

# Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Introduction</b>                           | <b>5</b> |
| 1.1      | Aim . . . . .                                 | 6        |
| 1.2      | Objective . . . . .                           | 6        |
| 1.3      | Overview . . . . .                            | 6        |
| 1.4      | Report Organization . . . . .                 | 6        |
| <b>2</b> | <b>Background Study</b>                       | <b>7</b> |
| 2.1      | Speech . . . . .                              | 7        |
| 2.2      | Principles of Speaker Recognition . . . . .   | 8        |
| 2.3      | Types of Speaker Recognition system . . . . . | 9        |
| 2.4      | Voice Recognition Algorithms . . . . .        | 10       |
| 2.5      | Speech Feature Extraction . . . . .           | 10       |
| 2.5.1    | MFCC . . . . .                                | 11       |
| 2.6      | Feature Mapping . . . . .                     | 13       |
| 2.6.1    | Overview . . . . .                            | 13       |
| 2.6.2    | Vector Quantization . . . . .                 | 14       |
| 2.7      | Applications of Speaker Recognition . . . . . | 16       |

|          |   |           |
|----------|---|-----------|
| 2.8      | Performance Factors for Speaker Recognition . . . . . | 16        |
| <b>3</b> | <b>Methodology</b>                                    | <b>18</b> |
| 3.1      | Data Collection . . . . .                             | 18        |
| 3.1.1    | Environment of Recording . . . . .                    | 18        |
| 3.1.2    | Device Used . . . . .                                 | 20        |
| 3.1.3    | Encoding . . . . .                                    | 20        |
| 3.1.4    | Male Female Ratio . . . . .                           | 20        |
| 3.1.5    | Age Group . . . . .                                   | 20        |
| 3.2      | Tools Used . . . . .                                  | 20        |
| 3.2.1    | Wavesurfer . . . . .                                  | 20        |
| <b>4</b> | <b>Implementation</b>                                 | <b>21</b> |
| <b>5</b> | <b>Simulation Results</b>                             | <b>23</b> |
| 5.1      | Datasets used: . . . . .                              | 23        |
| 5.2      | Experiments conducted: . . . . .                      | 23        |
| 5.2.1    | Number of MFCC coefficients . . . . .                 | 24        |
| 5.2.2    | Number of filterbanks . . . . .                       | 24        |
| 5.2.3    | Number of Centroids . . . . .                         | 26        |
| 5.2.4    | Test and Train Speech Duration . . . . .              | 27        |
| 5.2.5    | Average Execution Time . . . . .                      | 27        |
| <b>6</b> | <b>Conclusion</b>                                     | <b>29</b> |

# List of Figures

|     |  |    |
|-----|--|----|
| 2.1 | Application areas of Speech processing . . . . .                   | 7  |
| 2.2 | Speaker Identification System . . . . .                            | 8  |
| 2.3 | Speaker Verification System . . . . .                              | 9  |
| 2.4 | Voice Recognition algorithms . . . . .                             | 10 |
| 2.5 | Block Diagram of MFCC processor . . . . .                          | 11 |
| 2.6 | An example of 2D VQ . . . . .                                      | 14 |
| 3.1 | Voice Algorithm flow Chart . . . . .                               | 19 |
| 5.1 | Identification rate (in%) Vs. the No. of MFCC . . . . .            | 24 |
| 5.2 | Filterbank Versus Identification rate . . . . .                    | 25 |
| 5.3 | No. of Centroids Vs Identification rate . . . . .                  | 26 |
| 5.4 | Speech Duration(in sec) Versus Identification Rate(in %) . . . . . | 27 |
| 5.5 | SpeechDurationVsExecTime . . . . .                                 | 28 |

# List of Tables

|     |  |    |
|-----|--|----|
| 3.1 | Training Requirement . . . . .                                 | 18 |
| 5.1 | Identification rate (in%) Vs. the No. of MFCC . . . . .        | 24 |
| 5.2 | Identification rate (in%) Vs. the No. of filterbanks . . . . . | 25 |
| 5.3 | Number of Centroid Vs Identification Rate . . . . .            | 26 |
| 5.4 | Speech Duration Vs Identification Rate . . . . .               | 27 |
| 5.5 | Speech Duration Vs Execution Time . . . . .                    | 28 |