



Visualizing Carnatic music as projectile motion in a uniform gravitational field

Venkata Viraraghavan^{1,3}, Rahul Gavas¹, Hema Murthy², R Aravind³

¹Embedded Systems and Robotics, TCS Research and Innovation, Tata Consultancy Services

²Department of Computer Science and Engineering, Indian Institute of Technology Madras

³Department of Electrical Engineering, Indian Institute of Technology Madras
venkatasubramanian.v@tcs.com, rahul.gavas@tcs.com, hema@cse.iitm.ac.in,
aravind@ee.iitm.ac.in

Abstract

There are several examples of visualization of Western classical music, many of which use the score as a reference. By contrast, a standard descriptive notation that serves the same purpose for Indian music is not available. In this paper, we draw upon recent transcription approaches for Carnatic music (CM) and propose an automated visualization procedure. It is formalized by an analogy between the pitch movement and projectile motion in a uniform gravitational field. During constant-pitch notes (CPNs), the projectile moves horizontally on ledges at constant speed. Outside CPNs, reflectors are placed at the points where the pitch curve changes its direction. The positions of reflectors and ledges are quantized, and the motion between them is under the influence of gravity. We then suggest that the non-uniform scaling of CPNs, silence, and transients (i.e. music outside CPNs) in CM matches the analogy. Finally, the projectile motion equations are used to interpolate pitch curves for CM synthesis. Perceptual tests rate it on par with existing interpolation schemes. In addition, experts felt that visualization is needed to correct errors in descriptive notation.

Index Terms: Carnatic music, Music visualization, Projectile motion

1. Introduction

Carnatic music (CM) is a sophisticated form of Indian music, prevalent in south India. It is characterized by continuous pitch movement called *gamakas*, between musical notes (called *svaras*); even individual *svaras* are often rendered with *gamakas*¹. A mapping between *svaras* in CM and notes in Western music (WM) is shown in Table 1. The mapping is ambiguous when continuous pitch variation is involved, which makes up a significant portion ($\approx 50\%$) of CM renditions. Although some form of written notation has evolved to represent *gamakas* [2, 3], CM does not yet have a descriptive notation needed to enable computer synthesis. Traditionally, the music has been taught orally, and *gamakas* are learnt by example; prescriptive notation is thus more of an aid to memory. For WM, the staff notation can capture pitch, rhythm, polyphony, and basic ornamentation such as vibrato, and, by contrast, is descriptive enough for computer synthesis.

The staff notation already has the ingredients of music visualization. Pitch increases from bottom to top (y-axis) and time proceeds from left to right (x-axis). Yet, even semi-automated visualization is painstaking, especially if there is a mapping between the music and the video. In one such example [4], shapes are mapped to instrumental groups, and colors, to pitch; the shapes light up “outside-in” when the corresponding notes are

Table 1: *Svara* positions in an octave for Carnatic (CM) and Western music (WM). The tonic is set as C.

CM	S	R ₁	R ₂	G ₂	G ₃	M ₁	M ₂	P	D ₁	D ₂	N ₂	N ₃
WM	C	C#	D	D#	E	F	F#	G	G#	A	A#	B

played. An example of *kinetic art*, e.g. Beethoven’s fifth symphony [5], uses line riders that move on, and jump across, sloping curves/lines. In [6], metal marbles roll on stacked dominoes, fall between them, or move around circular magnets. In these examples, the videos are created with varying levels of manual intervention. Visualization may also be based on emotion [7], and serve as a tool to study the emotional impact of music [8].

Indian music, despite its profusion of continuous pitch variation, has surprisingly few visualization attempts. Of course, there are abstract examples such as the personification of a *rāga*, *svara*, etc. in paintings and sculptures, and Indian dance forms depict the meaning/interpretations of lyrics. However, these are not cases of visualizing the characteristics of the music. Instead, the textual descriptions of *gamakas*, which evoke images in the readers’ minds, are better examples. These descriptions are often in terms of fingers moving on a *vīṇā* [2], or along the fingerboard of a violin. The Gamaka box [9] is similar to the staff notation, and has a visual representation of *gamakas*. Like the descriptions, it mainly aids performers. On the other hand, as far as we are aware, audiovisual perception of Indian music by analysis has only one example [10], which plays back the unprocessed pitch curve synchronized to the music, with manually added metadata.

