A Learning Environment for Sign Language

Faculty of EEMCS
Delft University of Technology, The Netherlands
http://www-ict.ewi.tudelft.nl/gesture

A. Vanhoutte, I. Kamp, J. Arendsen, A.J. van Doorn, H. de Ridder
Faculty of Industrial Design

E. Wenners, M. Elzenaar, G. Spaai
The Dutch Foundation for the Deaf and Hard of Hearing Child
Amsterdam, The Netherlands
http://www.nsdsk.nl

C. Fortgens, M. Bruins
Royal Auris Group
Gouda, The Netherlands
http://auris.org

Abstract

We have developed a prototype for a learning environment for deaf and hard of hearing children. This demonstration consists of hands-on experience with the prototype. In total, there are three exercises: 1) an introduction of all pictures and corresponding signs, 2) multiple choice sign-to-picture and 3) performing the sign that corresponds to the picture shown on the screen. The live recognition from a wide-angle stereo camera provides immediate feedback for the third exercise where the sign must be performed.

Figure 1. Interaction with the learning application by touch screen.

1. Introduction

From 2003 to 2007, the Dutch Foundation of the Deaf and Hard of Hearing Child (NSDSK) and the TU Delft faculties Industrial Design (IO) and Electrical Engineering, Mathematics and Computer Science (EEMCS) have collaborated to create a prototype for an Electronic Learning environment (ELo) for children (age 3-5) to practice sign language vocabulary. This project has been funded partly by the VSB foundation. The goals of the project were to find out how current technology can be used and/or improved upon to provide useful feedback on sign production and to test the learning effects of an electronic sign learning environment on the target group. Important results have been published in [1] and [2].

Sign languages are natural languages that have evolved in communities of deaf people and their families, like Sign Language of the Netherlands (SLN, or in Dutch: Nederlandse Gebarentaal, or NGT). When a deaf child has hearing parents who use little or no sign language, the child gets little language input for the language acquisition process. The ELo aims to provide such input and practice opportunities, in support of the input and education the child gets from teachers, parents and peers.

ELo is not an academic experiment, but a real-world application that can teach us what is required from gesture recognition. One such requirement is a quick response, a feeling of real-time interaction, which has been a focus of our efforts. ELo consists of three exercises: 1) an introduction of all pictures and corresponding signs, 2) multiple choice sign-to-picture and 3) performing the sign that corresponds to the picture shown on the screen.
2. ELo Setup

The design and development of the setup involved many important aspects of usability and functionality. The application, the algorithm for sign language recognition and the enclosure - recognizable as a castle -, are all essential parts of the integral learning experience.

The live sign language recognition from a wide-angle stereo camera provides immediate feedback for the third exercise where the sign must be performed. The algorithm was developed by the Information and Communication Theory Group at EEMCS [1]. It performs robust 3D hand tracking (even in front of the face) by combining skin color detection, template tracking and stereo disparity. Sign classification is done by combining Dynamic Time Warping, feature selection and a statistical discriminant. The skin detection algorithm has been designed to work with various lighting conditions and backgrounds. However, practical experience has learned us that both illumination and background clutter are much more extreme in schools than in the average office or laboratory.

Therefore, we have decided to use an enclosure that provides optimal conditions for the algorithm. Surprisingly, the enclosure also provided some unexpected benefits. The castle enclosure, shown in figure 2, not only guarantees suitable lighting conditions and background for the recognition algorithm, but also contributes to multiple aspects of usability. Among these are to provide a comfortable and pleasurable environment (by the castle theme and the warm lighting) and to reduce distractions from family or classmates.

In many (class) rooms, sunlight cannot be sufficiently blocked to prevent it from influencing skin color detection. The textile of the cover of the enclosure is impenetrable by light. By completely blocking out any other light than the illumination of the setup itself, the need for re-initialization is eliminated and continuation of operation is greatly improved. The white walls and table ensure that scattered light has approximately the same color from any direction and the textile surface of the table prevents mirror reflections of the hands.

The touch screen makes it possible to interact with the application in an intuitive manner, without the need for any experience with computer interface devices such as keyboard or mouse. However, the angle of the touch screen is an important trade-off. If the angle is too low, the screen is not well visible due to reflections and reduced contrast, and the top of the screen is hard to reach for small children. If the angle is too high, it blocks the view of the cameras onto the hands.

The stereo camera has to look over the screen, while still getting a near-frontal view of the person. As it also needs to have a full view of the person (even when making large hand gestures), the viewing angle of the camera had to be large (approximately 100 degrees). This compromises resolution and stereo matching due to the large image distortion. It also means that nobody else can even be near the child without entering the camera view and confusing the color-based hand tracking.

The illumination consists of four fluorescent saving light bulbs of 11 Watt each. These bulbs do not get as hot as regular light bulbs. This is not only important for safety, but also to prevent high temperatures inside the enclosure. A high temperature is not only uncomfortable, but also results in the natural reaction to roll up sleeves, while the current hand-tracking algorithm can only deal with long sleeves. The lights are directed downwards to minimize the brightness to the user and to create soft light that is both pleasant and suitable for reliable skin-color detection. Placing the lights above the user was not an option as there they would be reflected in the screen, reducing visibility of the application.

3. Live Demonstration

The ELo setup is demonstrated for hands-on experience with the application and/or the real-time gesture recognition. Please note that the recognition system is not designed to work with short sleeves or skin-colored clothes (pink, orange, brown or red). However, we encourage to challenge the system with difficult situations during this demonstration.

References
