A component-based approach to study the effect of Indian music on emotions

Venkata Subramanian Viraraghavan\textsuperscript{1,2}, Arpan Pal\textsuperscript{1}, Hema Murthy\textsuperscript{3}, R Aravind\textsuperscript{2}

\textsuperscript{1}Embedded Systems and Robotics, TCS Research and Innovation Labs, Bangalore
\textsuperscript{2}Department of Electrical Engineering, Indian Institute of Technology Madras
\textsuperscript{3}Department of Computer Science and Engineering, Indian Institute of Technology Madras

\{venkatasubramanian.v, arpan.pal\}@tcs.com, hema@cse.iitm.ac.in, aravind@ee.iitm.ac.in

Abstract

The emotional impact of Indian music on human listeners has been studied mainly with respect to ragas. Although this approach aligns with the traditional and musicological views, some studies show that raga-specific effects may not be consistent. In this paper, we propose an alternative method of study based on the components of Indian Classical Music, which may be viewed as consisting of constant-pitch notes (CPNs) providing the context, and transients, the detail. One hundred concert pieces in four ragas each in Carnatic music (CM) and Hindustani music (HM) are analyzed to show that the transients are, on average, longer than CPNs. Further, the defined scale of the raga is not always mirrored in the CPNs for CM. We also draw upon the result that CPNs and transients scale non-uniformly when changing the tempo of CM pieces. Based on the observations and previous results on the emotional impact of the major and minor scales in Western music, we propose that the effect of CPNs and transients should be analyzed separately. We present a preliminary experiment that brings outs related challenges.

Index Terms: Carnatic music, Hindustani music, Emotion, Constant-pitch notes, Transients

1. Introduction

The effect of music on the mental states of humans has been studied extensively [1, 2]. Typical studies in Western music focus on Western classical forms of music and other popular forms such as pop and heavy metal (e.g. [3, 4]). In addition, the effect of other audio, e.g. natural sounds, is also studied [5]. In almost all studies, personal preferences are acknowledged [1].

When analyzing the effect of Indian classical music, the traditional approach focuses on rāgas [6, 7, 8, 9]. This can be attributed to Indian musicological texts (e.g. [10]) associating rāgas with specific rasas (emotions). Some rāgas are said to cause rain or have soporific effects. An example of the latter is the rāga nilambari, but in an experiment [6] it did not affect sleep quality any differently from the rāga kalyāni, which is certainly not counted as sleep-inducing. When seen in the context of widely varying personal preferences, and given that a very small percentage of listeners consciously look out for rāga information, most rāga-specific studies have limited application. In fact, one study [8] shows that rāga-specific effects are inconsistent. Instead, components of Indian music may be more relevant. For example, parts of vocal music in Indian films use continuous pitch variation (gamakas), perhaps more than the keyboard-dominated background music.

As an analogy, in Western classical music, the effect of different types of chords on the mind has been studied. For example, the authors of [11] show that minor and dissonant chords are perceived as more unpleasant and sad respectively compared to major chords. Just as this analysis was performed on one of the basic elements (harmony) of Western classical music, we maintain that a similar study for Indian music is necessary, but in terms of the components of melodic music.

In this paper, we analyze two components – context or constant-pitch notes (CPNs), and detail, or transients – and propose that the emotional impact of these components must also be studied separately. This approach holds promise for generalization in that other forms of music that employ continuous pitch variation may also have similar effects.

The rest of the paper is organized as follows. Previous results on the analysis of the CPNs and transients in Carnatic music (CM) are summarized in Section 2. In the same section, the analysis of similar rāgas in Hindustani music (HM) is presented, which, re-confirming the result in [12], shows the large presence of transients in HM. We then argue in Section 3 that the emotional impact of the context and detail of Indian music need to be studied separately. The associated challenges are brought out in a preliminary experiment, whose results form the basis of the future work suggested in Section 4.

2. Context and Detail in Indian Music

2.1. Database and Definitions

In our analysis, we use subsets of the CompMusic database described in [14]. In this database, the pitch is extracted using the algorithm described in [15] and the tonic of the melody is identified by the method in [16]. The pitch tracking algorithm also marks silence segments. Other segments of pitch that are within 0.3 semitones from the segment’s mean, and whose best-fit line has a slope less than 1 semitone per second, are marked as CPNs [17]. All non-silence and non-CPN segments are called transients. Points where the pitch curves of transients change direction are called stationary points (STAs). See Figure 1 for an example of these terms.

