Abstract

Early detection of speech-related disorders is of major importance in order for intervention and treatment to be effective. This work presents our effort to create a web-based computerized speech pathology assessment tool that users can easily access on-line in order to get an initial assessment of possible speech-related problems and seek further help.

Our system uses technologies such as HTML5, JavaScript, AJAX and Play! Web-Application Framework to provide an interactive front-end that can capture audio and video from the user’s browser and also uses back-end technologies such as Java and MatLab to implement algorithms that analyze the captured data and perform the assessment.

Introduction

According to the National Institute of Deafness and other Communication Disorders (NIDCD), approximately 7.5 million people in the United States have a speech or voice disorder [1].

While many disorders - such as dysarthria - are quite obvious when one listens to an affected speaker, many disorders - such as cluttering - are more mild in severity and may go undiagnosed, even into adulthood.

There is a lack of user-friendly applications which would aid to help identify possible speech impediments. If there is a match, a very high-level diagram of system purpose.

Fig. 1. A very high-level diagram of system purpose.

According to the National Institute of Deafness and other Communication Disorders (NIDCD), approximately 7.5 million people in the United States have a speech or voice disorder [1].

While many disorders - such as dysarthria - are quite obvious when one listens to an affected speaker, many disorders - such as cluttering - are more mild in severity and may go undiagnosed, even into adulthood.

There is a lack of user-friendly applications which would aid to help identify possible speech defects that a user may have by using a speech sample provided by the user. We are attempting to create such an application which could be accessible via the internet.

Such an application would not provide a definitive diagnosis for a disorder. However, the application would bring the consideration to the user that he or she may have a speech disorder and should seek an official diagnosis. This application could also relieve pathologists from certifying certain types of abnormalities. More advanced classification methods can be used for that purpose.

In the early stages of the project, the main focus is getting a user interface that performs the functions of recording a user voice sample and uploading the file in order to be processed.

Our application uses Dynamic Time Warping to measure dissimilarities between normal and abnormal utterances of certain phonemes. DTW is also well-known for its ability to find an optimal alignment between two time-dependent sequences. The experimental results (Fig. 2) confirm our initial assumption that matching a normal and an abnormal utterance of the same sound would yield a higher distance than compared to matching two normal utterances.

DTW may be useful for detecting abnormality in speech, however, since the distance values for the hoarse and flutter voice qualities are very similar, it may not be suitable for detecting the specific type of abnormality. More advanced classification methods can be used for that purpose.

According to the National Institute of Deafness and other Communication Disorders (NIDCD), approximately 7.5 million people in the United States have a speech or voice disorder [1].

While many disorders - such as dysarthria - are quite obvious when one listens to an affected speaker, many disorders - such as cluttering - are more mild in severity and may go undiagnosed, even into adulthood.

There is a lack of user-friendly applications which would aid to help identify possible speech defects that a user may have by using a speech sample provided by the user. We are attempting to create such an application which could be accessible via the internet.

Such an application would not provide a definitive diagnosis for a disorder. However, the application would bring the consideration to the user that he or she may have a speech disorder and should seek an official diagnosis. This application could also relieve pathologists from certifying certain types of abnormalities. More advanced classification methods can be used for that purpose.

In the early stages of the project, the main focus is getting a user interface that performs the functions of recording a user voice sample and uploading the file in order to be processed.

Our application uses Dynamic Time Warping to measure dissimilarities between normal and abnormal utterances of certain phonemes. DTW is also well-known for its ability to find an optimal alignment between two time-dependent sequences. The experimental results (Fig. 2) confirm our initial assumption that matching a normal and an abnormal utterance of the same sound would yield a higher distance than compared to matching two normal utterances.

DTW may be useful for detecting abnormality in speech, however, since the distance values for the hoarse and flutter voice qualities are very similar, it may not be suitable for detecting the specific type of abnormality. More advanced classification methods can be used for that purpose.

Methods

In the early stages of the project, the main focus is getting a user interface that performs the functions of recording a user voice sample and uploading the file in order to be processed.

The front-end of the application involves the use of the user’s microphone. The user is asked to repeat several recordings that are presented to him or her which are selected to exemplify certain traits of common speech disorders. The user records his or her attempt at replicating the sample recordings and the user’s samples are then processed.

Using MatLab for the processing of the signals allows us to use algorithms which take advantage of the more expressive MatLab language and library. Fig. 3 shows an overview of our system.

Fig. 3. A high-level diagram of the application system.

Future Work

Firstly, an analysis of suitability of algorithms must be completed. This will involve other algorithms that are suited for classification – such as Neural Networks and Hidden Markov Models.

Also, the system will have to be tested for stability. Currently, no more than a few active connections have been tested at a given time. If the system is to hold up to real-world use, it must be able to efficiently use its resources and coordinate many instances of JVMs.

Once we have a system that can perform the task to a sufficient degree, we will then begin subject testing and analysis. This will be a major factor in the efficacy of the system and will be the major result that will determine the usefulness of the system.

References
