Volume 2, Issue 6 June 2016 ISSN NO: 2454 - 7514

North Asian International Research Journal Consortium

North Asian International Research Journal Of Science, Engineering and Information Technology

Welcome to NAIRJC

ISSN NO: 2454 -7514

North Asian International Research Journal of Science, Engineering & Information Technology is a research journal, published monthly in English, Hindi. All research papers submitted to the journal will be double-blind peer reviewed referred by members of the editorial board. Readers will include investigator in Universities, Research Institutes Government and Industry with research interest in the general subjects

Editorial Board

Address: - Dr. Ashak Hussain Malik House No. 221 Gangoo, Pulwama, Jammu and Kashmir, India - 192301, Cell: 09086405302, 09906662570, Ph. No: 01933-212815,
Email: nairjc5@gmail.com, nairjc@nairjc.com, info@nairjc.com Website: www.nairjc.com

STUDY OF EMOTIONAL INTELLIGENCE TRACKER AND PREDICTOR FOR RDS

ABHINAV ANAND & ASST. PROF. MR. RAHUL KADIAN

CBS Group of Institutions, Maharshi Dayanand University, Haryana

ABSTRACT:

Automatic emotion recognition from voice and face has become a core discipline in machine learning and pattern recognition. From the machine perspective, recognizing the user's emotional state is one of the main requirements for computers to successfully interact with humans. The performance of human emotion recognition system can be improved my combining more input modalities for the interpretation. The emotions studied in the proposed work are neutral, happy, sad and angry. Prosodic and spectral features extracted from speech are used for discriminating the dialects and emotions. The spectral features of the speech are represented by Mel Frequency Cepstral Coefficients (MFCC) and prosodic features are represented by mode, pitch and energy contours. MFCC, as studied by earlier researchers, provide 75% efficiency results to get more efficient results is it required to use more features of the sound. MIRtoolbox in Mat lab is available to extract more efficient features of the sound as used in the proposed work .This simulation tool is used to prepare the complete trained dataset which explains about the features of sounds of a particular person. The way of speaking of all the persons have the slight difference. Some persons speak very lightly in the sad mood. Some speak in the similar tone as in neutral form. Trained dataset records the way of talking of the users. This trained dataset is used to study the emotions of the users that can be further used in machine learning. This system is intended to be used during human robot interaction, and it is integrated as part of the overall interaction system of the robot: the Robotics Dialog System.

Keywords: ***Emotion prosodic , Spectral features ,MFCC , MIR toolbox , RDS***

1. INTRODUCTION

The expression "emotion intelligence" does not yet show up in word references. The vast majority of the early research on knowledge concentrated on critical thinking on different things that were anything but difficult to find. It has for quite some time been acknowledged that different elements are vital for foreseeing somebody's capacity to succeed at work and in life. This feelings knowledge incorporates the ability to screen one's own specific and others' feelings and sentiments, to partition among them, and to use this information to guide one's thinking and action. Energetic understanding addresses the eager, individual, social, and survival estimations of learning. In this paper, a calculation is created for feeling acknowledgment in the discourse preparing utilizing the dataset. Machine learning calculation is produced for better results. This work is to use further in mechanical autonomy and machine learning. With the goal that machines can comprehend what to do as indicated by the feelings of the human nature. In this paper, a calculation is produced for feeling acknowledgment in the discourse preparing utilizing the dataset. Machine learning calculation is created for better results. This work is to use further in mechanical autonomy and machine learning. So machines can comprehend what to do as indicated by the feelings of the human instinct. The feelings can be concentrated on as the investigation of versification; particularly: the efficient investigation of metrical structure. It is a specific framework, hypothesis, or style of versification. These are additionally called Prosodic elements. These components are showed up when we set up sounds together in associated discourse. It is as imperative to show learners prosodic elements as fruitful correspondence depends as much on pitch, anxiety and beat as on the right elocution of sounds.

RDS: In this paper, the algorithm is produced for feeling acknowledgment in the discourse preparing utilizing the dataset. **Machine learning algorithm is developed for better results. This work is to use further in mechanical autonomy and machine learning. So machines can comprehend what to do as per the feelings of the human instinct. This system is intended to be used during human robot interaction, and it is integrated as part of the overall interaction system of the robot: the Robotics Dialog System.**

2. LITREATURE REVIEW

"Intelligence" did not show up in books before the twentieth century. "Insight" wasn't regular until after 1930.Generally the hints of spoken languages have been learned at two distinct levels: (1) phonetic segments of spoken words, e.g., vowel and consonant sounds, and (2) acoustic wave designs. A dialect can be isolated into a bit number of vital sounds, called phonemes (English has pretty nearly forty). An acoustic wave is an arrangement of changing vibration designs (by and large in air), in any case we are more habituate to "seeing" acoustic waves as their electrical simple on an oscilloscope (time presentation) or range analyzer (recurrence presentation) Likewise found in stable examination are two-dimensional examples called *spectrograms*, which show frequency versus time and speak to the sign vitality as the figure force or shading. Daniel Goleman advanced the term *'Passionate Intelligence' in 1995* in the title of his smash hit book, *Emotional Intelligence: Why It Can Matter More than IQ*. Goleman characterized enthusiastic insight as 'Understanding one's own emotions, sympathy for the sentiments of others and the regulation of feeling in a manner that upgrades living.' Not everybody concurs with Goleman's model of passionate knowledge, yet there is general assertion that passionate insight exists, that it is a variable in individual and expert achievement, and that it can be progressed.

The distribution of Goleman's book *Emotional Intelligence* in 1995 denoted the start of enthusiastic knowledge as something that was perceived by standard business scholars and authors.

Social-psychological limits help us arrange the social world by illuminating us about the individuals with whom we interface. Out of these limitations our understandings of others' musings and feelings are foremost. Hypothesis of brain **(ToM; Wellman, 1990)** concerns our gratefulness for individuals' psychological states, for example, convictions and information. Emotion Understanding (EU) alludes to our capacity to recognize obvious passionate responses, to anticipate others' enthusiastic responses, and to admire that individuals have both tangible and private passionate encounters **(Denham, 1986; Pons, Harris, & de Rosnay, 2004).** Youthful kids' initial comprehension of feelings (e.g., essential full of feeling point of view taking) and comprehension of perception (e.g., false-conviction comprehension) are at first unmistakable limits (Cutting & Dunn, 1999), with diverse relate**s (Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991).** Then again, ToM and EU are together fundamental for adult social discernment for instance a research of shrouded feelings **(Harris, Donnelly, Guz, & Pitt-Watson, 1986)** obliges a joint comprehension of subjective and full of emotions states. Moreover, ToM and EU can be both interestingly and mutually powerful in our thinking and choice making **(Pons et al., 2004; Wellman & Banerjee, 1991).**

Eric Brill, Radu Florian, John C. Henderson, Lidia Mangu proposed that best in class dialect models for discourse acknowledgment are in light of an exceptionally unrefined semantic model, to be specific molding the likelihood of a word on a little settled number of going before words. Notwithstanding numerous endeavors to join more modern datainto the models, the n-gram model remains the best in class, utilized as a part of for all intents and purposes all discourse acknowledgment frameworks. **Sameer Maskey, Julia Hirschberg in 2005** introduced aftereffects of an observational investigation of the helpfulness of distinctive sorts of elements in selecting extractive synopses of news telecasts for our Broadcast News Summarization System. Most content based synopsis frameworks depend upon lexical, syntactic, and positional data in figuring out which portions to incorporate in a rundown. They portrayed the element classes we use to foresee sentences to be separated and our system for selecting them, including lexical, auxiliary, and prosodic and talk highlights.

In the proposed work, **Mel Frequency Cepstral Coefficents (MFCCs)** highlight is likewise considered. MFCCs are a component broadly utilized as a part of programmed discourse and speaker acknowledgment. They were presented by **Davis and Mermelstein in the 1980's**, and have been best in class from that point onward. Before the presentation of MFCCs, Linear Prediction Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) and were the fundamental element sort for programmed discourse acknowledgment (ASR).

Lawrence R. Rabiner proposed Hidden Markov Models and Selected Applications in Speech Recognition" **Proceedings of the IEEE, 77 (2), p. 257–286, (1989-02-09)**

He attempt to carefully and methodically review of the theoretical aspects of this type of statistical modeling and shown how they have been applied to selected problems in machine recognition of speech.

K Sreenivasa Rao and **Shashidhar G Koolagudi** "**Identification of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech**", SYSTEMICS, CYBERNETICS

School of Information Technology, Indian Institute of Technology Kharagpur, Kharagpur - 721302, West Bengal, India. They have explored speech features to identify Hindi dialects and emotions. A dialect is any distinguishable variety of a language spoken by a group of people. Emotions provide naturalness to speech. In this work, five prominent dialects of Hindi are considered for the identification task. They are Chhattisgarhi (spoken in Central India), Bengali (Bengali accented Hindi spoken in Eastern region), Marathi (Marathi accented Hindi spoken in Western region), General (Hindi spoken in Northern region) and Telugu (Telugu accented Hindi spoken in Southern region).

