SPECTROGRAM BASED MUSICAL INSTRUMENT IDENTIFICATION USING HIDDEN MARKOV MODEL (HMM) FOR MONOPHONIC AND POLYPHONIC MUSIC SIGNALS

D.G. BHALKE[1], C. B. RAMA RAO[1], D. S. BORMANE[2], M. VIBHUTE[2]
[1]*NIT Warangal,* [2]*RSCOE, Pune, India*
bhalkedg2000@yahoo.co.in, cbrr@nitw.ac.in, bdattatraya@yahoo.com

Abstract: Spectrogram is generated for musical notes, which is used to calculate the spectral, temporal and modulation features. To detect the musical instruments from polyphonic and monophonic musical notes , 23 features are analyzed . Out of 23 features 12 specific features are used to generate feature vector . Hidden Markov model (HMM) is used to calculate the conditional instrument existence probability. In this work, ten musical instruments from wind and string categories are used for identification. The musical instruments are recognized using different HMM algorithms: Forward, Backward, Posterior decoding and Viterbi algorithm and their results are compared . Recognition accuracy achieved for monophonic musical notes are 91% and 87% for polyphonic musical notes.

Keywords: *Spectrogram, Feature vector, HMM, Musical Instrument Identification result..*

I. INTRODUCTION

Musical instrument recognition is an important task for many applications including automatic music transcription, musical information retrieval, analysis of music tune for copyright, computational auditory scene analysis, which is used in online music distribution services and portable digital music players.

In last decade, musical instrument recognition studies mainly dealt with solo musical sounds. In recent years, identification of musical instruments playing in polyphonic music is based on estimation of on-set time and fundamental frequency of each note, time domain waveform template matching, or block matching techniques are used.

The key idea of the technique used in this work is to visualize the probability that the sound of each of the target instruments exists at each time and each frequency. The technique uses spectrogram, which calculates the spectral, temporal and modulation features of musical instrument for every dominating fundamental frequency F0 in a frame. This approach made it possible to avoid errors caused by conventional method based on estimation of pitch, duration and timbre.

In addition, by using a Markov chain whose states corresponds to target instruments features for every possible F0, the identification of musical instruments for monophonic and polyphonic music can be achieved [1], [12].

The specific instrument existence probability is calculated using the hidden Markov model (HMM) since spectral and temporal characteristics of an instrument are considered while recognizing musical instruments.

At each frame, an observed spectrum of the input signal containing multiple musical instrument sound is a weighted mixture of harmonic structure [2] with every possible fundamental frequency F0. The weight (amplitude) of each harmonic structure represents how relatively predominant it is. Each HMM model consists of multiple states. The instrument existence probability $p(\omega_i|exist; t, f)$ can be calculated from the likelihoods of paths in the chain.

In this work, ten musical instruments are used for identification. They are classified as wind instruments: Flute, Brass-Tuba, Harmonica, Saxophone, Trumpet and string instruments: Guitar, Piano, Mandolin, Sitar, and Violin. In musical instrument identification, two phases are used, one is the feature vector training phase and the other is instrument identification phase. In feature vector training phase, feature vector for selected musical instrument is calculated, which is used for identification of musical instrument. In identification phase, musical note is played and from the identified features musical instrument existence probability is calculated using HMM. Four HMM algorithms: Forward, Backward, Posterior decoding and Viterbi algorithms are used for musical instrument identification.

II. INSTRUMENT IDENTIFICATION METHOD

Musical notes from wind and string family are used for instrument identification. The flow of musical instrument identification is as shown in figure 1.

Manuscript received February 18, 2011; revised June 23, 2011

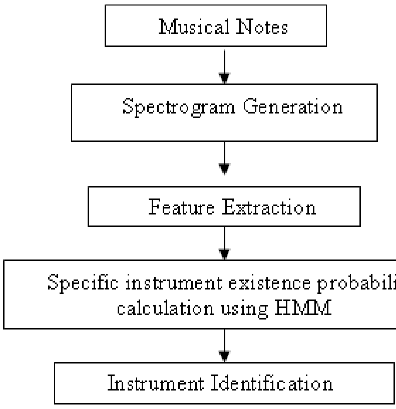


Figure 1: Flow of musical instrument identification.

II A. SPECTROGRAM GENERATION

The spectrogram is obtained by splitting the musical audio signal into short-time segments and by taking the Fourier transform of each segment [4]. To analyze music signals wideband spectrogram is convenient for investigating the characteristics of the musical note as it provides a spectral envelope.

