Title:

HAPTIC GLOVE: VIBROTACTILE DELIVERY OF FACIAL EXPRESSIONS FOR PEOPLE WHO ARE BLIND

Authors:

Sreekar Krishna, Shantanu Bala, Troy McDaniel, Steven McGuire, and Sethuraman Panchanathan.

Affiliation:

Center for Cognitive Ubiquitous Computing (CUbiC), School of Computing, Informatics and Decision Systems Engineering (CIDSE), Arizona State University, 699 S. Mill Ave, Tempe, AZ 85281.

Contact:

Sreekar.Krishna@asu.edu, Shantanu.Bala@asu.edu, troy.mcdaniel@asu.edu, and panch@asu.edu. **Ph:** 480 727 3612. **Fax:** 480 965 1885

Project Summary:

Introduction and Motivation:

Nonverbal communication cues make up nearly 65% of all human interpersonal communication [1]. Unless the communication is being carried out remotely (through telephone, instant messenger, social networking, etc), people tend to enrich their communication through the use of head and body postures or gestures. Some of the head based communicative cues include head nod, head shake, head tilt, eye gaze, and facial expressions, while, the body based communicative cues include hand movements, hand and leg crossing, overall body posture, appearance and proxemics (how people use and perceive the physical space around them to convey social, personal and intimate emotions). Most part of such head and body based cues are visual in nature and people who are blind cannot independently access this visual information, putting them at a disadvantage in daily personal and professional social encounters [2].

Of all these visual nonverbal cues, facial expressions play a very important role in everyday interpersonal communication. For example, during bilateral conversations, smile can convey agreement, while a raised eyebrow can convey questioning. Very subtle movements on the facial features can convey rich and varied cues like desire, confusion, acceptance, anger, frustration, excitement, surprise, interest and even deception. In a professional setting, people who are blind and visually impaired find these nonverbal cues to be very inaccessible. To compound these problems, sighted individuals are often unaware of their non-verbal cues and often do not (or

cannot) make appropriate adjustments when communicating with people who are blind. Most individuals who are visually impaired find alternate modes for understanding some of these cues, especially through heightened auditory perception of sounds created by body movement. But facial movements are too subtle for any such auditory coding, unless they are co-conveyed through prosody. Thus, any form of access to nonverbal cues can increase involvement of people who are visually impaired into the typical construct of human interpersonal communication.

We conducted an online survey of individuals who are blind or visually impaired to better understand their needs during social interactions [3]. Among the various requirements identified by the participants, access to facial expressions ranked on the top three priorities. In the past, we have worked towards addressing various needs [5] [4] through our Social Interaction Assistant project, which is a computer vision-based, wearable assistive device for individuals who are blind. The main components of this system include a pair of sunglasses with an embedded camera, on-body motion sensor, a computing element, and a vibrotactile belt that can convey elementary non-verbal communication cues [6] while also delivering rehabilitative feedback towards reducing stereotypic body mannerisms [5]. Adding to the capabilities of the Social Interaction Assistant, the Haptic Glove addresses the need to access facial expressions during interpersonal communicative conversations.

Design Considerations:

People who are blind rely on their auditory senses to understand and comprehend the environment around them. Assistive technologies that use audio cues to deliver information back to a user can cause sensory overload leading to the rejection of any benefits that a device might offer. Especially during social interactions and bilateral conversations, it is imperative that any device should not hinder the primary sensory channel of the user. To this end, the proposed facial expression delivery mechanism uses the somatosensory (touch) system to deliver facial expressions to the user while he/she is engaged in bilateral interactions. Currently, the device only delivers the 6 basic expressions (Smile, Anger, Disgust, Surprise, Sad and Fear) [1] along with indications of when the face reaches neutral expression. In future, we plan to encode the dynamic motion of the human facial features into vibrotactile patterns. This would allow indiscriminate access to the facial movements of the interaction counterpart. Due to the lack of space we do not discuss the details of the design process, but we introduce, in brief, the construction of the haptic device, the vibotactile structure of the expressions and describe an experiment carried out on the haptic glove.

Construction of the Haptic Glove:

The haptic glove has 14 tactors (vibration motors) mounted on the back of the fingers, one per phalange. The 14 motors correspond to the 14 phalanges (3 each on the index finger, middle finger, ring finger and the pinky with 2 on the thumb) on the human hand. A controller is also

integrated on the glove to allow control of the motor's vibration (magnitude, duration and temporal rhythm) through the USB port of a PC.
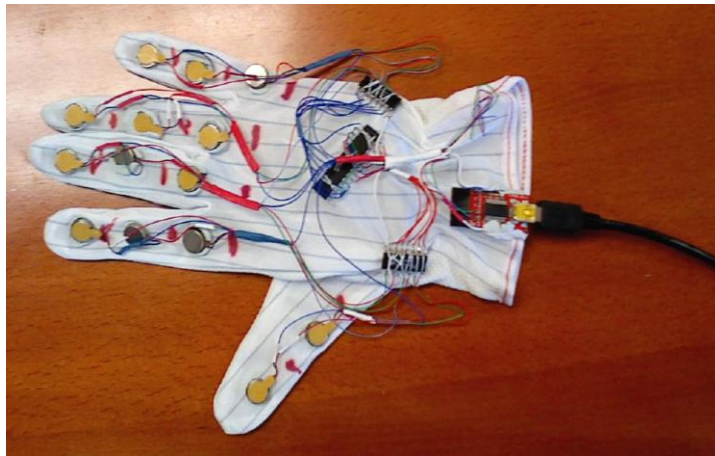

