Adaptive Robot Design with Hand and Face Tracking for Use in Autism Therapy

Laura Boccanfuso[1] and Jason M. O'Kane[1]

Department of Computer Science and Engineering, University of South Carolina, Columbia, South Carolina 29208, USA
boccanfu@cse.sc.edu, jokane@cse.sc.edu

Abstract. This research explores interactive games using hand and face tracking with a robot as a tool for autism therapy. The robot is equipped with a head and two arms, each with two degrees of freedom, and a camera. We trained a classifier to detect human hands and subsequently, used this classifier along with a standard face tracker to create two interactive games. In the first game the robot waits for the child to initiate an interaction by raising one or both hands. In the second game, the robot initiates interactions. These games are designed to increase attention, promote turn-taking skills and encourage child-led verbal and non-verbal communication through simple imitative play. This research makes two specific contributions: (1) We present a low-cost robot design which measures and adapts to a child's actions during interactive games and, (2) we train and test a hand detector, based on Haar-like features, which is usable in various kinds of human-robot interactions.

Keywords: Human-robot interaction, hand detection, hand tracking, adaptive robotics

1 Introduction

Robot-assisted autism therapy employs robots as social mediators for promoting and teaching communication skills in autistic children. Robots have been used effectively to engage autistic children in interactive game playing and research has demonstrated that robot-assisted autism therapy promotes increased speech and increased child-initiated interactions in children with Autism Spectrum Disorder (ASD) [4]. The goal of our research is to provide parents and therapists with an effective, widely usable, interactive robot that will broaden the impact of traditional therapies. Research in robot-assisted autism therapy typically emphasizes specific objectives for ideal human-robot interaction including an increased attention span, eye contact, proactive interaction with the robot initiated by the child, verbal and non-verbal cues, turn-taking, imitative game playing and overall use of language.

In this paper, we describe a simple interactive robot, named CHARLIE (CHild-centered Adaptive Robot for Learning in an Interactive Environment), which uses a turn-taking game for the purpose of engaging autistic children during therapy. See Figure 1. The robot is designed with a head and two arms, each

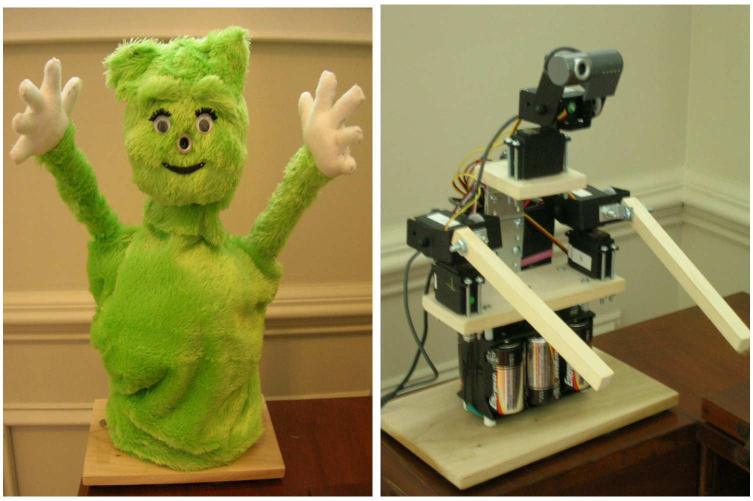

Fig. 1. CHARLIE. [left] Completed robot. [right] Internal structure.

with two degrees of freedom, and a camera for face and hand detection. The camera is mounted inside the robot head which moves, as needed, to maintain visibility of the face whenever possible. We show that basic commodity hardware is sufficient to implement face and hand tracking for interactive games designed for use in autism therapy. Ongoing research focuses on designing and conducting clinical studies with autistic children in order to test and measure the effectiveness of using CHARLIE for promoting turn-taking and basic communication skills. Due to its relatively low cost and inelaborate hardware, CHARLIE is intended to be accessible to a larger population of children than many of the robots currently used for autism therapy. Because of its hand tracking capability, CHARLIE can autonomously participate in a wider range of user-driven, interactive games where robot actions are determined by the actions of the child. Furthermore, the robot can automatically collect information about the child's interactions and provide a summary report for evaluation at a later time.

