



[Este artículo también está en Español.](#)

HCI applications for aiding children with mental disorders

by [Hossein Mobahi](#) and [Karrie G. Karahalios](#)

Introduction

This work is a part of an ongoing project that focuses on potential applications of HCI in aiding children with mental disorders. In this work we look at autism and bipolar disorder in particular. We believe HCI is promising here because it provides the possibility of capturing, analyzing, and influencing human perception and behavior. Treating mental disorders, particularly those discussed in this article, requires frequent individual sessions over long durations. An HCI assisted treatment allows for such personalized therapy and it can be expanded to allow for treatment in households that do not have ready access to such therapy.

Children with autism are socially impaired and usually do not attend to the people around them. Recently, a new trend for utilizing HCI technology in the treatment of autism has emerged. So far it has mainly focused on diagnosing of autism [20] or practicing social skills like imitation [19]. However, we choose a different route that up to our knowledge has not yet been explored. We are interested in developing a mediating tool through which the world would seem more sensible to those with difficulty perceiving social cues. We suggest attention control for this purpose by eye tracking and augmented/manipulated reality techniques.

Bipolar children suffer from abnormally strong mood swings. There has been no research reported in the literature for exploiting HCI in the treatment of bipolar disorder. We think there is much to be gain from HCI in this area as it provides tools for estimating one's emotional state. These tools can be used for monitoring mood. Moreover, affective computing, an area of HCI which explores the influence of emotional agents on people, shows potential for influencing the moods of patients [16].

Although the two disorders are very different, we can adopt a common strategy by abstraction. In fact, both problems have a desired and an actual mental state (mood or attention). The goal is to continually measure the actual state and influence the person such that his or her current state moves toward the desired one. This kind of abstraction is the subject of control and system theory where a process is studied as a inputs, outputs and feedbacks. A system theoretical approach also facilitates analysis and simulation of the system.

Related Work

Dautenhahn *et al* [19] in their Aurora project have investigated how an autonomous mobile robot can encourage children to become engaged in social interaction which demonstrates important aspects of human-human interaction. Children with autism prefer a predictable and structured environment. They favor objects over people because of the unpredictability in human behavior. The long-term goal of the Aurora project is to slowly advance the robot's behavior to guide the children towards the more complex interaction found in social human-human interaction.

Another ongoing project by Scaessellati [20] uses a humanoid robot, named Nico, to react to a child by detecting the direction of their gazes and make eye contact. Nico will hopefully create for children **social presses**, behaviors that trigger a social response in the child such as pointing to a distant location in the environment. It is possible to diagnosis autism by comparing a child's expected and actual responses.

Although no HCI approach has been reported in the literature for aiding bipolar patients, data collection methods used in the Wireless Mood Telemetry project by Kreindler *et al* [14] can benefit researchers. This project collects mood information from patients using wireless technology and computer software. Through cell phones and handheld computers, a series of mood rating questions were presented. This project provided 40,000 completed responses that can be used to study bipolar disorder.

Autism

Autism is a lifelong developmental disorder that causes difficulty in the areas of communication and social interaction. There is between a 0.02% and 0.05% chance of being born with autism however it is the third most common developmental disorder in children [4]. Unfortunately, there is no treatment or intervention strategy that cures autism. Detecting autism early and then helping children understand their social environment is currently the best course of action.

Since social impairment is an identifying attribute of children with autism, we will review two areas in which it manifests, social orienting and joint attention. **Social orienting** is the ability to adjust to natural social stimuli in the environment. For instance, while normal children show attraction to people, particularly another's voice and facial movements, children with autism do not [7]. **Joint attention** is the ability to properly divide their attention between people and their surroundings [13]. This capability includes sharing attention through eye gaze and following the attention of others by their gazes. Our study on autism is mainly focused on improving social orienting ability.

Methodology

Our proposed approach manipulates raw sensory information that the child would perceive by increasing saliency of social stimuli. For the sake of simplicity, we limit our study to visual attention. Possible hardware for sensing environment and projecting the manipulated information to child's eyes might be a light-weight head mounted camera and display set. The degree of saliency

intensification depends on his or her attention span. Someone who easily social stimulated does not should not receive as much intensification as a person who is insensitive to stimuli. Such social eyewear can be used when interaction people is crucial such as a teacher in an educational setting.

