study, human tracking is realized by a clustering-based object detection approach, which is specialized for point cloud data processing, and a frame-matching method using motion information.

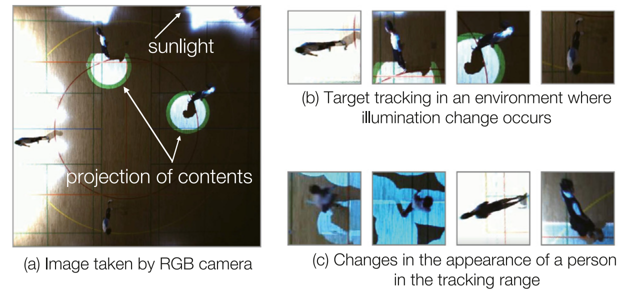

Fig. 3. Image of actual environment and changes in the appearance of a person in the tracking range captured by the ceiling camera.

3. Proposed method

3.1. Robust human tracking

Owing to the dynamic changes in brightness of the camera-shooting environment due to sunlight and projection, as well as the use of a wide-angle RGB camera, the appearance of a person may vary even when the person is stationary. Fig. 3 illustrates the actual environment and changes in the appearance of a person in the tracking range. Fig. 3(c) shows the change in appearance according to the change in the position of the person over the entire field.

We then propose a robust human-tracking method for situations where it is difficult to track using only an RGB camera. Specifically, four laser ranging image sensors, which are time-of-flight (ToF) sensors, were installed in the middle of the gymnasium ceiling, and the point cloud data provided by these sensors were used. A ToF sensor has an advantage in that it is easy to obtain robust measurement results against changes in ambient light, because it projects IR light onto the target for measurement.

Calibration was performed to convert the position coordinates expressed in the ToF sensor coordinate system into the ceiling camera image coordinate system. We confirmed that accurate calibration was realized by observing that all points were superimposed onto the position of the retroreflective material on the image captured by the ceiling camera.

Fig. 4 presents the outline of the person-tracking method, which consists of three steps: preprocessing, detection, and ID assignment.

At the preprocessing step, the set of point cloud data P obtained by the sensor coordinate system from the four ToF sensors, are converted into a floor coordinate system with the center of the shooting range as the center of the imaging range, and with a position of approximately 0.1 m above the floor surface on the z-axis as 0. It is indicated as $\boldsymbol{p}_i \in P$, $\boldsymbol{p}_i = (x_i, y_i, z_i)$, $i = \{1, \dots, N\}$, where N denotes the number of points. The preprocessed point group P', which comprises only the person/object existing on the floor, is obtained by removing the point group of $z_i < 0.0$ in the floor coordinate system.

At the detection step, the preprocessed point group P' goes through clustering for each individual person/object (person and object are not distinguished here, but a person is taken as an example). The points measured on the same person are considered to be close in distance. Therefore, we apply the Euclidean clustering method, that uses the distance between points. The point groups existing within 0.3 m are grouped to generate a cluster, which is denoted as $\boldsymbol{C}_i$, $i = \{1, \dots, K\}$, where K indicates the number of clusters at time t. Subsequently, the position of the centroid of each cluster, $\boldsymbol{c}_i = 1/m_i \sum_{\boldsymbol{p} \in C_i} \boldsymbol{p}$, is calculated, where m_i denotes the number of points that generate the ith cluster $\boldsymbol{c}_i$. Then, assuming that the person does not overlap vertically, the

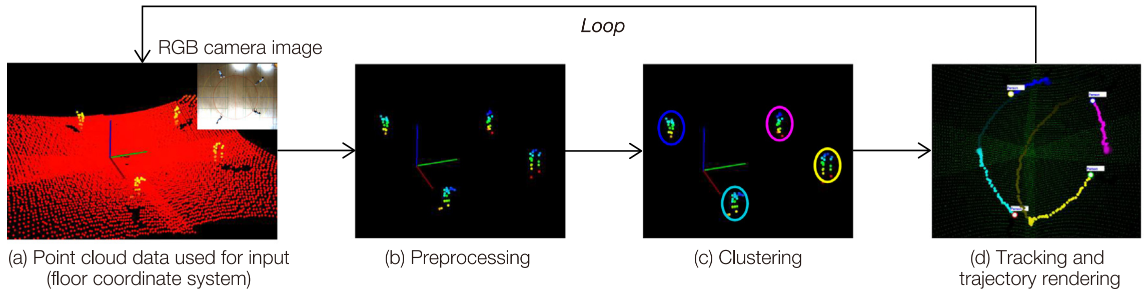

Fig. 4. Outline of the person tracking method, which consists of three steps: preprocessing, detection, and ID assignment.

height information is eliminated and the x, y coordinates of the centroid in the floor coordinate system are used as the position of the person.

The cluster trajectory is obtained by associating the cluster detected in the previous frame with the closest cluster detected in the current frame. First, a tracker, which is denoted as $\boldsymbol{r}$, is generated in the initial frame; $\mathbf{r}$ is composed of i that represents the cluster ID, and the position at time t is indicated by $\boldsymbol{r}_i(t)$. Let $i = k$, $\boldsymbol{r}_i(t) = \boldsymbol{c}_k$ in the initial frame. The number of clusters obtained in the previous frame $K(t-1)$ and that in the current frame $K(t)$ are compared, and their IDs are assigned through three different cases as follows.

In the case of $K(t) < K(t-1)$: It is considered that a person being tracked has moved out of the frame, or has merged into a cluster owing to the approach of multiple persons. In this case, the cluster q at time t ($\boldsymbol{c}_q(t)$) that gives the minimum distance $\hat{i}$ between $\boldsymbol{r}_i(t-1)$ and $\boldsymbol{c}_q(t)$ is assigned to be $\boldsymbol{r}_i(t)$. The function of $\hat{i}$ is indicated by Eq. (1).

$$\hat{i} = \underset{q \in K(t)}{\arg\min} |\boldsymbol{r}_i(t-1) - \boldsymbol{c}_q(t)| \tag{1}$$

$q = \{1, \dots, Q\}$ denotes the cluster ID of the current frame at time t.

In the case of $K(t) > K(t-1)$: It is considered that a new person has entered in the field of view. Therefore, a new tracking target is added. After updating the tracker to the nearest cluster using the same method as above, the surplus cluster is added as a new tracker.

In the case of $K(t) = K(t-1)$: The tracker is updated to the nearest cluster, but the combination that minimizes $\hat{g}$ in Eq. (2) is calculated from the combination of the tracker and cluster so that assignment failure does not occur. The combination of the tracker and cluster is denoted as $g = \{i, j\} \in G$. To improve the robustness of tracking, trajectory prediction using past frames is also introduced. Here, the current tracker position is expressed by $\boldsymbol{r}_i(t) = \boldsymbol{r}_i(t-1) + (\boldsymbol{r}_i(t-1) - \boldsymbol{r}_i(t-2))$,

$$\hat{g} = \underset{g \in G}{\arg\min} \sum_{g} |\boldsymbol{r}_i(t) - \boldsymbol{c}_q(t)| \tag{2}$$

This function minimizes the total distance difference of the entire assignment between the previous frame and current frame. Therefore, it has the effect of minimizing local assignment failures.

