DISTINGUISHING MUSICAL INSTRUMENT PLAYING STYLES WITH ACOUSTIC SIGNAL ANALYSES

Prateek Verma and Preeti Rao

Department of Electrical Engineering, IIT Bombay, Mumbai - 400076

E-mail: prateekv@ee.iitb.ac.in

In this work, we study the different playing styles of Indian classical instrumental music with respect to signal characteristics from flute solo concert performances by prominent artists. Like other Hindustani classical instrumentalists, flautists have evolved *Gayaki* (vocal) and *Tantrakari* (plucked string) playing styles. The production and acoustic characteristics of the music signal are discussed. Computational features that serve to discriminate the two styles are proposed.

1. Introduction

A variety of Indian musical instruments has been an integral part of the Hindustani classical music tradition. While vocal music remained the most performed, instrumental solos began to gain prominence in the early 20th century. String and wind instrument performers developed unique styles consistent with their instruments but also developed elements of vocal style (*gayakiang*) to match the preferences of the conservative listening community. Today, instruments such as the sitar, sarangi and bansuri (bamboo flute) among others are performed in both *gayaki* and *tantrakari* (plucked string) styles.The choice of style influences the instrumentalist's repertoire and concert structure as well as the actual rendering of the *raga* in terms of the realization of phrases and ornamentation. For instance, the *gayakiang*, imitating the singing voice, is more fluid and makes extensive use of the *mind* (glide), while the *tantrakariang* is characterized by melodic leaps. Similarly *gamaks* are used in *gayakiang* instead of the repeated note articulations required in plucked string instruments where the notes are not sustained. Thus the playing style influences the acoustic attributes of the music. The goal of this work is to study this relationship via the analysis of the audio signal.We choose the bamboo flute (*bansuri*) concerts of various well-known artists for this study.

Interestingly, the flute is the earliest known music instrument in human history [1]. PanditGhosh is credited with introducing the flute solo to the Hindustani classical music stage, and his ability to perform both *gayaki* and *tantrakariang* on his own specially constructed instrument with equal mastery.

In the next section, we present the concert audio recordings selected for this study. The goal was to assemble a dataset comprising examples of each style performed by well-known Hindustani flautists. A study of the acoustic characteristics of each style is presented next, motivating computational features that can discriminate the two styles.

2. Database and Annotation

After consultations with the musicians we narrowed down our database to the artists and ragas listed in Table 1. Pt. Ghosh and Chaurasia were chosen as the stalwarts who influenced the way the instrument was presented during their era, with their disciples who themselves are renowned artists. Pt. Ghosh is credited with introducing the flute solo to the Hindustani classical music stage, and his ability to perform both *gayaki* and *tantrakariang* on his own specially constructed instrument with equal mastery.He is widely regarded as father of modern flute playing and mainly played in Gayakitradition or the khayal style. His later performances influenced by his tutelage underAllauddin Khan incorporated the*Tantrakari* style. HariprasadChaurasia trained under the tutelage of Annapurna Devi who herself was a sitar player. He was influenced by the sitar style of playing and incorporated a lot of intricacies pertaining to the sitar into flute. His playing is characterized extensive use of tonguing which many experts believe brings the effect of sitar pluck into the flute.

Some of the audios in Table 1 are commercial CD recording and others taken from YouTube sources. Only the *alap* sections of the concerts were used. The audio thus comprised flute solos accompanied by the tanpura (drone). The audios were downsampled to 16 kHz mono with 16 bits/sample.Ten tracks of total 48 mins were manually annotated for section which belonged to *Tantrakari* or *Gayaki* based on heuristics arrived at by consultation with four musician experts. Table 1 shows the distribution of annotated segments. Certain ambiguous markings were not done as either as Tantrakari or Gayaki. The Jhinjhoti and Hemant concerts were labeled as Alap/Jod/ Jhala on the CD cover as well as unanimously labelled as Tantrakari style by all four experts.

