Subtle Expressivity in a Robotic Computer

Karen K. Liu
MIT Media Laboratory
20 Ames St. E15-120g
Cambridge, MA 02139 USA
kkliu@media.mit.edu

Rosalind W. Picard
MIT Media Laboratory
20 Ames St. E15-020g
Cambridge, MA 02139 USA
picard@media.mit.edu

ABSTRACT

A Robotic Computer, which moves its monitor "head" and "neck" but has no explicit face, is being designed to interact with users in a natural way for applications such as learning, rapport-building, interactive teaching, and posture improvement. In all these applications, the robot will need to move in subtle ways that express its state, and that promote appropriate movements in the user, but that don't distract or annoy. Toward this goal, we are giving the system the ability to recognize subtle expressions as well as the ability to have them. This paper describes the design of this system, initial findings on the perceived qualities of its expressions, and planned future work aimed at measuring behavioral effects of its expressiveness.

Keywords

Subtle expressiveness, robot, affect, emotion

1. INTRODUCTION: A Robotic Computer

The Robotic Computer is a computer monitor with a mechanical neck that physically interacts with the user by recognizing and responding to social-emotional cues. The system takes in perceptual data – visual data through an IBM Blue Eyes camera system to track user facial expressions (Kapoor & Picard, 2002) and posture data through a posture recognition system, to track behaviors related to user interest level (Mota, 2002) – and is capable of responding to these cues through mechanical and auditory expressions. The mechanical expressions include postural shifts like moving closer to the user, and "looking around" in a curious sort of way, while the auditory expressions, designed to be similar in spirit to the fictional Star Wars robot R2D2, are non-linguistic but aim to complement the movements, e.g. electronic sounds of surprise. The long-term plan is for the robot to learn relationships between the expressive cues it makes and the ones its user(s) make, as part of an effort to learn what kinds of actions and behaviors are appropriate for facilitating user task goals while not annoying or distracting the user.

There are three application areas that have motivated the creation of a Robotic Computer. The first application is the construction of a computerized learning companion (Kort et. al., 2001), that would try to help a child persist and stay focused on a learning task, possibly also mirroring some of the child's affective states in a way to increase awareness of the role those states play in propelling the learning experience. For example, if the child's face and posture show signs of intense interest in what is on the screen, the robotic terminal would hold very still so as not to distract the child. If the child shifts her posture and glance in a way that shows she is taking a break, the computer might do the same, and might note that as a good time to interrupt the child and provide scaffolding (encouragement, tips, etc.) to help with the learning progress. In so doing, the system not only acknowledges the presence of the child, and shows respect for her level of attentiveness, but also shows subtle expressions that, in human-human interaction, are believed to help build rapport and liking (e.g., LaFrance, 1982). By increasing likeability, we aim to make the system more enjoyable to work with and potentially facilitate task outcome, such as how long the child perseveres with the learning task.

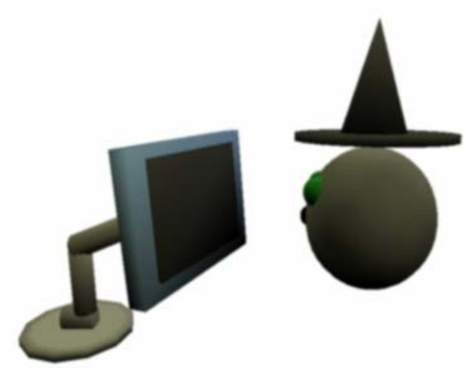


Figure 1: Virtual Robotic Computer Interacting with a Head Creature in a World Model

The second application area considers how a human can teach a robot new actions over time. In this application it is important that the robot's state be transparent to the user: for example, if the robot doesn't understand what action the person is requesting it to learn, it could look confused. If the person hasn't shown it anything in a while, it could try to look curious. If its goal is to learn as much as possible from the user, without annoying the user, then its communication of its state via subtle expressions, in tandem with recognizing the user's expressions, will be important in achieving success. Additionally, in this application there is an issue of shared control (Breazeal, 2003): the user and character must negotiate, and subtle expressions will assist in this process.

