A TWO-COMPONENT REPRESENTATION FOR MODELING GAMAKAS OF CARNATIC MUSIC

Srikumar K. Subramanian
National University of Singapore
srikumar.k.subramanian@nus.edu.sg

Lonce Wyse
National University of Singapore
lonce.wyse@nus.edu.sg

Kevin McGee
National University of Singapore
mckevin@nus.edu.sg

ABSTRACT

Continuous pitch movements called "gamakas" are key to Carnatic music, but are not included in its traditional notation system. Modeling the melodic and shape aspects of gamakas and how they are selected for phrases given as notation are interesting interrelated problems that have analogs in text to speech and expressive singing synthesis systems. Descriptive textual and graphical notation techniques have also been proposed as tools for studying gamakas. However, a concatenative model of phrase-level gamaka selection leads to a combinatorial explosion of possibilities to be selected from. We propose a "longest path" optimization algorithm for phrase-level selection of gamakas that solves this combinatorial explosion using a scoring function that expresses local continuity preferences; and a two-component "stage-dance" representation of gamakas which simplifies the scoring function by separating local melodic and shape continuity preferences. We analyzed a performance of a composition in this framework and found the approach to be able to imitate the phrase-level preferences exhibited by the performer.

1. INTRODUCTION

Continuous pitch movements feature in the expressive performance techniques of several musical genres. Known techniques in western classical music include vibrato, glissandi and coloratura. In the classical Indian music traditions known as "Hindustani music" in northern India and "Carnatic music" in southern India, such movements, called gamakas, are not only used for ornamental and expressive purposes, but also have grammatical significance. Modeling the selection, transformation and combination of gamakas is an interesting and challenging problem.

In current Carnatic music practice and pedagogy, the logic of gamakas and their relationship to the melodic system of rāgas are conveyed through a large body of compositions notated using a sparse "prescriptive notation" [1]. This notation expresses melody as a series of "notes" and does not make explicit the gamakas to be used in their performance. Since a literal interpretation of the prescription would make a melody comprising only discrete tones, such an interpretation does not constitute an acceptable performance. Trained musicians, however, are capable of interpreting the prescriptive notation in accordance with the rules of the rāga that a given composition is set in. The knowledge to construct gamakas appropriate for each melodic context remains implicit in the genre.

A trained musician is not only capable of providing various interpretations of the notes of a given melody, but also chooses them such that phrases are rendered as a connected whole. In a lyrical passage, a syllable might be associated with a sequence of such "notes" that are expected to be rendered with some degree of continuity between them. Such a notion of continuity can extend up to and even cross phrasal boundaries and is an important criterion for selecting appropriate gamakas to interpret prescriptive notation. The work presented here is an attempt at modeling gamakas by attending to this notion of continuity.

We now discuss previous work on modeling continuous pitch movements in related areas.

2. RELATED WORK

Though detailed notation of gamakas is not common practice, a tradition of using mnemonic symbols to annotate the solfa in prescriptive notation does exist and goes toward constructing an ontology for gamakas [2–5]. In particular, the categories of gamakas and terminology documented by Dikshitar in [2], such as *kampita*, *odukkal*, *orikkai* and *kaṇḍippu*, serve as a reference for later work [6,7]. Viswanathan's "descriptive notation" [8] represents gamakas using traditional solfa names for all the intermediate pitch classes involved in them, together with explicit time structure. Some ethnomusicologists have also used standard western staff notation or variants to represent melodic detail comparable to descriptive notation [9, 10]. While the symbolic approaches suggest a view of gamakas as movements relative to a reference pitch, the latter approaches tend to take an absolute view of movements by treating them in their entirety.

The linear syntactic appearance of traditional notation in Carnatic music suggests modeling gamaka combination as textual concatenation. The Gaayaka system for synthesizing Carnatic music features a textual notation for gamakas and a lookup-table based mechanism for determining gamakas given prescriptive notation. In analogous situations in Jazz, grammar based systems have been used to expand chord progressions into melodies or to create melodic accompaniment for a song [11–16]. Grammar engines for solving such problems in musical domains have also been developed [17].

