

A Hierarchical Stratagem for Classification of String Instrument

Arijit Ghosal, St. Thomas' College of Engineering & Technology, Kolkata, India

Suchibrotta Dutta, Royal Thimphu College, Thimphu, Bhutan

Debanjan Banerjee, Sarva Siksha Mission, Kolkata, India

ABSTRACT

Automatic recognition of instrument types from an audio signal is a challenging and a promising research topic. It is challenging as there has been work performed in this domain and because of its applications in the music industry. Different broad categories of instruments like strings, woodwinds, etc., have already been identified. Very few works have been done for the sub-categorization of different categories of instruments. Mel Frequency Cepstral Coefficients (MFCC) is a frequently used acoustic feature. In this work, a hierarchical scheme is proposed to classify string instruments without using MFCC-based features. Chroma reflects the strength of notes in a Western 12-note scale. Chroma-based features are able to differentiate from the different broad categories of string instruments in the first level. The identity of an instrument can be traced through the sound envelope produced by a note which bears a certain pitch. Pitch-based features have been considered to further sub-classify string instruments in the second level. To classify, a neural network, k-NN, Naïve Bayes' and Support Vector Machine have been used.

KEYWORDS

Chroma, Classification, Co-Occurrence Matrix, Instrument Classification, Note, Pitch

INTRODUCTION

In India several kinds of string type instruments are being employed since ancient times. This work takes care of complexity of classification of string instruments. To develop applications related to music and instrument classification, audio indexing, audio retrieval a good quality audio classification is essential. In recent past some advancement in the field of audio retrieval, audio classification is observed because of improvement of research in area of data mining as well as signal processing.

The oldest identified instrument classification system was identified as Chinese in the time period of the 4th century B.C. The system has discriminated different instrumental devices based on materials they are created. Aftermath, researchers are trying to propose fine instrument discrimination system. They have suggested several models but unfortunately none of those is accepted worldwide as a standard benchmark to fulfill the requirements of different applications. Hence, classification of instrument is an open research area. Researchers have successfully classified instruments into string, woodwind, percussion, keyboard, etc., but less work has been done for sub-classification of instruments. In this work, string instruments have been sub-classified without using Mel Frequency Cepstral Coefficients (MFCC) which is a first-rate aural feature but excessively used. A hierarchical approach is adopted here to classify string instruments. Initially string instruments are categorized into

DOI: 10.4018/IJWLTT.2020010101

This article, originally published under IGI Global's copyright on January 1, 2020 will proceed with publication as an Open Access article starting on January 28, 2021 in the gold Open Access journal, International Journal of Web-Based Learning and Teaching Technologies (converted to gold Open Access January 1, 2021), and will be distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

plucking, bowing and striking followed by further sub classification of plucking, bowing and striking type string instruments. The work is organized in the following way: previous works done related to classification of instruments is described in after introduction section. Proposed scheme is described in next section. The next section reflects investigational outcomes along with comparative analysis.

ASSOCIATED RESEARCH ACTIVITIES

Canvassers have paid efforts on several aural traits to discriminate instruments. Haralick (1992) has suggested a method for extraction of statistical features in the meadow of processing of images. These statistical features are very useful and they may be utilized in other domains too. Most of these works have used MFCC. A spectral trait for musical gadget categorization has been used by Agostini et al. (2003). Centroid bandwidth, pitch, skewness as well as zero crossing rates have been used by them. Peeters (2003) has proposed his own algorithm named IRMFSP for large database of musical database. Spectral features were also been explored by some researchers like Zhu et al. (2004). They have used spectrum of instruments. Their work was limited to jazz, pop and rock instrumentals only. Kaminskyj and Czaszejko (2005) have been able to classify mono type musical instrumental sounds with the help of 6 traits - cepstral coefficients, constant Q transform frequency spectrum, multidimensional scaling analysis trajectories, RMS amplitude envelope, spectral centroid and vibrato. Algorithms for automatically categorization of musical instrumental sounds have been proposed by Benetos, Kotti and Kotropoulos (2006). Hierarchical scheme was ventured by Essid et al. (2006). Support Vector Machine or SVM has been used there. They have used spectral features like MPEG-7 audio features, cepstral traits for example Mel Frequency Cepstral Coefficients or MFCC, temporal traits like autocorrelation coefficients and ZCR, wavelet traits, perceptual traits for example sharpness and loudness.

Hidden Markov Model (HMM) has been employed by Sinith and Rajeev (2007) while dealing with classical music of South India. Instruments are divided into wide-ranging types for example brass, string, percussion and woodwind by applying MPEG-7 audio features, perceptual features and Mel Frequency Cepstral Coefficients by Deng (2008). Gunasekaran and Revathy (2008) have worked with fusion of multiple classifiers. They have worked with temporal, perceptual, spectral, statistical and harmonic features. A 37-dimensional feature set based on MFCC as well as perception-based features has been used by Senan et al. (2009) while working with Malay musical instruments. Western and Chinese musical instruments were classified into 3 wide-ranging groups – percussion, wind as well as string using SVM based approach by Liu and Xie (2010). Kumari et al. (2010) have worked with musical instruments of North India for example: flute, dholak, sitar, mandar and bhapang. They have used a combination of MFCC and spectral features to design their feature set. Polyphonic instrumental signals were dealt with Barbedo and Tzanetakis (2010, 2011).

Wavelet and MFCC based hierarchical scheme were suggested by Ghosal et al. (2011) to categorize instrumental devices in wide-ranging groups like String, Woodwind, Percussion and Keyboard. But they have not explored further sub-classification of instruments. Grindlay and Ellis (2011) have dealt with music having polyphonic in nature. Müller and Ewert (2011) have employed Chroma Toolbox of Matlab. Impact of selection of traits and classifiers on classification accuracy has been observed by Chandwadkar and Sutaone (2012) aiming to recognition of musical instrumental devices. Chandwadkar and Sutaone have applied an amalgamation of spectral traits and Mel Frequency Cepstral Coefficients in their feature set. Nadgir and Joshi (2014) have worked with spectral, temporal and MFCC based features. Gaikwad et al. (2014) has applied Principal Component Analysis (PCA) on the combination of spectral traits and cepstrum features to classify Indian classical instruments. To design spectral traits the authors have used spectral range and amplitude together with MFCC. Probabilistic Latent Component Analysis (PLCA) has been used by Arora and Behera (2014) to recognize instruments. Dandawate et al. (2015) have classified Indian instrumental music-based Raga like Bhairavi, Bhairav, Yamanand Todi Raga. They designed their feature set using temporal features. Abeber and Weib

(2015) have recognized instruments using polyphonic type classical music. Abeber and Weib have divided instrumental devices into woodwinds, brass, piano, strings, in addition to speaking classes. Musical onset detection strategy was adopted by Kumar et al. (2015) to classify Carnatic percussion instruments into mridangam, ghatam, kanjira, thavil and morsing. A combination of MFCC based features and counter propagation neural network (CPNN) has been applied by Bhalke et al. (2016). By applying schemes applicable for processing talking signal, musical instruments are classified by Ghisingh and Mittal (2016). Spectral Centroid or SC, MFCC, signal energy and ZCR are used in their feature vector.

PROPOSED METHODOLOGY

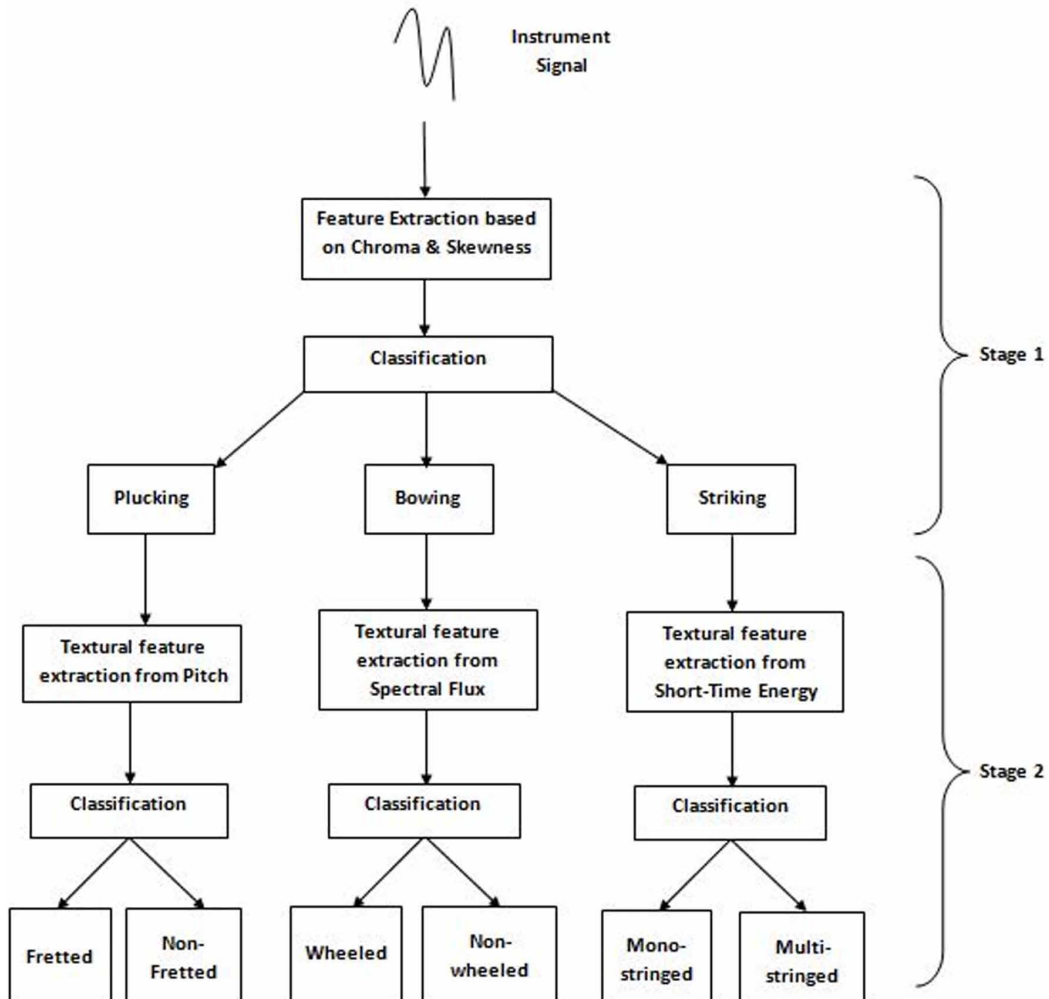
Precedent works tell that MFCC has been frequently used by different researchers in their works related with discrimination of different string type instrumental devices. MFCC is a brilliant acoustic trait but due to too much usage of it, this outstanding acoustic feature has lost its novelty. Moreover, previous research works reveal that most of the works was aimed to differentiate instrumental devices in different wide-ranging types for example woodwind, string, percussion etc. Comparatively very less work has been done to classify instruments up to different sub-categories. Inspiring from these two facts in this work a hierarchical scheme is proposed to classify instrumental devices in diverse sub-groups without using MFCC. Trait set used in this work is proposed as alternative of MFCC. Experimental result exhibits that a better classification result is possible by using proposed feature set compared to that of MFCC.