1.1. Our Contribution

In this paper, we draw an analogy between pitch curves in CM and projectile motion in a uniform gravitational field. A fully automated visualization procedure based on this analogy is then presented. Although applicable to Hindustani music too, the focus of this paper is CM. We suggest that the analogy can be interpreted meaningfully to explain the nature of speed-change in CM. We also introduce a parabolic interpolation scheme mimicking projectile motion and show that is perceptually comparable to existing interpolation schemes used in the CM synthesis.

1.2. Definitions and dataset

We reuse the following definitions [1] in this paper. Constant-pitch notes (CPNs) are segments of pitch that are at least 80 ms long and do not deviate from their mean by more than 0.3 semitones. *Transients* are non-silence sections of music that are not CPNs. Stationary points (STAs) are local maxima and minima of transients. *Anchors* are CPNs that are adjacent to STAs. Figure 1 shows some examples of these terms.

¹See [1] for further explanations of CM terms.

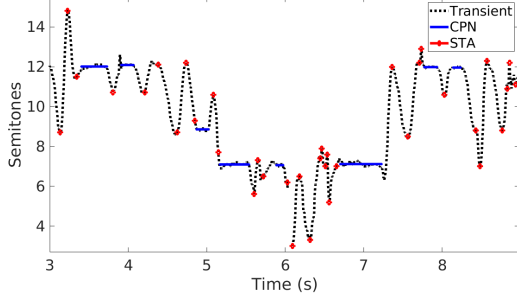


Figure 1: *Transients, CPNs, and STAs of a pitch curve.*

We use the dataset of [11] that has the pitch of the dominant source, typically the voice, of 474 renditions in 40 *rāgas* estimated according to [12] in 46 ms-windows every 4.44 ms $= 1/f_w$. This pitch is converted to semitones relative to the identified tonic [13] and CPNs are extracted according to [1, 14]. STAs are obtained as in [1] to remove redundant maxima and minima. The index of the sliding window, and the pitch in semitones of CPNs/STAs are relevant to the analogy. Note that, if CPNs/STAs are used to model the music, the analogy depends neither on the exact algorithms to extract CPNs/STAs, nor on those for transcription and quantization.

2. Visualizing CM as projectile motion

The analogy of CM pitch curves with projectile motion is developed after a brief recap of projectile motion.

2.1. Recap of projectile motion

The important equations of projectile motion in a uniform gravitational field are summarized below. Derivations and further details can be found in [15]. Let the magnitude of the acceleration due to the uniform gravitational field be g . ‘Downward’ is defined as the direction of the field. Displacement in the horizontal direction (x -axis) increases from left to right, and in the vertical direction (y -axis), from bottom to top. Let the horizontal component of the velocity of a projectile be $v_x(t)$ and let the vertical component be $v_y(t)$. Further, let the horizontal and vertical components of the initial velocity be u_x and u_y respectively. For upward (rightward) movement, u_y (u_x) is positive, and for downward (leftward) movement, u_y (u_x) is negative. Then, the components of the projectile’s displacement and velocity at time t are related by the following equations.

$$v_x(t) = u_x \quad (1)$$

$$x(t) = u_x t \quad (2)$$

$$v_y(t) = u_y - gt \quad (3)$$

$$y(t) = u_y t - \frac{1}{2}gt^2 \quad (4)$$

2.2. Analogy with projectile motion

In the proposed analogy with projectile motion, CPNs are modeled as ledges along which a projectile moves/rolls from left to right at a constant horizontal speed. The projectile is not shown on screen during silence. For Indian music in general, and CM in particular, the duration of music outside CPNs is invariably *higher* than the the duration of CPNs [16]. The rest of this section details the crux of the analogy: between projectile motion and the music traversing between STAs of transients.

