Design and Evaluation of Animated Everyday Objects with 6-DoF Social Body Gestures

Ching Yi, Tsai
National Taiwan University
Taiwan
r09944002@ntu.edu.tw

Cheng Hsun, Ho
National Taiwan University
Taiwan
b07902129@ntu.edu.tw

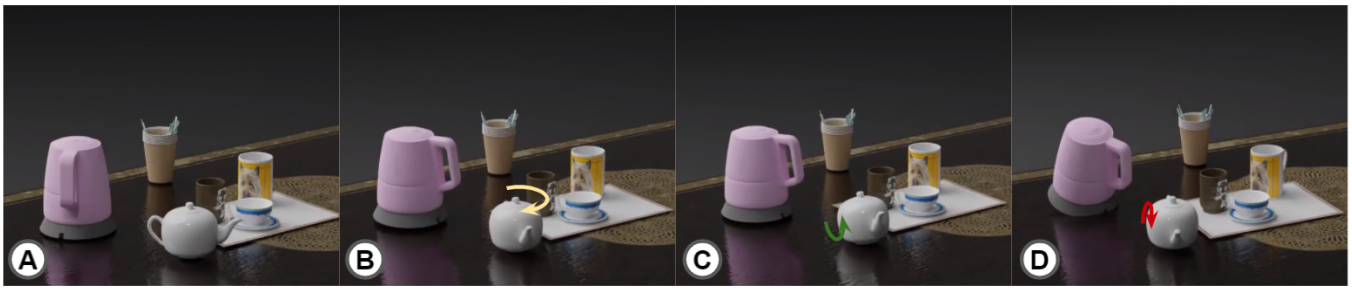

Figure 1: The figure shows one of our application scenario. When users approaching daily objects, the objects (B) turn to the users, and (C)(D) make a bow. Within these six degree of freedom movements, social entity and intents can be formed and expressed.

ABSTRACT

Body gestures play a crucial role in daily human communication. However, while more and more daily artifacts are becoming social entities under the trending of ubiquitous computing and the Internet of Things, they are limited to voice or screen pixel output (e.g., smart speaker). In this paper, we explore the possibility of applying body gestures to everyday objects. Firstly, we understand how people would like to perceive objects' gestures and intentions through an expression survey study with 12 daily scenarios. Then, we design body gestures within six degrees of freedom (6DoF) movements based on users' suggestions and prior work. Lastly, we conducted a 14-person study to examine whether these body gestures suffice to express different social intentions in everyday events. Our user study showed that the body gestures succeeded in providing the perception of various desired social intentions. Also, the 6DoF movements indicate non-intrusive solutions for augmenting everyday objects without redesigning each of them.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Heuristic evaluations**.

AIT'21, Taipei, Taiwan,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/1122445.1122456

KEYWORDS

Object with Intents, Body Gesture, Six Degree of Freedoms

ACM Reference Format:
Ching Yi, Tsai and Cheng Hsun, Ho. 2021. Design and Evaluation of Animated Everyday Objects with 6-DoF Social Body Gestures. In *Proceedings of NTU CSIE Course (AIT'21).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Humans are used to body gestures. For example, people might wave hands to the new-coming members, look away when strangers gaze closely, and follow others' back when they are far away. These body movements are intuitive and familiar in our everyday life.

Recent trends of ubiquitous computing and Internet of Things make electronics are woven in our everyday artifacts, free daily object from their traditional roles, and provides services that grow exponentially every year. For example, a smart speaker like Amazon Echo and Google Nest Audio [6] can process verbal communication, and some emerging smart furniture, including plant pot, lamps, refrigerator, are capable of reporting their status and react with the environment automatically. To sum up, daily objects are becoming more and more likely social agents instead of inanimate things.

Powerful as these objects may seem, their current output channel is limited to voice, text, or through a screen, lacking the intuitive social interaction of body movements and gestures. As social and psychological interaction has been foreseen and identified as crucial difficulties in human-agent interaction (HAI) [9], adding body gestures to everyday objects can be beneficial to enhance social entity for these smart daily objects.

