APPLICATION OF PITCH TRACKING TO SOUTH INDIAN CLASSICAL MUSIC

Arvindh Krishnaswamy

Center for Computer Research in Music and Acoustics
Dept of Electrical Engineering, Stanford University
arvindh@ccrma.stanford.edu

ABSTRACT

We present results of applying pitch trackers to samples of South Indian classical (Carnatic) music. In particular, we investigate the various musical notes used and their intonation. We try different pitch tracking methods and observe their performance in Carnatic music analysis. Examining our data, we find only 12 distinct intervals per octave among the notes that are played with constant pitch. However, there are pitch inflexions used sometimes that are not mere ornamentations - they are essential to the correct rendition of certain notes. Though these inflexions can be viewed as different versions of a particular note, they are certainly not equivalent to constant-pitch intervals like Just Intonation intervals, semitones or quartertones.

1. INTRODUCTION

How many intervals per octave does Indian classical music use? What are their numerical values? How much precision in intonation is required of an Indian musician? What ornaments (like vibrato, for example) do Indian musicians use to embellish the notes they play? These questions have been at the heart of vigorous debate and controversy in Indian music theory for centuries [1].

There are several musicologists who still maintain that Indian music uses more than 12 intervals per octave (many speak of 22), while others have disagreed with such claims. However, among the multitudes of books, articles and papers in the literature, only a handful of authors have presented quantitative data to substantiate their individual viewpoints, like Jairazbhoy [2] and Levy [1] for instance. Unfortunately, these two authors used equipment for pitch tracking that are now antiquated and were not able to present any detailed, high-resolution plots. Their focus was also restricted to North Indian music. While some of their conclusions are reasonable, their work seems to have had little effect overall on the mindset of most Indian musicians and musicologists today.

It is widely accepted that Carnatic music has a *basic* set of 12 intervals as shown in Table 1. (There are four sets of enharmonic intervals). The debate is over whether there are further variations of the 12 notes listed. For example, some people claim that 10 of the 12 notes (all except Sa and Pa) have two varieties each - higher and lower pitched versions, and give *constant* numerical interval values for them [3].

In this paper, we wish to present data which we feel will contribute to a better understanding of the musical notes and intervals used in Carnatic music. We focus on contemporary South Indian classical music, and we do not speculate about music in ancient India or its subsequent evolution. We also do not treat North Indian music directly, but our findings, discussed later, are applicable to all types of Indian music, including Indian pop music, which is a fast-growing and commercially viable industry today. For applications such as automatic transcription and computer music synthesis, the observations presented here would be essential.

2. PITCH TRACKERS

We tried three pitch tracking methods which were based on (i) the Short-Time Fourier transform (STFT), (ii) the time-domain autocorrelation function and (iii) a harmonic source separation technique [4]. We found that all three methods were capable of revealing the information we were after - clues about intonation and intervals present. All of the pitch trackers were verified to have negligible error for slow-varying, clean, monophonic signals: when synthetic signals (variable amplitude, variable pitch harmonic signals) were played and recorded using an inexpensive desktop computer's speakers and mic, the maximum error in the pitch estimates was always less than 0.3 cents in all three techniques. The pitch ranges of these test signals were similar to those in the melodies we analyzed in this paper.

For each audio segment, we computed the STFT using a Hamming window. Typical parameter values we used were: WindowSize = 25 - 50 ms and HopSize = 3 ms. These choices ensured that we were able to satisfactorily resolve spectral peaks of adjacent harmonics in all our signals, as well as obtain a reasonable time resolution. The DFTSize was usually 4 - 8 times the WindowSize (with zero padding) for better frequency interpolation. In each frame, we located the harmonics using peak detection and