In general, the implementation of a robust hand tracking system can greatly improve the quality of human-robot interaction, especially when the robot is intended for rehabilitative or therapeutic purposes. First, the progress and preferences of a user can be measured objectively by monitoring response times, length of engagement and number of user-led responses. This information, combined with the amount of verbal and non-verbal communication during a session, can provide valuable information pertaining to a user's progress. In the longer term, we expect the insight gained from this research to generate deeper understanding of the unique nature of robot interactions with the developmentally disabled, leading to broader innovations in robot software for therapy and assistance to this population.

The remainder of this paper is structured as follows. Section 2 is a review of related work. Then, we detail the fundamental methodology and approach underlying the robot and game design in Section 3. In Section 4, we present a description of the preliminary test design and results. We conclude the paper with a summary of our research and a brief discussion about future work in Section 5.

2 Background

Autism therapy ultimately seeks to promote human-to-human interaction. Over the past decade, the use of robots as social mediators has been explored as a tool for supplementing traditional autism therapies in order to teach and improve social skills. Robots are well-suited for interactive games with autistic children since they tend to be perceived as predictable, non-threatening, and are able to perform repetitive tasks consistently and reliably [10, 13]. Most importantly, an increase in basic social and interaction skills has been observed when using robots for turn-taking and imitation games [6].

Some of the most promising results from robot-assisted autism therapy include an increased attention span, eye contact, child-led speech, improved turn-taking and imitative game playing skills and overall use of language [5]. Minimally expressive robots such as KASPAR [15] have been used to explore the efficacy of robot-mediated therapy for autistic children. That research revealed that relatively low functioning autistic children, who would not normally seek physical or eye contact, directly engaged with the robot and, in some cases, proactively touched and gazed at co-present others during sessions with KASPAR.

Other research used for assisting autistic children has resulted in the design and development of various robotic systems. With Keepon [8], it was observed that a very simple robot interface could be used to engage the attention of autistic children and facilitate social interaction. Results obtained from the Bubblebot research [17] showed that human-robot and human-human interaction is increased with a responsive robot whose actions are contingent on user commands. The IROMEC project [12] identified three play scenarios and five distinct developmental areas most beneficial for collaborative, interactive play with autistic children [7].

The robot described herein incorporates key characteristics from each of the above studies. The toylike, non-humanoid appearance of the Keepon and the user-directed modality of the Bubblebot were used as the basis for the development of the robot architecture and the three types of play scenarios identified in the IROMEC study, (1) turn-taking, (2) sensory reward and (3) imitation were used to design the games detailed in this paper. The unique contribution made by this research is the low-cost design and additional functionality provided with the face and hand tracking system. With face and hand tracking, the robot will not only be able to participate in qualitatively different interactive games but it will also allow the robot to collect pertinent information regarding a child's specific progress that may be difficult or impossible to obtain otherwise.

3 Methodology and Approach

The approach taken for this research is based on the integration of robot and game designs that are known to be effective with autistic children. The five specific design components considered in this paper include robot hardware and design (Section 3.1), face and hand detection (3.2), face and hand tracking (3.3), interactive game design (3.4) and data collection (3.5) .

3.1 Robot Hardware and Design

Recent research has shown that robots with a simple interface are generally better received initially by children with autism, than robots with a more realistic, human-like appearance [14]. The implication is that low-tech robots, when designed appropriately for the particular needs of the autistic child(ren) they will serve and the context in which they will be used, can be used effectively to teach and promote social skills. In addition to the low cost mentioned above, CHARLIE'S physical design is intended to be toy like to create a friendly and approachable outward appearance and to more easily attract the attention of a child.

CHARLIE's hardware includes 6 servos, 3 pan-tilt platforms, an 8 channel servo controller, a consumer-grade web cam, and 2 D-cell battery packs. The robot's body is padded for safety, and its outer surfaces are covered with a bright green, fur-like material to achieve a non-threatening appearance. During active game play the child's attention is typically focused near CHARLIE's hands, so one LED is embedded in each of the hands to provide positive feedback during interactive games. Exclusive of the computing hardware, the retail cost of the robot's components is approximately 200 USD. In a production version of this robot, a computer could be integrated into the robot's body, or users could connect via USB to a standard laptop or desktop PC.

