Recognition of a Robot's Affective Expressions under Conditions with Limited Visibility

Moojan Ghafurian[1], Sami Alperen Akgun[2], Mark Crowley[1], and Kerstin Dautenhahn[1,2]

[1] Department of Electrical and Computer Engineering
[2] Department of Systems Design Engineering
University of Waterloo, Waterloo, Ontario, Canada
{moojan, saakgun, mcrowley, kerstin.dautenhahn}@uwaterloo.ca

Abstract. The capability of showing affective expressions is important for the design of social robots in many contexts, where the robot is designed to communicate with humans. It is reasonable to expect that, similar to all other interaction modalities, communicating with affective expressions is not without limitations. In this paper, we present two online video studies (72 and 50 participants) and investigate if/to what extent the recognition of affective displays of a zoomorphic robot is affected under situations with different levels of visibility. Recognition of five affective expressions under five visibility effects were studied. The intensity of the effects was more pronounced in the second experiment. While visual constraints affected recognition of expressions, our results showed that affective displays of the robot conveyed through its head and body motions can be robust and recognition rates can still be high even under severe visibility constraints. This study supported the effectiveness of using affective displays as a complementary communication modality in human-robot interaction in situations with visibility constraints, e.g. in the case of older users with visual impairments, or in outdoor scenarios such as search-and-rescue.

Keywords: Affective displays; impaired visibility; social robots; zoomorphic robots; Miro

1 Introduction

Emotional expressions have been designed on many different robots [9, 51, 44, 14] and using different modalities, such as, facial expressions [9, 44], light [11], sound [41], body gesture [19], motion [46], and other more abstract methods [32]. Each of these methods (or a combination of them) can be used to improve emotional displays of a robot and increase its social capabilities. This capability is especially beneficial in many context where humans interact directly with robots, such as in health and safety applications.

It is important to ensure that emotions expressed by a social robot can be perceived reliably by humans who are interacting with the robot or working alongside it (e.g., in healthcare or search and rescue applications of robots). Otherwise, robots that lack clearly identifiable expressions of emotions may be socially unacceptable to their users [6], and might be unable to sustain successful social interactions.

Sometimes end users are less familiar with robotics technologies, such as children and older adults, or medical staff. For such cases, expressions of emotions that can be perceived reliably becomes more crucial since those users often assume that the emotions expressed by the robot are real and genuine [27, 43]. In addition, many real-life applications of emotionally expressive robots include scenarios with limited visibility that might affect perception of emotions [23], possibly related to the different visual capabilities of different user groups (e.g., children [5] or older adults [16]). Also, depending on the specific application area, the environment in which the robot operates might introduce conditions that can result in limited visibility such as in search and rescue [55], firefighting [30], or service robotics [20].

Misinterpreting a robot's emotional displays can have a significant impact on the quality of the interaction itself, as well as the acceptability and usability of the robot across a variety of different application areas and user groups [34, 6]. Therefore, here, we investigate how situations limited by visibility constraints affect recognition of a robot's affective expressions. We investigated the robustness of recognition of affective expressions of a robot expressed through its head and body gestures. Our choice of robotic platform is the Miro robot, a zoomorphic robot designed not to look like any specific animal, but which can express emotions via many modalities, such as body language [12, 36]. This zoomorphic robot was selected since (1) animal-like robots are widely used in many applications in health (e.g., therapeutic situations and with older adults) and safety (e.g., search and rescue scenarios), and (2) a recent study designed and evaluated different affective displays for the Miro robot, all of which involved head and body gestures which are of interest in the present study [19].

In the following, we will first detail our research questions followed by a discussion of background and related work. Two experiments and their results are presented afterwards. We conclude with a discussion of the results and identifying limitations of this work.

Research Questions and Expectations: We investigate the robustness of human observers' recognition of affective expressions of a zoomorphic robot (including emotions and moods) conveyed through its body language, under different visibility conditions. Our specific research questions are as follows.

RQ1 To what extent is the recognition of a robot's affective expressions robust with regards to the visibility of the robot?

RQ2 How does the degree of visibility (moderate versus severe) impact the recognition of a robot's affective expressions?

Concerning RQ1, since to our knowledge this is the first study of its kind on this particular topic, RQ1 was exploratory in nature. However, we expected that some affective expressions might be easier to recognize than the others under the same visibility constraints. Regarding RQ2, we found it reasonable to expect that severe constraints on visibility will have a larger impact than moderate constraints. And again, we expected that some affective expressions might be more affected than others.[3]

2 Background and Related Work

In many application areas, social robots may operate in an environment that introduces constraints. For example a noisy environment can make sound less effective, or an environment with limited visibility can affect visibility of the robot's expressions, such as its facial expressions. These situations are very common in robots that assist in search and rescue (in which emotions have been recently proposed as a new modality [1]) or fire fighting missions. In these missions many conditions such as rain [28, 50], smoke [7], limited available light [7], or even temperature [22] can affect visibility of robots. Further, some users, such as older adults may be affected by a condition that can influence their vision, such as blurred vision. For robots to be successful in these situations, it is important to ensure that both robots and users can continue to communicate successfully despite these constraints.

From the perspective of robots, many studies have investigated improving performance of a robot in situations with limited visibility, for example in smoky environments [45, 49], dark areas [13], underwater [24], or to detect humans at different distances [39]. To this end, different machine learning methods, computer vision techniques, as well as sensors have been proposed for the robots to enhance their navigation and detection of humans in situations with similar visibility constraints [45, 39, 49, 13, 38]. For instance, Na et al. suggested using smart gloves to overcome limited visibility problem in water for gesture recognition [38].

However, visibility of emotional expressions of robots by humans is also important for the success of interactions between humans and robots, a research area that has so far attracted little attention, but will ultimately determine to what extent emotional expressions can be used successfully in real life situations. In general, visibility of robots has been studied in many situations where humans need to locate a robot [31]. For example, with the goal of designing a socially assistive service robot, McGinn et al. used lights to improve visibility of the robot and suggested that while it is commonly believed that a social robot with a head/face is most effective for interactions, having a head can limit the observation angle and result in a poor visibility of the robots' expressive states [35]. This can be particularly true if the robot's expressiveness is restricted to facial expressions. However, an approach that uses both head and body gestures to convey expressions could be beneficial in overcoming this challenge. However, even when the observation angle is not an issue, there are still other visual constraints as mentioned above that can affect the visibility of a robot's expressions.

[3] Note, we had originally envisaged to conduct both studies as in-person experiments, but due to COVID-19 this was not possible and we had to move the study online.

Therefore in this article we ask how such situations affect recognition of a robot's affective expressions.