The third application area is that of a health and posture improving system. As static sitting postures have become more prevalent in our workplaces (approximately 75 percent of the workforce has sedentary jobs) musculoskeletal problems -- in particular, low back pain and discomfort – have also increased (Faiks and Reinecke, 1998). Not only do these disorders contribute to higher medical expenses for corporations (Steelcase Inc., 2000), but they also decrease worker productivity and increase fatigue. Experts agree that movement and alternating postures during sitting is beneficial. For instance, the research of Holm & Nachemson (1983) suggests that the flow of nutrient-rich fluids to and from the intervertebral discs increases with spinal movement. Adams (1983) found that alternating periods of activity and rest, and posture change, further boosts the fluid exchange, helping to nourish these spinal discs. Grandjean (1980) found that alternating unloading and loading of the spine (through movement) is ergonomically beneficial because this process pumps fluid in and out of the disc, thereby improving nutritional supply. To date, most efforts such as those from Steelcase Inc. (www.steelcase.com) and Herman Miller (www.hermanmiller.com) have focused on researching the ergonomic design of office furniture and environments that remain passive, such as chairs, tables, keyboards, office spaces etc. Successful designs, such as the LEAP chair by Steelcase, Inc. are able to show longitudinal benefits to health and productivity (Allie, P. and Palacios, N., 2002). In this case, we want to explore how similar benefits can be achieved through interactive technologies. The aim of the Robotic Computer would be to encourage movement and proper posture of the user, without distracting him from his primary task. Again, we think the combined perception and use of subtle mechanical and acoustic expressions will assist in achieving these multiple goals.

But why design a robotic computer? There has been increasing amounts of evidence that a physically present robot character, rather than an animated character on a screen, offers interesting advantages for applications such these. "Presence" is a term used to describe several dimensions of how similar a given interaction is to an actual social interaction between people (Lombard & Ditton, 2001). Endowing a technology with a strong social presence precedes the ability of the technology to develop a solid social rapport with the user (provided that the social presence is "strong" in a pleasing way and not in an annoying way). Further, we believe that there are interaction advantages to the physical presence of a machine that shares the same space as the user. For example, an important aspect of the mentor-student relationship is shared reference through cues such as directing attention, mutual gaze, etc. These cues are more easily communicated when both parties occupy the same physical space. Not only would a physical presence be advantageous in our application areas, but utilizing the computer terminal as a robotic interface will allow for natural communication of information to the user while not increasing the number of interfaces that a user would have to interact with.

New challenges arise with designing effective expressions within the physical limitations of a robot's motors and hardware. The rest of this paper describes the design of the current Robotic Computer (Sec 2), initial results on its expressive capabilities and how users perceive these (Sec 3), and future work (Sec 4.)

2. SYSTEM DESIGN AND USER PERCEPTIONS

2.1. System Design

The Robotic Computer system is built on the c4 behavior architecture system developed at the MIT Media Lab (Burke et. al., 2001). The IBM Blue Eyes system (Kapoor & Picard, 2002) in conjunction with the posture recognition system (Mota, 2002) sends the location coordinates of the user's pupil positions, whether the user is nodding or shaking his head, and the user's posture (leaning forward, leaning right, leaning left, leaning back, slumping, and sitting upright) to a virtual head creature in an internal world model which adopts the appropriate state based on sensor data as shown in Figure 1. The behavior engine contains a mental representation of the Robotic Computer that consists of a "perception and memory" section, an "action selection" section, and a "motor and graphic" section as shown in Figure 2. As data comes in from the world model to the cognitive architecture of the system, the "perception and memory" section affects the "action selection" section to determine which behaviors the terminal will take. The actions are passed to the "graphic and motor" section that controls the virtual computer and physical motors of the robot. The Robotic Computer's behavior patterns and actions adapt over time as the system builds a working memory with the user and the likelihood of taking certain actions changes based on past interactions and adaptations in the internal motivation system of the terminal.