3.2 Face and Hand Detection

The Open Source Computer Vision Library (OpenCV) [2], a cross-platform library for real-time computer vision applications, was used for training the hand classifier and for the implementation of hand and face detection. OpenCV provides a facility for object detection based on an extended set of Haar-like features [9]. Informally, this method works by screening small portions of an image for visual characteristics of the target object. To train a classifier to identify a specific class of objects, OpenCV uses Adaptive Boosting (AdaBoost) [16] to create a cascade of boosted classifiers defined over these features.

Face detection is a well-studied problem [18, 19], and effective face classifiers are freely available through OpenCV. Conversely, robust and real-time hand detection in diverse environments, is a topic of continuing research. Numerous approaches for developing robust hand detectors have been explored [1, 11], but the resulting classifiers have not been made available to the research community. Further, hand classifiers that are freely available are not accurate or efficient enough for our application. In order to implement a hand detector suitable for our purposes, we trained a new hand classifier to detect hands in various lighting conditions, rotations, scales and finger positions. Approximately 750 positive hand images of various size, color and position and approximately 3300 negative images were collected and cropped to a uniform pixel size of 40x40. Representative examples are shown in Figure 2. To create additional positive training samples representing variations in lighting, rotation and scale, ten distortions were applied to 100 of those samples, yielding a total of approximately 1750 positive hand samples. We trained a twenty-stage cascade on these samples, yielding an error rate on the training set approaching zero. Section 4 presents a quantitative evaluation of the classifier performance.

Fig. 2. Images used to train the hand detector. [top] Positive examples. [bottom] Negative examples.

3.3 Face and Hand Tracking

Face and hand tracking were implemented using the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm [3]. CAMSHIFT incorporates the MEANSHIFT algorithm which is based on a nonparametric technique for climbing density gradients to find the peak of the probability distribution of the position of a given target object. For face and hand tracking, this translates to identifying the center of the target color distribution in a given video frame. The CAMSHIFT technique was implemented for this research because it is fast (making it ideal for use in real-time tracking applications), with relatively robust accuracy. This tracking method improves performance by eliminating the need to repeat the face and hand detection for each frame of the video. To overcome errors resulting from drift in the CAMSHIFT algorithm, the robot periodically repeats the full face detection process. In the event that the robot cannot detect the face, the robot head is reset to a neutral position and searches outward in an increasingly larger area.

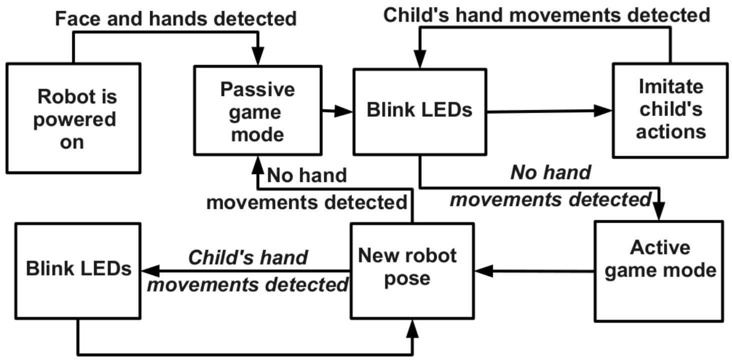

Fig. 3. State diagram for CHARLIE's interactive games.

3.4 Interactive Game Design

As detailed in Figure 3, the robot has two states: passive and active game modes. Within each of the two game modes, there are five poses: neutral (both hands down), left hand raised, right hand raised, both hands raised and peek-a-boo, as shown in Figure 4. In order to give the child initial control over the robot's actions, the default robot state is the passive game mode. Once the robot detects and begins tracking the child's face and hands, the robot indicates that it is ready to interact by moving to the neutral pose and blinking the LEDs in its hands three times. The robot then immediately enters the passive game mode and waits for the child to initiate a game by raising one or both hands. As the child's hand movements are detected, the robot responds by imitating the child's hand positions and lighting the LED in the corresponding hand while simultaneously detecting any additional hand movements. If ten seconds elapse without any detected hand movement, the robot will transition to the active game mode.