3 Experiment 1

In this experiment we evaluate the recognition of five affective expressions of the Miro robot under five different situations that *moderately* affect visibility through a crowdsourcing video study, which is a common approach in many of the related studies in human-robot interaction (e.g., [9, 32, 18]).

3.1 Methodology

Affective Expression Selection As discussed earlier, we study the affective expressions of the zoomorphic Miro robot. For this particular robot, prior work designed and evaluated eleven affective expressions (including emotions and moods) [19], inspired by various sources of information, including ethology, cartoon animations, etc. These affective expressions were conveyed through the robot's eyes, ears, tail, and body movements. The Miro robot has the potential to be used as a companion robot for older adults for applications such as pet therapy, and the emotional expressions (used in this study) have the potential to be implemented on the other zoomorphic robots. Further, in other applications, such as in search and rescue, body and head motions could be effective for communicating emotions, because these motions can, to some extent, be applied to appearance constrained robots used, e.g., in search and rescue [8].

Out of the proposed 11 affective expressions of Miro, we selected expressions that (1) were highly recognized by the majority of the participants who evaluated these expressions in a previous study [19], and (2) were found to be important in situations where the robot's visibility can be affected, such as in search and rescue situations [2]. This resulted in the selection of five affective expressions: happy, sad, tired, excited, and surprised.

Video Effect Selection After careful consideration of different situations that can affect the visibility of a robot, five visual effects were selected: dark, rain, smoke, blurred vision (called blur in this paper), and the robot located at a far distance (called zoom hereafter). Here are some real-life situations with limited visibility that can be represented by the selected video effects:

– **Dark:** Other than many common situations where the robot may operate in a dark room or a room with dimmed lights, during almost all types of search and rescue operations, the rescue might be conducted under diminishing light conditions. During wilderness search and rescue [21], maritime/sea search and rescue [37], or mountain rescue [28], the reason for darkness might just be related to the operation time (i.e., the missions need to start or continue at night). On the other hand, for other search and rescue situations such as urban search and rescue [3] or cave rescue [25], areas of interest can be dark

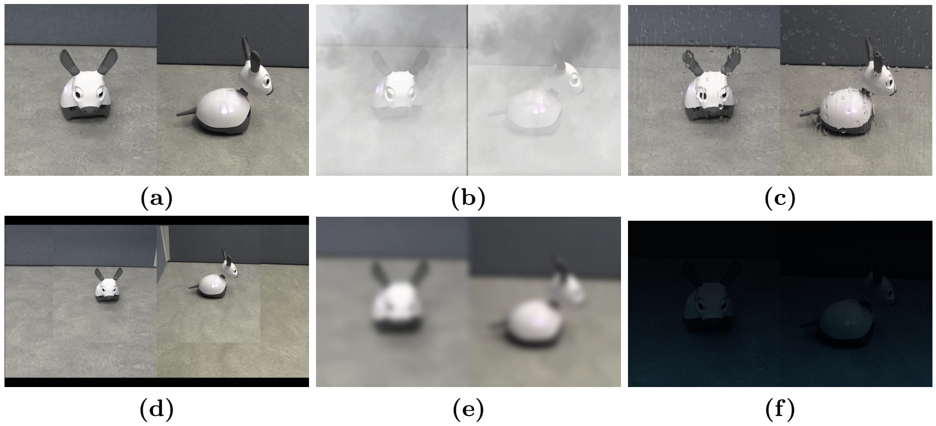


Fig. 1. Snapshots taken from the start of the tired expression for (a) actual videos used in training, (b) smoke effect, (c) rain effect, (d) zoom effect, (e) blur effect, and (f) dark effects

regardless of the operation time, since these types of rescues involve confined spaces. This can be the case for firefighters, too, since they sometimes have to operate in dark areas [7]. Therefore, this effect can represent many situations in which the robot is used to increase safety.

- **Smoke:** This effect, similar to dark, mostly represents situations in which the robot is used to increase safety. In general, smoke is the main factor that firefighters need to deal with [7]. Further, the areas involved in search and rescue situations might be affected by the smoke caused by either fire at the scene [10] or by dust of collapsed structures [33]. This effect also represents situations where the rescue workers need to deal with dense fog in open areas like mountain rescue [47].
- **Rain:** Similar to dark and smoke effects, heavy rain (and, with a similar effect, snow) can introduce many different challenges in search and rescue situations (for example in mountain search and rescue situations [28, 50]). Limited visibility due to rain is present in many maritime search and rescue operations since most of the time rescue workers need to operate either in or near to water [37].
- **Blur:** This effect can represent situations happening both in robots designed in the context of healthcare and those designed for safety. For example, blurred vision (or a similar condition) can affect many older adults, as well as other individuals, and constrain the visibility of the robots. Furthermore, a blurred vision may happen in firefighting scenarios, or search and rescue scenarios due to high humidity/temperature [22] or low-lighting conditions [42].
- **Zoom:** As many of the social robots may be used in large areas, for example in a care home or in a search and rescue scene, it is expected that the robot will not always remain close to its users. For example, a companion robot in a care home may be farther from some residents/caregivers and closer to the

others. Also, while rescue robotics involves situations that requires co-located interactions between humans and robots [15], the distance between the human and robot teammates varies significantly depending on the search area (e.g., they may need to operate at very distant locations during wilderness search and rescue [26] or forest fire missions [29]).

All the videos were adjusted to have the same length. The videos were modified in a way that each effect was identical for all the affective expressions. Figure 1 shows a snapshot of these effects.[4]

Design and Procedure Upon signing the consent form, the participants followed four steps of the study as below.

Step 1 - Training: The participants first saw a video of all five expressions of the robot, in which the affective expressions were clearly stated overlaying text on the video. The purpose of this step was to teach participants about the affective expressions of Miro, and to ensure that they would know each expression and its label under clear visibility conditions. The participants were told that this was a training part and we wanted them to learn the meaning of each expressions. They were able to replay this video as many times as they wished.

After watching the training video (as many times as needed), we tested the success of the training, i.e., participants' ability to recognize the expressions, to ensure that they paid sufficient attention to the video learned the five expressions. The expressions were shown to them one by one, and they were asked to indicate what each expression was. If they passed a test for each expression, that expression was removed from the testing set. Otherwise, it remained in the test set and we showed them that expression again at a later time and in a random order. This continued until the participants were either able to complete training successfully, or if they exceeded "15" test videos (i.e., watching each video three times on average). In the latter case, participants were allowed to continue with the study but their data was removed since it indicated that for those participants training was unsuccessful.

