Discovering variable length phrases from symbolic notation of Carnatic music

Ranjani, H. G. and Prof. T. V. Sreenivas
ranjanihg@ece.iisc.ernet.in, tvsree@ece.iisc.ernet.in

Problem

- Given symbolic transcript of a *rāga*, discover repetitive phrases

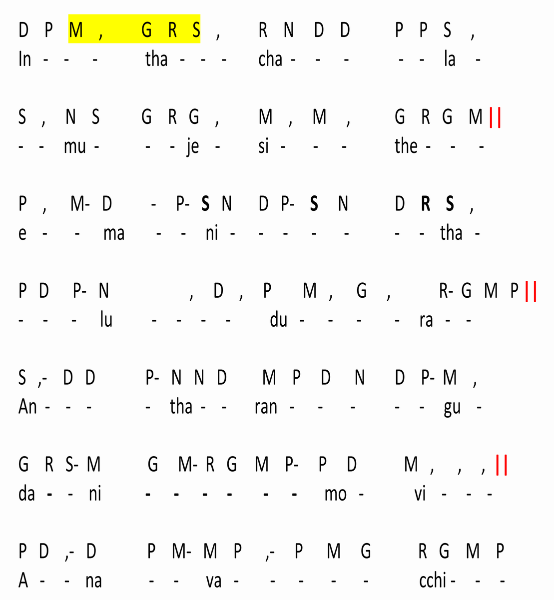

Figure 1: Sample symbol transcript of *Begada rāga*.

- Let transcript be denoted by $\underline{A} = [A_1, A_2, \ldots A_I]$.
- Any rhythm cycle, $A_i \triangleq [u_{t=1} u_{t=2} \ldots u_{t=T_{A_i}}]$, where swaras, $u_t \in V$, with $V = \{S, R, G, M, P, D, N, S\}$.

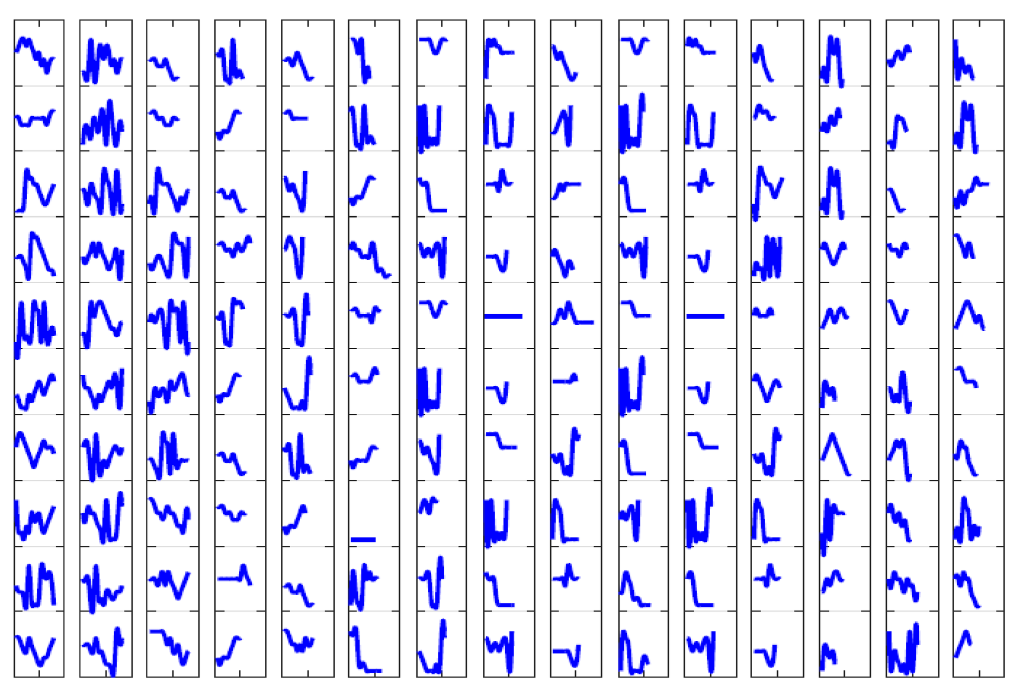

Figure 2: Rough pitch contours of more than 100 rhythm cycles from symbolic transcripts o*Begada rāga*.