Prior works have displayed objects that perform unexpected gestures, movements, and (mis)behavior [2] and tested users' perception of liveliness and objects' body ownership, thus proving

the movements of an object can produce a sense of emotional and living agents. EnsadLab developed an open-source building blocks kit, Misbehavior kit, that allows people to animate physical objects [4]. However, the kit requires redesign and embedding multiple actuators for animating objects, making the designed gestures inapplicable for different designs or objects in different forms.

In this work, we explore the possibility of applying 6DoF body gestures to everyday objects, where these simple movements would enable non-intrusive solutions (e.g., motion platform) for augmenting daily artifacts without redesigning the gestures nor the actuating hardware. Although these movements might be more straightforward than explicitly designed mechanisms, prior works [1] in non-humanoid robots have indicated perceivable emotions can be built on velocity changes or directional accelerations.

To understand and establish social intents behind body gestures before applying actual body movement, we understand how people would like to perceive objects' gestures and intentions through an 8-user expression survey study with 12 daily scenarios. We found 18 common social intentions and several body gestures design suggestions.

Then, we design body gestures within 3DoF and 6DoF movements based on users' suggestions, prior work [1], and most importantly, motions platforms' physical constraints for the 13 most mentioned social intentions.

Lastly, we conducted a 14-person study to examine whether these body gestures suffice to express different social intentions in everyday events. Although few intentions result in some confusion, our user study showed that the body gestures succeeded in providing the perception of various desired social intentions. In the end, we integrate our result with a virtual scene to allow more users to experience expressive daily objects with body gestures.

Our main result our as follows:

(1) We investigate what expression should daily object behave.
(2) We design 30 body gestures within 3DoF/6DoF movements for various daily scenarios.
(3) We evaluate these body gestures' expressiveness and compare how DoFs affect its performance.

2 RELATED WORK

We discussed the closely related research, their methodology, and the current shortcomings in this section.

2.1 Object with intent

Objects with intent are objects that show their intents through single or multi-modal output. They are designed to enhance the human-object interaction experience for various purposes. Rozendaal et al. developed a ball that can roll, shake and stay for stimulating and interacting with hospitalized children [12]. A set of desk-based artifacts was designed to interact with people working through movements and natural language, which aims to provide office workers with the best working conditions [15]. Objects that perform unexpected gestures, movements, and (mis)behavior were discussed, evaluation of object behavior interpretation was also proposed [2]. EnsadLab developed an open-source building blocks kit, Misbehavior kit, that allows people to animate physical objects [4]. The gestures and movements are currently designed for each object separately with completely different actuator layouts, which indicates that the gestures need to be re-designed every time for a new object or a new layout of actuators. We used the motion platform as a universal and non-intrusive solution for generality, which can be easily added to existing static objects without having to modify them.

2.2 Non-humanoid robots

Non-humanoid robots are robots that do not have a human-like appearance, e.g., Roomba. Saerbeck and Bartneck adjusted the speed acceleration and angular acceleration to create different motion characteristics for a Roomba, which influence the perceived effect by humans [14]. Angel-Fernandez J. et al. created a cylindrical robot with wheels and several body parts that uses different movement speeds, shoulder height, angular velocity, and body rotation to express human-like emotions [1]. Novikova and Watts proposed two categorized models for emotion modeling, both breaks down the modeling process into several layers, including a description of movement, associated body part, and associated emotion [10]. These results indicate that the use of changes of speed, angular velocity, and acceleration may allow daily objects to express emotion or intention, even if they do not have human-like appearances.

3 SOCIAL BODY GESTURE SURVEY FOR EVERYDAY OBJECT

As the original purpose for body gestures is to convey high-level expression or emotion, it is crucial to understand and establish social intents behind each gesture before applying actual body movement. Although this social intention seems to align with human expression easily, users' expectations for an expressive daily object might sometimes deviate from anthropomorphic beings to animistic ones (e.g., a pet).

