VOICE ACTIVATED DISPLAY OF AMERICAN SIGN LANGUAGE FOR AIRPORT SECURITY

Glenn Lancaster, Karen Alkoby, Jeff Campen, Roymieco Carter, Mary Jo Davidson, Dan Ethridge, Jacob Furst, Damien Hinkle, Bret Kroll, Ryan Leyesa, Barbara Loeding, John McDonald, Nedjla Ougouag, Jerry Schnepp, Lori Smallwood, Prabhakar Srinivasan, Jorge Toro, Rosalee Wolfe

School of Computer Science, Telecommunications and Information Systems
DePaul University
Chicago, IL
Email: asl@cs.depaul.edu, lancaster@cti.depaul.edu

1. Introduction

Airport security procedures pose accessibility problems for the Deaf. These procedures have become more complex and intrusive as security measures have been reassessed. The Americans with Disabilities Act (ADA)[1] and the Air Carriers Access Act (ACAA) [2] do specify accommodations that airlines must make. However, the requirements specified in these acts are minimal and do not help the Deaf in negotiating security checkpoints.

There are many audible clues and spoken directions at security checkpoints to guide hearing passengers, but the Deaf currently do not get comparable visual clues or equivalent directions. This can result in adversarial misunderstandings when, for example, a deaf person fails to stop after an alarm sounds or fails to respond to a verbal command.

At first glance it would seem that appropriately placed signs or monitors with printed English instructions would alleviate part of this problem similar to the original TTY device, developed by a deaf physicist Robert Weitbrecht. However, this assumes that the Deaf are fluent in English.

ASL is not just signed English. Although it shares some vocabulary with English, ASL is a distinct natural language with a radically different grammar[5][3][4]. With English effectively their second language, Deaf adults in the United States have an average reading level between the third and fourth grade[6].

Even the strictest interpretation of ACAA regulations does not require airport personnel to be trained in ASL. Indeed, knowledge of ASL outside the Deaf community is rare[5] in spite of its being the fourth most commonly used language in the United States[7]. Airport personnel have also affirmed that having full-time interpreters available at security areas is not feasible due to the prohibitive cost.[8]

For these reasons a synthetic English to ASL translation system would provide a feasible and more cost effective solution. However, until now no translation system has adequately addressed the problems. Previous systems use a word-for-word substitution which does not result in ASL.

Signs are typically depicted with no facial expressions. Speech recognition is speaker dependent and requires extensive voice training for each user.

To solve this problem we propose a system to facilitate Deaf access to airport security procedures by providing translations from spoken English to grammatical ASL via 3D computer graphic animations.

2. A New System

Four modules support the system: an ASL transcriber, a database of transcribed signs, a speech recognition and translation module, and a graphics display module. These modules were built using widely available programming tools, speech recognition development tools, and a commercially available graphics engine. In this paper we will only describe the translation and display modules.

2.1 Goals

The goal of the system is to provide quality translations from English into ASL. These translations must be accurate and timely, must be displayed prominently, and must be effective in enhancing Deaf access to the security screening process.

The impact of the system on security personnel should be minimal in order to ensure its effective use. Minimal instruction should be needed and the system should be essentially speaker independent. It should also respond in real time and should function well in noisy airport environments.

2.2 Speech Recognition and Translation

Natural language processing theory provides a range of techniques that typically involve tradeoffs between quality of translation versus simplicity and responsiveness.

Our initial translation system uses context-free grammars to constrain the speech recognition. Initialized with an appropriately designed grammar, the speech recognition component can still accommodate typical variations of input while also achieving the goal of robust speech recognition. One session of about 3 minutes of voice training for each particular speaker has been sufficient in our experience.

Full translation techniques of natural language processing are feasible with a constrained domain or sublanguage and can provide higher quality translations at the expense of complexity and slower performance. Although we have a candidate sublanguage, full translation techniques are not attempted in our initial system. Rather that system is best described as using "direct machine translation(MT)" [9]. A direct MT system typically is built with only one language pair in mind. It consists of several stages, each focused on one type problem. Stages might be 1) recognize words, 2) translate words, 3) rearrangements of subject, verb, object, 4) fix miscellaneous inflections, 5) generate translated output[9]. As an example of 3), the English sentence "Show me your ID and ticket." would be rearranged into the ASL signs and word order: #ID TICKET you-GIVE-me. See Figure 1. ASL signs are represented here in Baker-Cokely notation[10]. In this notation a named sign roughly corresponds to an English word and a pound sign before a gloss indicates finger spelling.