Table 2: *Mapping music to projectile motion, with parameter α*

Music	Projectile motion
Sliding window index, w	Horizontal distance in meters, x
Pitch in semitones, $p(t)$	Height, $y(w/u_x)$ in meters
Scaled sliding rate, αf_w	Horizontal speed in m/s , u_x
<i>Gamaka</i> rate, $r(t)$	Vertical speed in m/s , $v_y(t)$

We define a Transcription Phrase (TP) as the music between two CPNs/silence segments. Let there be K STAs in a TP occurring at window indices $\{w_1, w_2, \dots, w_K\}$, with their quantized pitch values in semitones: $\{q_1, q_2, \dots, q_K\}$. Further, we explain the analogy for the case of a TP starting and ending in two CPNs, whose pitch values are denoted by $q_0 = c_1$ and $q_{K+1} = c_2$. Their window indices are w_0 , denoting the end of the first CPN, and w_{K+1} , denoting the start of the second CPN. The corresponding time instances are $t_k = w_k/u_x$ for $k = \{0, \dots, K+1\}$. It is straightforward to extend this analogy to TPs that start and/or end in a silence-segment.

The mapping of the musical parameters to those of projectile motion is given in Table 2. It has been observed that the duration of transitions between STAs is usually within 200 ms [1], and longer traversals often do not take longer time. Therefore, it is impossible to map to free-fall ($u_y = 0$) varying magnitudes of transitions in the same time interval. Instead, for each pitch traversal from STA $k-1$ to STA k , a corresponding initial vertical velocity u_{k-1} has to be evaluated with $\Delta t = t_k - t_{k-1}$:

$$u_{k-1} = s_{k,k-1} \frac{|q_k - q_{k-1}|}{\Delta t} + \frac{1}{2}g\Delta t \quad (5)$$

where $s_{k,k-1} = 1$ for upward movement ($q_k \geq q_{k-1}$), or $s_{k,k-1} = -1$ for downward movement ($q_k < q_{k-1}$), and α is explained below. With $p(t_{k-1}) = q_{k-1}$ and $p(t_k) = q_k$, for $t_{k-1} \leq t < t_k$, equations (3) and (4) change to:

$$y(t) = u_{k-1}(t - t_{k-1}) - \frac{1}{2}g(t - t_{k-1})^2 \quad (6)$$

$$v_y(t) = u_{k-1} - g(t - t_{k-1}) \quad (7)$$

Two accompanying conditions are $\max(y(t)) = \max(q_{k-1}, q_k)$ and $\min(y(t)) = \min(q_{k-1}, q_k)$. If these conditions are not satisfied, the projectile will overshoot (or undershoot) q_k before returning to it at $t = t_k$. This behavior alters the nature of the pitch curve by introducing a new STA at the point where the projectile turns. It is especially detrimental to *gamaka*-heavy Carnatic music. The parameter α has been introduced to enable ensuring this condition empirically. In our implementation we chose $\alpha = 0.2$ to match $g = 9.81 m/s^2$ and $f_w = 225 Hz$. With u_k depending on the subsequent STA, the projectile must be imparted varying (but precise) velocities at each STA. We model these as collisions that alter the speed of the projectile so as to exactly reach the next STA. Note that outside the collisions, the projectile is completely under the influence of gravity, an important aspect of interpreting the analogy in Section 3.

The analogy above is used in constructing a video that is synchronized to the original song. The resulting audiovisual is the output of visualization of CM. The CPN-STA model, quantization of their pitch values, and the analogy in creating the audiovisual, set apart this work from [10].