To have a general understanding of how a vivid daily object should express and perform (should it be more like a servant, a pet, or just an inanimate machine) in different daily scenarios, we conducted a guessability study, in the style similar to Wobbrock et al. [16] and Ruiz et al. [13]. Overall, we surveyed 12 daily scenarios, where users are asked to advise the underlying social intention for the daily objects in each of them and design corresponding body gestures.

3.1 Design

As far as we know, there is no formal framework or procedures for designing body gestures or social intention on machines. However, according to Fong et al. [5], body gestures for robots can generally be based on either functionality or biological design. While the former enhances the affordance of each service the robot provides, the latter makes robots mimic human or animals behaviors. Thus, our survey mainly consists of a two-fold question, where participants were asked to speculate what social intention should daily object express and design the body gesture for it from two different perspectives:

(1) *Functionality Design Perspective*: thinking from a user's point-of-view for understanding what functions would such vivid object be expected, and

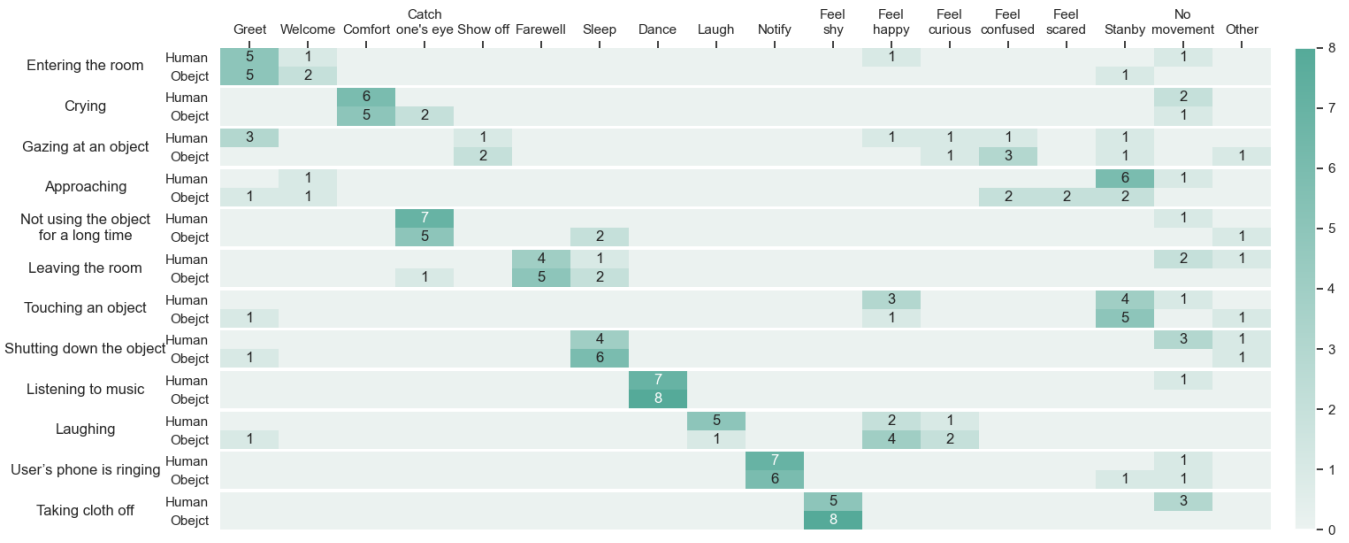

Figure 2: The distribution of 23 intentions designed by participants among 12 scenarios and 2 perspectives. The darker the grid represents the more participants reported the corresponding intention. Intentions that were mentioned only once are labelled as *other*.

(2) *Biological Design Perspective*: role-playing as a daily object to involve expression for intuitive behaviors.

The selected daily scenarios are based on Li's [8] work on designing body movement for an in-house Teddy bear. However, a few of them were modified for the interest of generality (e.g., *"Reaching your hand to pick it up"* were divided into two separate scenarios, *"Approaching the object"* and *"Touching the object"*) as shown in Table 1.

3.2 Participants

We recruited 8 participants (3 male, 5 female) by word-of-mouth, age from 21 to 24 years old, with an average of 22.13 (SD = 1.13).