Each image of the spectrogram is a time - frequency plane. The intensity of color of each point (t, f) in the image represents the probability $P(\omega_i; t, f)$ that a sound of the target instrument ω_i exists at time t and frequency f. Instrument existence probability is given by:

$$P(\omega_i;t,f) = P(\omega_i| \text{exist};t,f) \quad (1)$$

$P(\omega_i| \text{exist};t,f)$ is called instrument existence probability. It is the conditional probability that, if a sound of a certain instrument exists at time t and frequency f, then the instrument is ω_i. The instrument is identified by using HMM.

At each frame, an observed spectrum of the input signal contains multiple musical instruments sound as a weighted mixture of harmonic-structure with every possible fundamental frequency F0. The fundamental frequency F0 is estimated from most predominant harmonic structure, with weighted input.

II B. FEATURE EXTRACTION

Feature extraction is the most vital part of this work. From literature survey, 23 features [1][9],[7],[8] are selected for identification of musical instrument. These features are classified as spectral, temporal and modulation features. The feature vector is generated from specific features of instruments. This feature vector is used as instrument model. The features used for identification of musical instruments are as given in Table 1.

Table 1: Overview of 23 Features

No.	Features	Details
1	Spectral	1. Fundamental frequency F0 2- 9. Harmonic components (i =2, 3, ···, 9) 10-17. Relative amplitude of Harmonic components (i= 2, 3,., 9) 18. Spectral centroid (SC)
2	Temporal	19. Attack time 20. Decay time 21. Roll-off Rate
3	Modulation	22. Amplitude modulation index 23. Frequency modulation.

SPECTRAL FEATURES

The following spectral features are obtained from the musical notes:

1. Fundamental frequency F0-

Fundamental frequency (F0) is the physical term, defined only for periodic or nearly periodic sounds. F0 is the inverse of the period corresponding to pitch chosen.

Musical instrument notes (A- G) can be identified from F0, using the formula:

$$Note\_frequency = \frac{440}{e^{(n/12\times\ln 2)}} \quad (2)$$

where n is semitone. Semitone is the frequency ratio given by 12^{th} root of 2.

$$\frac{f_2}{f_1} = 2^{1/12} \quad (3)$$

2- 9. Harmonic components (i =2, 3, ··· , 9)

Harmonics or trajectory frequency is generally integral multiple of fundamental frequency. Trajectory frequency is given by,

$$Trajectory\_freq = F0 \cdot harmonic\_nr \quad (4)$$

10-17. Relative amplitude of Harmonic components (i= 2,3,.,9)

Relative amplitude of Harmonic components is calculated with respect to fundamental component. This feature indicates the spectral variations in musical instruments.

$$\mathrm{Re}lative\_Harm\_nr = \frac{Harm\_Ampl}{Fdt\_Ampl} \quad (5)$$

18. Spectral centroid (SC)

Spectral centroid measures the average frequency weighted by sum of spectrum amplitude within one frame. Spectral centroid measures the spectral energy distribution in steady state portions of tone [7][8].

$$SC = \frac{\sum f \cdot |X(k)|^2}{\sum |X(k)|^2}(Hz), \text{ k=0 to N/2} \qquad (6)$$

for X(k) = Spectrum magnitude of k^{th} frame .
f_x = Corresponding frequency of X .
N = Number of samples.

TEMPORAL FEATURES

It consists of features which gives information of the shape of the musical note [3].

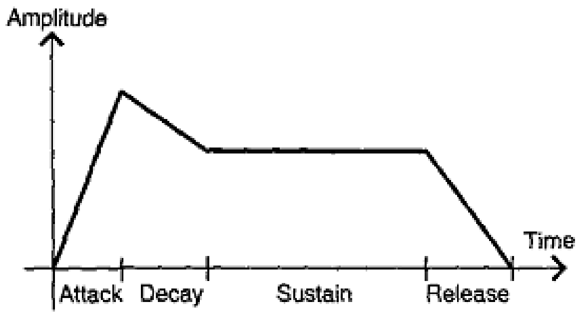


Figure 2: ADSR envelope.

19. Attack time

Attack time is the time difference between onset and end-of-attack. Onset time is the time at which the note begins to sound.

20. Decay time

Decay time is the time difference between end-of-attack and the forward position where the amplitude is 25% of the amplitude at end-of-attack.

21. Roll-off Rate

Roll-off Rate is the rate of decrease of amplitude from end-of-attack to the forward position where the amplitude is decreased by 25% of the amplitude at end-of-attack.

MODULATION FEATURES

Body resonance in musical instrument causes Amplitude Modulation. Finger or bow on string causes Frequency Modulation. The modulation features considered are as follows:

22. Amplitude Modulation Index

The amplitude modulation index is the change in pitch amplitude.

$$Mod.Index[\%] = \frac{\max.amp. - \min.amp.}{\max.amp. + \min.amp.} \cdot 100 \qquad (7)$$

23. Frequency Modulation

The frequency modulation is the ratio of change in frequency and the modulating frequency.

$$Freq.Modulation = \frac{Change\_in\_frequency}{F_s/2} \qquad (8)$$

Fs=Sampling frequency.

