Development and Evaluation of Emotional Robots for Children with Autism Spectrum Disorders

Myounghoon Jeon[1](✉), Ruimin Zhang[1], William Lehman[1], Seyedeh Fakhrhosseini[1], Jaclyn Barnes[1], and Chung Hyuk Park[2]

[1] Michigan Technological University, Houghton, USA
{mjeon, ruiminz, welehman, sfakhrho, jaclynb}@mtu.edu
[2] New York Institute of Technology, New York, USA
chung.park@nyit.edu

Abstract. Individuals with Autism Spectrum Disorders (ASD) often have difficulty recognizing emotional cues in ordinary interaction. To address this, we are developing a social robot that teaches children with ASD to recognize emotion in the simpler and more controlled context of interaction with a robot. An emotion recognition program using the Viola-Jones algorithm for facial detection is in development. To better understand emotion expression by social robots, a study was conducted with 11 college students matching animated facial expressions and emotionally neutral sentences spoken in affective voices to various emotions. Overall, facial expressions had greater recognition accuracy and higher perceived intensity than voices. Future work will test the recognition of combined face and voices.

Keywords: Social robotics · Emotion · Autism spectrum disorders

1 Introduction

Recognizing and understanding emotional cues while interacting with other people is vital for effective communication as these cues contain information about meaning, intention, and appropriate responses. People with Autism Spectrum Disorders (ASD) often lack the ability to decipher these cues and this challenge has been identified as one of the biggest barriers to their social inclusion [1]. To help children with ASD develop richer emotional interaction, researchers have used interactive robots and shown positive results [2–8]. Robots allow for a simplified, predictable, and reliable environment where the complexity of interaction can be controlled and gradually increased [9]. Robots can also work as embedded reinforcers of learning [7] and thus, they can form rapport with children.

Our project aims to enhance the emotional communication of children with ASD using social robots. The work described here is a portion of that larger project. In order to enhance communication, we must examine how children with ASD and neurotypical children understand and interpret emotion and how the robot can encourage better emotional interaction. As a test platform, we are using an iOS-based interactive robot,

C. Stephanidis (Ed.): HCII 2015 Posters, Part I, CCIS 528, pp. 372–376, 2015.
DOI: 10.1007/978-3-319-21380-4_63

Romo, which is non-humanoid, but has important human expressive characteristics (eyelids, mouth, voice, etc.) [10, 11] and thus, can have emotional communication with children (Fig. 1).

Fig. 1. Romo

For successful emotional interaction, the robot needs capabilities of emotion recognition and expression. On the emotion recognition side, we are developing software to recognize emotion in facial expressions extracted from real-time video of the individual during interaction. This portion of the research is in progress. For emotional expression, we conducted an experiment with college students to determine how Romo's facial expressions and a variety of equivalent affective voice recordings are interpreted.

2 Design and Implementation of Emotion Recognition and Expression

2.1 Emotion Recognition on Romo

With emotion recognition, we can monitor a child's affective state for intervention purposes. We can also verify if the child-robot interaction successfully yields the intended goal (e.g., enhancing social interaction). While researchers have suggested a variety of emotion recognition methods, no single method has been perfectly successful [12]. Also, it should be differentiated depending on users' characteristics, tasks, and environments. We have developed a multimodal emotion recognition system (facial + voice) for drivers with Traumatic Brain Injury (TBI) [13]. For that project, we developed a facial expression recognition system using the Support-Vector Machines (SVMs) algorithm, which could detect positive, negative, and neutral states. Our second-generation facial detection system for the current project has been developed using the Viola-Jones algorithm in Objective-C. It can detect more specific affective states than our previous version, such as happiness, surprise, and anger, etc. For higher recognition accuracy and additional affective states, we are currently updating our system using the standardized database sets (e.g., Cohn-Kanade [14] and MMI data-base [15]). Future work will utilize a dataset of children's emotional expressions that we are currently creating.