3.3 Task and Procedure

The survey was conducted in the form of online video meetings. After explaining the study procedure, participants advised a social intention for daily objects and performed a corresponding body gesture in each of the 12 selected daily scenarios. They were allowed to perform the gesture with their heads, through their hands, manipulating actual objects around them, or verbal descriptions, but limited within 6DoF movements. The following are the questions that we used during the survey:

- User's perspective (functionality): When you [in a daily scenario], what do you think such a vivid, expressive and friendly object should perform, what does it mean or want to express?
- Object's perspective (Biological): Imagine yourself as such a vivid, expressive, and friendly object, when a human [with a daily scenario], what would you perform, why, what do you want to express?

The order of the above two questions was counterbalanced among participants, and for each participant, we shuffled the order of the 12 scenarios. At the end of the survey, participants were asked to provide extra daily scenarios that they think body gestures would fit in for daily objects and design additional intention and gestures for them.

Figure 3: Participants designed the body gestures by *a)* moving around their hands, *b)* manipulating a cup and *c)* using a stuffed bear.

3.4 Result and Discussion

3.4.1 Intention Result. There were in total 23 intentions in the participants' responses among 12 scenarios and 2 perspectives. Figure 2 shows the distribution of the intentions.

3.4.2 Qualitative Gestures Feedback. Participants gave verbal descriptions on how the object should perform, which are surprisingly similar between different participants. "Jumping with small rotation around row axis" was reported for *Feel happy* gesture P(0,3,4,5). "Tilt backward with little jump" was reported for *Laugh* gesture P(1,6). "Tilt toward user" was reported for *Standby* gesture P(3,4). "Swing from side to side" was reported for *Farewell* gesture P(1,2,5).

3.4.3 Discussion. Several intentions were mentioned in multiple scenarios, such as *Sleep*, *Feel happy* and *Standby*, which allows the gestures designer to designs once and use in multiple scenarios.

Table 1: 12 scenarios used in expression survey.

(1) Approaching an object	(7) Taking clothes off
(2) Gazing at an object	(8) Listening to music
(3) Touching an object	(9) Entering the room
(4) Leaving the room	(10) User's phone is ringing
(5) Laughing	(11) Shutting down the object
(6) Crying	(12) Not using object for a long time

In some scenarios, the human's perspective and the object's perspective conflict with each other. In the case of *Approaching*, from the user's point of view, participants wanted the object to standby the most for easier usage. However, in the object's perspective, the object could feel confused, scared, or could standby.

4 3DOF/6DOF BODY GESTURE DESIGN

Our goal is to design the body gestures for objects that can be installed on a motion platform. However, we do not know if a 6DoF motion platform is required to express the selected intents or not distinctively, so we designed both 3Dof and 6DoF body gestures for each intent for comparison.

4.1 Design Heuristics

We selected 13 most mentioned intentions from the result of section 3.4.1 for gestures design. The base gestures were based on participants' qualitative feedback in section 3.4.2 with some combination and modification.

The speed and acceleration of the objects were assigned differently depending on the intention, which we hope to make it easier for users to distinguish between intentions [1, 14].

The translation and rotation are constrained by 3DoF/6DoF motion platform. The translation is limited to $\pm\frac{1}{2}$ length of the motion platform. Rotation in roll and pitch is limited to ±30°, comparing to the horizontal plane. Rotation in yaw is limited to 0° to 360°.

4.2 Body Gestures

We recorded the designed body gestures into 3 to 5 seconds video clips [1]. Fig 4 shows the keyframes of 13 intents' body gestures design. Translation and rotation of the body gestures are marked in straight arrows and curved arrows, respectively. Three colors indicate the direction of translation or the axis of rotation, red for the x-axis, green for the y-axis, and yellow for the z-axis.

5 FEASIBILITY EVALUATION

To evaluate whether the designed body gestures suffice for daily usage, we measured the validity of expressing social intentions through our body gestures. we conducted a short evaluation to let users label each 3DoF or 6DoF body gesture with one of the 13 underlying social intentions.

