Database Design for American Sign Language

Jacob Furst, Karen Alkoby, Andre Berthiaume, Pattaraporn Chomwong, Mary Jo Davidson, Brian Konie, Glenn Lancaster, Steven Lytinen, John McDonald, Lopa Roychoudhuri, Jorge Toro, Noriko Tomuro, Rosalee Wolfe

School of Computer Science, Telecommunications and Information Systems
DePaul University
243 South Wabash Avenue
Chicago, IL 60604-2301
(312)-362-5158
FAX: (312)-362-6116
asl@cs.depaul.edu

Introduction

American Sign Language (ASL) is a natural language used by members of the North American Deaf community and is the third or fourth most widely used language in the United States [Ster96]. At present deaf people rely on sign language interpreters for access to spoken English, but cost, availability and privacy issues make this an awkward solution at best. A digital sign language interpreter, which translates spoken English into ASL, would better bridge the gulf between deaf and hearing worlds.

Current technology for the translation of written English includes closed captioning on television and TDD. These are good first efforts at making spoken English more accessible to the deaf population, but do not represent a completely satisfactory solution. While ASL shares some vocabulary with English, there is no simple word-for-word translation. Further, research in linguistics shows that ASL's concise and elegant syntax differs radically from English grammar [Klim79][Vall93]. Because of the differences in the two languages, most native ASL signers read English at the third or fourth grade level [Holt94]. Again, a digital sign language interpreter, which translates written English into ASL, would better aid the deaf community.

Sign Language Interpreter

An English-to-ASL translator would convert written or spoken English into a three-dimensional graphic animation depicting ASL. As personal computers become lighter and cheaper, a digital sign language interpreter would provide an economical and flexible solution to the problem of deaf access to English.

Further, ASL synthesis technology could be used as a valuable tool for educators and researchers. By using three-dimensional graphics, signs can be viewed from any position, increasing understanding of the sign in space. By using animation, signs can be viewed through time, increasing understanding of the signs as part of complete sentences, rather than as simple vocabulary.

However, developing this kind of English-to-ASL translation system imposes serious technical challenges. Because of the unique modality of ASL as a sign language -- visual/gestural rather than aural/oral -- the translation system must store the linguistic and geometric aspects of ASL signs and generate graphic animations on screen in real time. Also, the sequence of signs in an ASL sentence must look smooth and natural. Therefore, building such an English-to-ASL translator requires expertise in a wide range of areas, including linguistics, machine translation, computer graphics, mathematics and kinesiology. This paper presents a nexus of the system: the database for storing geometric and linguistic information about signs and the transcription interface for creating signs.

Database Support

The unique nature of sign languages has largely prevented them from being stored electronically, unlike written languages such as English. The advantages of storing a language like American Sign Language (ASL) digitally are threefold: 1) the linguistic information enables researchers to make quick queries about lexical properties of the language, 2) the graphical information enables design programs to synthesize graphic animations of ASL signs, and 3) the dictionary information provides an online resource for English-ASL translations.

Our database scheme draws on the experiences of Dutch [Cras98], German [Pril89], and Japanese [Lu97] researchers who are working on similar projects for other sign languages. Our design is divided into two main tables: one to contain information about handshapes and one to contain other information used to generate complete signs. This includes position, orientation and shape of the hands as well as motions that comprise a sign. Both tables include raw geometric information as well as linguistic information.

The geometric information is used to create the three-dimensional graphical representation for a sign, while the linguistic information provides independence from the geometry and access to lexical information for researchers. The linguistic information for both handshapes and signs is derived from the works of Liddell & Johnson [Lidd89], Sandler [Sand89] and Brentari [Bren99] as well as geometric information from the interface.

The largest task in creating a database of this type is gathering the information once the database scheme has been created. While motion capture initially presents an appealing option for gathering data, it cannot record the linguistic aspects of a sign. It is analogous to scanning in a printed sentence and attempting to use the raw bitmap instead of extracting the characters. Further, the motion and position of a sign may change depending on its context. Another possibility would be to use a general animation package to transcribe signs. However, this approach also poses severe problems. Learning such a package requires a significant investment of time. Our students reported that working through the tutorials of a commonly used animation package took between 40 and 100 hours. Few members of the deaf community are willing to invest such a large amount of time in training before beginning the transcription process.