Greg Cox proposed On the **Relationship Between Entropy and Meaning in Music**: An Exploration with Recurrent Neural Networks.

He described, current neural network model which produces estimates of instantaneous entropy for music with multiple parts and use it to analyze a Haydn string quartet .Features found by traditional analysis to be related to tension are shown to have characteristic signatures in the model's entropy measures. Thus, an information based approach to musical analysis can elaborate on traditional understanding of music and can shed light on the more general cognitive phenomenon of musical meaning.

M. Chinna Rao, A.V.S.N. Murthy, Ch. Satya Narayana published **Emotion Recognition System Based On Skew Gaussian Mixture** Model and MFCC Coefficient published on ***I.J. Information Engineering and Electronic Business*, 2015, 4, 51-57.**

They used Skew Gaussian mixture model. The proposed model has been experimented over a gender independent emotion database. In order to extract the features from the speech signals cepstral coefficients are used. The developed model is tested using real-time speech data set and also using the standard and data set of Berlin. This model is evaluated in the presence of noise and without noise the efficiency of the model is evaluated and is presented by using confusion matrix

3. OBJECTIVE

Speech recognition is the analysis side of the subject of machine speech processing. The synthesis side might be called speech production. These two taken together allow computers to work with spoken language. This study focuses on emotion acknowledgment from a speech. The speech identification, in people, is a huge number of years old. On our planet it could be followed supported a large number of years to the dinosaurs.

- Creating emotional corpus in continuous Hindi speech
- Finding prosodic elements
- Comparative examination of prosodic elements of feelings undertaken

The first step in any automatic speech recognition system is to find features that are identify the components of the audio signal that are useful for recognizing the semantic substance and tossing the various stuff which conveys data like background noise, emotions and so forth.

The principle point to see about speech is that the sounds produced by a human are separated by the state of the vocal tract including tongue, teeth and so forth. This shape figures out what sound turns out. In the event that we can focus the shape precisely, this ought to issue us a precise representation of the phoneme being delivered. The shape of the vocal tract manifests itself in the envelope of the short time power spectrum, and the job of MFCCs is to accurately represent the objective.

4. ARCHITECTURE OF PROPOSED SYSTEM

In this proposed system we design and develop a machine learning classifier to analyses the emotions of the sounds using signals analysis. Dataset can created by recording the sounds of five males and five females. So there will be total ten persons.

From each person, there will be twenty recording. In proposed system five sentences in different emotions (joy, sad, anger, normal) could be spoken by 10 people.

Four Emotions:

(A) Normal mode

(B) Angry mode

(C) Joy mode

(D) Sad mode

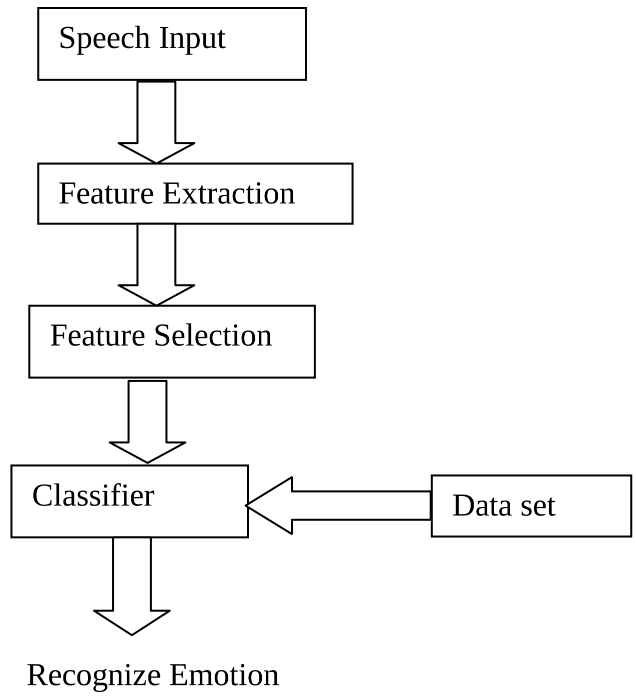
For each person 1(sentence)*4(emotions)=4 Files

For 5 different sentences for each person=4Files*5=20 Files

For 10 persons total no. of files=20Files*10=200 Files.

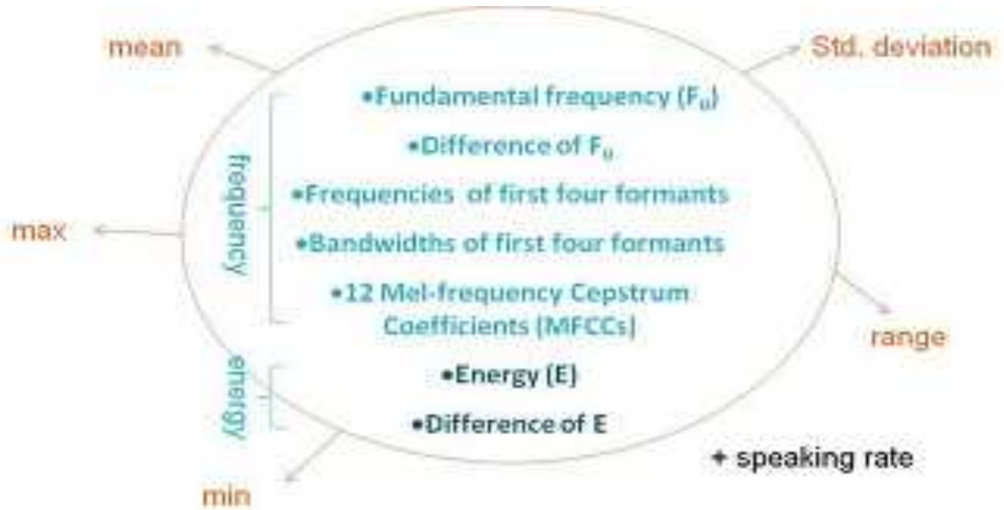
These sounds in different emotions can be recorded using mobile app Voice Recorder.

The files recorded in this application are in the MPEG4 format then they can be converted to wav file to use as inputs.

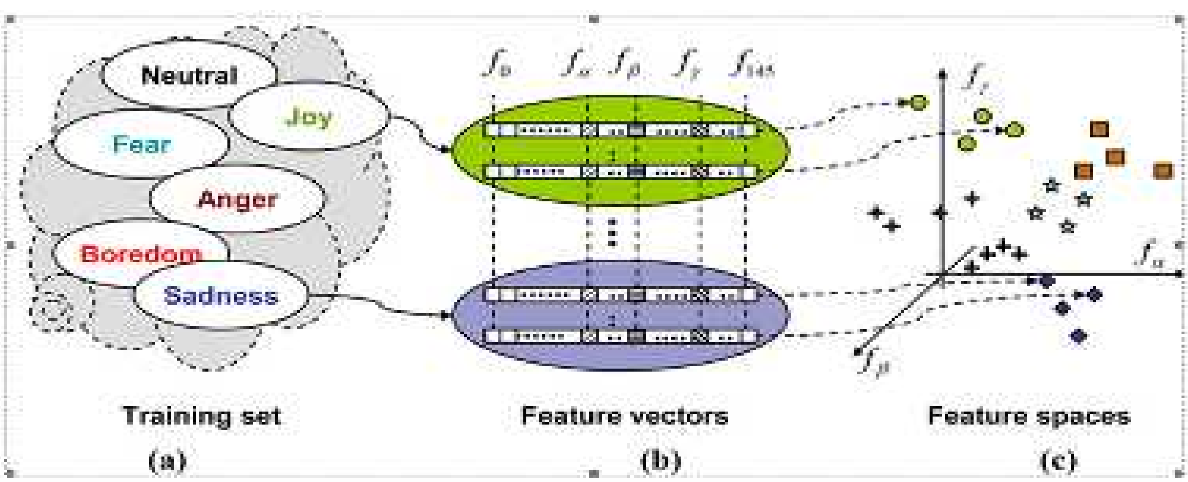


5. FEATURES

All the emotion recognition experiments described in this paper have been performed with the database, a collection of 200 wav files of 10 speakers. In the prosody based recognition system, a nine-feature vector (already used in [2]) was obtained for each conversation side: three features related to word and segmental durations - number of frames per word and length of word-internal voiced and unvoiced segments - and six features related to fundamental frequency - mean, maximum, minimum, range, pseudo-slope and slope are important can be evaluated using mat lab.