5.0.1 Participants. We recruited another 14 participants (13 male, 1 female) by word-of-mouth, age from 15 to 53 years old, with an average of 24.86 (SD = 8.97).

5.0.2 Task and Procedures. The whole study was conducted through an online video meeting, where participants were asked to fill out a Google Form and share their screen during the whole procedure.

After we explained the study procedure, participants were asked to start with either 3DoF or 6DoF session and end with the other. A set of related body gestures videos (same as those in Section 4) were presented in each session, and there is a drop-down text menu containing all 13 designed social intentions under each of them. Participants then selected the social intentions they believe that fit the video the most through the menu. We counter-balanced the order of 3DoF and 6DoF sessions across all participants. After the two sessions were finished, we revealed the answer to the participants and collected several qualitative feedback verbally, including:

(1) How do you recognize different social intentions?
(2) Are there any intention-gesture pair confusing with another? Why?
(3) Are there any intention-gesture easily recognize? Why?

5.0.3 Result. The result is shown as two confusion matrices in Figure 5, where the 6DoF condition has a noticeable higher accuracy than 3DoF across nearly all body gestures.

Overall, our body gestures succeeded in providing various required social intentions for daily scenarios, leaving a few gestures with confusions on "Feel Happy," "Say Goodbye," and "Dance."

Although the confusion here seems catastrophic for these target intentions at first glance, we believe they will perform better in an actual use case. For example, suppose users receive the "Say Goodbye" gesture on leaving the room. In that case, the intention can be more evident than watching the same body gesture on a white screen without an actual daily event happening. Another example here is that for performing "Dance," participants argue that without the actual music playing, it is hard to judge the cube's intention as music styles usually vary: "I listen to heavy metal music a lot, but the answer body movement is more like soft jazz, and thus I misunderstand it. (P6)," which also indicates the confusion might be addressed on actual usage.

Otherwise, most of the presented body gestures were answered as the desired target social intentions. Qualitative feedback includes "I can relate some of the videos as human movement, which serves as my reference (P4)" and "The rotation speed helps me separate different emotions (P9)." This suggests that our method succeeded in expressing various social intentions for daily scenarios.

6 BODY GESTURES IN DAILY OBJECT FORM

As the feasibility of our method has been shown, we integrate some of our 6DoF body gestures in the form of actual object shapes, including a pot, lamp, screen, keyboard, and printer in a virtual scene as shown in Figure 6. These daily objects perform body gestures on users moving forward, gazing at them, or touching specific colliders. We welcome anyone interested in our preliminary result to download the application scene from our shared drive [2].

