A Graphical Environment for Transcription of American Sign Language

Jorge Toro, Jacob Furst, Karen Alkoby, Roymieco Carter, Juliet Christopher, Brock Craft, Mary Jo Davidson, Damien Hinkle, Brian Konie, Glenn Lancaster, Steve Luecking, Ashley Morris, John McDonald, Eric Sedgwick, Noriko Tomuro, Rosalee Wolfe

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

Abstract

A system to interactively create and modify/edit American Sign Language signs is described. The system is grounded on the use of three-dimensional computer graphics to construct the signs. Usability tests have been conducted to obtain early feedback on the user experience with the system. The final goal is to build a personal digital translator for the deaf. Since ASL is a visual language, it is particularly important that the interface be visually efficient, and easy to use.

Keywords: Computer Graphics, Animation, American Sign Language

INTRODUCTION

American Sign Language (ASL) is a rich and varied natural language used by members of the North American Deaf community and is the third most widely used language in the United States [Ste96, Dea00]. To interact with the hearing world, the deaf community relies mainly on human interpreters. While ASL shares some vocabulary with English, it is not a direct translation of English words and sentence structure. It presents many of the same challenges of any language translation process, but adds the complexity of changing modality from aural/oral to visual/gestural [Alk99]. Since ASL is meant to be seen, visual clarity is a critical factor [Bak80].

Although researchers have studied the use of digital technology to simulate ASL [Mic99, Ste96, Su98], their approaches do not create the full range of the language. We believe that the use of computer graphics (CG) provides important advantages especially in the creation and presentation of ASL for conversation. As part of a personal digital translator, CG would provide greater access to conversations in the hearing world. For instance, in medical and legal matters it could facilitate confidential doctor-to-patient or attorney-to-client conversations without the need for an interpreter. An English-to-ASL translator would convert written or spoken English into three-dimensional graphic animations depicting ASL.

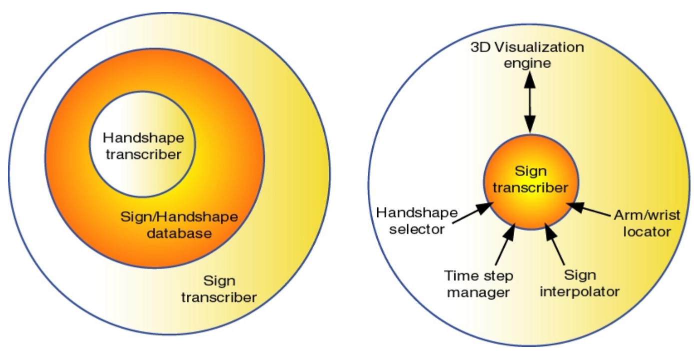

Figure 1. Architecture of the ASL transcribing system

The following is a description of a CG-based system to transcribe ASL signs. The system is an important step toward our goal of a digital translator for ASL.

THE MAIN ELEMENTS OF ASL

In ASL, a *sign* loosely corresponds to a word, but can express entire concepts and complex phrases. *Fingerspelling* is used to spell out proper names and technical terms [Kli79, Val93]. While additional elements may be present, there is a consensus among ASL linguists that the shape of the hand (handshape), as well as its location and movement are essential elements of a sign [Lid89]. Most signs are a sequence of these elements. *Handshapes* are particular configurations of the hand. A relatively small set of handshapes (40) generates the majority of signs in ASL [Ten98]. Comprehension of a sign depends on recognizing the handshape. Often, a slight change in a feature of a handshape could render it unrecognizable [Sto79].

ARCHITECTURE OF THE SYSTEM

The system (Figure 1) has a *handshape transcriber* module and a *sign transcriber* module. The *handshape transcriber* (Figure 2) is used to build the handshape data, which represents most of the geometric information contained in signs. This data is stored in the handshape database for use by the sign transcriber. This approach is useful since it allows users of the system to create a handshape only once and reuse it in different signs, making the sign construction process less complex. To create the handshapes, the transcriber uses a geometric model of the human hand that accurately simulates the complex behavior of the thumb [McD00].

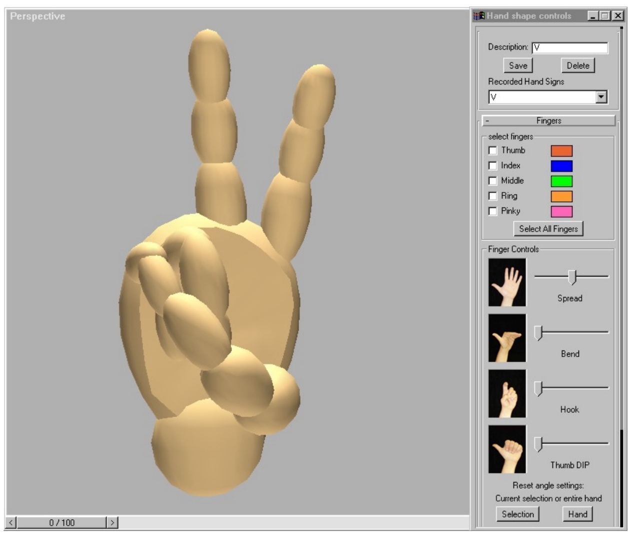

Figure 2. Handshape transcriber

The *sign database* scheme draws on the experiences of Dutch [Cra98], German [Pri89], and Japanese [Lu97], researchers who are working on similar projects for other sign languages. It is designed to contain detailed linguistic information for translation and geometric information suitable for creation of 3D ASL animations [Tom99, Fur00]. The geometric information includes position, orientation, and shape of the hands as well as motions that comprise a sign.