6. FEATURE EXTRACTION



7. RESEARCH AND METHODOLOGIES

MIRtoolbox: MIRtoolbox offers an incorporated arrangement of capacities written in Matlab, devoted to the extraction from sound records of musical components, for example, tonality, rhythm and other measurable investigation on signals. This tool stash proposes a huge arrangement of musical component extractors. MIRtoolbox is a Matlab tool stash devoted to the extraction of musical components from sound documents, including schedules for measurable analysis. The target is to offer a review of computational methodologies in the zone of *Music Information Retrieval.* The outline is taking into account a secluded structure: the diverse calculations are disintegrated into stages, formalized utilizing a negligible arrangement of rudimentary instruments. These building squares frame the essential vocabulary of the toolkit, which can then be openly explained in new unique ways. These rudimentary components coordinate all the distinctive variations proposed by option approaches - including new methods we have created, that clients can choose and parametrize. This engineered review of highlight extraction instruments empowers a promotion of the innovation offered by all the option techniques. Furthermore to the essential computational procedures, the toolkit additionally incorporates more elevated amount musical element extraction devices, whose option methods, and their numerous mixes, can be chosen by the client. MIRtoolbox is taking into account a situated of building hinders that can be parametrized, reused, reordered etc.

MIRtoolbox References: Olivier Lartillot, Petri Toiviainen, Tuo+mas Eerola, "A Matlab Toolbox for Music Information Retrieval", in C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker (Eds.)

Data Analysis, Machine Learning and Applications, Studies in Classification, Data Analysis, and Knowledge Organization, Springer-Verlag in 2008.

Olivier Lartillot, Petri Toiviainen, "A Matlab Toolbox for Musical Feature Extraction From Audio", International Conference on Digital Audio Effects, Bordeaux in 2007.

MIRtoolbox Features: MIRtoolbox incorporates around 50 sound and music elements extractors and measurable descriptors. MIRtoolbox requires the Matlab environment, form 7, and does not work exceptionally well with past forms of Matlab. This is expected specifically to the truth MIRtoolbox depends on multi-dimensional clusters and different yields, which appear to be elements presented by adaptation 7. MIRtoolbox additionally obliges that the Signal Processing Toolbox, one of the discretionary sub-bundles of Matlab, be appropriately introduced.

Regardless, a specific number of administrators can adjust to the nonappearance of this tool kit, and can deliver pretty much solid results. At the same time, for genuine utilization of MIRtoolbox, we emphatically suggest a fitting establishment of the Signal Processing Toolbox.

MIRtoolbox incorporates in its appropriation a few other openly accessible tool kits, that are utilized for particular calculations.

• The Auditory Toolbox, by Malcolm Slaney (1998), is utilized for Mel-band range and MFCC processing's, and Gammatone filterbank decay.
• The Netlab tool stash, by Ian Nabney (2002), where the schedules for Gaussian Mixture Modeling(GMM) is utilized for order (mirclassify).
• Finally, the SOM tool stash, by Esa Alhoniemi and partners (Vesanto, 1999), where just a routine for bunching in view of k-means strategy is utilized, in the mircluster capacity.

At that point open the "Set Path" environment accessible in Matlab File menu, tap on "Include with Subfolders...", peruse into the record chain of command and select the principle MIRtoolbox organizer, then snap "Open". You can then "Spare" and "Close" the Set Path environment.

Help and demos

To get a diagram of the capacities accessible in the tool stash, sort:

help mirtoolbox

A short documentation for every capacity is accessible utilizing the same help charge. For instance, type:
help miraudio

8. MIRTOOLBOX INTERFACE

All capacities are gone before by the mir- prefix with a specific end goal to evade clashes with other Matlab capacities. Every capacity is identified with a specific information sort: for case, miraudio is identified with the stacking, change and presentation of sound waveform. A sound record, how about we say a WAV document of name song_file.wav, can be stacked just by composing the summon:

miraudio('song_file.wav')

Operations and alternatives to be connected are shown by specific essential words, communicated as contentions of the capacities. Case in point, the waveform can be focused utilizing the "Inside" catchphrase:

miraudio('song_file.wav', 'Center')

which is comparable to any of these parameters:

miraudio('song_file.wav', 'Center', 'yes')

miraudio('song_file.wav', 'Center', 'on')

miraudio('song_file.wav', 'Center', 1)

though the inverse arrangement of parameters

miraudio('song_file.wav', 'Center', 'no')

miraudio('song_file.wav', 'Center', 'off')

miraudio('song_file.wav', 'Center', 0)

are redundant on account of the "Middle" alternatives as it is flip off of course in miraudio.

the distinctive alternatives can be consolidated in one single summon line:

miraudio('song_file.wav', 'Center', 'Examining', 11025)

9. IMPLEMENTATION

Experimental Parameters: The parameters considered for the proposed work are entropy; mode and pitch are separated for every sentence to anticipate significance of the sentences.

Pitch: The pitch of a sound can be said to be basically a measure of its frequency, there are circumstances in which a relentless frequency sound can be seen to be changing in pitch.

Sentence length: The term of a sentence, LEN, is utilized. LEN incorporates delay time in the sentence

Analysis of a Single Chord: This is the fundamental step which is identified with the pitch of the signs. It watches the periodicities contained in the sign by figuring the autocorrelation. The tops of the bend demonstrates the most essential periodicities

Entropy: Entropy is a measure of the width and consistency of the power range. Entropy, of course, is a measure of the "issue" of a framework.

EXPERIMENTS: The simulation tool for this task is made in Matlab. Each examination uses a couple autonomously recorded examples. Some are used for get ready and the others are used for testing. We make a point not to test with our planning data and not to get ready with our testing data. The undertakings consider parameter assortment and figuring substitution for key things, for instance, highlight extraction and class determination. The pack moreover is expected to gather and record noteworthy bits of knowledge on the accuracy of affirmation. The general believed is to change estimation and/or distinctive controlling parameters and rerun the standard test, seeing any change or degradation in the experiences that identify with successful affirmation. The results that I report are from runs (tests) using two voices (one male and one female). The essential code is found pitch method acquired Train Pitch. The model for find pitch limits is according to the accompanying:

FindPitch() Method

pitch = findpitch(dataset/voice_signal)

This function is to study the varying devices available in the toolbox for pitch extraction. It endeavor to think the pitch substance of diverse sound records using varying methods. All the sound documents capacities are joined in the MIRtoolbox

Analysis of a sign harmony

% Load the sound document .The records are browsed the dataset. Dataset contains five envelopes for unmistakable customers having sounds in differing emotions.

```
a = miraudio('dataset/u1/001.wav')
% Observe the periodicities contained in the sign by figuring the autocorrelation limit:
air conditioning = mirautocor(a);
ac= mirautocor(ac, 'Freq','Min',75,'Hz','Max',2400,'Hz');
% The peaks of the twist shows the most basic periodicities:
pac = mirpeaks(ac);
% But the two peaks toward the start and end of the twists are immaterial, so should be cleared:
pac = mirpeaks(ac, 'NoBegin','NoEnd');
% The relating pitch tallness are given by:
p = mirpitch(pac);
```

% The bona fide numerical data (here, the repeat of each pitch) is given by:

mirgetdata(p);

%% Upgrading the examination

air conditioning = mirautocor(a,'Compres',.5);

% the straggling leftovers of the handling can truly be performed using the simple course:

[p pac] = mirpitch(ac);

% Look and listen to the results.

mirgetdata(p);

%% Upgrading the examination, again

d=mirpitch(a);

array=mirgetdata(d);

max(array);

pitch=ans;

pitch interference, close all

%% Monody examination

As the tune is sung by one voice nobody however, you can use the Mono decision in mirpitch.

mirplay('laksin')

o = mironsets('laksin','Attacks','Contrast',.1)

mirplay(o)

sg = mirsegment('laksin',o)

mirplay(sg)

p = mirpitch(sg,'mono')

mirgetdata(p)

mirplay(p)

Train Pitch () Method: Train Pitch () function is calling the Find Pitch () method recursively for each user to find the pitch estimations of their sounds. The sound records contain unmistakable Hindi sentences in differing emotions. These contribute qualities are set away contribute table database. Table: Pitch, Entropy and Mode Tables:

FIELD	DESCRIPTION
Username	Username
Neutral	Sound in Neutral emotion
Happy	Sound in Joyful emotion
Sad	Sound in Sadness
Angry	Sound in Anger