- Multiple and unknown phrases
- Variable length phrases

Assumptions

- Rhythm cycle contains note sequences : concatenation of independent phrases
- Phrases are well within rhythm cycle
- Phrases are repeated across rhythm cycles/compositions

Experimental details

- Publicly available online database [http://www.shivkumar.org/music/] (notations by Dr. Shivakumar Kalyanaraman)
- Experiments on 12 *rāgas*: *Hari-Kambhoji, Bhairavi, Shankarābharana, Thōdi, Nāttai, Panthuvarāli, Madhyamāvathi, Khamas, Begada, Kalyani, Reethigowla* and *Sahana*
- Octave folded
- Each note of unit duration
- Training: $>$ 2000 note sequences; Testing: $>$ 1500 per *rāga*
- Performance measures: perplexity, semantic relevance

Conclusions

- Use of 7 notes as generally available in transcription
- Discovering grammatical structure of music
- Obtain phrases containing varied length sub-sequences
- Multigram perplexity lower than N-gram on training and test data
- Modified multigram for longer length sequences
- Appreciable number of musicological phrases captured

Formulation

- Any rhythm cycle $A = [u_1, u_2, u_3, \ldots, u_{T_A}]$ s.t. $p(A) = \prod_{k=1}^{Q_A} p(s_k) \triangleq \prod_{k=1}^{Q_A} \theta_k$
  where s_k is such that $|s_k| \leq N$ and for any $T_A > N$, $Q_A > 1$. and $s_1 = [u_{b_0}, \ldots u_{b_1}]$, $s_2 = [u_{b_1+1}, \ldots u_{b_2}]$ and $s_{Q_A} = [u_{b_{Q_{A-1}}+1}, \ldots u_{b_{Q_A}}]$, $b_0 = 1$ and $b_{Q_A} = T_A$.
- A typical segmentation on A gives : $A \equiv [s_1, s_2, s_3, \ldots s_{Q_A}]$
- $Z = \{b_k\}, k = 1 : Q_A$
- Estimate parameters, θ_k to maximize posterior $p(\underline{Z}|\underline{A};\theta)$: $\theta^* = \arg\max_\theta \left\{ \max_{\underline{Z}} \left[\log p(\underline{Z}|\underline{A};\theta^{old}) \right] \right\}$
- Constraint : $\sum_{k=1}^{Y} \theta_k = 1$ where, Y is total number of unique phrases
- Algorithm :
  - 1. Find Z^*, $$Z^* = \arg\max_{Z \in \mathcal{Z}} \log p(\underline{A}, \underline{Z}; \theta^{old}) = \arg\max_{\underline{Z} \in \mathcal{Z}} \log p(\underline{A}|\underline{Z}, \theta^{old}) p(Z; \theta^{old})$$
  - 2. Update parameters $$\theta_j^{new} = \frac{c_j^{Z^*}}{c^{Z^*}}$$