Speech intonation models in text to speech synthesizers model continuous pitch movements - “F0 contours” - which are analogous to gamakas. The Fujisaki model treats the F0 contour as the response of a dynamical system to a series of discrete commands and has been applied to both speech and singing synthesis [18]. Non-dynamical models such as the tilt-intonation model [19] and maximum-based descriptions [20] have also been proposed that parameterize frequency and movement shape in different ways. Paralleling these parameterizations, constrained two-dimensional Bezier splines have been proposed for Hindustani vocal music [21]. An interesting aspect of most modern speech synthesizers is that they represent the overall F0 contour of a spoken sentence or phrase using an “F0 mean” that is associated with speech syllables and an “F0 shape” that governs continuity of pitch movement between syllables that is combined with the mean to produce the resultant intonation. A system originally developed for speech synthesis has also been adapted for synthesizing expressive singing techniques such as vibrato and coloratura [22].

3. RESEARCH PROBLEM

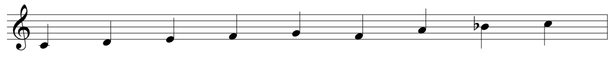
The known ontologies of gamakas treat melodic aspects of gamakas independent of their shape and dynamics. Their treatment of gamakas suggests that shape aspects might generalize over rāgas. However, representations based on these ontological categories of gamakas are ambiguous as specifications of realizable gamakas. The linear representations discussed in the previous section can represent gamakas of all kinds with enough precision for resynthesis, but the rules for parsing such linear representations into melodic and shape aspects are not known. There is therefore a need for a precise representation that separates these two aspects of gamakas.

The melodic and shape characteristics of a gamaka also influence its appropriateness for expressing a note in the context of a whole melodic phrase. The choice of specific gamakas depends on their continuity with those that follow and precede them. This aesthetic is expressed by musicians in the choice they make for each note of a phrase from the set of possibilities otherwise permitted in the same note-local melodic context in other phrases. Melodic and shape aspects of gamakas contribute in different ways to this notion of continuity.

From a modeling perspective, it is conceivable that the melodic and shape aspects of a performed gamaka are determined through different processes that draw on knowledge from different areas of musicology. For example, in speech prosody models, the F0 mean is determined as a per-syllable pitch level by using a semantic model of the text to be spoken whereas the F0 shape is encoded using curve parameters as in the tilt intonation model [19] or as the response of a dynamical system to discrete input as in the Fujisaki model [18], which claims to have a physiological basis.

For the above reasons, we propose a precise representation of gamakas as two components corresponding to the melodic and shape aspects of gamakas, and phrase-optimal selection of gamakas using local continuity preferences.

Ascent



C	*D*	*E*	*F*	*G*	*F*	*A*	*B*♭	*C*$^+$
sa	ri$_2$	ga$_3$	ma$_1$	pa	ma$_1$	da$_2$	ni$_2$	ṡa
sa	*ri*	*gu*	*ma*	*pa*	*ma*	*di*	*ni*	*ṡa*
S	R	G	m	P	m	D	n	Ṡ

Descent

C$^+$	*B*♭	*A*	*G*	*F*	*E*	*F*	*D*	*E*	*D*	*C*
ṡa	ni$_2$	da$_2$	pa	ma$_1$	ga$_3$	ma$_1$	ri$_2$	ga$_3$	ri$_2$	sa
ṡa	*ni*	*di*	*pa*	*ma*	*gu*	*ma*	*ri*	*gu*	*ri*	*sa*
Ṡ	n	D	P	m	G	m	R	G	R	S

Table 1. Ascent and descent pitch patterns for the rāga “Sahānā”.

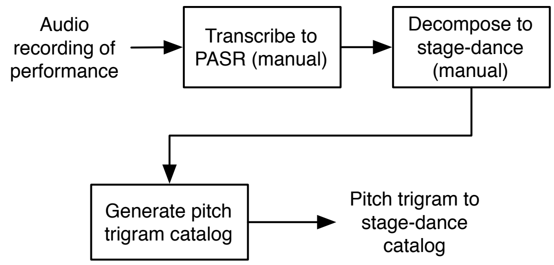


Figure 1. Analyzing a performance to construct the stage-dance representation.