II C. SPECIFIC INSTRUMENT EXISTENCE PROBABILITY CALCULATION USING HMM

Feature extraction is followed by specific instrument existence probability calculation using hidden Markov Model [5]. Markov Model is a statistical model for prediction. For a sequence $\{q_1, q_2, ..., q_n\}$, the first-order Markov assumption probability depends on observation q_{n-1} at time (n – 1) and is given by,

$$P(q_n | q_{n-1}, q_{n-2}, ..., q_1) = P(q_n | q_{n-1}) \qquad (9)$$

The second-order Markov assumption probability depends on observation q_{n-1} and q_{n-2}. An output sequence $\{q_i\}$ of such a system is a Markov chain.

For hidden values - according to Bayes' rule conditional probability is given by,

$$P(q_i | x_i) = \frac{P(x_i | q_i) \cdot P(q_i)}{P(x_i)} = P(x_i | q_i) P(q_i) \qquad (10)$$

Note that in equation (10), P(xi) is considered as negligible, since it is independent of sequence qi. The transition probabilities are the probabilities to go from state i to state j is,

$$a_{i,j} = P(q_{n+1} = s_j | q_n = s_i) \qquad (11)$$

A HMM allows for transitions from any emitting state to any other emitting state is called an ergodic HMM. In the other type of HMM, the transitions only go from one state to itself or to a unique follower is called a left-right HMM.[12]

The model is called Hidden Markov Model (HMM) because the sequence of states that produces the observable data is not available (hidden).

Entire Model is given by: $\lambda = (A, B, \pi)$ (12)

Emission probability distribution is continuous in each state and can be represented by a Gaussian mixture model. Emission probability distribution is continuous in each state and can be represented by a Gaussian mixture model.

$$E_j(O) = f(O; \mu_j, \sigma_j) \quad , j=1,N \qquad (13)$$

For continuous observation HMM, the joint probability of both sequence of observations, O and sequence of state Q occurring simultaneously in Lambda model is given by:

$$P(O,Q|\lambda)=\pi_{q_1}e_{q_1}(O_1)\prod_{i=2}^{L}a_{q_{i-1},q_1}e_{q_i}(O_1) \qquad (14)$$

The instrument existence probability is calculated from the likelihoods of paths in the Markov chain. The procedure of HMM algorithms: Forward, Backward, Posterior decoding and Viterbi algorithms [12] is explained below.

FORWARD PROCEDURE

The Forward and Backward Procedure is based on the technique known as dynamic programming. Dynamic programming makes calculations for a small instance, stores the result, and then uses it later. To apply dynamic programming, a recursive property is used that allows us to do calculations for the next instance based on the current one.

Let $\alpha_k(i)$ be the probability of the partial observation sequence $O1\rightarrow k = O_1O_2\ldots O_k$ to be produced by all possible state sequences that end at i-th state. Then the probability of the partial observation sequence is the sum of $\alpha_k(i)$ over all N states.

$$\alpha_k(i)=P(O_1O_2...O_k,q_k=i|\lambda) \qquad (15)$$

$$P(O_1O_2...O_k|\lambda)=\sum_{i=1}^{N}\alpha_k(i) \qquad (16)$$

The Forward Procedure is a recursive algorithm for calculating $\alpha_k(i)$ for the observation sequence of increasing length k. First, the probabilities for the single-symbol sequence are calculated as a product of initial i-th state probability and emission probability of the given symbol O1 in i-th state. Then the recursive formula is applied. Assume we have calculated $\alpha_k(i)$ for some k. To calculate say $\alpha_{k+1(2)}$, we multiply every $\alpha_k(i)$ by corresponding transition probability from i-th state to the second state, sum the products over all states, and then multiply the result by the emission probability of the symbol O_{k+1}. Iterating the process, we can calculate $\alpha_{k(L)}$, and then summing them over all states, we can obtain the required probability.

BACKWARD PROCEDURE

In a similar manner, we can introduce a symmetrical backward variable $\beta k(i)$ as the conditional probability of the partial observation sequence from O_{k+1} to the end to be produced by all state sequences that start at i-th state. The Backward Procedure calculates recursively backward variables going backward along the observation sequence.

$$\beta_k(i)=P(O_{k+1}O_{k+2}...O_L,q_k=i,\lambda) \qquad (17)$$

$$P(O_{k+1}O_{k+2}...O_L|\lambda)=\sum_{i=1}^{N}\pi_ie_i(O_{k+1})\beta_k(i) \qquad (18)$$

Both procedures are generally used for finding the optimal state sequence and estimating the HMM parameters.