2 Table Data | 2 Info

Limit rows First row 0

users	Neutral	Happy	Sad	Angry
U1	751.62711827627425	257.31969911945453	640.11101856522612	261.08295232773388
U1	542.35246965856566	580.67938170283774	214.40487230116187	502.26665131521432
U1	254.96182480304399	633.22224139853897	459.35330992971419	236.97737814401381
U2	502.42880737506073	522.61798344376777	557.27100256270718	444.40622992379161
U2	489.28332945908056	603.14844417669758	505.836183861828	615.15325273930493
U2	475.80302385939706	440.59457230610911	541.05585166604141	436.0865433840997
U3	487.86520176168676	735.2403050272784	253.53054394648129	658.43122884791501
U3	187.04133644948536	532.95873706316002	262.15546533173023	622.55988440999473
U3	236.41580272081784	329.85132797633105	198.40493254046328	528.22294670583824
U4	454.90216484009932	728.49802558423187	701.77484426720621	710.37645499904966
U4	498.96213585168056	505.64224992165271	516.41047219855341	318.68266109823559
U4	722.33470941787821	316.3915777547752	708.85365227163766	300.53759849790669
U5	215.83182399914108	277.6994063071507	253.53054394648129	658.43122884791501
U5	187.04133644948536	532.95873706316002	235.07280954874521	255.58805919641229
U5	218.23792700389815	282.40268553734438	198.40493254046328	243.9688193716658
(NULL)	(NULL)	(NULL)	(NULL)	(NULL)

Train Entropy () Method: This method reads the entropy of the sound files using mirentropy () method. The results are further converted to scalar entities and then stored in Entropy Table in the database. It is also likely that different dimensions of sound (e.g., pitch, rhythm, harmony) contribute differently to tension and to entropy

2 Table Data | 2 Info

Limit rows First row 0

usern	Neutral	Happy	Sad	Angry
U1	0.88343007697493134	0.84941426154835924	0.84005460511984816	0.92264950599140811
U1	0.83338056036747372	0.86182242651602847	0.90081624850964803	0.92717977749267342
U1	0.84120312816197119	0.86419716885487898	0.88636065066310413	0.88485661540378091
U2	0.86124391509454357	0.80809986058109215	0.81100643527291072	0.86582239455240817
U2	0.77881650320420093	0.79743156496825107	0.76031873900951985	0.84413480228474946
U2	0.84889511049082082	0.87128376451502053	0.80541531772685 06	0.87152427986886198
U3	0.85065645130731771	0.84474389631265532	0.83055798852242555	0.84389657196129397
U3	0.80287292994841787	0.80345609783489824	0.81490126840624 91	0.80527797927743713
U3	0.880207179427633 6	0.81392526610276883	0.87310520258723445	0.85451712462815443
U4	0.90451973706 75202	0.90498633479886004	0.91607495750181638	0.91295907363674333
U4	0.91370727297614813	0.89162844107968741	0.90520707325714322	0.91089773740865 08
U4	0.91408813749684723	0.91948948890805 0043	0.90734197051053445	0.89557744754519211
U5	0.87155694720802412	0.84678077813665273	0.83055798852242555	0.84389657196129397
U5	0.80287292994841787	0.80345609783489824	0.83143561606598149	0.81387419189593213
U5	0.88084157857494172	0.86745882047406381	0.87310520258723445	0.85013640551276692
(NULL)	(NULL)	(NULL)	(NULL)	(NULL)

TrainMode () Method: This method reads the mode of the sound files using mirmode () method. The results are further converted to scalar entities and then stored in Mode Table in the database.

Pitch Features () Method: these methods find the average pitch values for each user. These pitch values are stored in Sound_Featues Table in the database.

Filed Name	Description
UserID	User Id
Feature	Features of the sound
Neutral	When a person is Neutral
Happy	When a person is Happy
Sad	When a person is Sad
Angry	When a person is Angry

Table: Database Table to store features of the sound

This method reads the pitch values from the Pitch table and store the average of each emotion for each user in Sound_features table.

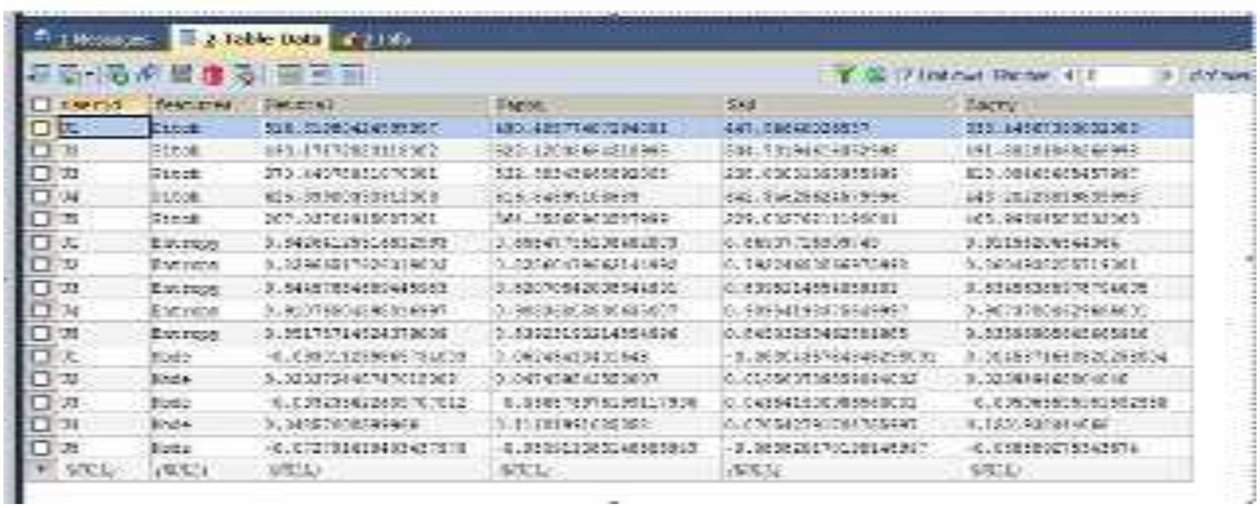


Figure: Sound Features stored in database

Above figure shows all the features of the sound stored in sound_features table

EntropyFeatures() Method: This method find the average entropy values for each user. These entropy values are stored in Sound_Featues Table in the database.

Mode Features () Method: This method finds the average mode values for each user. These mode values are stored in Sound_Featues Table in the database.

userid	features	Neutral	Happy	Sad	Angry
U1	Pitch	516.31980424595997	490.40577407294001	447.96640026537	333.44567393032003
U2	Pitch	490.17172023118002	522.12033664218995	534.73194616352998	491.88201868266998
U3	Pitch	370.44078031070001	532.68345668892005	238.03031393955999	613.08468665457997
U4	Pitch	625.39900330312003	516.84395108689	642.34625624579996	443.20223819839998
U5	Pitch	207.03702915087001	364.35360963587999	229.00276211190001	405.99604580533003
U1	Entropy	0.84264125516812993	0.85847795208482003	0.86907716909749	0.91156206544064
U2	Entropy	0.82966517626319003	0.82560479662144992	0.79224683066975998	0.86049388255719301
U3	Entropy	0.84457884689445983	0.82070842008344991	0.8395214654059101	0.83456388978794005
U4	Entropy	0.91076504898016997	0.90536808830635007	0.90954198375649997	0.90737808629686001
U5	Entropy	0.85175714524379009	0.83923193214854996	0.84503293462581985	0.83596905645665986
U1	Mode	-0.093011239869734003	0.06243413401648	-0.0030435784345298001	0.0045871660820258004
U2	Mode	0.023372448747615002	0.047459843550007	0.016563739859094002	0.023949468804048
U3	Mode	-0.093233622639707012	-0.066576975199117996	0.043541900986560002	-0.095965595981592988
U4	Mode	0.04857008599966	0.11101991085358	0.070542791704755997	0.125193804406
U5	Mode	-0.072791619403437979	-0.080911383148556993	-0.083820170198148997	-0.038889275343674
(NULL)	(NULL)	(NULL)	(NULL)	(NULL)	(NULL)

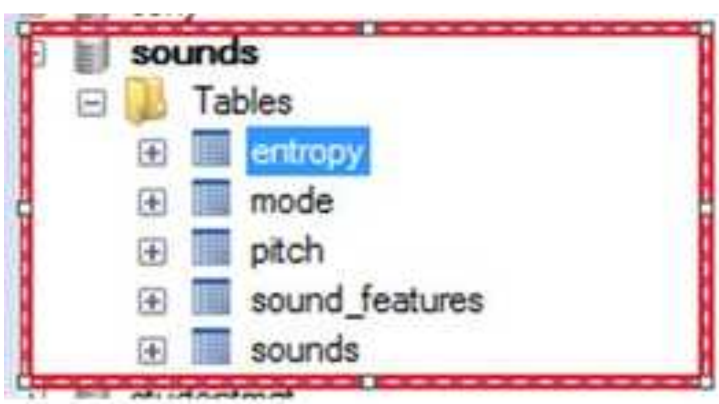
Figure: Mode values for different users

Pitch_compare() Method and Entropy_Compare() Method:

These methods are used to compare Pitch and Entropy differences for the different persons.

