

# Supplementary Material accompanying TISMIR submitted paper titled “Identifying Melodic Motifs and Stable Notes from Gestural Information in Indian Vocal Performances”

This document contains links to video snippets organized by the section number they correspond to in the main paper.

## Contents

### [Section 7S: Gesture-Based Classification Experiments](#)

[7S.1 Stable Note: Example of Singer Meher Paralikar](#)

[7S.2 Stable Note: Snippets showing misclassified segments](#)

[7S.3 Raga Phrase Detection Experiment](#)

[7S.3.1 Phrase r/P from Raga Shree](#)

[7S.3.2 Phrase nDN from Raga Bahar](#)

[7S.3.3 Phrase gmD from Raga Bageshree](#)

[7S3.4 Phrase P\R from Raga Nand](#)

[7S.4 Results for Singer-specific Gesture based phrase detection](#)

# Section 7S: Gesture-Based Classification Experiments

## 7S.1 Stable Note: Example of Singer Meher Paralikar

[Link to Video Folder](#)

The video snippets illustrate the characterization of stable and non-stable notes using gestures. If hand movements correspond to the singer's gestures we would expect that stable notes would be accompanied by relatively stationary hands, and melodic movements by moving hands. These clips illustrate that as a generalisation, this is what we find. However, unsurprisingly perhaps this correspondence does not always hold, and we can observe moments when a steady pitch is accompanied by hand movements (e.g. Video-2\_MP\_Alap1\_Shree\_137s\_157s, 4-5 secs). Also worth noting is the fact that in these cases the singer's left hand is held steady at abdomen, chest or shoulder height; his right hand, resting on his knee, is not involved in the gesture and does not help to distinguish stable notes. This may explain why Position information is useful in classifying Stable Notes.

Video-1: 20 s snippet from MP\_Alap1\_Bag (middle octave notes)

Video-2: 20 s snippet from MP\_Alap1\_Shree (higher octave notes)

## 7S.2 Stable Note: Snippets showing misclassified segments

[Link to Video Folder](#)

The video snippets show one case from either class that was misclassified by the SVM model. The first one shows a non-stable note classified as stable. Singer Sawani Shikhare (SS) moves her hands uniformly with very limited wrist motion, as a result of which the note is classified as stable. In the second snippet, we can see a stable note that is classified as non-stable, which may be because of the fact that there is some motion of the hands even when the note is stable. This is quite a common phenomenon in Hindustani vocal music: the singer holds a stable note while continuing to move the hand, as if signifying that the stable note is being stretched out.

Video-1: Non-stable note classified as stable (from SS\_Alap1\_Shree)

Video-2: Stable note classified as non-stable (from SS\_Alap1\_Shree)

## 7S.3 Raga Phrase Detection Experiment

### 7S.3.1 Phrase r/P from Raga Shree

[Link to Video Folder](#)

The videos show some examples of phrase r/P from singers Apoorva Gokhale (AG) and Sudokshina Chatterjee (SCh). We see a similarity in the gesturing style wherein the right wrist is involved in an upward motion during the glide.

We also see two instances of r/P by CC where the realisations differ significantly in the speed of rendering but display considerable similarity in the form of the hand movement.

The use of such hand-raising gestures to accompany this melodic movement, and the ways in which it is linked to the meanings ascribed to the raga, are discussed in detail in Leante (2009).

### 7S.3.2 Phrase nDN from Raga Bahar

[Link to Video Folder](#)

The videos show the phrase nDN from raga Bahar sung by singer Nishad Matange (NM). The first one is a pakad template while the next one is an alap instance. The gestures seem to be different in the two cases, which might be the reason for low performance in phrase nDN. The gestures have something in common - a high degree of mobility and an oscillatory movement from n down to D, then back up to N - but the movements occur in a different plane and with different hand positions, presumably making it difficult for a classifier to pick up the similarity.

### 7S.3.3 Phrase gmD from Raga Bageshree

[Link to Video Folder](#)

The examples from the pakad (query templates) show some variability: for AK and AG there is a clear gestural analogue to the rising pitch, SCh shows more lateral movement (out, to right of screen), and SM's movement is much more limited. Two examples from

AG's alap evidence the same feature seen in her pakad: recognition of this motif was in fact particularly good for this singer. (Both the singing and the hand movement seem to be more fluent in the alap version, where the middle note Ma is not clearly articulated, but there is enough similarity to allow successful detection.)

#### 7S3.4 Phrase P\R from Raga Nand

##### [Link to Video Folder](#)

This falling motif demonstrates the fact that we cannot expect consistent pitch height to hand height correlations, even in this dataset where gestures are assumed to refer to melody rather than rhythm or inter-performer communication. The clip of MP shows one possible gesture, in which the focus seems to shift from the singer's right hand to his left, the latter occupying a slightly lower spatial position. We have two examples from CC. One of these (56.5 - 59 s) is a near prototypical downward movement with the right hand falling while also moving inward. The other CC example, in contrast, illustrates a manual gesture which actually *rises* while the pitch falls. A possible interpretation of the latter case is that following a large fall from the 5th to the 2nd degree of the scale, the raised hand is marking the melody's resistance to a further fall to the tonic Sa, although this is only conjecture. What is clear is that singing-accompanying gesture is complex and multidimensional. Gestures accompanying this P\R motif seem to be more varied than those accompanying the Shree r/P described above, accounting for the lower detection accuracy.

