A METHODOLOGY OF NOTE EXTRACTION FROM THE SONG SIGNALS

A. K. Datta , R. Sengupta, N. Dey, D. Nag and A. Mukerjee

Scientific Research Department, ITC Sangeet Research Academy,
1, N S C Bose Road, Tollygunge, Kolkata 700 040
E.Mail: **srasrd@vsnl.net**

ABSTRACT

The paper presents an approach to automatic transcription of Hindustani vocal music as a transformation of acoustic signal representing singing to symbols representing notes. Algorithms are discussed that concern five sub problems. (i) Extraction, (ii) Smoothing and (iii) Segmentation of Pitch Profile, (iv) Tonic (Sa) Measurement and (v) Note Extraction.

In Indian classical singing the mapping of the F_0 sequences onto the notes is by no means a simple task. The F_0 profiles representing notes are not always flat and it may be quite problematic to determine when a profile represents a note and when they are really glides. Furthermore as the note perception is categorical and the notes are, in essence, learnt intervals, this mapping is not a trivial problem.

This paper presents an approach for annotation of aalap in north Indian classical vocal singing without using any musicological information except that of ratios representing notes. As the aalap portions are usually non-metrical, the analysis for determination of meters is not undertaken. Thus the annotation merely consists of detecting notes and their corresponding durations. Musicological information is used to verify the notations. 96% notes are correctly identified.

INTRODUCTION

Research in the realm of Indian classical music has long existence since ancient times. However, objective investigations in Indian music have fallen far behind those in western music. The present state of science and technology can provide ample scope for quantitative as well as qualitative approach to investigate srutis, swaras, intervals, octaves (*saptak*), consonance (*vaditya*), musical quality, rhythm etc. [1]. The performance of the 'ragas' in Indian Classical music is based on both explicit knowledge (which forms the so called grammar and is transmitted through the tradition of oral teaching and also through documented analysis e.g. old Shastra's, Bhatkhande's works etc.) and non-explicit knowledge which is born out of the musician's sensibilities and normally defies identification, categorisation etc. in the manual realm.

Written music is traditionally presented as a *musical notation* (score), which comprises the times, durations and notes of the sounds that constitute a piece. The aim of music transcription is to discover this musical notation embedded in an acoustic signal, so that a musician or a synthesizer program can reproduce and modify the original performance. Musical notation of a piece of composition further allows one to take recourse to computational analysis for understanding of various syntactic and semantic contents embedded in the piece. Analysis of musical structure of a composition with respect to its influence on human mind is an interesting topic. For this purpose one may follow the paradigm of similar research in language. The analysis of syntax and semantics play an important role in it. Drawing on this similarity one can see that role played by the notes in a music composition plays the roles that letters play in languages. In Indian classical singing syntax may be understood only in terms of notes, sometimes probably srutis, not absolute values of pitch. While notes may be considered as a mapping of the F_0 sequences, this is by no means a simple one. The F_0 profiles representing notes are not always flat and it may be quite problematic to determine when a profile represents a note and when they are really glides. Furthermore as the note perception is categorical and the notes are, in essence, learnt intervals this mapping is not a trivial problem.

Automatic transcription of music (A2M) is difficult and includes several sub problems. Figure 1 gives an idea of the parts of an automatic transcriber. F_0 estimation refers to a core part, which estimates the fundamental frequencies of the music signal. Usually, internal models are needed along with the acoustic signal to perform the analysis. Musical meter (a measure of rhythm) characterizes the temporal regularity of a music signal. Musical meter is a hierarchical structure, which consists of pulse sensations (periodicities) at different levels [2]. Generally F_0 estimation is preceded by an estimator of musical meter. This is shown in gray back ground because its details are not addressed in the present paper. Comparing with the role of a “language model” in speech recognition, musicological models play the same role in music.

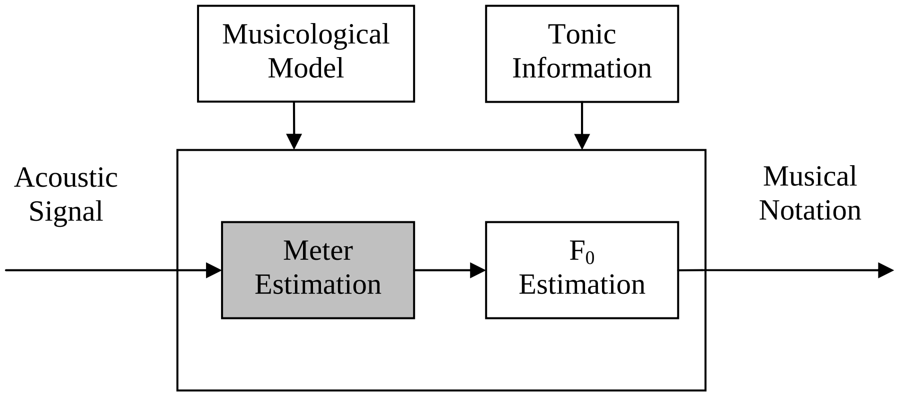

Figure 1. A General Schema for Music Annotation

Musicological information is very important for music transcription. This information could be readily used as a post-processor for a transcription system. However, to really benefit from the musical knowledge, it should be utilized during the analysis. Also, while a rule-based model reveals a lot about the human cognition, sometimes probabilistic models are advantageous in that they evaluate the likelihoods of several candidates.