POSTERIOR DECODING

The precise posterior decoding of the HMM states can be obtained by application of a forward–backward algorithm, which is used in many speech recognition. To choose the states those are individually most likely at the time when a symbol is emitted.

If $\lambda_k(i)$ is the probability of the model to emit k-th symbol being in the i-th state for the given observation sequence.

$$\lambda_k(i)=P(q_k=i|O,\lambda) \qquad (19)$$

$$\lambda_k(i)=\frac{\alpha_k(i)\beta_k(i)}{P(O|\lambda)},\quad k=1,L;\ i=1,N \qquad (20)$$

$$Q_k=\arg\_\max\{\lambda_k(i)\},\quad k=1,L;\ i=1,N \qquad (21)$$

VITERBI ALGORITHM

The Viterbi algorithm chooses one best state sequence that maximizes the likelihood of the state sequence for the given observation sequence.

It keeps track of the arguments that maximize $\delta_k(i)$ for each k and i storing them in the N by L matrix. This matrix is used to retrieve the optimal state sequence at the backtracking step.

$$Q^*=\arg\_\max\{P(Q|O,\lambda)\} \qquad (22)$$

$$\delta_k(i)=\max\left[P(q_1q_2...q_k=1,O_1O_2...O_k|\lambda)\right] \qquad (23)$$

III. SYSTEM IMPLEMENTATION

A technique based on spectrogram and musical instrument identification using HMM is implemented, which takes a musical audio signal, in .wav format as input (with sampling frequency 44.1 KHz) and recognizes the musical instruments. A database of ten instruments, which are classified as wind instruments and string instruments, is collected. A wide set of 23 numbers of audio features are used. The 18 spectral features, 3 temporal features, and 2 modulation features, are considered for classification of musical instruments. To deal with the diversity of the harmonic structure the weighted features are used to enhance the specific features of the instrument.

The block schematic diagram for feature vector training phase and instrument identification phase is shown in figure 3 and 4.

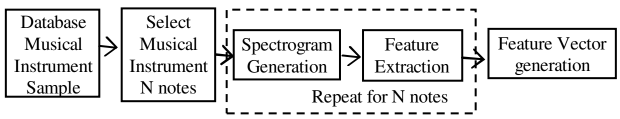


Figure 3: Feature Vector Training Phase.

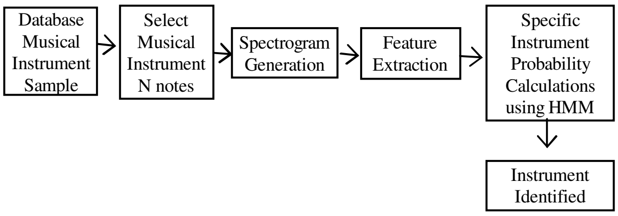


Figure 4: Instrument Identification Phase.

III A. CONDITIONAL INSTRUMENT EXISTENCE PROBABILITY USING HMM

Conditional Instrument Existence Probability function computes the probability of the specific features: Relative harmonics (i= 2…9) amplitude with respect to the fundamental component, attack time, decay time, roll-off rate and frequency modulation.

1. Feature vector x (t, f) is generated by musical instrument Mi note at time t, which is given by P (x (t, f) |Mi; t).

The probability of feature x generated by musical instrument Mi at time t is given by,

$$P (x \mid Mi; t) = P(x \mid Mi) * P(Mi \mid Mj) \tag{24}$$

2. Since music signals are random signals, the probability of feature x generated by musical instrument Mi is given by the Normal density or Gaussian density function, equation (25).

$$P\left(x \middle| M_i\right) = \frac{e^{(-1/2)((x-\mu)/\sigma)^2}}{\sigma\sqrt{2}\pi} \tag{25}$$

where co-variance $= \sigma^2 = \sum \frac{(x_i - \mu)^2}{n}$.

III B. FEATURE VECTOR TRAINING ALGORITHM

The features of all specified ten instruments are studied and specific twelve features (Relative harmonics amplitude with respect to the fundamental component, attack time, decay time, roll-off rate and frequency modulation) are selected for recognition of a particular musical instrument. Feature vector is generated for each instrument using feature vector training algorithm. Features are weighted to improve the recognition accuracy. Weights are applied to the dominant features of that particular instrument.

IV RESULTS

IV A. DATABASE

The musical notes of wind and string family instruments are recorded with sampling frequency 44.1 KHz in our Lab . The wind instruments used are Flute, Brass-Tuba, Harmonica, Saxophone, Trumpet .String instruments used are Guitar, Piano, Mandolin, Sitar, and Violin. The seven notes from A to G of each instrument are recorded using Yamaha-PSR-I425 Electronic keyboard. Specifications of keyboard are 61-key touch sensitive keyboard + 32 note polyphony, produce natural and realistic sounds of 514 instruments, supports MIDI format, compressive sound recording function.