String instruments produces sounds through the vibration made from its strings or chords. String instrumental devices may produce sounds in different methods. Depending on the way string instrumental devices produce sounds, they are divided into 3 main categories – plucking, bowing and striking. Again, based on the existence of fret, plucking type string instruments are further classified into two sub-categories – fretted and non-fretted. In plucking type string instruments strings or chords are plucked by using finger, thumb, or quills (now plastic plectra). Sitar, veena, and guitar falls into this category of string instruments. In case of bowing type string instruments, they are played with the help of a bow. Depending on how bow is played bowing type string instruments are further classified into two sub-categories – wheeled and non-wheeled. Dulcigurdy and the Hurdy-gurdy is example of wheeled bowing type string instruments and the cello and violin is example of non-wheeled bowing type string instruments. For wheeled bowing type of string instrument, the instrument produces sound by means of a turning wheel which acts as bow over the strings. When a string instrument is played by striking on strings they fall into the category of striking type string instruments. Santoor, hammered dulcimers are examples of striking type string instruments. Depending on number of strings, striking type string instruments can be further classified into two sub-categories – mono-stringed and multi-stringed. Berimbau is a striking type string instrument which has single string whereas Santoor has multiple strings.

This is noticed that bowing, plucking and striking forms of string instrumental devices are generally performed with diverse body stances in addition to notes produced by them are also of dissimilar kind. But the presence of fret, presence of wheel in bow and number of strings produces a difference in the sound quality of the string instrument. This observation has motivated to adopt a hierarchical scheme to classify string instruments into 3 groups in the first stage – plucking, bowing and striking based on the notes produced by these string instruments and then each type string instrument is further classified in the second stage.

Next sub-divisions describe trait mining and categorization techniques used respectively. Fig.1 describes the hierarchical arrangement of the proposed concept.

Figure 1. Hierarchical outline of suggested method



Audio Attribute Computation

This is recognized that to achieve high classification accuracy the feature set need to be designed judiciously. Also, at the same time, the length of the trait vector should not be so big that it will increase computation complexity.

Features for Bowing, Plucking and Striking String Instrument Classification: As this is known that bowing, plucking, and striking type string instrumental devices mainly differs from each other for the notes they generate. Their notes vary because of existence of diverse number of cords in those string type instrumental devices. Note that the small bit of sound is similar to concept of syllable in case of a language. Note that it can also be treated as a small lone musical incident. Notes produced from these three types of string instruments vary widely as the musical events are not same for them. For creation of music, notes are required to be played in orderly fashion. A note in case of string instruments is nothing but a sound wave generated due to vibration of strings having a specific wavelength. Notes are thought of as building blocks of instruments. This surveillance has instigated to suggest a trait set which should be able to represent nature of notes for different types of string instrumental devices like bowing, plucking and striking.

Chroma-Based Features

Chroma based features can reflect 12 discrete semitones or Chroma of the melodic octave. Chroma dependent traits are so powerful acoustic feature that it can be used to study the characteristics of instrumental music. Chroma features can reflect melodic and harmonic nature of instrument. From hearing perception, it can be felt that sounds produced by bowing, plucking and striking forms of string instrumental devices vary melodically and harmonically. Because of these Chroma traits are the most excellent traits to classify string instruments into different categories. Chroma values can be considered by the group

{C, C#, D, D#, E, F, F#, G, G#, A, A#, B}

Chroma assessments may be acknowledged by a group of numerals (1, 2, 3, ..., 12), where 1 indicates Chroma C, 2 indicates C# and so on. Chroma ring is shown in Figure 2. This circle is such a next of kin that it is impossible to be denoted by only Mel. Here lies the benefit of Chroma based traits compared to Mel based traits like MFCC.

This circle has been quantized into 12 positions for analysis of western tonal music. Chroma features indicate the intensity of each of these 12 positions. Chroma traits reflect perceptual dissimilarities inside an octave. Signals of instrumental devices are dynamic by nature. For better computation of Chroma features input auditory information is broken down with 50% partly covered frames. To keep away from losing of border line uniqueness of any envelope, the frames are overlapped. Chroma features are calculated for each of the frames. After that mean of all 12 positional values for all the frames are considered. This way Chroma depended traits contribute 12 discrete numerical values.

Figure 3, Figure 4, and Figure 5 correspondingly indicate the representations of Chroma values along with degree for bowing, plucking and striking-type string instruments.

From these plots it may be noticed that Chroma features vary widely for bowing, plucking, and striking kinds of string instrumental devices. To extract Chroma features the chromagram toolbox (2011) has been applied.

Skewness Based Feature

Spectral skewness reflects the third order moment of the spectral circulation. Undoubtedly it is a numerical trait. Skewness basically measures the unevenness existing in a certain normal circulation concerning to its mean location. Bowing, Plucking and Striking kinds of string instrumental devices carry certain quantity of unevenness in their respective normal circulation. Value of skewness sn is characterized by the underneath Equation (1).

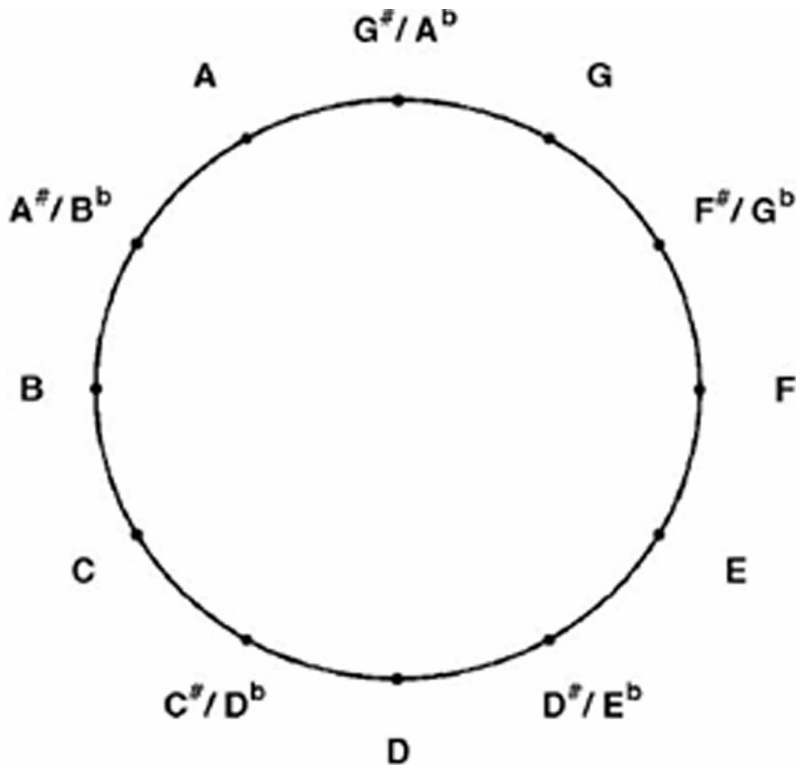
$$sn = \frac{E(b - \mu)^3}{\sigma^3} \quad (1)$$

In Equation (1) μ indicates mean of exemplified information b , σ indicates standard deviation of b , and the expected value of a quantity y is symbolized by $E(y)$.

Features for Bowing, Plucking and Striking String Instrument Classification

A) Features for Fretted and Non-Fretted Plucking String Instrument Classification: In the second stage of the proposed hierarchical scheme, plucking type string instruments are further classified into two sub-categories: Fretted plucking string instruments and non-fretted plucking string instruments. Fret is nothing but a raised element presents on the neck region of a string instrument. It is observed that sound quality of Fretted Plucking string instruments differs from non-fretted plucking string instruments. A fret generally represents one semitone and musical note or tones which exhibit a

Figure 2. Chroma circle

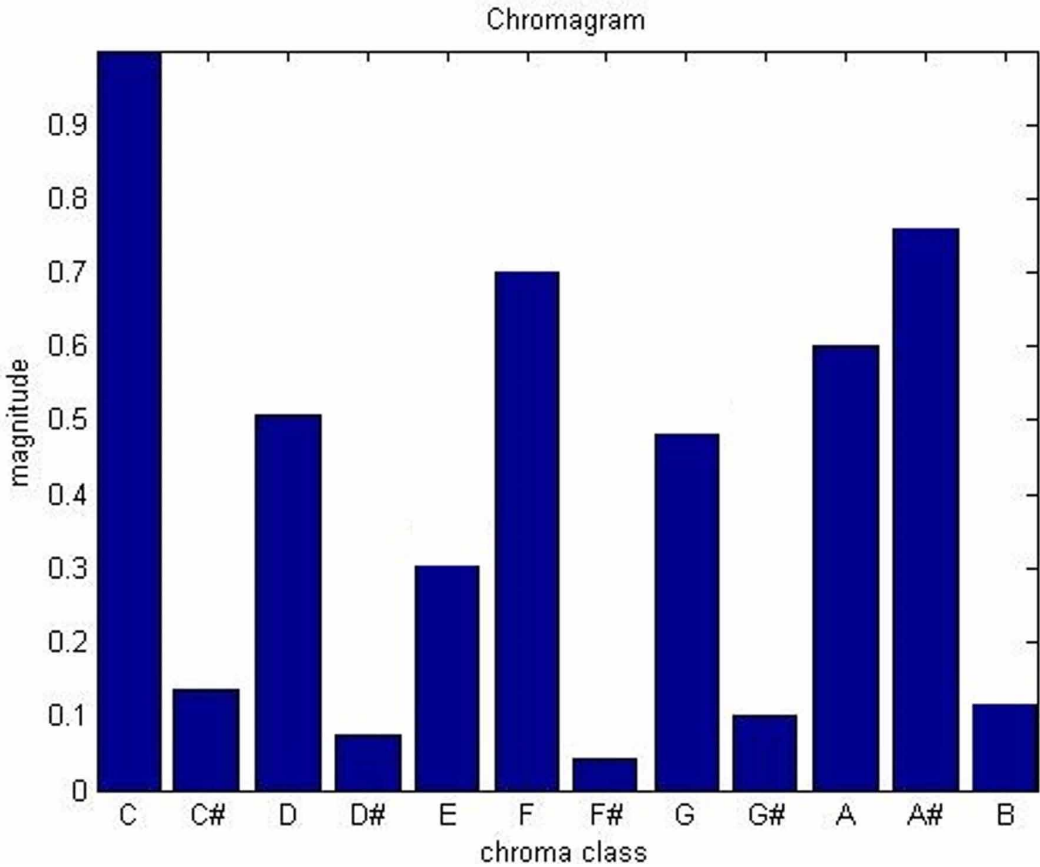


pitch. Frets in an instrument restrict pitches to the temperament defined by the fret positions. As a result, ranges of pitch for fretted and non-fretted plucking string instruments are noticeably different. This observation has motivated to employ pitch-based features for differentiating fretted plucking string instrument and non-fretted plucking string instrument. Previous research study reveals that some work has already been done using pitch-based features in this area. But they have used all the 88 pitch values in the feature vector. This results in a large dimensional feature vector. To decrease the dimension of trait vector principal component analysis or PCA is required to be applied. Use of PCA increases both computational complexity and classification time.

