CHAPTER 4
IMPLEMENTATION AND RESULTS

4.1 Introduction

This chapter shows the implementation details of speaker recognition system project. It also shows the steps required to achieve the complete speaker recognition process. Also, it introduces the project testing using different testing environments and the results after testing. It also introduces a GUI we implement to facilitate interaction with the system and make it very comfortable.

4.2 System implementation

In this section all the implementation details are presented including the software used associated with some facilitation figures.

4.2.1 Software components

Matlab

MATLAB is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, and numeric computation. Using the MATLAB product, you can solve technical computing problems faster than with traditional programming languages, such as C, C++, and FORTRAN.

MATLAB can be used in a wide range of applications, including signal and image processing, communications, control design, test and measurement, financial modeling and analysis, and computational biology. Add-on toolboxes (collections of special-purpose MATLAB functions, available separately) extend the MATLAB environment to solve particular classes of problems in these application areas.

MATLAB has extensive facilities for displaying vectors and matrices as graphs, as well as annotating and printing these graphs. It includes high-level functions for two-dimensional and three-dimensional data visualization, image processing, animation, and presentation graphics. It also includes low-level functions that allow you to fully customize the appearance of graphics as well as to build complete graphical user interfaces on your MATLAB applications [9].
Chapter 4

IMPLEMENTATION AND RESULTS

Figure 4.1 Matlab Interface
4.2.2 Hardware components

Headset and Microphone

It is used as voice input unit to computer and also as speaker for output voice.

![Headset and microphone](image1)

**Figure 4.2** Headset along with microphone

The headset is configured as wanted using special computer software shown below

![Sound configurations interface](image2)

**Figure 4.3** Windows sound configurations interface
4.3 Results

4.3.1 GUI implementation

Our speaker recognition system is designed using three different algorithms, and the results are taken from all of them.

The sequence of sub processes and steps are shown in a GUI starting from the welcoming message and till the Identification results.

All the algorithms are achieved through two main phases, the training phase and the testing phase. Also many other options are available such as if you want to check the voice you uttered or the plot of that voice and many other options according to the used algorithm.

When a user starts our speaker recognition system, he will be faced with a welcoming message.

![Welcoming message GUI](image)

**Figure 4.4 Welcoming message GUI**

Then the user should click next to proceed, and then he will be faced with another message containing the three available algorithm that provide three different ways for achieving complete speaker recognition system.
The user should simply choose one of the three algorithms provided by the system.

![Figure 4.5 Choosing message GUI](image)

As we mentioned above, the user should choose one of those algorithms available in the above message.

4.3.1.1 MFCC Algorithm

Let’s assume that the user has chosen the first technique –MFCC- by clicking on the button MFCC, the next message will appear.

![Figure 4.6 MFCC training phase GUI](image)
As we have mentioned before, all the techniques come through two basic phases, the training phase and the testing phase.

The message above shows the training phase of the MFCC algorithm. It contains four buttons, a recording button, which enables the user to enter his voice for training purpose, a playing button, which sounds the same voice entered by the user in the recording step, it is just for the user to ensure a satisfied and complete voice entering—with no noise and side distortion-, a plot showing button, and it just shows the signal of the recorded voice with respect to time, and the last button—proceed to testing—provides the basic calculations of MFCC algorithms mentioned in the last chapter.

When the user recorded his voice and then he pushed the plot showing message, the signal of his voice appears in a separated window. The next image is an example of an uttered voice signal, certainly “Hello” word.

Figure 4.7 “Hello “ voice signal for training
When the user click on proceed to testing, immediately he will be directed to the next message which is mainly provide the testing phase.

![Figure 4.8 MFCC testing phase GUI](image)

As shown above, this message also contains four buttons, a recording voice button, and it has the same function of recording button provided in the training phase, a playing button and plot showing button, which have the same function of playing and plot showing mentioned in the training phase respectively –but this time for the testing voice-, and the result showing button which performs the same calculations provided previously for the training voice, and then it takes a decision of authentication, i.e. either permitted or rejected.

Example of voice signal plotting –for testing- certainly “Hello” word is shown below.

![Figure 4.9 “Hello” voice signal for testing](image)
If the system detected that he is the same user, a permission message will be shown.

**Figure 4.10** Acceptance Result GUI for MFCC

Otherwise, a rejection message appears.

**Figure 4.11** Rejection result GUI for MFCC
4.3.1.2 FFT Algorithm

Now let’s assume the user has chosen –from the choosing message shown above- the second algorithm, which it is FFT.

When the user push FFT button appeared in the choosing message, he will be directed immediately to the training phase message of FFT algorithm. The message is shown below.

![FFT training phase GUI](image)

**Figure 4.12 FFT training phase GUI**

In the training phase of this algorithm, the user supposed to enter his voice ten times in order to perform the FFT calculations mentioned in the previous chapter.

It is clear that from the above figure there are ten buttons dedicated for recording the training voice, another ten plot showing buttons dedicated for each training voice, a training button, which performs calculations on the training voice, and a testing button, which directs the user to the testing message. Thus the user should record his voice ten times to create the reference voice and save it in the database. The user can also show plots for each voice he enters.
After that he should push on the training button to perform the required calculations on the ten samples of his voice. Then he should push on the testing button to be directed to the testing phase message which is shown next.

![FFT testing phase GUI](image)

**Figure 4.13 FFT testing phase GUI**

It is clear that the above figure contains four buttons, recording button, which enables the user to enter his voice for testing, a plot showing button and playing button which enable the user to show the signal of his entered voice and listen to it respectively, and a decision button which when pushed performs the required calculations on the testing voice and takes the right decision—either permission or rejection—, these calculations are also shown in the previous chapter—testing phase using FFT algorithm—.
The decision results either must be a permission and the next message appears.