The polyphonic Duo and Trio (combinations) samples are generated using computer mixing software, Audacity. The monophonic and polyphonic samples used in experiments are as given in Table 2 and Table 3.

Table 2: Monophonic Samples with sampling frequency Fs=44.1khz

Sr. No.	Name of Instrument	Abbreviations used for Instrument in this paper	Musical Notes used	Number of Samples used
	Flute	F	A - G	21
2	Guitar	G	A -G	21
3	Piano	P	A - G	21
4	Brass (Tuba)	B	A - G	21
5	Mandolin	M	A - G	21
6	Harmonica	H	A - G	21
7	Saxophone	Sx	A - G	21
8	Sitar	S	A - G	21
9	Trumpet	T	A - G	21
10	Violin	V	A - G	21

Table 3: Polyphonic samples with sampling frequency Fs=44.1khz

Sr. No.	Name of Instruments	Abbreviations used for combinations of Instruments	Musical Notes used	Number of Samples used
1	Flute + Guitar	F+G	A - G	21
2	Piano +Guitar	P+G	A -G	21
3	Brass+ Guitar	B+H	A - G	21
4	Harmonica + Saxophone	H+S	A - G	21
5	Mandolin + Saxophone	M+Sx	A - G	21
6	Trumpet + Violin	T+V	A - G	21
7	Flute + Piano + Guitar	F+P+G	A - G	21
8	Brass + Harmonica + Sitar	B+H+S	A - G	21
9	Mandolin + Saxophone + Trumpet	M+Sx+T	A - G	21

10	Mandolin + Trumpet + Violin	M+T+V	A - G	21

IV B. EXPERIMENTAL RESULTS

In this work two identification phases are used, feature vector training phase and instrument identification phase. The results of these two phases are as given below:

FEATURE VECTOR TRAINING RESULTS

The feature vector training phase is used to calculate the feature vector for a particular instrument. The spectral, temporal and modulation features are extracted from the musical notes of a instrument and mean values of specific features are calculated to generate the feature vector.

We have recognized both monophonic and polyphonic music samples in this work. We have used Trellis diagram show the possible musical instrument (state) for the features extracted from music sample. If the feature varies than there is transition from one state to another state. The joint likelihood of observation sequence of overall possible states is calculated to identify the specific musical instrument playing.

MONOPHONIC SOUND SAMPLES:

For each musical instrument note, the spectrogram is generated by using hamming window with window length 64. The spectral, temporal and modulation features are extracted from the spectrogram. The spectrogram and spectral features of C-note of the monophonic instruments: Flute, Guitar and Piano are shown in figures 5 – 10.

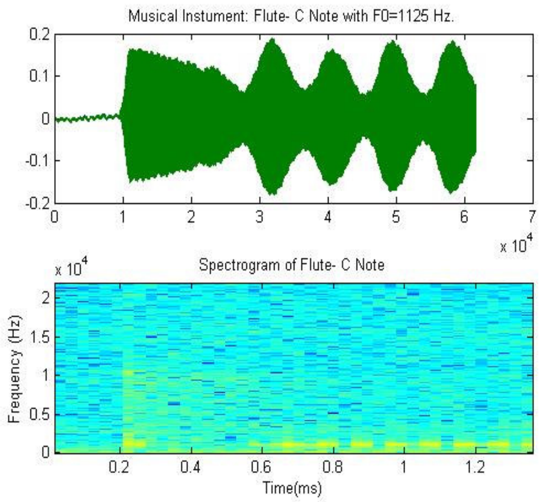


Figure 5: Spectrogram of Flute- C note.

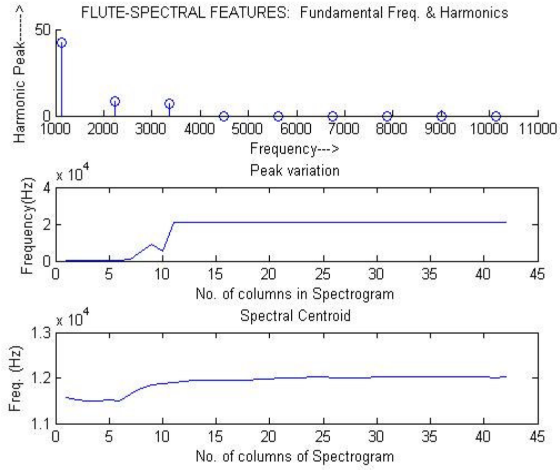
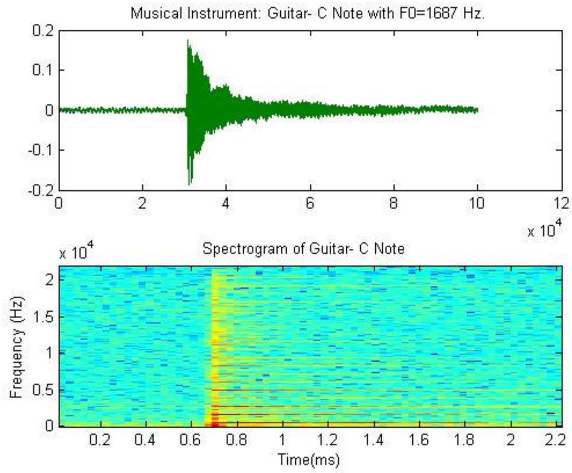