The extraction of notes from the pitch profile needs the knowledge of the tonic used by the singer.

This paper presents an approach for annotation of aalap in north Indian classical vocal singing without using any musicological information except ratio information. As the aalap portions are usually non-metrical, the analysis for determination of meters is not undertaken. Thus the annotation merely consists of detecting the sequence of notes and their corresponding durations. Musicological information is used to verify the notations. 96% notes are correctly identified.

METHOD OF ANALYSIS

A. Pitch Period Extraction from Signal

A method based on Phase-Space Analysis (PSA) was used for extracting pitch periods. PSA [3] uses the basic fact that, for a repetitive function, two points of the signal having a phase difference of 2π have the same displacement and, therefore, if plotted in a 2-D phase diagram would be a straight line with a slope of $\pi/4$. This allows a clear threshold and algorithmic logic for pitch extraction. Henceforth the pitch pattern files, extracted using the above method, will be referred to as ‘.cep file’.

B. Smoothing

The cep file contains pitch extracted only in the quasi-periodic region of the signal. The PD algorithm above uses a predefined specific range for possible pitch values e.g., in the present case the default range is fixed between 70 Hz to 700 Hz. Three types of error were observed. One is that the determined pitch value is approximately equal to half or double of the actual pitch value. This error is typical for PSA and is usually a rare and isolate occurrence. The other is pitch occurring outside the range. Third one is spike in pitch sequence. The smoothing operation for the first type of error is simply detection of it and doubling or halving the value as required. The two other types of error are detected through an examination of local pitch values and the erroneous values are replaced using linear interpolation from the neighboring valid data.

C. Segmentation of Pitch Profile

Even in a perfectly perceptible steady note all the pitch values are never identical. There are always some involuntary small variations in pitch, which are irrelevant to the perception of the corresponding

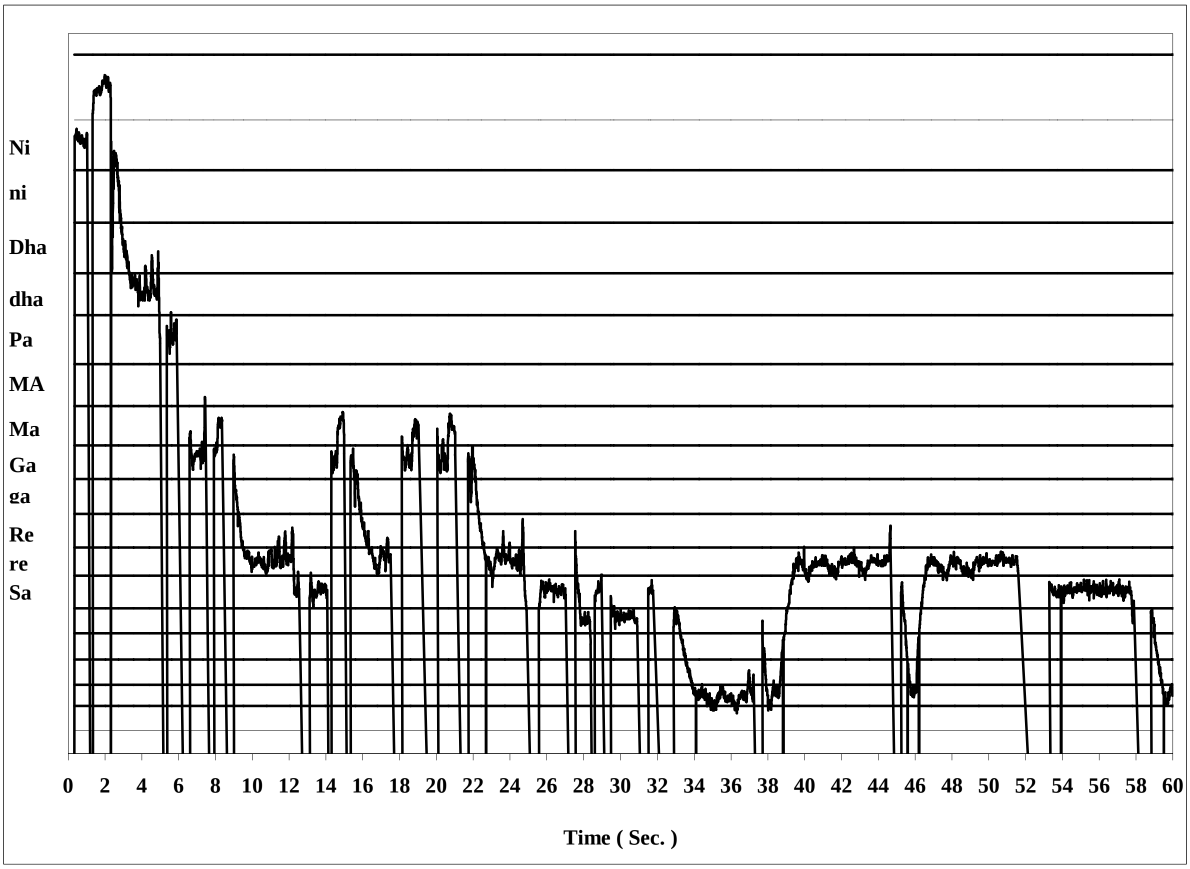

Figure 2. Example of a Pitch Profile along with Interval Boundaries

note whereas there are variations of pitch which are perceived as glissando. Thus the pitch profile is to be pre-segmented for facilitating the detection of actual notes from pitch movements. These segmented pitch sequences are put in .std files. Figure 2 shows a portion of 60 sec. duration of one .std file of raga Bhairav along with interval boundaries of notes corresponding to the tonic of the singer.