To overcome these drawbacks as well as to minutely study the characteristics of pitch, conception of co-occurrence matrix, which is very popular in the domain of image processing, has been applied. Co-occurrence matrix assists to observe the occurrence pattern of a certain trait. It is known that pitch is a recognized form of frequency. Initially input signal is divided into 88 frequency bands having center frequency equivalent to A0 to C8 pitches. These reflect successive MIDI pitches 21 to 108. For every sub-band, the mean of short-time mean-square power (STMSP) is calculated. This mean value indicates the power localized in that corresponding frequency sub-band. Thinking about all these sub-bands, 88-dimensional pitch trait vector $\{P_i\}$ is generated for the corresponding plucking string instrument signal.

Short-time mean-square power (STMSP) allocation over diverse frequency sub-bands is depicted in Figure 6 and Figure 7 respectively corresponding to fretted plucking string instrument and non-fretted plucking string instrument. STMSP for a band diverges with time. The mean values of STMSP corresponding to each band are considered. It is apparent that the deviation of strength of signal over different frequency bands differs for Fretted Plucking string instrument and Non-fretted Plucking string instrument.

Figure 3. Representation of Chroma values beside magnitude for Bowing kind string instrumental devices

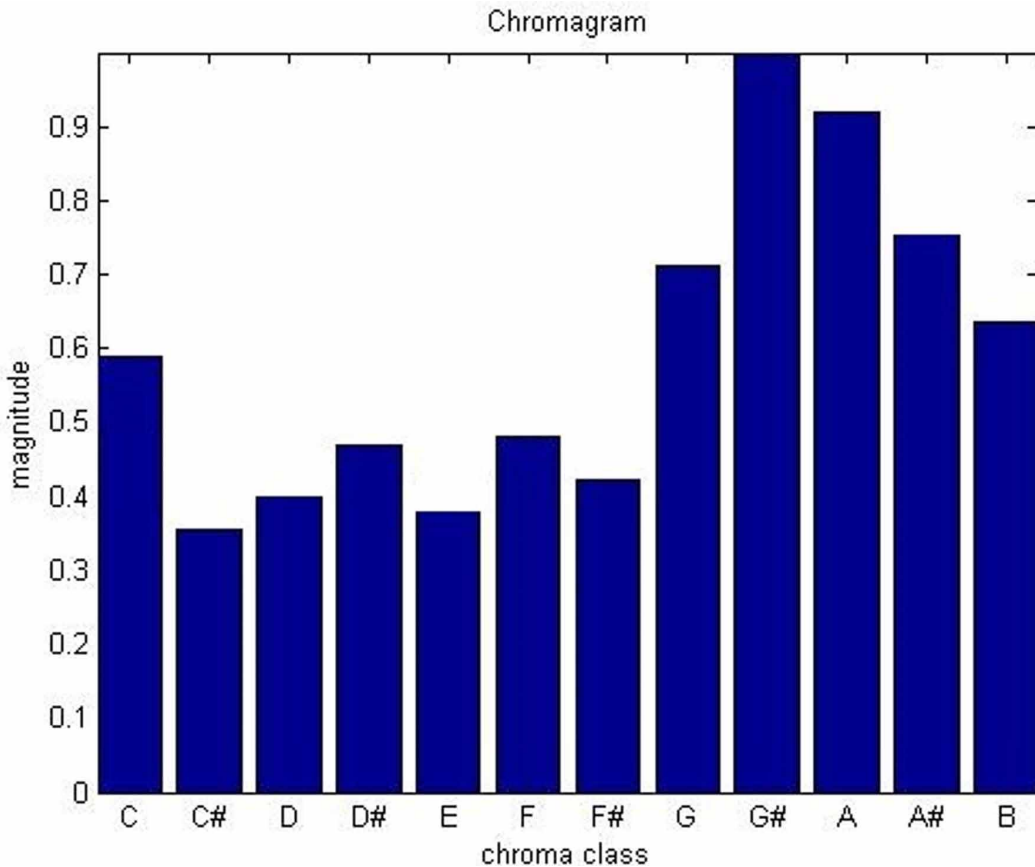


Positioning of different pitch values enclosed by a region reflects the characteristics of corresponding plucking string instrument category. Thus, a matrix C_n having dimension $A \times A$ (where, $A = \max\{P_i\} + 1$) is formed. A component in the matrix $C_n(x, y)$ specifies the number of incidences of pitch x and y in ensuing time occasions. From this co-occurrence matrix C_n 14 statistical features (Haralick & Shapiro, 1992) like angular second moment, variance, entropy, energy, correlation, contrast, homogeneity, etc., are computed. This 14-dimensional statistical feature set is of low-dimensional and at the same time it can represent the repetitive characteristics of pitch very well.

Co-occurrence matrix plots of pitch values corresponding to Fretted Plucking string instrument and Non-fretted Plucking string instrument are portrayed in Figure 8 and Figure 9 correspondingly. The plots clearly explain that occasion nature of pitch values is different for Fretted Plucking string instrument and Non-fretted Plucking string instrument.

B) Features for Wheeled and Non- Wheeled Bowing String Instrument Classification: In the second stage of the proposed hierarchical scheme, bowing type string instruments are also further classified into two sub-categories: Wheeled Bowing string instrument and Non- wheeled Bowing string instrument. Sounds can be generated from Bowing type string instrumental devices in two manners – either by rubbing a stick over the strings or by bowing by spinning wheel. This difference of playing bow results in a basic difference in their respective spectrum. It is known that changing of power spectrum in audio signal can be captured well by spectral flux. Spectral flux measures the disparity in spectrum amid two successive frames. Motivated by this observation spectral flux depended traits have been applied to sub-classify bowing type string instruments. Spectral flux can be calculated as:

Figure 4. Representation of Chroma values beside magnitude for Plucking kind string instrumental devices

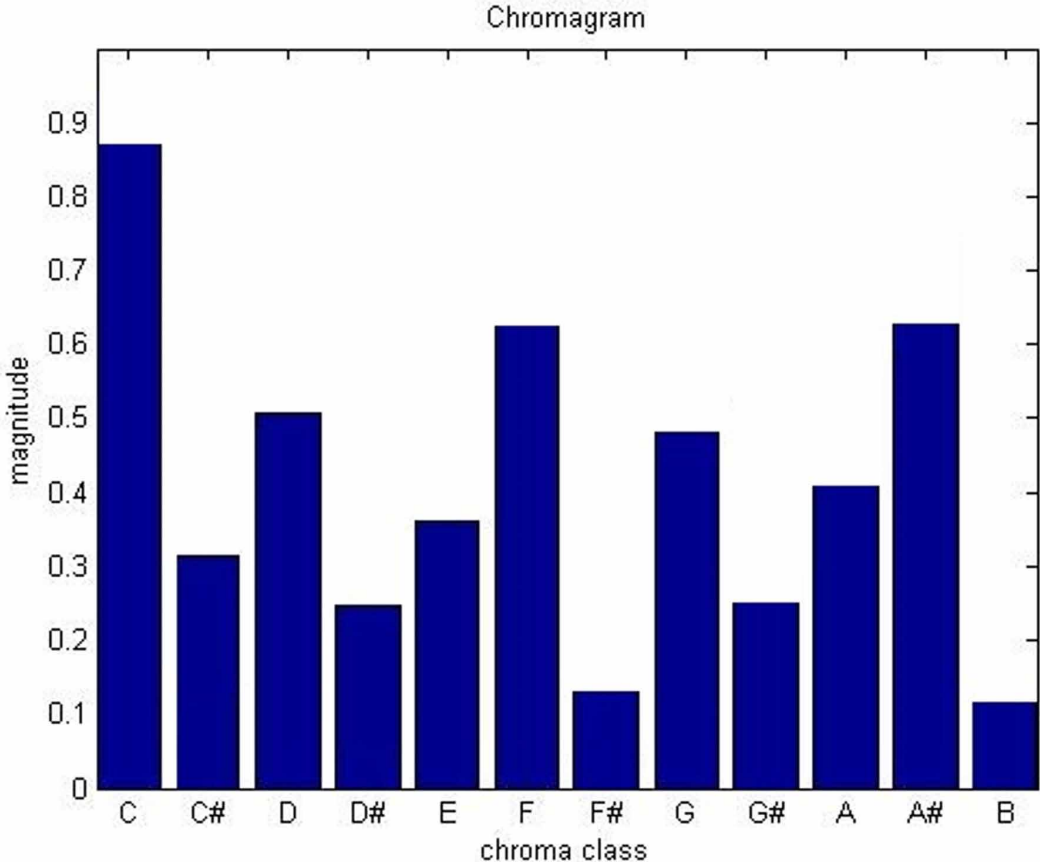


$$SFL(Y) = \sum_{a=0}^{n-1} SL(Y, a) - SL(Y-1, a) \quad (2)$$

where, n represents whole number of frames. Spectral flux intended for the Y^{th} spectrum is denoted by $SFL(Y)$. Suppose, the amount of a^{th} bin in Y^{th} spectrum is $SL(Y, a)$, then $SL(Y-1, a)$ is same for spectrum earlier than Y. The amounts of every bin of preceding spectrum are deducted from the amounts of its equaling bin in current spectrum and after that adding up those dissimilarities to generate a concluding value which is the aspired ultimate spectral flux for spectrum Y. For minute study of characteristics of spectral flux a co-occurrence matrix of it, CO_n having measurement $Ba \times Ba$ (where, Ba represents $\max\{SF_i\} + 1$) is formed. A component in the matrix $CO_n(m, p)$ points to the amount of incidences of spectral flux m and p in consecutive moment occurrences. From this co-occurrence matrix CO_n 14 statistical features (Haralick & Shapiro, 1992) like angular second moment, variance, entropy, energy, correlation, contrast, homogeneity, etc., are computed. This 14-dimensional statistical feature set is of low-dimensional and at the same time it can represent the repetitive characteristics of spectral flux very well.