## 7S.4 Results for Singer-specific Gesture based phrase detection

In the following tables, we report the F1-scores after stratified 10-fold cross-validation for the raga-phrase detection framework for each phrase. Bold values indicate results that are statistically significant when compared to a chance classifier which predicts 'Like' with probability equal to the percentage of like segments ( $p < 0.05$ ). In some of the cases, the singer did not sing the particular phrase in the pakad, or the alap instances were too few in number (less than 15 in total). In such cases, the singer's data was not considered in the experiment and the tables contain a blank '-' for them. Also, for ragas MM and Bahar, the overall number of segments was low because of which we decided to do a 3-fold cross validation rather than a 10-fold cross validation, since the implementation requires at least 1 example from each class in the testing data in every fold.

### Raga Shree, Phrase r/P

	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
Count	31	38	41	52	28	27	30	47	36	36	49
% Like	71	57.9	43.9	30.8	50	48.1	56.7	61.7	38.9	61.1	42.9
DTW-D (1)	83	71.2	0	11.8	80.0	48.3	74.4	<b>76.3</b>	81.5	75.9	0
DTW-I (1)	88.	73.3	0	40.0	59.3	23.5	73.7	<b>76.3</b>	63.6	75.9	0
DTW-LR (2)	76	71.2	48	43.5	76.9	43.5	72.2	<b>76.3</b>	<b>92.3</b>	75	58.8
DTW-Ind (36)	81	<b>78.0</b>	<b>77.8</b>	<b>85</b>	<b>82.8</b>	<b>69.6</b>	<b>82.4</b>	74.6	85.7	74.4	<b>66.7</b>
DTW-Ind-W (18)	<b>87.0</b>	60.5	<b>77.8</b>	78.6	75.9	64	74.3	73.3	89.7	<b>78.3</b>	42.9

### Raga Bageshree, Phrase gmD

	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
Count	30	28	34	35	19	23	28	32	24	36	42
% Like	60	60.7	47.1	31.4	68.4	60.8	46.4	56.3	58.3	50	57.1
DTW-D (1)	<b>87.2</b>	<b>77.3</b>	0	28.6	86.7	75.7	0	61.9	73.7	58.3	<b>72.7</b>
DTW-I (1)	75	75.6	0	26.7	77.4	75.7	0	<b>68.1</b>	73.7	41.2	<b>72.7</b>
DTW-LR (2)	<b>87.2</b>	76.2	55.2	60	77.4	70.6	34.8	<b>68.1</b>	<b>81.3</b>	62.5	70.8
DTW-Ind (36)	<b>87.2</b>	64.7	71.0	60.9	<b>88.9</b>	<b>83.9</b>	<b>72</b>	64.7	71.4	<b>73.7</b>	64
DTW-Ind-W (18)	78.9	60.6	<b>74.1</b>	<b>66.7</b>	84.6	78.6	60.9	66.7	72.2	66.7	65.3

### Raga Nand, Phrase P\|R

R	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
PIR Count	23	33	22	27	25	21	35	37	-	32	40
PIR % Like	78.3	42.4	40.9	25.9	48	47.6	42.8	43.2	-	53.1	42.5
PIR DTW-D (1)	<b>87.8</b>	0	0	54.5	14.3	<b>62.5</b>	0	0	-	69.4	9.5
PIR DTW-I (1)	<b>87.8</b>	0	46.2	61.5	37.5	16.7	47.6	64	-	72.3	0
PIR DTW-LR (2)	<b>87.8</b>	<b>66.7</b>	46.2	61.5	66.7	40	30	56	-	75.0	38.5
PIR DTW-Ind (36)	80	60	<b>55.6</b>	<b>72.7</b>	<b>72.7</b>	52.6	<b>64.5</b>	<b>66.7</b>	-	<b>80.0</b>	54.5
PIR DTW-Ind-W (18)	87.2	62.5	50	54.5	72	60	60	56.2	-	76.5	<b>60.6</b>

### Raga MM, Phrase nDN (3-fold CV)

	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
nDN Count	19	16	-	16	-	-	18	-	-	-	21
nDN % Like	68.4	56.3	-	56.3	-	-	61.1	-	-	-	28.6
nDN DTW-D (1)	81.1	78.6	-	72.2	-	-	75.6	-	-	-	0
nDN DTW-I (1)	81.1	72.2	-	72.2	-	-	18	-	-	-	0
nDN DTW-LR (2)	73.3	<b>95.2</b>	-	86.9	-	-	69	-	-	-	48.9
nDN DTW-Ind (36)	89.6	77.8	-	64.1	-	-	76.7	-	-	-	13.3
nDN DTW-Ind (18)	90.7	84.1	-	82.1	-	-	78.6	-	-	-	26.7

### Raga Bahar, Phrase nDN (3-fold CV)

	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
nDN Count	-	-	-	44	-	-	29	-	-	37	30
nDN %Like	-	-	-	38.6	-	-	62	-	-	56.8	36.7
nDN DTW-D (1)	-	-	-	61.8	-	-	75.1	-	-	71.7	11.1
nDN DTW-I (1)	-	-	-	25	-	-	75.5	-	-	71.7	52.4
nDN DTW-LR (2)	-	-	-	21.6	-	-	76.3	-	-	63.6	22.9
nDN DTW-Ind (36)	-	-	-	57.8	-	-	<b>81.1</b>	-	-	61	20.6
nDN DTW-Ind (18)	-	-	-	47	-	-	73.7	-	-	64.8	22.9