4. METHOD

We analyzed a linear PASR transcription of a *vīṇā* performance of the varṇam “Karuṇimpa” in the rāga Sahānā (table 1) by Smt. Rajēswari Padmanābhan, by separating the gamakas into two components. We then constructed a reduced two-component representation we call *stage* and *dance* by simplifying gamaka amplitudes and categorizing movement shapes based on timing (Figure 1). We used the representation in a system for interpreting phrases given in the rāga of the composition. We discarded the phrase-level discriminating choices made by the performer in the input to the system and introduced scoring functions that express local continuity preferences instead. The system uses a phrase-level “longest path” optimization algorithm for directed acyclic graphs to reconstruct the phrases of the original composition (Figure 2). We compared the system’s output to the gamaka choices made by the performer.

We begin with an overview of the PASR form we use to represent gamakas.

5. THE PASR FORM OF A GAMAKA

The PASR form, short for “Pitch Attack Sustain Release form”, treats a gamaka as a sequence of quasi-stationary pitches which we refer to as “focal pitches”, each of which is described using four numbers (p, a, s, r) – its *pitch number*, *attack*, *sustain* and *release* [23]. The *pitch number*

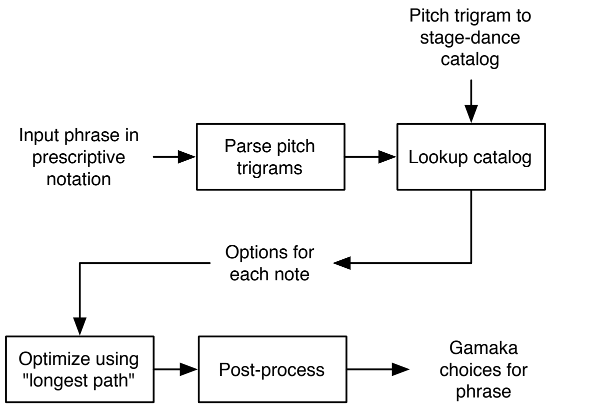


Figure 2. Automatic selection of gamakas for a given phrase.

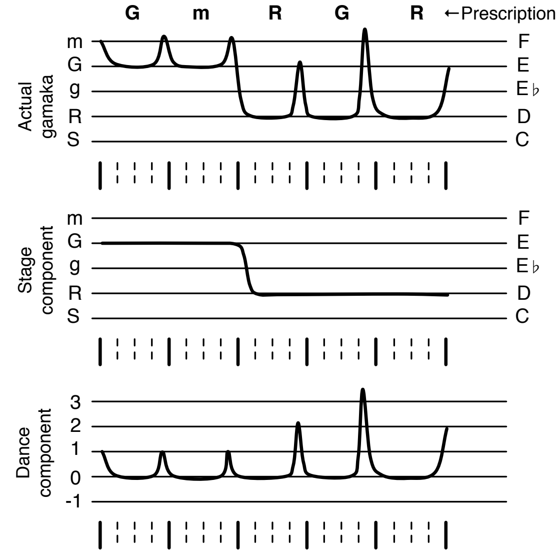


Figure 3. Illustration of decomposing a gamaka into "stage" and "dance" components.

provides the quasi-stationary value of the focal pitch and is expressed in semitones relative to the tonic of the piece being transcribed. The *attack* is the time spent approaching the stationary pitch from the preceding one, the *sustain* is the time spent at the focal pitch without movement and the *release* is the time spent moving away from the focal pitch. towards the next one.

The PASR form simplifies gamaka representation by including only the timing of the pitch movements and avoiding specific pitch interpolation schemes. Asymmetry in a movement can be encoded using appropriate attack and release values. We use sinusoidal interpolation to realize a gamaka given in PASR form as a pitch control signal for a synthesizer.

6. ANALYSIS

We manually analyzed the linear PASR transcription of the chosen performance in terms of two superposed pitch movements as shown in Figure 1, refactored the gamaka amplitudes and categorized the residual shape component to create a reduced representation of gamakas whose components we call *stage* and *dance*. These components respectively capture the movement of what might be called the "melodic centres" of a gamaka, and the shape of oscillatory movements that occur around these centres.

