

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

ECTET R·I·T M1D CASCI

Introduction

The Sign2 Project is a focused research and development effort whose three-fold goal is to:

- further establish and enhance the body of knowledge in physical movement/position to language translation,
- to conceptualize and engineer a prototype device that closes the communication gap between the deaf and the hearing, and
- to establish and build a statistical database from the prototype results useful to the research and development community.

ECTET R·I·T NTID 2005 Technology Symposium CASCI

Introduction

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

ECTET R·I·T NTID 2005 Technology Symposium CASCI

Introduction

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.

ECTET R·I·T NTID 2005 Technology Symposium CASCI

Concept

Functional Block Diagram

ECTET R·I·T NTID 2005 Technology Symposium CASCI

Process

PHASE I: ASL Fingerspelling

ECTET R·I·T NTID 2005 Technology Symposium CASCI

Process

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

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

where I is the original image and I' is the images in the statistical database,

Adaptive Statistical Database

	S_1	S_2	S_3	...	S_N
A	I_{A1}	I_{A2}	I_{A3}	...	I_{AN}
B	I_{B1}				
C	I_{C1}				
.					
Z	I_{Z1}				I_{ZN}

Error Matrix

	S_1	S_2	S_3	...	S_N
A	e_{A1}	e_{A2}	e_{A3}	...	e_{AN}
B	e_{B1}				
C	e_{C1}				
.					
Z	e_{Z1}				e_{ZN}

Error Sum

↓

This definition of error is used to compare the input subject image to the statistical database.

NTID 2005 Technology Symposium

Results

Sign2 Graphical User Interface

NTID 2005 Technology Symposium

Results

Process:

1. Extraction
2. Thresholding
3. Scaling
4. Comparison

Analysis:
The frame-to-frame error was used to determine transitions between letter formulation

NTID 2005 Technology Symposium

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.

NTID 2005 Technology Symposium

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)

NTID 2005 Technology Symposium

Acknowledgements

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

NTID 2005 Technology Symposium

References

1. Inventor Designs Sign Language Glove, USA Today Online, (www.usatoday.com), Tech section, Associated Press, August 4, 2003.
2. Murakami, Kouichi, and Taguchi, Hitomi. Gesture recognition using recurrent neural networks. *Journal of the ACM*, 1(1):237-242, January 1991.
3. Bobick, A. and J. Davis, 2001 "The recognition of human movement using temporal templates," *IEEE Transaction on Pattern Analysis & Machine Intelligence*, Vol 23, No. 3.
4. Glenn, C. M., Eastman, M., and Palwal, G., "A New Digital Image Compression Algorithm Based on Nonlinear Dynamical Systems", IADAT International Conference on Multimedia, Image Processing and Computer Vision, Conference Proceedings, March 2005.
5. Yin, Lijun, et al., "Synthesizing a realistic facial animation using energy minimization for model based coding", *Pattern Recognition*, Vol.34, No.11, Nov. 2001, Elsevier Science, pp2201-2213.
6. K. Fukunaga, K., and Koontz, W. G. L., "Application of the Karhunen-Loeve expansion to feature selection and ordering," *IEEE Trans. Comput.*, vol. C-19, no. 4, pp. 311-318, Apr. 1970.
7. ASL Fingerspelling Conversion, Where.com (www.where.com).
8. Alkoby, K., and Sedgwick, E., *Using a Computer to Fingerspell*. DeafExpo 99, San Diego, CA, November 19-22, 1999.
9. 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.
10. Bristo, M., et.al., *Access to Multimedia Technology by People with Sensory Disabilities*. National Council on Disabilities report to Congress, 1998.
11. Carter, R., et. al., *A Better Model for Animating American Sign Language*. Proceedings of the Technology and Persons with Disabilities, California State University Northridge, Los Angeles, CA, USA, March 2002.

ECTET  R-I-T NTID 2005 Technology Symposium 