

**Speech Recognition using a Xilinx FPGA and an ARM9 development board**

In general, Speech recognition refers to the process of translating spoken words or phrases into a form that can be understood, which for an electronic system usually means using mathematical models and methods to process and then decode the sound signal. Speech recognition technology is commonly found in modern consumer electronics, and most people who own a computer or an Apple iPhone are aware of the capabilities. Most people are also very aware of the limitations of such systems. Most require extensive training to be effective, and even then often don't work well in noisy environments or when the speech is hurried or otherwise different from the training samples.

If speech recognition was broken down into two sections, they would be 'pre-processing' and 'decoding'. The decoding stage usually decides how fast the speech recognition engine is, as it usually involves fairly complex statistical calculations. The speed of speech recognition can be improved either by improving the decoding methodology, or by increasing the speed at which the current calculations are performed.

A series of electronic development boards (The 'Micro Arcana') is currently being designed by the University of Southampton in order to provide undergraduates with versatile and powerful prototyping platforms. However, only a few sample projects exist for these boards, and a greater number of examples would be useful to demonstrate their abilities. This series includes an FPGA board ('La Papessa'), and an ARM9 processor based mini-computer capable of running Linux.

As an FPGA is capable of performing mathematical operations very fast, using one may be a way of increasing the speed of the decoding stages. In addition, Hidden Markov Models and Viterbi decoding (popular ways of decoding speech) benefit from a parallel architecture, such as that found on an FPGA. The 'La Papessa' board will be used to perform this task. As the FPGA on it is relatively small, part of the pre-processing will be performed by the ARM board, which already has microphone-in circuitry.

In order to test the capabilities of the system designed, a set of speech samples will be prepared, and then used to test the FPGA based system against a software package. This software package will be run on the ARM board so that it runs at the same clock speed as the engine that will be created.

If the FPGA based system proves to be faster than a pure software implementation, it may be extended in several directions. Initially, it will be built to recognise a limited vocabulary of words, to test the effectiveness of the method. It may then be improved to perform more complex analysis on speech, such as language or grammar based recognition. In addition, an FPGA may be useful for voice recognition - the act of recognising a specific speaker, rather than the words.