6.1 Separating gamaka components

Figure 3 shows a gamaka sequence being decomposed into stage and dance components. We used the control surface of the *vīṇā* as a guide to determine the two components that constitute a gamaka. On the *vīṇā*, a pitch movement can be performed either by sliding over the frets or by pulling on the string. A first approximation to a decomposition is to treat fret slides as the *stage* component and string pulls as the *dance* component. This mapping is not strict and would differ between performers based on style and technical facility. Therefore some degree of musical judgement is involved in such a decomposition.

6.2 Refactoring gamaka amplitudes

The notes of a rāga may be rendered as oscillations between two other pitch values. Therefore it is not surprising that the oscillations in the *dance* component occur associated with the sustain portions of the *stage* focal pitches in our transcription. To separate the rāga-specific end points of such movements from its shape, we factored out the amplitudes of the dance component oscillations and associated the amplitudes with the stage focal pitches. With this transformation, we saw a large reduction in the complexity of capturing the pitch movements constituting a gamaka, with 47% of the stage focal pitches having unique associated dance amplitudes and only 4.1% having two amplitudes, with the rest being stationary. With the amplitude values thus factored out of the dance component, only three values for the amplitudes remain – $+1, 0$ and -1 – which we denote using $^\wedge$, $-$ and $_\vee$.

6.3 Categorizing focal pitch shapes

A histogram of a shape characteristic for the focal pitches of the *dance* components showed the distribution to be trimodal. We used this to map dance component focal pitches to three categories – a) transient, b) stable and c) sustained – corresponding to how much time is spent sustained at the focal pitch value. The semantic distinction between transient focal pitches and the others is that transients can be inserted or dropped when sequencing gamakas.

We now describe reconstructing gamakas for phrases by selecting from the possible options for the notes constituting a phrase.

7. PHRASE INTERPRETATION

To create a phrase interpreter based on the performance transcription, we created a catalog of *stage* and *dance* components. The composition was divided into "notes" as specified in the composition's prescriptive notation, a pitch-class

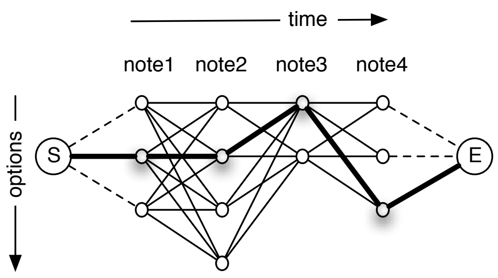


Figure 4. Illustration of finding the best choice of gamakas over a phrase as the "longest path" in a directed acyclic graph. The directions on the edges are aligned with the direction of time.

$\text{kampita}(-,-,n)$	$\text{kampita}(-,\wedge,n)$
$\text{kampita}(\wedge,-,n)$	$\text{kampita}(\wedge,\wedge,n)$
$\vee -$	$\vee \wedge$
$- \vee -$	$- \vee \wedge$
$\wedge - \vee -$ (i.e. *ravai*)	$\vee \wedge - \wedge$

Table 2. Simplified dance movement catalog. $\text{kampita}(s, e, n)$ denotes sequences of the form $[\wedge, -\wedge, -, ...]$ or $[-, \wedge, -, \wedge, ..]$ The word *kampita* used is suggestive of the traditional term, but generalizes to include *odukkal* and *orikkai* in the $n = 0$ case.

trigram context was computed for each note, with which a set of decomposed gamakas were associated. This trigram context is similar to the note context used in Gaayaka, except that we discard note timing information. Constructing such a catalog discards the discrimination expressed by the performer in choosing gamakas for a phrase, which we attempted to restore by constructing a scoring function for the phrase-level "longest path" optimization algorithm illustrated in Figure 4. The scoring function determines the weights on the graph and the gamaka sequence that maximizes the sum of these weights is chosen.

This architecture can be seen as the fusion of a grammar based approach, and a constraint-satisfaction approach. Vijayakrishnan's proposed formulation of the "grammar of Carnatic music" [24] based on "optimality theory" inspired our approach to gamaka selection as the optimal satisfaction of a set of potentially conflicting constraints.