THE SIGN TRANSCRIBER

The sign transcriber (Figure 3) relies on the handshape database and allows users to create both static and animated signs by specifying the location and motion of both hands in 3D space.

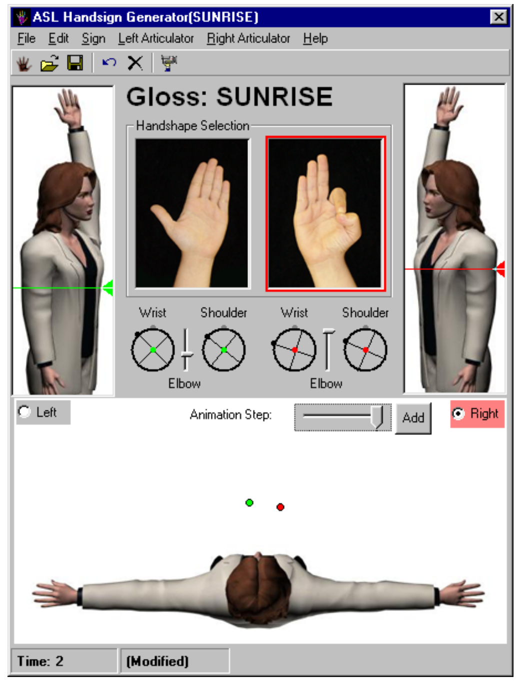

Figure 3. Sign Transcriber

The transcriber is composed of a graphic handshape selection tool, which lists the set of handshapes available for use (Figure 4). It also has controllers to locate the arm and wrist in 3D space, a time step manager to specify hand-position configurations at desired time intervals, and a sign interpolator to construct the intermediate positions between two or more signs.

To view the animations, the transcriber uses a *3D visualization engine* (Figure 5)*,* which uses a geometric representation of a human body to display the animations of the signs.

THE USER ENVIRONMENT

The design and construction of the interface has been done following a user-centered approach. Potential users of the transcriber have been involved from the early stages of design, to help us analyze important usability variables such as user preference, familiarity with the visual controls, and ease of use.

To minimize the use of modes, all the components of the transcriber have been placed on the main dialog of the interface (Figure 3). Each side of the body has its own set of controllers to guarantee a better visibility of the options available, and to allow the user to keep the locus of attention on the construction of the sign.

The interface has been conceived to take advantage of the user’s visual memory since ASL is a visual language. That is why the interface is highly visual, with controls that use the mouse as the primary input device. The user does not have to memorize command names or keywords, again reducing the learning curve and complexity of the system.

As mentioned before, the transcriber interacts with a 3D visualization engine to show the animations of the signs constructed with the transcriber. Here we faced a major challenge: to specify a 3D location using a 2D interface. This was not an easy task, especially since potential users of the system will not have previous experience with three-dimensional graphics packages. The next section describes our approach to this problem.

BUILDING A SIGN WITH THE SYSTEM

The sign transcriber captures the user actions on the interface and translates these actions into geometric data, which is sent to the 3D visualization engine for display. The communication between the interface and the graphic engine is done via Common Object Module (COM) technology.

The sign construction process can be separated into four major steps: locate the hands in space, select the handshapes, define intermediate hand-position configurations (if any), and instruct interpolator to animate the sign.

Hand location in space

To locate the hand in 3D space, the interface uses two controls (Figure 3). One controls the height and another one controls the location. The system uses a normalized triplet (r, θ, z) to represent the point in 3D space and to store it in the database. For visualization, the system transforms the triplet (r, θ, z) into a (x, y, z) value needed by the graphic engine. Using normalized coordinates allows the transcriber to store device-independent data that can be reused independently of the coordinate system the 3D engine is using.

Handshape Selection

The current handshapes assigned to both hands of the 3D model are displayed on two images located in the middle of the interface (Figure 3). To select a new handshape, the user clicks on the image of the one to be changed. This pops-up a list with a visual set of handshapes from which the user may select (Figure 4). The new handshape is now displayed on both the image and the 3D model.

In the current version of the transcriber, it takes a user a total of three mouse clicks to assign a handshape to one of the hands. This is enormously faster than specifying individual joint rotations on the hand manually.

