Fall 10-22-2012

Combining Emotion and Facial Nonamanual Signals in Synthesized American Sign Language

Jerry Schnepp
Bowling Green State University, schnepp@bgsu.edu

Rosalee Wolfe
DePaul University

John McDonald

Jorge Toro

Combining Emotion and Facial Nonmanual Signals in Synthesized American Sign Language

Jerry Schnepp, Rosalee Wolfe,
John C. McDonald

School of Computing, DePaul University
243 S. Wabash Ave.
Chicago, IL 60604 USA

{jschnepp,rwolfe,jmcdonald}@cs.depaul.edu

Jorge Toro

Department of Computer Science
Worchester Polytechnic Institute
100 Institute Road
Worcester, MA 01609 USA

jatoro@wpi.edu

ABSTRACT

Translating from English to American Sign Language (ASL) requires an avatar to display synthesized ASL. Essential to the language are nonmanual signals that appear on the face. Previous avatars were hampered by an inability to portray emotion and facial nonmanual signals that occur at the same time. A new animation system addresses this challenge. Animations produced by the new system were tested with 40 members of the Deaf community in the United States. For each animation, participants were able to identify both nonmanual signals and emotional states. Co-occurring question nonmanuals and affect information were distinguishable, which is particularly striking because the two processes can move an avatar's brows in opposing directions.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing – language generation, machine translation; K.4.2 [**Computers and Society**]: Social Issues – assistive technologies for persons with disabilities.

General Terms

Design, Experimentation, Human Factors, Measurement.

Keywords

Accessibility Technology, American Sign Language

1. INTRODUCTION

An automatic English-to-ASL translator would help bridge the communication gap between the Deaf and hearing communities. Text-based translation is incapable of portraying the language of ASL. A video-based solution lacks the flexibility needed to dynamically combine multiple linguistic elements. A better approach is the synthesis of ASL as animation via a computer-generated signing avatar. Several research efforts are underway to portray sign language as 3D animation [1][2][3][4], but none of them have addressed the necessity of portraying affect and facial nonmanual signals simultaneously.

2. FACIAL NONMANUAL SIGNALS

Facial nonmanual signals appear at every linguistic level of ASL [5]. Some nonmanual signals carry adjectival or adverbial information. Figure 1 shows the adjectival nonmanuals OO (small) and CHA (large) demonstrated by our signing avatar.

Nonmanual OO – "small size" Nonmanual CHA – "large size"

Figure 1: Nonmanual signals indicating size

Other nonmanuals operate at the sentence level [6]. For example, raised brows indicate yes/no questions and lowered brows indicate WH-type (who, what, when, where, and how) questions.

Affect is another type of facial expression which conveys emotion and often occurs in conjunction with signing. While not strictly considered part of ASL, Deaf signers use their faces to convey emotions [7]. Figure demonstrates how a face can convey affect and a WH-question simultaneously.

WH-question, happy WH-question, angry

Figure 2: Co-occurrence

3. SYNTHESIZING CO-OCCURANCE

We characterize linguistic facial nonmanual signals and affect poses as a set of facial muscle transformations which combine to create facial animations. We use a framework that represents syntax, lexical modifiers and affect as separate, but co-occurring influences on the position and timing of subordinate geometric components. This has the flexibility to synthesize novel utterances. See [8] for implementation details.

4. INITIAL EVALUATION

An initial study measured the perceptibility of affect in the presence of co-occurring nonmanual signals that could potentially interfere. For this, we created two pairs of sentences. Each pair consisted of one sentence with happy affect and one sentence with angry affect. The first pair combined the WH-nonmanual with each of these emotions. The second pair combined the CHA nonmanual with the same two emotions.

Twenty people participated in a face-to-face setting at Deaf Nation Expo in Palatine Illinois, and another twenty were recruited through Deaf community websites and tested remotely using SignQUOTE [9]below. All participants self-identified as members of the Deaf community and stated that ASL is their preferred language. In total, 40 people participated. Participants viewed animations of synthesized ASL utterances and were asked to repeat the sentence, rate its clarity, and identify the emotion in the animation using a five-point Likert scale. All testing was conducted in ASL.