Simulation Tool:

For simulation tool the dataset files are stored in the MATLAB project folder it is dataset. To read the files from dataset, Mirfolder command can be used. The database in My SQL can be used to store details of the wav files.



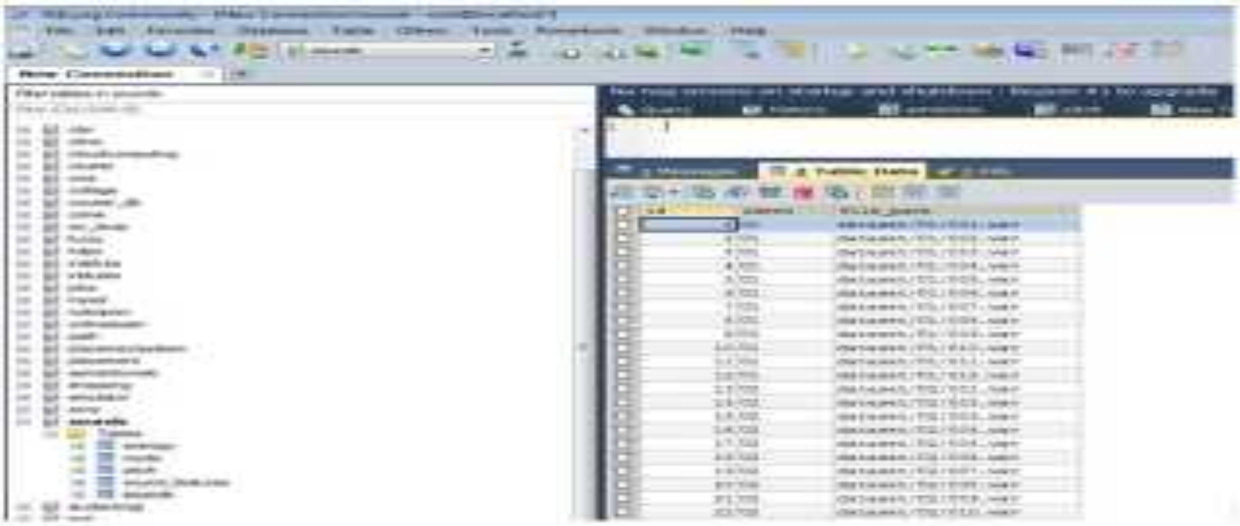


Figure: Sound Database in My SqL

Sounds Table stores the path of the sound files stored.

FIELD	DESCRIPTION
ID	ID
Username	Username
File_Path	File Path

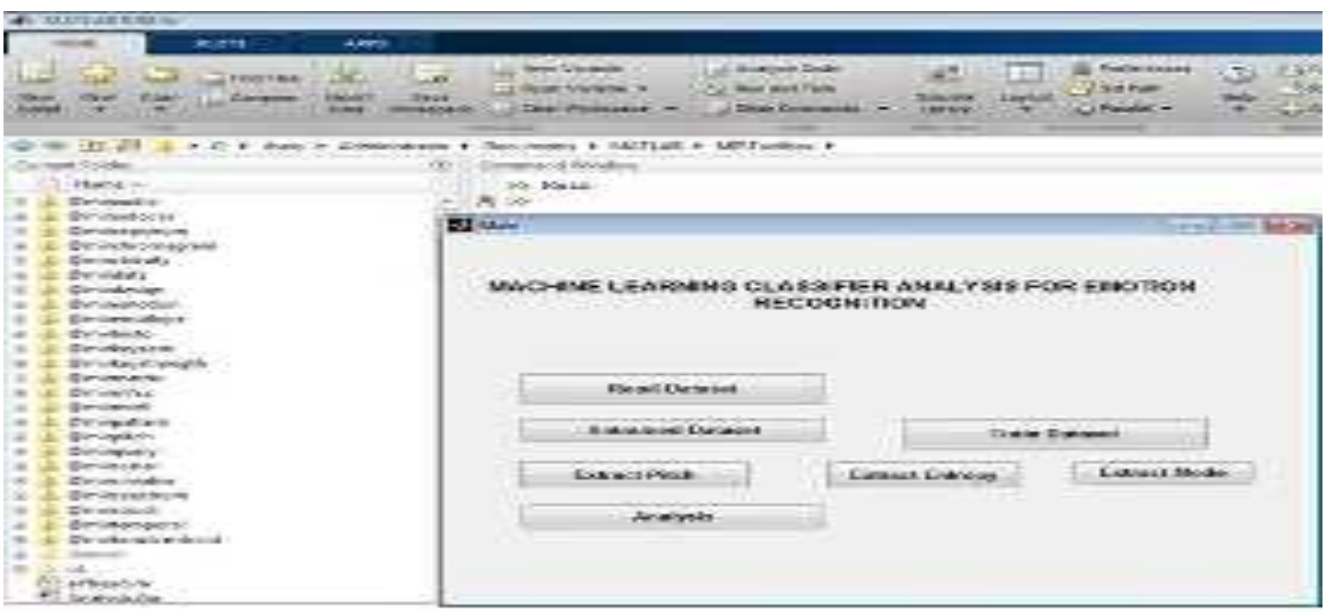


Figure: Main Interface of the Simulation Tool

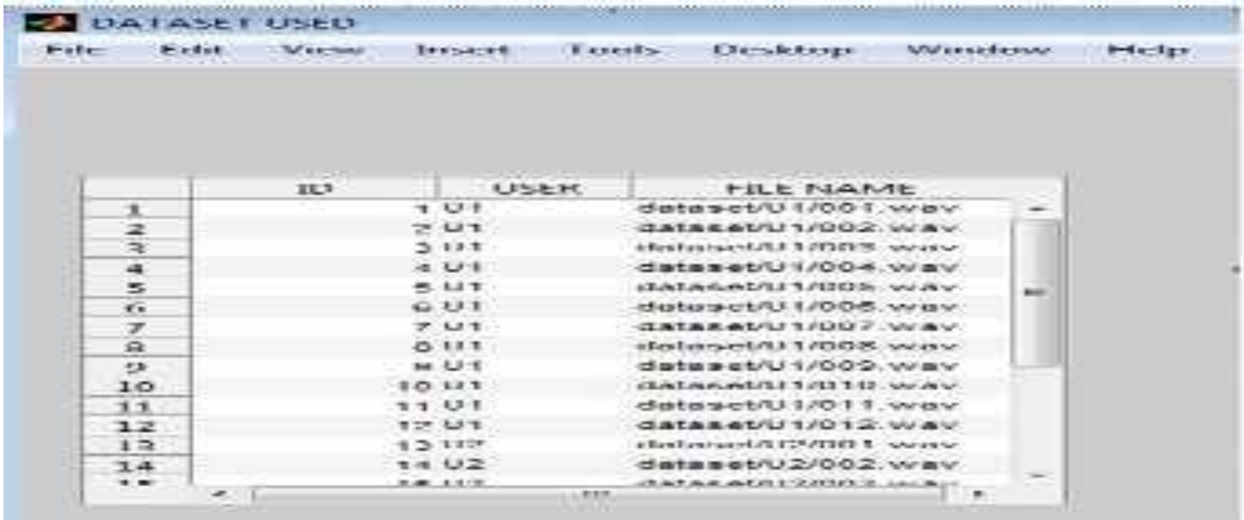


Figure: Database used

Read Dataset: Read Dataset button displays the list of wav files in different emotions.

Untrained Dataset: Untrained dataset button delete the features stored in the database.

Figure: Untrained Dataset

Train Dataset: Trained dataset read the wav files of all the users, calculate pitch, mode and entropy of all the emotion files. Pitch values are stored in pitch table, mode is stored in mode table and entropy is saved in entropy table.

Extract Pitch, Extract Mode and Extract Entropy: These buttons are used to prepare the complete trained dataset which explains about the features of sounds of a particular person. Way of speaking of all the persons has slight difference. Some persons speak very lightly in sad mood. Some speaks in the similar tone as in neutral form. Trained dataset records the way of talking of the users.

