Capturing attention with wind

Natalie Friedman[1], David Goedicke[1], Vincent Zhang[1], Dmitriy Rivkin[1],
Michael Jenkin[1], Ziedune Degutyte[2], Arlene Astell[2], Xue Liu[1], Gregory Dudek[1]

***Abstract*—Having a robot interact with people in a shared environment is complex. Both running into humans and loud audio warnings are inappropriate. Visual signalling may be appropriate but is only effective if the humans are looking at/attending to the robot vehicle. Are there effective and socially acceptable mechanisms that a robot can exploit to capture the attention of humans in a shared environment? Here we explore the potential of using controlled blasts of wind (haptic air) to capture attention in a socially acceptable manner.**

I. INTRODUCTION

When the way forward is blocked by people, a robot must interact with bystanders to capture attention and communicate the desire for people to move out of the way. What is an effective mechanism to communicate this intent in a social setting? Through a set of "in the wild" experiments [8] at informal social gatherings (office parties) we evaluated different attention-seeking behaviours for a mobile robot. A Wizard of Oz (WoZ) methodology was used in an office/cocktail party setting in which participants had the impression that a robot was moving autonomously following a line on the floor, while it was actually being controlled remotely [5, 4]. This allowed the robot to safely avoid colliding with people and to prototype the system before building a fully autonomous system. Responses to three combinations of different attention-seeking modalities (haptic air, visual, audio) were collected during the social events. The haptic air cue was a directed wind energy event created by a fan on the robot, the visual cue asking people to move was presented via text on a tablet screen mounted on the robot, and the audio cue was a pre-recorded distorted voice emitted from the robot. Behavioural responses from participants in the experiment, and post-interaction questionnaire responses were used to gauge the effectiveness and social appropriateness of the different strategies.

II. METHOD

An in the wild study was used to investigate capturing attention and communicating intent in a social setting. Data was collected at social events over a time span of 12 days. All participants signed legally-approved consent declarations which did not divulge the mechanisms to be examined in the experiment.

Conditions 1 (no modality) and 2 (haptic air only) were collected 12 days apart from Conditions 3 (haptic and visual only) and 4 (haptic air, visual and audio). Between Conditions 1 & 2, and 3 & 4, there was a one hour technical presentation on topics distinct from this study. All conditions used the same robot.

Participants Participants were recruited through convenience sampling: e-mail lists and word of mouth. There were a total of N = 25 unique participants across the social events, of which N = 23 adults (4 female) completed the questionnaire afterwards, while N = 19 were behaviourally coded. The others were bystanders to the robot at the parties and did not interact with the robot. The participant ages ranged from 21 to 45 (M = 29) and on average, were halfway between intermediate and advanced in their experience with robots (M = 3.5, SD = 1.0), where 1 = Fundamental Awareness and 5 = Expert [6].

Behavioural Measures Each condition was monitored by six time-synchronized cameras (two on the robot and four GoPro's on the walls). The video sequence was broken down into a sequence of events. An *event* is defined as beginning when the robot stops because a person is in the way of the robot's path until the path is clear and the robot can move. Within the events, we measured the interaction time, the participant moving outcome, and if the engagement was constructive.

Subjective Measures We utilized the Interpersonal Dominance Scale [3] (used in HRI studies [7, 1, 2]) which measures perception of an actor's behavior along five dimensions: poise, persuasion, conversational control, panache, and self-assurance. Responses to individual questions were grouped into the five dimensions of the Interpersonal Dominance Scale.

The robot was augmented with three different interaction mechanisms: **Haptic mechanism** The robot was augmented with a computer controlled USB-powered fan that could provide an air-delivered touch to participants near the robot. The air fan was chosen as the delivery mechanism for a haptic touch given its relative safety compared to physical touch. **Audio mechanism** Speakers mounted on the robot were used to provide audio cues for two seconds. Audio cues were provided via a pre-recorded audio file stored on the robot. The audio cue was a higher pitched abstracted version of "excuse me" that could be heard 2m from the robot. **Visual mechanism** An onboard screen was used to intermittently display the text message "Can you move off the line so I can get through?"

Manipulations The robot relied on three basic interaction modalities; haptic air, audio and visual to cue attention in a crowded space. These basic modalities were integrated in the following test conditions. **Condition 1: No Modality** When the robot was blocked, it waited until the person moves out of the way. **Condition 2: Haptic** This event involved turning the fan on for five seconds and then turning it off for five seconds. This pattern was repeated a maximum of four times. **Condition 3: Haptic and visual** The wizard intermittently

---

*This work was supported by Samsung.