The scores were expressed in two parts –

1. The *stage* continuity score was determined by extending the candidate gamakas over the joint duration of the two consecutive notes and comparing their pitch values. A speed bias factor that preferred less movement for shorter notes and another that preferred a match between the stage pitch and prescribed note's pitch were also used.

2. The *dance* continuity score between two candidate gamakas was determined by whether *kampita* gamakas occur at the boundary and are compatible. The gamakas were also biased to be around the same speed, if the associated notes were of the same duration.

8. RESULTS

We do not have expert musical evaluation to report at this point. Using only local melodic and shape continuity scores, the algorithm was able to select gamakas for phrases that matched the performer's choice 80% of the time (average between three sets), with at most 2 focal pitch values differing between the selected gamakas and the performed gamakas in these cases, allowing for the two end points of a phrase to differ due to inter-phrase continuity which our algorithm does not account for.

9. DISCUSSION

This study draws on a performance of one composition and we cannot claim generality between different composition types, or between performers or other rāgas. However, the simple form of the continuity rules permits us to look at the traits of Sahānā embedded in catalogs of stage and dance components extracted from the transcription. The stage catalog is particularly interesting as it summarizes key rāga traits such as the pitch classes and movement spans using which certain notes are to be performed. For instance, the ma of Sahānā is either performed with ga as the base or is rendered plain, and ni is almost always performed as a movement between da and sa. This is common knowledge and is available in "rāga lakshaṇā" literature. However, the stage catalog also includes the expressed stylistic preferences of the performer which we can now systematically study in terms of the difference with known musicology.

The categories into which gamakas fall as described by the dance catalog are less obvious than in the case of the stage catalog. Some of the well known gamaka types can indeed be mapped to dance components — *nokku* is the sequence $[(-, \text{sus}), (\vee, \text{tran}), (-, \text{tran})]$ and *orikkai* is the sequence $[(-, \text{sus}), (\wedge, \text{tran})]$. *Odukkal*, however, has the same dance component as *orikkai* but forces a slide in the stage component. This raises the interesting question of whether this research can identify new gamaka categories. Some of the cataloged components indeed do not appear to fit into any of the known categories, unless one wishes to label all of them as *kampita* or choose to treat them as combinations of other gamaka fragments. Compared to the literature, the dance component leads to clearer categorization of gamaka types especially with respect to timing information.

The "longest path" optimization makes expressing some kinds of relationships between rules simple using additive score combination for independent contributing preferences and multiplicative score combination for conjunctive preferences. However, as the number of such scoring functions grows, the interdependencies between them grows quickly. This places limitations on such manual study of performances. With the availability of transcriptions at the level of detail of PASR forms, perhaps through future automatic transcription software, it might be possible to systematically explore alternative theories and parameter spaces for gamakas.

One of the key difficulties with our approach so far is that it is unclear how the initial decomposition of the linear

PASR form into the two components (section 6.1) can be done automatically. Though pitch tracking algorithms are far from perfect for the task of creating the linear PASR in the first place, it is possible for someone with no knowledge of the genre to perform a transcription using human-in-the-loop analysis by synthesis. The decomposition into the stage and dance components, however, requires familiarity with the genre, since there is considerable ambiguity in the decomposition. We tried to reduce this ambiguity using the *vīṇā* and the prescriptive notation as reference.

Overall, we consider the optimal satisfaction of local constraints on gamakas to be a useful framework to explore the rules of Carnatic music through synthesis. The *stage-dance* representation makes it possible to express scoring functions that focus on rāga-specific local rules to be considered separately from dynamic criteria.

10. CONCLUSION

We presented a two-component gamaka representation that separates pitch aspects of a gamaka from the shape of the movement. We analyzed a varṇam in the rāga Sahānā and found that focal pitches for the *dance* component fall into three categories which simplifies the representation. We also found that the amplitudes of gamakas are determined by the pitch of the *stage* component given local melodic context. Finally, we showed how gamaka selection for interpreting prescriptive notation can be expressed as a phrase-level "longest path" optimizer operating on a catalog of gamakas keyed using a duration-free local melodic context. This representation can be used to automatically construct valid renditions of performances from sparse notation, and can also provide insight when compared to traditional literature on gamaka categories and choices made in performance.