Co-occurring matrix plots of spectral flux values corresponding to Wheeled Bowing string instrument and Non-wheeled Bowing string instrument are portrayed in Figure 10 and Figure 11 correspondingly. Plots clearly point to that occasion nature of spectral flux is different for Wheeled Bowing string instrument and Non-Wheeled Bowing string instrument.

Figure 5. Representation of Chroma values beside magnitude for Striking kind string instrumental devices



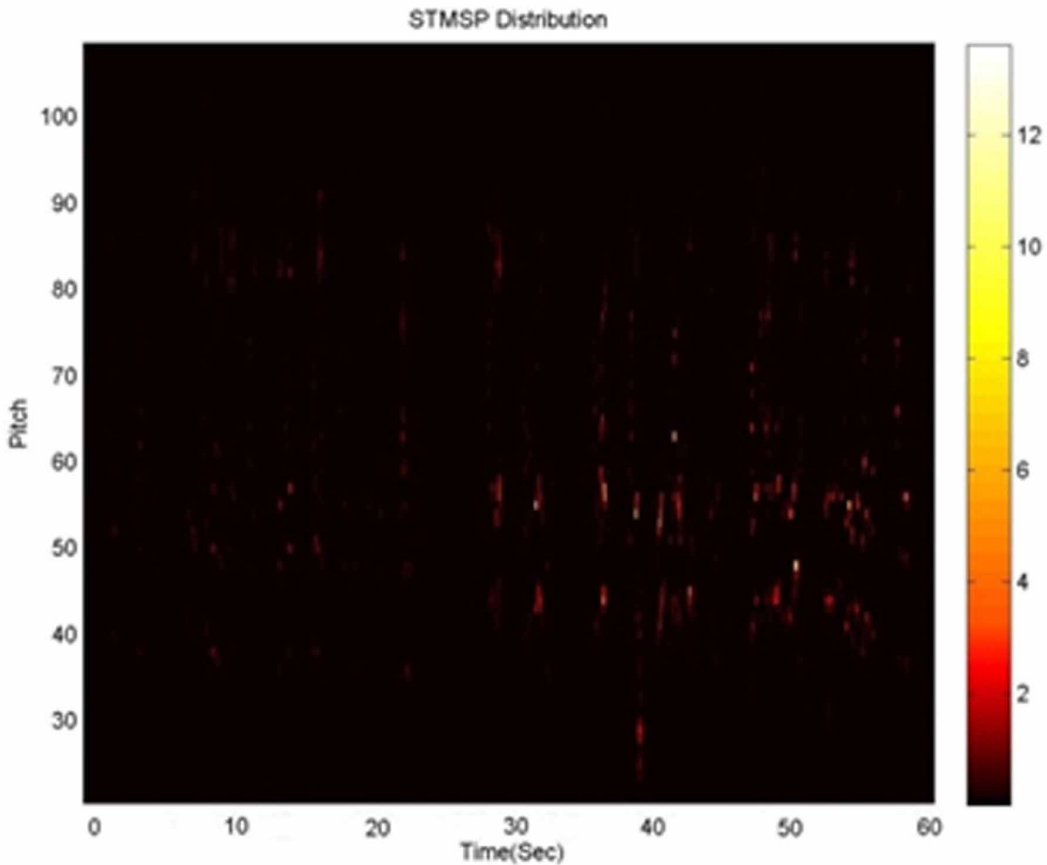
C) Features for Mono-Stringed and Multi-Stringed Striking String Instrument Classification:

In the second stage of the proposed hierarchical scheme, striking type string instruments are also further classified into two sub-categories: Mono-stringed Striking string instrument and Multi-stringed Striking string instrument. Striking type string instruments are of two types – mono-stringed and multi-stringed. Berimbau is a striking type string instrument which has only a single string and on the other side Santoor is a striking type string instrument having multiple numbers of strings. Due to presence of only a string, mono-stringed striking string instrument produces mono sound and multi-stringed striking string instrument produces stereo sound. Due to this, sound produced by these two different types of striking type string instruments exhibits different levels of energy. This difference of energy level can be well captured by short time energy (STE). STE can be calculated as:

$$En_n = \frac{1}{r} \sum_{m=1}^r [x_n(m)]^2 \tag{3}$$

where, En_n denotes the n^{th} frame energy, r is the frame length and $x_n(m)$ represents the m^{th} sample in the n^{th} frame. For minute study of characteristics of STE a co-occurrence matrix of it, CO_STE_n having dimension $T \times T$ (where, $T = \max\{En_n\} + 1$) is formed. A component in the matrix $CO_STE_n(g,h)$ points to the count of incidences of STE g and h in consecutive occasion cases. From

Figure 6. Plot of STMSP allocation over diverse sub-bands for fretted plucking type string instrument



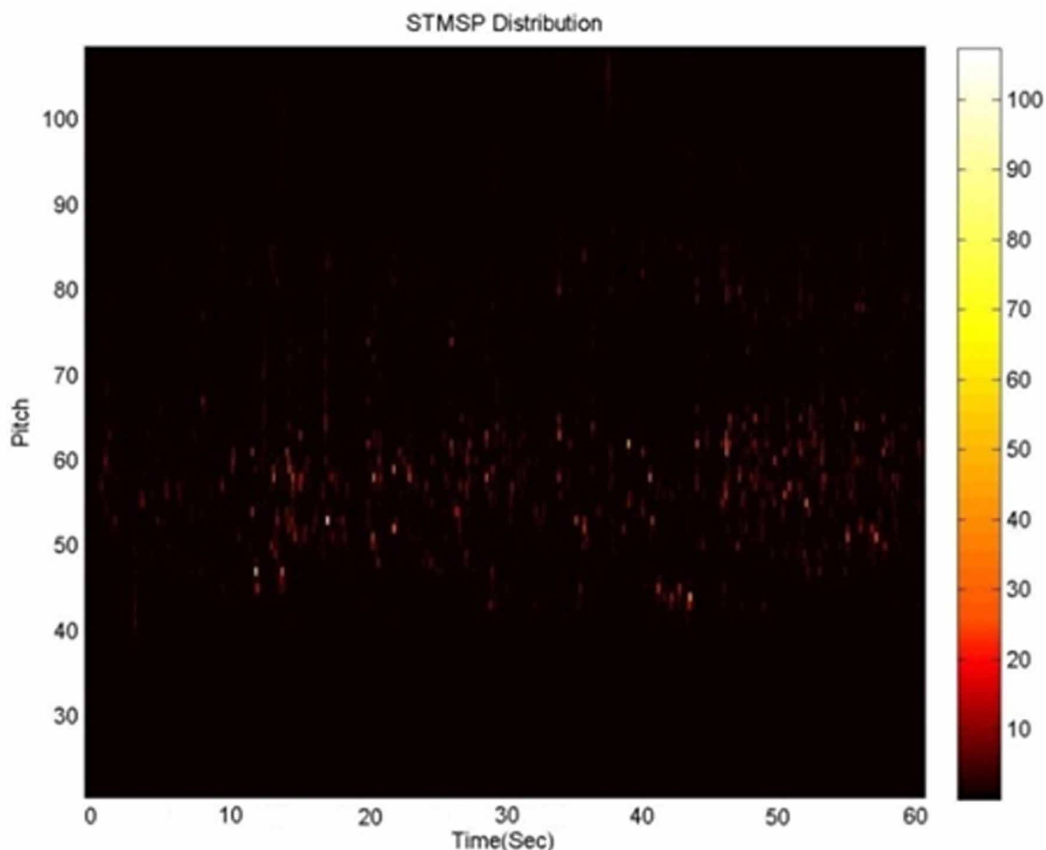
this co-occurrence matrix CO_STE_n 14 statistical features (Haralick & Shapiro, 1992) like angular second moment, variance, entropy, energy, correlation, contrast, homogeneity, etc., are computed. This 14-dimensional statistical feature set is of low-dimensional and at the same time it can represent the repetitive characteristics of STE very well.

Co-occurrence matrix plots of STE values corresponding to mono-stringed striking string instrument and multi-stringed striking string instrument are portrayed in Figure 12 and Figure 13 correspondingly. From these plots this is understandable that occasion nature of STE is different for mono-stringed striking string instrument and multi-stringed striking string instrument.

Classification

Most important intention of this effort is to propose small dimensional acoustic feature set which can discriminate string instruments into different categories and sub-categories hierarchically very well. In the first stage of the scheme string instruments are classified into plucking, bowing and striking types of string instrument based on 13 dimensional Chroma and skewness-based features. In the second stage statistical features based on pitch, spectral flux and short time energy have been used to further sub-classify bowing, plucking and striking type string instrumental devices. For emphasizing the potency of the projected acoustic trait group standard classifiers for example k-NN, neural network as well as Naïve Bayes have been used in the first stage and Support Vector Machine or SVM, Naïve Bayes in addition to a neural network has been used in the second stage of the hierarchical scheme.