Figure: Features are extracted in Sound_feature table on Training Dataset

Figure: Computing Pitch, Entropy and Mode

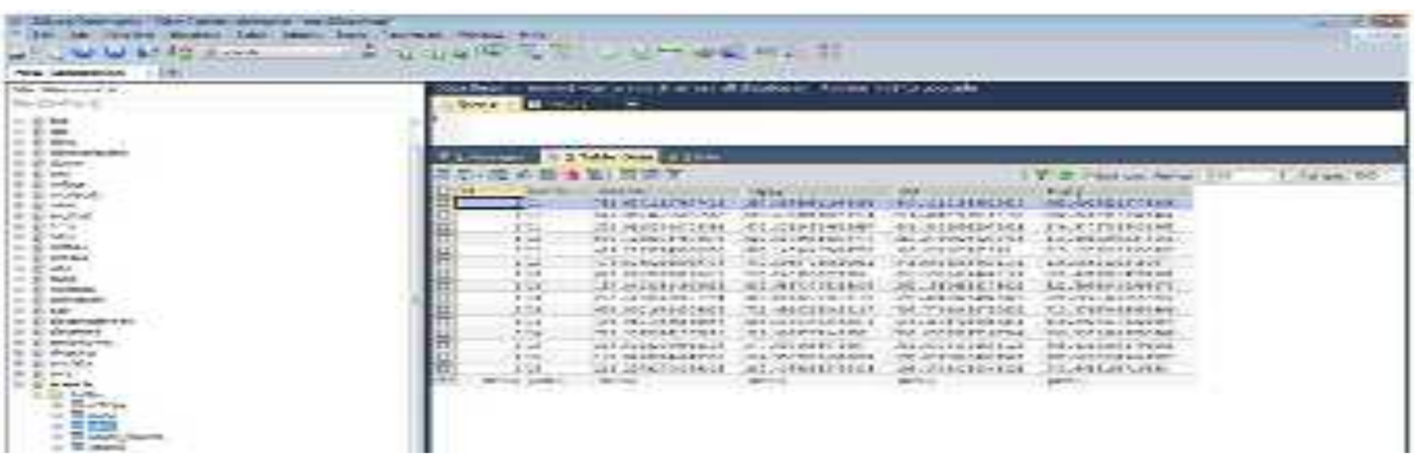


Figure: Features Extracted are stored in database

Sounds may be by and large portrayed by pitch, loudness, and quality. The apparent pitch of a sound is only the ear's response to frequency, i.e., for most reasonable purposes the pitch is only the frequency. The pitch perception of the human ear is comprehended to work fundamentally by the spot hypothesis, with some honing instrument important to clarify the remarkably high determination of human pitch recognition.

The pitch of a sound can be said to be essentially a measure of its frequency, there are circumstances in which a steady frequency sound can be seen to be changing in pitch. One of most reliably watched "psychoacoustic" impacts is that a maintained high frequency sound (>2 kHz) which is expanded relentlessly in power will be seen to be ascending in pitch, while a low frequency sound (<2kHz) will be seen to be dropping in pitch.

The impression of the pitch of short heartbeats contrasts from that of managed hints of the same measured frequency. On the off chance that a short beat of an unadulterated tone is rotting in abundance, it will be seen to be higher in pitch than an indistinguishable heartbeat which has enduring sufficiency. Meddling tones or clamor can bring about an evident pitch shift.

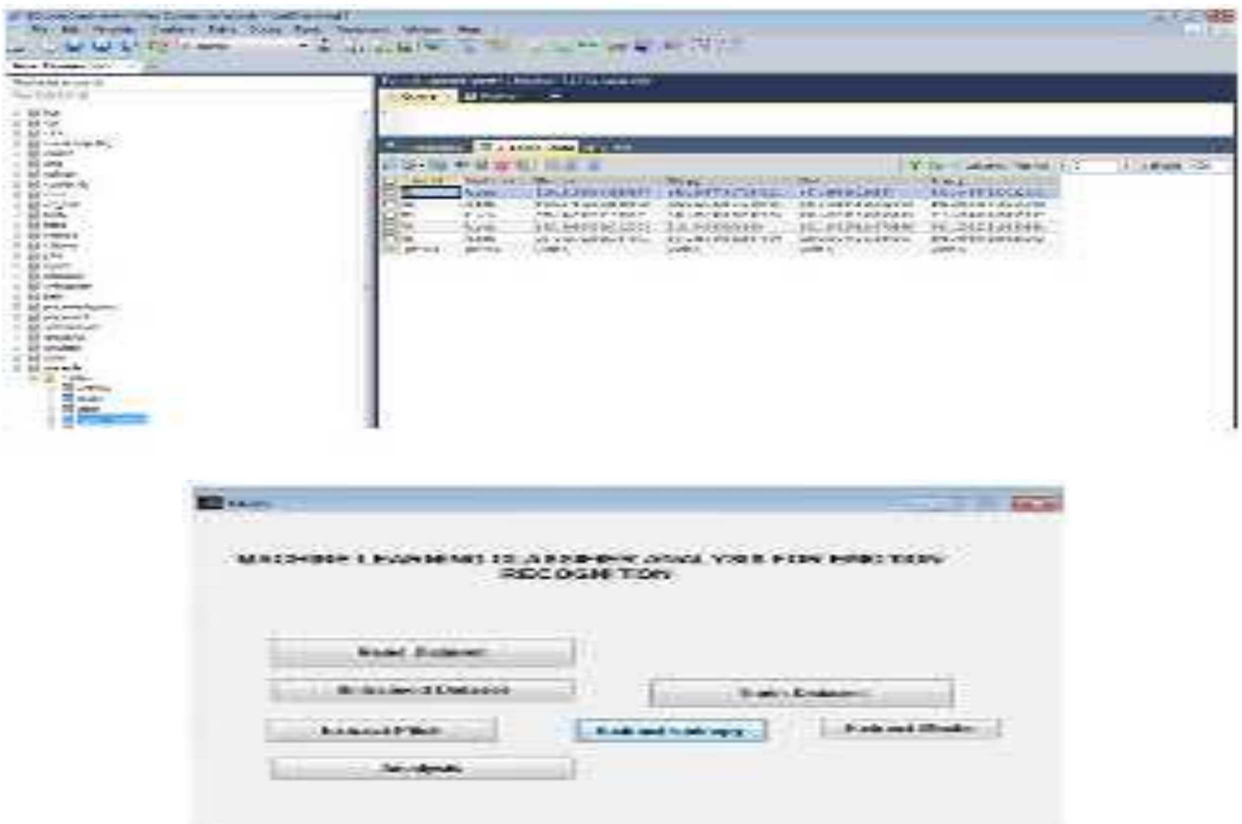


Figure: Click Extract Entropy to extract entropy of the recorded sounds

Entropy is a measure of the width and uniformity of the power spectrum. ***Entropy***, on the other hand, is a measure of the "disorder" of a system

userid	features	Neutral	Happy	Sad	Angry
U1	Pitch	516.31980424595997	490.40577407294001	447.96640026537	333.44567393032003
U2	Pitch	490.17172023118002	522.12033664218995	534.73194616352998	491.88201868266998
U3	Pitch	370.44078031070001	532.68345668892005	238.03031393955999	613.08468665457997
U4	Pitch	625.39900330312003	516.84395108689	642.34625624579996	443.20223819839998
U5	Pitch	207.03702915087001	364.35360963587999	229.00276211190001	405.99604580533003
U1	Entropy	0.84264125514812993	0.85847795208482003	0.86907716909749	0.91156204544064
U2	Entropy	0.82966517824319003	0.82560479662144992	0.79224683066975998	0.8604938255719301
U3	Entropy	0.84457884489445983	0.82070842008344991	0.83952148854059101	0.83456388978794009
U4	Entropy	0.91076504898016987	0.90536808830635007	0.90954195375649997	0.90737808629686001
U5	Entropy	0.85175714524379009	0.83923193214854996	0.84503293462581985	0.83596905645665986

Figure: Entropy recorded

A key measure of data is entropy, which is typically communicated by the normal number of bits expected to store or convey one symbol in a message. Entropy measures the instability included in anticipating the estimation of an irregular variable. Case in point, determining the result of a reasonable coins flip (two just as likely results) gives less data (lower entropy) than indicating the result from a move of a pass on (six similarly likely results).

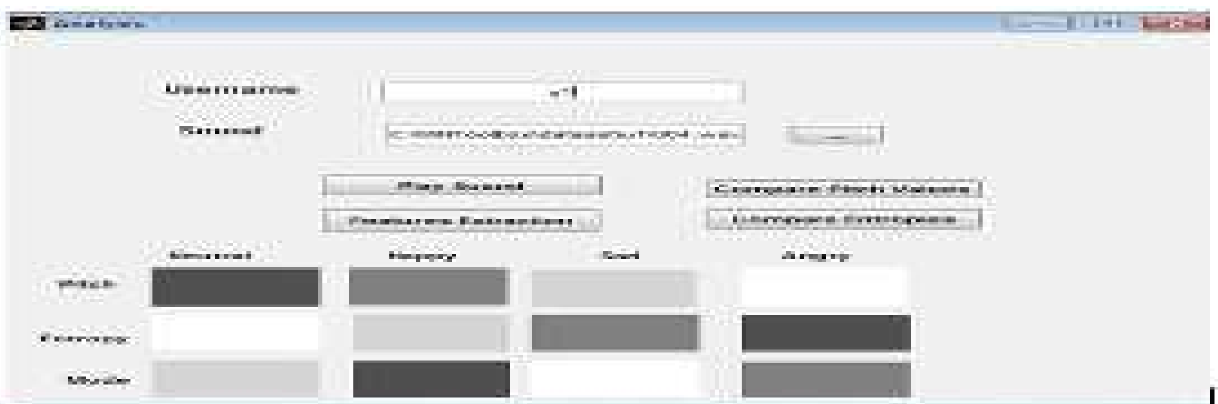


Figure: Features Extracted

userid	features	Neutral	Happy	Sad	Angry
U1	Pitch	516.31980424595997	490.40577407294001	447.96640026537	333.44567393032003
U2	Pitch	490.17172023118002	522.12033664218995	534.73194616352998	491.88201868266998
U3	Pitch	370.44078031070001	532.68345668892005	238.03031393955999	613.08468665457997
U4	Pitch	625.39900330312003	516.84395108689	642.34625624579996	443.20223819839998
U5	Pitch	207.03702915087001	364.35360963587999	229.00276211190001	405.99604580533003

