To appear in Frontiers of Research on Speech and Music (FRSM)
November 11-12, 2016, Baripada, India.

PERCEPTUAL ANCHOR OR ATTRACTOR: HOW DO MUSICIANS PERCEIVE RAGA PHRASES?

Kaustuv Kanti Ganguli and Preeti Rao

Department of Electrical Engineering
Indian Institute of Technology Bombay, Mumbai.
{kaustuvkanti,prao}@ee.iitb.ac.in

Abstract- *A raga performance in Hindustani vocal music builds upon a melodic framework wherein raga-characteristic phrases are presented with creative variations while strongly retaining their identity. It is therefore of interest, for both music information retrieval and pedagogy, to understand better the space of "allowed" variations of the melodic motifs. Our recent study of melodic shapes corresponding to a selected raga phrase showed that variations in the temporal extent of a passing note within a characteristic phrase was perceived categorically by trained musicians. The work is extended here to non-prototypical melodic phrases. Several synthetic but musically valid versions of the phrase are generated from the canonical form and presented to musicians in a pairwise discrimination rating task. Results demonstrated better discrimination performance in the non-prototypical context than in the prototypical context. We interpret this finding to indicate that a category prototype may function as a "perceptual magnet", effectively decreasing perceptual distance, and thus discriminability, between stimuli. This paper provides a few insights into the nature of musical phrase categories in terms of their raga-belongingness.*

Keywords- Raga-characteristic phrase, behavioral experiment, perceptual magnet effect.

1. Introduction

Musicians are trained to produce and recognize raga phrases. An interesting analogy would be to imagine a phrase as a spoken word in a language that musicians understand. We want to present a musician with many acoustic versions (each slightly modified to a different extent from the "canonical" form, e.g., what might be stored in their long-term memory). We would like to know whether they are sensitive to the differences and measure how the physically measured acoustic signal differences relate to perceived differences. To answer the question how this would be useful for us, we expect music learners to make mistakes akin to the deviations in certain melodic aspects. If we can predict how a good musician responds to such stimuli, we can give proper feedback to the learner (correct/slightly incorrect/very wrong etc.). The question we ask is whether trained Hindustani musicians perform a memory abstraction for the raga characteristic phrases. Our recent work [1] investigated, through acoustic measurements followed by behavioral experiments through listening, the possibility of a canonical form or "prototype" of a raga characteristic phrase. In our context, a prototype may be considered as the phrase that serves to establish the raga around the initial phase of the performance. The case study was conducted for a characteristic phrase DPGRS in raga Deshkar. We first determine all the distinct independent dimensions of actual physical variability by observing actual instances from concerts. We would like to verify whether the existence of a 'prototype' only applies to raga-characteristic phrases or it extends to any melodic pattern. The chief objective of the current work is to investigate via perception experiments whether a non-characteristic melodic shape behaves like a prototypical melodic motif.

Researchers in the past [2] have used the term melodic 'predictors', in the context of music similarity, to refer to high-level quasi-independent musical features (pitch distance, pitch direction, rhythmic salience, melodic contour, and tonal stability). The authors proposed the algorithmic (dis)similarity measure to be a function (multiple linear regression) of these melodic predictors. For our case, stimuli should be generated with the appropriate modifications of the given melodic shape. Thus to obtain a canonical form of a phrase, we need to observe several instances of the phrase to infer the dimensions in which the variations take place and to what extent. This is because we should be able to create artificial stimuli by extrapolating on the obtained trend in the given dimensions. We aim to understand the "extent of dissimilarity" (with respect to the canonical form of the phrase) vs. the magnitude of change in stimulus on each of the dimensions. Our recent study [3] used standard music information retrieval (MIR) tools to explore melodic structures in a data-driven way, to validate certain musicological hypotheses. Judicious use of both data and knowledge can lead to building a cognitively-based computational model that could simulate the human-judgment of melodic similarity. Apart from its pedagogical value, this approach has potential applicability as a compositional aid.

