Multimodal Sign Matching Game Interface for Children

Principal Investigators: Lale Akarun

Candidate Participants: Alp Kindiroglu, Marek Hruz, Pavel Campr, Zdenek Krnoul, İsmail Ari, Eray Berger, Barış Evrim Demiröz

Abstract: The aim of this project is to develop multimodal interfaces that aid in the sign language education of hearing impaired children. Sign Languages are the main communication medium for the hearing impaired, and uses hand signs and facial gestures to convey meaning. We aim to teach sign language to infants through the use of interactive game interfaces. Different sign languages such as Turkish and Czech will be considered for this project.

PROJECT OBJECTIVES

The main objective of this project is to design and implement a system that can interact with children over 5 years old through the use of sign languages. By using recognition and synthesis techniques, the system will help the user to mentally link concepts with their corresponding signs.

In our proposed scenario, we propose a game where children will match previously determined concepts to synthesized sign videos. The system will present the user a picture of the concept and show him two or three possible sign gestures. The user will then attempt to perform his/her choice.
Then the system will show the correct version of the sign, the user's performance and give a similarity score. Depending on the score, the system will suggest the user to make a correction.

Since the target user groups of this study are 5 to 9 years old children, the scenarios of the simple games will need to have to a certain degree of simplicity. Therefore a balance between the simplicity of games such as matching concepts to signs and story driven continuity will be sought.

The objectives of the project are the following:

- Designing a close to real time multimodal interface that performs and interacts with users through the use of sign languages
- Designing various modules of the system that is required to complete the given task.
    - Hand Gesture Recognition Module
    - Facial Expression Recognition Module
    - Sign Synthesis
    - Graphics Engine

BACKGROUND INFORMATION

Sign Language Recognition:
Deaf children, unlike hearing children of speaking parents often lack the access to a language at home which is necessary for developing linguistic skills. Usually the only exposure of these children to sign languages is signing at school. Linguists have identified a "critical period" for language development - a period during which a child must be exposed to and immersed in a language [1]. It is important that children are exposed to sufficient language examples during this period to aid in the development of lifelong language skills. Although originally thought to exist only for spoken languages, research has shown that this critical period also applies to sign language education[2].

An example of such software is Con-SIGN-tration [3]. In Con-SIGN-tration, children play a memory game which involves matching cards of sign language snapshots to cards with English words. However, except the study in [1] which makes use of color gloves, no games currently on the market allow children to communicate with the computer via their native language, to the best of our knowledge. Currently, games that prompt children to mimic signs have no measure of evaluation. However, this lack of repetition with feedback prevents children from benefiting fully from the software. In order to help the child to improve the clarity and correctness of their signs, we believe such reinforcement is of crucial importance.

Sign Language Recognition:
Vision based fingerspelling recognition task aims to understand concepts that are represented thorugh the use of sign language hand, arm and face gestures in image sequences. Automatic sign language recognition is still an active and challenging research topic. It is composed of several subtasks such as segmentation of hands, extraction of features from segmented hand images and classification and temporal modeling of hand features. Through the classification of features extracted from these images, sign gesture recognition can be achieved. Since an ideal method of segmenting skin color objects from images with complex backgrounds has not yet been proposed, recent studies hand gesture recognition make use of different methodologies.
Vision based fingerspelling uses two different approaches to represent segmented hands. One approach is the usage of appearance based descriptive features to model human hands. Common appearance based features used in hand representation tasks include Histograms of Oriented Gradients, Local Binary Patterns, Elliptic Fourier Descriptors or SURF descriptors. Another

approach to hand representation is the usage of generative 3D models to represent hands.

Sign Language synthesis:

Sign language synthesis can be used in two forms. The first is real-time generated avatar animation shown on computer screen that provides real-time feedback. The second form is pre-generated short movie clips inserted into graphical user interfaces.

The avatar animation module can be divided to two models: 3D animation model and a trajectory generator. The animation model of the upper part of human body currently involves 38 joints and body segments. Each segment is represented as one textured triangular surface. In total, 16 segments are used for fingers and the palm, one for the arm and one for the forearm. The thorax and the stomach are represented together by one segment. The talking head is composed from seven segments. The relevant body segments are connected by the avatar skeleton. Rotations for shoulder, elbow, and wrist joints are commutated by inverse kinematics in accordance with 3D positions of wrist joint in the space. Avatar's face, lips and tongue are rendered by the talking head system morphing the relevant triangular surfaces.

Detailed Technical Description

Technical Description

The project has the following work packages:

WP1. Design of the overall system

In this work package the design of the overall system will be implemented. The scenarios and scope of the project will be evaluated based on the capabilities of current modules.

WP2 Graphics Engine

Graphics Engine which will be used to integrate the concept design and support for multiple modalities will be implemented.

WP3. Sign Language recognition

Sign Language recognition will be implemented for a given subset of considered sign languages. Additional data for missing words may be collected in the workshop.

WP4. Sign Language Synthesis

Sign Language synthesis will be implemented.

