Harmonium Raga Recognition

Rajshri Pendekar, S. P. Mahajan, Rasika Mujumdar, Pranjali Ganoo

***Abstract*—This paper proposes a novel method for the identification of swaras and ragas in an Indian classical harmonium recital. Swaras are musical notes which are produced by pressing any key of the harmonium (an instrument analogous to the piano). Ragas are melodic combinations of swaras which capture the mood and emotions of the performance. We first deal with the segmentation of the audio signal by using two different methods for onset detection viz. use of spectral flux and by fundamental frequency estimation. Swara identification is proposed by using the pitch frequency of the swaras as a distinguishing feature. For this the frequencies associated with each *swara* are first identified and the exact mapping of the frequencies in the given signal is done by using devised database. In case of raga identification, an extensive database of commonly used *swara* permutations is structured and dynamic programming is used for template matching and hence raga recognition.**

***Index Terms*—Harmonium, onset detection, raga recognition, *swara* transcription.**

I. Introduction

Audio transcription and digitization of musical data is a very interesting and challenging research topic. Many techniques have been implemented for this in the context of Western music. However, Indian classical music differs from its western counterparts in the type of instruments being played and their harmonic nature. Hence the current methods cannot be effectively used to analyse pre-recorded Indian classical music, specifically harmonium performances which are an important accompaniment to any Indian musical recital. Fig. 1 shows the harmonium instrument.

Fig. 1. Harmonium.

A harmonium is a free-standing keyboard instrument like the piano. However the peculiarity of the harmonium is that sound is produced by air, supplied by hand-operated bellows. Due to this the musical notes produced, called '*swaras*', are continuous without any intermittent pauses unlike that in a piano. The harmonium has total 36 keys which when pressed produce a particular *swara* [1]. The comparison of the keys of a piano and the keys of a harmonium is as shown in Fig. 2(a) and Fig. 2(b) respectively.

Fig. 2(a). Keys of a piano.

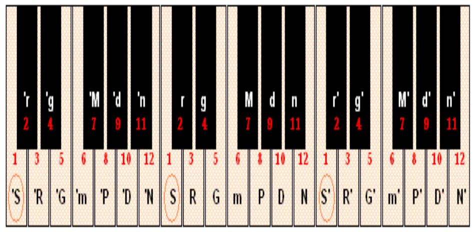

Fig. 2(b). Keys of a harmonium.

Classical Indian music is characterised by seven main musical notes ('pure' *swaras*) called the *'saptak'* viz. *Shadja, Rishab, Gandhar, Madhyam, Pancham, Dhaivat and Nishad* along with five intermediate notes known as altered notes or *'vikrit swaras'*. Further the *swaras* can be played in three octaves, the first or lower octave starting from 130 Hz; the middle octave starting from 260 Hz; and the upper octave from 520 Hz. A combination of five or more notes upon which a particular melody is based is called a '*raga*' [2]. Any classical concert always pertains to a particular *raga* and hence any audio recording requires identification of the *raga* being played.

Automatic *raga* identification can provide a basis for searching the songs based on particular *raga* for automatic generation of playlist. It can be used by novice musicians who find it difficult to distinguish *ragas* which are very similar to each other. It might also evolve into a system which checks how accurately a person is performing a certain *raga* on the harmonium [2], [3]. It can also act as a tutor for beginners for checking the performance skills.

The objective of this paper is to propose a solution for the automatic identification and transcription of the *swaras* as well as the *ragas* present in any pre-recorded stand-alone harmonium performances.

II. System Architecture

The aim of this system is to transcribe harmonium notes signal into a higher level of representation for indexing and retrieval applications. The information automatically extracted from the signal includes the identification of the

Manuscript received January 4, 2013; revised May 18, 2013.

The authors are with the Dept of Electronics and Telecommunication, College of Engineering, Pune 411005, India (e-mail: rpendekar@gmail.com, spm.extc@coep.ac.in).

DOI: 10.7763/IJMLC.2013.V3.336

swaras played and hence the *raga*, the onset time of each note and the scale in which it is played.