.Table: Pitch values in different emotions

userid	features	Neutral	Happy	Sad	Angry
U1	Entropy	0.84264125516812993	0.85847795208482003	0.86907716909749	0.91156206544064
U2	Entropy	0.82966517626319003	0.82560479662144992	0.79224683066975998	0.8604938255719301
U3	Entropy	0.84457884689445983	0.82070842008344991	0.8395214854059101	0.83456388978794005
U4	Entropy	0.91076504898016997	0.90536808830635007	0.90954198375649997	0.90737808629686001
U5	Entropy	0.85175714524379009	0.83923193214854996	0.84503293462581985	0.83596905645665986

Table: Entropy values in different emotions

userid	features	Neutral	Happy	Sad	Angry
U1	Mode	-0.093011239869734003	0.06243413401648	-0.0030435784345298001	0.0045871660820258004
U2	Mode	0.023372448747615002	0.047459843550007	0.016563739859094002	0.023949468804048
U3	Mode	-0.093233622639707012	-0.066576975199117996	0.043541900986560002	-0.095965595981592988
U4	Mode	0.04857008599966	0.11101991085358	0.070542791704755997	0.1251938044066
U5	Mode	-0.072791619403437979	-0.080911383148586993	-0.083820170198148997	-0.038889275343674

Table: Mode values in different emotions

10. RESULTS AND CONCLUSION:

Figure: Graphical representation of Different Emotions

Above figure shows the comparison between basic concepts of emotions (happy, sad, tender, anger, fear) and emotion dimensions (activity, valence, tension).The pitch values for the neutral and sad emotions are likely to be same. Entropy values for the angry sounds are higher as compared to sad and neutral emotions.

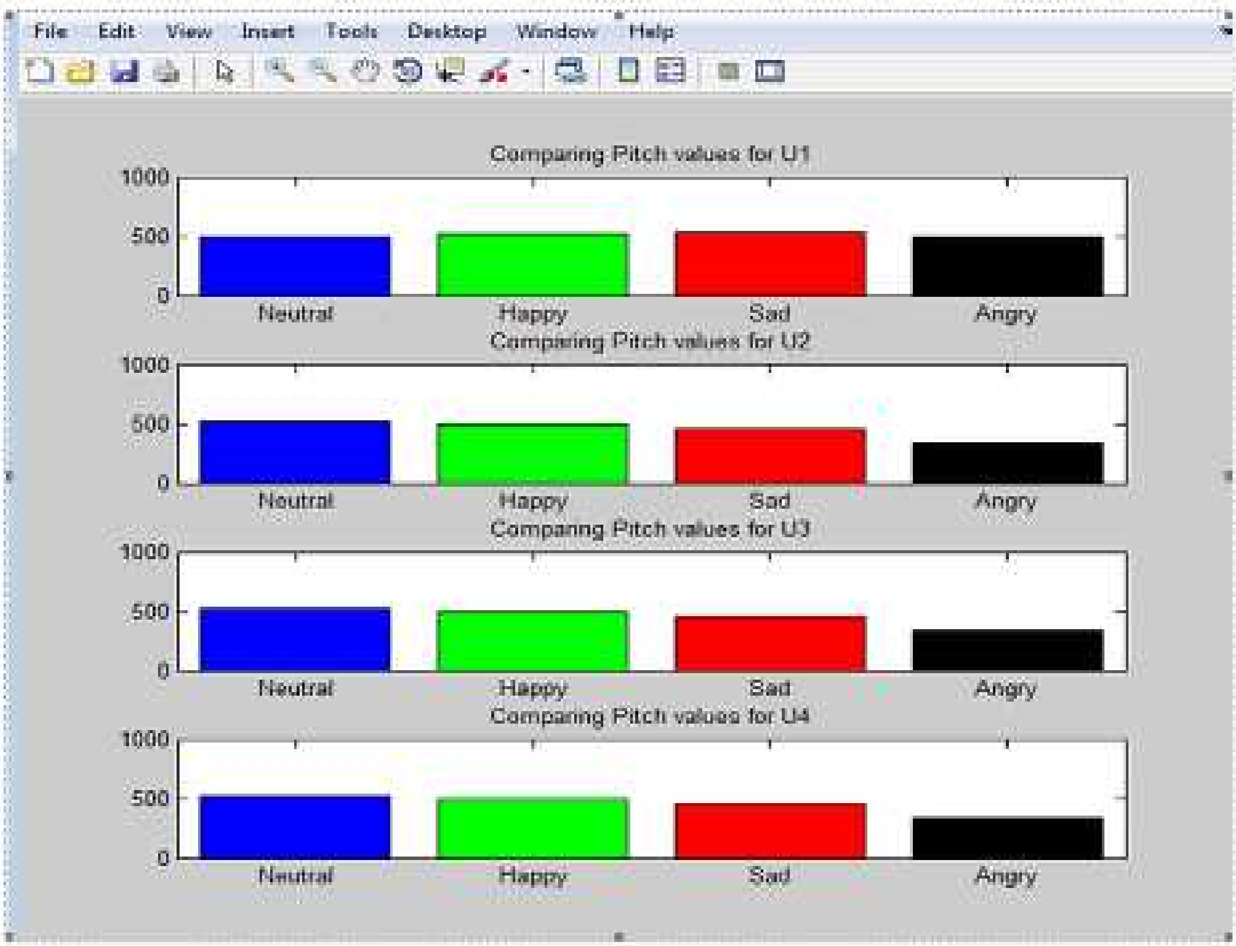


Figure: Pitch is lower in Angry Sound

The proposed system have found that the emotions, Neutral, joy and anger, are portrayed at a higher frequency than emotions such as sadness

- Anger: Anger can be divided into two types: "anger" and "hot anger". In comparison to neutral speech, anger is produced with a lower pitch, higher intensity, more energy (500 Hz) across the over the vocalization, higher first formant (first solid created) and speedier attack times at voice onset (the start of speech). "Hot anger", in contrast, is produced with a higher, more varied pitch, and even greater energy (2000 Hz).
- Happy: Fear can be divided into two types: "joy" and "anxiety". In comparison to angry speech, happy emotions have a higher pitch, little variation, higher energy, and a slower speech rate with more pauses.
- Sadness: In comparison to neutral speech, sad emotions are produced with a slightly higher pitch, less intensity but more vocal energy (2000 Hz), longer duration with more pauses, and a lower first formant.

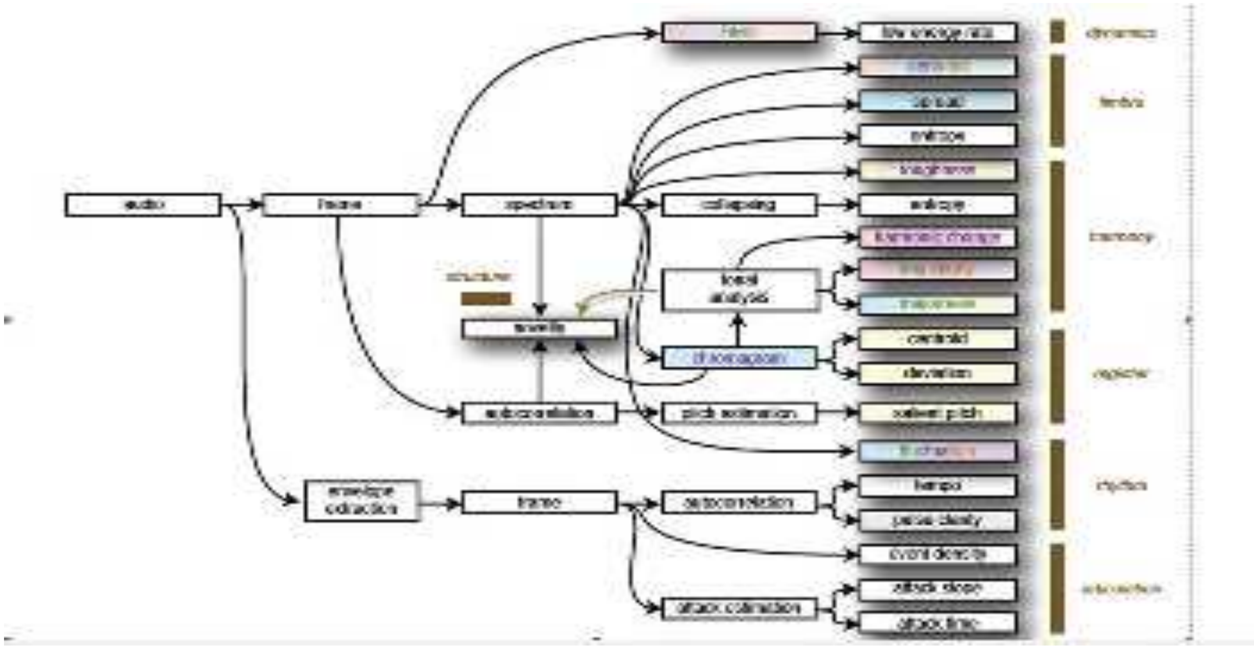


Figure: Features in sound

Audio and musical features founding the prediction of emotion. The colors associated to each feature corresponds to the correlating emotions, following the same color code that in the previous figure.