To reduce the large time investment for transcription, we have customized an animation system to accommodate ASL. Initial usability studies have been quite promising. On average, it takes 10 minutes for native ASL signers to learn enough about the system to input hand shapes, and transcribing a handshape takes an average of 82 seconds.

Transcription Interface

As with the database schema, the transcription system has a bi-level structure. The lower level, the handshape transcriber, is used to build the handshape data, which represents most of the geometric information contained in signs. The upper level, the sign transcriber, relies on the handshape database and allows users to specify the location and motion of both hands. The transcribers utilize familiar user interface elements such as checkboxes, selection lists and slider bars and place an emphasis on graphical, rather than textual labeling. Labels identify symbolic and linguistic features of signs and hide the underlying mathematic information.

The handshape transcriber [McD2000] allows users to specify handshapes. Users can specify linguistically significant aspects of hand shape one finger at a time (e.g. bend of a finger) as well as aspects of finger groupings (e.g. spread between fingers). The thumb is configured with its own set of controls, mirroring the unique use of the thumb in signing. Once created, both the linguistic and geometric properties of the handshape are saved to the handshape database.   The linguistic information can be saved in a format consistent with Liddell & Johnson, Sandler, or Brentari.

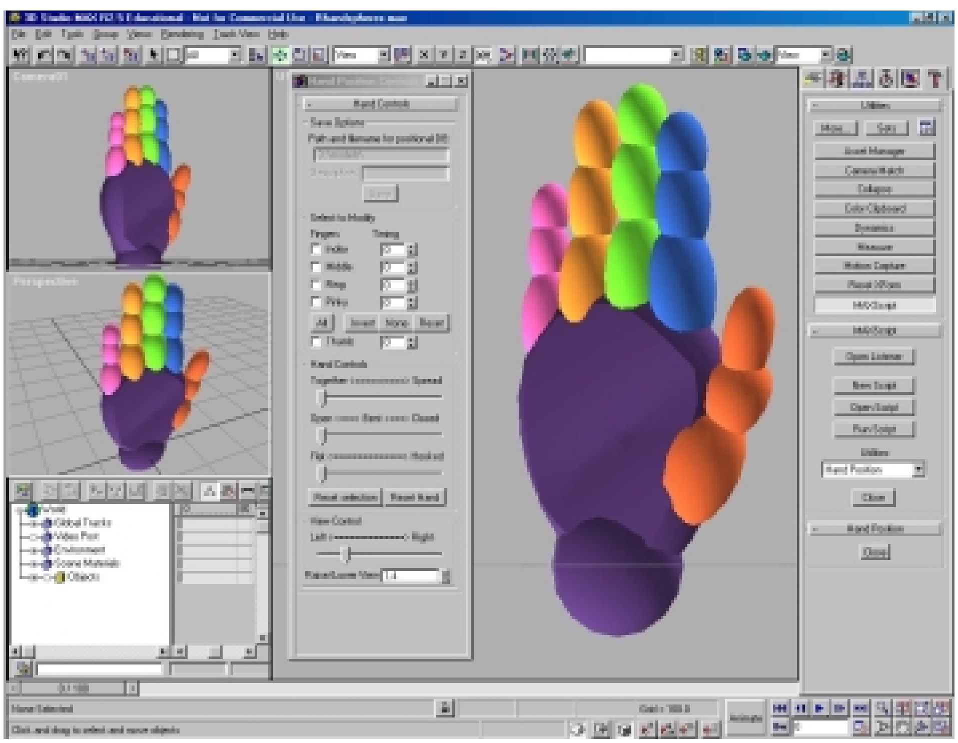

Figure 1 - Image of the handshape transcriber

The sign transcriber [Wolf99] allows users to specify complete ASL signs. Users can specify handshape, location and orientation for each hand. Signs with motion are entered by specifying this information for various time steps. The sign transcriber provides an animation option, allowing the user to preview a sign at any point during its creation.