During the active game mode, the robot initiates a new game and attempts to engage the child by raising or lowering one or both arms, or beginning a game of peek-a-boo. Each pose assumed by the robot in the active game state is selected randomly in order to avoid repetitive patterns of poses. When a positive outcome is detected (the child successfully imitates the robot's pose), positive sensory feedback is generated by the robot. A positive sensory response entails the robot lighting a small LED in the hand corresponding to the raised hand or hands of the imitated pose. As with the passive game mode, the robot will wait ten seconds for the child's response. If ten seconds elapses and a positive response has not been detected, the robot will transition back to the passive game mode, waiting again for the child to initiate a new game.

3.5 Data Collection

There are two distinct kinds of user interaction information collected by the robot. Information pertaining to the user's overall progress such as (1) the total length of active engagement (time spent actively engaging in either passive or active mode), (2) number of child-led actions and (3) the number of successful interactions is continuously captured during each session. At the end of the session, this information is used to create a user progress report for analysis and for future sessions with the same child. The second type of user information, such as the length of the intervals between interactions, is used for controlling the robot state.

Fig. 4. CHARLIE poses. [top left] Left hand high. [top center] Right hand high. [top right] Both hands high. [bottom left] Neutral. [bottom right] Peek-a-boo.

4 Preliminary Tests

As a proof of concept for CHARLIE's effectiveness, preliminary tests were conducted with a small group of typically developing children. See Figure 5. A

Fig. 5. Children Interacting with CHARLIE.

relatively large age range (4-11 years) was selected primarily to test the reaction times of the robot when used with children of varying levels of ability. Each child participated in an 8-10 minute session, in which both game modes were tested and the accuracy of the hand and face detectors was measured. The duration of each game mode was recorded to ensure that adequate time is given for the child to respond before a transition is made to the alternate game mode and the effectiveness of the positive sensory feedback (LEDs in hands indicating successful detection) was assessed.

It is important to note that preliminary tests were conducted strictly to test the speed and accuracy of the face and hand detector and to assess the appropriateness of CHARLIE's timed responses during game play. Future studies with autistic children will be designed to test and measure each child's progress in terms of specific turn-taking through imitation objectives including length of engagement, number of child-led actions and verbal or nonverbal communication.

The accuracy of the face detector and tracker was determined by calculating the ratio of successful face detection time to the total session time. The face detector averaged an accuracy of 86% across all sessions and users. This accuracy rate is artificially low because it includes as misses the aggregate time when participants moved outside of the video frame. The accuracy of the hand detector and tracker was calculated similarly. In a typical session, users averaged 33 child-initiated hand movements and imitated 16 robot movements per minute. The hand detector accurately detected the child's hands an average of 92% of the total session time, with 244 hits out of 265 total hand events.

Nearly all of the children expressed a preference for the passive game mode, where the robot imitates the child's hand actions, and their comments were supported by the significantly greater amount of time each of those children spent in the passive mode compared to the active mode during their respective sessions. Our hypothesis is that autistic children interacting with CHARLIE may also prefer the passive game mode, since this affords the child the greatest amount of control over the robot. We consider these preliminary results as an important proof-of-concept in preparation for controlled tests with autistic children.

5 Future Work and Conclusions

This research resulted in the design and development of a low-cost, adaptive robot and a dual-mode interactive game for use in robot-assisted autism therapy. One of the aims of this research was to create a robot that is financially accessible to a greater population of therapists and families with autistic children in order to broaden the impact of traditional therapies. The second objective was to develop a hand detector enabling a larger scope of interactive games in which the robot can engage autonomously. Achieving this second objective also allows for real-time collection of important user interaction information specific to the preference and progress of each child undergoing autism therapy. Collectively, these contributions produce a new robot which is designed to be child-centered, adaptive to user preference, and to fulfill a key supportive role for therapists by automatically generating user progress reports.

Work on the hand classifier is ongoing in order to produce a more robust hand detector with improved accuracy. To further improve the accuracy of hand and face detection and to explore the remote detection of user stress levels, we are researching the use of infrared sensing to collect physiological information using a prototype similar to the one used for this research. Continuing work is also being conducted to identify and develop measures of success so that user information collected during an interactive game can be used to assess the child's readiness for more advanced, child-initiated games such as collaborative group play and story-telling.