WP5. System Integration and Module testing

The modules implemented in WP2-WP5 will be tested and integrated in the system designed in WP1.

WP6. Usability Testing and System Demo

Usability tests for different scenarios will be performed on the integrated system.

Resources Needed

- The training databases for the recognition tasks should be ready before the project. Additional data will be collected for adaptation and test purposes.
- Prototypes or frameworks for each module should be ready before the starting date of the project. Since the project duration is short, this is necessary for successful completion of the project.
- A high fps, high resolution camera to capture sign language recognition is required.
- A dedicated computer for the demo application is required.
- Staff with enough expertise is required to implement each of the tasks mentioned in the detailed technical description.

- A modern IDE for C/C++ language will be used for development.

PROJECT MANAGEMENT

One of the co-leaders for each week will be present during the workshop.
Each participant will have a clear task that is parallel with their expertise.
Required camera hardware will be provided by the leaders.

WORK PLAN AND IMPLEMENTATION SCHEDULE

	Week 1	**Week 2**	**Week 3**	**Week 4**
WP1. Design of the overall system				
WP2. Graphics Engine				
WP3. Sign Language recognition				
WP4. Sign Language synthesis				
WP5. System Integration and Module testing				
WP6.Usability Testing and System Demo				
Documentation				

BENEFITS OF THE RESEARCH

The deliverables of the project will be the following:

D1: Sign Language recognition module

D2: Sign Language synthesis module

D3: Facial Expression Recognition module

D4: Graphics Engine

D5: Game Demo

D6: Final Project Report and presentation

PROFILE OF TEAM LEADER

Lale Akarun is a professor of Computer Engineering in Bogazici University. Her research interests are face recognition and HCI. She has been a member of the FP6 projects Biosecure and SIMILAR, COST 2101: Biometrics for identity documents and smart cards, and FP7 FIRESENSE. She currectlycurrently has a joint project with Karlsruhe University on use of gestures in emergency management environments, and with University of Saint Petersburg on Info Kiosk for the Handicapped. She has actively participated in eNTERFACE workshops, leading projects in eNTERFACE06 and eNTERFACE07, and organizing eNTERFACE07.

Selected Papers:

- Pinar Santemiz, Oya Aran, Murat Saraclar and Lale Akarun , Automatic Sign Segmentation from Continuous Signing via Multiple Sequence Alignment, Proc. IEEE Int. Workshop on Human-Computer Interaction, Oct. 4, 2009, Kyoto, Japan.
- Oya Aran, Lale Akarun, "A Multi-class Classification Strategy for Fisher Scores: Application to Signer Independent Sign Language Recognition, Pattern Recognition, accepted for publication.
- Cem Keskin, Lale Akarun, " Input-output HMM based 3D hand gesture recognition and spotting for generic applications", Pattern Recognition Letters, vol. 30, no. 12, pp. 1086-1095, September 2009.
- Oya Aran, M.S. Thomas Burger, Alice Caplier, Lale Akarun, "A Belief-Based Sequential Fusion Approach for Fusing Manual and Non-Manual Signs", Pattern Recognition, vol.42 no.5, pp. 812-822, May 2009.
- Oya Aran, Ismail Ari, Alexandre Benoit, Pavel Campr, Ana Huerta Carrillo, Franois-Xavier Fanard, Lale Akarun, Alice Caplier, Michele Rombaut, and Bulent Sankur, "Signtutor: An Interactive System for Sign Language Tutoring". IEEE Multimedia, Volume: 16 Issue: 1 Pages: 81-93, Jan-March 2009.
- Oya Aran, Ismail Ari, Pavel Campr, Erinc Dikici, Marek Hruz, Siddika Parlak, Lale Akarun & Murat Saraclar, Speech and Sliding Text Aided Sign Retrieval from Hearing Impaired Sign News Videos , Journal on Multimodal User Interfaces, vol. 2, n. 1, Springer, 2008.
- Arman Savran, Nese Alyuz, Hamdi Dibeklioğlu, Oya Celiktutan, Berk Gokberk, Bulent Sankur, Lale Akarun: "Bosphorus Database for 3D Face Analysis", The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008), Roskilde, Denmark, 7-9 May 2008.
- Alice Caplier, Sébastien Stillittano, Oya Aran, Lale Akarun, Gérard Bailly, Denis Beautemps, Nouredine Aboutabit & Thomas Burger, Image and video for hearing impaired people, EURASIP Journal on Image and Video Processing, Special Issue on Image and Video Processing for Disability, 2007.