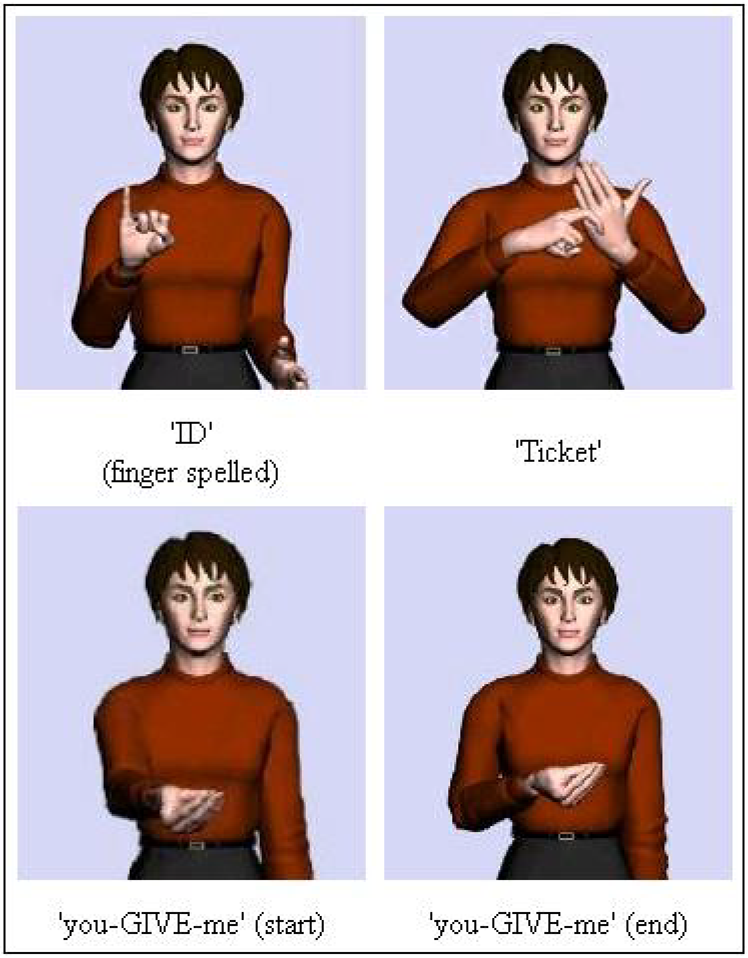

Figure 1: Frames from the ASL translation of 'Show me your ID and ticket.'

2.3 Graphic Display

The signing model used in the graphics display of ASL translations has undergone many changes from an amorphous bubble person to the present realistic model. Extensive early work concentrated on accurate modeling and control of hands in order to display the subtle differences in hand shapes required in ASL signs. Using several commercial graphics development tools, the ASL team has added significant reality and linguistic features to our system including the ability to transcribe and display all realistic hand and arm motions associated with ASL signs[11][12][13][14]. All signs are grammatically correct and have been extensively reviewed by members of the Deaf community. Additionally considerable effort has gone into incorporating facial expressions, head and shoulder movements. These "non-manual" features of signs are extremely important to the meaning associated with ASL signs.

3. Airport System Configuration

Our system will be deployed at airports to provide translations at four critical steps in the security check-in process. Display monitors will be located at these locations and security personnel will be equipped with microphones to activate the system. Each step will have an associated set of security related scenarios. See Figure 2. The first step occurs at the initial screening when the passenger's ticket and identification are requested. The security guard speaks any of the requests associated with this location. The spoken request is translated and displayed on the monitor in ASL to assist the Deaf passenger.

The second translation point occurs at the entry to the metal detector and X-ray machine. Here the security guard may typically instruct the passenger to place keys, coins or devices into a tray before passing through the detector. Step three occurs just past the metal detector where the alarm may sound. Several scenarios are associated with this step. If the alarm sounds, the monitor will display the ASL sign for "stop". The passenger may then be instructed to remove keys or other metal and walk back through. Alternatively, the passenger may be directed to a different area for a further search. The last step occurs when the security check is completed, where a final ASL translation will be displayed.

Figure 2. Airport Security Site Configuration

4. Results And Conclusions

At the conference we will present a video of our system for assisting the Deaf through airport security procedures. Researchers in computer graphics, human computer interaction, and machine learning have participated in its development in consultation with experts in speech recognition. Significantly, this system has been developed at each stage in conjunction with members of the Deaf community.

Our system is end to end operational and provides translations from English speech to grammatical ASL. We used inexpensive wired head microphones and our direct translation scheme with good results. The display system has been extensively user tested and has consistently produced signs with high rates of comprehension by native ASL signers, especially after the addition of numerous improvements and aids in the sign transcriber.

5. Future Work

Our system has yet to be tested in an airport environment. Background noise there is a concern for the speech recognition module. We intend to test this facet more thoroughly in realistic noisy environments. We also need to determine if security personnel will be willing to use these microphones.

In the future we will be working on more sophisticated translation schemes while maintaining speaker independence for speech recognition. We also plan to use advanced techniques in the graphics modules to enhance the high quality of non-manual features.

These efforts should lead to greater reliability and quality for our present system and allow us to expand its use to other airport sites such as ticket counters or lost luggage and to other application domains.

References

- [1] United States Department of Justice, Americans with Disabilities Act, Public Law 336 of the 101st Congress, enacted July 26, 1990.
- [2] United States Department of Transportation, Title 14 CFR, Part 382, 1990.
- [3] E. Klima & U. Bellugi, *The signs of language* (Harvard University Press, 1979).
- [4] C. Valli & C. Luca, *Linguistics of American Sign Language* (Gallaudet University Press, 1993).
- [5] S. Baker-Shenk & D. Cokely, *American Sign Language, a teacher's resource text on grammar and culture* (the green book) (Gallaudet University Press, 1980).
- [6] J. Holt, *Stanford Achievement Test - 8th Edition for Deaf and Hard of Hearing Students : Reading Comprehension Subgroup Results.*
- [7] National Institute on Deafness and other Communication Disorders, Website. Current as of September 2002. http://www.nidcd.nih.gov/health/pubs_hb/asl.htm.
- [8] Personal communication.
- [9] D. Jurafsky & J. H. Martin, *Speech and Language Processing, An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. (Prentice Hall, 2000)
- [10] S. Baker-Shenk & D. Cokely, *American Sign Language, A Teacher's Resource Text on Grammar and Culture.* (the green book) (Gallaudet University Press, 1980).
- [11] J. Toro, et. al., *An improved graphical environment for transcription and display of American Sign Language*. Information 4 (4) October 2001, 533-539.
- [12] E. Sedgwick, et. al., *Toward the effective animation of American Sign Language,* Proceedings of the 9th International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Medial, February 6 - 9, 2001, 375-378.
- [13] J. McDonald, 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, 2000, 306 - 313.
- [14] R. Wolfe, et. al., *An interface for transcribing American Sign Language*, SIGGRAPH 99 Sketches, August 11, 1999.