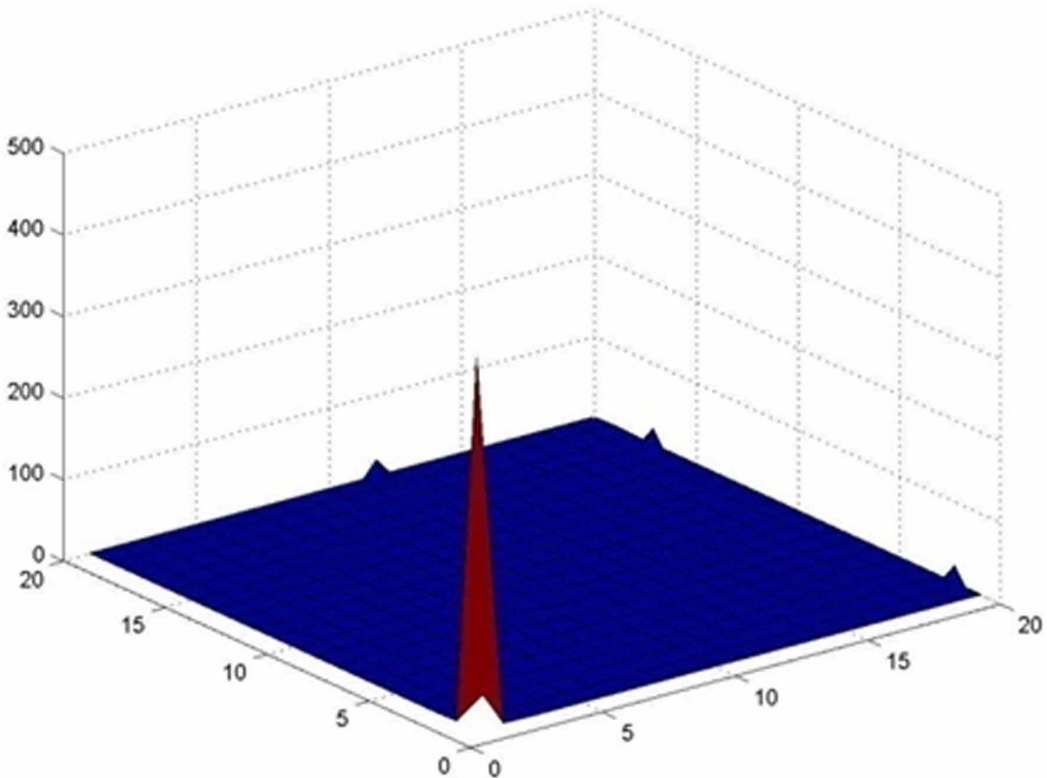
Figure 7. Plot of STMSP allocation over diverse sub-bands for non-fretted plucking type string instrument



To implement neural network multi-layer perceptron (MLP) has been applied in both the stages. There are very few standard audio dataset available for instrumental sounds and unfortunately either this dataset is paid or does not contain wide variety of instrumental sounds applicable in this work. So, an audio dataset comprising of 3600 audio files have been formed by collecting audio files from assorted resources for example Internet, audio CD, live shows, etc. Each type of string instruments bears equal division in the custom audio dataset. In the first stage the whole dataset was divided into equal parts – training and testing data each consisting of 1800 audio files representing 600 files each for plucking, bowing and striking. In the second stage, 14-dimensional statistical feature set extracted from the co-occurrence matrix of pitch, spectral flux and STE has been used with a dataset of 1200 audio files each for plucking, bowing and striking.

To implement multi-layer perceptron or MLP type neural network or NN in first stage 13 neurons has been put in the input layer indicating 13 features values – 12 values based on Chroma and 1 skewness based feature, 3 neurons has been put in the output layer indicating three categories of audio files – plucking, bowing and striking, 7 neurons has been put in the only hidden layer. Consequently, in the second stage 14 neurons are put in the input layer indicating 14 statistical traits intended from the co-occurrence matrix of pitch, spectral flux and STE respectively for sub-classifying bowing, plucking as well as striking type string instrumental devices. 2 neurons be there in the output layer reflecting 2 sub-categories of bowing, plucking as well as striking type string instrumental devices. 5 neurons be there in the single concealed stratum. k-NN classifier has been applied only in the first

Figure 8. Representation of co-occurrence matrix of pitch for Fretted Plucking type string instrument

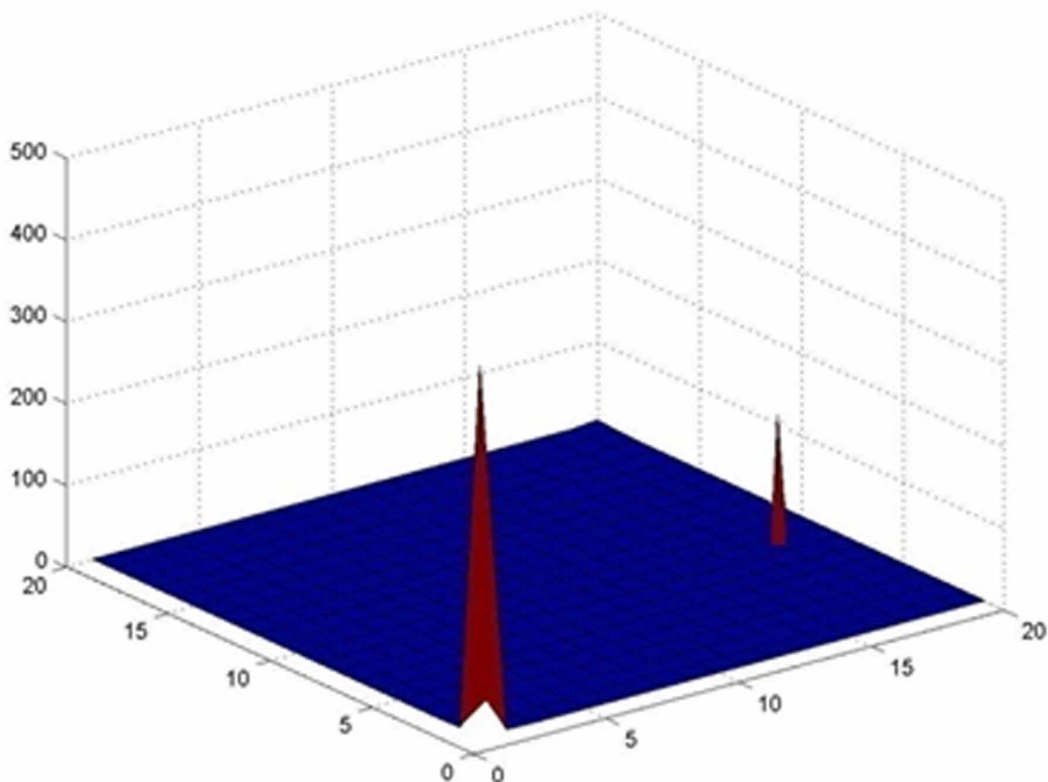


stage with the following configuration: $k=3$, distance metric as “City Block” and “Nearest Neighbour” law for breaking a tie. Naïve Bayes’ classifier has been used in both stages. Support Vector Machine or SVM has been configured in the second stage with kernel type as “Quadratic”.

INVESTIGATIONAL OUTCOMES

An audio dataset has been adopted in this hierarchical scheme consisting of 3600 audio cases having identical sharing for every type of class. Bowing, plucking as well as striking string instrumental devices have equal contribution inside this dataset that is there are 1200 audio files for each of these categories. In the first stage all the 3600 audio files are used and in the second stage 1200 audio files indicating all types of plucking, bowing and striking string instruments have been used while sub-classifying them. There are 600 audio files corresponding to fretted and non-fretted plucking string instruments. Similarly, there are 600 audio files also corresponding to wheeled and non-wheeled bowing string instruments and there are 600 audio files corresponding to mono-stringed and multi-stringed striking string instruments. All of these audio files have duration of 90 seconds with sampling frequency 22050 Hz and all of them are of 16-bit per sample. The custom dataset has been built by collecting audio files from a variety of foundations for example Internet, audio CD, live shows etc. Intentionally to reflect proper real-life scenario, some of these audio files are considered as noisy also. Among the dataset applicable for a certain stage in this proposed scheme, 50% is utilized for guidance intention with the remaining 50% has been used intended for test intention. For both stages same convention is followed. Again, after performing experiment, guidance and test information group is reversed and the same experiment is repeated with same feature set along with

Figure 9. Representation of co-occurrence matrix of pitch for Non-fretted Plucking type string instrument



same configuration of classifier. Average of these two steps is considered as final classification result and the same is tabulated in Table 1 for bowing, plucking as well as striking type string instrumental devices classification (stage 1). Experimental result for sub-classifying bowing, plucking as well as striking type string instrumental devices (stage 2) is tabulated in Table 2 and Table 3 in addition to Table 4, respectively.

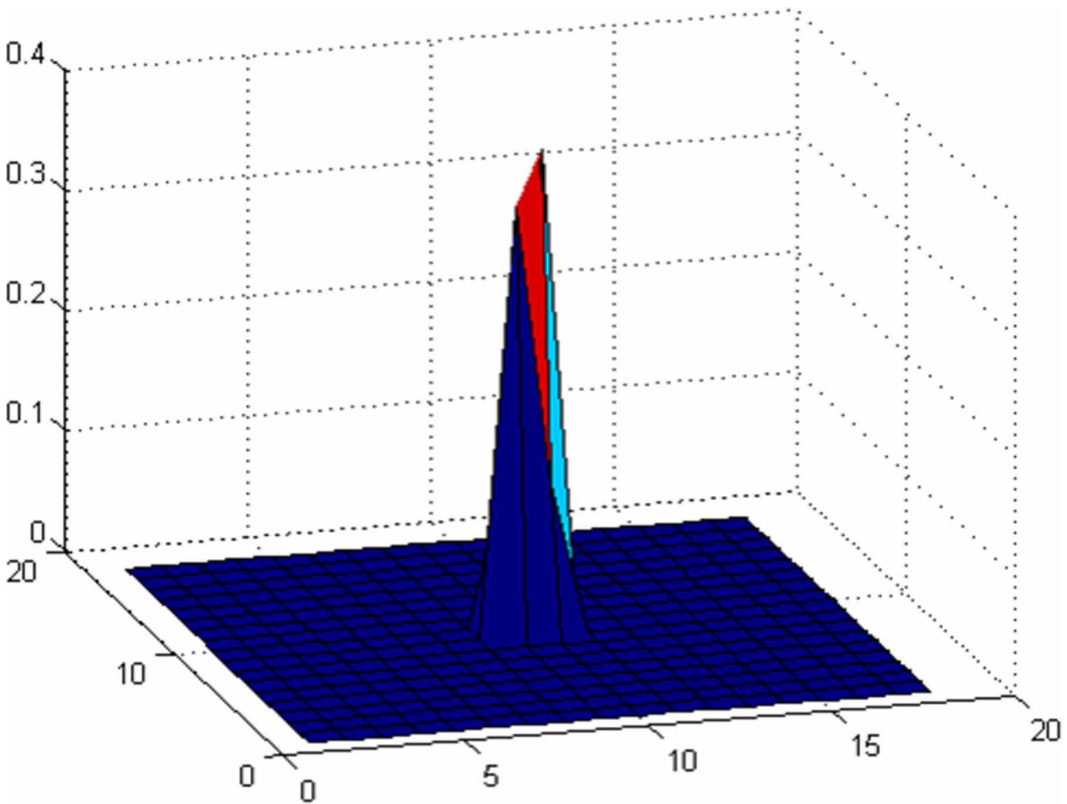
PRESENTATION COMPARISON USING AUC

Area Under Curve (AUC) is utilized also as a performance measurement criterion instead of accuracy percentage. From Figure 14, Figure 15, and Figure 16 it is clear that ROC curve for Naïve Bayes exhibits the best result among Naïve Bayes, neural network (MLP) and k-NN classifiers.