5. RESULTS

For each animation, every participant repeated the utterance correctly. This included all of the processes that occurred on the face. Seventy-eight percent rated the WH-Happy animation as clear or very clear while sixty-five percent indicated that the WH-Angry animation was clear or very clear. For both animations combining the CHA nonmanual signal with either happy or angry affect, seventy five percent of participants indicated the animations were clear or very clear.

The majority of participants perceived the intended affect in each animation. Figure 3 displays the perceived affect for the WH-Happy and WH-Angry animations. Data for the perceived affect of the CHA-Happy and CHA-Angry animations are similar.

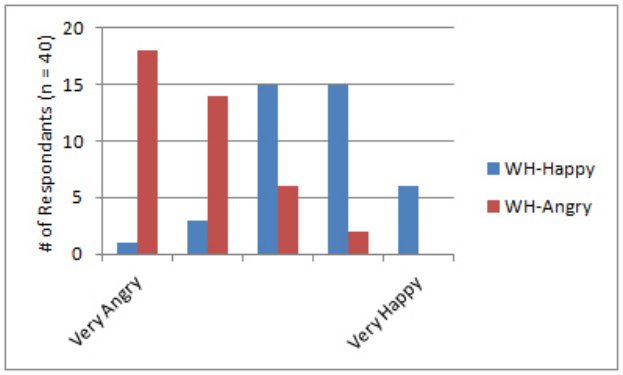

Figure 3: Perception of emotion in the presence of a WH-question nonmanual signal

6. DISCUSSION AND FUTURE WORK

In the case where the WH-nonmanual occurs simultaneously with happy affect, the brows are influenced by both in a competing manner. The WH-nonmanual tends to pull the brows downward, but a happy affect tends to push the brows upward. Despite these opposing influences, seventy-eight percent rated the animation as clear or very clear. This shows that the new technique has promise for portraying both affect and co-occurring nonmanual signals that are recognizable to members of the Deaf community.

Going forward, we plan to develop and evaluate additional nonmanual signals and follow up with more rigorous testing.

7. REFERENCES

[1] VCom3D. n.d. *Sign Smith 3.0 user manual.* http://www.vcom3d.com/docs/Sign_Smith_Studio3.0_UsersGuide.pdf.

[2] Huenerfauth, M., Lu, P. and Rosenberg, A. 2011. Evaluating importance of facial expression in American Sign Language and pidgin signed English animations. In *Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, UK, October 22 - 24, 2011). ASSETS'11. ACM, New York, NY, 99-106. DOI= http://doi.acm.org/10.1145/2049536.2049556.

[3] S. Gibet, N. Courty, K. Duarte, and Le Naour, T. 2011.The signcom system for data- driven animation of interactive virtual signers: Methodology and evaluation. ACM Transactions on interactive intelligent systems. 1, 1 (Oct. 2011), 1-26. DOI= http://doi.acm.org/10.1145/2030365.2030371.

[4] Elliott, R., Glauert, J. and Kennaway, J. 2004. A framework for non-manual gestures in a synthetic signing system. In *Proceedings of the Second Cambridge Workshop on Universal Access and Assistive Technology* (Cambridge, UK, March 22 - 24, 2004). CWUAAT 04. 127–136. http://www-edc.eng.cam.ac.uk/cwuaat/04/39-pat-elliott_cwuaat-2004.pdf.

[5] Valli, C., Lucas, C. and Mulrooney, K. 2005. *Linguistics of American Sign Language: An Introduction.* 4th ed. Washington, DC: Gallaudet University Press.

[6] Neidle, C., Kegl, J., Maclaughlin, D., Bahan, B., and Lee, R. 2000. *The Syntax of American Sign Language: Functional Categories And Hierarchical Structure*. Cambridge, Massachusetts: MIT Press.

[7] Weast, T. 2008. Questions in American Sign Language: A quantitative analysis of raised and lowered eyebrows. Doctoral dissertation. The University of Texas, Arlington.

[8] Schnepp, J. 2012. A representation of selected nonmanual signals in American Sign Language. Doctoral dissertation. DePaul University, Chicago, Illinois.

[9] Schnepp, J., Wolfe, R., Shiver, B., McDonald, J., and Toro, J. 2011. SignQUOTE: A remote testing facility for eliciting signed qualitative feedback. In *Proceedings of the Second International Workshop on Sign Language Translation and Avatar Technology* (Dundee, UK, October 23, 2011) http://vhg.cmp.uea.ac.uk/demo/SLTAT2011Dundee/4.pdf.