Field tests with a population of autistic children are being designed and planned for Fall 2010. Discussions with clinicians currently working with autistic children from the South Carolina Department of Disabilities and Special Needs and the South Carolina Autism Treatment Network most recently resulted in the recommendation that young autistic children undergoing early intervention, Applied Behavioral Analysis (ABA) would be good candidates for interacting with CHARLIE. A cornerstone of ABA relies on the assessment and documentation of interventions to ensure their efficacy and to promote progress from one session and from one therapist to the next. In addition, it was recommended that two physically distinct robot prototypes be tested. To complement the "soft and fuzzy" appearance of CHARLIE, a robot with a more mechanical, robotic outward appearance will be constructed in order to appeal to a broader scope of children. Ultimately, three general elements will be considered for evaluating CHARLIE's design: (1) children's response to CHARLIE's physical appearance, (2) overall engagement of the child undergoing therapy with CHARLIE and, (3) effect on turn-taking and communcation skills. Specific measures of success are currently being developed.

Acknowledgements We acknowledge support from the National Science Foundation (NSF Award No. 1004899) for this research. This research is also partially supported by a grant from the South Carolina Developmental Disabilities Council. We would like to thank Ruth K. Abramson, Harry H. Wright, Alicia V. Hall and Elizabeth Wilkinson for their time, insight and support.

References

1. A.L. C. Barczak, F.D.: Real-time hand tracking using a set of co-operative classifiers based on haar-like features. Res. Lett. Inf. Math. Sci. 7, 29–42 (2005)
2. Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools pp. 120–126 (Nov 2000)
3. Bradski, G.R.: Computer vision face tracking for use in a perceptual user interface. Intel Technology Journal (1998)
4. Dautenhahn, K.: Robots as social actors: Aurora and the case of autism. In: Proc. Cognitive Technology Conference. pp. 359–374 (1999)
5. Dautenhahn, K., Werry, I.: Issues of robot-human interaction dynamics in the rehabilitation of children with autism. In: Proc. International Conference on the Simulation of Adaptive Behavior. pp. 519–528 (2000)
6. Duquette, A., Mercier, H., Michaud, F.: Investigating the use of a mobile robotic toy as an imitation agent for children with autism. In: International Conference on Epigenetic Robotics (2006)
7. Ferrari, E., Robins, B., Dautenhahn, K.: Therapeutic and educational objectives in robot assisted play for children with autism. In: Proc. IEEE International Symposium on Robot and Human Interactive Communication (2009)
8. Kozima, H., Nakagawa, C., Yasuda, Y.: Children-robot interaction: a pilot study in autism therapy. Progress in Brain Research 164, 385–400 (2007)
9. Lienhart, R., Maydt, J.: An extended set of Haar-like features for rapid object detection. In: Proc. IEEE International Conference on Image Processing. pp. 900–903 (2002)
10. Marti, P., Pollini, A., Rullo, A., Shibata, T.: Engaging with artificial pets. In: Proc. Conference on European Association of Cognitive Ergonomics. pp. 99–106 (2005)
11. P. Buehler, M. Everingham, D.H.A.Z.: Long term arm and hand tracking for continuous sign language tv broadcasts. In: British Machine Vision Conference (2008)
12. Patrizia, M., Claudio, M., Leonardo, G., Alessandro, P.: A robotic toy for children with special needs: from requirements to design. In: Proc. IEEE International Conference on Rehabilitation Robotics. pp. 918–923 (2009)
13. Pierno, A.C., Maria, M., Lusher, D., Castiello, U.: Robotic movement elicits visuomotor priming in children with autism. Neuropsychologia 46, 448–454 (2008)
14. Robins, B., Dautenhahn, K., Boekhorst, R., Billard, A.: Robots as assistive technology - does appearance matter? In: Proc. IEEE International Workshop on Robot and Human Interactive Communication. pp. 277–282 (2004)
15. Robins, B., Dautenhahn, K., Dickerson, P.: From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot. In: Proc. International Conference on Advances in Computer-Human Interactions (2009)
16. Schapire, R.E.: The boosting approach to machine learning: An overview. In: Denison, D.D., Hansen, M.H., Holmes, C., Mallick, B., Yu, B. (eds.) Nonlinear Estimation and Classification. Springer (2003)
17. Tapus, A., Tapus, C., Matarić, M.J.: User-robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. Intelligent Service Robotics Journal pp. 169–183 (Apr 2008)
18. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. vol. 1, pp. 511–518 (2001)
19. Viola, P., Jones, M.: Robust real-time face detection. International Journal of Computer Vision 57(2), 137–154 (2004)