Comparative Analysis

Feature sets for stage 1 and stage 2 are proposed as an alternative of MFCC. So, classification performances of both the feature sets are compared with MFCC. The same dataset is used to find the comparative classification performance with same configuration of classifiers of respective stages. Instead of using proposed feature sets MFCC has been used to find the comparative analysis of classification performance.

Figure 10. Spectral flux co-occurrence matrix plot for wheeled bowing type string instrument



Comparative Analysis for Plucking, Bowing and Striking String Instrument Classification

The classification accuracy of Chroma and skewness-based feature set in stage 1 for classifying string instruments into plucking, bowing and striking is compared with MFCC. The classification accuracy using MFCC as feature is tabulated in Table 5.

From Table 5 it is very clear that proposed feature set can classify string instrumental devices in 3 wide-ranging classes – plucking, bowing as well as striking better than MFCC. As most of the previous approaches have worked with MFCC, performance of proposed feature set is compared with MFCC as feature along with some previous approaches.

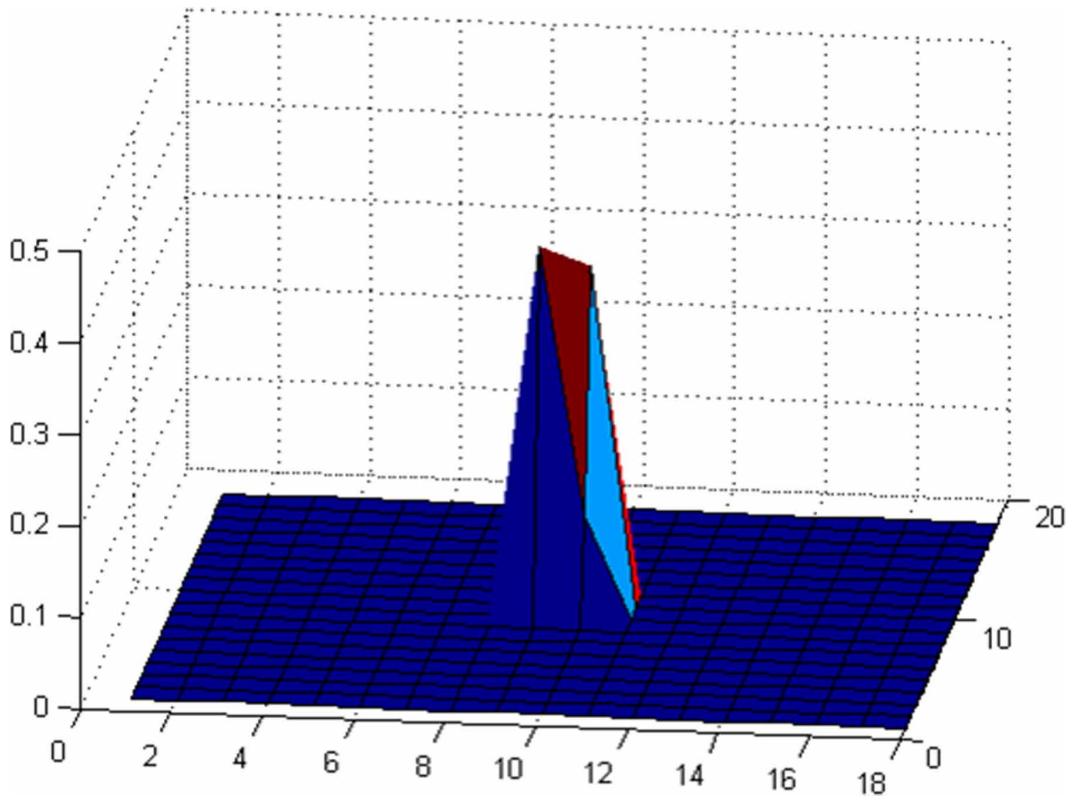
Classification performance of proposed Chroma and skewness-based feature set has been compared with some other previously work done. The information set used here, has been employed to implement the arrangement projected by Kumar et al. (2015), Bhalke et al. (2016) in addition to Ghisingh and Mittal (2016).

From Table 6 this is noticed that the projected trait group can classify string type instrumental devices into Plucking, Bowing and Striking better than other previous approaches.

Comparative Analysis for Fretted and Non-Fretted Plucking String Instrument Classification

Similarly, performance of statistical feature-based feature set in stage 2 is compared with MFCC. The same dataset used in stage 2 in this work has been used with same configuration of classifiers. The

Figure 11. Spectral flux co-occurrence matrix plot for non-wheeled bowing type string instrument



classification accuracy for classifying plucking string instruments further into fretted and non-fretted sub-category by using MFCC is tabulated in Table 7.

From Table 7 it is clear that proposed statistical trait group carries out enhanced categorization than MFCC. As MFCC is excessively in previous audio related works, this comparative analysis is done. Though it is observed that very few works have been done related with sub-classification of instruments still to check the discriminating power of the trait set proposed in stage 2 for classifying fretted and non-fretted plucking string instrument has been compared with the work suggested by Peeters (2003). He has used his own algorithm named as IRMFSP for classifying musical instruments sounds. In his work he has followed a hierarchical approach where string instruments were placed in one level of hierarchy.

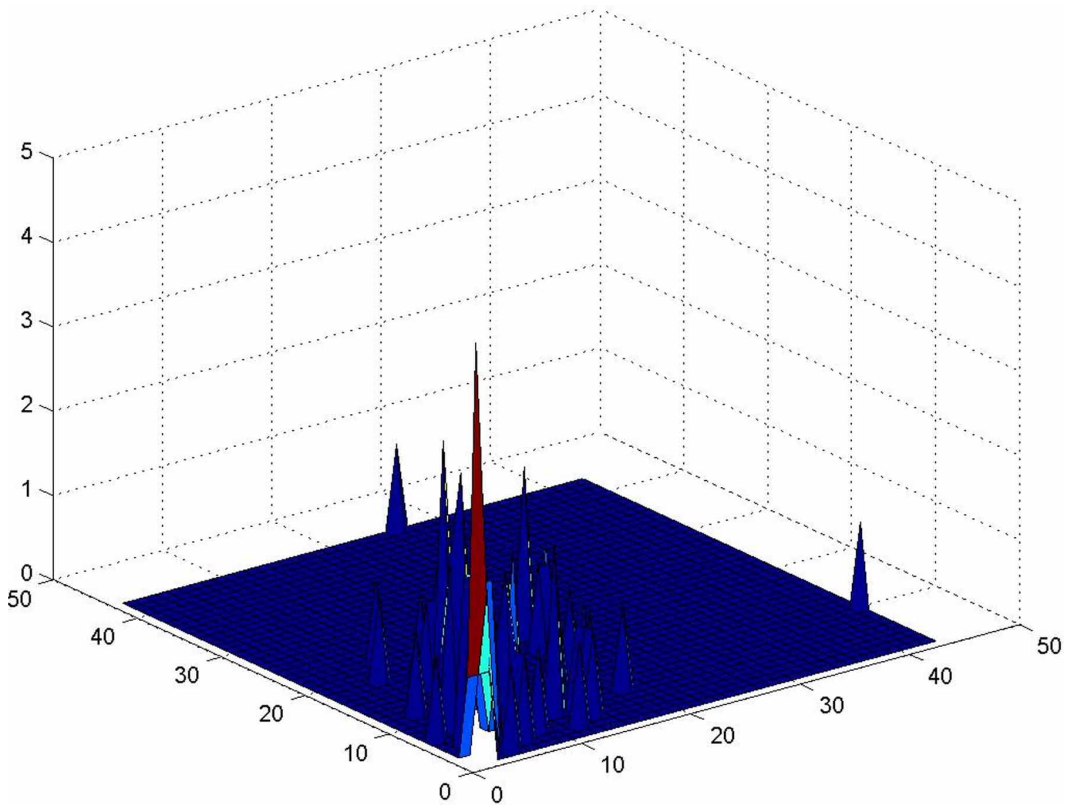
From Table 8 this can be observed that trait set used in stage 2 for classifying fretted and non-fretted plucking string instrument can discriminate plucking type string instruments better than the other approach.

Comparative Analysis for Wheeled and Non-Wheeled Bowing String Instrument Classification

Similarly, performance of statistical feature-based feature set for classifying bowing string instrument in stage 2 is compared with MFCC. The same dataset used in this work has been used with same configuration of classifiers. The classification accuracy for classifying bowing string instruments further into wheeled and non-wheeled sub-category by using MFCC is tabulated in Table 9.

From Table 9 it is clear that proposed statistical trait group executes healthier than MFCC. As MFCC is excessively in previous audio related works, this comparative analysis is done. Though it

Figure 12. STE co-occurrence matrix plot for mono-stringed striking type string instrument



is observed that very few works have been done related with sub-classification of instruments still to check the discriminating power of the trait set proposed in stage 2 for classifying wheeled and non-wheeled bowing string instrument has been compared with the work suggested by Peeters.

From Table 10 it can be observed that trait set used for classifying bowing string instrument in stage 2 can discriminate bowing type string instruments better than the other approach.

Comparative Analysis for Mono-Stringed and Multi-Stringed Striking String Instrument Classification

Similarly, performance of statistical feature-based feature set for classifying striking string instrument in stage 2 is compared with MFCC. The same dataset used in this work has been used with same configuration of classifiers. The classification accuracy for classifying striking type string instruments further into mono-stringed and multi-stringed sub-category by using MFCC is tabulated in Table 11.

From Table 11 it is clear that proposed statistical trait group carries out well against MFCC. As MFCC is excessively in previous audio related works, this comparative analysis is done. Though it is observed that very few works have been done related with sub-classification of instruments, still to check the discriminating power of the trait group proposed in stage 2 for classifying striking type string instrument, has been compared with the work suggested by Peeters (2003).

From Table 12 it can be perceived that trait set used in stage 2 for classifying striking type string instrument can discriminate striking type string instruments better than the other approach.