D. Tonic (Sa) Measurement

A skilled musician was requested to listen to the signal files one after another to detect the position of Sa in the file. By frequency domain analysis at the detected region, Sa for the respective signal file was calculated. All these Sa values (in Hz.) are put into a file, ‘safile.txt’.

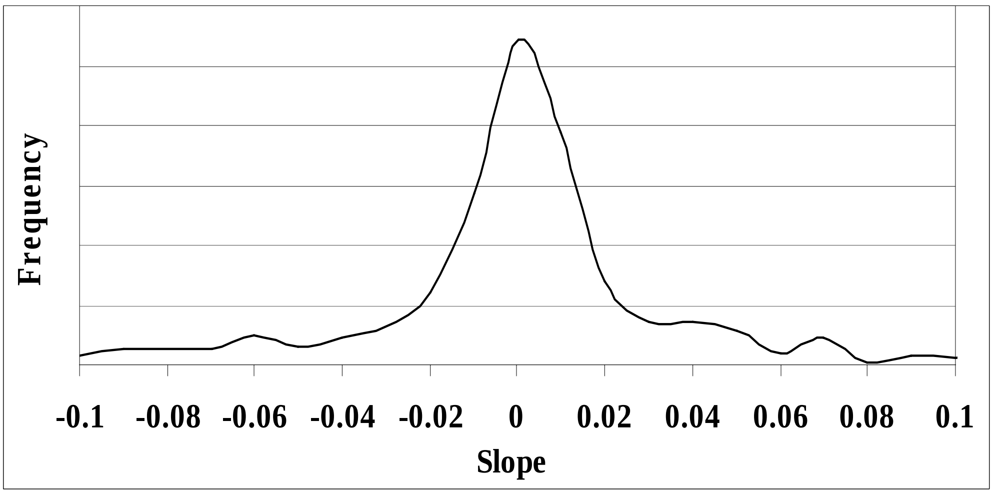

Figure 3. Frequency Distribution of Slopes of Pitch Sequences

E. Note Extraction

The pitch values in the sequences in .std files are divided by the measured value of the tonic for each song before deciding the note of the sequence. These sequences may need to be subdivided further depending on the condition whether the whole sequence is in one note interval [4] or more. For each of the above-mentioned subsequences or sequences, as the case may be, the slope of best-fitted line is calculated. If the slope be beyond a certain range (+0.02 to -0.02) or the duration of the sequence is less than a threshold value (0.1 sec. arbitrarily decided as the maximum duration of a touch note) the sequence is rejected otherwise the sequence is given the corresponding note name. The threshold for the slope is determined from the frequency distribution of slopes of the note sequences (Figure 3). Extracted notes, their respective start and end time in milliseconds are presented in Table 1 for the pitch profile shown in figure 2. Bold entry in the table indicates incorrect notes with respect to the raga (Bhairav) rendered. A single prime sign after the note name has been used for the note of lower octave and double prime for upper octave.

Notes	**Time (mSec.)**		**Notes**	**Time (mSec.)**		**Notes**	**Time (mSec.)**		**Notes**	**Time (mSec.)**	
	Start	**End**		**Start**	**End**		**Start**	**End**		**Start**	**End**
Ni	339	1030	Ga	14287	14616	Sa	25559	27021	**Dha'**	**36919**	**37026**
Sa"	1318	2312	Ma	14616	14988	Ni'	27793	28362	dha'	37026	37204
Ni	2466	2660	Ga	15604	15720	Sa	28601	28954	dha'	38051	38156
dha	3312	4850	**Re**	**15904**	**16324**	Ni'	29661	30913	dha'	38326	38657
Pa	5348	5884	re	16324	17500	Sa	31503	31767	re	39413	39973
Ga	6662	7380	Ga	18147	18685	Ni'	32947	33219	re	40236	44621
Ga	7904	8037	Ma	18685	19028	**Dha'**	**33547**	**33861**	dha'	45735	45981
Ma	8095	8350	Ga	20058	20623	dha'	33903	34857	re	46575	49051
re	9552	12137	Ma	20623	21017	dha'	35217	36195	re	49156	51597
Sa	12249	12472	**ga**	**21819**	**21921**	Pa'	36248	36363	Sa	53283	57784
Sa	13184	14065	re	22704	24660	dha'	36363	36919	dha'	59513	60137

Table 1: Notes with Start & End Time of the Pitch Profile shown in Figure 2.