The structure of the rest of the paper is as follows. We first review previous studies on human similarity-judgment in speech and music literature. Next we discuss relevant music concepts to decide on a suitable raga characteristic phrase for a case study. The following section discusses the preparation of suitable stimuli for a set of behavioral rating experiments and results. Finally, we summarize our findings and propose planned future work.

2. Background

Vempala and Russo [2] observed that melodic contour (directions of pitch change) was an important predictor in similarity of phrases with single note alterations. Authors motivated the importance of a cognitive basis for melodic similarity. Mullensiefen and Frieler [4, 5] showed that automated similarity measurements performed well in folk-song similarity from symbolic scores. From these papers we learn: (i) importance of using tested musically trained subjects, (ii) designing of stimuli, (iii) specifying the task and rating scale, and (iv) drawing conclusions about the predicting power of various representation-cum-similarity measures. In speech literature, authors [6, 7] in the past have reported categorical perception (CP) in the perception of phonemes. Prosodic phrases are also shown to be categorical in nature [8]. Authors have taken an analysis-by-synthesis framework to generate synthetic stimuli and perform perception experiments to investigate a possible presence of a `prototype' representation of a prosodic phrase. Several authors have noted top-down effects of musical expectancy interacting with lower perceptual processes. Among the first attempts of studying CP in music, Burns and Wards [9] observed that melodic intervals (sequential presentation of tones) elicited categorical perception effects in certain experimental paradigms. Barrett [10], McMurray et. al. [11] found that in the case of major chords, musical expectancy actually narrows a category, i.e., discrimination becomes sharper near the prototype (it acts like a perceptual anchor). This brings us to an interesting theory called the perceptual magnet effect (PME), where a prototype is expected to act either as a perceptual attractor or an anchor. This means that the sensitivity of a listener to discriminate between stimuli is either decreased (attractor) or enhanced (anchor) around a prototype, i.e., the perceptual space is warped with distinct behaviors in regions around prototypical shapes and non-prototypical shapes. The question to ask for the case of raga phrases is whether the prototypes act like attractors or like anchors.

2.1. Relevant music concepts

A raga performance can be thought of as a sequence of melodic motifs or characteristic phrases. The precise phrase intonation is so crucial that it acts as a major cue to raga identification by listeners and is well accepted as the foundational unit of a raga in the pedagogical tradition as well [12]. Though a characteristic phrase (lit. *pakad*) of a raga often holds a unique canonical form, considerable variability is observed among the instances of the same phrase in a raga performance. This variation usually involves multiple dimensions, such as pitch, time, timbre, energy dynamics etc. [13]. It is implied that these phrases are still highly recognizable by trained listeners [14]. In two dimensions (pitch vs. time), the captured similarity among phrases is either local or global: there can be micro-tonal variation on a particular note, or the relative tonal duration structure may vary as well. Repeated use of the same melodic motif brings out its inherent variability, but it is difficult to estimate the boundary of this variability space from the concert audio data. We propose a methodology to be able to gauge the fine limits of variability allowed for a melodic phrase within a raga framework. Our previous work [15] showed that musicians are able to abstract stylistic features in the raga alap to classify Hindustani and Carnatic music. This paper investigates whether the same idea extends to melodic phrases (and their improvisations) in terms of the identifiability of the corresponding raga.

We exploit musicological knowledge to choose the stimuli for this experiment. We choose a phrase which is characteristic of the raga by itself without any further context. There are certain regions in the melody which act like an unbreakable unit like a gestalt. One such phrase is the GRS phrase in raga Deshkar. Here the ‘R’ is a small step within the glide between ‘G’ to ‘S’, but its presence is a must. Thus this phrase is a good choice for a case study for melodic similarity. Though one might argue that the typicality of raga Deshkar lies in the GRS portion of the DPGRS phrase, but the context of DP provides the identity to be independently qualified as a raga Deshkar phrase. With our experience of observing raga phrases, raga Deshkar seems to be the most invariant across different artists which makes it a good choice for case study. We extract many instances of the GRS phrase from real-world concert audios by eminent Hindustani vocalists to study the systematic melodic variations.