In *Proc. IEEE ICASSP 2003, Apr 6-10, Hong Kong, China.*

Carnatic	Western	Just Ratios			Cents
Sa	P1	1	/	1	0.0000
Ri_1	m2	16	/	15	111.73
Ri_2 or Ga_1	M2	9	/	8	203.91
Ga_2 or Ri_3	m3	6	/	5	315.64
Ga_3	M3	5	/	4	386.31
Ma_1	P4	4	/	3	498.04
Ma_2	+4	17	/	12	603.00
Pa	P5	3	/	2	701.96
Da_1	m6	8	/	5	813.69
Da_2 or Ni_1	M6	5	/	3	884.36
Ni_2 or Da_3	m7	9	/	5	1017.6
Ni_3	M7	15	/	8	1088.3

Table 1. The 12 basic Carnatic music intervals along with Western equivalents. Also shown are values in a possible Just Intonation tuning system.

refined the frequency estimates using quadratic interpolation. A pitch estimate for each time frame was obtained by using a weighted average (using the local SNR) of the fundamental pitch implied by each partial. This way, we were able to obtain accurate and robust pitch estimates. In general, we found that while the STFT based method was sufficient for capturing the overall pitch contour of slow passages, it was difficult to track fast phrases which had a high frequency slew rate. This difficulty was basically due to inherent smoothing in the STFT: for resolving spectral peaks, the window size needed to be much longer than the duration of one time sample.

Our autocorrelation based pitch tracker (ACPT) is similar to [5] with the following variation: we wish to produce a pitch estimate for each time sample using only those data *centered* around that sample to avoid any delays in the pitch estimate. As shown in Fig. 1, if we are considering the time sample denoted by "Center," we would pick segments such as "L1" and "R1" for computing the correlation value. If we wanted more overlap in the segments or smoothing, we could select the samples as shown by "L2" and "R2." We would consider various sets of time-shifted segments, each with a different lag, to find the ones with the highest correlation and refine the corresponding pitch period estimate using quadratic interpolation. This method gives better time resolution than the STFT method when following fast passages since we have a pitch estimate for each time sample and also because the amount of smoothing can be more easily controlled. However, it works best only with monophonic signals with little noise since there is no filtering performed to suppress noise or other unwanted signals. (In the STFT method, there are some implicit filtering mechanisms, if we adopt the filter-bank interpretation, which helps to reduce the effect of noise and interfering signals somewhat).

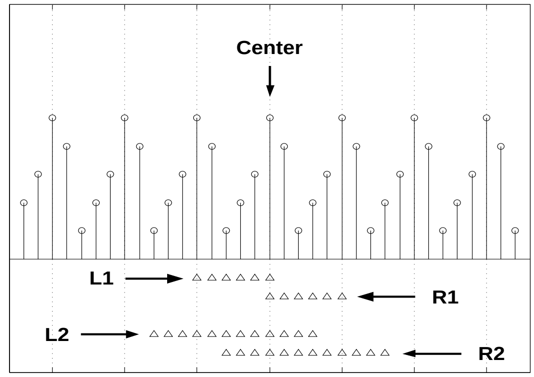

Fig. 1. Segment selection in the ACPT.

Artist	Instrument	Raga
T. N. Krishnan	Violin	Thodi
T. R. Mahalingam	Flute	Sankarabaranam
U. Srinivas	Mandolin	Sriranjani
Ravikiran	Chitraveena	Chintamani
Unnikrishnan	Vocal	Panthuvarali

Table 2. Source of the audio clips presented.

We also tried a variation of the technique described in [4] where we extract from an audio signal the "analytic signal" corresponding to a partial using "frequency demodulation" and low-pass filtering. Each partial in a harmonic set gives an estimate of the fundamental pitch, and a weighted average of these estimates is used as the final pitch value. Though this method is the finest of all three tried, it was also the most expensive computationally. We did not require the resolution offered by this method to make our case in this paper, but, nevertheless, this method seems to be the most promising for accurate pitch tracking of fast passages in the presence of noise and other interfering signals. (In most Indian music concerts and recordings, there are usually two or more musical instruments playing at once).

