An Image Processing Technique for the Translation of ASL Finger-Spelling to Digital Audio and Text

Chance M. Glenn, Sr., Divya Mandloi, Kanthi Sarella, and Muhammed Lonon
The Laboratory for Advanced Communications Technology
Rochester Institute of Technology
Rochester, New York USA

Paper presented at
Instructional Technology and Education of the Deaf International Symposium
National Technical Institute for the Deaf
Rochester, New York June 2005

Introduction

The Sign2 Project is a focused research and development effort whose three-fold goal is:
(a) further establish and enhance the body of knowledge in physical movement/position to language translation,
(b) to conceptualize and engineer a prototype device that closes the communication gap between the deaf and the hearing, and
(c) to establish and build a statistical database from the prototype results useful to the research and development community.

PURPOSE:
To bridge the communication gap between deaf signers and hearing, non-signers

Hearing, unable to sign
Deaf, able to sign

The first phase of this project is to develop a fully image-processing approach to the translation of ASL finger-spelling. The image-processing approach was taken as opposed to other techniques such as data gloves and more exotic techniques because:
• it is a more natural approach to the problem
• it is less intrusive to the signer
• data reduction techniques are readily available in the form of image compression and feature extraction, and
• image processing techniques can be integrated with standing and developing technologies such as PDAs, smart-phones, video-phones, high-tech kiosks, etc.

Concept

PHASE I: ASL Fingerspelling
**Process**

We use the definition of the mean square error, that is,

\[ \text{MSE} = \frac{1}{L} \sum_{l=1}^{L} \sum_{w=1}^{W} \left( I(l,w) - I'(l,w) \right)^2 \]

where \( I \) is the original image and \( I' \) is the image in the statistical database.

**Results**

**Analysis:**
The frame-to-frame error was used to determine transitions between letter formation.

**Future Work**

**Short-term**
- Expand the statistical data base
- Increase reliability of letter extraction/determination process
- Improve algorithm for more natural environments

**Long-term**
- Extend this procedure for full signing
- Develop a deeper working relationship with NTID and deaf research community
- Broaden the test subject base for analysis and development
- Develop embodiments that will aid in education and general communication.

**Conclusions**

- We have developed a procedure and a system to distinguish ASL fingerspelling from a purely image processing approach
- We have successfully used the Sign2 process to resolve several words from varying subjects with a high degree of reliability
- We are developing a process for extending Sign2 to full fledged ASL
- Our goal is to provide technology to help bridge the communication gap between the deaf and the hearing.
- See demonstration at the poster session (3:00 PM Location: LBJ [060] 2nd FLOOR STREET )

**Acknowledgements**

- ECTET Department (TET Master’s Students, ET undergraduate)
- CASCI (support, facilities)
- Northstar Center (student funding)
- NTID (collaboration)
References


Paper presented at
Instructional Technology and Education of the Deaf International Symposium
National Technical Institute for the Deaf
Rochester, New York
June 2005
M1D 3