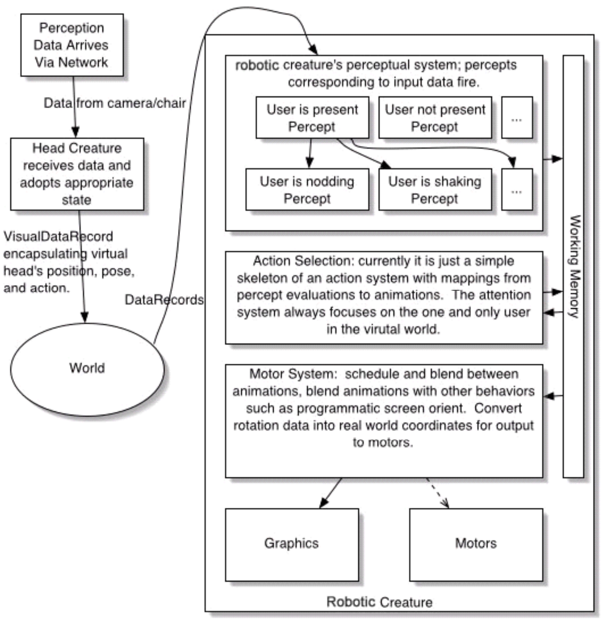


Figure 2: Robotic Computer System Architecture

2.2 Subtle Expressivity in the Robotic Computer

Unlike almost all other expressive robots built, this robot has no facial features that can be used to convey expression. Animation studios such as Disney have looked at how to convincingly portray life in inanimate objects; however, such essential animation tools such as "squash and stretch" (Thomas & Johnson, 1981) cannot be applied to a robotic terminal with limited degrees of freedom, as the motors and hardware are not malleable. What types of subtle behaviors can we use to convey attitude and emotional and cognitive state then? The Robotic Computer communicates all expression through the use of a mechanical neck with five degrees of freedom and the use of different mechanical sounds similar to those of R2D2. In order to evaluate if the robot has recognizable expressions, virtual simulations of the behaviors were created. These subtle expressions through use of audio and posture movements can be utilized to provide feedback and readable behavior to the user. Other questions would include how to effectively use these subtle expressions to manage expectations (help prevent annoyance, frustration, boredom in both user and computer)? The Robotic Computer could also manage the pace of the interaction through varying movement or audio speed as Kismet, the sociable robot, did (Breazeal, 2003).

2.3 User Perceptions

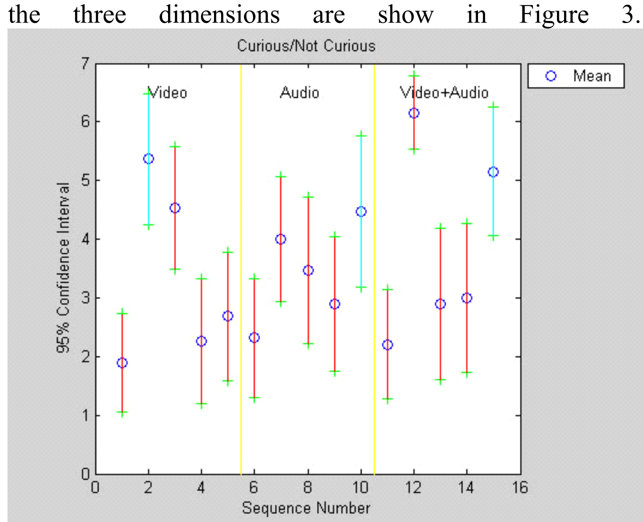
A preliminary study to measure if the first expressive behaviors of the Robotic Computer could accurately be recognized was conducted. While it is important to also measure if the expression actually made the user feel the intended way, this study only looked at user perceptions of the agent's behavior. The study consisted of an online survey where nineteen subjects viewed fifteen different video clips of an animated version of the Robotic Computer and/or heard sound sequences that the terminal would emit. Each of the five different video or audio sequences was designed to convey a certain expressive behavior. The fifteen sequences can be found at http://www.media.mit.edu/~kkliu/roco.html.

