Discovering variable length phrases from symbolic notation of Carnatic music

Ranjani H. G.

Advisor: Prof. T. V. Sreenivas,
Dept of ECE,
Indian Institute of Science

April 26, 2016

Problem Formulation

- Given symbolic transcript, discover variable length phrases for a *rāga*

D P M , G R S , R N D D P P S ,
In - - - tha - - - cha - - - - - la -

S , N S G R G , M , M , G R G M ||
- - mu - - - je - si - - - the - - -

P , M- D - P- **S** N D P- **S** N D **R** **S** ,
e - - ma - - ni - - - - - - - - tha -

P D P- N , D , P M , G , R- G M P ||
- - - lu - - - - - du - - - - ra - -

S ,- D D P- N N D M P D N D P- M ,
An - - - - tha - - ran - - - - - - gu -

G R S- M G M- R G M P- P D M , , , ||
da - - ni - - - - - - - - mo - vi - - -

P D ,- D P M- M P ,- P M G R G M P
A - - na - - va - - - - - cchi - - -

Sample symbol transcript of *Begada rāga*

Problem Formulation

- Given symbolic transcript, discover variable length phrases for a *rāga*

D P M , G R S , R N D D P P S ,
In - - - tha - - - cha - - - - - la -

S , N S G R G , M , M , G R G M ||
- - mu - - - je - si - - - the - - -

P , M- D - P- **S** N D P- **S** N D **R S** ,
e - - ma - - ni - - - - - - - - tha -

P D P- N , D , P M , G , R- G M P ||
- - - lu - - - - - du - - - - ra - -

S ,- D D P- N N D M P D N D P- M ,
An - - - - tha - - ran - - - - - - gu -

G R S- M G M- R G M P- P D M , , , ||
da - - ni - - - - - - - - mo - vi - - -

P D ,- D P M- M P ,- P M G R G M P
A - - na - - va - - - - - cchi - - -

Sample symbol transcript of *Begada rāga*

- Multiple phrases exist - unknown
- Variable number of notes

Multigram

- Let transcript be denoted by $\underline{A} = [A_1, A_2, \ldots A_I]$; sequence of rhythm cycles

Multigram

- Let transcript be denoted by $\underline{A} = [A_1, A_2, \ldots A_I]$; sequence of rhythm cycles
- Consider any rhythm cycle $A = [u_1, u_2, u_3, \ldots, u_{T_A}]$; where $u_t \in V$, with $V = \{S, R, G, M, P, D, N\}$.

$$p(A\ ) = \prod_{k=1}^{Q_A} p(s_k) \triangleq \prod_{k=1}^{Q_A} \theta_k \tag{1}$$

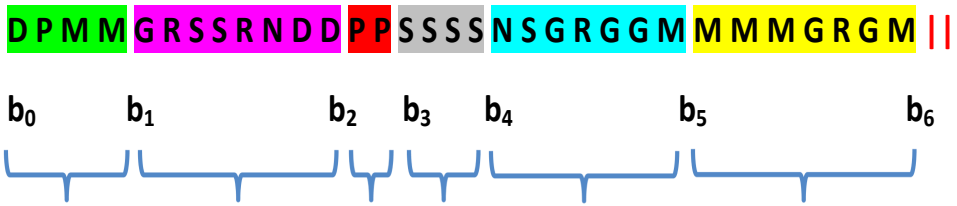
Multigram

- Let transcript be denoted by $\underline{A} = [A_1, A_2, \ldots A_I]$; sequence of rhythm cycles
- Consider any rhythm cycle $A = [u_1, u_2, u_3, \ldots, u_{T_A}]$; where $u_t \in V$, with $V = \{S, R, G, M, P, D, N\}$.

$$p(A\ ) = \prod_{k=1}^{Q_A} p(s_k) \triangleq \prod_{k=1}^{Q_A} \theta_k \tag{1}$$

- Segmentation on A results in

$A \equiv [s_1, s_2, s_3\ , \ldots s_{Q_A}]$.