3. AUDIO SAMPLES AND PLOTS

For our experiments, we used several audio clips drawn from commercial recordings of professional musicians who are at the top of their field, and in this paper we show pitch tracks for five of them, each from a different instrument/musician as listed in Table 2.

We plotted the pitch in cents relative to the tonic (Sa), and the solid horizontal lines in each plot corresponds to a Just Intonation value given in Table 1. The values of Equal Temperament intervals are simply integer multiples of 100 cents. Finally, all the pitch tracks have been labeled with the notes played by the artist.

Figure 2 shows a typical STFT, and the pitch inflexions

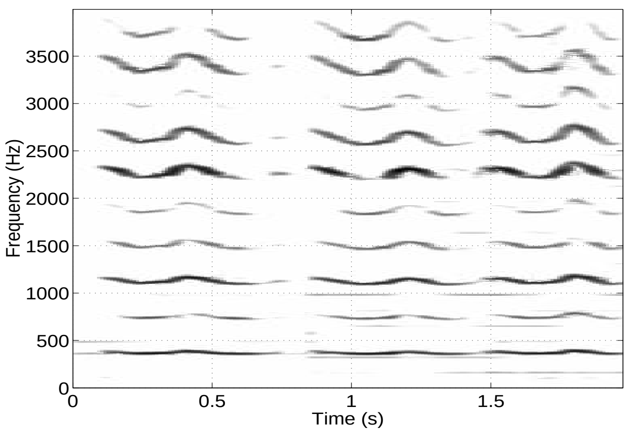

Fig. 2. A typical STFT - T. N. Krishnan - Thodi

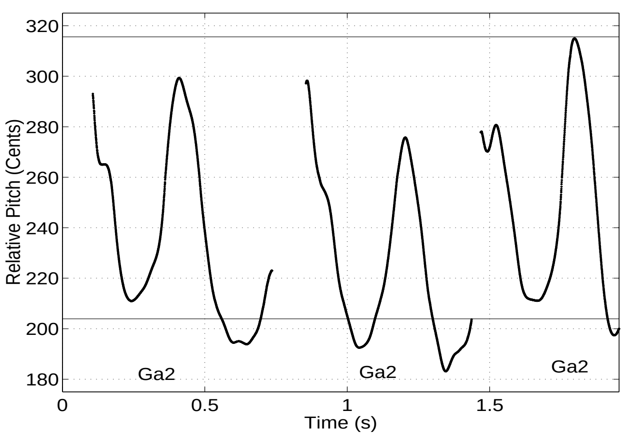

Fig. 3. T. N. Krishnan - Thodi - From STFT

of interest are clearly visible, especially in the higher partials. The corresponding pitch estimate is shown in Fig. 3. (The gaps in the plot are due to silence). Figures 5 to 7 were also generated using the STFT method.

Figure 4 was obtained using the autocorrelation method, and it reveals more minute details than all the other plots. For example, there are artifacts visible near 300 and 400 cents due to the frets on the instrument used by the artist. Such features, which are also audible, are smoothed out by the STFT sometimes.

The audio samples were manually segmented into notes which were further classified into (i) “constant-pitch notes” and (ii) “notes with ornaments” by the author (who has, thus far, 20 years of experience learning, performing and teaching Carnatic music on the violin). By “constant-pitch notes,” we are referring to those notes that were *intended* by the musician to be played with unchanging pitch and without any ornaments like vibrato or portamento.