EXPERIMENTAL DETAILS

Thirty seven singers (thirty male and seven female) of Hindustani music were asked to render four ragas namely Bhairav (That –Bhairav), Darbari Kannada (That – Asavari), Mian-ki-Malhar (That – Kafi) and Todi (That – Todi), which included aalap, vistar, taan and gamaka. In most cases the notes extended on both sides of the middle octave. Direct digital recording was done in a noise proof studio having a reverberation time of 0.1 sec. via standard sound card (full Duplex PnP) in Pentium-IV PC (2.6GHz.) to avoid possible phase distortions present in most of the analog recording [5]. The digitisation of the signal was done at the rate of 22050 samples /sec (16 bits/sample). Only the voice of the singer was recorded and the accompanying instruments were excluded. For our analysis only the aalap part of each singer was selected from each raga. Pieces of aalap for each singer for a raga were taken out from the complete aalap deleting the bandish part. These constituted the aalap signal files for a singer for each raga. For each raga a singer had one aalap signal of ~2 to 3 minutes. Total 111 aalap signal files (90 for male and 21 for female) were thus selected for analysis, which constituted our database.

Pitch periods were extracted by using the PD algorithm mentioned in an earlier section. In the present study our requirement was to determine time periods for individual glottal cycle in the continuum of a quasi-periodic speech signal. This process on a wav file produces a .cep file, which contains pitch values and the corresponding time in the wav file. As already mentioned this file may contain erroneous pitch values, which are subsequently removed through the smoothing operation outlined in the last section. Even after all these corrections there may still be some errors which reveals themselves in the contour as sharp local fluctuations. We presume these to be irrelevant in the perception of tonality or pitch movement. To remove these we simply replace the $(i+1)^{th}$ pitch x_{i+1} by the i^{th} pitch x_i when $|(x_{i+1} – x_i)| > x_i *0.1$ & $x_{i+1}>0$. This creates a p-file. From p-file sub-sequences are created with all consecutive pitch in a sequence, which is terminated when $| x_{i+1} – M | > M/30$ where

$M = (1/i)\ \Sigma\ x_i$. If the duration of any sequence be less than 100 msec. then the sequence is rejected. These create total 111 .std files, which constitute our database for analysis.

Next part of the work was to find tonic for individual audio sample file. This has been done by the help of an expert musician. The musician first detected the correct position of tonic in the file by carefully listening to the song. Then by FFT analysis at the detected position of the signal using a standard software package, fundamental frequency was measured. This has been done for all of 111 .wav files and the results were put into a text file, safile.txt.

Segmentation of pitch sequences in .std files has been done in the following way. Pitch values are divided by the respective tonic of the file and the ratio obtained was folded back to middle octave (value between 1 & 2), if necessary. Then this value was accorded a note position after comparing with the theoretical intervals of notes obtained from scale information and interval boundary concept described earlier. This was done using total 8 systems of 12-note ratios and 8 system of 22-sruti ratios resulting in total 16 sets of outcomes. If in a sequence a pitch period belongs to a note interval different from its predecessor, a new sequence is considered to start. Using this criterion, the sequences of pitch periods in .std files were broken into subsequences, wherever necessary. Duration and slope of best-fit straight line of each sequence/subsequence, obtained thus, was calculated. If duration is less than 100 milliseconds or the value of slope is more than 0.02, the sequence was rejected and marked as non-note.

Using restricted grammar of all the four ragas, notes detected were classified into ‘Correct’ and ‘Incorrect’ categories. Total time of all correct, incorrect and non-notes were found out. Percentages of correct/incorrect notes were calculated with respect to total duration of all notes detected.

RESULTS AND DISCUSSIONS

Table 3 shows the distribution of time in milliseconds for all the sung notes, as detected, in 111 songs under analysis along with the non-note sequences using interval ratios under different available schemes. Total time of singing is 14085794 milliseconds. The correctness of the notes for the table is assessed by using the restricted grammar [6] given in Table 2 which shows the allowed notes for a

Raga	**Suddha Swara**	**Vikrit Swara**
Bhairav	Sa, Ga, Ma, Pa, Ni	re, dha
Darbari Kannada	Sa, Re, Ma, Pa	ga, dha, ni
Mian-ki-Mallhar	Sa, Re, Ma, Pa, Dha, Ni	ga, ni
Todi	Sa, Pa, Ni	re, ga, MA, dha

Table 2. Raga-wise swara position for the restricted grammar

12 note ratio system	**Note**		**Non-note**	**Percentage of notes**	
	Correct	**Incorrect**		**Correct**	**Incorrect**
Altered Consonance[7]	9873981	488015	3723798	95.29033	4.709667
Barveg[8]	9622115	488386	3975293	95.16951	4.830487
Bhandarkar[8]	9944851	472272	3668671	95.46639	4.533615
Equi-tempered [7]	9964139	474608	3647046	95.45339	4.546605
Experimentally derived ratio system [9]	9939621	483014	3663158	95.36572	4.634285
Parijata[8]	9661359	467976	3956459	95.37999	4.62001
Western compilation[8]	9960713	461284	3663797	95.57394	4.426063
22 note ratio system					
Altered Consonance[7]	9253866	617433.6	4214495	93.74516	6.254837
Consonance[7]	9333556	598677	4153561	93.97238	6.027618
Deval[8]	9920911	479748.9	3685134	95.38732	4.612677
Equi-tempered[7]	9283166	598335.1	4204293	93.9449	6.055104
Nagoji Row[8]	9723801	532361.8	3829631	94.80935	5.190653
Experimentally derived ratio system[10]	9682658	575271.9	3827865	94.39193	5.60807
Western compilation[8]	9914703	493534.5	3677557	95.25823	4.741768

Table 3. Result of analysis of labeled notes

particular raga.