[1] Designed body gestures playlist

[2] Download application scene

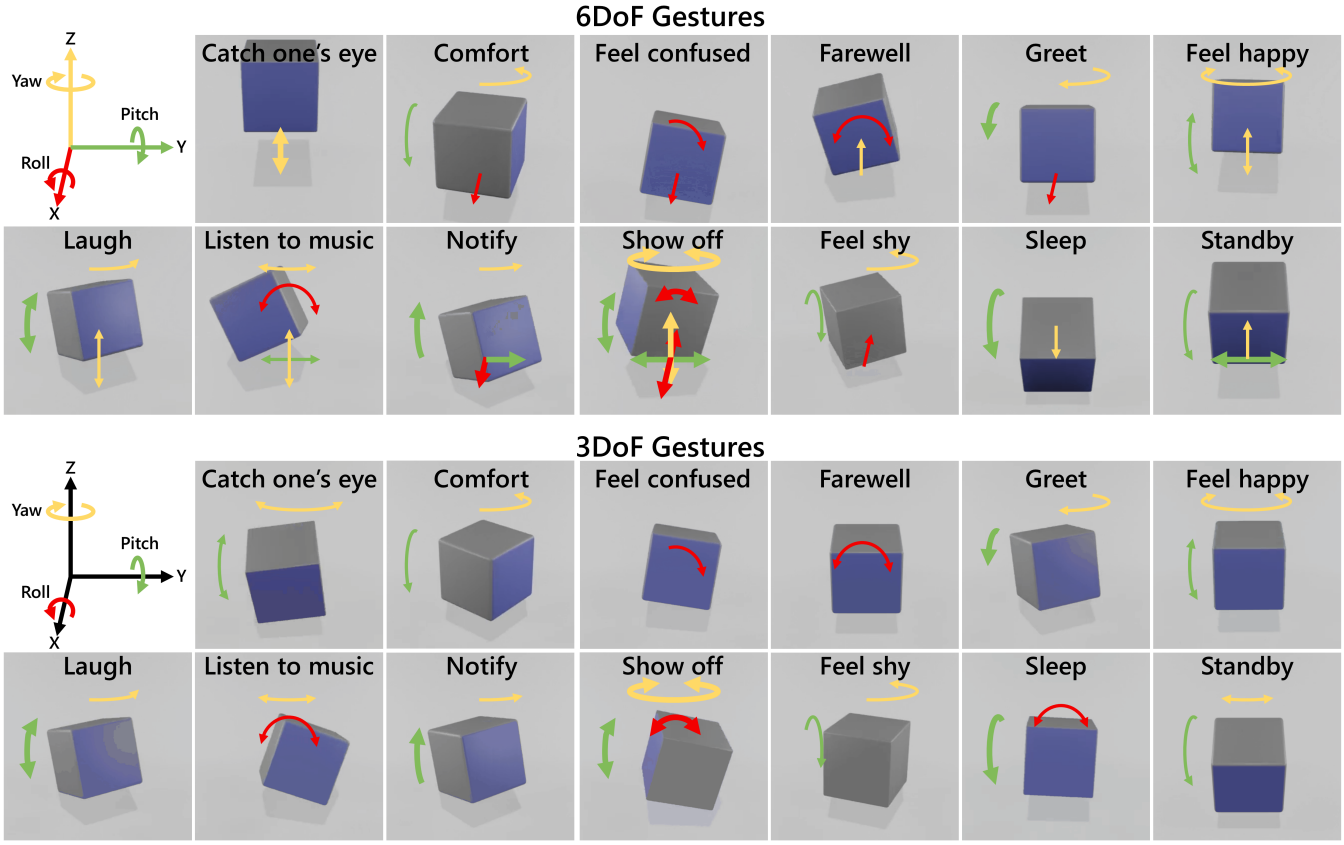

Figure 4: Body gestures designed for 13 intentions. Translation and rotation of the body gestures are marked as straight arrows and curved arrows, respectively. Three colors indicate the direction of translation or the axis of rotation, red for the x-axis, green for the y-axis, and yellow for the z-axis. There is two wideness of arrows, the wider one indicates a greater acceleration in movement, and the thinner one indicates a smaller one. 6DoF gestures are composed of both translations and rotations; 3dof gestures are only composed of rotations.

Figure 6: The Unity application scene where the pot, screen, keyboard, and printer interact with users through body gestures.

7 LIMITATIONS AND FUTURE WORKS

7.1 Influence of Pandemic

Due to the strict COVID-19 safety regulations, we have made some give-and-takes on our experimental methods amid our project, and we will leave a few of them here to suggest how some improvement can be made.

One apparent limitation is that the online-video study, the Biological Design Perspective in our Expression Survey, was once considered a VR setup. Participants role-play daily objects in an immersive environment and advise specific social intentions or gestures. Another limitation is validating the distinguishability of body gestures through animation instead of an actual motion platform with physical objects. Although our current method seems reasonable enough, the existence of actuators might still affect users' perceived gestures.