This step was essential, because (1) while these affective expressions were previously designed and their recognition was evaluated correctly in the majority of instances, it is not reasonable to assume that all participants would recognize the affective expressions 100% correctly, and (2) without training, and thus ensuring that all participants will recognize the affective expressions correctly when visual constraints are not present, it is not possible to know whether any recognition errors observed in our study may be due to individual participant's difficulties in evaluating those affective expressions (which is not the focus of this study) or is due to visual constraints (which is the focus of this study). Furthermore, in

[4] A combination of widely available video editing software (i.e., iMovie, Wondershare Filmora, and HitFilm Express) was used to create the video effects. The choice of software depended on the effect. Please contact the authors if you are interested to see the videos.

many contexts, such as search and rescue, it is expected that those interacting with the robots will get prior information and training regarding the interpretation of the robots' affective expressions, if they are used as a communication modality.

Step 2 - Evaluating Video Effects: After the training phase, participants read instructions about the next phase, and they were informed that the videos are intentionally not as visible as those that they saw in the training phase, in order to avoid the remote participants to worry that reduced visiblity of the videos might be due to internet issues, etc.

In this step, the participants evaluated 10 videos with different effects. They were allowed to only select one option as their response. The first 5 videos were selected in a way that both the order of expressions (EM_1 to EM_5) and effects (EF_1 to EF_5) were random. Afterwards, they saw another combination as below:

– The order of effects were similar to what they originally saw
– The order of affective expressions were the same as before, except for one shift, moving the first expression to the last.

In other words, participants first saw $[EM_1, EF_1]$ to $[EM_5, EF_5]$, where EM_n and EF_n were randomly selected from the set of five expressions (i.e., happy, sad, surprise, excited, and tired) and effects (i.e., rain, smoke, blur, dark, and zoom). After completing these five combinations, the participants then saw $[EM_2, EF_1]$, $[EM_3, EF_2]$, $[EM_4, EF_3]$,$[EM_5, EF_4]$, $[EM_1, EF_5]$ (a total of 5 videos). This way, (a) all participants saw 10 videos, (b) evaluated each effect and each expression twice, and (c) the combinations (i.e., order of effects, affective expressions, and the combination of effect+expression) were counterbalanced. As it is suggested by previous work (e.g., see [17]) and to avoid fatigue, we did not expose each participant to all 25 videos.

Step 3 - Test: After evaluating the 10 videos, we tested participants to see how well they recalled the original expressions, to ensure that the errors in the recognition of the affective expressions were actually due to the video effects, as opposed to the participants forgetting the expressions. In this step, participants saw the original 5 videos that they saw in the training phase, and were asked to select the correct expression.

Step 4 - Questionnaire: A Questionnaire was designed to obtain information on participants' demographics, as well as their perception of the effects and their familiarity with pets and emotions (which are shown to affect recognition of a zoomorphic robot's expressions [19]). The questionnaire asked about participants' (1) age, (2) gender, and (3) ethnicity. We also included questions asking participants (4) if they had pets, (5) if they had a friend with a pet, (6) if they wear glasses or contact lenses, and if yes, if they were wearing them at the time of the study, (7) whether they liked pets, (8) whether they were scared of pets, (9) how good they thought they are in recognizing pets' emotions, and (10) if they had difficulty understanding pets emotions. Finally, participants were asked to report how difficult it was for them to understand Miro's expressions under each of the five effects (questions 11-15).

Questions 7 through 15 were responded to on a continuous scale. Questions 9 and 10 were designed to complement each other, as an attention check. One additional attention check question was added, which asked participants if they thought that cats and dogs were popular pets in North America (we assumed that the correct answer to this question would be YES and removed the data from those whose answers were anywhere from the centre of the continuous scale to "Completely Disagree").

Step 5 - Human Emotion Understanding: As a previous study suggested that understanding human emotions can be correlated with understanding Miro's affective expressions [19], in this step we asked participants to evaluate 14 images of human emotions (i.e., angry, fearful, disgusted, happy, neutral, sad, and surprised; two images of each). All images were taken from the FacesDB dataset.[5] Gender, age, and skin color were counterbalanced in the two images representing the same emotion, as well as among all the 14 images. In order to avoid fatigue we decided on 14 images. This step was added as a previous study showed that the recognition of affective expressions of the selected robot can be significantly affected by the ability to recognize human emotions [19].

All the steps were performed on the same day and one right after the other.

Participants: A total of 116 participants who had an approval rate of over 97% based on at least 100 HITs completed the study on Amazon Mechanical Turk. Data from 37 participants were removed as they failed the attention checks. This left 79 participants (44 male, 34 female, 1 unknown; age range [19,58], average: 35.2 yrs). The study received full Ethics clearance from the University of Waterloo's Research Ethics Committees.

4 Results - Experiment 1

The results for recognition of each affective expression under each effect is shown in Table 1 and Figure 2 summarizes the results. The correct expression was selected significantly more than the others for all expressions and effects. The accuracy was very similar for all the effects and were either 86% or 87%. We also studied the accuracy in recognition of each affective expression. While all accuracies were high (82% to 92%), the lowest accuracy belonged to surprised (82%), and the highest to sad (92%). It is in fact reasonable for surprise to have a lower accuracy when visibility is affected, since the associated behavior is not as continuous as in the other expressions and the behavior change is fast and sudden.

To study how other factors affected recognition of expressions, we fit a linear model predicting the number of correctly recognized expressions of Miro based on (a) participants' understanding of human emotions measured through the images, (b) their reported difficulty in understanding pets' emotions, (c) their accuracy in the post-test results, and (d) how long training took for them. Results are shown in Table 2.

[5] http://app.visgraf.impa.br/database/faces/

A higher reported difficulty in understanding pets' emotions significantly and negatively affected recognition of Miro's expressions ($se = 0.001, t = -2.077, p < .05$). Participants' performance in the post-test was also a significant predictor for recognition of expressions: as expected, those who performed better in their test results had a better recognition of Miro's affective expressions under different visibility constraints ($se = 0.150, t = 7.046, p < .0001$). Performance in training was also a significant predictor of the correct responses

Table 1. Results for the recognition of each affective expression for each visibility effect. The last column shows the overall accuracy for each effect. The last row shows the overall accuracy for each affective expression. Through binomial tests, we show if each choice was selected significantly more than the others: ***: $p < .001$, **: $p < .01$. The differences in number of results for affective expressions are due to (1) the random assignment of sets of effects+emotions, and (2) removal of data from those who failed the study.