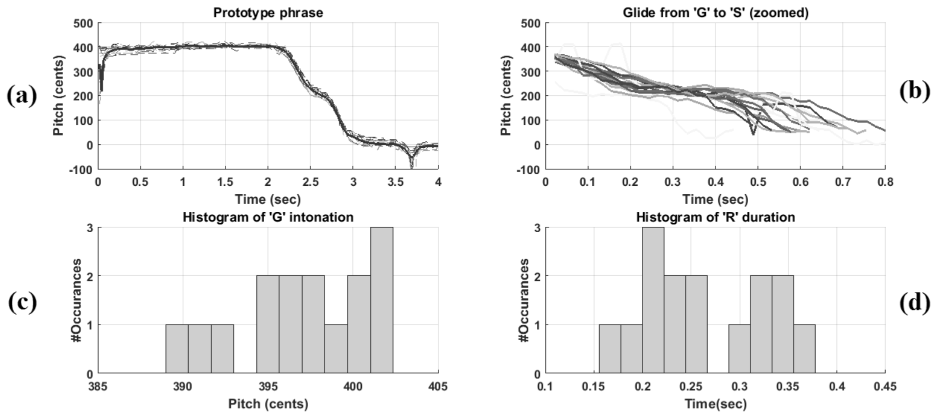

Figure 1. Acoustic measurements on fifteen GRS phrases in raga Deshkar performance (alap and madhyalaya bandish) by Ajoy Chakrabarty. (a) All phrases with the centroid (proposed prototype) in bold, (b) glide from ‘G’ to ‘S’ including the passing ‘R’ note, (c) histogram of the mean intonation of the ‘G’ note segments, and (d) histogram of the duration of the ‘R’ note segments.

3. Method and Material

The steps for stimulus creation from audio, pitch contour stylization, and model space variations is borrowed from our previous work [16]. The first group of stimuli belong to the characteristic DPGRS phrase of raga Deshkar [16]. The second group of stimuli is a descending melodic sequence DPMGRS which is not a characteristic phrase of any particular raga. Figure 2 illustrates the comparison of the two phrases. The experiment belongs to the ‘similarity rating’ paradigm. This involves measurement of melodic distance in an AX phrase-pair configuration in a differential discrimination setup. Given the resynthesized melodies A and its variant X, subjects are asked to rate whether the stimuli are same or different. We hypothesize that musicians would be less sensitive to the dissimilarities between the AX pair if either of them is one of the prototypes (or close to the prototype centre). Similarly the AX pair away from the prototype is speculated to be more sensitive to small differences in musicians’ rating. We report results obtained from eight trained Hindustani musicians’ response, four of them are common to both stimulus groups. All responses were recorded in the same environment, apparatus, and settings.

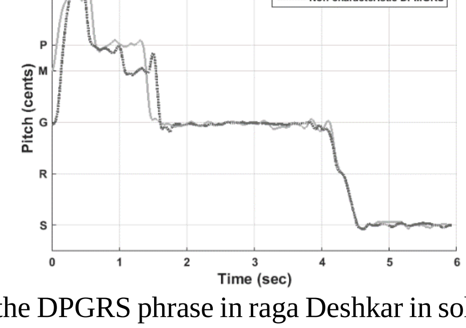

Figure 2. Stylized contour for the DPGRS phrase in raga Deshkar in solid line and the DPMGRS phrase (not-so-characteristic phrase for any raga) in dotted line. Note that the GRS phrase is totally overlapping (except for a 0.2 sec shifted ‘G’ onset) and the difference lies only in the context of the GRS phrase. This allows us to apply the same transformations as discussed in [1, 16].

