Exploring the Correspondence of Melodic Contour With Gesture in Raga Alap Singing

Shreyas Nadkarni[1], Sujoy Roychowdhury[1], Preeti Rao[1], Martin Clayton[2]
[1]Indian Institute of Technology Bombay, India, [2]Durham University, UK

Motivation

Meher Paralikar

- Indian vocal music has featured prominently in the area of Musical Gesture studies, due to its rich use of manual gesture

Dataset & Preprocessing

- **11 Hours** of Audiovisual Data
- Videos of **11 singers** (5 male, 6 female), each performing **9 ragas**: Bageshree, Bahar, Bilaskhani Todi, Jaunpuri, Kedar, Marwa, Miyan ki Malhar, Nand, Shree
- 2 alaps (average of 3 minutes) and 1 short pakad recorded for each singer-raga pair

Previous Empirical Work

- Paschalidou (2016) — movement and audio features in relation to the concept of 'effort' in *dhrupad* (Hindustani music)
- Clayton et al (2022) — classified 12-sec *khyal* (Hindustani music) excerpts using wrist movement data alongside audio
- Pearson and Pouw (2022) — alignment of kinematic extrema with acoustics changes in Karnatak music

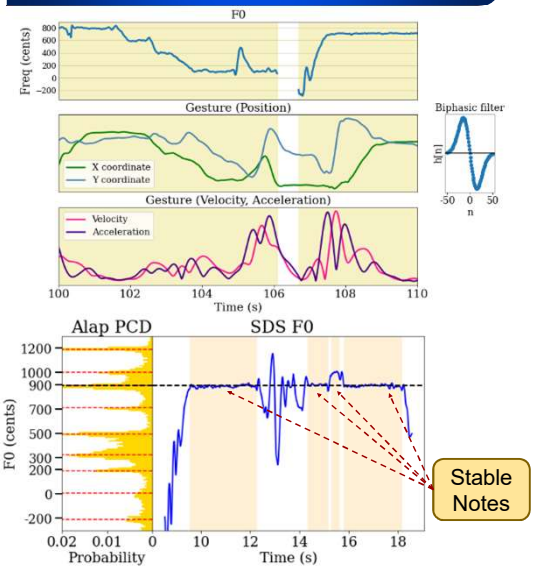
Our Question

Can wrist movement data be used to identify specific melodic features of Indian vocal music?

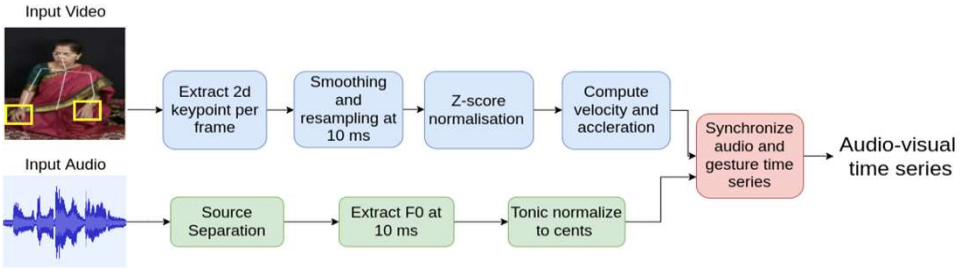
i. stable notes (bottom-up)

ii. raga-specific phrases (top-down)