		Happy	*Sad*	*Surprised*	*Tired*	*Excited*	*Accuracy*
Rain	Happy	**22*****	0	0	1	3	**86%**
	Sad	0	**29*****	0	3	0	
	Surprised	4	0	**26*****	2	3	
	Tired	0	5	0	**28*****	0	
	Excited	1	0	0	0	**31*****	
Blur	Happy	**35*****	0	0	1	3	**87%**
	Sad	0	**27*****	0	1	0	
	Surprised	1	2	**22*****	0	4	
	Tired	0	5	0	**35*****	0	
	Excited	3	0	0	1	**18*****	
Dark	Happy	**28*****	1	0	0	1	**86%**
	Sad	0	**39*****	0	4	0	
	Surprised	0	2	**30*****	1	1	
	Tired	0	3	0	**18*****	0	
	Excited	3	1	5	0	**20****	
Smoke	Happy	**28*****	0	3	0	2	**87%**
	Sad	0	**25*****	0	1	0	
	Surprised	1	1	**24*****	2	3	
	Tired	0	4	0	**26*****	0	
	Excited	2	0	2	0	**34*****	
Zoom	Happy	**26*****	0	1	0	3	**87%**
	Sad	0	**26*****	0	3	0	
	Surprised	0	1	**27*****	1	0	
	Tired	0	5	0	**27*****	1	
	Excited	4	0	1	0	**32*****	
	Accuracy	88%	92%	82%	85%	85%	

(a) (b)

Fig. 2. Overall results for (a) recognition of each effect, and (b) recognition of each affective expression

Table 2. Linear model predicting the number of correct responses to the affective expressions.

Covariate	**Estimate**	**SE**	**t**	**Pr** $(>\|t\|)$
Intercept	7.078	1.328	5.331	$< .0001$
humanUnderstanding	-0.075	0.101	-0.750	0.455
petEmotionDifficulty	-0.001	0.001	-2.077	$< .05$
test	1.057	0.150	7.046	$< .0001$
training	-0.276	0.069	-4.028	0.0001

$(se = 0.069, t = -4.028, p < 0.0001)$, the faster the participants finished the training (i.e., less errors they had), the better they performed in recognizing Miro's expressions in the videos with effects. This can be because (a) Miro's expressions were more intuitive for those who learned them quicker and finished faster, thus they recognized them better also in other conditions, or (b) those who finished the training faster paid more attention to the study in general and Miro's expressions shown in videos in particular. We did not observe any effect of participants' age or gender on recognition of the affective expressions.

Finally, as the difficulty level of the effects can affect recognition of affective expressions, we asked participants to report how difficult it was to recognize each affective expression under each of the five effect conditions. Figure 3 shows the results. Recognition was reported to be relatively easy for all the effects. After controlling for the other effects (i.e., using linear models and for the factors shown in Table 2), only for the rain effect the reported difficulty level was a predictor of the accuracy of recognition of the expressions $(se = 0.000, t = -3.308, p = .001)$. We did not observe any effect of the reported difficulties on the recognition accuracies for the other visibility effects.

5 Experiment 2

In this experiment we investigated how the recognition of the affective expressions changes in extreme situations, when visibility of the expressions are *severely* affected.

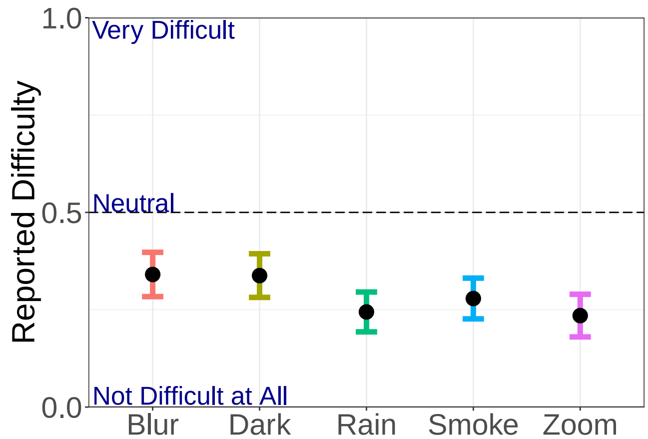
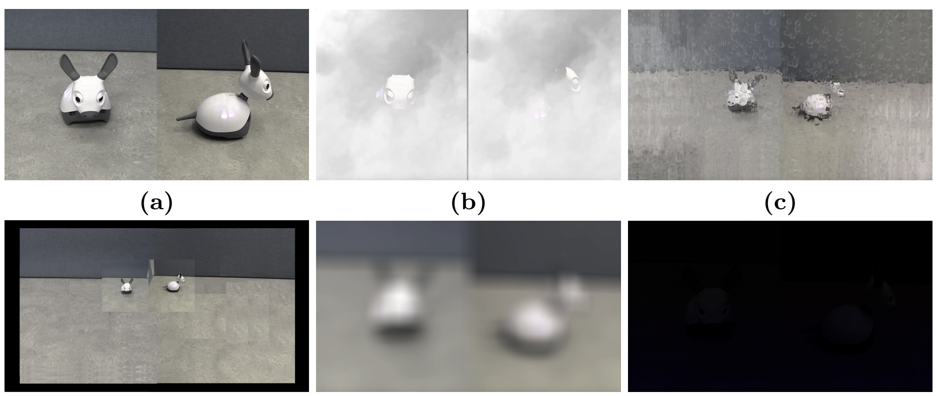


Fig. 3. Reported difficulty of each effect on the recognition of the robot's affective expressions



(a) (b) (c)
(d) (e) (f)

Fig. 4. Snapshots taken from the start of the tired expression for (a) actual videos used in training, (b) smoke effect, (c) rain effect, (d) zoom effect, (e) blur effect, and (f) dark effect

5.1 Methodology

The procedure and all the steps of this experiment were similar to Experiment 1. Participants followed the exact same five steps, except that new videos with more extreme effects were used and, as our intention was to use severe constraints on visibility in this experiment, as compared to moderate effects in Experiment 1, participants were given specific instructions to adjust their screens' brightness prior to evaluation of Miro's expressions in the videos.

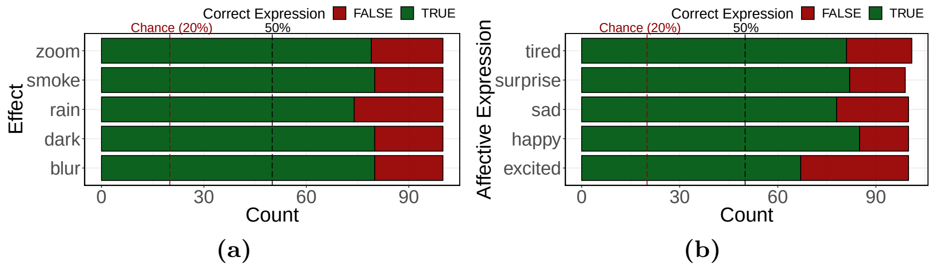


Fig. 5. Overall results for (a) recognition of each effect, and (b) recognition of each affective expression