4. Results and Discussion

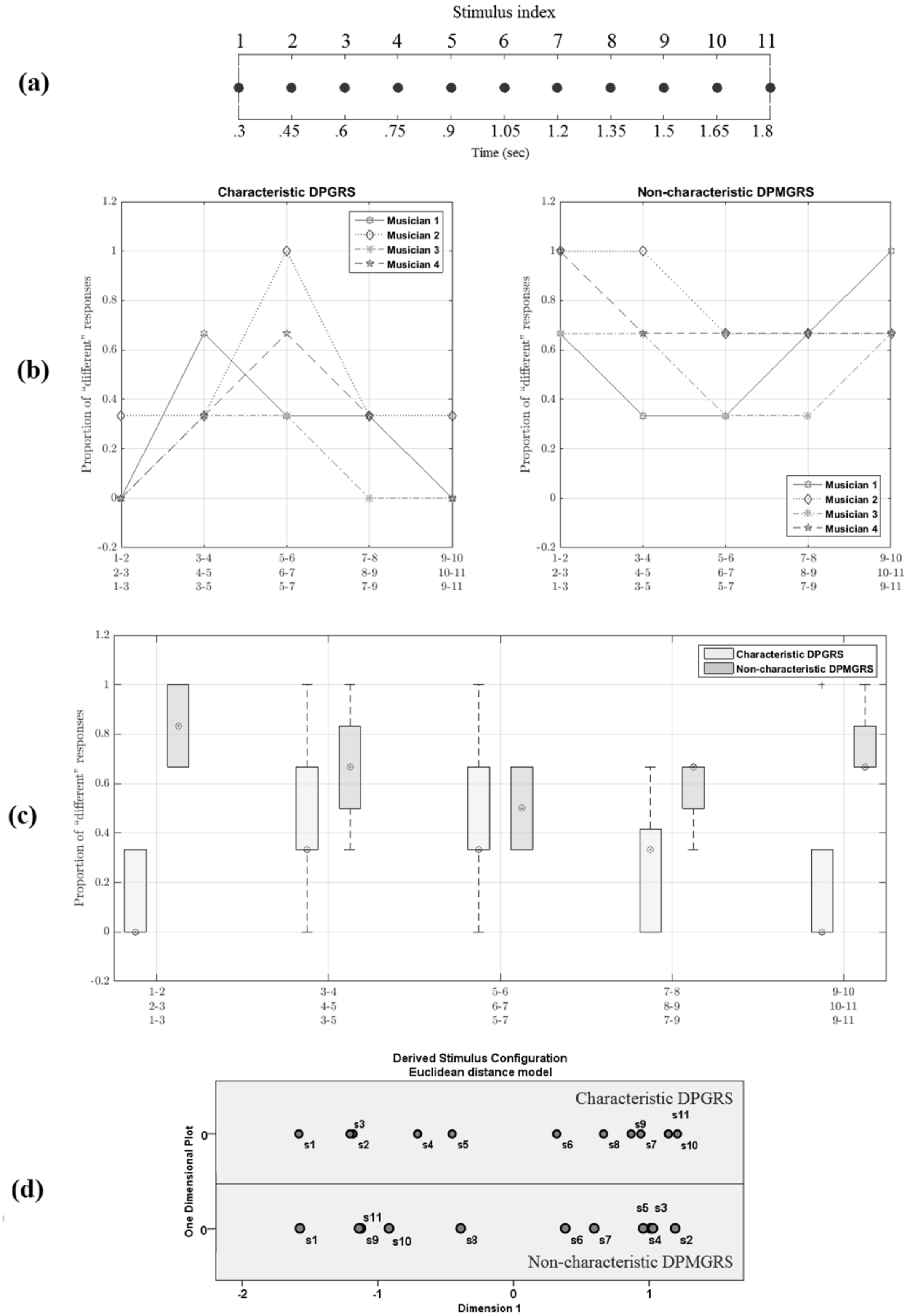

Figure 3. Individual and group plots for the response obtained from 8 musicians. (a) Model space: stimuli indices and corresponding duration of passing 'R' note, (b) individual responses for 4 (common) musicians for characteristic DPGRS phrase (left) and non-characteristic DPMGRS phrase (right), (c) comparison of the boxplots for the aggregate response, and (d) perceptual space: multidimensional scaling (1-D) of the (dis)similarity matrices of the average response for the two stimulus groups.