3.2. Floor projection algorithm

We propose an algorithm to provide floor projection feedback according to the local distance and density of individuals. The content of floor projection feedback is expressed by a function indicated as $\boldsymbol{H}_i(t)$. At the centroid of the cluster $\boldsymbol{c}_i(t)(x_{c_i}(t), y_{c_i}(t))$, a circular figure with different diameters and colors is drawn on the floor according to the size of the cluster $s_i(t)$ or the distance between the clusters $d_{ij}(t)$.

$$\boldsymbol{H}_i(t) = h(\boldsymbol{c}_i(t), \mu(s_i(t)), \eta(\min d_{ij}(t), s_i(t))) \tag{3}$$

where $\mu(\cdot)$ and $\eta(\cdot, \cdot)$ are functions that determine the diameter and color schemes of the projected circular figure, respectively. Here, μ and η give the following values depending on particular thresholds:

$$\mu = \begin{cases} \mu_1 = 1.0 & (s_i < \theta_S) \\ \mu_2 = 2.0 & (\theta_S \leq s_i < \theta_L) \\ \mu_3 = 3.0 & (\theta_L \leq s_i) \end{cases} \tag{4}$$

$$\eta = \begin{cases} \eta_a = \text{Alert color} \\ \eta_s = \text{Success color} \end{cases} \tag{5}$$

The threshold of each function is denoted as θ_S, θ_L. θ_S and θ_L are the minimum and maximum thresholds of $s_i(t)$, respectively. In this study, the thresholds θ_S and θ_L are determined empirically as 0.3 m and 0.8 m, respectively. If $s_i(t) < \theta_S$, then one person is supposed to be included in the cluster. If $\theta_L \leq s_i(t)$, then it is assumed that four persons are gathered at one location. The alert color of the projected circular figure η is different among each game. The success color is set to green in this study.

3.3. Design of interactive floor projection

The design of the floor projection feedback was determined through discussions with school teachers and by referring to our previous findings (Oki et al., 2019). Examples of the content of floor projection feedback are shown in Figs. 6 and 7. Details are provided in Section 5. One suggestion by the teachers is that simple figures would be better as reinforcers in these tasks, because complex figures would impose considerable cognitive burden on the students to derive the meaning. These are consistent with our prior findings that simple figures, such as circles or lines, are sufficient to support the cognition and change in physical behaviors of students with ND (Oki et al., 2019). The color scheme of the floor projection feedback was determined through discussions with the school teachers. Simple objects also make it easier to focus on the difference in the behavior of students with ND and their reactions to the projection.

The previous feedback algorithms used in FUTUREGYM (Oki et al., 2019; Takahashi et al., 2018a, 2018b, 2018c) gives only individual feedback, and these were static or simply moved at a fixed pace without any group behavior detection. On the other hand, the proposed algorithm in this study provides projection feedback according to the degree of gathering with individuals around oneself, that is, the local distance and density among multiple individuals, by means of detection of the individuals in the field. The diameter and color of the projected circle changes accordingly.

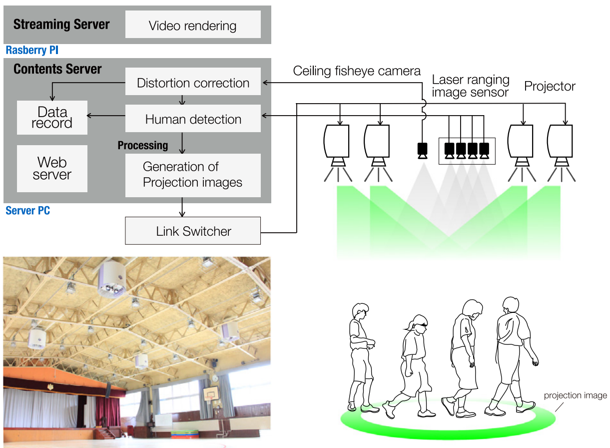

Fig. 5. Configuration of interactive floor projection system of FUTUREGYM.

4. System overview

The configuration of the proposed interactive floor projection system is demonstrated in Fig. 5. The maximum floor projection size is 8.5 × 11.34 m (approximately 558 inches of total screen size), which is realized by eight different digital light processing (DLP) projectors (Panasonic, PT-DW100 W). The projectors are facing downward, with the lens height from the floor being 4.7 m. An RGB camera with a fish-eye lens (Point Grey BFLY-PGE-13S2M-CS) was fixed in the center of the ceiling at a height of 5.2 m. The projection image generated from a Windows server computer (CPU:3.30 GHz, RAM: 16G, OS: Windows 10 Pro) is sent to another server and monitored by four joint displays using a graphic board (NVIDIA, NVS 510). The image on the display is projected by projectors on the ceiling. The gymnasium floor was optimized by repainting with a polyurethane resin coating (Bona Mega EX Matt). This was chosen for anti-reflection in order to project a clear image, and at the same time to maintain appropriate friction of the floor when used for physical educational purposes for children with ND (Takahashi et al., 2018c).

A software image blender (Japan Material, GeoBox G-602) is used so that the four projected images can be stitched without gaps at the borders. The required specifications of human tracking in FUTUREGYM are specified so that the measurement stability and sufficient resolution can be ensured when installed at a ceiling height of 5.2 m. Four laser ranging image sensors (ToF sensors) (Nippon Signal, InfiniSoleil FX8) were installed in the center of the ceiling next to the RGB camera at a height of 5.26 m. The ToF sensor data are acquired at 10 fps, and the floor projection feedback is sequentially projected on the floor at 10 fps. They are also stored and analyzed offline to evaluate the level of achievement of collective PAs (Oki et al., 2019).

5. Experiment

5.1. Collective physical activities

We designed collective PAs called *Walking Apart Game* (WA), *Walking Together Game* (WT), and *Delivery Service Game* (DSG). Each game is conducted with the cooperation of four people at once.

The predetermined tasks in WA and WT were that four participants walk in the same direction around an 8 m-diameter specific circle route in the same counterclockwise direction. The 8 m-diameter specific circle route was a court line originally painted red on the floor of the gymnasium. Meanwhile, the participants were asked to maintain the same distance apart from each other (90 degrees) in WA, as shown in Row(a) of Fig. 6. On the other hand, the participants were asked to keep a close distance to each other in a gathered formation in WT, as shown in Row(c) of Fig. 6. The main difference between WA and WT was the distance between individuals. WA and WT activities themselves were designed to see the basic skills of collective physical activities. These fundamental skills to walk and act while keeping a distance to other participants are needed in a number of collective physical activities at the school such as walking together, lining up, keeping distance to each other in dance classes or to keep physical distance, and sports for people with disabilities such as line soccer, mini port ball. Regarding the floor projection feedback, the main difference was the diameter of the projected circle, μ, and especially the alert color, η, in a failure condition. Nothing will be projected for failure in WA, and a red circular figure will be projected in WT, respectively.

The DSG was based on what the teachers in the cooperating school and a professional specialist of adapted physical activity devised for students with special needs as a game to develop the ability of cooperative games, being aware of others, and nonverbal communication.

The relationship between the success and failure of the predetermined task in each activity and the image examples of projection feedback are shown in Fig. 6.

5.1.1. WA (experiment 1)

The participants were asked to maintain the same distance apart from each other (90 degrees) in WA, as shown in Row(a) of Fig. 6. For WA, the proposed algorithm was applied as follows:

$$\boldsymbol{H}_i(t) = h(\boldsymbol{c}_i(t), \mu_2, \eta(\min d_{ij}(t))) \tag{6}$$

$$\eta = \begin{cases} \eta_a = \text{none} & (\min d_{ij} < \theta_A) \\ \eta_s = \text{green} & (\theta_A \leq \min d_{ij}) \end{cases} \tag{7}$$

where $\mu_2 = 2.0$ m, θ_A is a constant value determined empirically to work at 1 m in the actual field. One session time was $\tau = 60$ s. When performing the predetermined task successfully, a green–white circular figure η_s = green was continuously projected around each participant. When more than two of them were too close to each other, which was a failure condition, it was observed from the participants as nothing projected η_a = none. The reason for not adopting a red circular figure for the feedback of failure in WA was that the teachers suggested that it could make them stop walking rather than being an alert.