The difference between the 12-ratio system and 22-ratio system is that in the later scheme the subsequences are first given sruti values and then clubbed in note intervals. As expected the difference between the results of the two groups is not very significant. The data in column ‘non-notes’ represent the time of all those sequences, which are deemed to be either glissando or touch notes. The data in column ‘Incorrect notes’ gives the total time of the subsequences, which have note labels not conforming to the grammar of the raga. The score around 96% is considered significant. However some of the errors seem to be due to oscillation of pitch around a note, which is perceived in a different manner of cognition.

Table 4 shows example of annotation of two songs each from the four ragas. The wrong notes are indicated in bold letters. Rejected sequences/subsequences resulting in non-notes occurring at various places of the pitch profile are not shown in the table. Successive notes, after rejecting non-notes, are presented row-wise and are to be read left to right and top down.

Bhairav 1

Ni	Sa"	Ni	dha	Pa	Ga	Ga	Ma	re	Sa	Sa	Ga	Ma	Ga	**Re**	re	Ga	Ma
Ga	Ma	**ga**	re	Sa	Ni'	Sa	Ni'	Sa	Ni'	**Dha'**	dha'	dha'	Pa'	dha'	**Dha'**	dha'	dha'
dha'	re	re	dha'	re	re	Sa	dha'	Ni'	Sa	Ga	Ma	re	re	re	Sa	dha	dha
dha	Ma	Pa	Pa	Ga	Ma	dha	dha	Ma	Ma	Pa	Ga	Ga	Ga	re	Ga	re	re
Ma	Ga	Pa	Pa	dha	Ma	Pa	Ga	Ga	Ma	Ga	Pa	Ga	Ma	Ga	re	re	Sa
Sa	Ga	Ma	Sa	dha	Pa	dha	dha	Ma	Pa	Ma	dha	Pa	Pa	Pa	dha	Ma	Pa
Ga	re	re	Ga	Ma	Ga	Ma	Ga	re	**Re**	re	Sa	Sa					

Bhairav 2

Sa	re	Ga	Ma	Ma	Ga	Ga	Pa	Ma	Ma	Pa	Ga	Ga	Ma	re	re	Sa	Ni'
dha'	dha'	dha'	Sa	Sa	re	Pa'	Sa	Ni'	Sa	re	Sa	**Dha'**	dha'	Pa'	**MA'**	Pa'	dha'
dha'	dha'	Pa'	Ni'	Sa	Ni'	Ni'	Sa	Sa	re	Ga	re	Sa	re	Sa	Ga	Ga	Ga
Ga	Ga	Ga	re	re	Ma	Ga	Ma	Ga	Ga	Ma	Ga	Ga	Pa	Ma	Ga	Ma	Ga
Pa	Ga	Ma	Ga	Ga	Ma	re	Sa	Ga	Ga	re	Sa	Sa	Sa	Ma	Ga	Ma	Ga
Ma	Ga	Pa	Pa	Ga	Ma	Ga	Ma	Ga	re	Ma	Ga	Ma	dha	Pa	Ga	Ma	Ga
Ma	dha	Ma	Ma	Pa	Ga	Ga	Ma	Pa	dha	Ma	Ma	Ga	Ga	Ma	re	re	Ga
re	Sa	Sa	Ni'	Sa	Ni'	Ma	dha	dha	Pa	dha	dha	dha	dha	Ni	dha	dha	Pa
Pa	Ma	dha	Pa	dha	Ni	Ni	dha	Sa"	Sa"	Ma	Pa	dha	Ni	Sa"	Pa	Ni	re"
Sa"	Sa"	Ni	Sa"	dha													