The aim of the experiment, as discussed earlier, is to find a mapping between the model space and the perceptual space. Figure 3 (a) shows the model space where the stimulus index 1 corresponds to the 'prototype' shape of the characteristic DPGRS phrase in raga Deshkar. The variations in the model space only bears elongation of the passing 'R' note, the corresponding duration thereof is shown in the lower pane. The 'R' duration of the rightmost (index 11) stimulus is 6 times as that of the prototype and we safely assume this to be a non-prototypical shape. Figure 1 (d) shows the perceptual space as obtained from MDS of the (dis)similarity matrix of the musicians' response. The results of the differential discrimination experiment is shown (Figure 3 (b) and (c)) in the form of "proportion of 'different' responses" for closely spaced stimuli in the model space. The interpretation of the figures is summarized as follows: for each column (e.g., first column {1-2,2-3,1-3}) we show the proportion of the stimulus-pair marked as 'different' for all pairwise comparisons across 3 stimuli in the model space. The median being close to 0 indicates poor discriminability and vice-versa.

4.1. Characteristic DPGRS

Figure 3 (c) shows averaged response for 8 Hindustani musicians. We observe the median to be close to 0 for the left-most column and gradually increasing to the right. The low discriminability around the {1-2,2-3,1-3} pair indicates presence of a 'prototype'. The median for the subsequent columns increase gradually, but again decreases at the last column {9-10,10-11,9-11}. This indicates possible presence of another prototype around {9-10,10-11,9-11}. We perform a multidimensional scaling (MDS) on the disparity matrix for Hindustani musicians and project to a 1-dimensional space to compare with the model space of variations along 'R' duration. Figure 3 (d) shows that there is a warping in the perception space while the order of stimuli is preserved. The interpretation is that the prototype at the left works as a perceptual attractor where the discriminability is poor for trained Hindustani musicians. The clustering of stimuli around the right-most end is suggestive of another prototype in the hypothesized non-prototype region. The interviews with the subjects confirmed that they perceived raga Bhoopali for stimuli with an extended 'R', which confirms the decrease in discriminability at the right-most column.

4.2. Non-characteristic DPMGRS

The individual differences as shown in Figure 3 (b) are rather interesting. The proportion of 'different' responses is observed to be close to 1 for the left-most and right-most columns, indicating high discriminability and hence nullifying the presence of a possible 'prototype'. The response dips for the intermediate columns, but the 4 musicians show difference in the locations of the dips. Upon interviewing the musicians, one striking find was that each of them assumed the non-characteristic DPMGRS phrase to belong to a particular raga (either of Shuddh Kalyan, Yaman, Maru Bihag, or Vachaspati). This is an interesting phenomenon, because the non-characteristic DPMGRS was recorded not to be typical of any particular raga (a neutral descending sequence with a high overlap with the characteristic DPGRS phrase). Musicians seem to have anchored to one particular raga (closest to each individual's opinion) and their perception was guided by this assumption. Thus, despite finding a unique 'prototpe' for this phrase, we are able to confirm the perceptual attractor effect.

5. Conclusion & Future work

Our findings suggest that trained Hindustani musicians perceive melodic phrases categorically, with less sensitivity to small changes around the prototype region. This supports the hypothesis that prototypes work as a perceptual attractor where musicians are less sensitive to small variations and they tend to perceive the phrase holistically. That perception of a raga characteristic phrase showed perceptual magnet (attractor) effect, we carried out the same set of experiments with the same GRS phrase in a context where it is not characteristic of any particular raga. The results indicate the same (perceptual attractor) effect, with the constraint that the location of the 'prototype' depends on each musician's assumption (of the closest raga). In sum, this paper gathers evidences how imparting music knowledge into a data-driven computational model helps modelling human-judgment of melodic similarity. This cognitively-based model can be useful in music pedagogy, as a compositional aid, or building retrieval tools for music exploration and recommendation.