Former eNTERFACE projects:

- Aran, O., Ari, I., Benoit, A., Carrillo, A.H., Fanard, F., Campr, P., Akarun, L., Caplier, A., Rombaut, M. & Sankur, B, "SignTutor: An Interactive Sign Language Tutoring Tool", Proceedings of eNTERFACE 2006, The Summer Workshop on Multimodal Interfaces, Dubrovnik, Croatia, 2006.
- Savvas Argyropoulos, Konstantinos Moustakas, Alexey A. Karpov, Oya Aran, Dimitrios Tzovaras, Thanos Tsakiris, Giovanna Varni, Byungjun Kwon, "A multimodal framework for the communication of the disabled", Proceedings of eNTERFACE 2007, The Summer Workshop on Multimodal Interfaces, Istanbul, Turkey, 2007.
- Ferda Ofli, Cristian Canton-Ferrer, Yasemin Demir, Koray Balcı, Joelle Tilmanne, Elif Bozkurt, Idil Kızoglu, Yucel Yemez, Engin Erzin, A. Murat Tekalp, Lale Akarun, A. Tanju Erdem, "Audio-driven human body motion analysis and synthesis", Proceedings of eNTERFACE 2007, The Summer Workshop on Multimodal Interfaces, Istanbul, Turkey, 2007.

- Arman Savran, Oya Celiktutan, Aydın Akyol, Jana Trojanova, Hamdi Dibeklioglu, Semih Esenlik, Nesli Bozkurt, Cem Demirkır, Erdem Akagunduz, Kerem Calıskan, Nese Alyuz, Bulent Sankur, Ilkay Ulusoy, Lale Akarun, Tevfik Metin Sezgin, "3D face recognition performance under adversarial conditions", Proceedings of eNTERFACE 2007, The Summer Workshop on Multimodal Interfaces, Istanbul, Turkey, 2007.
- Pavel Campr, Erinç Dikici, Marek Hruz, Alp Kindiroglu, Zdenek Krnoul, Alexander Ronzhin, Hasim Sak, Daniel Schorno, Lale Akarun, Oya Aran, Alexey Karpov, Murat Saraclar, Milos Zelezny, "Automatic Fingersign to Speech Translator", Proceedings of eNTERFACE 2010, The Summer Workshop on Multimodal Interfaces, 2010.

STAFF PROPOSED BY THE LEADER

The actual staff will be determined later however the following staff can be provided by the leaders:
Two PhD students from Bogazici University, working on Sign Language recognition
One PhD student from Bogazici University, working on Graphics Engine
Two MS/PhD students from University of West Bohemia working on sign synthesis and recognition

OTHER RESEARCHERS NEEDED

MS or PhD students with good C/C++ programming knowledge. The students will work on concept design, graphics engine and multimodal system integration.

REFERENCES

[1] H. Brashear, V. Henderson, et al. American sign language recognition in game development for deaf children. Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility 2006

[2] R. I. Mayberry and E. B. Eichen. The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. Journal of Memory and Language, 30:486–498, 1991.

[3] IDRT. Con-sign-tration. Product Information on the World Wide Web, Institute for Disabilities Research and Training Inc.

[4] Zafrulla, Z., Brashear, H., Hamilton, H., and Starner, T. (2010).A novel approach to American Sign Language (ASL) Phrase Verification using Reversed Signing. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition.

[5] Juan Pablo Hourcade. 2008. Interaction Design and Children. Found. Trends Hum.-Comput. Interact. 1, 4 (April 2008)

[6] Helene Brashear, Valerie Henderson, Kwang-Hyun Park, Harley Hamilton, Seungyon Lee, and Thad Starner. 2006. American sign language recognition in game development for deaf children. In Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06). ACM, New York, NY, USA, 79-86

[7] Oya Aran, Ismail Ari, Alexandre Benoit, Pavel Campr, Ana Huerta Carrillo, Franois-Xavier Fanard, Lale Akarun, Alice Caplier, Michele Rombaut, and Bulent Sankur, “Signtutor: An Interactive System for Sign Language Tutoring”. IEEE Multimedia, Volume: 16 Issue: 1 Pages: 81-93, Jan-March 2009.

[8] Konstantinos Moustakas, Laila Dybkjaer, Oya Aran, Dimitrios Tzovaras and Niels Ole Bernsen, Using Modality Replacement to Facilitate Communication Between Blind and Hearing Impaired People, IEEE Multimedia, accepted, 2010

[9] S. Argyropoulos, K. Moustakas, A. Karpov, O. Aran, D. Tzovaras, T. Tsakiris, G. Varni & B. Kwon, A Multimodal Framework for the Communication of the Disabled, Journal on Multimodal User Interfaces, vol.2, pp.105-116, Springer, 2008.

[10] A.A.Kindiroglu, H.Yalcin, O.Aran, M.Hruz, P.Campr, L.Akarun, A.Karpov, “A Multi-Lingual Fingerspelling Recognition for Handicapped Kiosk”, PRIA2010, Proceedings of Pattern Recognition

and Image Analysis: New Information Technologies, St. Petersburg, December 2010

[11] Pavel Campr, Erinç Dikici, Marek Hruz, Alp Kindiroglu, Zdenek Krnoul, Alexander Ronzhin, Hasim Sak, Daniel Schorno, Lale Akarun, Oya Aran, Alexey Karpov, Murat Saraclar, Milos Zelezny, “Automatic Fingersign to Speech Translator”, Proceedings of eNTERFACE 2010, The Summer Workshop on Multimodal Interfaces, 2010

[12] CopyCat: http://www.cats.gatech.edu/research