Darbari Kannada 1

Sa"	Sa	Sa	ni'	**Dha**'	dha'	**Dha**'	dha'	**Dha**'	dha'	ni'	Pa'	**Dha**'	dha'	**Dha**'	dha'	**Dha**'	dha'
ni'	ni'	Re	dha'	ni'	Re	ni'	ni'	Sa	ni'	Sa	Re	Re	Sa	ni'	ni'	Sa	ni'
Sa	Re	Re	Sa	Re	**Ga**	ga	**Ga**	ga	ga	Ma	Sa	Re	Sa	Sa	**Dha**'	ni'	Sa
dha'	ni'	Sa	ni'	ni'	Re	Re	ni'	Sa	Re	Sa	ni'	Sa	Re	Re	Sa	ni'	ni'
Sa	Re	dha'	ni'	Pa'	Sa	Sa	Re	Sa	Re	ga	Re	**Ga**	ga	Re	**Ga**	Re	ga
Ma	Sa	Re	Re	ga	**Ga**	ga	Ma	Pa	Ma	Pa	dha	Pa	**Dha**	dha	**Dha**	dha	ni
Pa	Pa	Pa	Ma	**Ga**	ga	ga	ga	Ma	Sa	Re	**Dha**'	dha'	Re	dha'	ni'	Sa	Re
Re	Sa	Re	**Ga**	ga	Ma	Sa	Sa	Sa	dha'	dha'	ni'	Re	Sa	Ma	Pa	**Dha**	dha
ni	Ma	Pa	Pa	dha	Sa"	ni	ni	Sa"	**Dha**	dha	Re"	Sa"	ni	Sa"	Re"	Sa"	ni
dha	ni	Pa	Ma	Pa	dha	Sa"	ni	Re"	Sa"	dha	Sa"	Pa	Pa	dha	**Dha**	ni	ni
Pa	Pa	ni	ni	Sa"	Sa"	Re"	Sa"	ni	Sa"	Re"	ni	Sa"	ni	Re"	Sa"	Sa"	dha
ni	Re"	Sa"	Re"	**Ga**"	**Ga**"	Sa"	Re"	dha	ni	Pa	Pa	dha	Ma	Pa	dha	ni	Pa
Ma	Pa	dha	**Ga**	Re	ga	Ma	Re	ga	Re	Re	ga	Ma	Ma	Re	Sa	Sa	ni'
Dha'	dha'	dha'	ni'	Re	Sa												

Table 4.: Annotation of Notes for two Samples each of four Ragas

Darbari Kannada 2

Ma	Pa	ga	Re	ga	ga	ga	Ma	Re	Re	Sa	Sa	Sa	Sa	Pa'	**Dha'**	dha'	Pa'
dha'	dha'	**Dha'**	dha'	Sa	ni'	Sa	dha'	Sa	Re	Re	ga	Re	ga	ga	Ma	Re	Sa
ni'	Sa	Sa	ni'	Sa	**Ni'**	ni'	Sa	Re	Sa	Re	ga	Re	ga	Re	ga	ga	ga
ga	Ma	Re	ga	Re	Sa	**Ni'**	Sa	Ma	Pa	Pa	Ma	Pa	Ma	Pa	**Dha**	ga	Re
Ga	ga	Re	ga	ga	Ma	Pa	ga	ga	ga	Ma	Re	Sa	Sa	Sa	Re	Re	ga
ga	Ma	Re	Re	Sa	Sa	ni'	**Ni'**	ni'	Sa	ni'	Re	ga	Re	ga	Ma	Pa	Ma
Pa	**Dha**	dha	Pa	dha	Pa	dha	Pa	dha	ni	Pa	Pa	Pa	Pa	Sa"	dha	dha	**Dha**
dha	dha	ni	Pa	Pa	ni	Ma	**Ga**	ga	ga	ga	ga	ga	**Ga**	ga	Re	ga	ga
ga	Ma	Re	Sa	ni'	Sa	Re	**Ga**	ga	Re	ga	Ma	Pa	ga	Ma	Re	Sa	**Ni'**
Sa	dha'	Sa	Re	Re	**Ga**	ga	Re	ga	Ma	Re	Sa	ni'	Sa	Sa			

Mian-ki-Malhar 1

Sa	ga	ga	Ma	Re	Re	Pa	Ma	Pa	Pa	ni	Dha	Dha	Dha	Ni	Sa"	Dha	ni
Dha	ni	Pa	Pa	**Ga**	ga	Re	ga	ga	Ma	Re	Sa	Dha'	ni'	ni'	Dha'	Dha'	Ni'
Sa	Re	Ma	**Ga**	ga	Re	**Ga**	ga	ga	Ma	Re	Re	Ni'	Sa	Ni'	Sa	Ni'	Re
Sa	Dha'	ni'	ni'	Dha'	Dha'	ni'	Dha'	Dha'	Ni'	Ni'	Ni'	Sa	Ni'	Re	Re	Dha'	ni'
Dha'	Pa'	**MA'**	ni'	Ni'	ni'	Dha'	Dha'	ni'	Dha'	Ni'	Ni'	Sa	Re	Sa	Ni'	Sa	Ma
Ga	ga	Re	**Ga**	ga	Re	**Ga**	ga	**Ga**	ga	**Ga**	ga	Ma	Re	Sa	Sa	Re	Ma
ga	Ma	Ma	Re	Re	Re	Pa	Pa	Ma	**Ga**	ga	**Ga**	ga	**Ga**	ga	Re	Sa	Ni'
Re	Sa	Re	Re	Re	Pa	**MA**	Pa	**MA**	Ma	Pa	Ma	Pa	**MA**	Pa	Pa	ga	**Ga**
ga	ga	Ma	Re	Pa	Pa	Pa	ni	Dha	ni	Dha	ni	Dha	ni	Dha	ni	Dha	Dha
Dha	Sa"	Ni	Sa"	Dha	ni	Pa	Ma	Re	Ma	Ma	Re	Pa	Pa	ni	Dha	ni	Dha
Dha	Dha	Ni	Sa"	Dha	Pa	Dha	Dha	Ni	Ni	Sa"	Ni	Sa"	Ni	Sa"	Pa	**Ga**	Ma
Pa	Pa	Dha	Dha	Ni	Ni	Sa"	Ni	Sa"	Dha	Ni							