From a control theory perspective, we can consider the child as the system we want to control ([Figure 1](#)). Therefore, the child's retinal image becomes the system input and his or her focus of attention becomes its output. In normal children, the system generates the appropriate output spontaneously. However, the attention deficit in children with autism results in an incorrect output, which we wish to correct using a controller.

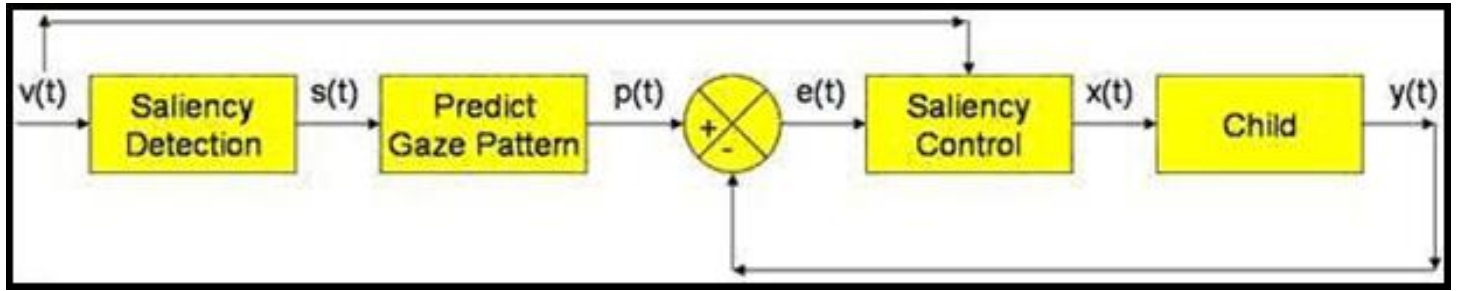


Figure 1: Child-Machine interaction with attention controller in the loop.

In this scheme, the controller gets the original visual information $v(t)$, manipulates it according to the error signal $e(t)$, and then feeds its processed version to the system. The signal $e(t)$, which will be defined later, guides the controller to process visual information such that it can grab the subject's attention and get it to the desired point. The controller's output should be projected over the subject's visual sensor, eyes, such that environmental disturbance is minimized. A very light-weight and comfortable head mounted display and camera set can be a good option for hardware.

The proposed controller works in closed loop; it gets feedback from system's output to adjust its own behavior. Therefore, we need some way to measure the subject's output. Unfortunately, measuring the focus of attention directly is not straightforward. Instead, we can estimate attention from gaze direction because the latter is easier to measure and is interconnected with attention. There are well established vision algorithms available for tracking eyes and estimating gaze [1, 18] that can be used here.

The controller needs to know the desired output. This can be determined either manually by a human supervisor that interactively changes the attention set point or autonomously by a gaze trajectory model. In the latter case, one way to roughly predict gaze trajectory is to use the model of two component attention [2]. In this model, focus of attention is the result of the interaction of top-down and bottom-up components.

The top-down process biases attention according to cognitive knowledge. In the case of autism, this knowledge should be comprised of socially stimulating concepts like human faces. This is because such concepts are not by themselves stimulating to autistic people thus making a face detector an essential part of this model.

The bottom-up component biases attention based on low-level visual features such as color, intensity, and motion. One way to implement this mechanism is by using a saliency map. This idea was introduced by Koch and Ullman [11]. A saliency map is a 2-dimensional image that encodes the saliency of items in the visual scene. A neural network can be trained to detect the most salient object [11]. In Koch and Ullman's model, a saliency map is obtained from a combination of feature maps. Each feature map acts as a filter that responds to a certain low level visual feature of the retinal image.

The idea here is that we can influence saliency in a given location by manipulating its feature maps. For instance, magnifying the intensity level in that location could increase saliency of a location. This scheme can generate low level pressure on a child's attention through this bottom-up process. The goal is to apply this pressure in socially stimulating locations that autistic children do not normally find attractive.







Figure 2: Manipulating saliency of human face through intensity.

Manipulating color saturation and intensity, which strongly influence the saliency map, is easily achieved in HSI color-space. Basically, we can use saturation and intensity channels as feature maps of saliency. For manipulation, each feature map is added up with a corrective map obtained from the top-down process, e.g. face detection. Eventually, corrected maps are combined again to construct the color image. [Figure 2](#) shows an image whose intensity is manipulated by the face detection map.