Video Effects Considering that participants will use different computers, screens, settings, etc., and to ensure that the videos are suitably manipulated and reflect extreme situations, we did two pilot tests. The first test was conducted by asking five people in the authors' research group (who were not involved in this study) to evaluate different effects. Three extreme versions of the effects (changing in intensity) were created and the pilot group were asked to select the video which they thought was most extreme, but they could still see Miro. According to this feedback, five videos were selected and used in the second pilot study. The second pilot study was conducted on Amazon Mechanical Turk, where we asked seven participants to complete the experiment with the new videos. We then referred to the difficulty levels that the participants reported for each effect. As the difficulty levels were initially not as high as expected, the videos were further manipulated and the effects were intensified. Figure 4 shows snapshots of the final videos that were used in this experiment.[6]

Participants: 92 participants were recruited on Amazon Mechanical Turk. The same qualification criteria as used in Experiment 1 were applied (i.e., participation was limited to North America, and those with an acceptance rate of over 97% based on at least 100 HITs). Data from 42 participants was removed as they failed the attention checks (attention checks were identical to those in Experiment 1). This left 50 participants (35 male, 15 female; age range [21,61], average: 34.7 yrs). The study received full Ethics clearance from the University of Waterloo's Research Ethics Committees.

6 Results - Experiment 2

Results for the new, more extreme effects are shown in Figure 5 and Table 3. While the accuracies were still high, in three situations (i.e., Rain-Excited, Dark-Excited, and Smoke-Sad) the correct expression was no longer selected significantly more than the other options, and in the Rain-Excited situation it was not

[6] Please contact the authors if you are interested to see the videos.

Table 3. Results for the recognition of each affective expression for each visibility effect. The last column shows the overall accuracy for each effect. The last row shows the overall accuracy for each expression. Through binomial tests, we show if each choice was selected significantly more than the others: ***: $p < .001$, **: $p < .01$, *: $p < .05$. The differences in the number of results for affective expressions are due to (1) the random assignment of sets of effects+emotions, and (2) removal of data from those who failed the study.

		Happy	***Sad***	***Surprised***	***Tired***	***Excited***	***Accuracy***
Rain	Happy	**16****	1	0	2	3	**72%**
	Sad	1	**21*****	0	4	0	
	Surprised	1	0	**18*****	1	0	
	Tired	0	3	1	**15****	1	
	Excited	**7**	0	0	0	4	
Blur	Happy	**14****	0	1	1	3	**80%**
	Sad	1	**17****	0	3	1	
	Surprised	1	1	**13*****	0	0	
	Tired	0	1	0	**15*****	1	
	Excited	5	1	0	0	**21****	
Dark	Happy	**21*****	0	1	0	0	**84%**
	Sad	0	**15****	0	2	0	
	Surprised	2	0	**16*****	0	1	
	Tired	2	1	0	**19*****	0	
	Excited	2	1	3	0	**9**	
Smoke	Happy	**18*****	0	1	0	1	**80%**
	Sad	1	**10**	0	5	0	
	Surprised	1	0	**21*****	0	2	
	Tired	0	3	1	**16****	0	
	Excited	3	0	2	0	**15****	
Zoom	Happy	**16*****	0	0	0	1	**77%**
	Sad	1	**15****	0	2	0	
	Surprised	1	3	**14****	1	1	
	Tired	1	1	0	**16*****	1	
	Excited	8	0	0	0	**18***	
	Accuracy	85%	79%	81%	83%	68%	

even selected significantly more than random. The accuracies in recognition of the expressions dropped from the previous experiment, with excited having the largest drop and the least accuracy among the other expressions.

A linear model with the same dependent variables as the previous experiment (i.e., human emotion understanding results, reported difficulty in understanding pet emotions, training results, and test results) was fit to predict the number of correct expressions recognized by the participants. Two effects were consistent with the previous experiment: a higher test performance resulted in a better recognition of Miro's expressions ($se = 0.266, t = 3.079, p < .01$) and the longer it took the participants to learn the expressions in the training phase, the lower their recognition rates were ($se = -0.014, t = -3.539, p < .001$). The results are shown in Table 4.

Finally, Figure 6 shows the results for the reported difficulties of the effects. The effects were perceived to be more difficult than those in Experiment 1. However, after controlling for the other effects (i.e., using linear models and for those factors shown in Table 4), reported difficulty levels were not significant predictors for the accuracy of recognition of expressions.

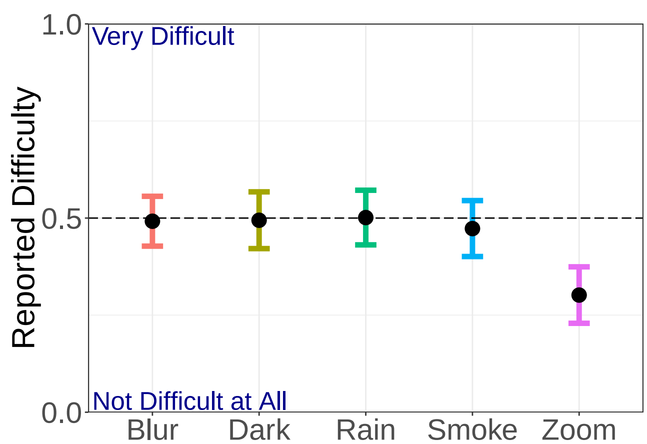


Fig. 6. Reported difficulty of each effect on the recognition of the robot's affective expressions

7 Results - Combined

To study the effect of the intensity of the video effects (moderate vs. severe) on the recognition of the robot's affective expressions, and to check that the difficulty levels changed indeed to more severe conditions in Experiment 2, we pooled the data. Note that while intensity of the effects was not intended to be a condition, the two studies were conducted very close and within 1 month

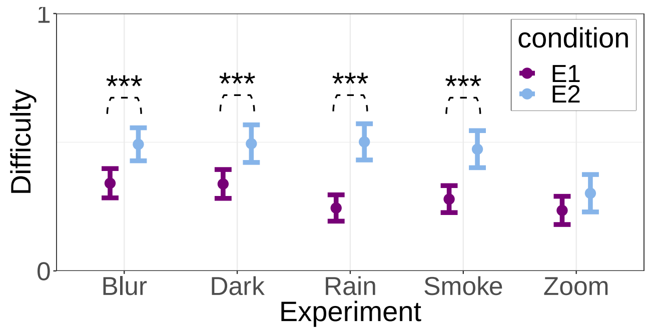


Fig. 7. Comparison between the reported difficulty levels in Experiment 1 and Experiment 2. 95% confidence intervals are visualized. $***$: $p < .001$. Significance was calculated using linear models that took other confounding factors such as results related to human understanding into account.

of each other, and anything except for the intensity of videos were identical in both. Therefore, pooling the data and doing a between-participants condition would be valid in this situation.