5.1.2. WT (experiment 2)

The participants were asked to keep a close distance to each other in a gathered formation in WT, as shown in Row(c) of Fig. 6. For WT, the proposed algorithm was applied as follows:

$$\boldsymbol{H}_i(t) = h(\boldsymbol{c}_i(t), \mu(s_i(t)), \eta(s_i(t))) \tag{8}$$

$$\mu = \begin{cases} \mu_1 = 1.0 & (s_i < \theta_S) \\ \mu_2 = 2.0 & (\theta_S \leq s_i < \theta_L) \\ \mu_3 = 3.0 & (\theta_L \leq s_i) \end{cases} \tag{9}$$

$$\eta = \begin{cases} \eta_a = \text{red} & (s_i < \theta_S) \\ \eta_s = \text{green} & (\theta_S \leq s_i) \end{cases} \tag{10}$$

where $\theta_S = 0.3$ m and $\theta_L = 0.8$ m. One session time was $\tau = 60$ s. When performing the predetermined task successfully, a green–white circular figure η_s = green was continuously projected around the participants. The diameter of the projected circular figure was $\mu_2 = 2.0$ m or $\mu_3 = 3.0$ m. When anybody falled out, which was a failure condition, a red circular figure η_a = red with a diameter of $\mu_1 = 1.0$ m appeared.

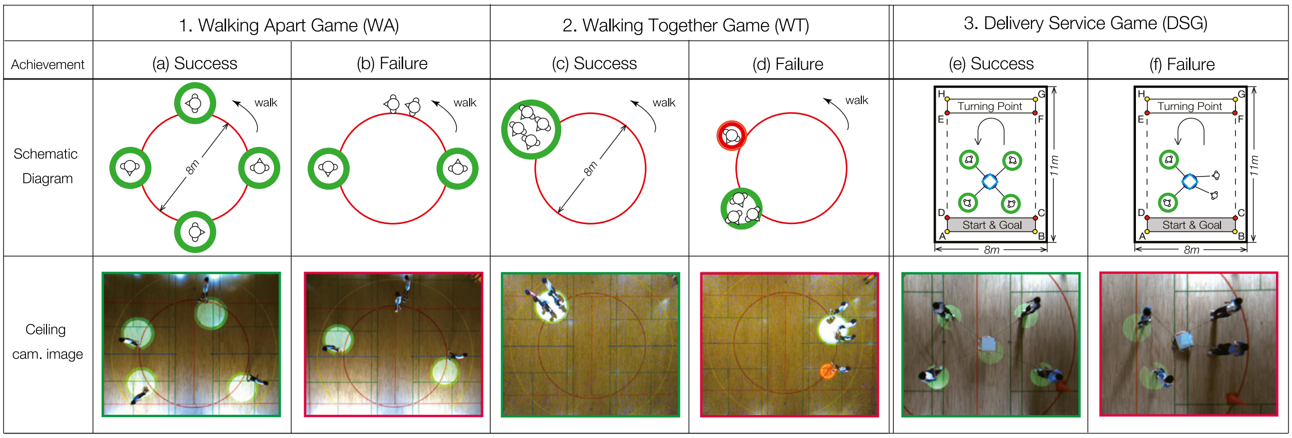

Fig. 6. Schematic diagram of predetermined task and ceiling camera image of projection algorithm (contingent feedback) for each physical activity, Walking Apart Game (WA), Walking Together Game (WT),and Delivery Service Game (DSG). The color and size of the feedback changes to provide an alert when the predetermined task comes to a failure condition by the group.

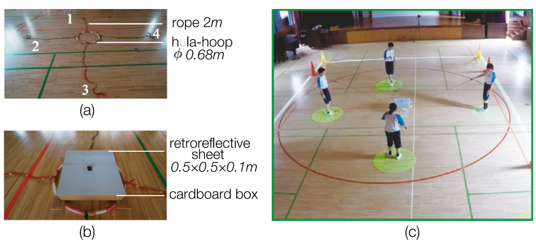

Fig. 7. Material, course, and image from the second floor of the gym for DSG. (a) Four ropes tied to the hula-hoop (Outer and inner circumference: 0.68 m, 0.64 m, respectively. (b) Box placed on the hula-hoop.

5.1.3. DSG (experiment 3)

The Delivery Service Game (DSG) is an educational game that has been developed for special needs education. Here, we use this game for an application of the proposed floor projection algorithm. Its aim is to move one's body while being aware of others and to develop the power of nonverbal communication through cooperation of four people and adjustment to each other's pace. In addition, the manipulation of tools with peers, such as adjusting how to lift and pull a rope while being aware of each other, is experienced. We modified the specific rules according to the environment of FUTUREGYM. The predetermined task in the DSG was that four people carry an object placed on a hula hoop, by pulling the ropes together that were tied to the hula hoop. Details are as follows: The tools were prepared by tying a rope to four points on a hula hoop ($\phi = 0.68$ m). A cardboard box ($W \times D \times H = 0.5 \times 0.5 \times 0.1$ m) was placed on the hula hoop as shown in Fig. 7(b). A retroreflective sheet was attached on top of the cardboard box. Four people made a group, and each person held one rope. The hula hoop was lifted by holding the rope to avoid the box dropping, and then moved from the area surrounded by the lower area of Fig. 6(e), (f) (starting point=goal point) to the area surrounded by the upper area (turning point), and returned to the area surrounded by the lower area. The faster they reach the goal without dropping the box, the better they have performed.

If the four do not always lift the hoop to the same height, the hoop will tilt and the cardboard box will fall from the hula hoop. The tips for success are pulling the rope outwards from the hoop with the same force at 90° to the hoop, as well as holding the rope so that the two people on a diagonal line will stand in a straight line. This requires that the angle between each other is always 90°, and that the distances between the four are always constant. For that reason, we hypothesized that the proposed floor projection algorithm can be effective in supporting the achievement of success, that is, reaching the goal faster without dropping the box. To this end, we projected a green circle around each person when they were at an appropriate distance from each other, and projected nothing as an alert when they got too close to each other, which was similar to WA.

For DSG, the proposed algorithm was applied as follows:

$$\boldsymbol{H}_i(t) = h(\boldsymbol{c}_i(t), \mu_1, \eta(\min d_{ij}(t))) \tag{11}$$

$$\eta = \begin{cases} \eta_a = \text{none} & (\min d_{ij} < \theta_B) \\ \eta_s = \text{green} & (\theta_B \le \min d_{ij}) \end{cases} \tag{12}$$

where $\mu_1 = 1.0$ m, θ_B is a constant value determined empirically to work at 2 m in the actual field. One session time τ was defined when they reach the goal. When performing the predetermined task successfully, a green circular figure $\eta_s =$ green was continuously projected around each participant. When more than two of them were too close to each other, which was a failure condition, it was observed from the participants as nothing projected $\eta_a =$ none.

5.2. Participants

The detailed demographic information of the participants with ND are presented in Table 1. We recruited a total of 16 students in higher secondary education (15–17 years old). Every participant had prior experience playing in FUTUREGYM in recent years. The study was carried out with the approval of the Ethics Review Board of the Education Bureau at the University. The special needs school with FUTUREGYM was affiliated with the University. The students signed consent forms on entering the school to participate in the experiments.

5.2.1. Experiment 1

The students who participated in WA were high school 2nd graders with mild/moderate ASD and ID (n=8, FSIQ: 46.3 ± 5.3, one adolescent could not be scored, FSIQ=Full Scale IQ measured by WISC-IV). They were divided into two groups (Group1 (G1), Group2 (G2)). When dividing the eight participants of each grade into each group of four, we asked the teachers to divide them into two groups so that the average FSIQ (or level of doing activities) would be almost the same between each group. The members were fixed to each group.

Table 1
Participant and group information.

