## Contents

### 1 Introduction

1.1 Introduction ............................................................. 1
1.2 Objective of the Book .................................................. 2
1.3 Organization of the Book .............................................. 2

### 2 Background and Literature Review

2.1 Approaches for Detection of Vowel Onset Points .............. 5
  2.1.1 VOP Detection Using Excitation Source, Spectral Peaks, Modulation Spectrum, and Their Combination .... 6
2.2 Speech Processing in Mobile Environment ...................... 9
  2.2.1 Speech and Speaker Recognition Under Coding ........... 10
  2.2.2 Speech Recognition Under Background Noise ............. 11
2.3 Recognition of CV Units of Speech in Indian Languages .... 12
2.4 Time Scale Modification ............................................. 13
2.5 Summary ................................................................. 14

### 3 Vowel Onset Point Detection from Coded and Noisy Speech

3.1 Speech Databases for VOP Detection ............................ 17
  3.1.1 TIMIT Database ................................................ 18
  3.1.2 Broadcast News Database ...................................... 18
3.2 VOP Detection Method for Coded Speech ....................... 18
  3.2.1 Extraction of Glottal Closure Instants Using ZFF Method ......................................................... 20
  3.2.2 Sequence of Steps in the Proposed VOP Detection Method .......................................................... 21
  3.2.3 Choice of Frame Size .......................................... 24
  3.2.4 Choice of Frequency Band .................................... 27
3.3 Performance of the VOP Detection Method in the Presence of Speech Coding ................................................. 30
  3.3.1 VOP Detection from Continuous Speech Under Coding ... 30
  3.3.2 VOP Detection from CV Units Under Coding ............ 32
3.4 VOP Detection Method for Noisy Speech ........................................ 35
  3.4.1 Formant Extraction Using Group Delay Function ...................... 35
  3.4.2 Sequence of Steps in the Proposed VOP Detection Method for Noisy Speech ........................................ 37
3.5 Performance of the VOP Detection Method in the Presence of Background Noise .................................. 37
  3.5.1 VOP Detection from Continuous Speech Under Noise ............... 38
  3.5.2 VOP Detection from CV Units Under Noise .......................... 39
3.6 Summary ........................................................................ 41

4 Consonant–Vowel Recognition in the Presence of Coding and Background Noise ........................................ 43
  4.1 Consonant–Vowel Unit Databases ............................................. 44
  4.2 Two-Stage CV Recognition System .......................................... 45
    4.2.1 Motivations for the Proposed CV Recognition Approach ..... 45
    4.2.2 Proposed CV Recognition Approach .................................. 46
    4.2.3 Framework ................................................................ 47
    4.2.4 Performance of the CV Recognition System ....................... 47
  4.3 Impact of Accuracy in VOP Detection on CV Recognition .......... 50
  4.4 Performance of CV Recognition System Under Coding .............. 51
    4.4.1 Isolated CV Units Recognition Under Coding ..................... 51
    4.4.2 CV Units Recognition from Continuous Speech Under Coding ......................................................... 53
  4.5 Performance of CV Recognition System in the Presence of Background Noise ........................................ 54
  4.6 Application of Combined Temporal and Spectral Processing Methods for CV Units Recognition Under Background Noise .................................................. 56
    4.6.1 Combined TSP Method for Enhancement of Noisy Speech ......................... 56
    4.6.2 CV Units Recognition Under Different Background Noise Cases Using Temporal and Spectral Preprocessing Techniques .................. 62
  4.7 Summary ........................................................................ 63

5 Spotting and Recognition of Consonant–Vowel Units from Continuous Speech ........................................ 65
  5.1 Two-Stage Approach for Detection of Vowel Onset Points ........... 66
    5.1.1 Sequence of Steps in the Proposed VOP Detection Method ......................................................... 66
    5.1.2 Choice of Deviation Threshold for Determining the Uniformity in the Epoch Intervals ..................... 67
    5.1.3 Performance of the Proposed Two-Stage VOP Detection Method .................................................. 69
  5.2 Performance of Spotting and Recognition of CV Units in Continuous Speech ........................................ 70
## Contents

5.3 Spotting and Recognition of CV Units from Coded Speech ........ 73  
5.4 Spotting and Recognition of CV Units from Noisy Speech .......... 74  
5.5 Summary ........................................................................ 75  

6 Speaker Identification and Time Scale Modification Using VOPs ................................................................. 77  
6.1 Speaker Identification in the Presence of Coding Using  
Vowel Onset Points ............................................................ 78  
6.1.1 Speech Databases ....................................................... 78  
6.1.2 SI System Using AANN Models ............................... 78  
6.1.3 Effect of Speech Coding on Speaker Identification ........... 80  
6.1.4 Proposed Speaker Identification Method ...................... 82  
6.1.5 Performance of the Speaker Identification System  
Using Features Extracted from Steady Vowel Regions ............ 84  
6.2 Nonuniform Time Scale Modification Using Instants  
of Significant Excitation and Vowel Onset Points ................. 84  
6.2.1 Duration Analysis of Vowels in Fast and Slow Speech ...... 85  
6.2.2 Determination of Different Speech Segments ............... 87  
6.2.3 Proposed Method for Time Scale Modification ............. 88  
6.2.4 Evaluation of the Proposed Nonuniform Time  
Scale Modification Method ............................................. 92  
6.3 Summary ........................................................................ 93  

7 Summary and Conclusions ................................................. 97  
7.1 Summary of the Present Work ........................................ 97  
7.2 Contributions of the Present Work .................................. 100  
7.3 Directions for Future Work .......................................... 100  

Appendix A MFCC Features....................................................... 103  

Appendix B Speech Coders ...................................................... 107  
B.1 Global System for Mobile Full Rate Coder (ETSI GSM 06.10) ... 107  
B.2 GSM Enhanced Full Rate Coder (ETSI GSM 06.60) ........... 107  
B.3 Codebook Excited Linear Prediction (CELP FS-1016) .......... 107  
B.4 Mixed Excited Linear Prediction (MELP TI 2.4 kbps) ............ 108  
B.5 Degradation Measures ................................................... 109  

Appendix C Pattern Recognition Models ................................. 111  
C.1 Hidden Markov Models .................................................. 111  
C.2 Support Vector Machines .............................................. 112  
C.3 Auto-Associative Neural Network Models ....................... 113  

References ........................................................................ 115