MFCCs Results:

Figure: MFCC for the sound of first user

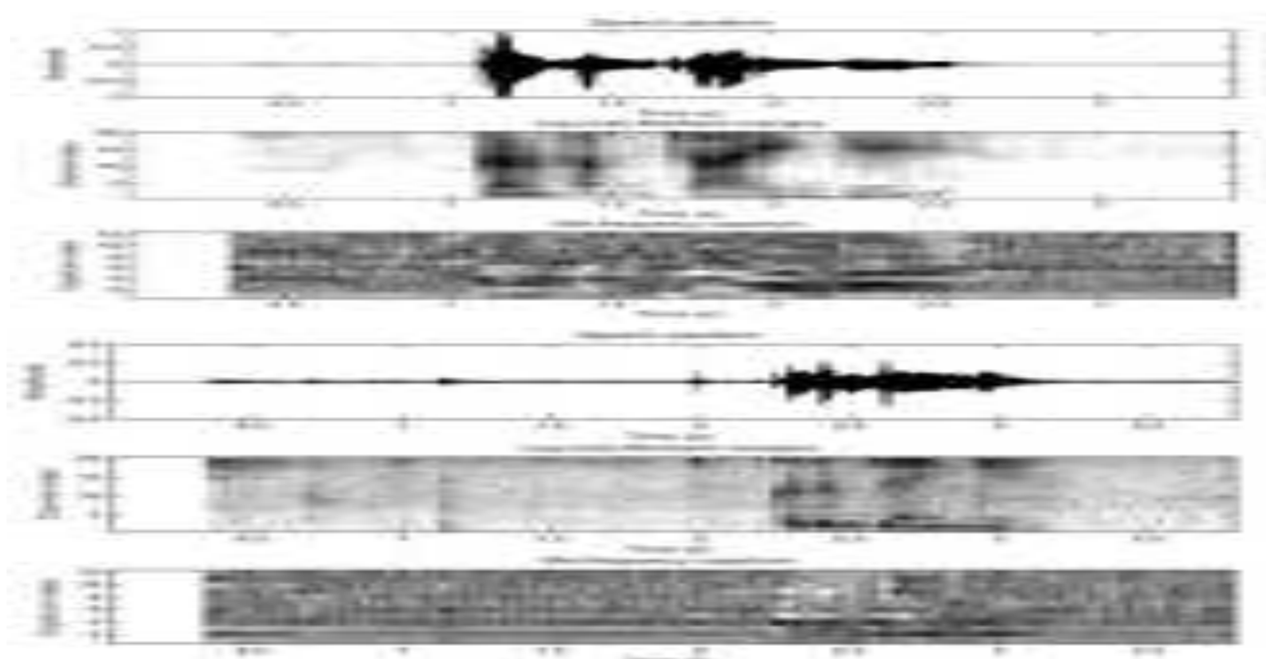


Figure: MFCC for the sound of first user

11. CONCLUSION

MFCC offers a description of the spectral shape of the sound. The frequency bands are positioned logarithmically (on the Mel scale) which approximates the human auditory system's response more closely than the linearly-spaced frequency bands. This features as studied earlier are also considered in the proposed work are not providing good results. From the features extraction from the sound using MIRtoolbox a machine can find the differences between the speech of the person in contrast to his emotion. This work is proposed to be used by machines to understand what a person wants exactly to be performed. The comparison between basic concepts of emotions (happy, sad, tender, anger, fear) and emotion dimensions (activity, valence, tension).The pitch values for the neutral and sad emotions are likely to be same. Entropy values for the angry sounds are higher as compared to sad and neutral emotions. Way of speaking of all the persons has slight difference. Some persons speak very lightly in sad mood. Some speaks in the similar tone as in neutral form. Trained dataset records the way of talking of the users. This trained dataset is used to study the emotions of the users that can be further used in machine learning.

12. REFERENCES

[1] K Sreenivasa Rao and Shashidhar G Koolagudi " Identification of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech", SYSTEMICS, CYBERNETICS AND INFORMATICS VOLUME 9 - NUMBER 4 - YEAR 201126 ISSN: 1690-4524

[2] Vayrynen, Eero, "Emotion recognition from speech using prosodic features". University of Oulu Graduate School" , University of Oulu, Faculty of Information Technology and Electrical Engineering, Department of Computer Science and Engineering; Infotech Oulu Acta Univ. Oul. C 487, 2014

[3] Somnath Roy, Jawaharlal Nehru University Centre for Linguistics Jnu, New Delhi-110067 Nishant Sinha IMS Engineering College CSE Department Ghaziabad UP, " Duration Modeling in Hindi " International Journal of Computer Applications (0975 – 8887) Volume 97– No.6, July 2014

[4] Leena Marya,*, B. Yegnanarayanab, " Extraction and representation of prosodic features for language and speaker recognition", Received 6 April 2006; received in revised form 20 February 2008; accepted 24 April 2008

[5] Chung, S., Hirose, K., and Minematsu, N., "Improvement of N-gram language modeling of Japanese using BUNSETSU boundary information and its application to large vocabulary continuous speech recognition," *Report for the Spring Meeting, ASJ*, 1, 65-66, 2004. (in Japanese)

[6] Hirose, Y., Ozeki, K., and Takagi, K., "Effectiveness of prosodic features in dependency analysis of read Japanese sentences," *Natural Language Processing*, 8 (4), 71-89, 2001.

[7] Inoue, A., Mikami, T., and Yamashita, Y., "Prediction of sentence importance for speech summarization using prosodic features," *Proc. Eurospeech*, *Geneva*, 1193-1196, 2003.http://www.speech-recognition.de/textbook.html (2003－05-18)

[8] Longbiao Wang, Norihide Kitaoka, Seiichi Nakagawa, "*Robust Distant Speaker Recognition Based on Position Dependent Cepstral Mean Normalization*", 2005

[9] "*Dynamic time warping*", 16 Dec 2006, http://en.wikipedia.org/wiki/Dynamic_time_warping

[10] Stuart N Wrigley, "*Speech Recognition by Dynamic Time Warping*", http://www.dcs.shef.ac.uk/~stu/com326/sym.html

[11] "*Formant*"http://en.wikipedia.org/wiki/Formant(2006-12-16)

[12] "*Signal Energy vs. Signal Power*"http://cnx.org/content/m10055/latest/(2004-08-12)

[13] Lawrence R. Rabiner, "*A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*" *Proceedings of the IEEE*, 77 (2), p. 257–286, (1989-02-09)

[14] Greg Cox, On the Relationship Between Entropy and Meaning in Music: An Exploration with Recurrent Neural Networks, Department of Psychological and Brain Sciences, Indiana University 1101 E. Tenth St., Bloomington, IN 47405 USA

[15] Emotion Recognition System Based On Skew Gaussian Mixture Model and MFCC Coefficient by M. Chinna Rao, A.V.S.N. Murthy, Ch. Satya Narayana ***I.J. Information Engineering and Electronic Business*, 2015, 4, 51-57** Published Online July 2015 in MECS (http://www.mecs-press.org)DOI: 10.5815/ijieeb.2015.04.07

Publish Research Article

Dear Sir/Mam,

We invite unpublished Research Paper,Summary of Research Project,Theses,Books and Book Review for publication.

Address:- Dr. Ashak Hussain Malik House No-221, Gangoo Pulwama - 192301
Jammu & Kashmir, India
Cell: 09086405302, 09906662570,
Ph No: 01933212815
Email:- nairjc5@gmail.com, nairjc@nairjc.com , info@nairjc.com
Website: www.nairjc.com