

## **MACHINE LEARNING AND ITS APPLICATION TO SPEECH IMPEDIMENT THERAPY**

Duration of the project: February 2002 – February 2004

Home page of the project: <<http://oasis.inf.u-szeged.hu/speechmaster>>

### **Co-ordinator:**

#### **University of Szeged, Department of Informatics, Szeged**

*Home page:* <<http://www.inf.u-szeged.hu>>

*Address:* H-6720, Szeged, Arpád tér 2. *Tel.:* (+36)-62-544-127

*Project/team leader:* András Kocsor, <[kocsor@inf.u-szeged.hu](mailto:kocsor@inf.u-szeged.hu)>

### **Consortium member(s):**

#### **University of Szeged, Training Teaching School, Szeged**

*Home page:* <<http://www.inf.u-szeged.hu>>

*Team leader:* János Bácsi, <[bacsi@jgytf.u-szeged.hu](mailto:bacsi@jgytf.u-szeged.hu)>

#### **Kindergarten, Primary school and Boarding school, Kaposvár**

*Home page:* <<http://www.deafkvar.hu>>

*Team leader:* Jenő Mihalovics, <[titkar@deafkvar.hu](mailto:titkar@deafkvar.hu)>

Keywords: artificial intelligence, machine learning, real-time phoneme recognition, teaching of reading, speech impediment therapy

## **1. Introduction**

The project consists of two main parts: a research part and an application part. The research part is devoted to investigating modern machine learning techniques which can be used in the development of numerous info-communication systems. The machine learning algorithms developed have been made available as open-source software and we hope that they will stimulate the development of systems that requires these kind of modules. One such example is speech recognition, which is the heart of the “SpeechMaster” program, developed in the second part of our project.

The “SpeechMaster” package (Fig. 1.) seeks to apply speech recognition technology to speech therapy (Fig. 2.) and the teaching of reading (Fig. 3). The role of speech recognition is to provide visual phonetic feedback. In the first case it is intended to replace the missing auditive feedback of the hearing impaired, while in the latter case it is to reinforce the ‘correct’ association between the phoneme-grapheme pairs. With the aid of a computer children can practice without the need for the continuous presence of the teacher. This is very important because the therapy of the hearing impaired requires a long and tedious fixation phase. Furthermore, experience shows that most children prefer computer exercises to conventional drills.

Both applications – teaching reading and speech impediment therapy – require a real-time response from the system in the form of an easily comprehensible visual feedback. With the simplest display setting visual feedback is provided by means of flickering letters, their identity and brightness being adjusted to the speech recognizer's output. Figures 2-3 show the user interface of “SpeechMaster” in the teaching reading and speech therapy applications,

respectively. As one can see, in the first case the flickering letter is positioned over a traditional picture for associating the word and word sound, while in the



**1. The "SpeechMaster" program**

second case it is combined with a web camera image to help the impaired student learn the proper articulator positions.

## **2. Machine Learning Algorithms**

Our software package containing the following machine learning algorithms was implemented in the first part of the project. These packages are freely available for download from the net.

- Artificial Neural Network,
- Gaussian Mixture Model,
- Support Vector Machine,
- Projection Pursuit Learner.

## **3. Feature Transformation Algorithms**

Before executing a learning algorithm, additional feature space transformations may be applied on the data. The role of these methods is really twofold. First, they may improve classification performance, and second they may also reduce the dimensionality of the data. This is because these techniques emphasize more important features and suppress or even eliminate less desirable ones. Such linear transformations have been studied for quite sometime. A relatively new idea currently being developed seeks to create their non-linear counterparts using the so-called kernel method. The most important contribution from our team to this topic is the development of the Springy Discriminant Analysis. In addition, our software package contains the following transformation algorithms.

Linear transformations:

- Principal Component Analysis,
- Independent Component Analysis,
- Linear Discriminant Analysis,
- Springy Discriminant Analysis

Non-linear algorithms:

- Kernel Principal Component Analysis,
- Kernel Independent Component Analysis,
- Kernel Linear Discriminant Analysis,
- Kernel Springy Discriminant Analysis.

#### 4. Text and Speech Corpora

All the machine learning algorithms listed above require training data as input. In the speech recognition application it means that a large amount of speech recordings must be collected and processed. We also carefully designed the contents of this speech corpus before making the recordings. To this end we studied the vocabulary of 13 reading textbooks used in the reading teaching of children. For the speech training of the hearing impaired, the vocabulary used in the speech drills of the deaf was collected and analyzed. This study resulted in a frequency dictionary for both the reading textbooks and the hearing impaired exercises. These lists influenced the selection of the words that were recorded for the speech corpus. Altogether 700 recordings were made which were phonetically segmented and labeled and formed the proper input for the machine learning algorithms.

#### 5. The “SpeechMaster” program

The “SpeechMaster” program was designed after listening to the advice from teachers who had a lot of experience of reading teaching and/or the speech training of the hearing impaired. It contains exercises for each main step of the methodologies traditionally employed in these areas. The user interface of the program was designed so that it would be user-friendly for children with interesting, playful graphics. Although the reading teaching part also has the usual letter-to-phoneme and word-to-letter association drills using images, the core of the software is the real-time phoneme recognition component. Here the pronounced phones immediately appear on the screen in the form of flickering letters, their identity and brightness being adjusted to the speech recognizer’s output. The speech recognizer itself of course makes use of the machine learning and transformation algorithms developed in the first part of the project, and is constantly being improved.

#### 6. Methodical Descriptions

The “SpeechMaster” program will be tested in several elementary schools all around the country, and in several schools for the hearing impaired. On the basis of these tests, we will assemble the best methodology of how to use the software package in reading teaching and in improving the speech of the hearing impaired. These methodologies and, of course, the software itself will be freely available so that all the teachers will have the opportunity to try it out for themselves.



2. Speech impediment therapy



3. Teaching of reading