

LIST OF FIGURES

Figure 1.1 Extracting information from speech	2
Figure 2.1 Speech Processing	5
Figure 2.2 Typical speaker verification setup.....	6
Figure 2.3 Conceptual presentation of speaker identification.....	8
Figure 2.4 Schematic diagram of the closed-set speaker identification system	9
Figure 2.5 Human speech production system	11
Figure 3.1 Human vocal system	17
Figure 3.2 Voiced Speech and Unvoiced Speech	23
Figure 3.3 A speech signal waveform for sentences “zero”	25
Figure 3.4 Spectrum of an user, speaking the same sentence “zero”	25
Figure 3.5 Spectrum of another user speaking the same sentence	26
Figure 3.6 A speech spectrogram for an user speaking the sentence “zero”	27
Figure 3.7 Speech waveform (top plot) and associated spectrogram (bottom plot) of the word “down”.	27
Figure 4.1 Typical VQ based closed-set speaker identification system.....	33
Figure 4.2 Diagram of the real time identification system	34
Figure 4.3 Illustration of match score saturation.....	35
Figure 4.4 Speaker Identification	36
Figure 4.5 Components of speaker verification system	37
Figure 4.6 Two distinct phases to any speaker verification system	37
Figure 4.7 The desicin matrix fot the system.....	38
Figure 4.8 Threshold selection for minimizing errors in speaker verification. Our system needs to work in small window, thus rendering the process as a sensitive one.....	38
Figure 5.1 Speech Analysis Filter	45
Figure 5.2 Speech Synthesis Filter	45
Figure 5.3 Block diagram of Linear Predictive Cepstral Coefficient.....	46
Figure 5.4 Mel Scale plot	47
Figure 5.5 Block diagram of Mel-Frequency Cepstral	48

Figure 5.6 Speech signal varying over time (quasi- stationary).....	50
Figure 5.7 Framing the signal.....	50
Figure 5.8 Hamming window.....	51
Figure 5.9 Time Domain Signal and its Equivalent Frequency Representation.	52
Figure 5.10 Mel Spaced FilterBank.	54
Figure 5.11 Mel Spectrum	54
Figure 5.12 Highly Correlated Mel-Spectral Vectors Decorrelated into 13 MFCCs.	55
Figure 5.13 Mel Cepstrum.	56
Figure 5.14 Linear Acoustic Model of Human Speech-Production	58
Figure 5.15 A block diagram representation of the short-term real cepstrum Computation	59
Figure 5.16 The real cepstrum computed for the voiced phoneme, /ae/ in the word “pan.”.	60
Figure 5.17 The First 20 Coefficients of the Real Cepstrum for the Phoneme /ae/.	61
Figure 6.1 Conceptual diagram illustrating vector quantization codebook formation. One speaker can be discriminated from another based of the location of centroids	65
Figure 6.2 Distribution of quantization levels for non-linear 3-bit quantizer.....	66
Figure 6.3 Vector Quantization encoder.....	67
Figure 6.4 Vector Quantization decoder.....	67
Figure 6.5 Vector quantization partitioning of two- dimensional vector space; centroids marked as dots.....	69
Figure 6.6 Flow diagram of the LBG diagram.....	71
Figure 6.7 Clustering Balla of the same Colour together	72
Figure 6.8 illustrates the k- clustering method of Figure 6.9. notice how similar data is grouped together.....	75
Figure 6.9 K-means clustering	76
Figure 7.1 Speaker Recognition System Menu	78
Figure 7.2 Dialog Box for selecting a new sound file	79
Figure 7.3 File “s2,wav” is saved in the database SOUNDS.DAT with ID number 2.....	80
Figure 7.4 Selecting a sound file to play (e,g s2,wav).....	81
Figure 7.5 The waveform for sound file s2.wav	82
Figure 7.6 Entering ID of the sound file to be displayed.....	83

Figure 7.7 Displaying waveforms of all files in the database.....	84
Figure 7.8 The system recognize the selecting file s2.wav.....	85
Figure 7.9 Typical power spectrum output from the program.....	88
Figure 7.10 Typical output from the program showing the effects of windowing	88
Figure 7.11 Displaying information about the files in the database.....	91
Figure 7.12 Message box after selecting option 11.....	92
Figure 7.13 Displaying HELP with option 12	93
Figure 7.14 Identification rate as the noise level and number of centroids (k) are varied ..	104
Figure 7.15 Sound waveform with and without small added noise (variance 0.001) ..	104
Figure 7.16 Sound waveform with and without small added noise (variance 0.01) ..	105
Figure 7.17 Sound waveform with and without small added noise (variance 0.04) ..	105