Activity	Group no.	Participant no.	CA (yrs.)	WISC-IV score				Gender
				FSIQ		VCI		
WA	G1	P1–P4	16–17	46.3 ± 5.3[a]	49.0 ± 4.8	54.6 ± 8.2[a]	56.0 ± 6.3	4 M
	G2	P5–P8			42.7 ± 3.8[a]		52.7 ± 11.6[a]	3 F, 1 M
WT	G3	P9–P12	15–16	52.1 ± 15.9[a]	50.3 ± 17.2	56.4 ± 13.1[a]	52.3 ± 9.4	2 F, 2 M
	G4	P13–P16			54.7 ± 17.2[a]		62.0 ± 17.3[a]	2 F, 2 M
DSG	G5	P1, 3, 4, 7	16–17	45.4 ± 5.1[a]	46.5 ± 5.9	53.3 ± 5.9[a]	55.0 ± 7.5	1 F, 3 M
	G6	P11, 13, 14, 15	15–16		44.0 ± 4.4[a]		51.0 ± 2.0[a]	2 F, 2 M

[a]One adolescent could not be scored, F = Female, M = Male.

5.2.2. Experiment 2

As shown in Table 1, the students who participated in the WT were high school 1st graders (15–16 years old) with mild/moderate ASD and ID (n=8, FSIQ: 52.1 ± 15.9, one adolescent could not be scored). In the same way as in Experiment 1, they were divided into two groups: Group3 (G3) and Group4 (G4).

5.2.3. Comparative study

For reference, we conducted the same group activities with typically developing students (TD) of almost the same chronological age. We recruited 8 typically developing students of high school 1st graders (15–16 years old), who were divided into two groups (T1 and T2). The members were fixed among each group. Both T1 and T2 conducted both WA and WT.

5.2.4. Experiment 3

We conducted a preliminary experiment by 8 students with ND. As shown in Table 1, four students were recruited from high school 2nd graders who participated in WA, called group G5. Another four students were recruited from high school 1st graders who participated in WT, called group G6 (n=8, FSIQ: 45.4 ± 5.1, one adolescent could not be scored).

5.3. Experiment design

The points common to Experiment 1 and 2 for students with ND regarding the number of session days, duration from one session to another, and session order per day, were as follows: Four sessions of ABAB design (A1, B1, A2, and B2) for the two groups were conducted per day. The A phase = without projection, will be called the baseline condition and the B phase = with projection, will be called the intervention condition, respectively. Each session was done one group after another so that each group could take a rest after participating in a session. The first group was alternated by day.

5.3.1. Experiment 1

The 5-day intervention experiment was conducted over 2 weeks, and 2 months later, additional intervention experiment was conducted for 1 day. For G1, at least one student was not able to participate due to absence from school or leaving the session on their own will for 3 days. For G2, at least one student was not able to participate due to absence from school or leaving the session on their own will for 2 days. The data of 3 out of 6 days was used for analysis. For TD students, two sessions of AB design (A1, B1) for the two groups were conducted on a single day.

5.3.2. Experiment 2

For students with ND, four sessions of ABAB design (A1, B1, A2, and B2) for the two groups were conducted per day. The 5-day intervention experiment was conducted over 2 weeks. For 2 days, at least one student was not able to participate due to absence from school or leaving the session on their own will. Therefore, the data of 3 out of 5 days was used for analysis. For TD students, two sessions of AB design (A1, B1) for the two groups were conducted on a single day.

5.3.3. Experiment 3

We conducted a preliminary experiment of ABA design for G5 and ABABA design for G6. The experiment was ended with an A phase, because we wanted to see if conducting the B phase would have a short-term effect or not. One-day intervention experiment was conducted.

5.4. Experiment procedure

The experiments were facilitated mainly by an experimenter with the assistance of the school teachers. In addition, one or two support teachers gathered in the gym to watch over the class. When participants did not want to join the session, they were able to tell the teachers and were not forced to participate.

5.4.1. Experiment 1 & Experiment 2

For both the students with ND and TD, the procedure in each session was as follows:

A three-minute instruction was given before the first session. Before each session, the experimenter reminded them of the predetermined task. The order of the participants of each group of how to line around the specific circle route was random in both WA and WT. In addition, in the WT, they were allowed to decide the formation of how to gather close together at the starting line. Examples of formation were standing in a vertical or horizontal straight line against the specific red circle court line, or two participants standing in front with two in the back forming a square shape.

Then, the experimenter verbally starts and ends the 1 min session. In addition, during the session, neither the experimenter nor teachers gave instructions or cheers to the participants, so that the data will demonstrate the effect of floor projection only.

Minimum explanations of floor projection was given before session B1 and B2. To what color and diameter size the floor projection feedback changes according to their physical movement, was not explained; it was left to their own findings as embedded support (i.e., ways for students to learn through experience as opposed to being taught concepts explicitly). This is because we have learnt from advice from teachers and previous studies (Oki et al., 2019; Takahashi et al., 2018a, 2018c) that if we tell them how the projection works, instead of the predetermined task, they will focus on making the feedback work successfully.

5.4.2. Experiment 3

The teachers showed demonstrations to the participants and explained the tips verbally. The tools are placed at the starting point. The order of the participants of each group of how to stand around the hula hoop was random. At the moment the box falls from the hula hoop, it was required for them to stop the activity, put the box back onto the hula hoop and restart from that point. During the session, the teachers and other four students waiting for their turn, were allowed to cheer them. Minimum explanations of floor projection was given before session B1 and B2. To what kind of color and to what size of diameter

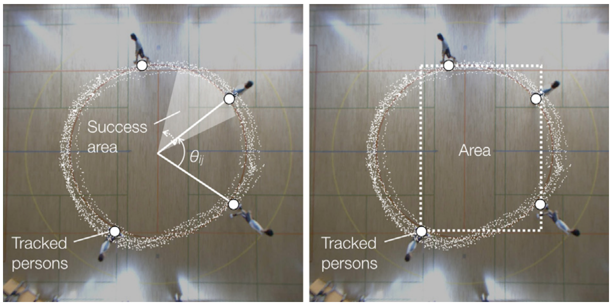

Fig. 8. Schematic diagram of evaluation method of WA (left), and WT (right). The trajectory plot of the x, y coordinate of a whole one-minute walk by four participants is superimposed as white dots on the ceiling RGB camera image.

the floor projection feedback changes according to their physical movement was not explained. That part was left to their own findings.

5.5. Evaluation method

The evaluation was performed based on the positions of the participants using the centroid of the cluster $\boldsymbol{c}_i(t)(x_{c_i}(t), y_{c_i}(t))$ obtained by the ToF sensors. After extracting the data, the trajectory of $\boldsymbol{c}_i(t)$ was plotted. The white dots in Fig. 8 indicate the plots of one session of four people superimposed on the RGB camera image.

5.5.1. Experiment 1

The central angle between one participant and others θ_{ij} were calculated (Fig. 8(left)). The success condition for WA was defined as when the angle between two persons was 90 $\pm$ 22.5 deg. ($67.5 < \theta_{ij} < 112.5$). This reflects the success of the predetermined task, that the four students keep the same distance apart from each other around the specific circle route. We then obtained the accumulated time of the success period averaged by each participant. This evaluation method provided only one data point for one session. The larger the success time, the more the students were able to keep the same distance apart from each other. The level of achievement was based on the time length of the success per session (longer is better).

5.5.2. Experiment 2

The success condition for WT was defined as the square area that surrounds the four students, which was calculated by the maximum and minimum coordinates of $x_{c_i}(t)$ and $y_{c_i}(t)$ among the four participants. That is, $(\max x_{c_i}(t) - \min x_{c_i}(t)) \cdot (\max y_{c_i}(t) - \min y_{c_i}(t))$ m^2, as illustrated in Fig. 8(right). The square area data was taken at 10 fps. One session was 60 s. Here, we obtain the level of achievement based on the averaged area per session. The smaller the calculated area, the more it indicates that the success level is high, since the required task was that four students walk together closely.