The system architecture is then based on 4 major parts:

- Feature extraction and *Swara* Segmentation
- *Swara* Identification module
- Raga Identification
- Database of Ragas

A. Feature Extraction and Swara Segmentation

Any music recognition algorithm requires the segmentation of the audio signal into separate events to detect different musical notes. For this, initially the audio signal was sampled at a rate of 16 KHz and then divided into different frames [4].

1) Design challenges

The harmonium is different from other keyboard instruments in the sense that, the note or *swara* is played continuously. Also the time duration for which a *swara* is played varies according to the mood of the artist and the requirements of the performance. Hence proper onset detection and end-point detection is a very critical aspect for this system.

2) Onset detection

In this paper, we tested two different approaches for onset detection viz. using spectral flux and using fundamental frequency estimation [5]-[7].

a) Spectral flux method

In this method the change in spectral energy from each frame to the next is calculated using spectral flux. Spectral flux (SF) is defined as the squared difference between the normalized magnitudes of successive spectral distribution. It is given by equation (1).

$$\mathrm{SF} = \sum_{k=0}^{K-1} (|X_n(k)| - |X_{n-1}(k)|)^2 \qquad (1)$$

where n-1 and n are the frame indices; k is FFT bin number. $X_n(k)$ is the FFT of n^{th} frame.

The spectral energy peaks obtained after processing the signal using spectral flux method denote the onset of new note. But this method can lead to false detection of the *swara* and is hence, less accurate. To achieve greater accuracy, it is necessary to go for a different approach.

b) Fundamental frequency estimation

This approach relies on joint time-frequency domain analysis of each frame [4].The underlying assumption in most music processing schemes is that the properties of the music signal change relatively slowly with time. This assumption leads to a variety of short-time processing methods in which short segments of the music signal are isolated and processed as if they were short segments from a sustained sound with fixed properties. This is repeated usually periodically as often as desired. Often these short segments, which are sometimes called analysis frames, overlap one another. In present investigation we have taken analysis frame of 30 ms with 20 ms overlap giving a frame rate of 100 frames/second since the harmonium fundamental frequency range is from 100 Hz - 1000 Hz. The criterion for selection of 30 ms frame size is to have at least 2-3 cycles of fundamental frequency of *swara* in order to have accurate estimation. The fundamental frequency or pitch frequency for each frame is obtained by applying autocorrelation function.

Autocorrelation function provides a measure of self-similarity between different alignments of the sequence. Autocorrelation of signal is expressed as in equation (2).

$$r_{xx}(\tau) = \sum_{m=0}^{M-1} x(m)\, x(m+\tau) \qquad (2)$$

where,

$r_{xx}(\tau)$ = Autocorrelation of signal $x(m)$.

$x(m)$ = m^{th} speech sample, and τ=sample lag.

Autocorrelation based fundamental frequency estimation algorithm is reliable and the pitch frequencies achieved are quite accurate. Table I shows the fundamental frequencies of 36 *swaras* in an ascending order. The 36 swaras came from 3 octaves namely lower, middle and upper. The table also list the popular name of *swara* and symbol, which is used for string detection algorithm. Each *swara* is represented by two characters, the first one indicating actual position in the octave and the second letter (L, M, U) representing the lower, middle or upper octave respectively. Position of 12 swaras in an octave in an ascending Order: S, r, R, g, G, M, m, P, d, D, n, N; where capital letters S, R, G, M, P, D, N indicate normal swaras, lower case letters 'r, g, d, n' indicate komal *swaras* and 'm' indicates teevra *swara*.