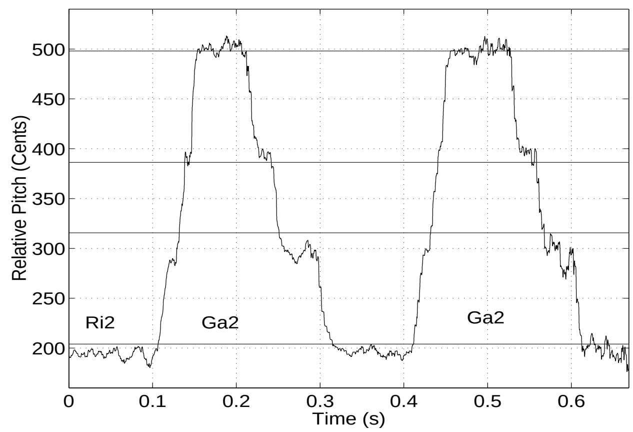

Fig. 4. U. Srinivas - Sriranjani - From ACPT

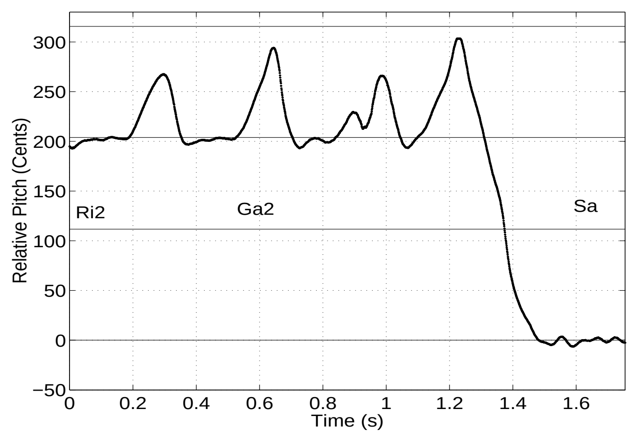

Fig. 5. Ravikiran - Chintamani - From STFT

4. OBSERVATIONS AND DISCUSSION

In the audio clips we analyzed, we noticed that intonation was highly variable, as can also be seen in the plots presented here. Levy [1] has made similar observations. Given the magnitude of these variations, we feel that it is unreasonable to expect an Indian musician to be able to produce, consistently and distinctly, two intervals that are only around 20 cents apart, like some of those listed in [3]. We did not find any quartertones in our audio samples either. Rather, we found a total of only 12 clearly distinguishable intervals among the “constant-pitch notes,” and their values were also close to the 12 semitones used in Western music. It should be further noted that the mandolin used in Carnatic music today has 12 frets per octave, which provides more evidence supporting our previous claim.

There were lots of instances, however, where a note would be played as an inflexion above or below one of the 12 basic intervals, like the Ga_2 in Fig. 5 or an inflexion between two of the 12 intervals like in Figures 3 and 4.

