CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS ii
CONTENTS iii
LIST OF TABLES viii
LIST OF FIGURES ix
CHAPTER 1: INTRODUCTION 1
CHAPTER 2: SPEAKER RECOGNITION CONCEPTS 4
2.1 Overview 4
2.1.1 Problem Statement 8
2.2 Biometrics 10
2.3 Relevant Studies 12
2.4 Summary 13
CHAPTER 3: SPEECH 14
3.1 Overview 14
3.2 Nature of Speech 14
3.3 Speech Processing 14
3.3.1 Speech Signal Acquisition 15
3.3.2 Speech Production 15
3.4 Designing Effective Speech 16
3.5 When to Use Speech 17
3.6 Challenges 18
3.6.1 Transience: What did you say? 19
3.6.2 Invisibility: What can I say? 20
3.6.3 Asymmetry 20
3.6.4 Speech Synthesis Quality 20
3.6.5 Speech Recognition Performance 21
3.6.6 Recognition: Flexibility vs. Accuracy 21
3.7 Voiced and Unvoiced Speech 22

3.8 Technical Characteristics and Analysis of the Speech Signal 22
3.8.1 Bandwidth 23
3.8.2 Oscillogram (Waveform) 23
3.8.3 Fundamental Frequency (Pitch) 24
3.8.4 Spectrum 24
3.8.5 Spectrogram 26
3.8.6 Cepstrum 28
3.9 Summary 29
CHAPTER 4: SPEAKER IDENTIFICATION SYSTEM 30
4.1 Overview 30
4.2 Speaker Recognition 30
4.3 Speaker Identification 31
4.3.1 VQ Based Speaker Identification 32
4.3.2 Real Time Speaker Identification 32
4.3.3 Speaker Pruning 34
4.4 Principles of Speaker Identification 35
4.5 Verification versus Identification 36
4.6 Steps in Speaker Recognition 38
4.6.1 Extraction Feature 39
4.6.2 Classification 39
4.6.2.1 Text Independent Recognition 40
4.6.2.2 Text Dependant Recognition 40
4.7 Summary 43
CHAPTER 5: SPEECH FEATURE EXTRACTION AND VECTOR QUANTIZATION 44
5.1 Overview 44
5.2 Speech Feature Extraction 44
5.2.1 Linear Predictive Coding (LPC) 44
5.2.2 Linear Predictive Cepstral Coefficient (LPCC) 46
5.2.3 Mel-Frequency Cepstrum Coefficients (MFCC) 47
5.2.3.1 Sampling 49

5.2.3.2 Framing and Windowing 49
5.2.3.3 Hamming Window 51
5.2.3.4 Fast Fourier Transform (FFT) 52
5.2.3.5 Mel Frequency Warping 52
5.2.3.6 Discrete Cosine Transform 55
5.2.3.7 Cepstrum 57
5.3 Cepstral Analysis 57
5.4 Summary of Feature Extraction Technniques 61
5.5 Summary 63
CHAPTER 6: FEATURE MATCHING 64
6.1 Overview 64
6.2 Speech Feature Matching 64
6.3 Quantization 65
6.4 Vector Quantization 66
6.4.1 Distortion Measure 68
6.4.2 Clustering the Training Vectors 70
6.5 K-Means Clustering 71
6.5.1 Clustering Overview 72
6.5.2 Non-Hierarchical Clustering 73
6.5.3 K-means Method 73
6.5.4 K-means Implementation 74
6.6 Summary 76
CHAPTER 7: MATLAB BASED SPEAKER RECOGNITION 77
7.1 Overview 77
7.2 The Speaker Recognition Program 77
7.2.1 Option 1: Load a New Sound File From Disk 78
7.2.2 Option 2: Play a Sound File From Disk 81
7.2.3 Option 3: Display a Sound Waveform From Disk 82
7.2.4 Option 4: Display a Sound Waveform From the Database 83
7.2.5 Option 5: Display All Sound Waveforms in the Database at the Same Time 84

7.2.6 Option 6: Speaker Recognition 84
7.2.7 Option 7: Display Sound Power Spectrum 86
7.2.8 Option 8: Display Sound With and Without Windowing 87
7.2.9 Option 9: Sound Database Information 90
7.2.10 Option 10: Display Information of a Sound File in the Database 92
7.2.11 Option 11: Delete Sound Database 92
7.2.12 Option 12: Help 93
7.2.13 Option 13: Exit 93
7.3 Steps in Speaker Recognition 94
7.3.1 Feature Extraction 94
7.3.1.1 Frame Blocking 94
7.3.1.2 Windowing 95
7.3.1.3 Fast Fourier Transform (FFT) 95
7.3.1.4 Mel Frequency Wrapping 96
7.3.1.5 Cepstrum 96
7.4 Speech Feature Matching 96
7.4.1 Vector Quantization 96
7.4.2 Distance Measure 97
7.4 Results 98
7.4.1 Modifying the Centroids (Code Book Size) 100
7.4.2 Speaker recognition in the Presence of Noise 101
7.5 Summary 106
CHAPTER 8: CONCLUSIONS 107
REFERENCES 109
APPENDIX A (The Main Program) 113
speaker.m 113
APPENDIX B 123
pspectrum.m 123
APPENDIX C 125
compwind.m 125

APPENDIX D 127
Cmatrix.m 127
APPENDIX E 129
Disteu.m 129
zero.m 129
APPENDIX F 130
mel.m 130
APPENDIX G 131
vqglbg.m 131
noise.m 131