TABLE I: THE FUNDAMENTAL FREQUENCIES OF 36 *SWARAS* IN AN ASCENDING ORDER

Sr. no.	Octave	Swara Name	Symbol	Frequency (Hz)
1	*Lower*	*Shadja*	SL	130
2	*Lower*	*Komal Rishab*	rL	141
3	*Lower*	*Shuddha Rishab*	RL	148
4	*Lower*	*Komal Gandhar*	gL	158
5	*Lower*	*Shuddha Gandhar*	GL	167
6	*Lower*	*Shuddha Madhyam*	ML	176
7	*Lower*	*Tivra Madhyam*	mL	186
8	*Lower*	*Pancham*	PL	196
9	*Lower*	*Komal Dhaivat*	dL	209
10	*Lower*	*Shuddha Dhaivat*	DL	222
11	*Lower*	*Komal Nishad*	nL	235
12	*Lower*	*Shuddha Nishad*	NL	251
13	*Middle*	*Shadja*	SM	260
14	*Middle*	*Komal Rishab*	rM	280
15	*Middle*	*Shuddha Rishab*	RM	296
16	*Middle*	*Komal Gandhar*	gM	314
17	*Middle*	*Shuddha Gandhar*	GM	332
18	*Middle*	*Shuddha Madhyam*	MM	351
19	*Middle*	*Tivra Madhyam*	mM	372
20	*Middle*	*Pancham*	PM	394
21	*Middle*	*Komal Dhaivat*	dM	419
22	*Middle*	*Shuddha Dhaivat*	DM	443
23	*Middle*	*Komal Nishad*	nM	470
24	*Middle*	*Shuddha Nishad*	NM	500
25	*Upper*	*Shadja*	SU	520
26	*Upper*	*Komal Rishab*	rU	558
27	*Upper*	*Shuddha Rishab*	RU	590
28	*Upper*	*Komal Gandhar*	gU	626
29	*Upper*	*Shuddha Gandhar*	GU	664
30	*Upper*	*Shuddha Madhyam*	MU	702
31	*Upper*	*Tivra Madhyam*	mU	744
32	*Upper*	*Pancham*	PU	787
33	*Upper*	*Komal Dhaivat*	dU	837
34	*Upper*	*Shuddha Dhaivat*	DU	884
35	*Upper*	*Komal Nishad*	nU	941
36	*Upper*	*Shuddha Nishad*	NU	995

Fig. 3(a) shows a typical signal *x*(*m*), consisting of 12 *swaras* of middle octave in ascending order. The signal is then divided into frames and the corresponding pitch frequency of each frame is as shown in Fig. 3(b). The range of pitch frequencies for middle octave is from 260 Hz to 500 Hz.

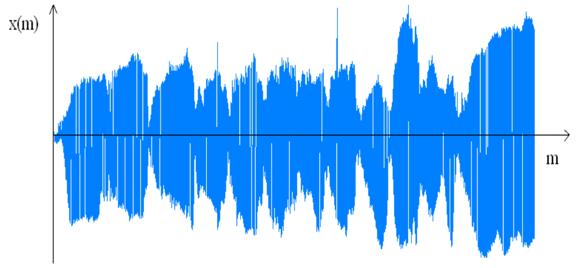

Fig. 3(a). Middle octave 12 swaras in ascending order.

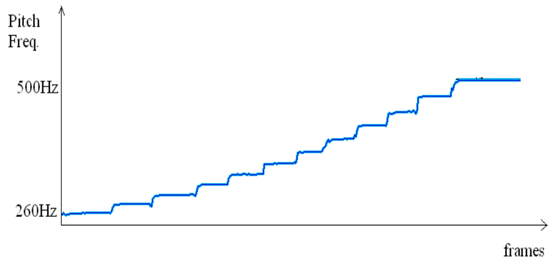

Fig. 3(b). Corresponding pitch frequencies of frames (Range of 260-500Hz).

The gradient function applied to the pitch frequency gives the proper onset detection peaks as shown in Fig.4.

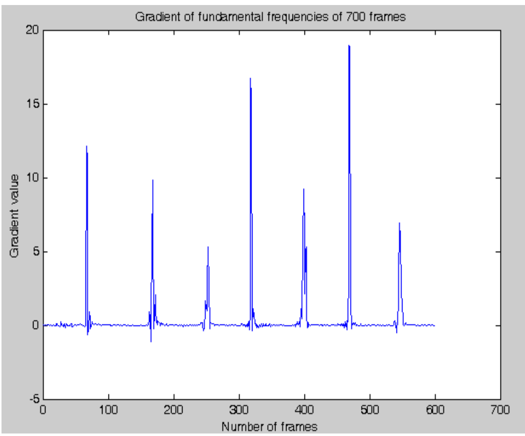

Fig. 4. Gradient of fundamental/pitch frequencies showing onset detection of 7 swaras.