[1] Samsung AI Center, Montreal, Canada
[2] Samsung AI Center, Cambridge, UK

displayed the text message ”Can you move off the line so I can get through?” while the haptic air cue was provided. An event involved playing this text message for 10 seconds. When the message was not displayed, a blank screen was shown. For the first five seconds of this display the fan turned on providing a haptic air stimulus, and then turned off for the last five seconds. When this sequence was repeated, the visual display stayed on. The sequence was repeated a maximum of four times. **Condition 4: Haptic, visual and audio** Synchronized with the start of the audio cue, the haptic air and visual display described under Condition 3 were provided. This sequence was presented up to four times.

After five minutes, if the robot was not able to compel the person to move out of the way, either because it could not attract the person’s attention or because the person did not understand that they needed to move, the interaction was deemed a ”failure.” In such a case, if possible, the robot moved around the person and the robot continued on its way to the next interaction.

Procedure Each data collection session began with an informed consent process. Following this, we invited people to socialize and eat. For each condition, the robot was tele-operated at speeds up to 1m/s on the pre-determined path for 20-30 minutes. The robot followed a 31m continuous path that was marked on the floor using tape; this ensured consistency of the robot’s path across conditions. The robot tele-operator controlled the robot to follow this path during the social events as well as controlling the attention seeking mechanisms. After the completion of the route 3-4 times, the robot departed the scene. The actual number of cycles of the route depended upon the end of the social event, robot connection, or the robot having to wait for too long for a participant to move. At the end of each condition participants completed a questionnaire that probed their experiences while interacting with the robot at the party.

III. RESULTS

Behavioural Results Condition 1, No modality (C1) is essentially a control condition within which the robot makes no effort to engage with humans in the space and simply waits until the space is clear to move. Although there were examples in which the robot was successful in having people move out of the way, a much more common outcome was the experimenter labelling this motion a failure and tele-operating the robot around the obstruction. Given that there was a lack of prompted interactions from the robot, we do not report the reactions from the participants for C1 here. In C2, C3 and C4 after an interaction, if a person moved out of the way, it was counted as a success. If after the modality interaction, a person did not move out of the way, it was counted as a failure. There was a trend for the success rate of these interactions to increase, rising from 18% for C2 to 57% in C4.

Subjective Results Individual participant responses to the questionnaires that make up each questionnaire dimension of the Interpersonal Dominance Scale were averaged per condition. There was an increasing perceived self assurance, panache, conversational control, poise and influence of the robot as the complexity of the manipulations increased. The No Modality (C1) condition resulted in the lowest score across all five measures while the haptic air, visual and audio (C4) condition results in the highest score across all conditions.

IV. DISCUSSION

For a robot to navigate in congested social environments it must be able to capture the attention of individuals in the space and communicate its intention to move through that space. Here, we examined how well haptic, visual and audio cues might provide these tools. We found that as the number of interaction modalities increased, the interactions tended to be shorter and modalities were more effective. Additionally, participant’s perception of the robot showed increased self assurance, panache, conversational control and influence over the interaction. Adding additional modalities not only resulted in more effective and efficient interactions, it also potentially made the robot appear more socially effective.

REFERENCES

[1] V. Ahumada-Newhart and J. S. Olson, “Going to school on a robot: Robot and user interface design features that matter,” *ACM Trans. Comput.-Hum. Interact.*, vol. 26, no. 4, pp. 25:1–25:28, June 2019.

[2] J. K. Burgoon, J. A. Bonito, B. Bengtsson, C. Cederberg, M. Lundeberg, and L. Allspach, “Interactivity in human–computer interaction: A study of credibility, understanding, and influence,” *Computers in Human Behavior*, vol. 16, no. 6, pp. 553–574, 2000.

[3] J. K. Burgoon, M. L. Johnson, and P. T. Koch, “The nature and measurement of interpersonal dominance,” *Communications Monographs*, vol. 65, no. 4, pp. 308–335, 1998.

[4] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, “Wizard of oz studies - why and how,” *Knowledge-Based Systems*, vol. 6, no. 4, pp. 258–266, 1993.

[5] J. F. Kelley, “An iterative design methodology for user-friendly natural language office information applications,” *ACM Transactions Information Systems*, vol. 2, no. 1, pp. 26–41, Jan. 1984.

[6] U. N. I. of Health (NIH), “Competencies proficiency scale,” https://hr.nih.gov/working-nih/competencies/competencies-proficiency-scale, Accessed: 2019-07-20.

[7] I. Rae, L. Takayama, and B. Mutlu, “The influence of height in robot-mediated communication,” in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. Tokyo, Japan: IEEE Press, 2013, pp. 1–8.

[8] S. Savanovic, M. P. Michalowski, and R. Simmons, “Robots in the wild: observing human-robot social interaction outside the lab,” in *Proceedings of the IEEE international Workshop on Advanced Motion Control*, Istanbul, Turkey, 2006.