Results

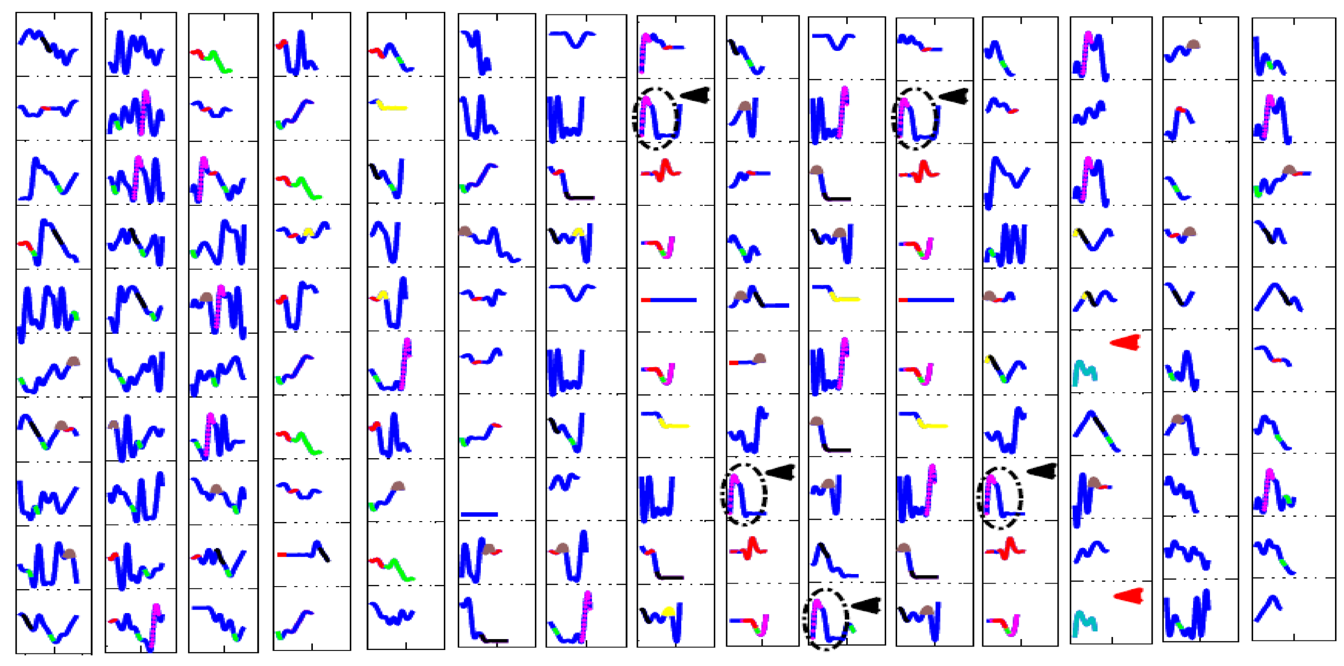

Figure 3: Rough pitch contours of more than 100 *āvarthanas* from training data of *rāga Begada* (in blue) and top ten frequently occurring phrases (sorted aided by other colors) as discovered by 8-multigram. Two characteristic phrase(s) are highlighted using (black and red) arrowheads.

- N determines maximum length of sub=sequence
- Propose a modified 2-stage approach:
  - Obtain $\{s_k\}_{k=1}^{Y}$ containing $\leq N$ length phrases, using multigram training
  - Create new vocab: $V' = \left\{ V \cup \{s_i : |s_i| = N,\ \theta_i > P_{thr}\},\ \forall i \in \{s_i\}_{i=1}^{Y} \right\}$.
  - Replace any occurrence of s_i in data with its corresponding entry from V'
  - Obtain $\{s'_j\}_{j=1}^{Y'}$ containing $N + N'$ length phrases through a second stage of multigram training

N-gram								
	Training				Testing			
Rāga	$N=5$	$N=6$	$N=7$	$N=8$	$N=5$	$N=6$	$N=7$	$N=8$
Bh	2.80	**2.79**	2.81	2.93	**17.55**	33.5	61.45	90.25
Nt	3.07	3.08	2.83	**2.81**	**8.4**	26.7	90.2	152.65
Pa	2.97	2.73	**2.72**	2.94	7.62	10.34	9.17	**5.77**
Sb	2.77	2.55	2.50	**2.45**	**11.57**	25.6	51.6	70.55
Th	2.76	2.43	2.25	**2.19**	**8.48**	16.36	27.02	34.35
Hk	2.47	2.29	2.29	**2.22**	**9.7**	29.08	62.61	60.49
Mv	2.93	**2.68**	2.77	3.06	**7.2**	10.69	10.25	7.55
Kh	2.53	2.32	2.19	**2.15**	**6.19**	10.61	12.23	11.46
Bg	2.63	2.50	2.45	**2.35**	**59.64**	236.08	186.75	85.97
Kl	2.97	2.96	3.17	**2.10**	**156.31**	770.04	1667	1946.5
Sh	2.30	2.12	2.02	**1.98**	**30.7**	212.28	1163	2324
Rg	2.67	2.51	**2.44**	2.49	**31.95**	239.13	1136	1777

N-multigram model								
	Training				Testing			
Raga	$N=5$	$N=6$	$N=7$	$N=8$	$N=5$	$N=6$	$N=7$	$N=8$
Bh	1.91	1.72	1.56	**1.43**	2.65	2.66	2.67	**2.63**
Nt	1.93	1.73	1.55	**1.43**	**2.18**	2.27	2.29	2.33
Pa	1.98	1.77	1.62	**1.48**	**2.82**	2.92	2.99	3.02
Sb	1.90	1.72	1.52	**1.36**	**2.50**	2.52	2.61	2.55
Th	1.92	1.77	1.54	**1.41**	2.50	**2.44**	2.53	2.55
Hk	1.82	1.59	1.40	**1.31**	2.47	**2.47**	2.56	2.62
Mv	1.86	1.65	1.48	**1.33**	**2.16**	2.18	2.25	2.24
Kh	1.83	1.67	1.44	**1.33**	**2.50**	2.59	2.62	2.76
Bg	1.84	1.57	1.50	**1.30**	**2.68**	2.86	2.99	3.01
Kl	1.87	1.65	1.51	**1.41**	**2.81**	2.96	3.17	3.24
Sh	1.74	1.59	1.42	**1.33**	2.50	**2.45**	2.58	2.64
Rg	1.83	1.70	1.49	**1.41**	**2.46**	2.54	2.73	2.78