Figure 1: (From [13]) CPN- and STA-examples in a pitch curve.
In this paper, a subset of four Carnatic rāgas (Figure 2) and four Hindustani rāgas (Figure 3) are chosen for analysis. They have an approximate correspondence: e.g. Mōohanam in Figure 2(a) corresponds to Bhūp in Figure 3(a), and similarly for sub-figures (b), (c) and (d). The total number of pieces analyzed is hundred (48 Carnatic pieces and 52 Hindustani pieces). To maximize the data analyzed, we consolidate all pieces in a rāga.

2.2. Analysis of some classical rāgas
As reported in [17, 13], the histogram of the mean pitch-values of CPNs longer than 150 ms, when folded to one octave, show sharp clusters that reflect a rāga’s properties. By contrast, the analysis of STAs is harder. A useful picture emerges only if the local maxima and minima of the transients are treated separately [13]. The CPN and STA plots for Carnatic rāgas are shown in Figure 2, and for Hindustani rāgas, in Figure 3.

For both CPNs and STAs, we found significant peaks – at least 0.01 times the maximum – and removed the false alarms manually. These peaks are noted in Table 1 for several rāgas. In the table, an entry ‘C’ indicates that the scale-note is a significant peak only in the CPN histogram of that rāga, and ‘S’, in any STA histogram but not in the CPN histogram. Scale-notes with significant peaks in both CPN- and STA-histograms are marked ‘CS’. The other scale-notes that are not locations of peaks, are marked with a dash. For reference, the cells for unexpected scale-notes in each rāga are shaded. Note that there are no peaks at unexpected scale-notes for CPNs.

The penultimate column of the table gives the average duration in milliseconds of a CPN (\(L_{CPN}\)) in that rāga, while the last column shows the value of \(\eta\), the ratio of the average duration of a contiguous chain of transients (CCTs) to \(L_{CPN}\). A CCT is any non-silence, non-CPN segment that is both preceded and followed by either a silence-segment or a CPN.

2.3. Dynamic behavior
It is important to note that the combination of CPNs and transients changes with time. An example for a piece in the rāga bhairavi is given in Figure 4. The percentage of CPNs, \(p_{CS}[n]\), in a window of length \(W_a\) ending at sample \(n\) is given by:

\[
p_{CS}[n] = \frac{\sum_{i<n} N_{W_a,n} \cdot L_{CPN}}{\sum_{i=0}^{n-1} N_{W_a,n} \cdot L_{CPN}} \times 100
\]

where \(N_{W_a,n}\) is the number of CPNs in the window, \(L_{CPN}\), the duration of CPN \(i\). \(N_{W_a,n}\) is the number of non-silence segments in the window, and \(L_i\), the duration of segment \(j\).

Observe that \(p_{CS}[n]\) in four-second windows \(W_a\) corresponds to 4 seconds) varies from as little as 10% to nearly 100% through the piece. In fact, in CM, it has been observed [18] that CPNs tend to scale by a much larger fraction \((8.7 \times \text{for the tōḍī example in [18]})\) than transients (only 2.6 \(\times\) for the same example) compared to the overall ratio \((3.5 \times)\) of the duration of similar phrases in two speeds. Thus, slow segments of CM tend to have a higher percentage of CPNs than fast phrases.

The following observations can now be made.

1. In all the CM and HM rāgas, the average transient (CCT) duration is larger \((1.4 < \eta < 2.1)\), by up to 100% (e.g. mōohanam in Table 1), than the average CPN duration. Further, in short windows of a music piece, CPNs can form as low as 10% of the music (Figure 4). Clearly, these transients play a bigger role [17] than, say vibratos and glissandi in Western music.

2. While the scale of a rāga comes out in the histogram of CPNs for the HM examples considered, it does not in some of the CM examples. This behavior is the norm rather than an exception for several CM rāgas.

2.4. Applicability to other forms of Indian Music
The two classical genres, Hindustani and Carnatic, are only two examples of Indian music, including myriad folk forms and film music in various languages, i.e. Bollywood and regional variants. In addition, Western styles including classical, pop, rock, rap etc. have been adapted. However, it is reasonable to expect that any form of (Indian) music can be explained by a combination of the context-detail model for Indian classical music and the Western classical view of harmonic music.

3. Emotions induced by Indian music
The scale-notes that typically appear as long CPNs in CM belong to the major scale and thus can be (generally, not always) associated with positive affect [19] (see [20] for the definition of positive and negative affect). The notes that are typically avoided as CPNs (with the exception of N3) are generally associated with negative affect. While this is true for the four Carnatic rāgas considered in this paper, there are others where it is not. For example, āḥhēri can be expected to have G2 and N2 as CPNs. Nevertheless, as a thumb rule, HM does not avoid these notes as CPNs, while CM seems to actively do so in several rāgas with such notes.