5.5.3. Experiment 3

The evaluation was done by observing videos using ELAN (ELAN, 2020). The required time from the beginning to achieving our goal, τ, was measured. The number of boxes dropping in each session was also counted. Smaller time and number of dropped boxes corresponds to better achievement.

6. Result

6.1. Experiment 1

The results of WA sessions for 3 days (WA-Day1, WA-Day2, WA-Day3) were illustrated in Fig. 9. The unfilled and filled markers in Fig. 9(a), (b) represent the results of the baseline condition (A phase = without projection), and the intervention condition (B phase = with projection), respectively. There exist no error bars because the evaluation method used here provides only one data point for one session. The horizontal dotted lines in each day, which represent the mean (average) of success time for that day regardless of A or B, also indicate an increase in success time with each passing day.

Next, based on the results of Fig. 9(a) and (b), a total of 12 plots of the baseline condition (A phase) of group G1 and G2, and 12 plots of the intervention condition (B phase) of group G1 and G2 are given in Fig. 9(c). The IQR 75 percentile–25 percentile were plotted as filled gray markers and the others were plotted with unfilled markers in Fig. 9(c). A Wilcoxon signed-rank test was conducted to determine whether adolescents with ND performed better in the intervention than in the baseline condition. The null hypothesis is that there is not a difference between the intervention and the baseline condition. The statistical tests were conducted to reveal only the effect of having projection or not during the activities among all groups and all days. The phases did not differ significantly ($p=0.23>0.05$). The mean for the intervention condition ($M=23.40$, $SD=10.36$) was not significantly different from that at baseline ($M=19.76$, $SD=5.50$). These findings do not support the idea that the intervention condition was more effective than the baseline condition. The cross markers in Fig. 9(c) are the average success times of T1 and T2. The success times for the baseline and intervention conditions were 52.58 and 59.68 s, respectively. In addition, since the predetermined task achievement level was close to 100% (60 s) in both the baseline and intervention conditions, this indicates that the TD students were able to conduct the task successfully even without the guidance by projection.

6.2. Experiment 2

The results of WT sessions for 3 days (WT-Day1, WT-Day2, WT-Day3) were presented in Fig. 10. The averaged area per session was calculated by plots of 60 s per each session (600 plots). Since there was a detection error for G3-Day1-A1, the results of only 43 s (430 plots) were plotted for G3-Day1-A1&B1. The unfilled and filled markers represent the results of the baseline condition (A phase = without projection) and the intervention condition (B phase = with projection), respectively. Taking a look at WT-Day1 of G3 and G4 altogether (Fig. 10(a) and (b)), there exist four sets of AB in total. It can be observed from the figure that the squared area size decreases in the intervention condition for three out of four sets of AB (75%) on WT-Day1. Taking a look at WT-Day2 and WT-Day3 of G3 and G4 together (Fig. 10(a) and (b)), the squared area size shows a decrease in the intervention condition for four out of four sets of AB (100%) on both WA-Day2 and WA-Day3. This indicates that the participants got closer to each other in the intervention condition compared to the baseline condition at a fairly high probability (more than 75%) even from the first day. Next, as the same way as for WA, in Fig. 10(c), 12 plots are used for the baseline condition and the intervention condition, which are a collection of 6 plots of the baseline condition or the intervention condition of one group (G3), which are seen in Fig. 10(a), and 6 plots of the baseline condition or the intervention condition of the other group (G4), which are seen in Fig. 10(b). The IQR 75 percentile–25 percentile were

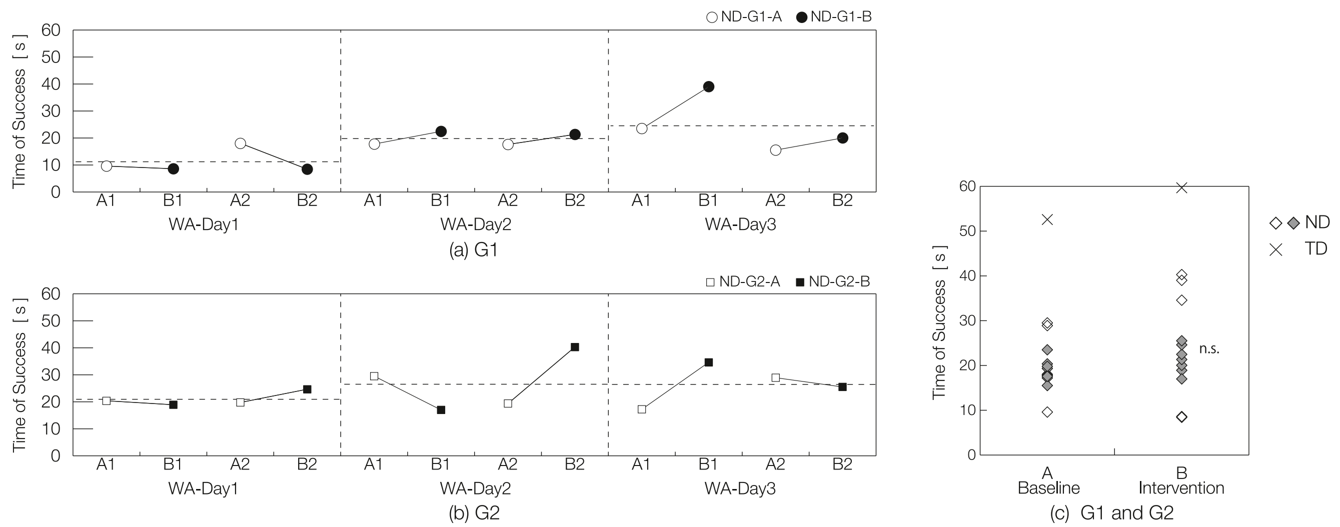

Fig. 9. The time of success when conducting WA activity without (A1, A2) and with (B1, B2) projection for three days for (a) G1 and (b) G2, respectively. For (a) and (b), there exist no error bars because the analysis method used here provides only one data point for a single session. The horizontal dotted lines in each day represent the mean (average) for that day. (c) The time of success when conducting WA activity without (A) and with (B) feedback over all sessions of two groups. For (c), gray markers represent data within 25–75 percentile for each condition. p=0.23>0.05 (n=12). For WA, the data will be counted as success when the angle between two participants is 90 ± 22.5 (67.5 to 112.5) degrees. The success rate was higher when the success time was longer. Since the session was 60 s, the success time would be 60 s if they achieved a 100% success rate.