Frame prior to the onset of next *swara* is considered as a last frame (end-point) of previous *swara*. Hence no separate end-point detection is necessary. The pitch frequencies thus obtained are comparable to those obtained by using Praat software [8]. As can be seen, sharp peaks are obtained which correspond to the onset of a new *swara*. This approach has very good accuracy and there is hardly any false detection. Hence, we decided to continue with this approach.

B. *Swara Identification*

As said earlier, in the feature extraction algorithm, pitch frequency of *swaras* is extracted. These frequencies are further used to identify each incoming musical note as a particular *swara*. For this, initially, a database is created which contains all the possible *swaras* with their corresponding values of pitch frequencies. The different pitch frequencies present in the audio file are mapped to the *swaras* according to prepared database [1]. The tolerance permitted for fundamental frequency of each *swara* is ±2Hz from the standard frequency given in table I. This tolerance range has given almost 100% accurate identification accuracy which is an essential part for highly accurate *raga* identification [3], [9].

C. *Raga Identification*

Every *raga* has two distinct patterns viz. the first pattern which is played one way going up (from *swaras* of lower frequencies to those of higher frequencies) called *'arohi'* and the other which is played one way going down called *'avarohi'*.

In live concerts, the entire *aarohi* or *avrohi* is very rarely presented; instead different rhythmic sequential combinations of the *swaras* from the *aarohi* or *avrohi* of that *raga* are played on the harmonium. These melodic combinations though apparently random, are always within the grammar restrictions of that particular *raga* with the result that some combinations are unique to that *raga* [2]. This concept is used for making the structure of *raga* which has as many numbers of templates as allowed by the grammar of that *raga*. Identification of *raga* problem is thus brought down to a simple template matching algorithm. In this algorithm dynamic programming method is used to compute the minimum edit distance between the input *swara* string and all the templates present in the *ragas* database structure as given in II (D).

D. *Database of Ragas*

A database is specially structured for eight different popular *ragas* viz. *'Bheempalas', 'Bageshri', 'Yaman', 'Bhoop', 'Durga', 'Des', 'Kafee' and 'Khamaj'* along with their identifying melodic combinations. A template is grammatically valid *swara* combination. A *raga* is a series of such templates. Popularly used *swara* combinations by the artists for the four *ragas* under investigation are listed below. Each template is separated by curly brackets, which represents a silence portion in the input audio. A *raga* structure used in *raga* identification algorithm is as shown below.

Raga Bheempalas :{{'nL'}, {'SM'}}, {{'gM'}, {'MM'}, {'PM'}, {'gM'}, {'MM'}}, {{'gM'}}, {{'RM'}, {'SM'}},{{'gM'}, {'RM'}, {'SM'}},{{'MM'}, {'PM'}} ,{{'gM'}, {'MM'} ,{'PM'}}, {{'MM'}}, {{'RM'}}, {{'SM'}}, {{'nL'} , {'SM'}, {'gM'}},{{'PM'} , {'nM'}},{{'DM'}, {'PM'}}, {{'nM'}, {'DM'}, {'PM'}}, {{'gM'}, {'MM'}}, {{'nL'}, {'SM'}, {'gM'}, {'MM'}}, {{'PM'}, {'nM'} , {'SU'}}, {{'nM'} , {'SU'},{'gU'},{'RU'},{'SU'}}

Raga Bageshri : {{'nL'}, {'DM'}, {'nL'}, {'SM'}} , {{'AB'}, {'nL'}, {'DL'}, {'nL'}, {'RM'}, {'nM'}, {'SM'} }, {{ 'DM'}, {'nL'} , {'SM'}, {'gM'} ,{ 'RM'}, {'SM'}} , {{ 'SM'}, {'gM'}, {'gM'}, {'MM'}} , {{ 'MM'}, {'gM'}, {'RM'}, {'SM'}, {'nM'}, {'SM'}} , {{'SM'}, {'gM'}, {'MM'}, {'DM'}} , {{'GM'}, {'MM'}, {'DM'}} , {{'MM'}, {'DM'}, {'MM'}, {'DM'}} , {{'gM'}, {'DM'}, {'gM'}, {'MM'}} , {{'SM'}, {'gM'}, {'MM'}, {'gM'}, {'RM'}, {'SM'}} , {{'gM'}, {'MM'}, {'DM'}, {'nM'}, {'nM'}, {'DM'}} , {{'MM'}, {'nM'}, {'DM'}} , {{'gM'}, {'DM'}, {'MM'}} , {{ 'SM'}, {'MM'}, {'gM'}, {'RM'}, {'SM'}} , {{'gM'}, {'MM'}, {'DM'}, {'nM'}, {'SU'}} , {{'MM'}, {'DM'}, {'nM'}, {'SU'}} , {{'SU'}, {'nM'}, {'RU'}, {'SU'}, {'SU'}, {'nM'} , {'DM'}, {'MM'} }, {{'CD'}} , {{ 'gM'}, {'MM'}} ,{{'gM'}, {'RM'}, {'SM'}}