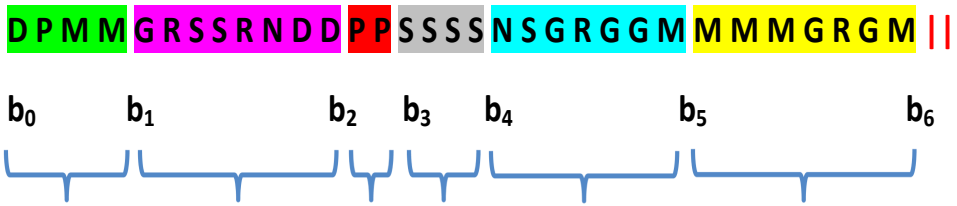


Multigram

- Let transcript be denoted by $\underline{A} = [A_1, A_2, \ldots A_I]$; sequence of rhythm cycles
- Consider any rhythm cycle $A = [u_1, u_2, u_3, \ldots, u_{T_A}]$; where $u_t \in V$, with $V = \{S, R, G, M, P, D, N\}$.

$$p(A\ ) = \prod_{k=1}^{Q_A} p(s_k) \triangleq \prod_{k=1}^{Q_A} \theta_k \tag{1}$$

- Segmentation on A results in



$A \equiv [s_1, s_2, s_3\ , \ldots s_{Q_A}]$.

- Set of boundaries $\{b_k\}$ represented by r.v. Z

Parameter estimation: Segmental K-means

- Estimate parameters, θ_k to maximize posterior $p(Z|A;\theta)$

$$\theta^* = \arg\max_{\theta} \left\{ \max_{\underline{Z}} \left[\log p(\underline{Z}|\underline{A};\theta^{old}) \right] \right\}$$

Parameter estimation: Segmental K-means

- Estimate parameters, θ_k to maximize posterior $p(Z|A;\theta)$

$$\theta^* = \arg\max_{\theta} \left\{ \max_{\underline{Z}} \left[\log p(\underline{Z}|\underline{A};\theta^{old}) \right] \right\}$$

- Constraint : $\sum_{k=1}^{Y} \theta_k = 1$ where, Y is total number of unique sub-sequence entries

Parameter estimation: Segmental K-means

- Estimate parameters, θ_k to maximize posterior $p(Z|A;\theta)$

$$\theta^* = \arg\max_{\theta} \left\{ \max_{\underline{Z}} \left[\log p(\underline{Z}|\underline{A};\theta^{old}) \right] \right\}$$

- Constraint : $\sum_{k=1}^{Y} \theta_k = 1$ where, Y is total number of unique sub-sequence entries
- Algorithm
  - 1. Find Z^* such that

$$\begin{aligned} Z^* &= \arg\max_{Z \in \mathcal{Z}} \log p(A, Z;\theta^{old}) \\ &= \arg\max_{Z \in \mathcal{Z}} \log p(A|Z,\theta^{old}) p(Z;\theta^{old}) \end{aligned} \tag{2}$$

Parameter estimation: Segmental K-means

- Estimate parameters, θ_k to maximize posterior $p(Z|A;\theta)$

$$\theta^* = \arg\max_{\theta} \left\{ \max_{\underline{Z}} \left[\log p(\underline{Z}|\underline{A};\theta^{old}) \right] \right\}$$

- Constraint : $\sum_{k=1}^{Y} \theta_k = 1$ where, Y is total number of unique sub-sequence entries
- Algorithm
  - 1. Find Z^* such that

$$\begin{aligned} Z^* &= \arg\max_{Z \in \mathcal{Z}} \log p(A, Z; \theta^{old}) \\ &= \arg\max_{Z \in \mathcal{Z}} \log p(A|Z, \theta^{old}) p(Z; \theta^{old}) \end{aligned} \tag{2}$$