The corrective map also depends on another factor, a scalar that we call the force factor. In fact, the corrective map is the top-down map whose response strength is scaled by the force factor. The force factor is the output of the attention controller shown in [Figure 1](#). The controller receives an error signal, denoted by $\mathbf{e}(\mathbf{t})$, which is the norm of the difference between the measured gaze and the top-

down attention's predication. It then outputs a force to minimize error. The computation of the force factor from error depends on the type of the controller. We suggest a classic Proportional Integral Derivative (PID) controller for the early investigation because of its simplicity and effectiveness.

Bipolar Disorder

Bipolar disorder is a condition where depressive episodes and elevated and manic mood swings happen recurrently in the patient. This condition not only causes problems for the patients, but also for the people near to them. About one percent of the United States adult population suffers from bipolar disorder [9]. Bipolar disorder is also associated with a high suicide rate [8]. Treatment for bipolar disorder helps mitigate the amplitude of the mood swings. Until recently, it was believed that bipolar disorder only occurred in adults. Researchers have now found that children can develop this disorder as well [15].

Sprott [17] has suggested a simple model that represents mood by a second order differential equation. In his model, shocking events with strong emotional impact are modeled by an impulse that results in mood oscillation. In fact, empirical studies on subjects who experienced emotional shock such as lottery winners and accident victims [3] report similar mood oscillations. This consistency improves plausibility of Sprott's model. Like other second order systems, this model depends on the damping factor parameter, which defines how fast the oscillation settles.

Sprott believes that in emotionally healthy people the damping factor is large enough to minimize the effect of shock. He thinks this parameter is smaller in emotionally sensitive people and negative in bipolar people. A negative damper actually acts as an amplifier. Therefore, small emotional noises are drastically enlarged to the peaks (depression and mania) in bipolar patients. Daugherty *et al* [6] proposed a more advanced model however, it works in essentially the same manner due to its second order nature.

The damping parameter makes the reviewed models flexible in order to cover people with diverse emotional characteristics and different severities of bipolar disorder. This parameter can be estimated for each individual by fitting his or her real data into the model. For adults, this data is usually collected by self-assessment using a questionnaire. However, self-assessment with children does not provide accurate results. Here HCI can help measure the emotional state of the patient.

Methodology

Our proposed method uses image processing and pattern recognition techniques to recognize child's emotional state. This data is then fitted to an emotion model which predicts future states of the user. In order to stabilize user's mood, the system tries project the opposite mood of the child on his or her state. This projection is achieved by using an attractive and influential synthetic character capable of expressing some basic emotions. This time hardware is simpler; a normal monitor and a camera mounted somewhere in front of the user are enough.

There are two broad categories for emotion measurement using a computer:

- Observable signals such as voice, facial expressions, and body movement.
- Physiological signals such as heart beat and skin conductance.