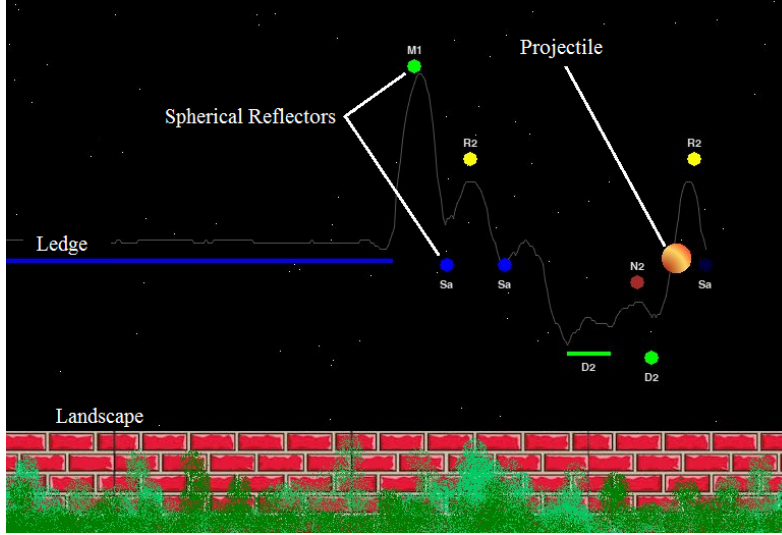


Figure 2: Snapshot of visualization canvas with ledges, reflectors, and projectile in motion for an excerpt in the *rāga bhairavī*. Svaranames unobtrusively mark ledges and reflectors; the original pitch contour is also shown as faint line.

2.3. Video construction

The video is generated in python with animation reused from the Pygame tool [17] on the lines of [18]. The projectile (orange ball of Figure 2) moves as described in Section 2.2. The bottom of the video shows a landscape of bricks and bushes to provide a feel of the earth’s surface. It is needed to convey that the projectile is moving under gravity. Other elements include the ledges and reflectors of Section 2.2, wherein the ball rolls straight on the former and bounces against the latter. Ledges are shown using straight lines and reflectors, using circles. The music piece is divided into chunks at most three seconds long, and each is shown using a different canvas, devoid of the elements of its predecessor. In each canvas, stars are added to give the feel of night time, and their positions are randomized to indicate the change of canvas. The regions of silence, as detected by [12], are shown as static canvases without ledges, reflectors, or the ball. The colors of the ledges and reflectors are based on note-triads as follows. S, G_3, P : blue; $M_1, D_2, (S)$: green; and $(P), N_3, R_2$: yellow; notes in brackets already have colors assigned in a previous triad. The remaining notes are assigned colors as follows. G_2, N_2 : brown; and R_1, M_2, D_1 : red. The color of the ball was chosen based on feedback from colleagues in our labs. A ledge/reflector lights up as the ball approaches it, and stays lit.

3. Interpretation of the analogy

Our analogy has an interpretation that can explain, to a first approximation, the observed scaling of CM across speeds. Consider the three speeds of rendering a single *svara*, R_i , in *kēdāragouḷa rāga* [19]. The pitch contours (estimated using [20]) are shown in Figure 3a. Transients are manually marked with red curves and CPNs, with blue lines. It is fairly evident from the figure that the transients and CPNs do not scale the same way. To aid seeing this better, Figure 3b has the pitch contours of the faster renditions shifted so that the transients are aligned. The figure clearly shows that CPNs are scaled down much more than the transients. This is analogous to speech, where the duration of consonants is preserved across different