  - 2. Update parameters

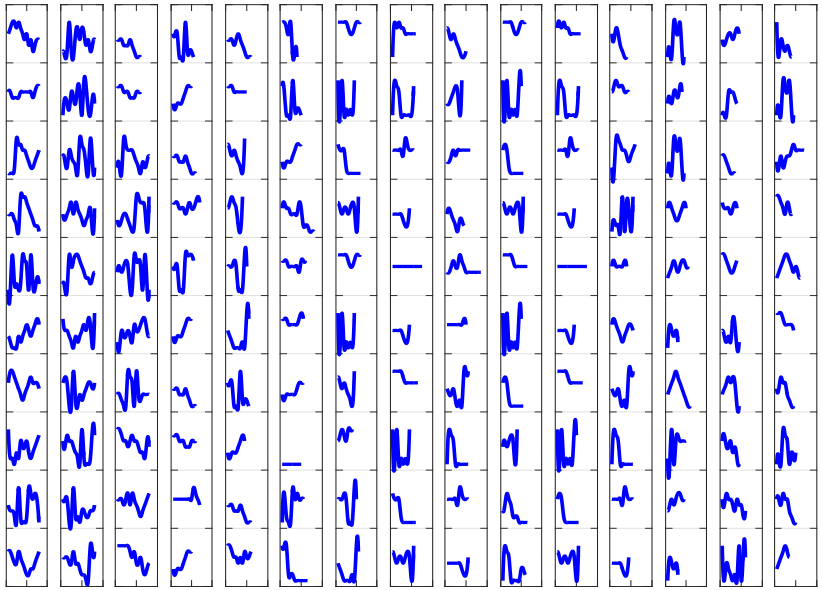
$$\theta_j^{new} = \frac{c_j^{Z^*}}{c^{Z^*}} \tag{3}$$

where Z^* maximizes posterior

Multigram attributes

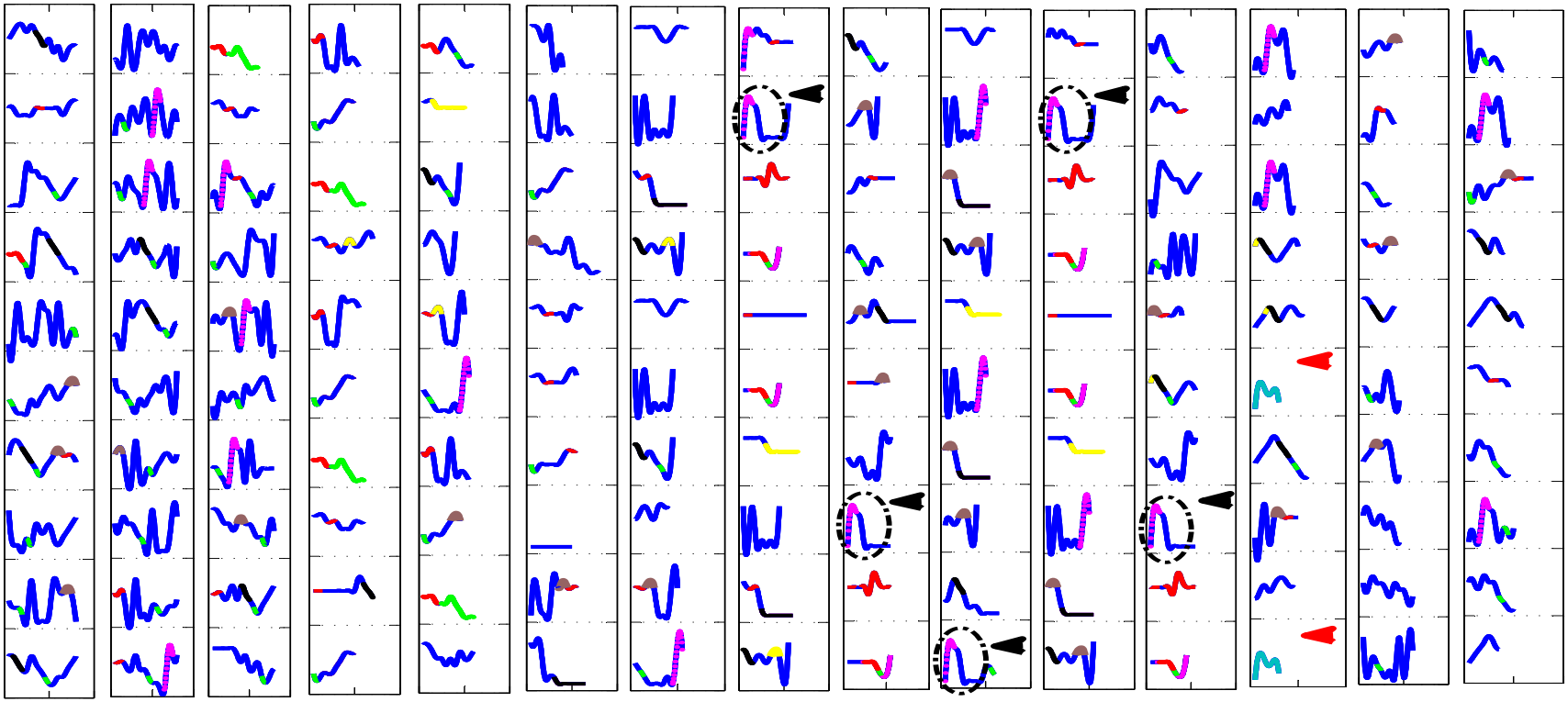
- Convergence criteria : Boundaries do not change
- $\{\theta\}$ - Variable length multinomial distribution
- Normalized count over number of segments
- Phrase entries themselves can change across iterations
- Total number of phrases can change across iterations