Mian-ki-Malhar 2

Sa	Ni'	Re	Re	Sa	ni'	Pa'	Ma'	Pa'	ni'	Dha'	Ni'	Sa	Re	**Ga**	Re	Re	Pa
Ma	Re	ga	Re	ga	ga	Ma	Re	ga	Sa	Ni'	Sa	Sa	Re	Re	ni'	Pa'	Dha'
ni'	Dha'	ni'	Dha'	Ni'	Ni'	Sa	Re	Sa	Dha'	ni'	Dha'	Ni'	Sa	Re	Re	Sa	Pa'
ni'	Dha'	Ni'	Ni'	Sa	Sa	Sa	Sa	Dha'	ni'	Dha'	ni'	Ma'	**MA'**	Pa'	Pa'	ni'	Dha'
ni'	Dha'	ni'	Dha'	Ni'	Dha'	Ni'	Ni'	Ni'	Sa	Ni'	Sa	Dha'	ni'	Ma'	Pa'	ni'	Dha'
ni'	Ni'	Sa	Ni'	Dha'	ni'	Ni'	Dha'	Ni'	ni'	Dha'	Dha'	ni'	Dha'	ni'	Ni'	ni'	Ni'
ni'	Ni'	Sa	Sa	Ni'	Sa	Sa	Ni'	Re	Sa	Re	Ni'	Sa	ni'	Dha'	ni'	Dha'	Dha'
Ni'	ni'	Ni'	Dha'	ni'	Dha'	ni'	Dha'	ni'	Dha'	ni'	Dha'	Ni'	Ni'	ni'	ni'	Ni'	Sa
Sa	ni'	Sa	Ni'	Sa	Sa	Re	ni'	Sa	Re	Re	ga	**Ga**	Re	ga	Re	ga	ga
Re	ga	Ma	Re	Sa	Ni'	Sa	Re	Sa	ni'	Dha'	Ni'	Sa	Sa	Re	Sa	Re	Re
Pa	**MA**	Pa	Pa	Re	Re	ga	ga	Re	**Ga**	ga	Ma	Re	Re	ga	Re	Re	Ma
Re	Sa	Pa	Pa	Pa	Pa	ga	Ma	Re	Re	**Ga**	**Ga**	ga	Ma	Re	Sa	Re	ni'
ni'	Ma	Dha	Dha	Dha	Ni	Ni	Ni	ni	**MA**	Ma	Pa	Pa	Re	ga	ga	Pa	Pa
Re	ga	**re**	ga	Re	ga	Ma	**Ga**	Ma	Re	Sa	Sa	ni'	Dha'	Ni'	Sa		

Table 4 (continued).: Annotation of Notes for two Samples each of four Ragas

Some of the incorrect notes are examined with reference to the pitch profiles. Some of such notes are found to belong to two classes namely a glissando type profile or a sharp rise or fall. These errors indicate that the selected threshold value of ± 0.02 for slope needs some modification. This can be done through readjustment of this value. However there are still some notes which are incorrect in the sense that a vikrit note is detected as pure note or vice-versa. This is not surprising. While the learnt intervals are quite specific and categorical the perception of notes is likely to be very much influenced by the semiotics of the total ambience of the music. It is quite possible that the syntax and semantics of the piece influences perception.

Furthermore the minimum threshold of 100 ms for a perceptible note also needs some reconsideration as this excludes some short duration notes from annotation.

Todi 1

dha'	dha'	dha'	dha'	Ni'	Sa	re	re	re	Sa	Sa	Ni'	Ni'	Sa	re	re	dha'	Pa'
Pa'	dha'	Ni'	Ni'	Pa'	Ni'	Sa	Sa	Sa	Sa	re	Sa	re	re	re	dha'	Sa	Sa
Sa	dha'	dha'	Ni'	re	dha'	Pa'	dha'	Ni'	Ni'	Ni'	Sa	Ni'	Sa	Sa	Sa	re	re
re	re	dha'	Ni'	re	re	re	re	re	Sa	Sa	Ni'	Sa	Ni'	re	re	Sa	dha'
dha'	re	re	re	re	Sa	Sa	ga	Sa	Ni'	Ni'	Sa	re	re	Sa	re	re	Sa
Sa	re	Sa	ga	Ni'	Sa	dha'	dha'	ga	re	dha'	re	ga	ga	ga	re	ga	Ni'
Sa	Sa	re	Sa	Sa	re	re	re	re	Sa	re	re	Sa	re	re	Sa	Sa	Sa
Re	Ni'	Sa	Sa	re	re	dha	dha	Pa	Pa	Pa	Pa	MA	dha	re	Sa	re	re
Re	re	re	Sa	Sa	Sa	Sa	Ni'	re	Ni'	re	ga	re	re	Ni'	dha	dha	Pa
Pa	MA	Pa	MA	dha	MA	Pa	MA	dha	re	re	re	re	Sa	Sa	re	re	Sa
Re	re	Sa	Sa	Sa	re	MA	dha	MA	ga	MA	MA	Sa"	Sa"	Sa"	dha	re"	dha
Pa	MA	dha	re	re	ga	re	Sa	re	re	Sa	re	re	dha	dha	dha	re	re
Re	dha	dha	re	ga	re	ga	re	dha'	**Re'**	**Ga**'	dha	Ni'	Pa	MA	dha	Pa	MA
Dha	dha	Sa	Sa	Sa	re	Sa	Sa	re	re								