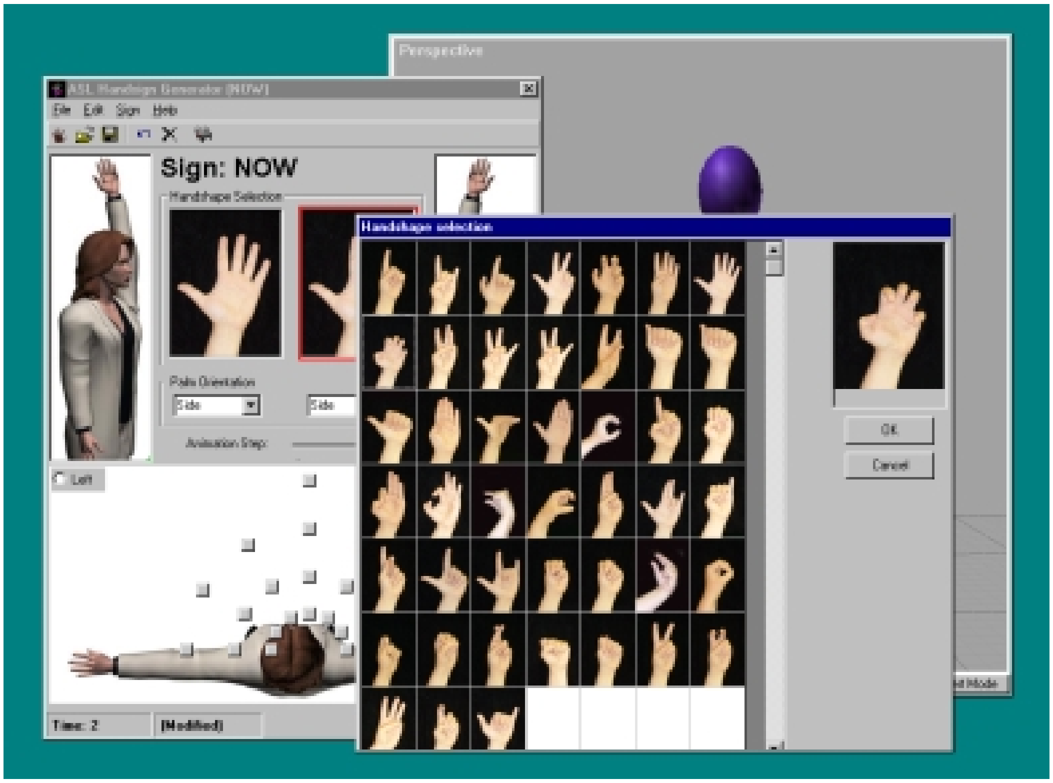

Figure 2 - Image of the sign transcriber

References

[Bren98] Brentari. *A Prosodic Model of Sign Language Phonology*. Cambridge, Massachusetts: MIT Press, 1998.

[Cras98] Onno Crasborn, Harry van der Hulst, and Els van der Kooij, SignPhon: A database tool for cross-linguistic phonological analysis of sign languages. Sixth International Conference on Theoretical Issues in Sign Language Research, 1998.

[Holt94] Holt, J., "Stanford Achievement Test - 8th Edition for Deaf and Hard of Hearing Students: Reading Comprehension Subgroup Results". http://www.gallaudet.edu/~cadsweb/sat-read.html

[Klim79] Edward Klima and Ursula Bellugi, *The Signs of language*. Harvard University Press, 1979.

[Lidd89] Liddell and Johnson. American Sign Language: The Phonological Base, *Sign Language Studies*, 64, 195-277.

[Lu97] Shan Lu, Seiji Igi, Hideaki Matsuo and Yuji Nagashima, Towards a Dialogue System Based on Recognition and Synthesis of Japan Sign Language, Gesture and Sign Language in Human-Computer Interaction, 1997.

[McD2000]

[Pril89] Prillwitz, Siegmund et al., HamNoSys Version 2.0; Hamburg Notation System for Sign Languages, *International Studies on Sign Language and Communication of the Deaf*; Signum 1989.

[Sand89] Sandler. *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Dordrecht, Holland: Foris Publications, 1989.

[Ster96] Martin Sternberg, *The American Sign Language Dictionary*. Multicom, 1996. (CD ROM)

[Vall93] Clayton Valli and Ceil Luca, *Linguistics of American Sign Language*. Gallaudet University Press, 1993.

[Wolf99] Wolfe, R. et al., "Interface for Transcribing American Sign Language". In Proceedings of the 26th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), Los Angels, 1999.