2.2 Emotion Expression on Romo

After estimating a child's affective state, a robot is required to respond in an emotionally appropriate manner. For facial expression, we tested the standard expressions provided by the Romo app including curious, excited, happy, neutral, sad, scared, and sleepy (Fig. 2). For voice expression, a young male adult and a young female adult recorded emotion-independent sentences [16] ("Can I get some food?", "It's time to go.", and "What are you doing?") using affective voice types for the same seven emotion. We controlled Romo's face expression and voice expression using the Romo software development kit.

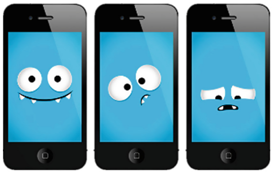

Fig. 2. Happy, curious, and sad (from left to right)

3 User Evaluation of Emotion Expression on Romo

3.1 Method and Procedure

This paper focuses on the evaluation of users' recognition of ROMO's facial and voice expressions, rather than Romo's recognition performance, which as mentioned is still in development. The study sought to understand how emotion expressed by Romo is interpreted through voice and facial cues, so that we can develop salient emotional cues that can be used in Romo and other interactive robots. Clear cues will provide a good foundation for our efforts to improve the emotional understanding of children with ASD.

Eleven college students (ages 18–22, 2 females and 9 males) participated in this experiment. No information was collected regarding participants' ASD or neurotypical status. We used a within subject design in which each participant was subject to face-only and voice-only conditions. The expression could be seen/heard multiple times by tapping the screen. There were 42 trials for voice (7 emotions, 3 different phrases, and 2 genders). For consistency, there were also 42 trials for face, each of the seven emotions repeated six times. After presentation of the stimulus, the participant was asked to choose one out of seven emotions that the stimulus conveyed and rate how strong the emotion was on a scale of 1 ("Not at all") to 7 ("Very"). The face and voice conditions were presented separately and the presentation order was alternated. Within each condition, the trials were randomized for each participant.

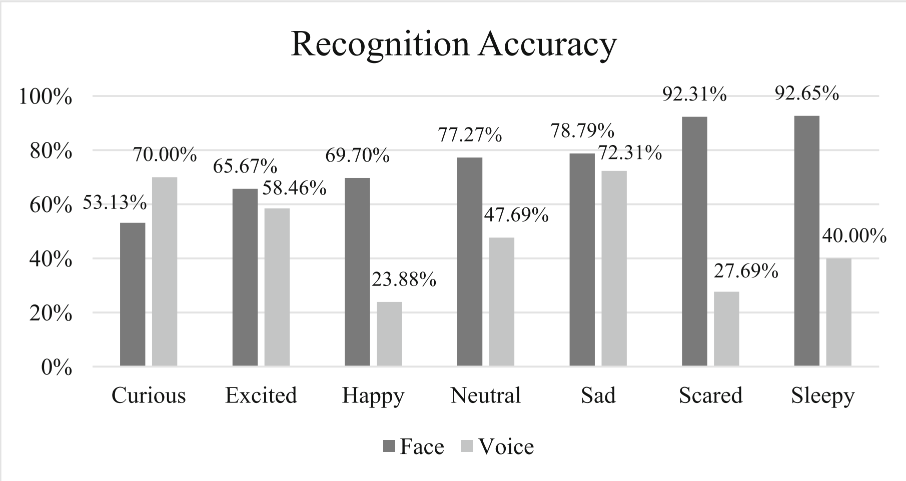

Fig. 3. Recognition accuracy for voice and face across all conditions

3.2 Results

A paired samples t-tests revealed a significant difference of recognition accuracy between faces (M = 75.70, SD = 14.25) and voices (M = 48.57, SD = 19.13), $t(6) = -2.46$, $p < .05$. For all emotions except curious, faces had higher recognition accuracy, particularly the happy, scared, and sleepy conditions (Fig. 3). There was also a trend toward more intense ratings of faces (average 5.78) compared to voices (average 4.57).