Todi 2

Sa	re	**Re**	re	ga	re	**Re**	re	Sa	Sa	Ni'	dha'	Sa	Ni'	Sa	re	re	ga
re	Sa	Sa	Sa	re	ga	**Re**	re	Sa	dha'	Ni'	Sa	re	re	ga	re	Sa	re
ga	re	Sa	re	Sa	Sa	Ni'	Sa	Sa	Ni'	re	Sa	Sa	Pa'	dha'	MA'	Sa	Ni'
re	Sa	re	re	Sa	re	ga	ga	re	**Ma**	MA	dha	dha	Pa	dha	dha	MA	dha
MA	MA	Pa	dha	Pa	MA	Pa	ga	re	re	Sa	ga	re	ga	re	Sa	re	Sa
re	re	**Re**	re	**Re**	ga	**Re**	re	Sa	Sa	Ni'	re	Sa	re	Sa	re	re	re
Re	re	**Re**	re	Sa	Pa'	Pa'	MA'	dha'	**ni**'	**Ma**'	dha'	Sa	Ni'	Sa	re	re	Sa
re	**Re**	re	**Re**	re	re	ga	**Re**	re	Sa	Sa	re	MA	dha	re	MA	Pa	dha
re	re	re	Sa	MA	Pa	dha	MA	Ni	dha	MA	Ni	Pa	MA	re	re	ga	re
Sa	Ni'	Sa	re														

Table 4.(continued) : Annotation of Notes for two Samples each of four Ragas

REFERENCE

1. Datta A K and Kundu R, "A psycho-perceptual hypothesis for Indian Musical scale", Pre-Proceedings of the Int. workshop on recent trends in Speech, Music and Allied Signal Processing , Delhi, 9-11,1991

Datta A K et.al, "Relevance of Consonance in Indian Musical Scale : Theory and Practice", J. Acoust. Soc. India, Vol. XXIII, 1995

Datta A K et al, "Ratios and Intervals used in North Indian Classical Music - a Study Based on Recorded Performance", J. Acoust. Soc. India, Vol. XXIV, 1996

Datta A K et al., "Pitch Analysis of Recorded Vocal Performances in Hindustani Music: Evidence of a Personal Scale", J. Acoust. Soc. India, Vol. XXV, 1997

Datta A K et al., "On Scientific Approaches to the Study of Vaditya in Indian Music", Proc. Fifth Int. Workshop on Recent Trends in Speech, Music and Allied Signal Processing, 14-15 December, 2000, Thiruvanathapuram, India

Datta A K et al, "Study of srutis in Indian musical scales and relevance of schools in recent times using Signal Processing techniques", Proc. Frontiers of Research on Speech and Music (FRSM-2003), 15-16 February, 2003, IIT Kanpur, India

Datta A K et al, " Sruti usuage by old and contemporary singers in khyal: An objective approach", Proc. Frontiers of Research on Speech and Music (FRSM-2004), 8-9 January, 2004, Annamalai University, Tamilnadu

2. Lerdahl, F., Jackendoff, R., “A Generative Theory of Tonal Music”, MIT Press, Cambridge, MA, 1983.
3. A. K. Datta, “Generation of Musical Notations from Song using State-Phase for Pitch Detection Algorithm”, J. Acoust. Soc. India, Vol. XXIV, 1996.
4. A. K. Datta, R Sengupta, N Dey and D Nag , “Ratios and Intervals used in North Indian Classical Music - A Study based on Recorded Performance”, J. Acoust. Soc. Ind., XXIV, III-4.1, 1996.
5. J Laver et al, “Acoustic Waveform Perturbations and Voice Disorders”, J. Voice, 6(2), 1992.
6. A K Datta et.al, “Sruti usage by old and contemporary singers in Khyal: An objective approach”, Proc. FRSM-2004, Annamalai University, Tamilnadu, January 8-9, 2004.
7. A K Datta, R Sengupta, N Dey and D Nag, “Relevance of Consonance in Indian Musical Scale : Theory and Practice”, J. Acoust. Soc. Ind., XXIII, 110, 1995
8. Rao Sahib M.Abraham Pandither, “Karunamirtha Sagaram – On srutis”, Asian Educational Services, New Delhi, 1984.

A. Daniloue, “Introduction to the Study of Musical Scales”, Oriental Books Reprints Corpn., New Delhi, 1979.

D. A. Lentz, “Tones and Intervals of Hindu Classical Music”, Nebraska Univ. Studies, New Series No. 24, Lincoln Univ. Press, 1961.

9. R. Sengupta, N. Dey, D. Nag , A. K. Datta and A. Mukerjee, “Automatic tonic (sa) detection algorithm in Indian classical vocal music”, Presented in the National Symposium on Acoustics, NSA-2005, 14-16 December, 2005, Bangalore.
10. A K Datta et.al, “Evaluation of srutis in Hindustani Music from recorded performances”, Proc. FRSM- 2005, Utkal University, Bhubaneswar, January 6-7, 2005.