Figure 13. STE co-occurrence matrix plot for multi-stringed striking type string instrument

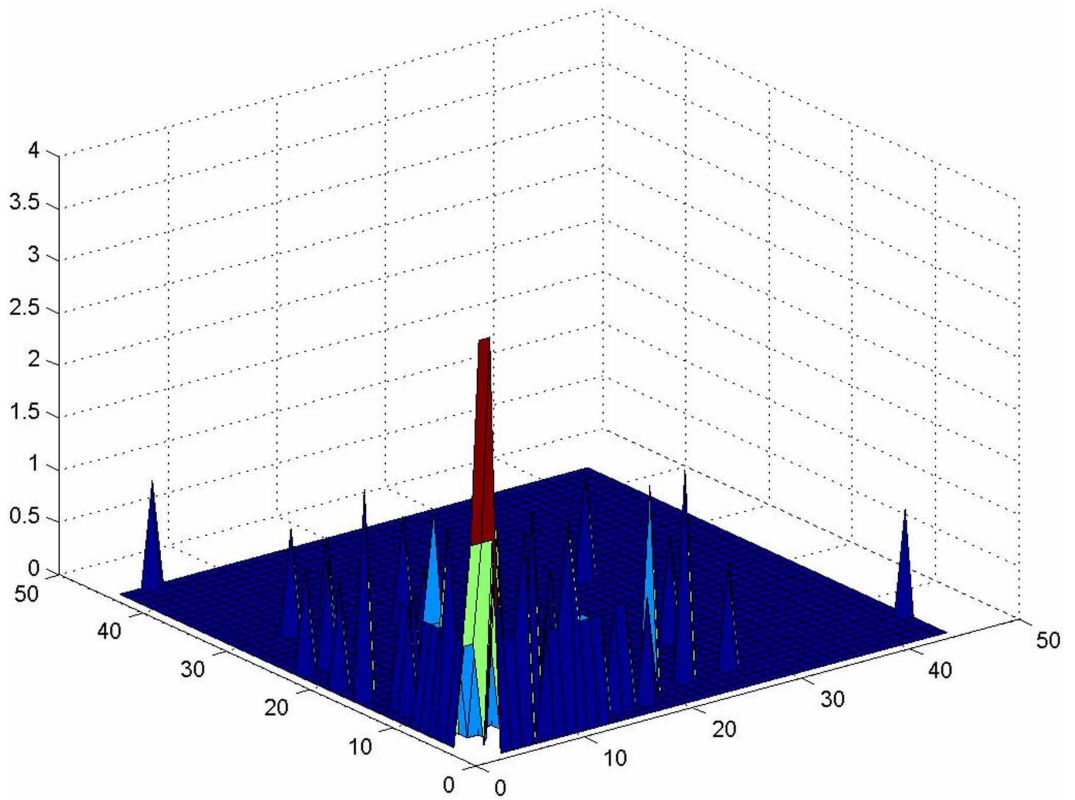


Figure 14. ROC curve intended for Naive Bayes for plucking type string instruments

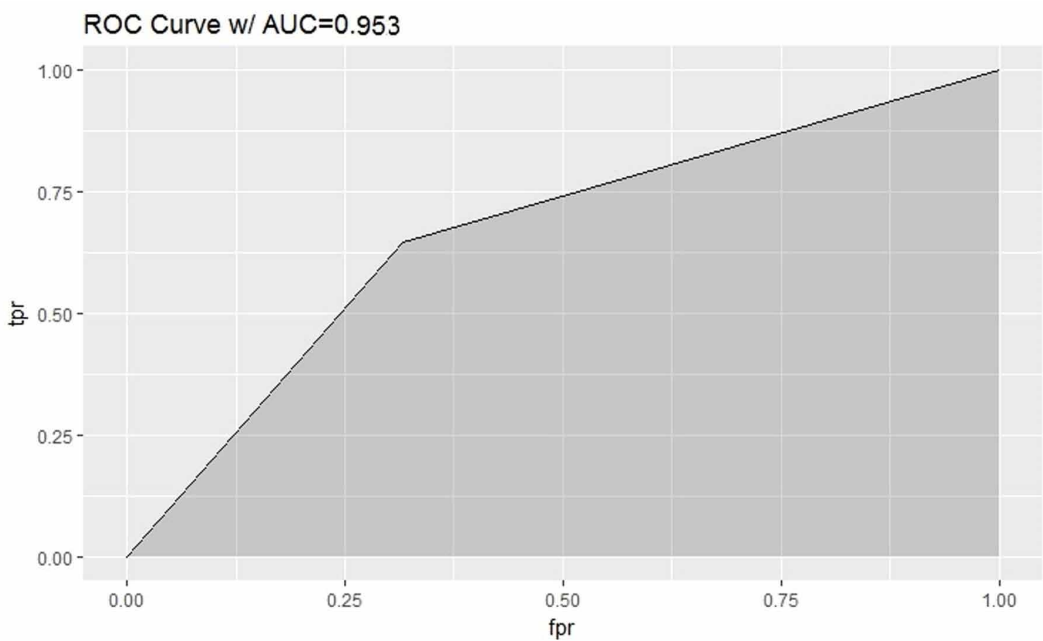


Figure 15. ROC curve intended for neural network (MLP) for plucking type string instruments

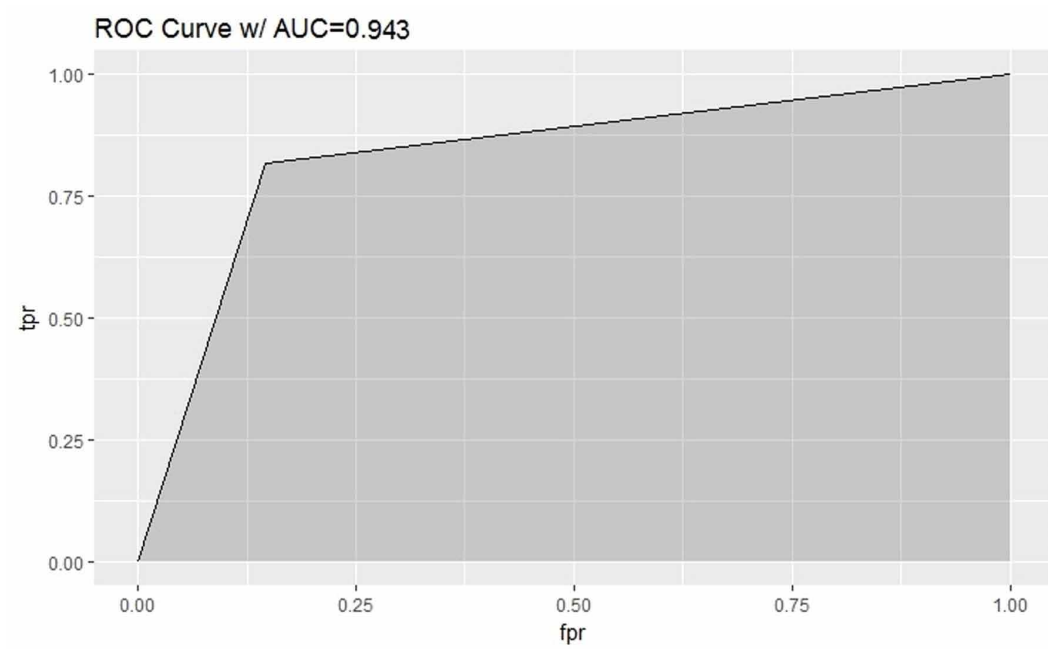


Figure 16. ROC curve intended for k-NN for plucking type string instruments

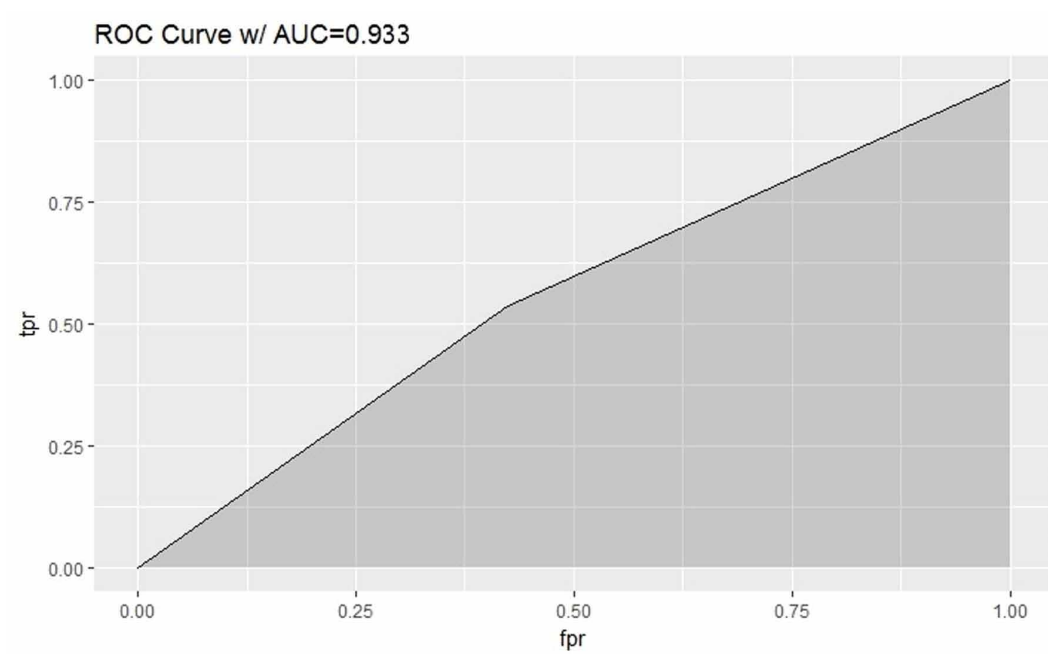


Table 1. Taxonomy correctness for categorization of String type instrumental devices using proposed feature for stage 1

Categorization Scheme	Percentage accuracy of taxonomy for projected effort		
	Plucking	Bowing	Striking
Naïve Bayes	95.3%	94.3%	94.3%
Neural Network (Here MLP)	94.3%	93.3%	93.3%
k-NN	93.3%	92.3%	91.3%

Table 2. Taxonomy correctness for categorization of plucking string type instrumental devices using proposed feature for stage 2

Categorization Scheme	Percentage accuracy of taxonomy for projected effort	
	Fretted	Non-fretted
Naïve Bayes	95.5%	95.0%
SVM	94.5%	94.0%
Neural Network (Here MLP)	93.0%	92.0%

Table 3. Taxonomy correctness for categorization of bowing string type instrumental devices using proposed feature in stage 2

Categorization Scheme	Percentage accuracy of taxonomy for projected effort	
	Wheeled	Non-wheeled
Naïve Bayes	95.0%	94.5%
SVM	94.0%	93.0%
Neural Network (MLP)	92.5%	91.5%