Table 5 shows the results. The recognition rates were significantly lower in Experiment 2, and experimental condition (i.e., the severity of the visibility effects) was a significant predictor of the correctly recognized expressions ($se = -0.937, t = -3.477, p < .001$).

Table 4. Linear model predicting the number of correct responses to the affective expressions.

Covariate	**Estimate**	**SE**	**t**	**Pr** $(> \|t\|)$
(Intercept)	5.916	2.147	2.755	$< .01$
humanUnderstanding	0.153	0.170	0.903	0.371
petEmotionDifficulty	0.000	0.001	0.262	0.794
test	0.819	0.266	3.079	$< .01$
training	-0.414	0.117	-3.539	$< .001$

Table 5. Linear model predicting the number of correct responses to the robot's affective expressions under different visibility conditions. Condition shows whether data belonged to Experiment 1 (E1) or Experiment 2 (E2).

Covariate	**Estimate**	**SE**	**t**	**Pr** $(> \|t\|)$
Intercept	6.801	1.171	5.806	$< .0001$
humanUnderstanding	0.033	0.0885	0.375	0.708
conditionE2	-0.937	0.270	-3.477	$< .001$
petEmotionDifficulty	-0.001	0.001	-1.023	0.308
training	-0.323	0.061	-5.246	$< .0001$
test	0.952	0.136	6.982	$< .0001$

The video effects were also rated as significantly harder in Experiment 2 ($se = 19.8632, t = 8.314, p < .0001$). Comparing the pair of effects, Rain, Blur, Dark, and Smoke were all rated to be significantly more difficult in Experiment 2 as compared with Experiment 1. However, while the robot's size (shown in the video) was halved from Experiment 1 to Experiment 2, the difference between the difficulty ratings was not significant for zoom (See Figure 7).

8 Discussion

This study investigated the feasibility of using a zoomorphic robot's affective expressions as a communication method in situations where users' vision or visibility of the robot are affected. Two experiments investigated recognition of five affective expressions (happy, excited, sad, tired, and surprised) under five situations with visual constraints (blurred vision, rain, far distance, smoke, and dark). The intensity of effects differed across the two experiments and were significantly intensified in the videos shown in Experiment 2. The two experiments helped us understand how visibility constraints affect recognition of emotions, while we controlled for the other confounding factors that may affect recognition in general.

Although, as expected, the accuracies dropped when the effects were intensified, the results suggested that the recognition of the affective expressions, as conveyed through head and body gestures of a zoomorphic robot, can be relatively robust to visual constraints, once the observer has received training in recognizing those expressions when they are clearly visible. This could be a great advantage in many application areas of HRI, as other communication methods may not be as effective under similar visual constraints. For example, text displayed on a screen mounted on the robot might not be readable if visibility is impaired, especially in conditions with severe effects (e.g., those in Experiment 2). Communication through voice can be affected by hearing impairments (e.g., in social robots used for older adults or situations with a high background noise that occur in search and rescue situations). If we had only relied on facial expression, e.g., by using a robot with a large repertoire of facial expressions (as are used in many humanoid robots), recognition accuracies might have been much lower since in many of the visibility conditions the face of the robot was not even visible.

Thus, our results highlight that affective expressions of a social robot, which are conveyed through its head and body gestures, can be an effective modality to improve interactions with a social robot in many application areas affected by low visibility (as well as other situations such as when hearing is impaired).

Among the five expressions, recognition of excited was affected the most when the effects were intensified, and it was mostly confused with happy: an expression that involved similar body movements, but executed with a lower speed. While future work is needed to understand why this happened, one explanation could be that intensity of the effects can obscure more frames of the video, thus reducing the number of frames in which the participants can see the robot and its movements (similar to motion illusions that happen in real scenarios [53]), and therefore, the sense of speed might have been affected as a result of those obscured frames. This could suggest that affective expressions differing only in speed can be mostly affected by visual constraints, and the one with faster motions is likely to be confused with an expression that has similar movements but at a lower speed.

Interestingly, for the more severe effects (Experiment 2), we did not find the reported difficulty to be a predictor of accuracy in recognition of the expressions. This might suggest that the participants might have thought that their selections were correct, and therefore reported the difficulty to be less than it actually was. As another possible explanation, as difficulty was rated only for the specific observed effects (and not as a direct comparison with other difficulty levels), people possibly judged the difficulties based on their perception of how accurate their responses were, and how well they learned/recognized Miro's expressions in general. In this case, those who did poorly in training and testing would then have reported the difficulties to be higher. Our data supports this possibility, as we found correlations between the reported difficulties of different effects and participants' performance in test and training.

Note, while we mainly focused on the situations where social robots are used for search and rescue or for the care of older adults, these results may be applicable to a wider range of applications where emotions can strengthen communication with social robots, such as in education, healthcare, and entertainment.

9 Limitations and Future Work

Our study had several limitations. First, being an online, remote study, despite our efforts (i.e., adding instructions on adjusting brightness, using Google Chrome only, using a laptop's screen and not a phone or a connected large screen, etc.), we could not guarantee that all the participants indeed saw the videos similarly. However, with the size of the data, we can expect that variations have been similar across different effects and expressions. Another limitation is that the participants did not see the actual robots as the study was online. However, this is a common approach in studies in human-computer/robot interaction, which is shown to be comparable to the direct recruitment methods [4] and can reduce biases such as experimenter bias [40]. This approach has become even more common as a result of COVID-19, and it helped us with further controlling the effects. For example, it would be extremely difficult to simulate smoke and rain effects and ensure that the effects are similar for all the participants in a laboratory and in person setting. Also, it has been suggested in previous studies that there might be a high agreement between video studies and live studies (e.g., see [54]), in particular in cases where there is no direct interaction (e.g. more "observation" than "contingent engagement") involved, and in these situations video studies could be beneficial [48]. However, in another study, the results of an online experiment could not be replicated through an in person experiment [52]. While in that case the task was playing a game with a robot and did require interactions, future work is needed to study how visibility of the affective expressions change when the participants interact with the physical robot in similar but real-world conditions. Also, while we did our best to ensure that the effects in the second study were representative of extreme situations, it is possible that more extreme situations happen depending on the application areas. Finally, this study investigated five affective expressions in five different situations with visual constraints. Future work is needed to investigate recognition of other expressions and in other situations with limited visibility. Future work could also consider comparative studies with differently embodied robots with different affective expressions. A particular challenge here could be to create affective expressions for already existing robot designs (e.g., those used in service or search and rescue). Lastly, the affective expressions used in our study were conveyed through the robot's eyes, ears, tail, and body movements. Future work is needed to understand whether emotions and affective expressions that are shown using other methods (e.g., light) can be as recognizable, or if these results are only valid for affective expressions that are conveyed through motion.

