Annotation of Nonmanual Signals for Automatic Sign Language Generation

Rosalee Wolfe, John C. McDonald, Melissa Bialek
School of Computing, DePaul University, Chicago, IL USA
{wolfe, jmcdonald}@cs.depaul.edu, mbialek1@mail.depaul.edu

Language: English
Subject area: 3, using annotated material for working on translations

Effective translation from spoken to signed language requires a translation target that is visual in modality. Thus, the target must be a medium that supports the portrayal of human motion. Researchers are currently exploring a rich variety of media and sign generation systems that could serve as a translation target. Video recordings of signers enjoy the highest degree of visual realism, including all nonmanual signals. Coupled with annotation software such as ELAN [1] or iLex [2] , video recordings can provide the basis for corpus-based translation systems. However, video suffers from inflexibility. It cannot produce new, naturally flowing sentences from previously recorded ones [3] or nor can it anonymize the signer [4].

Annotated motion capture [5] [6] can drive an avatar, which addresses the anonymization problem, but even with the most sophisticated mocap editing techniques [7], the resulting massive amounts of data prevent the production of new animations in real time. Approaches that utilize less data can achieve real-time animation, but at the cost of lower realism. These approaches include interpolation from a library of signs defined by sparse key-frame data [8] [9] [10], constraint-based procedural animation [11] or a combination of both approaches [12].

A commonality among these systems is a reliance on annotation. Transcription systems such as HamNoSys [13] and annotation guidelines, including the ECHO [14] and Auslan corpus guidelines [15], have been a rich resource for developing sign generation systems. However, most current annotation guidelines are theory neutral, and define annotation codes in terms of phonetic and/or phonological forms only.

Unfortunately, a theory-neutral coding does not account for co-occurring linguistic and pragmatic processes that can influence the same anatomic feature. For example, consider an American Sign Language (ASL) signer who is producing the sign PAH in a yes-no question with a surprised affect. A slight upwards motion of the brows often co-occurs with the jaw drop in PAH. In addition, an upwards motion of the brows is obligatory for a yes-no question, and the brows also rise with a surprised affect. However, a theory-neutral coding system only notes that the eyebrows are raised. However, translation requires knowledge of what process or processes cause the brows to rise.

Other current projects are developing techniques for converting annotation codes to geometry and understanding the interaction between nonmanuals and manuals. Literal geometric interpretations of annotation codes do not yield satisfying results. We hope to discuss this as well as consider methods for separating NMS, particularly those on the spine, from manual parameters. This is not straightforward, as the spine influences hand position.

Our current approach to signed language generation allows for co-occurring linguistic and pragmatic nonmanual processes, and we are investigating methods for interpreting NMS codes in a way that produces visual results that are legible, acceptable and natural in appearance. At the workshop, we hope to explore annotation alternatives that indicate how co-occurrences jointly influence selected facial features.

Bibliography

1 Crasborn, Onno and Sloetjes, Han. Enhanced ELAN functionality for sign language corpora. In *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora* (2008), 39-43.

2 Hanke, Thomas. iLex-A tool for Sign Language Lexicography and Corpus Analysis. In *LREC* (Las Palmas 2002), ELRA.

3 Gil, Sangyong and Park, Jong. Revisiting Concatenative Video Synthesis with Relaxed Constraints. In *Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT)* (Dundee, UK 2011), ACM.

4 Kipp, Michael, Heloir, Alexis, and Nguyen, Quan. Sign language avatars: animation and Comprehensibility. *Intelligent Virtual Agents. Lecture Notes in Computer Science*, 6895 (2011), 113-126.

5 Kennaway, J. R., Glauert, R. W., and Zwitserlood, Inge. Providing signed content on the Internet by synthesized animation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 14, 3 (2007), 15-1 - 15-26.

6 Huenerfauth, Matt. Improving spatial reference in American sign language animation through data collection from native ASL signers. *Universal Access in Human-Computer Interaction. Applications and Services* (2009), 530-539.

7 Awad, Charly, Courty, Nicolas, Duarte, Kyle, Le Naour, Thibaut, and Gibet, Sylvie. A combined semantic and motion capture database for real-time sign language synthesis. *Intelligent Virtual Agents,* (2010), 432-438.

8 Efthimiou, Eleni, Fotinea, Stavroula-Evita, and Sapountzaki, Galini. Feature-based natural language processing for GSL synthesis. *Sign Language & linguistics*, 10, 1 (2007), 1-21.

9 San-Segundo, Rubèn, Montero, J. M., Córdoba, R., Sama, F., Fernández, F., and D'Haro, L. F. Design, Development and Field Evaluation of a Spanish into Sign Language Translation System. *Pattern Analysis and Applications*, 15, 2 (2012), 203-224.

10 Ebling, Sarah. Evaluating a Swiss German Sign Language Avataramong the Deaf Community. In *Third International Symposium on Sign Language Translation and Avatar Technology (SLTAT)* (Chicago, IL 2013).

11 Braffort, Annelies and Dalle, Patrice. Sign language applications: preliminary modeling. *Universal access in the information society*, 6, 4 (2008), 393-404.

12 Wolfe, Rosalee, Cook, Peter, McDonald, John C., and Schepp, Jerry. Linguistics as structure in computer animation: Toward a more effective systhesis of brow motion in American Sign Language. *Sign Language & Linguistics*, 14, 1 (2011), 179-199.

13 Hanke, Thomas. HamNoSys -- Representing sign language data in language resources and language processing contexts. In *Fourth International Conference on Language Resources and Evaluation (LREC 2004). Representation and Processing of Sign Languages Workshop* (Paris 2004), European Language Resources Association, 1-6.

14 Nonhebel, Annika, Crasborn, Onno, and van der Kooij, Els. *Sign language transcription conventions for the ECHO Project*. http://sign-lang.ruhosting.nl/echo/docs/ECHO_transcr_conv.pdf, 2004.

15 Johnston, Trevor. *The Auslan Corpus Annotation Guidelines*. Centre for Language Sciences, Department of Linguistics, Macquarie University, Sydney, Australia, 2009.