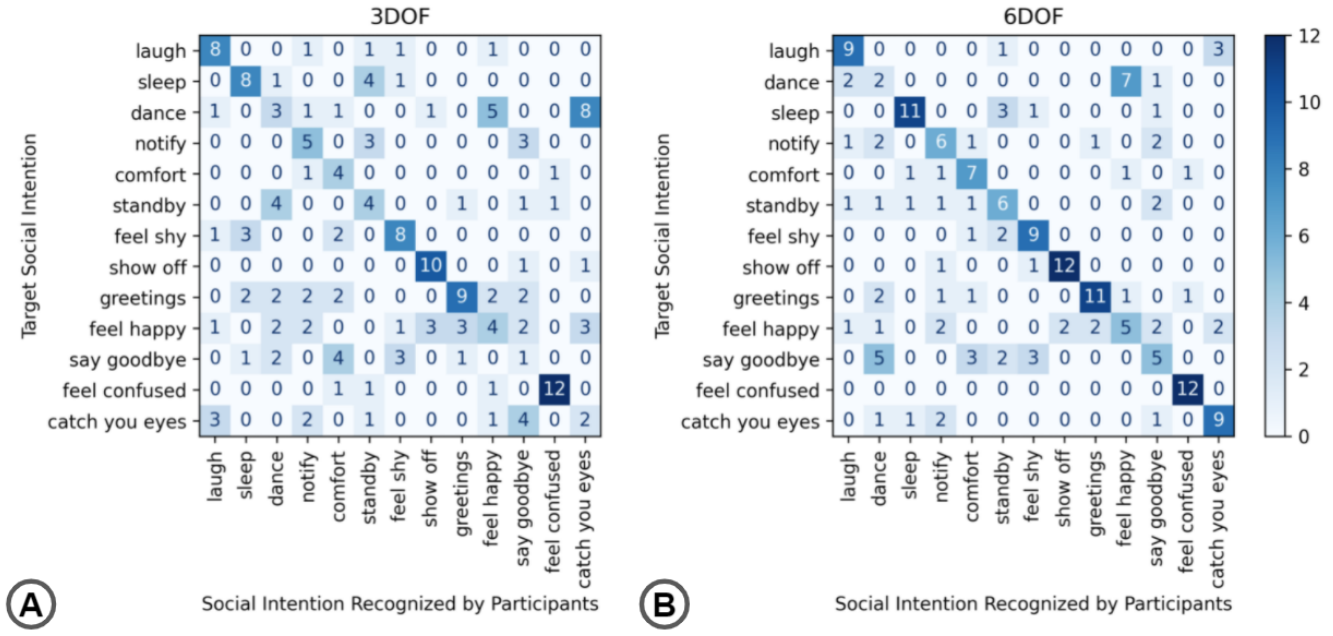

Figure 5: Confusion matrix derived from the results of the distinguishability experiment. Each row shows number of the actual target social intention answered as the test intentions.

7.2 Combination with Other Modality

Although body gestures on a daily object alone can generate sufficiently many expressions and possibilities, most daily objects (especially for electronic products) are usually compatible with more than one modality and output channel. As Rozendaal et al. [12] has shown an expressive lamp with different brightness, we are considering combining body gesture with screen light and voice feedback. Ultimately, the concept of a behavioral object or object with intents can be expressed through multimodal outputs.

7.3 Calibrate Hardware Platform

Ultimately, the project's goal is to apply body gestures within 6DoF on a daily object. The best way to implement it might be an end-to-end hardware system with sensing capacity and motion actuators on physical objects. While we currently have some promising results indicating the feasibility of our method, finding a proper sensing channel or paradigm and testing the performance of the motion platform still require tons of effort. Some possible candidates might include depth camera for users' body tracking [7] or NFC-based sensing system [11] for detection on motion platform.

7.4 To Generate Body Gesture

In our project, we adopt and collect body gestures based on human labor. However, as mentioned in [12], daily objects usually encounter interactions within considerable flexibility, and such flexibility might lead to more complicated events, exceeding the 12 daily scenarios we are currently testing with. While facial expression can be generated through generative adversarial network models, [17], we are exploring the possibility of generating body gestures through an automatic model. However, great efforts are awaiting such destinations, ranging from finding the training dataset to mappings between daily objects' skeletal constraints. Previous works on human-object mapping for animation and virtual body control [3] might shed light on a possible solution.