Analysis



Rough pitch contours of more than 100 rhythm cycles from symbolic transcripts of *rāga Begada* (in blue)

Analysis



Rough pitch contours of more than 100 rhythm cycles from symbolic transcripts of *rāga Begada* (in blue) and top ten frequently occurring phrases (sorted aided by other colors) as discovered by 8-multigram. Two musicological phrase(s) are highlighted using (black and red) arrowheads.

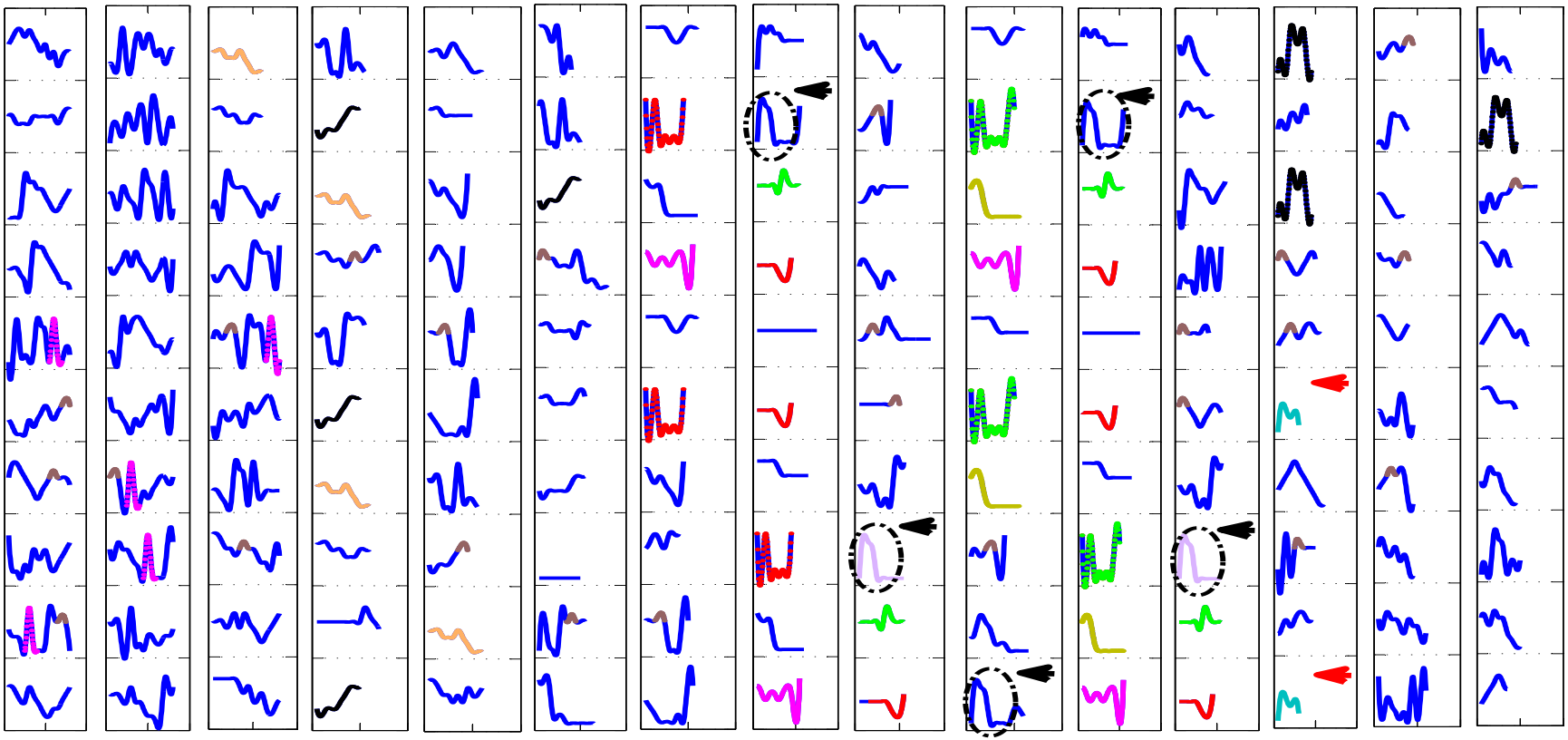
Modified multigram

- Sub-sequences limited by N

Modified multigram

- Sub-sequences limited by N
- Propose a modified 2-stage approach:
  - Obtain phrase set ($\leq N$ length phrases), using multi-gram model
  - Create new vocab:
    $V' = \left\{ V \cup \{ s_i : |s_i| = N,\ \theta_i > P_{thr} \},\ \forall i \in \mathcal{D}^r_{N-multi} \right\}.$
  - Replace any occurrence of s_i with its corresponding entry from V'
  - Obtain new set of phrases of maximum $N + M$ length phrases

Analysis



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

- Conclusions
  - Use only 7 notes (irrespective of pitch position)