Raga Yaman : {{'SM'}, {'NL'}, {'DL'}, {'NL'}, {'SM'}} , {{'NL'}, {'RM'}, {'SM'}} , {{'NL'}, {'RM'}, {'GM'}} , {{'RM'}, {'GM'}} ,

{{'NL'}, {'RM'}} , {{'DL'}, {'NL'}, {'RM'}, {'SM'}} , {{'NL'}, {'RM'}, {'GM'}} , {{'mM'}, {'GM'}} , {{'RM'}, {'mM'}} , {{'RM'}, {'GM'}} , {{'RM'}, {'SM'}, {'NM'}, {'RM'}, {'GM'}, {'mM'}, {'PM'}} , {{'EF'}} , {{'mM'}, {'GM'}} , {{'RM'}, {'GM'}, {'RM'}, {'mM'}, {'GM'}} , {{'RM'}} , {{'NL'}, {'GM'}, {'RM'} , {'SM'}} ,{{'PM'}, {'mM'}, {'GM'}, {'mM'}, {'DM'}, {'PM'}} , {{'GM'}, {'mM'}, {'DM'}} , {{'mM'}, {'DM'}, {'NM'}} , {{'DM'}, {'PM'}} , {{'mM'}, {'GM'}} , {{'RM'}, {'GM'}} , {{'NL'}, {'RM'}} , {{'DL'}, {'NL'}, {'RM'}, {'SM'}} , {{'GM'}, {'mM'}, {'DM'}, {'NM'}, {'SU'}} , {{'GH'}, {'NM'}} , {{'DM'}, {'NM'}, {'DM'}, {'PM'}} , {{'mM'}, {'GM'}} , {{'mM'}, {'RM'}, {'GM'}} , {{'RM'}, {'SM'}}

Raga Bhoop : {{'SM'}, {'DL'}, {'SM'}} , {{'DL'}, {'PL'}, {'DL'}, {'DL'}, {'SM'}} , {{ 'SM'}, {'RM'}, {'SM'}} , {{'AB'}, {'DL'}, {'DL'}} , {{'SM'}, {'RM'}} , {{'RM'}, {'RM'}, {'RM'}, {'GM'}, {'RM'}, {'SM'}} , {{'DL'}, {'DL'}, {'PL'}, {'DL'}, {'DL'}, {'SM'}} , {{'SM'}, {'RM'}, {'GM'}, {'PM'}} , {{'GM'}, {'RM'}, {'GM'}, {'PM'}} , {{'IJ'}}, {{'KL'}}, {{'MN'}} , {{'OP'}}, {{'QR'}}, {{'MN'}}, {{'PM'}, {'GM'}} , {{'UV'}} , {{'KL'}} , {{'OP'}} , {{'WX'}}, {{'PM'}, {'GM'}} , {{'SM'}, {'RM'}, {'GM'}, {'PM'}, {'DM'}} , {{'PM'}, {'GM'}} , {{'RM'}, {'SM'}} , {{'DL'}, {'PL'} , {'DL'}, {'SM'}} , {{'SM'}, {'RM'}, {'GM'}, {'PM'}, {'DM'}, {'SU'}} , {{'GH'}, {'DM'} ,{'PM'}} , {{'DM'}, {'SU'}} , {{'DM'}, {'PM'}, {'GM'}} , {{'WX'}} , {{'YZ'}}, {{'BA'}}, {{'SM'}, {'RM'}} , {{'DL'}, {'SM'}}