Point 1 of Section 2.3 shows that results on the emotional effect of forms of music employing profuse continuous pitch variation cannot be analyzed only with CPNs. Further, rāga-specific emotional impacts may show more intra-rāga variation than between rāgas [8]. A possibility is that CPNs and STAs have different impacts.

Given that CM is able to induce feelings such as pathos (e.g. pantunvarāḷī rāga) in listeners, even though point 2 holds, it is possible that these feelings may be induced only with fleeting movements to these scale-notes as STA-targets. To gauge this aspect, we conducted a simple experiment. This preliminary exercise and the associated challenges (given in bold font) in designing a more comprehensive experiment are described next.

3.1. Experiment Design
The following three types of segments, denoted by the letters in brackets, in the Carnatic rāgas bilahāri, yadukalākambhōji, and pantunvarāḷī were created automatically from violin ālāpanas.

1. Concatenate only CPNs of the major scale (M)
2. Concatenate all CPNs irrespective of the scale (C)
3. Concatenate segments with transients while ensuring that CPNs of only the major (T) scale were present

The effect of major-scale CPNs alone is thus the control set-up. The effect of the non-major-scale CPNs is measured in the second clip, and the effect of transients alone, in the third. We used original violin pieces with transients (T), but synthesized CPN-only pieces (M & C). The synthesized pieces sound artificial, but in our judgment, the fragmentation effects in the CPN-only pieces were worse in the original. Next, from these segments, smaller clips (4 to 7 seconds long), denoted by \(M_r\), \(T_r\), and \(C_r\), were manually constructed by choosing one or more similar sections of music. That is, the starting- and ending-notes, and the range of notes were similar in all the clips.
Figure 2: CPN and STA histograms for four Carnatic rāgas. Each rāga has 12 pieces with its total duration being more than one hour. The notes in each rāga are listed along with its name. Blue curves are used for CPNs; red curves are used for STAs that are maxima; and yellow curves, for STAs that are minima. Each curve has been independently normalized to add up to 1.

Figure 3: CPN and STA histograms for four Hindustani rāgas. Each rāga has at least 10 pieces, with its total duration being more than one hour. In 12 pieces, the vilambit (slower section) is treated separately. The color-coding and normalization are as in Figure 2.
of a rāga. As expected, for the rāgas yadukalakāmbhōji and bilahari, the clip $C'$ was nearly identical to the clip $M'$, and were not used. Eleven listeners, with varying tastes in music, rated every clip twice according to whether they perceived/associated the clip with five levels of valence (ratings in brackets): Very Sad (-2), Sad (-1), Neutral (0), Happy (1), Very Happy (2).

3.2. Indicative Results

Thirty five of the 154 ratings were consistent – that is, the ratings for the two times a clip was played, differed by less than $\delta = 1$. The counts of consistent ratings is given in Table 2. Consistent differences, relative to the reference clip, $M'$, are also counted. The second clip is seen as more sad or happy than the reference if the difference, $\text{rating}((T'_{\text{clip}}) - \text{rating}(M'))$, respectively is $<-2$, or $>2$. Any other difference is considered as ‘no change in the perceived valence.’

Table 2 confirms that major-scale-notes clips ($M'$) are usually felt as happy, and that non-major-scale CPNs (pantuvarāli $C'$) are usually felt as sad. However, it appears that fewer listeners perceive non-major-scale STAs (pantuvarāli $T'$) to be as sad as their CPN counterparts ($C'$, does this suggest conditioning?). The yadukalakāmbhōji $T'$ clip is rated as more sad than $M'$ four times out of 12, possibly due to its slow pace.

4. Discussion

The results in Section 3.2, though only indicative, pose some research questions that can form the basis of future work. The inconsistency in rating may be due to the small duration of the pieces: the authors of [21] ignore annotations of the first five seconds of music. With intra-rāga variation [8], the clip-duration should be long enough for the emotion to be perceived, but short enough to evoke a single emotion. Further, again in [8], the rate of onset-events in music was shown to correlate with emotional arousal. In the light of our observations, it is important to also include, apart from CPNs, events in the transient part of the pitch, which often do not have onsets.

We reiterate that the effects of the context and detail of melodic music need to be studied separately. However, many challenges (highlighted in Section 3.2) need to be overcome in designing experiments that conclusively separate these effects.

5. Acknowledgements

This research was partly funded by the European Research Council under the European Unions Seventh Framework Program, as part of the CompMusic project (ERC grant agreement 267583).
6. References