Preliminary Results

After each sequence, the user was asked to rate how expressive the computer appeared along five different dimensions. The dimensions used a seven-point scale and are listed below:

- Not Welcoming (1) – Welcoming (7)
- Sad (1) – Happy (7)
- Not Curious (1) – Curious (7)
- Not Confused (1) – Confused (7)
- Not Surprised (1) – Surprised (7)

For purposes of analyzing the data on the same scale, the happy/sad data was inverted to Not Sad (1) – Sad (7), as the sequence was designed with a sad expression in mind. For each dimension, the subjects were also given a "Not Applicable" choice. After performing a two-tailed t-test on this data, we found that there was significant recognition of the behavior sequences that were designed to express curiosity, sadness, and surprise. The results for the three dimensions are show in Figure 3.



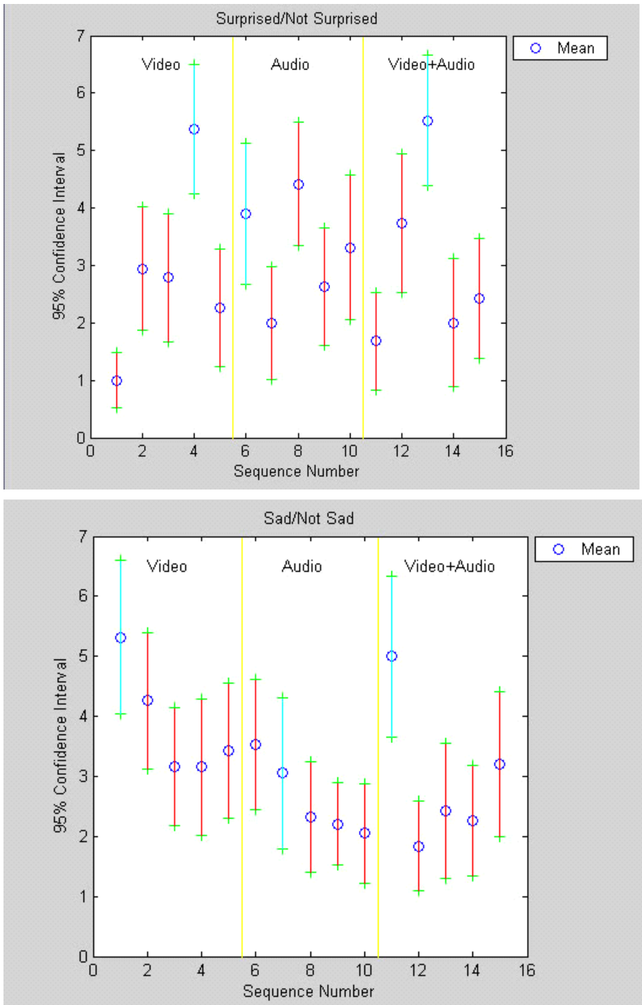


Figure 3: Mean and 95% Confidence Interval for Each Sequence on Curious, Surprised, and Sad Dimensions. Sequences designed for respective dimensions shown in blue.

The intended expression for each sequence is shown in Figure 4.

Sequence	Intent	8	Confused
1	Sad	9	Welcoming
2	Curious	10	Curious
3	Confused	11	Sad
4	Surprised	12	Confused
5	Welcoming	13	Surprised
6	Surprised	14	Welcoming
7	Sad	15	Curious

Figure 4: Intended Expression for Each Sequence

It was not surprising to find very little significant differences for the confused and welcoming expressive behaviors. As both confused and curious behavior could be seen as questioning behavior, there was expected overlap in the perception of how confused the computer appeared to be. As for a welcoming expression, this behavior could easily be interpreted as happy, which would explain some spread in the results. In the future, we are interested in measuring change in user behavior as well as perceived expression.

3. FUTURE WORK

Future directions include looking at how this Robotic Computer can create its own subtle expressions by learning over time which behaviors were effective in its different interactions as well as measuring the behavioral effects of its expressiveness. One major challenge will be to learn when an appropriate intervention time is in order to avoid annoying the user with a moving terminal. For example, when the posture recognition system (Mota, 2002) detects that the child is taking a break, this may signal a good time for the robotic computer to request permission for itself to take a break and then exercising its motors in a manner that resembles stretching. We plan on measuring effects related to the three application areas mentioned above in following questions:

- Does interacting with a robotic terminal that recognizes and responds to affective state cause a user to persist in the learning and teaching interaction?
- Does the system's movement encourage your own movement and alternating postures as you work at a terminal?
- Can the machine learn when its expression was ineffective or annoying? Can it be taught new behaviors? Can it learn when to hold still as people are working on a task?
- Do I like the computer more? Do I enjoy working with it more?