Intermediate Locations

In ASL, the process of signing not only involves static configurations of the body, but also transitions between signs (movement). A sign can start in a particular configuration of the body and end in a different one. The time step manager allows defining multiple configurations of the body, which are all part of a sign. To define a new configuration, the user clicks on the **Add** button located at the top of the Arm Location image (Figure 2). At this point, the 3D visualization engine resets the human model to its default position for the user to define the new handshapes and hand locations for the new configuration. The status bar located at the bottom gives the user information about which intermediate sign is being edited.

Animating Signs

Once the configurations of a sign have been defined, the sign interpolator will create a smooth transition between them. The system uses cubic interpolation and forward kinematics to simulate the transition from one handshape to another and to animate the transitions between signs [Sed01].

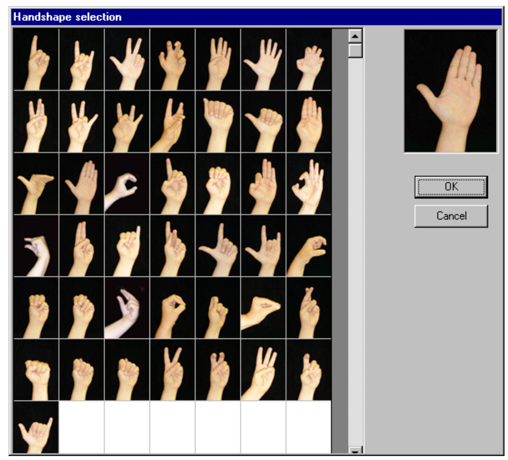

Figure 4. Handshape selection tool

Figure 5. 3D Visualization Engine

FUTURE WORK

With the current version of the transcriber, it is possible to form simple declarative sentences from hand shape, location, orientation, and movement. However, in order to form interrogative and imperative sentences in ASL, one must consider facial expressions [Bri99]. The use of facial expressions will allow us to incorporate differences in sentence type (e.g. question vs command).

Although the geometric model of the hand has been improved, user testing indicates the need for a more realistic hand model.

Usability of the interface is another important issue. Although the interface was designed to reduce the learning curve as much as possible and to make the sign construction process less complex, it is important to perform more usability tests to guarantee its efficacy and efficiency for the future users of the system. We are planning to revise and improve the affordances of some of the controls currently used on the interface.

The current 3D engine used by the transcriber comes from a commercial software vendor. For the future, we expect to construct a native graphic engine to avoid the limitations and programming complexities imposed by the current one.

REFERENCES

[Alk99] Alkoby, Karen. A Survey of ASL Tenses. *Proceedings of the 2nd Annual CTI Research Symposium.* Chicago, Illinois, November 4, 1999. http://bach.cs.depaul.edu/ctiphd/ctirs99online/alkoby.html

[Bak80] Baker-Shenk, C. and Cokely, Dennis, *American Sign Language: A Teacher's Resource Text on Grammar and Culture,* Clerc Books, Washington D.C., 1980.

[Bri96] Bridges, B. and Metzger, M. *Deaf Tend Your: Non-Manual Signals in American Sign Language.* Silver Spring, Maryland: Calliope Press, 1996.

[Cra98] 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.

[Dea00] Deafworld web site, http://dww.deafworldweb.org/int/us/

[Fur00] Furst, J., et al. "Database Design for American Sign Language." Proceedings of the ISCA 15th International Conference on Computers and Their Applications (CATA-2000). 427-430.

[Lid89] Liddell, S., Johnson, R. American Sign Language: The Phonological Base, *Sign Language Studies*, 64, 195-277, 1989

[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.

[Pri89] 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.

[Kli79] Klima, E. and Bellugi, U., *The Signs of Language*. Harvard University Press, 1979.

[Mcd00] McDonald, J., et al. "An Improved Articulated Model of the Human Hand." Proceedings of the 8th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media. pp. 306 - 313.

[Mic99] Personal Communicator CD, Michigan State University Communication Technology Laboratory, 1995,1999.

[Sed01] Sedgwick, E. et al. "Towards the Effective Animation of American Sign Language", Submitted to the 8th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media, 2001.

[Ste96] Sternberg, M., *The American Sign Language Dictionary*, Multicom, 1996. (CD ROM)

[Sto79] Stokoe, W. *A Field Guide of Sigh Language Research*. Linstok Press, Silver Springs, USA, 1979.

[Su98] Su, S.A., *"VRML-based Representations of ASL – Fingerspelling on the World-Wide Web.*

[Ten98] Tennant, R. and Brown, M. *The American Sign Language Handshape Dictionary*, Clerc Books, Washington D.C., 1980.

[Tom99] Tomuro, N., et al, "An Alternative Method for Building a Database for American Sign Language", *Technology and Persons with Disabilities Conference 2000.* California State University at Northridge, Los Angeles, CA March 20-25, 2000.

[Val93] Valli, C. and Luca, C., *Linguistics of American Sign Language*, Gallaudet University Press, 1993.