Acknowledgments

Many thanks to Martin Henz for guidance leading towards the phrase level optimization algorithm and to Shivkumar Kalyanaraman for making digital copies of notations available online.

11. REFERENCES

[1] C. Seeger, "Prescriptive and descriptive Music-Writing," *The Musical Quarterly*, vol. 44, no. 2, pp. 184–195, Apr. 1958.

[2] S. Dikshitar, *Sangita Sampradaya Pradarsini*, Ettayapuram Samasthanam, 1904.

[3] V. Shankar, *Shyama Sastry's Compositions*, 2nd ed. Madras: C. S. Ayyar, 1979.

[4] R. Iyengar, *Sri Kriti Mani Malai*, 2nd ed. Madras: Sabarmati, 1965.

[5] A. Mallikarjuna Sharma, "Kampitas and gamakas," in *AMS Easy Methods*. AMS Foundation, 2007, pp. 25–31.

[6] S. Gopalam, *Facets of Notation in South Indian Music*. Delhi: Sundeep Prakashan, 1991.

[7] R. S. Jayalakshmi, "Gamakas explained in sangita-sampradaya-pradarsini of subbarama diksitar," Ph.D. dissertation, Madras University, Dec. 2002.

[8] T. Viswanathan, "The analysis of rāga ālāpanā in south indian music," *Asian Music*, vol. 9, no. 1, pp. 13–71, Jan. 1977.

[9] E. t. Nijenhuis, *Varnam: Selected Concert Studies for the South Indian Lute*. Naarden, The Netherlands: Sarasvati Bhavan, 2001.

[10] R. Widdess, "Involving the performers in transcription and analysis: A collaborative approach to dhrupad," *Ethnomusicology*, vol. 38, no. 1, pp. 59–79, 1994.

[11] J. W. Ulrich, "The analysis and synthesis of jazz by computer," in *In Fifth International Joint Conference on Artificial Intelligence*, 1977.

[12] B. Pennycook, D. R. Stammen, and D. Reynolds, "Toward a computer model of a jazz improvisor," in *Proceedings of the International Computer Music Conference*, 1993, p. 228–228.

[13] G. Ramalho and J. G. Ganascia, "Simulating creativity in jazz performance," in *Proceedings of the National Conference on Artificial Intelligence*, 1994, p. 108–108.

[14] J. Gillick, K. Tang, and R. M. Keller, "Machine learning of jazz grammars," *Computer Music Journal*, vol. 34, no. 3, p. 56–66, 2010.

[15] J. Biles, "GenJam: a genetic algorithm for generating jazz solos," in *Proceedings of the International Computer Music Conference*, 1994, p. 131–131.

[16] R. Keller and D. Morrison, "A grammatical approach to automatic improvisation," in *Proceedings, Fourth Sound and Music Conference, Lefkada, Greece*, 2007.

[17] B. Bel and J. Kippen, "Bol processor grammars," *Understanding Music with AI—Perspectives on Music Cognition*, p. 366–401, 1992.

[18] H. Fujisaki and K. Hirose, "Modelling the dynamic characteristics of voice fundamental frequency with application to analysis and synthesis of intonation," in *Proceedings of 13th International Congress of Linguists*, 1982, p. 57–70.

[19] P. Taylor, "The tilt intonation model," in *Fifth International Conference on Spoken Language Processing*, 1998.

[20] T. Portele and B. Heuft, "The maximum-based description of f0 contours and its application to english," in *Fifth International Conference on Spoken Language Processing*, 1998.

[21] B. Battey, "Bézier spline modeling of Pitch-Continuous melodic expression and ornamentation," *Computer Music Journal*, vol. 28, no. 4, pp. 25–39, Dec. 2004.

[22] G. Berndtsson, "The KTH rule system for singing synthesis," *Computer Music Journal*, vol. 20, no. 1, p. 76–91, 1996.

[23] S. Subramanian, L. Wyse, and K. McGee, "Modeling speed doubling in carnatic music," in *Proceedings of the International Computer Music Conference 2011*, University of Huddersfield, UK, Aug. 2011, pp. 478–485.

[24] K. G. Vijayakrishnan, *The Grammar of Carnatic Music*, Har/MP3 ed. Mouton De Gruyter, Dec. 2007.