4 Discussion and Future Work

The data show that emotion conveyed by facial expression tended to be more recognizable and more intense than that in voice. However, there was no clear pattern of differences in terms of the traditional valence and arousal dimensions, which requires further research.

We plan to combine face and voice stimuli into one condition to test the strength of the combined emotion. We will utilize both matched face and voice affect and contradictory affects. The mismatched condition is expected to indicate whether faces or voices produce stronger emotional cues when combined. The emotion detection program will be expanded to recognize a wider variety of emotions and to appropriately detect children's emotional states.

Acknowledgements. This material is based upon work supported by the National Institutes of Health under grant No. 1 R01 HD082914-01.

References

1. Paul, R., Shriberg, L.D., McSweeny, J., Cicchetti, D., Klin, A., Volkmar, F.: Brief report: relations between prosodic performance and communication and socialization ratings in high functioning speakers with autism spectrum disorders. J. Autism Dev. Disord. **35**, 861–869 (2005)
2. Feil-Seifer, D., Mataric, M.: Robot-assisted therapy for children with autism spectrum disorders. In: Proceedings of the 7th International Conference on Interaction Design and Children, pp. 49–52. ACM (2008)
3. Michaud, F., Théberge-Turmel, C.: Mobile robotic toys and autism. In: Dautenhahn, K., Bond, A., Cañamero, L., Edmonds, B. (eds.) Socially Intelligent Agents, pp. 125–132. Springer, Berlin (2002)
4. Scassellati, B.: How social robots will help us to diagnose, treat, and understand autism. In: Thrun, S., Brooks, R., Durrant-Whyte, H. (eds.) Robotics Research, pp. 552–563. Springer, Berlin (2007)
5. Scassellati, B.: Affective prosody recognition for human-robot interaction. In: Microsoft Research's External Research Symposium. Redmond, WA, USA. Citeseer (2009)
6. Stanton, C.M., Kahn, P.H., Severson, R.L., Ruckert, J.H., Gill, B.T.: Robotic animals might aid in the social development of children with autism. In: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 271–278. IEEE (2008)
7. Kim, E.S., Berkovits, L.D., Bernier, E.P., Leyzberg, D., Shic, F., Paul, R., Scassellati, B.: Social robots as embedded reinforcers of social behavior in children with autism. J. Autism Dev. Disord. **43**, 1038–1049 (2013)
8. Werry, I., Dautenhahn, K., Harwin, W.: Investigating a robot as a therapy partner for children with autism. In: Proceedings of AAATE 2001 (2001)
9. el Kaliouby, R., Picard, R., Baron-Cohen, S.: Affective computing and autism. Ann. N. Y. Acad. Sci. **1093**, 228–248 (2006)
10. DiSalvo, C.F., Gemperle, F., Forlizzi, J., Kiesler, S.: All robots are not created equal: the design and perception of humanoid robot heads. In: Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, pp. 321–326. ACM (2002)
11. Adams, A., Robinson, P.: An android head for social-emotional intervention for children with autism spectrum conditions. In: D'Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) ACII 2011, Part II. LNCS, vol. 6975, pp. 183–190. Springer, Heidelberg (2011)
12. Calvo, R.A., D'Mello, S.: Affect detection: An interdisciplinary review of models, methods, and their applications. IEEE Trans. Affect. Comput. **1**, 18–37 (2010)
13. Jeon, M., Walker, B.: Emotion detection and regulation interface for drivers with traumatic brain injury. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI11) (2011)
14. Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. In: Proceedings. Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000, pp. 46–53. IEEE (2011)
15. Valstar, M., Pantic, M.: Induced disgust, happiness and surprise: an addition to the MMI facial expression database. In: Proceedings of International Conference Language Resources and Evaluation, Workshop on Emotion, pp. 65–70 (2010)
16. Jeon, M., Rayan, I.A.: The effect of physical embodiment of an animal robot on affective prosody recognition. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part II, HCII 2011. LNCS, vol. 6762, pp. 523–532. Springer, Heidelberg (2011)