Table 4. Taxonomy correctness for categorization of striking string type instrumental devices using proposed feature in stage 2

Categorization Scheme	Percentage accuracy of taxonomy for projected effort	
	Mono-stringed	Multi-stringed
Naïve Bayes	95.5%	94.5%
SVM	94.0%	93.5%
Neural Network (MLP)	93.5%	92.5%

CONCLUSION

Here a hierarchical scheme is projected to categorize string type instrumental devices in different classes and sub-classes. As MFCC is a very frequently used acoustic feature, an alternative feature to MFCC is proposed in this work. As the notes of different varieties of string type instrumental devices like bowing, plucking and striking are different chroma based features are used to classify them. But due to presence of fret sound quality of different types of plucking string instruments differs. Also,

Table 5. Taxonomy correctness for String instrumental devices classification with MFCC as feature

Categorization Scheme	Percentage accuracy of taxonomy for projected effort		
	Plucking	Bowing	Striking
Naïve Bayes	93.3%	91.3%	91.3%
Neural Network (Here MLP)	92.3%	91.3%	90.3%
k-NN	91.3%	90.3%	90.3%

Table 6. Percentage accurateness of relative investigation of projected effort against other efforts

Efforts Name	Plucking	Bowing	Striking
Kumar, Sebastian and Murthy (2015)	91.3%	90.3%	91.3%
Bhalke, Rao and Bormane (2016)	90.3%	89.3%	90.3%
Ghisingh and Mittal (2016)	91.3%	90.3%	91.3%

Table 7. Taxonomy correctness for Plucking String type instrumental devices categorization using MFCC as feature

Categorization Scheme	Percentage accuracy of taxonomy for projected effort	
	Fretted	Non-Fretted
Naïve Bayes	92.5%	91.5%
SVM	92.0%	91.0%
Neural Network (MLP)	91.5%	90.5%

Table 8. Taxonomy correctness for Plucking String type instrumental devices categorization with other work

Name of the Method	Percentage accuracy of taxonomy for projected effort	
	Fretted	Non-Fretted
Peeters (2003)	91.0%	90.0%

Table 9. Taxonomy correctness for Bowing String type instrumental devices categorization using MFCC as feature

Categorization Scheme	Percentage accuracy of taxonomy for projected effort	
	Wheeled	Non-Wheeled
Naïve Bayes	92.5%	91.5%
SVM	92.0%	91.5%
Neural Network (MLP)	91.5%	91.0%

Table 10. Taxonomy correctness for Bowing String type instrumental devices categorization with other work

Name of the Method	Percentage accuracy of taxonomy for projected effort	
	Wheeled	Non-Wheeled
Peeters (2003)	91.5%	90.5%

Table 11. Taxonomy correctness for Striking String type instrumental devices categorization using MFCC as feature

Categorization Scheme	Percentage accuracy of taxonomy for projected effort	
	Mono-Stringed	Multi-Stringed
Naïve Bayes	93.5%	92.0%
SVM	92.0%	91.5%
Neural Network (MLP)	91.5%	90.0%

Table 12. Taxonomy correctness for Striking String type instrumental devices categorization with other work

Name of the Method	Percentage accuracy of taxonomy for projected effort	
	Mono-Stringed	Multi-Stringed
Peeters (2003)	90.5%	90.0%

the way of playing a bow is different in the bowing type string instruments and the number of strings differs in case of striking type string instruments. Due to this phenomenon a hierarchical scheme is proposed to classify string instruments. From the classification result it can be concluded that proposed feature set in every stage performs better than previous approaches as well as the proposed feature set can be considered as an alternative of MFCC. Computationally all the proposed feature sets are simple and are of low-dimension. To show up categorization power of suggested trait sets popular and effortless taxonomy schemes for example Neural Network, Support Vector Machine, Naïve Bayes' and k-NN have been applied.

REFERENCES

- Abeber, J., & Weib, C. (2015). Automatic Recognition of Instrument Families in Polyphonic Recordings of Classical Music. In *Proceedings of the 16th International Society for Music Information Retrieval Conference*.
- Agostini, G., Longari, M., & Pollastri, E. (2003). Musical instrument timbres classification with spectral features. *EURASIP Journal on Applied Signal Processing*, 2003, 5–14.
- Arora, V., & Behera, L. (2014). Instrument identification using PLCA over stretched manifolds. In *Proceedings of the 2014 Twentieth National Conference on Communications (NCC)* (pp. 1-5). IEEE. doi:10.1109/NCC.2014.6811270
- Barbedo, J. G. A., & Tzanetakis, G. (2010). Instrument identification in polyphonic music signals based on individual partials. In *Proceedings of the 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)* (pp. 401-404). IEEE. doi:10.1109/ICASSP.2010.5495794
- Barbedo, J. G. A., & Tzanetakis, G. (2011). Musical instrument classification using individual partials. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(1), 111–122. doi:10.1109/TASL.2010.2045186
- Benetos, E., Kotti, M., & Kotropoulos, C. (2006). Musical instrument classification using non-negative matrix factorization algorithms. In *Proceedings of the IEEE International Symposium on Circuits and Systems*. doi:10.1109/ISCAS.2006.1692967
- Bhalke, D. G., Rao, C. B. R., & Bormane, D. S. (2016). Automatic musical instrument classification using fractional Fourier transform based-MFCC features and counter propagation neural network. *Journal of Intelligent Information Systems*, 46(3), 425–446. doi:10.1007/s10844-015-0360-9
- Chandwadkar, D. M., & Sutaone, M. S. (2012). Role of features and classifiers on accuracy of identification of musical instruments. In *Proceedings of the 2012 2nd National Conference on Computational Intelligence and Signal Processing (CISP)* (pp. 66-70). IEEE. doi:10.1109/NCCISP.2012.6189710
- Dandawate, Y. H., Kumari, P., & Bidkar, A. (2015). Indian instrumental music: Raga analysis and classification. In *Proceedings of the 2015 1st International Conference on Next Generation Computing Technologies (NGCT)* (pp. 725-729). IEEE. doi:10.1109/NGCT.2015.7375216
- Deng, J. D., Simmermacher, C., & Cranefield, S. (2008). A study on feature analysis for musical instrument classification. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics*, 38(2), 429–438. doi:10.1109/TSMCB.2007.913394 PMID:18348925
- Essid, S., Richard, G., & David, B. (2006). Hierarchical classification of musical instruments on solo recordings. In *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2006* (Vol. 5). IEEE. doi:10.1109/ICASSP.2006.1661401
- Gaikwad, S., Chitre, A. V., & Dandawate, Y. H. (2014). Classification of Indian classical instruments using spectral and principal component analysis based cepstrum features. In *Proceedings of the 2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies (ICESC)* (pp. 276-279). IEEE. doi:10.1109/ICESC.2014.52
- Ghisingh, S., & Mittal, V. K. (2016). Classifying musical instruments using speech signal processing methods. In *Proceedings of the 2016 IEEE Annual India Conference (INDICON)* (pp. 1-6). IEEE. doi:10.1109/INDICON.2016.7839034
- Ghosal, A., Chakraborty, R., Dhara, B. C., & Saha, S. K. (2011). Automatic Identification of Instrument Type in Music Signal Using Wavelet and MFCC. In K. R. Venugopal & L. M. Patnaik (Eds.), *Computer Networks and Intelligent Computing* (pp. 560–565). Springer. doi:10.1007/978-3-642-22786-8_71
- Grindlay, G., & Ellis, D. P. W. (2011). Transcribing multi-instrument polyphonic music with hierarchical eigen instruments. *IEEE Journal of Selected Topics in Signal Processing*, 5(6), 1159–1169. doi:10.1109/JSTSP.2011.2162395
- Gunasekaran, S., & Revathy, K. (2008). Recognition of Indian Musical Instruments with Multi-Classfier Fusion. In *Proceedings of the International Conference on Computer and Electrical Engineering ICCEE 2008* (pp. 847-851). IEEE. doi:10.1109/ICCEE.2008.159

- Haralick, R. M., & Shapiro, L. G. (1992). *Computer and Robot Vision*. Addison-Wesley Longman Publishing Co.
- Joshi, M., & Nadgir, S. (2014). Extraction of feature vectors for analysis of musical instruments. In *Proceedings of the 2014 International Conference on Advances in Electronics, Computers and Communications (ICAIECC)* (pp. 1-6). IEEE. doi:10.1109/ICAIECC.2014.7002391
- Kaminskyj, I., & Czaszejko, T. (2005). Automatic recognition of isolated monophonic musical instrument sounds using knnc. *Journal of Intelligent Information Systems*, 24(2/3), 199–221.
- Kumar, P. A. M., Sebastian, J., & Murthy, H. A. (2015). Musical onset detection on carnatic percussion instruments. In *Proceedings of the 2015 Twenty First National Conference on Communications (NCC)* (pp. 1-6). IEEE.
- Kumari, M., Kumar, P., & Solanki, S. S. (2010). Classification of North Indian Musical Instruments using Spectral Features. *Computer Science & Telecommunications*, 29(6), 11-24.
- Liu, J., & Xie, L. (2010). Svm-based automatic classification of musical instruments. In *Proceedings of the 2010 International Conference on Intelligent Computation Technology and Automation (ICICTA)* (Vol. 3, pp. 669-673). IEEE. doi:10.1109/ICICTA.2010.64
- Müller, M., & Ewert, S. (2011). Chroma Toolbox: MATLAB implementations for extracting variants of chroma-based audio features. In *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR)*.
- Peeters, G. (2003). Automatic classification of large musical instrument databases using hierarchical classifiers with inertia ratio maximization. In *Proceedings of the Audio Engineering Society Convention 115*. Audio Engineering Society.
- Senan, N., Ibrahim, R., Nawi, N. M., & Mokji, M. M. (2009, December). Feature extraction for traditional Malay musical instruments classification system. In *Proceedings of the 2009 International Conference of Soft Computing and Pattern Recognition* (pp. 454-459). IEEE. doi:10.1109/SoCPaR.2009.94
- Sinith, M. S., & Rajeev, K. (2007). Pattern recognition in South Indian classical music using a hybrid of HMM and DTW. In *Proceedings of the International Conference on Conference on Computational Intelligence and Multimedia Applications* (Vol. 2). IEEE. doi:10.1109/ICCIMA.2007.337
- Wikipedia. (n.d.). Musical Instrument Classification. Retrieved from http://en.wikipedia.org/w/index.php?title=Musical_instrument_classification&oldid=325554329
- Zhu, J., Xue, X., & Lu, H. (2004). Musical genre classification by instrumental features.