![Acceptance Result GUI for FFT](image1)

**Figure 4.14 Acceptance Result GUI for FFT**

Or it is a rejection and the following message will appear.

![Rejection result GUI for FFT](image2)

**Figure 4.15 Rejection result GUI for FFT**

### 4.3.1.3 VQ Algorithm

This is the third and last algorithm in our system. So let’s assume the user has chosen it to perform speaker recognition system. The procedures are exactly as the above, training then testing and the result will be permission if the speaker is the same and a rejection if another speaker tried to use the system.
Figures of the GUI used at VQ algorithm is shown next.

![VQ training phase GUI](image)

**Figure 4.16  VQ training phase GUI**

All the buttons have the same functions and usages of the early mentioned algorithm. But here we have two additional buttons, the acoustic vector button and the acoustic vector with VQ button, which show acoustic vector features -of the training voice - without and with VQ respectively. Example of acoustic features vector without VQ is shown next.

![Acoustic features vector without VQ](image)

**Figure 4.17  acoustic features vector without VQ**
An example of acoustic features vector with VQ is shown next.

![Figure 4.18 acoustic features vector with VQ](image)

**Figure 4.18 acoustic features vector with VQ**

When the user pushes proceed to testing button, immediately he will be directed to the testing phase message which is shown below.

![Figure 4.19 VQ testing phase GUI](image)

**Figure 4.19 VQ testing phase GUI**
The same buttons mentioned in previous algorithms messages—testing phase—are also used in this algorithm message—testing phase-. Thus the user should enter his voice, and optionally can show plots of his voice signal, and also he can listen to it to ensure satisfaction and completion of uttering (without noise and/or distortion).

Finally the user should push on the yellow buttons-Show Results- to see if he will be permitted or rejected. The permission and rejection messages are shown next respectively.

![Acceptance Result GUI for VQ](image1.png)

**Figure 4.20** Acceptance Result GUI for VQ

![Rejection result GUI for VQ](image2.png)

**Figure 4.21** Rejection result GUI for VQ
4.3.2 Testing Results

Now, for each algorithm we will calculate its efficiency. We will present the same user uttering the same word used for training, and then we will calculate the percentage acceptance. Then we will present the same user and different word than that used in the training phase and also the percentage correctness will be shown. Further, we will introduce different user uttering the same word and different user uttering different word than that used for training and for each the percentage acceptance will be calculated. All the above combination will be introduced in each algorithm.

4.3.2.1 MFCC

Table 4.1 MFCC testing results

<table>
<thead>
<tr>
<th>Users</th>
<th>The same user uttering the same word</th>
<th>The same user uttering different word</th>
<th>Different user uttering the same word</th>
<th>Different user uttering different word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of acceptance</td>
<td>82%</td>
<td>0%</td>
<td>16.7%</td>
<td>0%</td>
</tr>
</tbody>
</table>

4.3.2.2 FFT

Table 4.2 FFT testing results

<table>
<thead>
<tr>
<th>Users</th>
<th>The same user uttering the same word</th>
<th>The same user uttering different word</th>
<th>Different user uttering the same word</th>
<th>Different user uttering different word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance percentage</td>
<td>70.5%</td>
<td>0%</td>
<td>20%</td>
<td>0%</td>
</tr>
</tbody>
</table>
4.3.2.3 VQ

Table 4.3 VQ testing results

<table>
<thead>
<tr>
<th>Users</th>
<th>The same user uttering the same word</th>
<th>The same user uttering different word</th>
<th>Different user uttering the same word</th>
<th>Different user uttering different word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptance percentage</td>
<td>97.1%</td>
<td>86.6%</td>
<td>16.6%</td>
<td>0%</td>
</tr>
</tbody>
</table>

4.4 Discussion

The efficiency, Quality and accuracy attributes for our system differ from one algorithm to another. This was shown clearly on the above tables which contain the correctness percentage for each algorithm applying all the available combinations. Here we will discuss the efficiency of each algorithm.

4.4.1 MFCC

It is clear that from the above table of MFCC correctness percentage, MFCC is a text dependent algorithm, i.e. in order to give permission to user, he should be the same user and utter the same word used for training. This of course has advantage and disadvantage. Its advantage is that it gives the system security with two dimensions, voice dimension and password dimension. Its disadvantage that it rejects the right user if he uttered different word.

However MFCC algorithm considered as the basic algorithm in speaker recognition, but it is also faced with some accuracy limitations.

4.4.2 FFT

FFT is also text dependent algorithm. As we mentioned above this may be an advantage or disadvantage.

As we mentioned in the previous chapter, the user in the training phase should enter his
voice ten times. It is not restricted to ten, but whenever the number of times increases, recognition accuracy also increases. That is because a good amount of changes and effects that may occur to the training voice will be considered as the number increase. It is the worst algorithm in terms of accuracy.

4.4.3 VQ

It is a text independent algorithm, i.e. the authenticated user will be accepted even if he uttered different word. However it is a text independent algorithm, but the acceptance percentage when uttering different word is less than that when uttering the same word. That is because the value of distortion- illustrated in the previous chapter- increases when uttering different word.

It is the best algorithm regarding accuracy. This algorithm is an improvement for the MFCC algorithm. It uses VQ based clustering along with MFCC approach, thus it is the best between the other two algorithms.

**General issues should be considered to enhance system accuracy:**

- User should utter the voice accurately in the given time interval in order to avoid signal sides distortion.
- External noise should be avoided by placing the system in a very quiet area so as to provide the required accuracy.
- Internal noise also should be avoided by using a high quality (fine) microphone for entering the voices.