10 Conclusion

In many contexts, affective expressions can complement human-robot interaction, improve quality and efficiency of interactions, and increase users' enjoyment. But similar to other modalities of interaction, there might be limitations in conveying affective expressions. This study evaluated recognition of a zoomorphic robots' affective expressions in different conditions with limited visibility and showed that expressions of a robot reflected through its head and body gestures can be an effective modality of communication for social robots in situations where vision is impaired.

ACKNOWLEDGMENT

This research was undertaken, in part, thanks to funding from the Canada 150 Research Chairs Program and funding from the Network for Aging Research at the University of Waterloo. We would like to thank the members of the Social and Intelligent Robotics Research Laboratory (SIRRL) at the University of Waterloo for their comments on the video effects.

References

1. Akgun, S.A., Ghafurian, M., Crowley, M., Dautenhahn, K.: Using emotions to complement multi-modal human-robot interaction in urban search and rescue scenarios. Accepted for publication in Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20) (2020)
2. Akgun, S.A., Ghafurian, M., Crowley, M., Dautenhahn, K.: Using emotions to complement multi-modal human-robot interaction in urban search and rescue scenarios. In: Proceedings of the 2020 International Conference on Multimodal Interaction. pp. 575–584 (2020)

3. Baker, M., Casey, R., Keyes, B., Yanco, H.A.: Improved interfaces for human-robot interaction in urban search and rescue. In: 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583). vol. 3, pp. 2960–2965. IEEE (2004)
4. Bartneck, C., Duenser, A., Moltchanova, E., Zawieska, K.: Comparing the similarity of responses received from studies in amazon mechanical turk to studies conducted online and with direct recruitment. PLOS ONE **10**(4), 1–23 (04 2015)
5. Beck, A., Cañamero, L., Hiolle, A., Damiano, L., Cosi, P., Tesser, F., Sommavilla, G.: Interpretation of emotional body language displayed by a humanoid robot: A case study with children. International Journal of Social Robotics **5**(3), 325–334 (2013)
6. Beer, J.M., Prakash, A., Mitzner, T.L., Rogers, W.A.: Understanding robot acceptance. Tech. rep., Georgia Institute of Technology (2011)
7. Bennett, M.V., Matthews, I.: Life-saving uncooled ir camera for use in firefighting applications. In: Infrared Technology and Applications XXII. vol. 2744, pp. 549–554. International Society for Optics and Photonics (1996)
8. Bethel, C.L., Murphy, R.R.: Non-facial and non-verbal affective expression for appearance-constrained robots used in victim management. Paladyn, Journal of Behavioral Robotics **1**(4), 219–230 (2010)
9. Breazeal, C.: Emotion and sociable humanoid robots. International journal of human-computer studies **59**(1-2), 119–155 (2003)
10. Chen, A.Y., Peña-Mora, F., Plans, A.P., Mehta, S.J., Aziz, Z.: Supporting urban search and rescue with digital assessments of structures and requests of response resources. Advanced Engineering Informatics **26**(4), 833–845 (2012)
11. Collins, E.C., Prescott, T.J., Mitchinson, B.: Saying it with light: A pilot study of affective communication using the miro robot. In: Conference on Biomimetic and Biohybrid Systems. pp. 243–255. Springer (2015)
12. Collins, E.C., Prescott, T.J., Mitchinson, B., Conran, S.: Miro: a versatile biomimetic edutainment robot. In: Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology. pp. 1–4 (2015)
13. Dang, Q.K., Suh, Y.S.: Human-following robot using infrared camera. In: 2011 11th International Conference on Control, Automation and Systems. pp. 1054–1058. IEEE (2011)
14. Dautenhahn, K., Bond, A.H., Canamero, L., Edmonds, B.: Playing the emotion game with Feelix: What can a LEGO robot tell us about emotion?, vol. 3. Springer Science & Business Media (2002)
15. Delmerico, J., Mintchev, S., Giusti, A., Gromov, B., Melo, K., Horvat, T., Cadena, C., Hutter, M., Ijspeert, A., Floreano, D., et al.: The current state and future outlook of rescue robotics. Journal of Field Robotics **36**(7), 1171–1191 (2019)
16. D'Onofrio, G., Fiorini, L., Hoshino, H., Matsumori, A., Okabe, Y., Tsukamoto, M., Limosani, R., Vitanza, A., Greco, F., Greco, A., et al.: Assistive robots for socialization in elderly people: results pertaining to the needs of the users. Aging clinical and experimental research **31**(9), 1313–1329 (2019)
17. Fleischer, A., Mead, A.D., Huang, J.: Inattentive responding in mturk and other online samples. Industrial and Organizational Psychology **8**(2), 196 (2015)
18. Gácsi, M., Kis, A., Faragó, T., Janiak, M., Muszyński, R., Miklósi, Á.: Humans attribute emotions to a robot that shows simple behavioural patterns borrowed from dog behaviour. Computers in Human Behavior **59**, 411–419 (2016)
19. Ghafurian, M., Lakatos, G., Tao, Z., Dautenhahn, K.: Design and evaluation of affective expressions of a zoomorphic robot. In: International Conference on Social Robotics. pp. 1–12. Springer (2020)