Neuro-musicologists take a few principled approaches to study the effect of incongruity in music stimuli via EEG experiments, the aim being modelling the musical expectancy. One planned future work is to validate our findings with a neurophysiological experiment. The immediate direction would be to expand the behavioral experiments with a diverse subject-base (e.g., Carnatic musicians). MDS with 1 dimension, for the characteristic DPGRS phrase, preserved the stimulus order from the model space to the perceptual space. This motivated us to try the same measure for the non-characteristic DPMGRS phrase, the outcome is not the same though. One possible reason could be non-availability of enough dimensions to model the (dis)similarity space. However, it is difficult to interpret the axes for a higher dimensional MDS and hence is posed as a future work.

Acknowledgement

This work received partial funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ ERC grant agreement 267583 (CompMusic).

References

[1] K. K. Ganguli and P. Rao, “Exploring melodic similarity in Hindustani classical music through the synthetic manipulation of raga phrases,” in Proc. of the Cognitively-based Music Informatics Research Workshop (CogMIR), New York, August 2016.

[2] N. N. Vempala and F. A. Russo, “A melodic similarity measure based on human similarity judgments,” in Proc. of Int. Conf. on Music Perception and Cognition (ICMPC), July 2012.

[3] K. K. Ganguli, S. Gulati, X. Serra, and P. Rao, “Data-driven exploration of melodic structures in Hindustani music,” in Proc. of Int. Soc. for Music Information Retrieval (ISMIR), August 2016.

[4] D. Mullensiefen and K. Frieler, “Measuring melodic similarity: Human vs. algorithmic judgments,” in Proc. of Interdisciplinary Musicology, 2004.

[5] D. Mullensiefen and K. Frieler, “Modelling experts’ notion of music similarity,” *Musicae Scientiae*, 2007.

[6] B. Pajak, P. Piccinini, and R. Levy, “Perceptual warping of phonetic space applies beyond known phonetic categories: evidence from the perceptual magnet effect,” *Journal of the Acoustical Society of America*, 136(4), 2107-2107, 2014.

[7] A. M. Liberman, K. S. Harris, H. S. Hoffman, and B. C. Griffith, “The discrimination of speech sounds within and across phoneme boundaries,” *Journal of Experimental Psychology*, vol. 54, no. 5, pp. 358–368, November 1957.

[8] J. Rodd and A. Chen, “Pitch accents show a perceptual magnet effect: Evidence of internal structure in intonation categories,” in Proc. of Speech Prosody, 697-701, 2016.

[9] E. Burns and W. Ward, “Categorical perception—Phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic intervals,” *Journal of the Acoustical Society of America*, 63, 456–468, 1978.

[10] S. Barrett. “The perceptual magnet effect is not specific to speech prototypes: new evidence from music categories,” *Speech hearing and language: Work in progress*, 11 (1999): 1-16.

[11] B. McMurray, J. L. Dennhardt, and A. Struck-Marcell, “Context Effects on Musical Chord Categorization: Different Forms of Top-Down Feedback in Speech and Music?,” *Cognitive Science*, 32(5), 893-920, 2008.

[12] K. K. Ganguli, “How do we ’See’ & ’Say’ a raga: A Perspective Canvas,” *Samakalika Sangeetham*, vol. 4, no. 2, pp. 112–119, October 2013.

[13] K. K. Ganguli and P. Rao, “Tempo dependence of melodic shapes in Hindustani classical music,” in Proc. of Frontiers of Research on Speech and Music (FRSM), March 2014, pp. 91–95.

[14] P. Rao, J. C. Ross, K. K. Ganguli, V. Pandit, V. Ishwar, A. Bellur, and H. A. Murthy, “Classification of melodic motifs in raga music with time-series matching,” *Journal of New Music Research (JNMR)*, vol. 43, no. 1, pp. 115–131, April 2014.

[15] A. Vidwans, K. K. Ganguli, and P. Rao, “Classification of Indian classical vocal styles from melodic contours,” in Proc. of the 2nd CompMusic Workshop, 2012, pp. 139-146, Turkey.

[16] K. K. Ganguli and P. Rao, “Discrimination of melodic patterns in Indian classical music,” in Proc. of National Conference on Communications (NCC), February 2015.