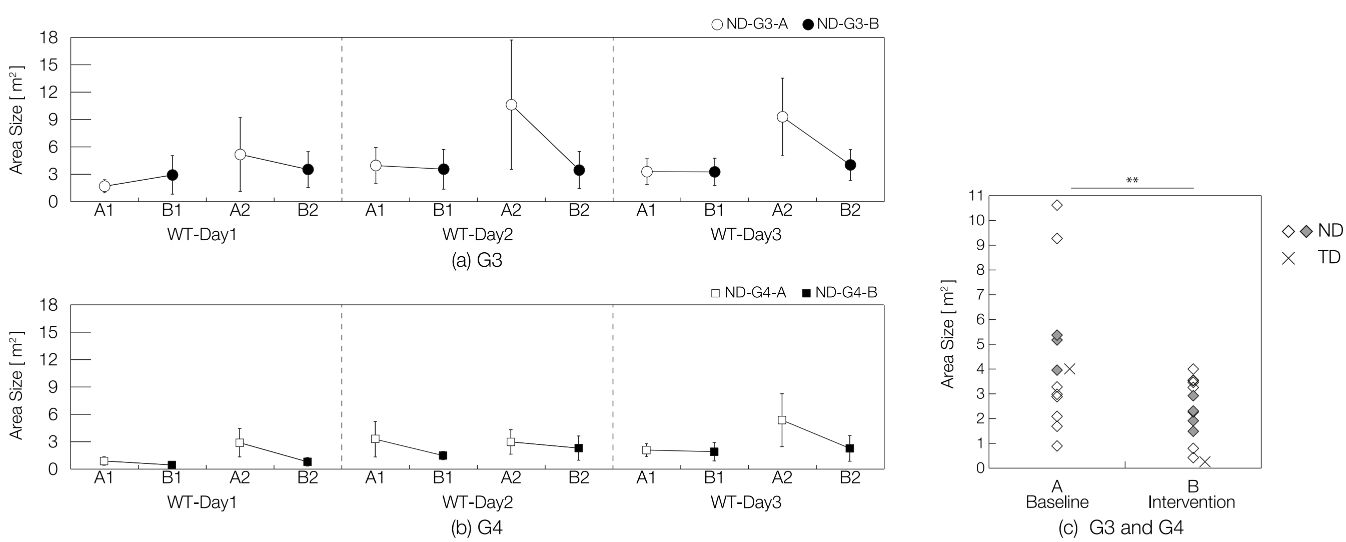

Fig. 10. The area size surrounded by 4 participants when conducting WT without (A1, A2) and with (B1, B2) feedback over three days for (a) G3 and (b) G4, respectively. For (a) and (b), Error bars indicate ± standard deviation. (c) The area size surrounded by 4 participants when conducting WT without (A) and with (B) feedback over all sessions of two groups. For (c), gray markers represent data within 25–75 percentile for each condition. p=0.007<0.05, d=0.85 (n=12). For WT, smaller area size indicates higher success rate.

plotted as filled gray markers and the others were plotted with unfilled markers in Fig. 10(c). A Wilcoxon signed-rank test was conducted to determine whether adolescents with ND performed better in the intervention than in the baseline condition. The phases differed significantly (p=0.007<0.05). The mean for the intervention condition (M=2.49, SD=1.16) was significantly smaller (success trend was significantly higher) than the baseline condition (M=4.29, SD=2.95), and the effect size was high (d=0.85). These findings support the idea that the intervention condition was more effective than the baseline condition, which was an indication of a positive effect of the proposed floor-projection feedback. The cross markers in Fig. 10(c) are the average squared area sizes of T1 and T2. The degree of decrease from baseline to intervention was sharper than that of adolescents with ND.

6.3. Experiment 3

The results are shown in Fig. 11. For G5, both the time duration τ and the number of dropped boxes were smaller in the intervention than in the baseline condition, which might indicate the efficacy of our proposed algorithm. For G6, the number of dropped boxes was smaller in the intervention condition. Although the time duration increased slightly during the intervention, the deviation from the baseline condition was small in the intervention condition. From observations during the experiment and through ELAN, we observed that conducting this activity was challenging for students with ND. It required patience, but that seemed to lead to a sense of accomplishment after reaching the goal, because we observed that they were praising each other and showed positive behaviors, such as high fives and saying "hooray" after reaching the goal. This task is not so difficult

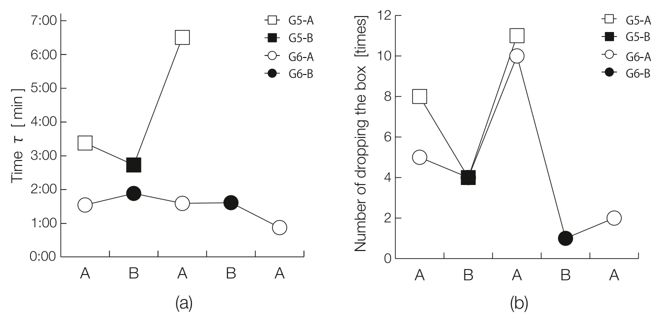

Fig. 11. Examples of time duration and number of dropped boxes in DSG for two groups. (a) Time duration starting from the start to achieving the goal, and (b) number of dropped boxes from the start to achieving the goal.

for teachers or adolescents who understand verbal instructions easily.

7. Discussion

7.1. Floor projection algorithm for WA

For WA, the results do not support the hypothesis because we were not able to make it clear whether the increase in success time in the intervention condition was made by the floor projection or not. We observed that some students did not understand the task to walk at an appropriate pace because they were not looking at both the person in the front and behind so as to keep the same distance apart from them. In WA, participants need to pay attention to a wider range in order to recognize and confirm each other. In the intervention condition, it is required to keep attention to a wider range, sometimes even using peripheral vision to see both the floor projection feedback and the distance from yourself to others.

Even in the case of TD students, the effect of the proposed projection algorithm cannot be discussed, since the success level was high even in the baseline condition. This means that the predetermined task, which was verbally explained, was not difficult for the TD students to follow, which was also an indication that the feedback did not interfere with their behavior in the intervention condition. There was no need to perform continuous interventions, because the predetermined task was successfully performed with ease.

7.2. Floor projection algorithm for WT

For WT, it was suggested that the hypothesis was supported, because there existed a significant decrease in the result comparing the mean square size area of eight students with ND of the whole three days in the intervention condition compared to the baseline condition. We consider that participants used the interactive floor projection as a cue to help the success of the predetermined task.

Evidence suggests that individuals with ND have reduced personal space (Asada et al., 2016) in dyadic situations, specifically many are in face-to-face situations. It was not possible to make clear from this experiment and further investigation is needed whether the inherent proximity problem affected the results or not.

In the case of TD students, it was shown that the degree of decrease of the squared area size from the baseline to intervention condition was sharper than that of adolescents with ND. As we can observe that the squared area size for ND (4.36 m^2) and TD (3.66 m^2) were close values in the baseline condition, it could be said that the degree of how close to gather was not clear from the verbal instructions even for the TD students. However, it can be considered that the TD students understood the contingent behavior of the projection feedback, which was given according to their movement, and used that visual feedback as an indicator of how close to gather.

7.3. DSG

From observations, the task in DSG seemed difficult, because it utilized an object, making the elements for conducting the activity more complicated. However, from the one-day intervention results, the feasibility of the proposed algorithm was suggested. It was shown in the DSG that the similar application of the algorithm as WA was effective to achieve a better DSG level.

We consider that having a physical object in between the students has a possibility to impact the activity. The task cannot be completed unless the students work cooperatively together keeping the same distance apart from each other as well as in WA. In DSG, the presence of an object determines a focal point for the participants, and the task of carrying a box is easily shared among the four participants. Being able to carry the box directly means “achievement” of the predetermined task, and the moment the box falls, it directly means “failure”, so it is easy for the four people to share the situation of failure, and it is possible to try at that time to stop the activity and restart. It is noted that, floor projection feedback can be combined in the form of extending existing school activities whether or not a tool is used.

By using the developed system, the places and objects to pay attention to, such as where the students should look at and where to put their hands on, can be visually emphasized. We consider that the instructions of the teachers given to students can be extended.

7.4. Behavioral repertoire

From the experimental results, we consider WT as one of the activities that were already in their repertoire of behavior of collective physical activity. However, the degree of gathering was not clear in order to successfully finish the predetermined task. We consider that the proposed projection algorithm performed effectively as an index of the degree of gathering. On the other hand, it is possible that WA was not in their repertoire of behavior. When we teach them to learn a new task that is not in their repertoire, it might be that practices with projection contents that do not change according to their movement but move at a fixed pace and size would be effective. In our previous study, the effect of showing the trajectory to achieve a given task was verified in a cleaning task scenario (Takahashi et al., 2018a).

