PART A

* + 1. Emotion recognition is the one way through which it is possible to know the emotion of a person by either facial or verbal methods. There are any methods developed in science to know the emotion of a person by computational methodologies. Each person has their own identical way of expressing their emotions. This method had number of applications in psychology. Emotion recognition can be detected using different techniques like, signal processing, computer vision(facial), machine learning and so on.
    2. The main reason for selecting speech processing is the originality in the features. Candidate can fake the expressions in the facial method of detecting the emotion but this cannot be done in speech. Image processing for emotion recognition is done by capturing a image of the candidate and analysis is done based on that image. There are chances of missing a few important characteristics required for recognising a emotion. Though noise id found in speech signal there several methods to remove the noise and continue with the speech processing.

1.2.1 Parameters for speech detection

1. Pitch: Pitch of the speaker is one of the important aspect for speech recognition. Every individual has a different pitch which various from each circumstance. If it is important is identify the speaker then pitch is the key parameter. Identifying the speaker emotion is also one application of pitch application.

Example: If a person is tired then the pitch of the person is low whereas if the person is happy then the pitch is expected to be high. The variation in the pitch helps in emotion detection.

2. Prosodic parameters: this is the duration of phonetic units. When it is necessary to identify what the speaker is talking then assuring a boundary between the words is important. Later this can also be used for selecting the required portion from the huge database.

3.Intensity : The intensity in speech also matters a lot in the speech processing because when the person intensity is low it becomes difficult to recognise any parameter in the speech. The intensity of the person can vary in the recording of the speech or based on the database available to the algorithm. This parameter is important is know the emotion of a person.

4. Duration: The database given for the algorithm must be long enough for the algorithm is recognise and train the system accordingly. The gap in between in each phonics also plays a vital role in differentiating the words

5. Environmental parameters: The Signal to noise ratio is expected to be less. This can be achieved by removing all the environmental disturbances in the speech and selecting what is exactly is important.

1.2.2 Challenges in speech detection:

1. Defining a emotion: It is very important to define a specific emotion characteristics. Before classifying the algorithm should be trained for specific emotion. Each emotion should have its own characteristics.

2. Environmental noise: Noise is one variable that needs to analyse well and there are number of techniques to remove it.

3. Selection of appropriate technique: When the technique must be selected its important to select the specific algorithm it can be for feature extraction or for classification or detection.

1.3.1 Feature extraction techniques:

1.3.2 Classification Approaches:

1. Support vector machine (SVM)( Cortes, Corinna; Vapnik, Vladimir N. (1995).): It is supervised learning for classification.

2. Hidden markov Models: It is a statistical model of assuming that the hidden states are modelled based on Markov process.

3. Gaussian mixture model: This method is statistical method of finding the sub models in the main model.

4. Neural network(Warren McCulloch and Walter Pitts(1943) ) : This the artificially created network that will help to solve problem. To solve very complex problems this is efficient.

5. K nearest Neighbor( Altman, N. S. (1992)): A dataset is classified based on the maximum vote by its neighbours in the feature space.

1.4 stance

Part B:

2.1 There are number research done in recent times for speech recognition .







2.2.2 Classification technique:

1. Support vector machine (SVM)( Cortes, Corinna; Vapnik, Vladimir N. (1995).): It is supervised learning for classification.

2. Hidden markov Models: It is a statistical model of assuming that the hidden states are modelled based on Markov process.

3. Gaussian mixture model: This method is statistical method of finding the sub models in the main model.

4. Neural network(Warren McCulloch and Walter Pitts(1943) ) : This the artificially created network that will help to solve problem. To solve very complex problems this is efficient.

5. K nearest Neighbor( Altman, N. S. (1992)): A dataset is classified based on the maximum vote by its neighbours in the feature space.

2.3.1 real time audio samples

Audio samples are created by recording the voice of batch mates in the year and a recording is done for 30 sec. A minimum of 30 sec is required because the system should have to extract enough features from the signal. Different recordings are done for training and testing set.

There are 6 sets of trainings and testing data which are recorded and converted to wav file format for MATLAB process.

2.3.2 model

Firstly audio samples are converted to wav format for further use in MATLAB. Later the signal is passed to few steps like:

1.Pre Processing

2. Extraction of features

3. Classification of the speaker.