Figure 6: Spectral features of Flute- C note.



. Figure 7: Spectrogram of Guitar-C note.

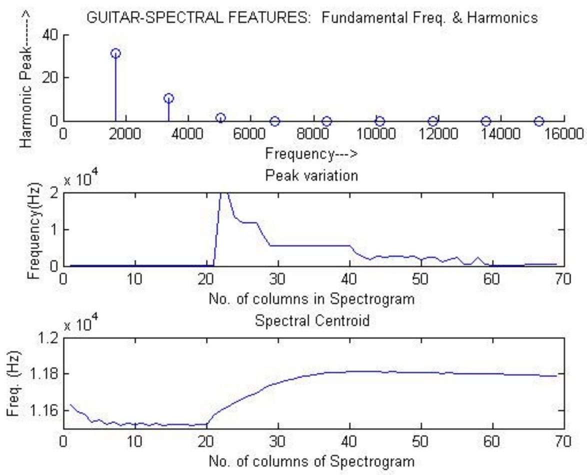


Figure 8: Spectral features of Guitar-C note.

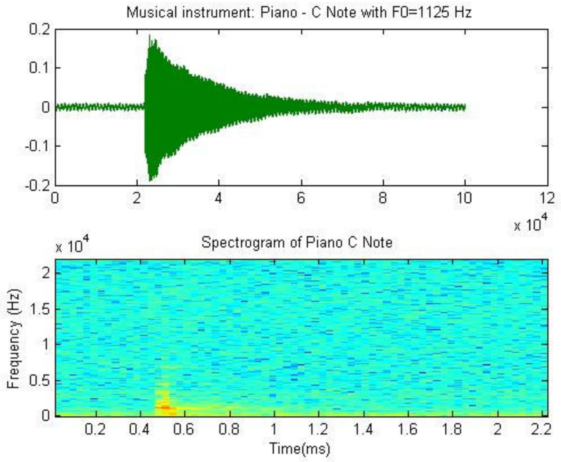


Figure 9: Spectrogram of Piano C- note.

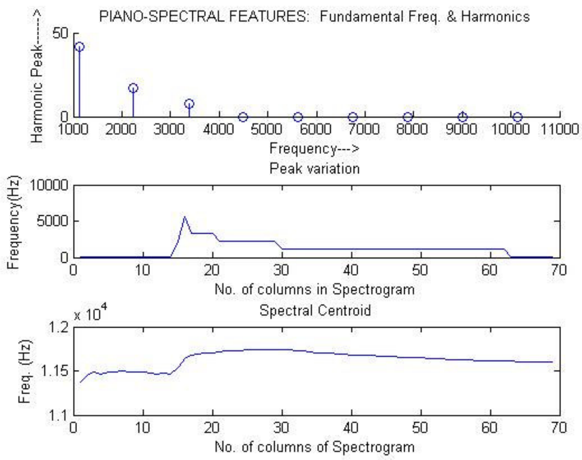


Figure 10: Spectral features of Piano C- note.

Both monophonic and polyphonic sound samples are used during experiments.

POLYPHONIC SOUND SAMPLES:

From the polyphonic note, duo and trio (combinations), the spectrogram is generated by using hamming window with window length 64. The spectral, temporal and modulation features are extracted from the spectrogram. The spectrogram, spectral features and Trellis diagram of the Trio sample: Flute + Guitar + Piano – B note are shown in figures 11-13.

For Instrument identification, the four HMM algorithms: Forward algorithm, Backward algorithm, Posterior decoding, and Viterbi algorithms are used. These algorithms calculate the most likely path through the Hidden Markov Model specified by state transition probability matrix.