7.5. Future work

Although VR environment using tablet games and floor-wall projection has been proposed as a method for interactively providing visual assistance, its effectiveness in intervention has not yet been sufficiently verified, and further investigation is needed in a school setting (Khowaja et al., 2020). In this study, we provided a new study so as to verify the difference between the baseline and intervention condition based on experiments on a short term. We saw a lot of potential in DSG as it was a practical application of collective physical activities which is something much more likely to occur in a physical education classroom.

Future work includes increasing the number of participants of the same chronological age and days of intervention, to investigate the probability of generalizing the obtained system result and behavior. In addition, conducting a follow-up condition would be useful to check the long-term effect to verify its effectiveness in intervention. By deploying this in the gymnasium of the school as in this study, it will open up the possibility of conducting a long-term intervention research that verifies the efforts of various subjects.

The strong demand from school teachers to support collective PAs was one of the motivations to propose this algorithm. When difficulties lie in guiding the students only through verbal words in how to move with their peers or the degree of behavior in organized PAs that are conducted under specific rules and predetermined tasks, the proposed feedback given by interactive floor projection showed a potential to be a complementary and effective tool.

We hope that the support of organized collective PAs in the FUTUREGYM will lead to inclusion of playing with neurodiverse peers. We have already been conducting such attempts through physical games and activities in the FUTUREGYM. Before playing inclusion games with TD students, the students with ND could practice the tasks through this algorithm, and then share a higher level of PA.

8. Conclusion

In this paper, we proposed an algorithm to provide floor projection feedback according to the local distance and density of individuals, in order to support the cognition of spatial–temporal structures of groups of individuals so that it will lead to the success of a certain predetermined task during organized PAs. Each game was conducted with the cooperation of four people at once. Observation and evaluation of behavioral changes in adolescents with ND when they were active with or without the floor projection feedback were conducted. We observed that the proposed algorithm can be implemented in different organized collective PAs. The proposed algorithm demonstrated significantly positive results for the support of WT, and although no significant positive effect was observed for WA, it was shown in DSG that a similar application of the algorithm as WA was effective for a better achievement level in DSG. We consider that not only DSG, but also activities that are performed based on the local distance and density of individuals has a potential to be generalized, and that the proposed algorithm can be utilized for tasks in which two or more people must act while considering the spatial–temporal relationship with other individuals.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was partially supported by JST CREST Social Imaging and Social Signals, Japan (No. JPMJCR19A2). The authors would like to thank the teachers and students of the special needs school at Otsuka for their participation in the activities and support of the experiments, and also Taku Hachisu, Ph.D., for the valuable comments.

Selection and participation

All the study's participants were students of Otsuka Special-Needs School, University of Tsukuba, located in Tokyo, Japan. The study took place in the gymnasium of their school where FUTUREGYM was implemented. Data related to the study were collected after approval from the Ethics Review Board of the University of Tsukuba Education Bureau of Laboratory Schools, following all the regulations and recommendations for research with students. The participants and legal guardians signed consent forms to participate in the experiments. The experiments were conducted during the classroom activity under the guidance of teachers and experimenters. We provided the teachers with consent forms and information sheets of the experiments and of WISC-IV tests, which they distributed to legal guardians prior to the work being carried out. In addition, participants were able to withdraw their consent for the data collection at any time without affecting their participation in the activity. When participants did not want to join the session, they were able to tell the teachers and were not forced to participate.

References

American Psychiatric Association *Diagnostic and statistical manual of mental disorders* (2013). (5th ed.) (DSM-5). Arlington, VA.

Arim, R. G., Findlay, L. C., & Kohen, D. E. (2012). Participation in physical activity for children with neurodevelopmental disorders. *International Journal of Pediatrics*, *2012*(460384), http://dx.doi.org/10.1155/2012/460384.

Asada, K., Tojo, Y., Osanai, H., Saito, A., Hasegawa, T., & Kumagaya, S. (2016). Reduced personal space in individuals with autism spectrum disorder. *PLoS One*, *11*(1), Article e0146306. http://dx.doi.org/10.1371/journal.pone.0146306.

Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. In *IEEE intl conference on image processing (ICIP)* (pp. 3464–3468). http://dx.doi.org/10.1109/ICIP.2016.7533003.

Bryson, S., Wainwright-Sharp, J., & Smith, I. M. (1990). Autism: A developmental spatial neglect syndrome?. In J. T. Enns (Ed.), *Advances in psychology, 69. The development of attention: research and theory* (pp. 405–427). Elsevier Science.

Chester, V. L., & Calhoun, M. (2012). Gait symmetry in children with autism. *Autism Research and Treatment*, *2012*(576478), http://dx.doi.org/10.1155/2012/576478.

Choi, W., & Savarese, S. (2014). Understanding collective activities of people from videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(6), 1242–1257. http://dx.doi.org/10.1109/TPAMI.2013.220.

Cibrian, F. L., Ortega, D. H., Escobedo, L., & Tentori, M. (2015). Reflections from a long-term deployment study to design novel interactive surfaces for children with autism. In J. Bravo, R. Hervás, & V. Villarreal (Eds.), *Lecture notes in computer science*: *vol. 9456*, *Ambient intelligence for health. AmIHEALTH 2015* (pp. 167–176). http://dx.doi.org/10.1007/978-3-319-26508-7_17.

Cibrian, F. L., Tentori, M., & Martinez-Garcia, A. I. (2016). Hunting relics: A persuasive exergame to promote collective exercise in young children. *International Journal of Child-Computer Interaction*, *32*(3), 277–294. http://dx.doi.org/10.1080/10447318.2016.1136180.

Coulter, R. A. (2009). Understanding the visual symptoms of individuals with autism spectrum disorder (ASD). *Optometry and Vision Development*, *40*(3), 164–175.

Donnellan, A. M., Hill, D. A., & Leary, M. R. (2012). Rethinking autism: implications of sensory and movement differences for understanding and support. *Frontiers in Integrative Neuroscience*, *6*(124), http://dx.doi.org/10.3389/fnint.2012.00124.

ELAN (Version5.9) [Computer software]. (2020). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

Fahle, M., & Poggio, T. (2002). *Perceptual learning*. MIT Press.

Gabbard, C. P. (2012). *Lifelong motor development*. San Francisco, CA: Pearson Benjamin Cummings.

Garzotto, F., & Gelsomini, M. (2018). Magic Room: A smart space for children with neurodevelopmental disorder. *IEEE Pervasive Computing*, *17*(1), 38–48. http://dx.doi.org/10.1109/MPRV.2018.011591060.

Gelsomini, M., Leonardi, G., & Garzotto, F. (2020). Embodied learning in immersive smart spaces. In *Proc. ACM SIGCHI', vol. 20* (pp. 1–14). http://dx.doi.org/10.1145/3313831.3376667.

Graf, R., Benawri, P., Whitesall, A. E., Carichner, D., Li, Z., Nebeling, M., et al. (2019). iGYM: An interactive floor projection system for inclusive exergame environments. *CHI PLAY*, *19*, 31–43. http://dx.doi.org/10.1145/3311350.3347161.

Greenspan, S. I., & Wieder, S. (1997). Developmental patterns and outcomes in infants and children with disorders in relating and communicating: A chart review of 200 cases of children with autistic spectrum diagnoses. *Journal of Developmental and Learning Disorders*, *1*, 87–142.

Hoberman, P., Pares, N., & Pares, R. (1999). El Ball del Fanalet or Light-pools. In *Proc. of intl conference on virtual systems and multimedia (VSMM'99)* (pp. 270–276).

Johnson, C. C. (2009). The benefits of physical activity for youth with developmental disabilities: a systematic review. *American Journal of Health Promotion*, *23*(3), 157–167. http://dx.doi.org/10.4278/ajhp.070930103.