Raga Durga : {{'SM'} , {'RM'} , {'MM'}} , {{'MM'} , {'RM'} , {'SM'}} , {{'AB'} , {'DL'} , {'SM'}} , {{'DL'} , {'SM'} , {'RM'}} , {{'MM'}} , {{'RM'} , {'MM'} , {'PM'}} , {{'MM'} , {'PM'} , {'DM'} , {'MM'} , {'RM'} , {'DL'} , {'SM'}} , {{'SM'} , {'RM'} , {'MM'} , {'PM'}} , {{'RM'} , {'MM'} , {'PM'} , {'DM'}} , {{'MM'}} , {{'RM'}} , {{'MM'}} , {{'RM'} , {'SM'}} , {{'DL'} , {'SM'}} , {{'SM'} , {'RM'} , {'MM'} , {'PM'} , {'DM'} , {'SU'}} , {{'GH'} , {'DM'} , {'PM'}} , {{'MM'} , {'PM'} , {'DM'} , {'MM'} , {'RM'}} , {{'SM'} , {'RM'} , {'SM'} , {'dM'} , {'SM'}}

Raga Des : {{'NL'} , {'SM'}} , {{'NL'} , {'SM'}} , {{'NL'} , {'SM'} , {'RM'}} , {{'NL'} , {'SM'} , {'RM'} , {'MM'} , {'GM'} , {'RM'}} , {{'GM'} , {'NL'} , {'SM'}} , {{'RM'} , {'MM'} , {'PM'} , {'NM'} , {'SU'}} , {{'GH'} , {'Nm'} , {'DM'} , {'PM'}} , {{'MM'} , {'GM'} , {'RM'}} , {{'GM'} , {'NL'} , {'SM'}}

Raga Kafee : {{'SM'} , {'nL'} , {'DL'} , {'PL'} , {'DL'} , {'nL'} , {'SM'}} , {{'RM'} , {'gM'} , {'RM'} , {'SM'}} , {{'SM'} , {'RM'} , {'gM'} , {'RM'} , {'gM'} , {'SM'} , {'RM'} , {'nL'} , {'SM'}} , {{'SM'} , {'RM'} , {'gM'} , {'MM'} , {'gM'} , {'RM'} , {'SM'}} , {{'MM'} , {'gM'} , {'RM'} , {'gM'} , {'MM'}} , {{'SM'} , {'RM'}} , {{'RM'} , {'gM'}} , {{'gM'} , {'MM'}} , {{'MM'} , {'PM'} , {'gM'} , {'RM'} , {'SM'}} , {{'SM'} , {'RM'} , {'gM'}} , {{'RM'} , {'gM'} , {'MM'}} , {{'gM'} , {'MM'} , {'PM'}} , {{'MM'} , {'PM'} , {'DM'} , {'PM'}} , {{'DM'} , {'PM'}} , {{'AZ'}, {'gM'} , {'RM'}} , {{'BY'}} , {{'CD'}} , {{'PM'} , {'MM'} , {'gM'} , {'RM'}} , {{'RM'} , {'gM'}} , {{'SM'} , {'RM'}} , {{'nL'} , {'SM'}}

Raga Khamaj : {{'SM'} , {'GM'} , {'MM'} , {'PM'}} , {{'GM'} , {'MM'} , {'GM'}} , {{'SM'} , {'GM'} , {'MM'} , {'GM'}} , {{'RM'} , {'SM'}} , {{'GM'} , {'MM'} , {'PM'} , {'DM'} , {'nM'} , {'DM'} , {'PM'}} , {{'SM'} , {'GM'} , {'MM'}} , {{'GM'} , {'MM'} , {'PM'}} , {{'MM'} , {'PM'} , {'DM'}} , {{'PM'} , {'DM'} , {'nM'} , {'DM'} , {'PM'}} , {{'GM'} , {'MM'} , {'GM'}} , {{'RM'} , {'SM'}} , {{'NL'} , {'SM'} , {'GM'} , {'MM'} , {'PM'}} , {{'GM'} , {'MM'} , {'PM'} , {'DM'} , {'NM'} , {'SU'}} , {{'nM'} , {'DM'} , {'PM'} , {'MM'} , {'GM'} , {'RM'} , {'SM'}}}} ;

III. Experimentation and Performance Analysis

From the input recordings of the harmonium the individual *swara* is identified by using pitch frequency as a feature. A string of *swaras* (between to pauses) is considered as an input template. In this way the subsequent input swaras are grouped into templates which then are checked against the templates available in database. For this, dynamic programming is used. The following algorithm shows the *raga* identification process.