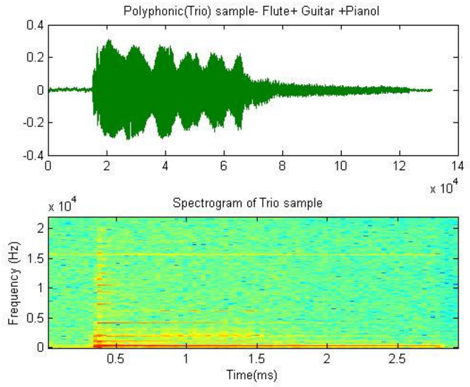


Figure 11: Spectrogram of Polyphonic Trio (Flute + Guitar+ Piano)-B Note.

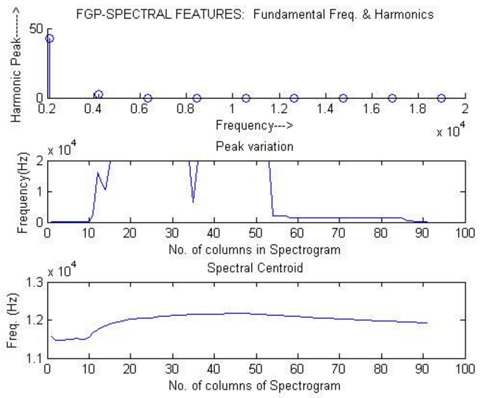


Figure 12:Spectral features of Polyphonic Trio (Flute + Guitar+ Piano)-B Note.

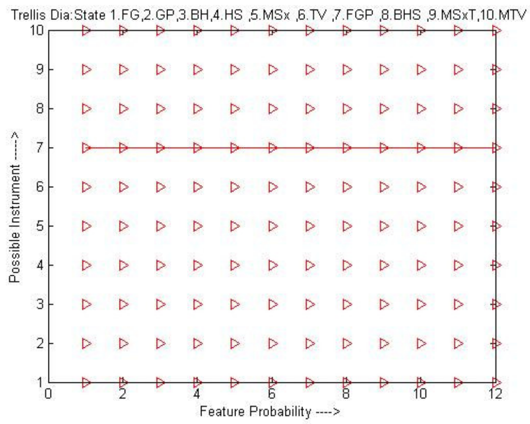


Figure 13: Trellis diagram of Polyphonic Trio (Flute + Guitar+ Piano)-B Note.

Transition (I, J) is the probability of transition from state I to state J (i.e. from one instrument to the other instrument). Since ten instruments are considered in this work, so the order of Transition matrix is (10 X 10).

Table 4: Transition matrix = (10 X 10)

1	0	0	0	0	.	.	0
0	1	0	0	0	.	.	0
0	0	1	0	0	.	.	0
.	.	.	.	.	.	.	.
0	0	0	0	0	.	.	1

The elements in the matrix are chosen as per the Left-Right HMM transition matrix, which indicate that only one state or instrument is considered at a time. There are no transitions from one instrument to the other instrument. The Transition matrix used in this work is as shown in Table 4.

Table 5: Confusion Matrix for monophonic samples

Instruments	Samples classified as									
	F	G	P	B	M	H	Sx	S	T	V
Flute	21	_	-	-	-	-	-	-	-	-
Guitar	02	19	-	-	-	-	-	-	-	-
Piano	-	-	21	-	-	-	-	-	-	-
Brass (Tuba)	-	-	-	18	02	-	-	01	-	-
Mandolin	-	-	02	-	18	01	-	-	-	-
Harmonica	-	-	-	-	-	19	-	-	02	-
Saxophone	-	-	-	-	-	02	18	-	-	01
Sitar	-	-	01	-	-	-	-	20	-	-
Trumpet	-	-	-	-	-	-	04	-	17	-
Violin	-	-	-	-	-	02	-	-	-	19

Table 6: Confusion Matrix for polyphonic samples

Instruments	Samples classified as									
	F+G	P+G	B+H	H+S	M+Sx	T+V	F+P+G	B+H+S	M+Sx+T	M+T+V
F+G	20	01	-	-	-	-	-	-	-	-
P+G	-	21	-	-	-	-	-	-	-	-
B+H	-	-	19	-	02	-	-	-	-	-
H+S	-	-	-	19	-	-	-	02	-	-
M+Sx	-	-	02	-	16	02	-	01	-	-
T+V	-	01	-	-	-	17	-	-	-	03
F+P+G	02	02	-	-	-	-	17	-	-	-
B+H+S	-	-	03	-	-	-	-	18	-	-
M+Sx+T	-	03	-	-	-	-	-	-	18	-
M+T+V	-	-	-	-	-	03	-	-	-	18

In the Confusion matrix as shown in Table 5 and Table 6, each entry is the number of samples that the row instrument is classified as the column instrument. Polyphonic Musical instruments from top to bottom, and left to right are: Flute + Guitar, Piano + Guitar, Brass(Tuba)+ Harmonica, Harmonica + Sitar, Mandolin + Saxophone, Trumpet + Violin, Flute + Piano + Guitar, Brass + Harmonica + Sitar, Mandolin + Saxophone + Trumpet, Mandolin + Trumpet + Violin.