20. Giambattista, A., Teixeira, L., Ayanoğlu, H., Saraiva, M., Duarte, E.: Expression of emotions by a service robot: a pilot study. In: International Conference of Design, User Experience, and Usability. pp. 328–336. Springer (2016)
21. Goodrich, M.A., Morse, B.S., Gerhardt, D., Cooper, J.L., Quigley, M., Adams, J.A., Humphrey, C.: Supporting wilderness search and rescue using a camera-equipped mini uav. Journal of Field Robotics **25**(1-2), 89–110 (2008)
22. Greatbatch, I., Gosling, R.J., Allen, S.: Quantifying search dog effectiveness in a terrestrial search and rescue environment. Wilderness & Environmental Medicine **26**(3), 327–334 (2015)
23. Hortensius, R., Hekele, F., Cross, E.S.: The perception of emotion in artificial agents. IEEE Transactions on Cognitive and Developmental Systems **10**(4), 852–864 (2018)
24. Islam, M.J., Ho, M., Sattar, J.: Understanding human motion and gestures for underwater human–robot collaboration. Journal of Field Robotics **36**(5), 851–873 (2019)
25. Jackovics, P.: Standard of operation for cave rescue in hungary. International Fire Fighter **2016**(9), 84–86 (2016)
26. Jones, B., Tang, A., Neustaedter, C.: Remote communication in wilderness search and rescue: Implications for the design of emergency distributed-collaboration tools for network-sparse environments. Proc. ACM Hum.-Comput. Interact. **4**(GROUP) (Jan 2020). https://doi.org/10.1145/3375190, https://doi.org/10.1145/3375190
27. Kahn Jr, P.H., Kanda, T., Ishiguro, H., Freier, N.G., Severson, R.L., Gill, B.T., Ruckert, J.H., Shen, S.: "robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. Developmental psychology **48**(2), 303 (2012)
28. Karaca, Y., Cicek, M., Tatli, O., Sahin, A., Pasli, S., Beser, M.F., Turedi, S.: The potential use of unmanned aircraft systems (drones) in mountain search and rescue operations. The American journal of emergency medicine **36**(4), 583–588 (2018)
29. Karma, S., Zorba, E., Pallis, G., Statheropoulos, G., Balta, I., Mikedi, K., Vamvakari, J., Pappa, A., Chalaris, M., Xanthopoulos, G., et al.: Use of unmanned vehicles in search and rescue operations in forest fires: Advantages and limitations observed in a field trial. International journal of disaster risk reduction **13**, 307–312 (2015)
30. Kim, J.H., Starr, J.W., Lattimer, B.Y.: Firefighting robot stereo infrared vision and radar sensor fusion for imaging through smoke. Fire Technology **51**(4), 823–845 (2015)
31. Kitade, T., Satake, S., Kanda, T., Imai, M.: Understanding suitable locations for waiting. In: 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 57–64. IEEE (2013)
32. Korcsok, B., Konok, V., Persa, G., Faragó, T., Niitsuma, M., Miklósi, Á., Korondi, P., Baranyi, P., Gácsi, M.: Biologically inspired emotional expressions for artificial agents. Frontiers in psychology **9**, 1191 (2018)
33. Larochelle, B., Kruijff, G.J.M., Smets, N., Mioch, T., Groenewegen, P.: Establishing human situation awareness using a multi-modal operator control unit in an urban search & rescue human-robot team. IEEE (2011)
34. Li, J., Davison, D., Schadenberg, B., Chevalier, P., Alcorn, A., Williams, A., Petrovic, S., Dimitrijevic, S.B., Shen, J., Pellicano, L., et al.: Usability of a robot's realistic facial expressions and peripherals in autistic children's therapy. arXiv preprint arXiv:2007.12236 (2020)

35. McGinn, C., Bourke, E., Murtagh, A., Donovan, C., Lynch, P., Cullinan, M.F., Kelly, K.: Meet stevie: a socially assistive robot developed through application of a 'design-thinking'approach. Journal of Intelligent & Robotic Systems **98**(1), 39–58 (2020)
36. Mitchinson, B., Prescott, T.J.: Miro: a robot "mammal" with a biomimetic brain-based control system. In: Conference on Biomimetic and Biohybrid Systems. pp. 179–191. Springer (2016)
37. Mu, L., Zhao, E.: The optimization of maritime search and rescue simulation system based on cps. In: Big Data Analytics for Cyber-Physical Systems, pp. 231–245. Springer (2020)
38. Na, ., Walker, C., Kvasić, I., Antillon, D.O., Mišković, N., Anderson, I., Lončar, I.: Towards advancing diver-robot interaction capabilities. IFAC-PapersOnLine **52**(21), 199–204 (2019)
39. Pan, Y., Gao, F., Qi, C., Chai, X.: Human-tracking strategies for a six-legged rescue robot based on distance and view. Chinese Journal of Mechanical Engineering **29**(2), 219–230 (2016)
40. Paolacci, G., Chandler, J., Ipeirotis, P.G.: Running experiments on Amazon mechanical turk. Judgment and Decision Making **5**(5), 411–419 (2010)
41. Ritschel, H., Aslan, I., Mertes, S., Seiderer, A., André, E.: Personalized synthesis of intentional and emotional non-verbal sounds for social robots. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). pp. 1–7 (2019)
42. Rivera, A., Villalobos, A., Monje, J., Mariñas, J., Oppus, C.: Post-disaster rescue facility: Human detection and geolocation using aerial drones. In: 2016 IEEE Region 10 Conference (TENCON). pp. 384–386. IEEE (2016)
43. Sabelli, A.M., Kanda, T., Hagita, N.: A conversational robot in an elderly care center: an ethnographic study. In: 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 37–44. IEEE (2011)
44. Saldien, J., Goris, K., Vanderborght, B., Vanderfaeillie, J., Lefeber, D.: Expressing emotions with the social robot probo. International Journal of Social Robotics **2**(4), 377–389 (2010)
45. Sales, J., Marin, R., Cervera, E., Rodríguez, S., Pérez, J.: Multi-sensor person following in low-visibility scenarios. Sensors **10**(12), 10953–10966 (2010)
46. Sharma, M., Hildebrandt, D., Newman, G., Young, J.E., Eskicioglu, R.: Communicating affect via flight path exploring use of the laban effort system for designing affective locomotion paths. In: 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 293–300. IEEE (2013)
47. Silvagni, M., Tonoli, A., Zenerino, E., Chiaberge, M.: Multipurpose uav for search and rescue operations in mountain avalanche events. Geomatics, Natural Hazards and Risk **8**(1), 18–33 (2017)
48. Soegaard, M., Dam, R.F.: The encyclopedia of human-computer interaction. The encyclopedia of human-computer interaction (2012)
49. Starr, J.W., Lattimer, B.: Evaluation of navigation sensors in fire smoke environments. Fire Technology **50**(6), 1459–1481 (2014)
50. Stott, S.: Critical Hours: Search and Rescue in the White Mountains. University Press of New England (2018)
51. Velásquez, J.D.: An emotion-based approach to robotics. In: Proc. International Conference on Intelligent Robots and Systems. vol. 1, pp. 235–240. IEEE (1999)
52. Wallkotter, S., Stower, R., Kappas, A., Castellano, G.: A robot by any other frame: Framing and behaviour influence mind perception in virtual but not real-world

environments. ACM/IEEE International Conference on Human-Robot Interaction pp. 609–618 (2020). https://doi.org/10.1145/3319502.3374800
53. Weiss, Y., Simoncelli, E.P., Adelson, E.H.: Motion illusions as optimal percepts. Nature neuroscience **5**(6), 598–604 (2002)
54. Woods, S.N., Walters, M.L., Koay, K.L., Dautenhahn, K.: Methodological issues in hri: A comparison of live and video-based methods in robot to human approach direction trials. In: ROMAN 2006-the 15th IEEE international symposium on robot and human interactive communication. pp. 51–58. IEEE (2006)
55. Zhao, J., Gao, J., Zhao, F., Liu, Y.: A search-and-rescue robot system for remotely sensing the underground coal mine environment. Sensors **17**(10), 2426 (2017)