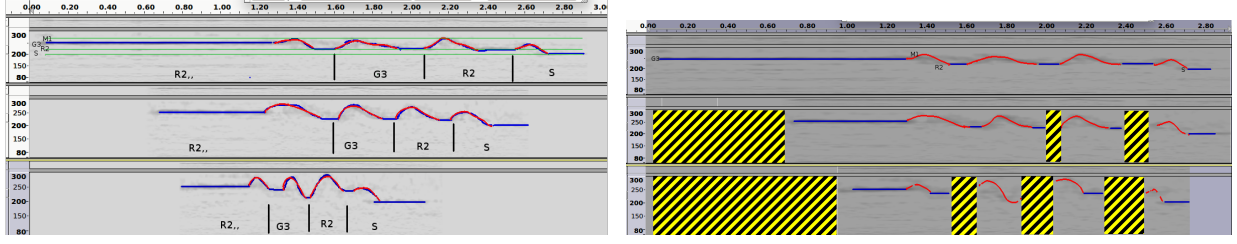
speeds, while the duration of vowels suffers significantly. In order to identify words that are spoken, consonants are necessary. The importance of vowels cannot be undermined, though: the distinction between beet and bit is just the vowel.

The example discussed so far was analyzed manually. The scalability of this approach is established by analyzing a particular class of items, called *varṇams*, where some lines (about a third of the composition) are usually rendered in two different speeds. In our database, we found six major *rāgas* with *varṇams* rendered in them by five musicians; these numbers are given in Table 3. For each *varṇam* rendition, the total duration of CPNs, transients and silence segments in the first and second speeds are obtained. Next, the ratios of the durations in the first speed to the corresponding durations in the second speed were found. The overall ratio is defined as the total duration in the first speed to that in the second². Table 3 shows clearly the large difference in the ratios of the CPNs, transients, and silence³. Specifically, CPNs and silence are extended when speed is halved, and in [21], transients are only moved outside them to halve speed.

An intuitive analogy with projectile motion is now drawn: If CPNs are viewed as ledges, they can be extended or contracted depending on the speed of the music, while STAs continue to stay as reflectors. The projectile can move along the ledges for longer when the music-speed is lowered because they provide the normal reaction to balance its weight. However, as soon as it leaves a ledge, it follows projectile motion under uniform downward acceleration, explaining the invariance of transient-duration across music-speeds. For algorithms such as [21], the nature of these reflectors does not change with music-speed. If transients are changed, reflectors will change accordingly; yet, the motion between them continues to be projectile motion. This interpretation also suggests that projectile motion curves can be used in CM synthesis. Such synthesis outputs and the visualization are evaluated next.

²The second speed is actually twice that of the first, but it is common for the lines of a *varṇam* to be repeated in the first speed and not in the second. This explains why the overall ratio is well above 2 in Table 3

³*Rāgas* with one rendition are not affected by variability in tempo across pieces (first four rows) and other *rāgas* show similar behavior.



(a) Aligned at the centers of the phrase in each speed.

(b) Silence (hatching) introduced for alignment at transients.

Figure 3: Pitch contours of the rendering of the phrase $R_2G_3R_2S$ in the *rāga kēdāragouḷa*, in three speeds. Transients have been manually marked with red curves, and CPNs, with blue lines. The x-axis is time in mm:ss format and the y-axis, pitch in Hz.

Table 3: Duration and ratios of constant-pitch segments and transients in two speeds of the first parts of several varṇams in six rāgas.

Rāga	No. of varṇams	Duration in the 1 st speed (seconds)				Ratio (1 st speed to 2 nd speed)			
		CPNs	Transients	Silence	Overall	CPNs	Transients	Silence	Overall
<i>Tōḍī</i>	1	37.7	69.7	10.5	117.9	8.66	2.61	3.79	3.49
<i>Bhairavī</i>	1	56.0	100.2	64.1	220.3	10.27	2.58	5.44	3.93
<i>Kāmbhōji</i>	1	52.1	79.3	29.1	160.5	6.2	1.9	2.8	2.7
<i>Śankarābharāṇam</i>	1	43.6	87.1	11.7	142.4	12.3	3.48	2.33	4.23
<i>Sahānā</i>	3	144.4	237.2	78.9	460.5	9.5	2.8	3.8	3.8
<i>Kalyāṇī</i>	2	114.6	220.1	59.6	394.3	10.1	3.0	4.5	4.1

Table 4: Average ratings out of 5, their signed and absolute differences, and preferences for the cosine (Cos) and projectile motion (PM) interpolation schemes.