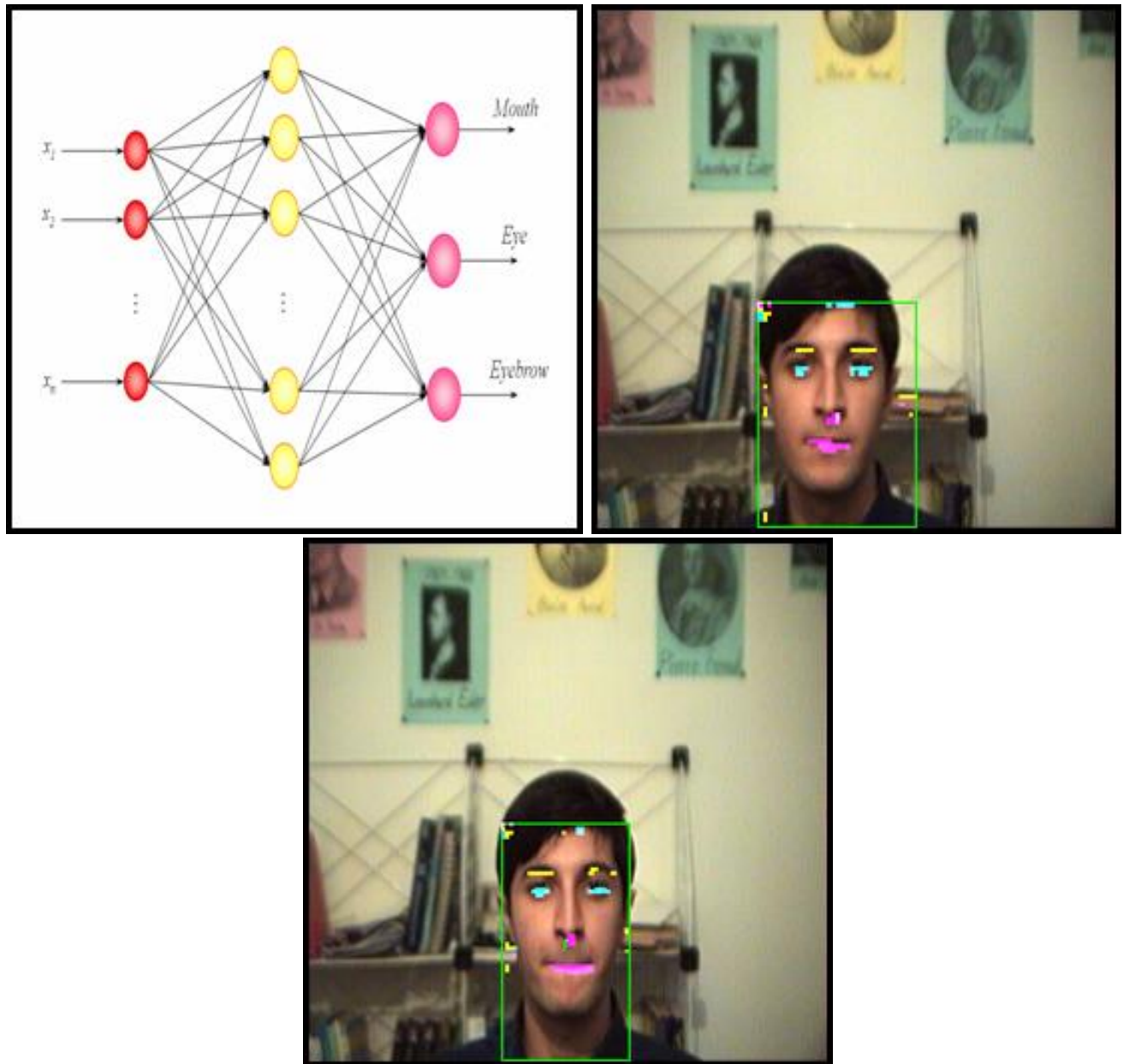


Figure 3: Results of our software for emotion recognition.

Due to its intrusive nature, the latter approach is less preferred. We use a non-intrusive computer program for emotion recognition with real-time performance [12] (Figure 3). It localizes the face using color and motion cues and then detects facial features via a neural network. These features are tracked over time and classified to discrete emotions using another neural network. This program helps to collect real data about subject's emotional state. Hence, the next step is to fit this data into Sprott's model for estimating the damping factor of a subject.

We used a 3-layer Perceptron for detection of facial features. Once face is detected, it is scaled to a 35x35 image and scanned by a 7x7 sliding window. The neural network has 49 input nodes wired to

the grey value of pixels within the sliding window. The output of the network determines the type of facial feature that falls in the sliding window. There were eight neurons for hidden layer and the network was trained by error back propagation algorithm.

Due to the attraction of cartoons and animation in children, a synthetic character can possibly be an effective interface for interaction with a child. Consequently, it can be a means for influencing the child's emotional state. Recently, Creed *et al.* began an investigation in the HCI domain about the influence of synthetic emotions on a user's attitudes and behavior [5]. Once such an influential character is created, it needs an engine to control its emotion expression.

Now that the mood model and its parameter(s) are determined for a given individual, future states of the person can be predicted. Using this prediction, we suggest developing an emotional engine based on a similar oscillator but in anti-phase. It should be investigated whether the latter oscillator can cancel out or reduce mood swings and stabilize the mood of a patient.

Conclusion

We proposed autism and bipolar disorder are two mental conditions where HCI techniques can potentially provide aid. We addressed social problems in children with an emphasis on attention deficit. We then developed a model for controlling attention. This model automatically looks for socially stimulating things like human faces in the environment. Once found, it projects a manipulated view of the environment where those things look more salient. Hardware requirements are a camera for gaze tracking and a comfortable head mounted display.

There are simple mathematical models of bipolar disorder in literature, which can be used by for developing tools for bipolar children. We linked these models to HCI capacities for reading and projecting emotions. The idea here was to use a mood model for predicting swings and projecting the opposite mood through a synthetic character to a subject's emotional state. A camera for capturing child's facial expressions is required.

