ERCIM "Alain Bensoussan"
Fellowship Scientific Report

Fellow: Dimitrios Ververidis
Visited location : Fraunhofer (FHG) IAIS Institute, Bonn, Germany
Duration of Visit: 1st July 2010 -> 31 March 2011

I - Scientific activity

In FHG, my research was focused on two topics, namely on the vocal tract visualization of the German phonemes, and on a platform that generates recognition systems.

A package to visualize the vocal tract was constructed. Vocal tract is manipulated by a vector of control points according to S. Maeda model. Animation can be achieved by providing a sequence of vectors over time, e.g. from Matlab. A pre-estimated codebook of certain phonemes allows for a sequence of phonemes animation. The vocal tract is manipulated with control points, denoted as articulatory model parameters, that may vary over time to produce an animation. The articulartory model parameters used, were proposed in by S. Maeda in 1982, and 1990 where the number of control points was minimized by using factor analysis on vocal tract x-ray animations, to particularly reducing the control points of the tongue. Code in Fortran was provided to solve the differential equations. The code was translated to C++ and Matlab by S. Ghosh at 2000. Later, the model was augmented to provide control for the vocal chords tension and to synthesize turbulent sounds by Huckvale at 2005. In our work, the mannequin developed has increased detail in teeth, lips, epiglottis, lungs, and stomach. The breathing is also simulated by an additional parameter. We consider that the provided package will enhance the vocal tract drawings to appear in manuscripts and help to understand the origins of speech production. The manuscript is submitted [1].

The second part of my work was to develop a platform that allows for fast implementation of recognizers that process sensor data. It consists of a Mother system that runs on the designer's side and a Kid system that is on the user's side. The Mother system is written in Matlab and its task is to design a recognition system employing the ground truth. The output of the Mother is a cartridge that contains all the necessary information that describes the recognizer. The Kid system is a small-scale replica of the mother system written in Python whose main task is to run the cartridge. Thus, an arbitrary number of recognition systems can be produced by changing the cartridge. Experiments for several problems such as voice/silence/music recognition and cars/silence recognition are presented. The platform is at evaluation stage [2].

II- Publications

1. D. Ververidis, Daniel Schroeder, Joachim Koehler, "The vocal tract Latex package," submitted in *Elsevier Speech Communication,* March 2011.

2. D. Ververidis, Daniel Schroeder, Joachim Koehler, "An engineering approach to design systems for pattern recognition," to be submitted in *Elsevier Pattern recognition*.