Rāga	Rating		Difference	
	Cos	PM	Signed	Absolute
<i>Kāmbhōji</i>	3.3	2.9	-0.4	0.8
<i>Śankarābharāṇam</i>	3.2	3.1	-0.1	0.8
<i>Tōḍī</i>	3.4	3.6	0.1	0.2
<i>Bhairavī</i>	3.4	2.9	-0.5	0.6
Preference	29%	17%	54% differed by ≤ 0.5	

4. Evaluation and Results

The mapping from musical movements to projectile motion in Section 2.2 can be used as an interpolation scheme for CM by treating (6) as the pitch. We have thus introduced a quadratic interpolation scheme for *gamakas*. Four pieces were transcribed according to [14] to obtain quantized CPNs and STAs. The pitch curve between CPNs and STAs was interpolated using cosine curves, and using the projectile motion equations described in Section 2.2. These were played in a randomized order and listeners were requested to rate the synthesized samples for adherence to the *rāga*. Eight experts conversant with identification of *rāgas*, took the test⁴. The results in Table 4 show that it is not easy to tell apart quadratic interpolation based on projectile motion from cosine interpolation (54% of the ratings had no preference for either). We observe again that the CPN-STA model is compatible with many interpolation schemes. The first three degrees of polynomials (linear [14], quadratic in this paper, cubic [22, 23]) and cosine curves [24, 14], all seem to work.

We interviewed two performing musicians among the experts to obtain feedback on using the video for correcting the descriptive notation comprising CPNs and STAs. Although they

acknowledged that the detail was sufficient, they did not identify any errors. We then asked them to consider instructing their student to practise the *gamakas* as transcribed. In this role of a teacher, they corrected errors in the video, e.g. in the *rāga śankarābharāṇam*, a $P-G_2-P$ -*gamaka* was corrected to $P-G_3-P$. The initial hesitation in making corrections is explained by the tolerance for errors in STAs, which are not perceived in the synthesized audio [14]. Visualization thus helps annotators in verifying transcription outputs based on music theory/practice when perceptual listening tests are inconclusive.

5. Conclusions and future work

In this paper, an analogy was drawn between pitch movements in CM and projectile motion in a uniform gravitational field. CM was modeled as consisting of CPNs and transients, the latter characterized by STAs. CPNs were mapped to ledges, and STAs, to reflectors. Continuous pitch variation outside the ledges and reflectors was mapped to projectile motion. The analogy can explain the non-uniform scaling of CPNs and transients across speeds. Projectile motion curves were then used to interpolate between CPNs and STAs in CM synthesis. Experts evaluated this scheme and the cosine interpolation scheme comparably. They felt the need for visualization (the analogy is optional) in evaluating transcriptions. In doing so, they should be asked to examine short segments (say three seconds long) while playing the role of a teacher correcting their student.

6. Acknowledgments

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). We thank Ms. Meena Singh and other colleagues, who gave initial feedback on the visualization, the participants of the perceptual tests, and Mr. Jom Kuriakose for his help in setting up the tests.

⁴https://www.iitm.ac.in/donlab/pctestmusic_video/index.html?owner=venkat2&testid=test1&testcount=4