Modified N'-multigram model								
	Training				Testing			
Rāga	$N=5$ $N'=5$	$N=6$ $N'=6$	$N=7$ $N'=7$	$N=8$ $N'=8$	$N=5$ $N'=5$	$N=6$ $N'=6$	$N=7$ $N'=7$	$N=8$ $N'=8$
Bh	1.62	1.55	**1.53**	1.63	2.86	2.75	2.69	**2.65**
Nt	**1.55**	1.62	1.62	1.64	2.64	2.36	**2.35**	2.36
Pa	1.76	1.64	**1.59**	1.61	**2.93**	2.97	2.99	3.02
Sb	1.50	1.41	1.43	**1.34**	2.86	2.76	2.63	**2.59**
Th	1.39	1.30	1.31	**1.29**	2.85	2.61	2.58	**2.55**
Hk	1.32	1.30	1.34	**1.26**	2.72	**2.52**	2.59	2.62
Mv	1.69	1.69	1.66	**1.57**	2.37	**2.19**	2.25	2.27
Kh	1.49	1.38	1.41	**1.33**	2.82	2.70	**2.63**	2.77
Bg	1.52	1.56	1.56	**1.48**	**2.75**	2.86	2.99	3.01
Kl	1.70	**1.66**	1.74	1.73	**3.01**	3.05	3.19	3.24
Sh	1.31	1.20	1.22	**1.20**	2.92	2.75	**2.63**	2.67
Rg	1.49	1.47	**1.42**	1.46	2.76	**2.72**	2.75	2.79

Figure 4: Perplexity values of N-gram, N-multigram and modified (N, N')-multigram on training and testing symbolic music data for the *rāgas* considered.

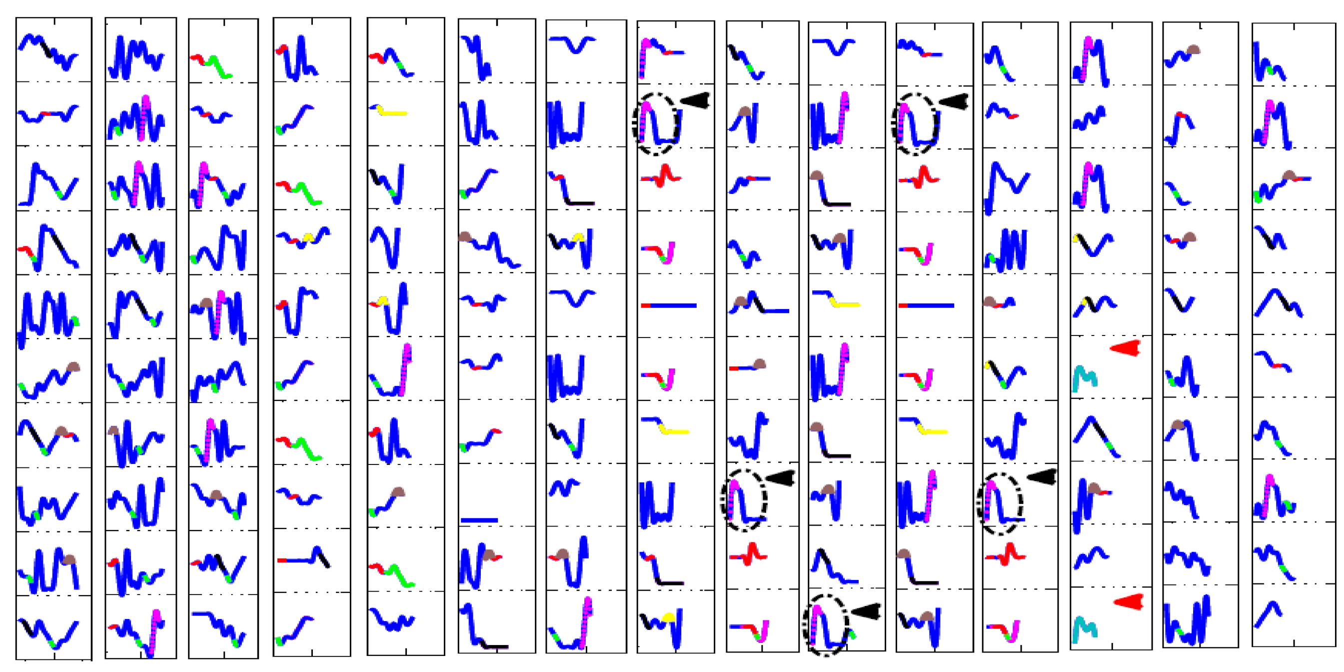

Figure 5: Rough pitch contours of more than 100 *āvarthanas* from training data of *rāga Begada* (in blue) and top ten frequently occurring phrases (sorted aided by other colors) as discovered by modified N'-multigram with $(N, N') = (8, 8)$. Two characteristic phrase(s) are highlighted using (black and red) arrowheads.