Figure 1: *Accessible Caption*: Haptic Glove: The figure shows a glove made out of stretchable material with 14 motors on the back of the glove with each motor corresponding to one phalange of the 5 digits. A microcontroller, two motor drivers and 1 USB controller (4 ICs) are also integrated on to the back of the glove with an ultra thin flexible USB cable leaving the glove.

Haptic Encoding of Expressions:

In order to encode the 6 basic expressions and neutral facial posture into haptic cues, we resorted to popular emoticon representations of these basic expressions. For example, smile is popularly represented by a smiley which was translated to a vibratory pattern of index finger top phalange, followed by middle finger bottom, followed by ring finger top phalange. The entire vibration sequence was completed within 750 milliseconds (The duration was arrived at after careful pilot studies with participants). The table below gives the vibration finger and phalange location in comma separated sequence for all 7 facial expression postures.

Expression	Comma separated vibration sequence. All sequences are 750ms long First letter indicates the finger – I for index, M for middle and R for Ring Second letter indicates the phalange – T for top, M for middle and B for bottom
Smile	IT, MB, RT
Sad	IB, MT, RB
Surprise	MT, IM, MB, RM, MT
Anger	IM, IB, MM, MB, RM, RB
Neutral	IM, MM, RM
Disgust	RB, MB, IB
Fear	IT, MT, RT, MT, IT, MT, RT

Experiment:

The above expressions were conveyed to 10 participants one of whom is blind. The participants were trained on the expressions and 70 stimulations (10 trails of each expression) were presented sequentially with 5 seconds gap for the user to respond. The table below represents the results as a 7x7 confusion matrix where each cell entry corresponds to how many times (on average) users when given the row expression as stimulation responded with the column expression as their answer. Following this average number, separated by a comma is the average time taken for answering.

		Response						
		Angry	**Disgust**	**Fear**	**Smile**	**Sad**	**Surprise**	**Neutral**
Stimulation	**Angry**	88, 2.28	-	-	1, 4.45	-	11, 2.99	-
	Disgust	-	95, 1.89	2, 1.52	-	2, 3.44	-	1, 1.67
	Fear	-	1, 2.64	98, 1.66	-	-	1, 3.62	-
	Smile	-	-	-	88, 2.12	5, 4.18	7, 2.8	-
	Sad	4, 2.54	-	-	2, 2.32	82, 2.67	10, 2.68	2, 2.38
	Surprise	10, 3.12	-	2, 2.71	-	1, 3.48	87, 2.52	-
	Neutral	2, 2.69	-	1, 2.94	3, 3.6	8, 3.56	-	86, 2.23

Table 1: Rows represent the stimulation provided to the users and the columns represent the response provided by the user. Each cell has two numbers. The first number represents the percentage recognition of a specific stimulation and a corresponding response. The second number represents the average time taken in seconds for that specific stimulation and response. Ideally this matrix should have 100% recognition along the diagonal with as low a time as possible.

References:

[1] M.L. Knapp and J.A. Hall, *Nonverbal communication in human interaction*, Wadsworth/Thomson Learning, 2005.

[2] S. Sacks, L. Kekelis, R. Gaylord-Ross, *The development of social skills by blind and visually impaired: exploratory studies and strategies*, AFB Press, New York, 1992.

[3] S. Krishna, D. Colbry, J. Black, V. Balasubramanian and S. Panchanathan. "A Systematic Requirements Analysis and Development of an Assistive Device to Enhance the Social Interaction of People Who are Blind or Visually Impaired." Workshop on Computer Vision Applications for the Visually Impaired (CVAVI 08), A satellite workshop of European Conference on Computer Vision (ECCV 08), Marseille, France, Oct 18, 2008.

[4] S. Krishna, G. Little, J. Black and S. Panchanathan, "A wearable face recognition system for individuals with visual impairments," In Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 106-133, 2005.

[5] S. Krishna, C. K. Narayanan and S. Panchanathan, "Detecting Stereotype Body Rocking Behavior through Embodied Motion Sensors", RESNA Annual Conference, New Orleans, Lousiana, June 2009.

[6] T. McDaniel, S. Krishna, V. Balasubramanian, D. colbry, and S. Panchanathan, "Using a Haptic Belt to convey Non-Verbal communication cues during Social Interactions to Individuals who are Blind," Ottawa, Canada: 2008.