Khowaja, K., et al. (2020). Augmented reality for learning of children and adolescents with autism spectrum disorder (ASD): A systematic review. *IEEE Access*, *8*, 78779–78807. http://dx.doi.org/10.1109/access.2020.2986608.

Kientz, J. A., Hayes, G. R., Goodwin, M. S., Gelsomini, M., & Abowd, G. D. (2020). In R. M. Baecker (Ed.), *Synthesis lectures on assitive, rehabilitative, and health-preserving technologies, lecture #13, Interactive technologies and autism* (2nd ed.). Morgan and Claypool Publishers.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Lecture notes in computer science*: *Vol. 1404*, *Spatial cognition* (pp. 1–17). Berlin, Heidelberg: Springer.

Kozlowski, L. T., & Bryant, K. J. (1977). Sense of direction, spatial orientation, and cognitive maps. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(4), 590–598. http://dx.doi.org/10.1037/0096-1523.3.4.590.

Lancioni, G. E., Cuvo, A. J., & O'Reilly, M. F. (2002). Snoezelen: An overview of research with people with developmental disabilities and dementia. *Disability and Rehabilitation*, *24*(4), 175–184. http://dx.doi.org/10.1080/09638280110074911.

Malinverni, L., & Parés, N. (2018). Learning from failures in designing and evaluating full-body interaction learning environments. In *Proc. ACM CHI EA', vol. 17* (pp. 1065–1074). http://dx.doi.org/10.1145/3027063.3053352.

Monastero, B., & McGookin, D. (2018). Traces: Studying a public reactive floor-projection of walking trajectories to support social awareness. In *Proc. ACM SIGCHI'18* (p. 487). http://dx.doi.org/10.1145/3173574.3174061.

Mora-guiard, J., Crowell, C., Parés, N., & Heaton, P. (2016a). Lands of Fog: Helping children with autism in social interaction through a full-body interactive experience. In *Proceedings of the 15th international conference on interaction design and children* (pp. 262–274). http://dx.doi.org/10.1145/2930674.2930695.

Mora-guiard, J., Crowell, C., Parés, N., & Heaton, P. (2016b). Sparking social initiation behaviors in children with autism through full-body interaction. *International Journal of Child-Computer Interaction*, *11*, 62–71. http://dx.doi.org/10.1016/j.ijcci.2016.10.006.

Mueller, F. F., Byrne, R., Andres, J., & Patibanda, R. (2018). Experiencing the body as play. In *Proc. of ACM SIGCHI'18* (p. 210). http://dx.doi.org/10.1145/3173574.3173784.

Oki, M., Bourreau, B., Takahashi, I., & Suzuki, K. (2019). CANVAS: A drawing tool for AR-aided special needs education using interactive floor projection. In *Proc. of IEEE SMC', vol. 19* (pp. 2639–2644). http://dx.doi.org/10.1109/SMC.2019.8914003.

Olley, G. J. (2005). Curriculum and classroom structure. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 863–881). New Jersey: John Wiley and Sons.

Parés, N., Carreras, A., Durany, J., Ferrer, J., Freixa, P., Gómez, D., et al. (2005). Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment. In *Proc. ACM IDC', vol. 05* (pp. 110–116). http://dx.doi.org/10.1145/1109540.1109555.

Parsons, S. (2015). Learning to work together: Designing a multi-user virtual reality game for social collaboration and perspective-taking for children with autism. *International Journal of Child-Computer Interaction*, *6*, 28–38. http://dx.doi.org/10.1016/j.ijcci.2015.12.002.

Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research*, *46*, 430–443. http://dx.doi.org/10.1046/j.1365-2788.2002.00425.x.

Phillips, D., Hannon, J. C., & Molina, S. (2015). Teaching spatial awareness in small-sided games. *A Journal for Physical and Sport Educators*, *28*(2), 11–16. http://dx.doi.org/10.1080/08924562.2014.1001101.

Pontifex, M. B., Fine, J. G., da Cruz, K., Parks, A. C., & Smith, A. L. (2014). The role of physical activity in reducing barriers to learning in children with developmental disorders. *Monographs of the Society for Research in Child Development*, *79*(4), 93–118. http://dx.doi.org/10.1111/mono.12132.

Quill, K. A. (1995). Visually cued instruction for children with autism and pervasive developmental disorders. *Focus on Autistic Behavior*, *10*(3), 10–20. http://dx.doi.org/10.1177/108835769501000302.

Redmon, J., & Farhadi, A. (2018). YOLOV3: An incremental improvement. arXiv:1804.02767.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *NIPS'15: Proceedings of the 28th international conference on neural information processing systems* (pp. 91–99). arXiv:1506.01497.

Ringland, K. E., Zalapa, R., Neal, M., Escobedo, L., Tentori, M., & Hayes, G. R. (2014). SensoryPaint: A multimodal sensory intervention for children with neurodevelopmental disorders. In *Proc. ACM Ubicomp* (pp. 873–884). http://dx.doi.org/10.1145/2632048.2632065.

Sharmin, M., Hossain, M. M., Saha, A., Das, M., Maxwell, M., & Ahmed, S. (2018). From research to practice: Informing the design of autism support smart technology. In *Proc. ACM CHI'18, vol. 102* (pp. 1–16). http://dx.doi.org/10.1145/3173574.3173676.

Siegl, C., Lange, V., Stamminger, M., Bauer, F., & Thies, J. (2017). FaceForge: Markerless non-rigid face multi-projection mapping. *IEEE Transactions on Visualization & Computer Graphics*, *23*(11), 2440–2446. http://dx.doi.org/10.1109/TVCG.2017.2734428.

Takahashi, I., Oki, M., Bourreau, B., Kitahara, I., & Suzuki, K. (2018a). Designing interactive visual supports for children with special needs in a school setting. In *Proc. of DIS'18* (pp. 265–275). http://dx.doi.org/10.1145/3196709.3196747.

Takahashi, I., Oki, M., Bourreau, B., Kitahara, I., & Suzuki, K. (2018b). An empathic design approach to an augmented gymnasium in a special needs school setting. *International Journal of Design*, *12*(3), 111–125.

Takahashi, I., Oki, M., Bourreau, B., Kitahara, I., & Suzuki, K. (2018c). FUTUREGYM: A gymnasium with interactive floor projection for children with special needs. *International Journal of Child-Computer Interaction*, *15*, 37–47. http://dx.doi.org/10.1016/j.ijcci.2017.12.002.

Taub, D. E., & Greer, K. R. (2000). Physical activity as a normalizing experience for school-age children with physical disabilities. Implications for legitimation of social identity and enhancement of social ties. *Journal of Sport and Social Issues*, *24*(4), 395–414. http://dx.doi.org/10.1177/0193723500244007.

Tayla, A., April, B., Kirsten, D., & Jeanette, G. (2017). Physical activity interventions for children with social, emotional, and behavioral disabilities – A systematic review. *Journal of Developmental & Behavioral Pediatrics*, *38*(6), 431–445. http://dx.doi.org/10.1097/DBP.0000000000000452.

Thapar, A., Cooper, M., & Rutter, M. (2017). Neurodevelopmental disorders. *Lancet Psychiatry*, *4*(4), 339–346. http://dx.doi.org/10.1016/S2215-0366(16)30376-5.

Tissot, C., & Evans, R. (2003). Visual teaching strategies for children with autism. *Early Child Development and Care*, *173*(4), 425–433. http://dx.doi.org/10.1080/0300443032000079104.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. http://dx.doi.org/10.3758/BF03196322.

Zhou, Y., Xiao, S., Tang, N., Wei, Z., & Chen, X. (2016). Pmomo: Projection mapping on movable 3D object. In *Proc. ACM CHI', vol. 16* (pp. 781–790). http://dx.doi.org/10.1145/2858036.2858329.