3. Production and Acoustic Characteristics

A sound, whether speech or instrumental, can be described along three main dimensions: production, acoustic signal properties and perception. This section will discuss production aspects of sound influte and ways of articulating different types of sound along with their acoustic analyses. A flute is open at both the ends, with length of the flute determining the frequency of the standing waves generated by the air due to blowing in the flute. Different notes of higher octave can be articulated with the same fingering position by changing the width of the lip aperture or by changing the strength of blow. The reason for production of harmonics in flute is the ability to produce integral multiples of standing waves with subdued amplitudes when blowing is same. Breath noise is mainly due to the randomness occurring at the lip aperture which is difficult to model for synthesis of flute sounds. The various notes produced are due to the change of length of the resonating chamber with adjustment in fingering patterns to render all the *swars*. The sequence of notes played forms the melody of the music. As such, the flute offers a certain flexibility in terms of how notes can be connected within a melody. Gayaki (or singing) style calls for smooth transitions between notes together with ornamentation typical of vocal music. To produce various ornamentation, production techniques and skills are required which come with immense practice. For example, a glide (mind) is articulated with slowly lifting or resting the fingers while making transitions from one note to the other instead of any abrupt movement. The rate of moving the fingers controls the rate of the glide. Further, the articulation of continuous notes requires steady blowing. Oscillatory continuous repetitive movements such as vibrato are also characteristic of voice.Vibrato is produced by the flautist by sudden bursts of air during continuous blowing with the rate of vibrato being controlled by the rate of bursts and extent with the amount of burst. Indeed, the flute was con sidered to be a vocal style instrument before artists were influenced by styles of other instrument such as sitar and tabla.

Artist	Raga	Duration	f/T/G(in s)
Hari Prasad Chaurasia	**Jhinjhoti**	**18:54**	**352/465/267**
	Hansadhwani	1:50	43/14/33
	Yaman	2:29	57/12/40
	Sindh Bhairav	3:30	50/20/66
RupakKulkarni	**Hemant**	**6:50**	**124/102/111**
RonuMajumdar	Vibhas	5:39	47/11/44
PannalalGhosh	Yaman	2:18	39/3/66
	Shri	2:20	45/0/29
	Pilu	1:30	133/35/105
NityanandHaldipur	ShuddhaBasant	2:46	94/0/41

Table 1: Dataset in terms of Artist , Raga and the duration of flute played without any style cue(f), Tantrakari style (T), and Gayaki style (G).

In instruments such as sitar, *Tantrakari* style is played by extensive use of right hand and making discrete left hand movements. Rendering of discrete pitches with minimum in between notes with small/ large jumps is characterized as *Tantrakari.* Also irrespective of pitch movements, sometimes truing sound (changing the blowing patterns with abruptness) is done extensively is also characteristic of *Tantrakari* style. Abrupt breaks in blowing steady notes with or without the use of tongue is characterized as *Tantrakari* style.To produce *Tantrakari* style ornaments, one way is to change the style of articulation. The changes in the blowing patterns arise by using tonguing along with blowing. Truing sound is a rapid tonguing pattern in order to change the blowing pattern.In order to articulate the same note we can have changes in the blowing patterns, i.e. simple blow, use of tongue at the onset of a note, use of blowing variations similar to rhythm patterns (e.g. 1-2, 1-2 or 1-1-2 etc), i.e. stopping and blowing after regular intervals. Another way to articulate a steady continuous note is to periodically touch other notes discretely for short duration which is a not possible in vocal due to the nature of the vocal cords.

Based on discussions with musicians, five categories were used for each of the two styles (T and G) as follows: (i)T1: Blowing same note with stops (ii) T2: Steady notes with abrupt movements (iii) T3: Discreteness without silences (iv) T4: Discreteness with rapid jumps (v) T5: Truing with any pitch movement (vi) G1: Oscillatory repetitive movements (vii) G2: Non-Oscillatory repetitive movements (viii) G3: Glides (ix) G4: Non-specific continuous patterns (x) G5: Vibrato.

The Table 2 shows the distribution of different ornaments across the clips. All the four music experts were unanimous in characterizing Truing as Tantrakari though there were some reservations regarding the T3/T4 category vs the G3/G4 category.

The different sound production methods discussed previously influence the signal characteristics of the flute. Both pitch and amplitude are expected to be affected. We carry out pitch and energy contour analysis on the audio recording as follows. The extraction of pitch was carried out using Melodia plug-in of Sonic visualizer [3] .For automatically extracting the pitch of the flute, detection of sections in performance is required. The Melodia plug-in uses in built singing voice detection which doesn`t work very accurately for flute sounds. For this the voicing tolerance was changed manually for different clips so that the tracking error was minimized.The energy variation with time is obtained by computed the whort-time energy of the audio signal. We discuss the acoustic signal peculiarities observed with the different production methods next.

3.1 Vibrato

The sinusoidal repetitiveness in pitch is clearly visible from the pitch contour calculated every 10ms and shown as a continuous pitch variation with time in Figure 1. In order to represent the temporal dynamics of the pitch contour, we segment it into 1.5 s duration chunks and take a1024-pointFFT as in [4]. Figure 1 shows the magnitude spectra so obtained for a vibrato segment and for a non-vibrato segment. We can clearly see that a peak near 5 Hz in the former which is the rate of vibrato whereas the non-vibrato spectrum is lowpass in nature.