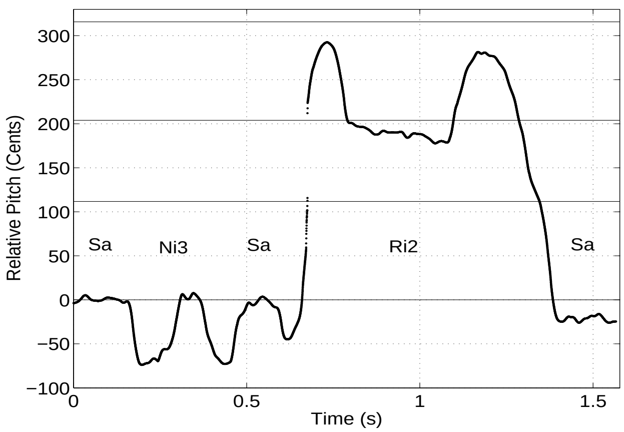

Fig. 6. T. R. Mahalingam - Sankarabaranam - From STFT

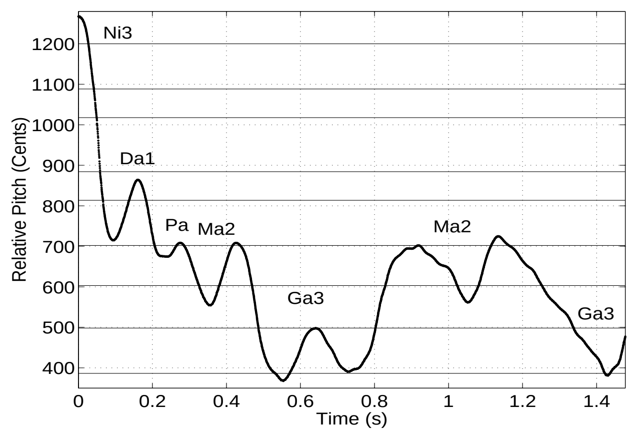

Fig. 7. Unnikrishnan - Panthuvarali - From STFT

Note that all three figures show 3 different "versions" of the "same" note, Ga_2. In addition, different notes could be played using similar pitch inflexion shapes and limits, like the Ga_2 in Fig. 5 and the Ri_2 in Fig. 6. Sometimes, what is vocalized or notated is not even close to the actual pitch that is played or sung, like the Ni_3 in Fig. 7.

These inflexions or "gamakas" are not always optional embellishments: in some cases, it is *required* that the notes are rendered in this manner, like the Ga_2 in Chintamani. However, since these gamakas are not constant in pitch, it is incorrect to assign them constant interval values like some authors have done [3]. These gamakas are also probably the main reason why some people feel that Carnatic music uses more than 12 notes. For instance, we just discussed 3 versions of Ga_2, none of which sound similar to each other or even resemble the "constant-pitch" variety (which is "flat" and somewhere around 300 - 315 cents).

Gamakas exhibit, to a certain extent, temporal structure and regularity, and to properly characterize them, they need to be analyzed both in pitch *and time*. The set of all possible such inflexions is finite however, and it should be possible to produce an exhaustive and categorized list.

We also note that the intervals in Just Intonation are very close to equally tempered intervals with respect to the variances in intonation that we observed. While our data do not conclusively prove that Carnatic music uses Just Intonation, they do not rule it out either. (Studies of Western musicians are similar [6]). The best way to determine the interval values of the 12 basic notes in Carnatic music is by recording and analyzing slow passages played by several trained musicians in a relaxed and controlled environment. However, it is also possible that different (accomplished) musicians would intonate the same notes differently from one another, in which case we can only hope to infer the allowable limits of intonation rather than come up with a single numerical value for an interval.

Finally, we state that it is a given that the artists whose recordings we have analyzed are of the highest calibre and that the inconsistencies in intonation we have observed should be interpreted as being perceptually negligible in the context of a musical performance rather than indicative of any weakness in their abilities. Despite the apparent variability in intonation, it still takes many years of rigorous practice for a typical musician to satisfactorily reproduce these gamakas within perceptually acceptable limits.

5. ACKNOWLEDGEMENTS

We wish to thank Prof P. P. Vaidyanathan and Prof. Julius Smith who provided many useful comments on this paper.

6. REFERENCES

[1] Levy, *Intonation in North Indian Music: A Select Comparison of Theories with Contemporary Practice*, Biblia Impex Private Limited, New Delhi, India, 1982.

[2] Jairazbhoy and Stone, "Intonation in present-day north indian classical music," *Bulletin of the School of Oriental and African Studies*, vol. 26, pp. 119–132, 1963.

[3] Sambamurthy, *South Indian Music*, vol. 4, The Indian Music Publishing House, Madras, India, 1982.

[4] Wang, *Instantaneous and Frequency-Warped Signal Processing Techniques for Auditory Source Separation*, Ph.D. thesis, EE Dept., Stanford University, 1994.

[5] Talkin, "A robust algorithm for pitch tracking (rapt)," in *Speech Coding and Synthesis*, Kleijn and Paliwal, Eds., chapter 14. Elsevier Science, 1995.

[6] Burns, "Intervals, scales and tuning," in *The Psychology of Music*, Deutsch, Ed., chapter 7. Academic Press, 2nd edition, 1999.