4. CONCLUSION

A Robotic Computer, having no explicit facial features, has been designed to recognize and exhibit a variety of subtle expressions. Additionally, its ability to communicate a set of expressions via motion and auditory cues has been evaluated with nineteen subjects. This paper has emphasized the importance of combining display of expressions with recognition of expressions, with the expectation that an appropriate interaction between these can lead to beneficial effects and behaviors in a variety of tasks.

ACKNOWLEDGMENTS

We thank Cynthia Breazeal, Jesse Gray, Ashish Kapoor, Cory Kidd, John McBean, and David Lafferty for their participation in building the Robotic Computer. This research was supported in part by NSF ROLE grant REC-0087768.

References

1. Adams, M. A. (1983). The Effect of Posture on the Fluid Content of Lumbar Intervertebral Discs. Spine, v8:n6.
2. Allie, P. & Palacios, N. (2002). Steelcase Leap Chair's Impact on Office Work Effectiveness, Productivity and Health. Steelcase Inc. Study Summary.
3. Breazeal, C. (2003). Social Interactions in HRI: The Robot View. To appear in IEEE Transactions in Systems, Man, and Cybernetics. MIT Media Lab. Cambridge, MA.
4. Burke, R., Isla, D., Downie, M., Ivanov, Y., Blumberg, B. (2001). CreatureSmarts: The Art and Architecture of a VirtualBrain. In Proceedings of the Game Developers Conference (pp.147-166). San Jose, CA.
5. Faiks, F. & Reinecke, S. (1998). Investigation of Spinal Curvature While Changing One's Posture during Sitting. Contemporary Ergonomics. M.A. Hanson, Taylor & Francis.
6. Grandjean, E. (1980). Fitting the Task to the Man. London: Taylor and Francis.
7. Holm S. & Nachemson A. (1983). Variations in Nutrition of the Canine Intervertebral Disc Induced by Motion. Spine, v.8, no.8.
8. Kapoor, A. & Picard, R. (2002). Real-Time, Fully Automatic Upper Facial Feature Tracking In the Proceedings of the 5th International Conference on Automatic Face and Gesture Recognition Washington D.C.
9. Kort, B., Reilly, R., & Picard, R. (2001). An Affective Model of Interplay between Emotions and Learning: Reengineering Educational Pedagogy – Building a Learning Companion. In Proceedings of IEEE International Conference on Advanced Learning Technologies. Madision, USA.
10. LaFrance, M. (1982). Posture Mirroring and Rapport. In M. Davis (Ed.), Interaction Rhythms: Periodicity in Communicative Behavior (pp. 279-298). New York: Human Sciences Press, Inc.
11. Lombard, M., Ditton, T.B., Crane, D., Davis, B., Gil-Egul, G., Horvath, K. and Rossman, J. (2000). Measuring Presence: A Literature-Based Approach to the Development of a Standardized Paper-and-Pencil Instrument. In *Presence 2000: The Third International Workshop on Presence*. Delft, the Netherlands.
12. Mota, S. (2002). Automated Posture Analysis for Detecting Learner's Affective State. MS Thesis. MIT Media Lab. Cambridge, MA.
13. Reeves, B. & Nass, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. New York: Cambridge University Press.
14. Reinecke, S., Bevins, T., Weisman, J., Krag, M.H. and Pope, M.H. (1985). The Relationship between Seating Postures and Low Back Pain. Rehabilitation Engineering Society of North America, 8th Annual Conference, Memphis, TN.
15. Steelcase Inc. (2000). Musculoskeletal Disorders (S10647). Grand Rapids, MI: Steelcase Inc.
16. Thomas, F., & Johnson, O. (1981). The Illusion of Life: Disney Animation. New York: Hyperion.