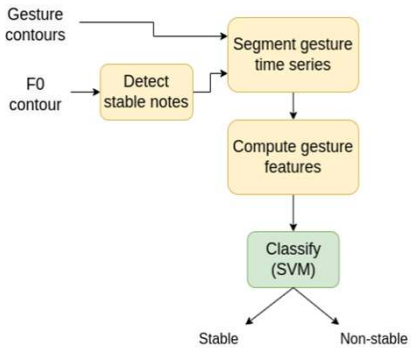
Audio-visual Time Series



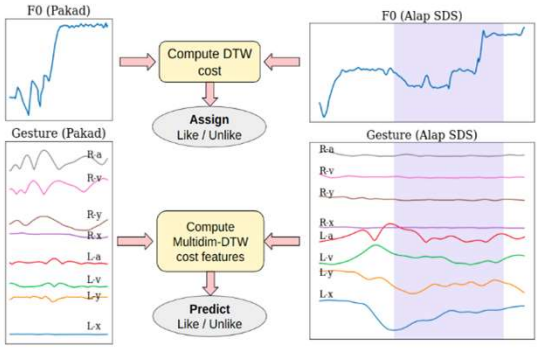
Audio-visual Processing Pipeline



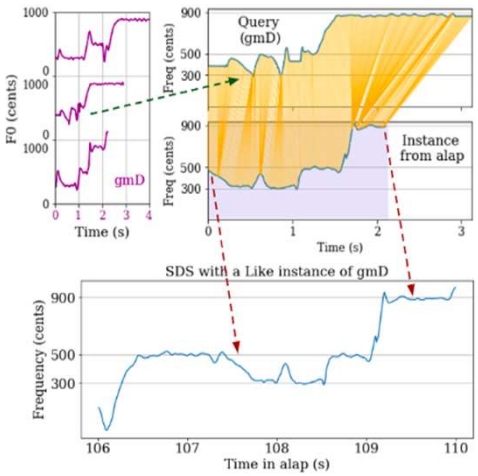
Stable Note Prediction with Gesture

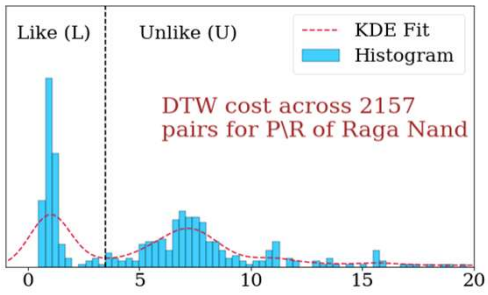


Like/Unlike Labelling & DTW



Raga Phrase Prediction with Gesture





Four Types of DTW Cost Features

- Dependent DTW (**DTW-D**): 1 cost from 8-D search
- Independent DTW (**DTW-I**): 1 cost from summing 8 one-dimensional searches
- **DTW-Ind**: 8 costs from one-dimensional searches
- **DTW-Left-Right**: 2 costs from two four dimensional searches (4 left wrist + 4 right wrist)

Results

Stable Note Classification

Using the 4 gesture kinematic features per wrist: Mean & S.D. of velocity, Mean & S.D. of acceleration

Singer	All	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
Count	20897	1242	1987	2382	2274	1822	2111	1769	1563	2069	2083	1595
% Stable	38.9	53.6	36.7	44.1	34.0	51.5	47.3	32.8	34.7	22.5	43.6	30.1
F1 Score (%)	65.7	81.1	63.6	69.5	68.2	72.5	71.6	65.8	65.2	60.5	75.1	49.2

Raga Phrase Prediction

Phrase	Like	Unlike	Chance Accuracy	DTW_D (1)	DTW_I (1)	DTW_{Ind} (8)	DTW_{LR} (2)
gmD	944	827	50.2	52.2	48.6	51.8	**52.4**
r/P	1035	1268	50.5	**55.3**	47.1	**56.1**	**55.1**
P\R	817	1340	53.0	**65.0**	45.7	**65.2**	**65.1**

Conclusion

- Raga-characteristic phrases and other low-level melodic events used to test the hypothesis of gestural consistency across and within singers.
- Experimental results indicate that there is significant kinematic information linked to the selected melodic events.
- We confirm the importance of computed velocity and acceleration profiles in the gesture representation.

References

1. Paschalidou, "Effort inference and prediction by acoustic and movement deS. scriptors in interactions with imaginary objects during dhrupad vocal improvisation," Wearable Technologies, vol. 3, p. e14, 2022.
2. Clayton, Rao, Shikharpur, Roychowdhury and Li, "Raga classification from vocal performances using multimodal analysis," in Proc. ISMIR, Bengaluru, 2022.
3. Pearson and Pouw, "Gesture–vocal coupling in Karnatak music performance: A neuro–bodily distributed aesthetic entanglement," Annals of the New York Academy of Sciences, vol. 1515, no. 1, 2022.