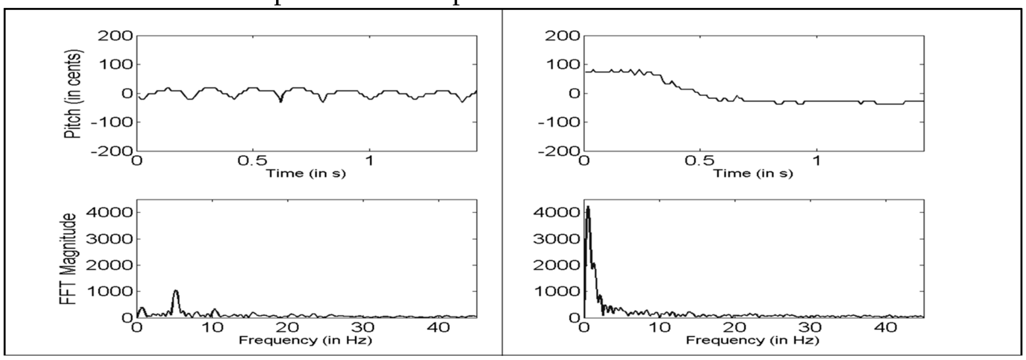

Figure 1: Comparison of 1024-pt FFT of 1.5s chunk of data for vibrato characteristics

3.2 Blow- Stop and Truing

Since the intensity of the tanpura drone is relatively constant throughout, the energy variations in the audio signal correspond that of the flute.The energy contour is obtained for a hop of 10ms and window of 25 ms. The figures below show the acoustic characteristics for the blow-stop and truing modes. We see the spectrogram followed by the energy contour and magnitude spectrum computed from the energy contour with a 1024-point FFT. The periodicity in energy versus time in the blow-stop mode is evident in the spectrum which shows harmonicity. In order to quantify the periodicity, ratio of the peak in 2-10 Hz to the DC values can be taken.In the truing mode, the flautist performs very rapid energy modulation while playing different notes as evident from the pitch contour below. The rate of energy modulations also increase due to the use of tongue.The peak in the FFT plot has clearly shifted to the 10-20 Hz range indicating the rapid movements which are inherent to truing style of play as evident from the Figure 2.

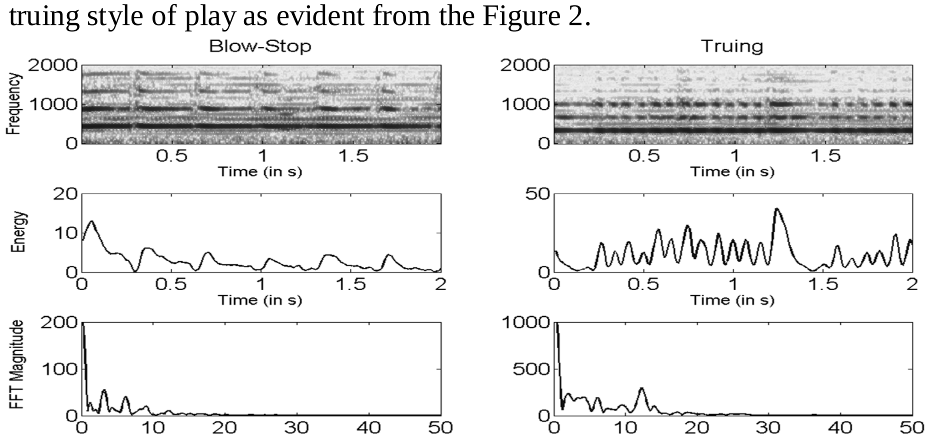

Figure 2: Depiction of spectogram variations along with the energy contour and the 1024-pt FFT on chunks of size 2s for blow-stop (on left) and Truing(right).

4. Feature Design

From the Table 2, we observe that the presence of truing serves to distinguish Tantrakari from Gayaki style rather any other single characteristic. From Figure 2, we see that truing can be detected by analysis of the short-time energy contour.

Artist	Raga	Vibrato	Truing	Blow -Stop	T3/T4	G3/G4
Hari Prasad Chaurasia	**Jhinjhoti(T)**	✓	✓	✓	✓	✓
	Hansadhwani	✓	×	✓	✓	✓
	Yaman	×	×	×	✓	✓
	Sindh Bhairav	✓	×	×	✓	✓
RupakKulkarni	**Hemant(T)**	×	✓	✓	✓	✓
RonuMajumdar	Vibhas	✓	×	×	✓	✓
PannalalGhosh	Yaman	✓	×	×	✓	✓
	Shri	✓	×	×	×	✓
	Pilu	✓	×	×	×	✓
NityanandHaldipur	ShuddhaBasant	✓	×	×	×	✓

Table 2: Manually detected ornaments across the database.