V. CONCLUSION

The specific features such as Relative amplitude of harmonics with respect to the fundamental component, Attack Time, Decay Time, Roll-Off Rate and Frequency Modulation are used to characterize and distinguish the features of the musical instruments. As per the result shown in Table 7, the recognition accuracy is increased by 10% with use of specific features in the feature vector. It further increased by 10% with use of the weighted specific features.

For recognition of instruments from polyphonic music, the feature vectors are used. Experimental results show that the recognition accuracy using weighted features has increased.

During the experiments it is found that the identification results using Viterbi algorithm are more accurate than Forward, Backward and Posterior Decoding algorithms.

From the results it is observed that the identification result for monophonic instruments is 91 % and that for polyphonic instruments is approximately 87%. This was deduced from ten musical instruments results. The results are closer to reference papers' results as shown in comparison Table 8.

Table 7: Recognition Results

Sr. No.	Results	Instrument Recognition (%)	
		Monophonic	Polyphonic
1.	Using General features	72%	60%
2.	Using Specific features	80%	71%
3.	Using weighted features	91%	87%

Table 8 : Comparative Results

Sr. No.	Author [Ref. no.]	Method used	Recognition result
1.	Tetsuro Kitahara et al [1]	HMM	Polyphonic: 73-97%
2.	Tetsuro Kitahara et al [2]	LDA	Polyphonic: 95%
3.	Harya Wicaksana [3]	Composite Neural network	Monophonic: 94%
4.	Tetsuro Kitahara et al [6]	Used Discriminate function	Category: 90%
5.	Anttieronen et al [7]	Energy Separation algorithm	Monophonic: 80% Category:

			94%
6.	Christian Simmermacher [10]	ANN	Monophonic: 94%
7.	Ozbek et al [11]	SVM	Monophonic: 80%

In future development we will be concentrating on integrating the recognizer into a system to process more complex sound mixtures and modifications of existing system into an Intelligent Recognizer. Also the number of features can be reduced by identification of the family of the instrument and the specific instrument.

REFERENCES

[1] Tetsuro Kitahara, Masataka Goto, Kazunori Komatani, Tetsuya Ogata and Hiroshi G. Okuno, "Instrogram: A New Musical Instrument Recognition Technique Without Using Onset Detection Nor F0 Estimation", *2006 IEEE Proceedings*.

[2] Tetsuro Kitahara, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno1 "Instrument Identification in Polyphonic Music: Feature Weighting to Minimize Influence of Sound Overlaps", *EURASIP Journal on Advances in Signal Processing, volume 2007.*

[3] Harya Wicaksana, Septian Hartono, & Foo Say Wei "Recognition of Musical Instruments", *2006 IEEE Proceedings* .

[4] Guoshen Yu and Jean-Jacques Slotine "The Audio Classification from Time-Frequency Texture" *in 25 Sept.2008 paper.*

[4] Matthias Eichner, Matthias Wolff, and Rudiger Hoffmann, "Instrument classification using Hidden Markov Models" *in 2006 paper.*

[5] Teisuro Kitahara, Masaiaka Goto, and Hiroshi G. Okunot, "Musical Instrument Identification Based on Fo-Dependent Multivariate Normal Distribution" *in 2003 IEEE Proceedings.*

[6] Antti Eronen and Anssi Klapuri, "Musical Instrument Recognition Using Cepstral Coefficients and Temporal Features" *in 2000 IEEE Proceedings.*

[7] Da Deng, Christian Simmermacher and Stephen Cranefield, "A Study on Feature Analysis for Musical Instrument Classification", *in The Information Science Discussion Paper Series,* Number 2007/04, August 2007, ISSN 1177-455X.

[8] Antti Eronen, "Automatic Musical Instrument Recognition", *Master of Science Thesis, Tampere University of Technology,* Department of Information Technology.

[9] Christian Simmermacher, Da Deng and Stephen Cranefield, "Feature Analysis and Classification of Classical Musical Instruments: An Empirical Study", *The Information Science Discussion Paper Series,* Number 2006/10, May 2006, ISSN 1172-6024.

[10] Mehmet Erdal Ozbek, Claude Delpha and Pierre Duhamel, "Musical Note And Instrument Classification with Likelihood-Frequency-Time Analysis And Support Vector Machines", *EURASIP Journal on Advances in Signal Processing*, vol. 2007.

[11] Lawrence R. Rabiner , "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proceeding of the IEEE,* vol. 77, No.2, February 1989.