7. References

- [1] V. S. Viraraghavan, R. Aravind, and H. Murthy, "A statistical analysis of gamakas in carnatic music," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, 2017, pp. 243–249.
- [2] S. Dikshita, *Sangita Sampradaya Pradarshini*, 1904, 2008 web version. [Online]. Available: <http://www.ibiblio.org/guruguha/ssp.htm>
- [3] T. K. G. Rao, *Compositions of Mudduswami Dikshitar*. Gana-mandir Publications, 2006.
- [4] "Beethoven, Symphony 3 ("Eroica"), 4th movement (finale)," accessed: 2019-05-18. [Online]. Available: <https://www.youtube.com/watch?v=ZkP3XVkJxL4>
- [5] "Line Riders - Beethoven's 5th Symphony," accessed: 2019-05-18. [Online]. Available: <https://www.youtube.com/watch?v=vcBn04IyELc>
- [6] "Marbles, Magnets, and Music (Synchronized)," accessed: 2019-05-18. [Online]. Available: <https://www.youtube.com/watch?v=GCTyFp0PpHE>
- [7] J. Grekow, "Emotion based music visualization system," in *International Symposium on Methodologies for Intelligent Systems*. Springer, 2011, pp. 523–532.
- [8] M. Fukino, Y. Hirata, and K. Aihara, "Coarse-graining time series data: Recurrence plot of recurrence plots and its application for music," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 26, no. 2, p. 023116, 2016.
- [9] "Gamaka Box- the first ever Comprehensive Notation System for Indian Music," <http://www.rameshvinayakam.com/research>, accessed: 2016-11-13.
- [10] "Music in Motion, The Automated Transcription of Indian Music, National Centre for Performing Arts (NCPA)," accessed: 2019-05-18. [Online]. Available: <https://autrimncpa.wordpress.com/>
- [11] S. Gulati, J. Serrà, K. K. Ganguli, S. Sentürk, and X. Serra, "Time-delayed melody surfaces for rāga recognition," in *ISMIR*, 2016, pp. 751–757.
- [12] J. Salamon and E. Gómez, "Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, pp. 1759–1770, 2012. [Online]. Available: <http://mtg.upf.edu/system/files/publications/SalamonGomezMelodyTASLP2012.pdf>
- [13] S. Gulati, A. Bellur, J. Salamon, V. Ishwar, H. A. Murthy, and X. Serra, "Automatic tonic identification in indian art music: approaches and evaluation," *Journal of New Music Research*, vol. 43, no. 1, pp. 53–71, 2014.
- [14] V. S. Viraraghavan, H. Murthy, and R. Aravind, "Precision of sung notes in carnatic music," in *Proceedings of the 19th International Society for Music Information Retrieval Conference*, 2018.
- [15] D. Halliday, J. Walker, and R. Resnick, *Fundamentals of physics*. John Wiley & Sons, 2013.
- [16] V. S. Viraraghavan, A. Pal, H. Murthy, and R. Aravind, "A component-based approach to study the effect of indian music on emotions," in *Proc. Workshop on Speech, Music and Mind 2018*, 2018, pp. 31–35.
- [17] "The Python pygame tool," accessed: 2019-05-18. [Online]. Available: <https://www.pygame.org/>
- [18] G. U. Navalyal and R. D. Gavas, "A dynamic attention assessment and enhancement tool using computer graphics," *Human-centric Computing and Information Sciences*, vol. 4, no. 1, p. 11, 2014.
- [19] "'Samidayajuda'- Kedaragaula varnam in different speeds," accessed: 2017-06-14. [Online]. Available: <https://www.youtube.com/watch?v=EKNSaCTbMIU>
- [20] "Audacity 2.1.2 [Audio editor] (2016)," Latest version freely available from: <http://www.audacityteam.org/download/>, accessed: 2017-04-16.
- [21] V. S. Viraraghavan, A. Pal, R. Aravind, and H. Murthy, "Non-uniform time-scaling of carnatic music transients," *arXiv preprint arXiv:1711.02318*, 2017.
- [22] B. Battey, "Bezier spline modeling of pitch-continuous melodic expression and ornamentation," *Computer Music Journal*, vol. 28, no. 4, pp. 25–39, 2004.
- [23] H. Ranjani, D. Paramashivan, and T. V. Sreenivas, "Quantized melodic contours in indian art music perception: Application to transcription," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, 2017, pp. 174–180.
- [24] S. K. Subramanian, L. Wyse, and K. McGee, "A two-component representation for modeling gamakas of carnatic music," in *Proc. of 2nd CompMusic Workshop*, 2012.