We hope this work has been able to shed light on the potential of HCI in aiding children with mental disorders and encouraging researchers in exploring this area.

Acknowledgment

We would like to thank Anthony Bergstrom and Matthew Yapchaian for their constructive feedbacks. Anthony also did a great help in revising this article.

References

- 1
Baluja, S., Pomerleau, D., "Non-intrusive gaze tracking using artificial neural networks", *Technical Report CMU-CS-94-102*, Carnegie Mellon University, 1994
- 2
Braun, J., Julesz, B., "Withdrawing attention at little or no cost: detection and discrimination

tasks", *Percept Psychophys*, 1998,60:1-23

Brickman, P., Coates, D., Janoff-Bulman, R. , "Lottery winners and accident victims: Is happiness relative?", *Journal of Personality and Social Psychology*, 36, 917-927, 1978.

Christian, J., "Autism and Related Disorders Handbook", *USD School of Medicine and Health Science Center for Disabilities*, 2002, <http://www.usd.edu/cd/autism/Autism%20Handbook.pdf>

Chris Creed, Russell Beale, "Using Emotion Simulation to Influence User Attitudes and Behavior", *To Appear in HCI workshop at BCS 2005*

Daugherty, D., Roque-Urrea, T., Urrea-Roque, J., Snyder, J., Wirkus, S., Porter, M. A., "Mathematical Models of Bipolar Disorder", *eprint arXiv:nlin/0311032*.

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., Brown, E., "Children with autism fail to orient to naturally occurring social stimuli", *Journal of Autism and Developmental Disorders*, 28, 479-485, 1998.

Dilsaver, S.C., Chen, Y., Swann, A.C., Shoaib, A.M., Krajewski, K.J., , "Suicidality in patients with pure and depressive mania", *American Journal of Psychiatry*. 151, 1312-1315, 1994.

Griswold, K., "Management of Bipolar Disorder", *American Family Physician*, September 15, 2000.

Itti, L., "Models of Bottom-Up and Top-Down Visual Attention", *California Institute of Technology*, Jan 2000.

Koch, C., Ullman, S. ,. "Shifts in selective visual attention: towards the underlying neural circuitry", *Human Neurobiology*, 4(4), 219-227, 1985.

Mobahi, H., "Building an Interactive Robot Face from Scratch", *Bachelor of Engineering Final Project Report*, Azad University, Tehran-South Campus, Tehran, Iran, May 2003.

Mundy, P., Sigman, M., Ungerer, J., Sherman, T., "Defining the social deficits of autism: The contribution of non-verbal communication measures", *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 27, 657-669, 1986.

Norcia, N., "Innovative research advances knowledge and treatment of mood disorders", *Sunnybrook & Women's News*, Vol. 14, Aug. 16, 2004.

Papolos, D., Papolos, J., *The Bipolar Child*", New York: Broadway Books, December, 1999.

Picard, R. W., *"Affective Computing*, MIT Press, Cambridge, MA, 1997.

Sprott, J. C., "Dynamical Models of Happiness", *Nonlinear Dynamics, Psychology, and Life*

- 18 *Sciences* 9, 23-36 ,2005.
- Stiefelhagen, R., Yang, J., Waibel, A., "Tracking eyes and monitoring eye gaze", *Workshop on Perceptual User Interfaces*, Banff, Canada, 1997.
- 19 Werry, I., Dautenhahn, K., "Towards Interactive Robots in Autism Therapy: Background, Motivation and Challenges", *Pragmatics and Cognition* 12(1), pp. 1-35, 2004.
- 20 Yoon, H., "Using Robots to End AI: Artificial Inference", *Yale Scientific Magazine*. 78, 2 (Winter 2004).

Biographies

Hossein Mobahi (hmobahi2@uiuc.edu) is a graduate student in the Computer Science department at the University of Illinois at Urbana-Champaign. His research interests include human-robot interaction, computer vision and pattern recognition.

Karrie G. Karahalios (kkarahal@cs.uiuc.edu) is an assistant professor in the Computer Science department at the University of Illinois at Urbana-Champaign. Her work focuses on the interaction between people and the social cues they perceive in networked electronic spaces. She completed a S. B. in electrical engineering, an M.Eng. in electrical engineering and computer science, an S.M. in media arts and science, and a Ph.D. in media arts and science at MIT.