8 CONCLUDING REMARKS

In this paper, we explore the possibility of applying 6DoF body gestures to everyday objects. Firstly, we understand how people desire objects' gestures and intentions through 12 daily scenarios. Secondly, we design body gestures within 3DoF and 6DoF movements based on users' suggestions and prior work. Lastly, we conducted a 14-person study to measure the feasibility of these body gestures on expressing high-level intentions. Our result showed that the body gestures succeeded in providing the perception of various desired social intentions. Besides, the 6DoF movements indicate non-intrusive solutions for augmenting everyday objects without redesigning each of them.

REFERENCES

[1] Julian Angel-Fernandez and Andrea Bonarini. 2016. Robots Showing Emotions: Emotion Representation with no bio-inspired body. *Interaction Studies* 17 (12 2016), 408–437. https://doi.org/10.1075/is.17.3.06ang

[2] Samuel Bianchini, Rémy Bourganel, Emanuele Quinz, Florent Levillain, and Elisabetta Zibetti. 2015. *(Mis)behavioral Objects*. Springer International Publishing, Cham, 129–152. https://doi.org/10.1007/978-3-319-13018-7_8

[3] Jiawen Chen, Shahram Izadi, and Andrew Fitzgibbon. 2012. KinEtre: Animating the World with the Human Body. *UIST'12 - Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, 435–444. https://doi.org/10.1145/2380116.2380171

[4] EnsadLab EnsadLab. [n.d.]. https://misbkit.ensadlab.fr/misb-what/

[5] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 3 (2003), 143–166.

https://doi.org/10.1016/S0921-8890(02)00372-X Socially Interactive Robots.

[6] Google Google. [n.d.]. Nest Audio. https://store.google.com/product/nest_audio

[7] Thomas Helten, Meinard Müller, Hans-Peter Seidel, and Christian Theobalt. 2013. Real-Time Body Tracking with One Depth Camera and Inertial Sensors. *Proceedings of the IEEE International Conference on Computer Vision*, 1105–1112. https://doi.org/10.1109/ICCV.2013.141

[8] Jamy Li and Mark Chignell. 2011. Communication of Emotion in Social Robots through Simple Head and Arm Movements. *International Journal of Social Robotics* 3, 2 (01 Apr 2011), 125–142. https://doi.org/10.1007/s12369-010-0071-x

[9] Donald A Norman. 1994. How might people interact with agents. *Commun. ACM* 37, 7 (1994), 68–71.

[10] Jekaterina Novikova and Leon Watts. 2014. A Design Model of Emotional Body Expressions in Non-Humanoid Robots. In *Proceedings of the Second International Conference on Human-Agent Interaction* (Tsukuba, Japan) *(HAI '14)*. Association for Computing Machinery, New York, NY, USA, 353–360. https://doi.org/10.1145/2658861.2658892

[11] Daniel Rodriguez, Mohammad A. Saed, and Changzhi Li. 2021. A WPT/NFC-Based Sensing Approach for Beverage Freshness Detection Using Supervised Machine Learning. *IEEE Sensors Journal* 21, 1 (2021), 733–742. https://doi.org/10.1109/JSEN.2020.3013506

[12] Marco C. Rozendaal, Boudewijn Boon, and Victor Kaptelinin. 2019. Objects with Intent: Designing Everyday Things as Collaborative Partners. *ACM Trans. Comput.-Hum. Interact.* 26, 4, Article 26 (June 2019), 33 pages. https://doi.org/10.1145/3325277

[13] Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-Defined Motion Gestures for Mobile Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 197–206. https://doi.org/10.1145/1978942.1978971

[14] Martin Saerbeck and Christoph Bartneck. 2010. Perception of affect elicited by robot motion. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 53–60. https://doi.org/10.1109/HRI.2010.5453269

[15] Sjoerd Stamhuis, Hans Brombacher, Steven Vos, and Carine Lallemand. 2021. Office Agents: Personal Office Vitality Sensors with Intent. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 174, 5 pages. https://doi.org/10.1145/3411763.3451559

[16] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[17] Huiyuan Yang, Zheng Zhang, and Lijun Yin. 2018. Identity-Adaptive Facial Expression Recognition through Expression Regeneration Using Conditional Generative Adversarial Networks. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*. 294–301. https://doi.org/10.1109/FG.2018.00050