- All the input templates provided by the user are compared with each *swara* template of all the *ragas* in the *raga* structure.
- A confusion matrix is estimated for each input template and each template of different r*agas* in the database.
- From this a minimum string edit distance is obtained for each comparison.
- The distance obtained for all templates of each *raga* are summed and the *raga* for which minimum value of summation is obtained is the recognized *raga*.

Estimated *swara* sequence:

'SM','MM','gM','RM','SM','nL','SM','MM','gM','MM','PM',
'MM','nM','DM','PM','MM','SM','MM','gM','RM','nL','SM',
'gM','MM','PM','MM','nM','DM','PM','MM','PM','ML','gL',
'nM','PM','SU','nM','SU','nM','SU','RU','gL','SU','nM','DM',
'PM','MM','PM','SU','nM','SU','nM','DM','PM','MM','nM',
'DM','PM','gM','MM','RM','SM','MM','gM','RM','nL','SM',
'gM','MM','PM','MM','nM','DM','PM','MM','PM','gM','MM',
'PM','gM','MM','PM','nM','PM','nM','SU','nM','SU','MM','M
U','gU','RU','nM','SU','PM','nM','SU','RU','SU','nM','DM','P
M',
'MM','PM','gL','MM','PM','nM','SU','nM','DM','PM','MM',
'gM','RM','SM','MM','gM','RM','nL','SM','gM','MM','PM',
'MM','nM','DM','PM','MM','gM','RM','SM','nL','SM','gM',
'RM','SM','MM','GM','MM'

Templates formed :

Here two templates are separated by semicolon.

'SM';'MM';'gM','RM','SM';'nL','SM';'MM';'gM','MM','PM';
'MM';'nM','DM','PM';'MM';'SM';'MM';'gM';'RM';'nL','SM',
'gM','MM';'PM','MM','nM';'DM','PM';'MM','PM';'ML';'gL',
'nM';'PM','SU','nM','SU';'nM','SU';'RU';'gL','SU','nM','DM',
'PM';'MM','PM';'SU','nM','SU';'nM','DM','PM';'MM';'nM',
'DM','PM';'gM','MM';'RM','SM';'MM';'gM';'RM';'nL','SM',
'gM','MM';'PM','MM','nM';'DM','PM';'MM','PM';'gM','MM',
'PM','gM','MM';'PM','nM';'PM','nM','SU';'nM','SU';'MM';'M
U';'gU';'RU','nM','SU';'PM','nM','SU';'RU','SU','nM','DM','P
M';
'MM','PM';'gL','MM','PM';'nM','SU';'nM','DM','PM';'MM';
'gM','RM','SM';'MM';'gM';'RM';'nL','SM','gM','MM';'PM',
'MM','nM';'DM','PM';'MM';'gM','RM','SM';'nL','SM','gM';
'RM','SM';'MM';'GM','MM';

Corresponding edit distance from each *raga*:

Raga Bheempalas: 20
Raga Bageshri: 64
Raga Yaman: 87
Raga Bhoop: 79
Raga Durga: 69
Raga Des: 15
Raga Kafee: 57
Raga Khamaj: 73

Minimum edit distance : 20
Recognised Raaga: 'Raag Bheempalas'

IV. Conclusion and Future Work

The designed system using fundamental-frequency as a feature for *swara* identification is fairly accurate within a tolerance band of ±2 Hz. The estimation of onset detection of *swara* is very precise giving a 100% identification accuracy which encouraged the further research for sequence of *swara* (*raga*) identification. A dynamic programming based approach for *raga* identification gives 100% results for the database of four *ragas* which consist of 25 pre-recorded stand-alone harmonium performances of each *raga* played on various harmoniums by different artists. The system performance is evaluated for normal harmonium performance that is around 10-12 *swaras*/second.