We model the behaviour of the energy contour using two features i.e. the sub-band Peak ratio (SPR) and the Sub-band FFT ratio (SFR) as given in Eqn.1 and Eqn.2. It was found that the truing feature is rendered with oscillations in the energy contour in the range of 10-20 Hz.

For *E(k)*, the short-time Fourier Transform of the energy contour, the peak ratio and energy ratio is defined as,

$$\text{Sub-band Peak Ratio (SPR)} = \frac{max(|E(k)|_{k_{10Hz} \le k \le k_{20Hz}})}{max(|E(k)|_{k_{20Hz} \le k \le k_{30Hz}})} \quad (1)$$

$$\text{Sub-band FFT Ratio (SFR)} = \frac{\sum_{k_{10\ Hz}}^{k_{20\ Hz}} |E(k)|}{\sum_{k_{20\ Hz}}^{k_{30\ Hz}} |E(k)|} \quad (2)$$

For the experimentation 54 Gayaki sections each of 2s data were chosen along with 61chunks belonging to Truing category. From a scatter plot of the two features, the distinction is clearly visible. Tantrakari clips have pronounced peak in the range of 10-20 Hz whereas the cases of non-tantrakari clips the FFT of energy contour will be similar as compared to the 20-30 Hz range. This separation can be easily modelled using a generative or discriminative classifier as evident from the Figure 3.

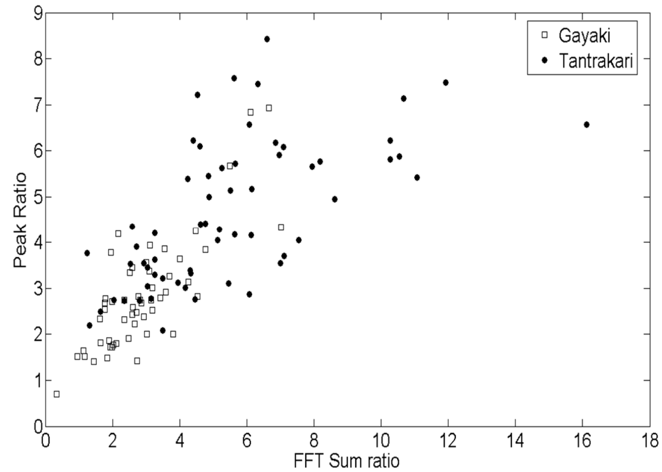

Figure 3: Scatter plot for 2s chunks of data for two features for Tantrakari and Gayaki styles.

5. Conclusions and Future Work

Style classification is studied for alap clips of flute recordings. Acoustic signal characteristics distinguishing the two styles are presented forming the basis for the extraction of features that can be applied to the task of automatic style detection from concert audio recordings. Improvements are required in pitch detection especially if the work is extended beyond the alap section to other concert sections where table accompaniment is expected. More extensive testing within a formal classification framework is needed to establish the relevance and adequacy of the proposed features. The approach followed in this work can be extended to style detection across musical instruments.

ACKNOWLEDGEMENTS

The authors thank the musicians for useful discussions and helpful insights: Pt. Nayan Ghosh, Dr. Vishwas Kulkarni, Prof. A. R. Shastri, Kaustuv Kanti Ganguli and Seema Vidwans. This work received partial funding from the European Research Council under the European Union`s Seventh Framework Programme (FP7/2007-2013)/ ERCgrant agreement 267583 (CompMusic).

REFERENCES

[1] N. J. Conard, M. Malina, and S. C. Münzel, "New flutes document the earliest musical tradition in southwestern Germany." *Nature*460.7256 (2009): 737-740.

[2] Bansuri (Flute), Spicmacay, "http://old.spicmacay.com/articles/bansuri-flute"

[3] J. Salamon, and E. Gómez. "Melody extraction from polyphonic music signals using pitch contour characteristics." *Audio, Speech, and Language Processing, IEEE Transactions on* 20.6 (2012): 1759-1770.

[4] C. Gupta, and P. Rao. "Objective assessment of ornamentation in Indian classical singing." *Speech, Sound and Music Processing: Embracing Research in India*. Springer Berlin Heidelberg, 2012. 1-25.