  - Discover variable length phrases
  - Possible representative feature for symbolic music
  - Some discovered phrases also correlate with musicological phrases
  - Capture grammatical structure of music

- Conclusions
  - Use only 7 notes (irrespective of pitch position)
  - Discover variable length phrases
  - Possible representative feature for symbolic music
  - Some discovered phrases also correlate with musicological phrases
  - Capture grammatical structure of music

QUESTIONS?

Performance: Perplexity

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

Perplexity values of N-gram, N-multigram on training and testing symbolic music data for the *rāgas* considered.

Performance: Perplexity

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			
Raga	$N=5$ $N'=5$	$N=6$ $N'=6$	$N=7$ $N'=7$	$N=8$ $N'=8$	$N=5$ $N'=5$	$N=6$ $N'=6$	$N=7$ $N'=7$	$N=8$ $N'=8$
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

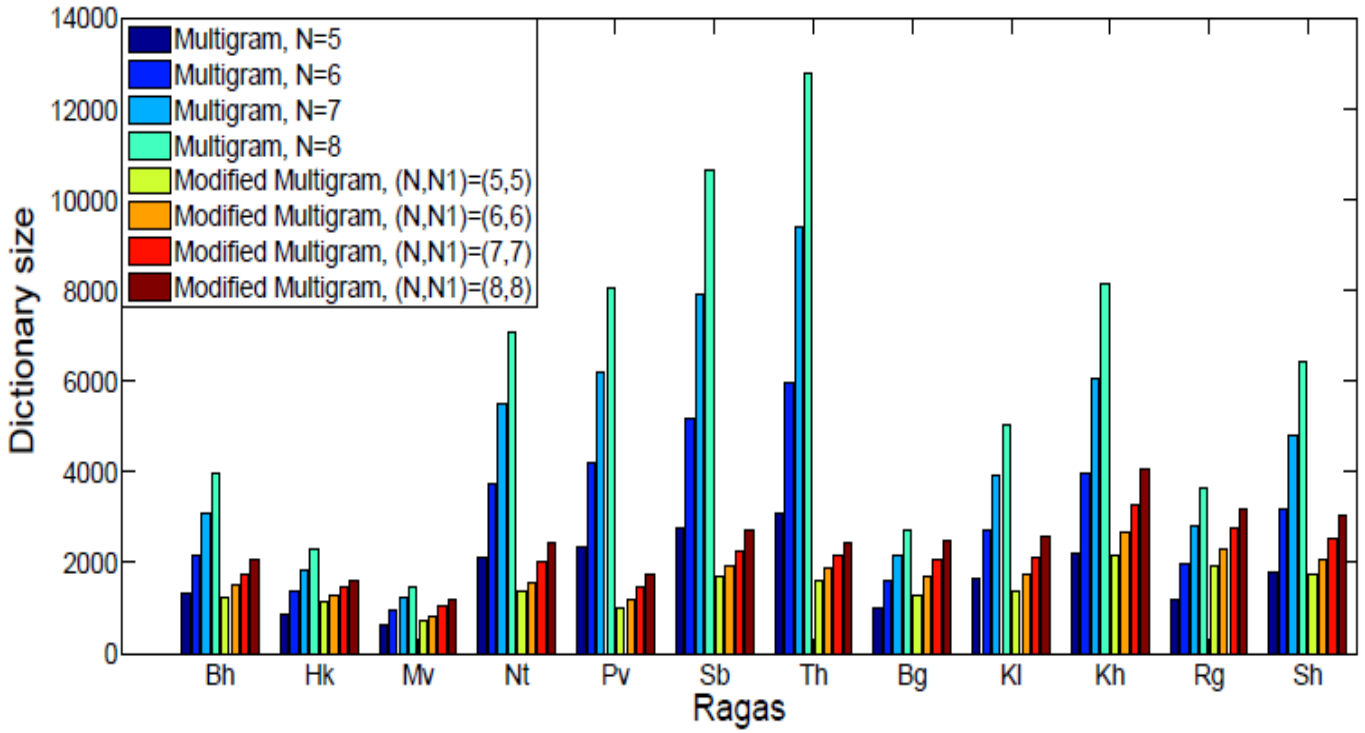
Perplexity values of N-multigram, (N, M) modified multigram on training and testing symbolic music data for the *rāgas* considered.

