

DEVELOPMENT OF A TEXT-INDEPENDENT SPEAKER
IDENTIFICATION SYSTEM

BY

ABIMBOLA ADEOLA FISUSI

B.Sc. Electronic & Electrical Engineering, OAU, Ile-Ife.

TP02/03/H/0829

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER OF
SCIENCE

IN

ELECTRONIC AND ELECTRICAL ENGINEERING

FACULTY OF TECHNOLOGY

OBAFEMI AWOLOWO UNIVERSITY

ILE-IFE, NIGERIA.

2007



ABSTRACT

Access control to confidential information and facilities is conventionally through the use of passwords, smart cards or keys that can be stolen or forgotten. In this work, a software-based access control system that identifies users by the unique features in their voices, which cannot be easily breached, was developed with a view to overcome the limitations of conventional access control methods.

The implementation of the system involved two phases - the training and testing phases. During the training phase, speech samples were collected from seven male and five female speakers by recording their voices with a microphone connected to a computer system. Speaker's unique features were extracted from the speech samples in form of Mel-Frequency Cepstrum Coefficients (MFCCs) which estimated the unique shape of the spectral envelope of each user. The extracted MFCCs features were used to build speaker models in form of codebooks for the speakers using the Vector Quantization (VQ) approach. The speaker models created were stored in a database. During the testing phase, another set of speech samples were collected from the same set of speakers of the training phase. MFCC features were extracted from each of the testing speech samples and compared with the codebooks created during the training phase. For each testing speech sample, the speaker whose codebook gave the lowest average distortion was identified as the true speaker. Codebooks of different sizes ranging from 16 to 256 were used to perform the identification task. In the same vein, the performance of the system as a text-dependent system was also evaluated using the same words for both training and testing phases.

The performance of the text-independent speaker identification system was evaluated by comparing the speakers' testing phase MFCCs with their training phase codebooks. The results showed that the identification rate of the text-independent speaker identification system increased with increase in codebook size. "The identification rate of the system was 57.14% when 16-vector codebooks were used as speaker models. It was 71.43% and 85.71% for 32-vector and 64-vector codebooks respectively. The speaker identification system was able to achieve 100% identification rate at codebook sizes of 128 and 256. The average distortions of speakers from testing speech samples were found to decrease with increase in codebook size. The 128-vector codebooks are preferred over 256-vector codebooks because the time required to perform the identification task is shorter for 128-vector codebooks although both give 100% identification rate.

In conclusion, the developed text-independent speaker identification system could distinguish between speakers correctly and provide security for confidential information and facilities better than conventional methods if the codebook size used for the identification task is large enough.