The current algorithm for *raga* identification does not assign any priorities to the templates in each *raga* in the database. A method to give preference or weighted probability to the most popularly used combinations in a *raga* can be developed in further research. As the computational time requirement for template matching program is proportional to number of templates in database a synchronous search algorithm will be developed which will process all the templates simultaneously which will make it independent of number of templates in the database and hence enabling a real-time applications.

In future, this system may also be used for more generic purposes by using a filtering mechanism initially which will filter out only the audio signals pertaining to the harmonium from a medley of signals normally present in an Indian Classical orchestra.

REFERENCES

[1] R. Sridhar and T.V. Geetha, "Swara identification for South Indian classical music," in Proc. *9th International Conf. on Information Technology*, December 2006, pp. 143-144.

[2] S. Belle, R. Joshi, and P. Rao, "Raga Identification by using Swara Intonation," *Journal of ITC Sangeet Research Academy*, vol. 23, December 2009.

[3] R. Sudha, "A tool for identifying ragas using MIDI (musical instrument devices) for CMIR (classical music information retrieval)," in *Proc. Eighth IEEE International Conf. on Dependable, Autonomic and Secure Computing*, 12-14 December, 2009, pp. 837-842.

[4] G. Velikic, E. L. Titlebaum, and M. F. Bocko, "Musical note segmentation employing combined time and frequency analyses," in *Proc. IEEE International Conf. on Acoustics, Speech and signal processing*, 17-21 May 2004, vol. 4 , pp. 277-280.

[5] H. L. Tan, Y. W. Zhu, L. Chaisorn, and S. Rahardja, "Audio onset detection using energy-based and pitch-based processing," in *Proc. IEEE International Symposium on Circuits and systems*, May 30-June 2, 2010, pp. 3689–3692.

[6] M. Marolt, A. Kavcic, M. Privosnik, and S. Divjak, "On detecting note onsets in piano music," *IEEE MELECON 2002*, pp. 385–389, May 7-9, 2002.

[7] N. Degara, A. Pena, M. E. P. Davies, and M. D. Plumbley, "Note onset detection using rhythmic structure," *Acoustics Speech and Signal Processing*, pp. 5526-5529, March 14-19, 2010.

[8] A website of audio processing software Praat. [Online]. Available: http://www.fon.hum.uva.nl/praat/

[9] R. Sudha, A. Kathirvel, and RMD. Sundaram, "A system tool for identifying ragas using MIDI," in *Proc. Second International Conf. on Computer and Electrical Engineering*, December 28-30, 2009, vol. 2, pp. 644–647.

Rajshri O. Pendekar was born in 1990 in Penda. She is graduated from College of Engineering, Pune. She obtained the bachelor's degree of technology (B.Tech) in Electronics and Telecommunication in 2011. She joined FinIQ Consulting Pvt. Ltd., Pune as a graduate trainee after completion of graduation and currently she is working on modeling and development of financial products. Her areas of interest are signal processing, robotics etc.

Shrinivas P. Mahajan was born on April 15, 1971. He received bachelor's degree in Electronics and Telecommunication Engineering in 1991 from the Marathwada University Aurangabad. He received masters and ph.D. degree in Electronics and Telecommunication Engineering from Pune University in 1996 and 2011, respectively. He is working as an associate professor in College of Engineering, Pune. He has teaching experience more than 21 years. He is working as post graduate teacher since last 12 years. His research interests are in the area of digital signal processing, speech processing and pattern classification.

Rasika Mujumdar was born in 1988. She is graduated from Cummins College of Engineering, Pune with bachelor's degree of technology (B.Tech) in Electronics and Telecommunication in 2011. She received her master's degree in Electronics and telecommunication from College of Engineering, Pune in 2011.

Pranjali Ganoo was born in 1989 in Pune). She is graduated from College of Engineering, Pune with the bachelor's degree of technology (B.Tech) in Electronics and Telecommunication in 2011. She is currently studying in Dartmouth College and will complete her master's degree in Dec. 2013.