Performance: Correlation with musicological phrases

	Rāga (short form)	Phrases (Some accepted)
1.	*Begada*	{NSGRGG}, {GMPDPSS}, {SNDPSS}◀◀, {MMMMGRSS}◀, {RNDP}, {SPMGMR}◀◀, {GMPGMR}, {SNRS}◀
2.	*Reethigowla*	{GMNDM}, {GMPMGRS}, {GMNNS}◀◀, {NSGRS}, {NSGGMM}◀◀, {NPNNS}, {NNS}◀ {SNDMPDMGRS}, {NDM}◀◀, {SGRGGM}◀
3.	*Shahana*	{PMDNSS}, {NDPDPM}, {NSRGGRR}◀, {RGMPPP}◀, {RSNSDPMDD}, {MPDPMM}, {NSD}, {GMR}, {PMD}◀. {MGMRGRS}◀
4.	*Khamas*	{SMGM}◀, {GMNDNP}, {NDPMGMM}, {DNSNSS}, {NDPMPP}, {DNSS}◀, {SMMGRS}◀, {MGM}◀◀, {MGMNDD}◀◀
	Rāga (short form)	Phrases (Common)
1.	*Bhairavi*	{SGRGM}, {GRGMPDP}◀◀, {PMNNDP}, {MPMNDP}, {SSRNDN}, {MMPDP}, {SNRSGRN}◀◀
2.	*Nattai*	{MGMPP}◀◀, {MGPMGMR}◀◀, {SNSRRS}, {SNPMPNSSG}, {SNRRSNPMG}, {GMPNSNPMG}
3.	*Panthuvarali*	{NSRSSN}◀◀, {PMGGRSRG}, {GMDPPM}◀◀, {DND}, {NDPMGRGG}, {GMDN}◀, {RGRS}
4.	*Shankarabharana*	{RGMPMG}, {SRGSSN}, {MGMPDDP}, {SRGMPPMGGG}, {SDDNPPGMM}
5.	*Thodi*	{GMMGGGR},{PMG MPDDPMGRS}◀◀, {MPDDPM}◀, {PDNDPDPM}◀, {GMMDPPM}, {PDNDPM}◀, {RNS}◀ {DDDNSRSRGNSNDN}◀
6.	*Hari-Kambhoji*	{RGMG}, {RPMGRGRSSS}, {RGMPPPDNDP}, {MGRGRSS}, {SNDP}◀
7.	*Madhyamavathi*	{PMRMRS}◀◀, {RMP}◀, {MRPMRMRRRS}, {SMRPMNP}, {MPNN},{MRS}◀◀, {PRSRRR}◀
8.	*Kalyani*	{RGMPD}, {PMGGRS}◀◀, {PMGDP}◀, {MGRGRSR}, {RNNDP}

Phrases as marked by musicians. Those found by N-multigram marked by black left arrowhead, phrases found by 2-stage (N,M) multigram marked by red left arrowhead.

Performance: Total size of phrases discovered



A comparison of phrase set sizes for the considered rāgas obtained using multigram and the 2-stage multigram models for N and M ranging from 5 to 8.