In Search of Sañcāras: Tradition-Informed Repeated Melodic Pattern Recognition in Carnatic Music

Thomas Nuttall[1], Genís Plaja[1], Lara Pearson[2], Xavier Serra[1]

ISMIR 2022 BENGALURU

Overview

We identify and group regions of variable-length, repeated, melodic patterns (sañcāras) in audio recordings of multiple Carnatic Music performances using a combination of transposition invariant features learnt by a Complex Autoencoder (CAE), predominant pitch tracks extracted using a Frequency-Temporal Attention Network (FTA-Net) and expert domain knowledge from practicing musicians/musicologists of the tradition.


Akkarai Sisters Subhalakshmi and Swarnalatha. Photograph: thehindu.com

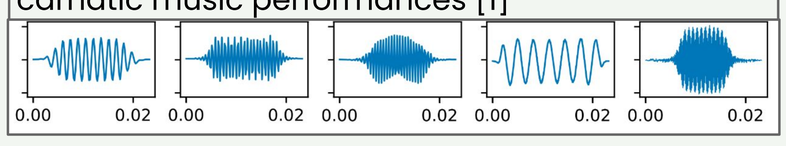
Characteristics of the Tradition

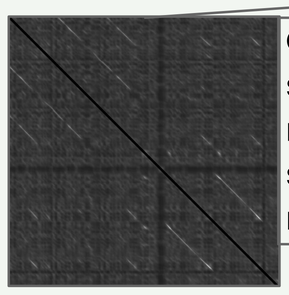
Certain tradition characteristics are **explicitly addressed** in the extraction process:

1. Sañcāras are often separated by features such as silence and long periods of stasis.
2. There may be multiple plausible segmentation points between longer phrases and shorter sub-segments
3. When repeated, a sañcāra may be repeated be immediately preceded or followed by different melodic material.
4. When a sañcāra is repeated, it is often elaborated on, with the insertion of additional svaras and gamakas.
5. There can be tempo variations between instances of the same sañcāra.
6. Single performances within a concert can range from between approximately 6 to 60 minutes in length.
7. Unique instrumentation - typically a vocalist accompanied by violin, mridangam and tambura.

Pipeline

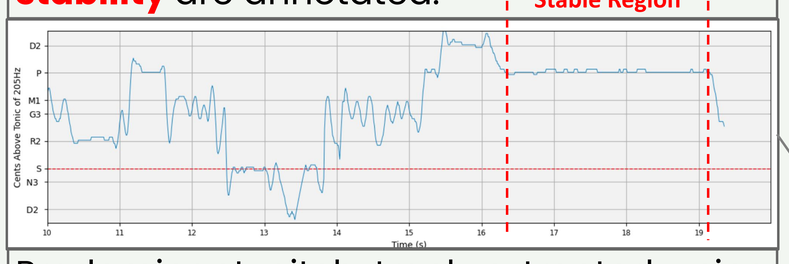
Melodic features embeddings from complex autoencoder trained on dataset of 100s of carnatic music performances [1]



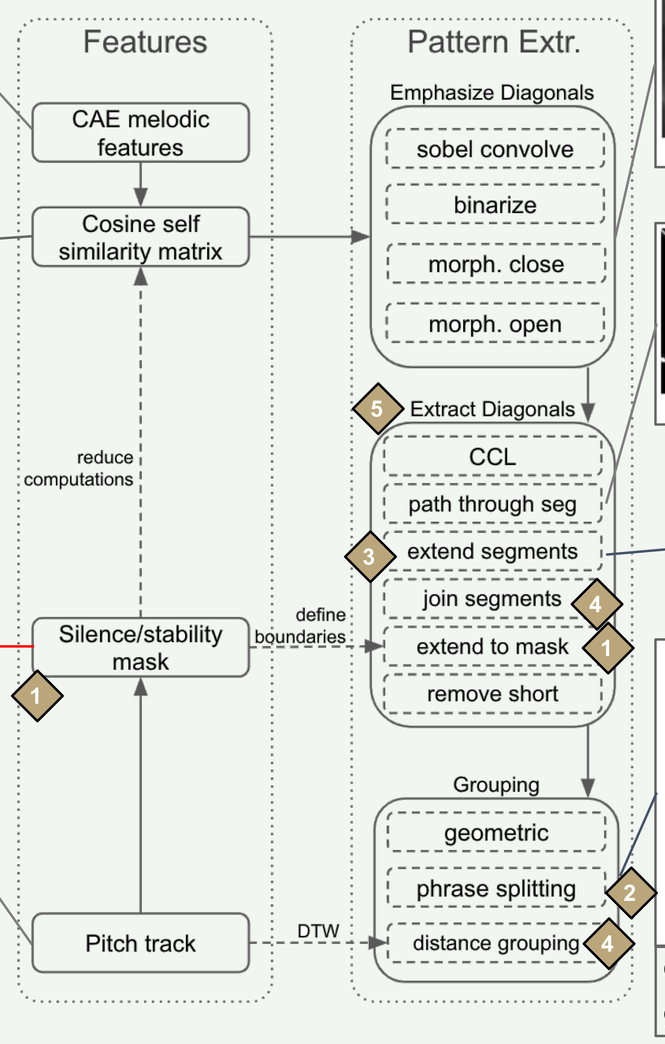


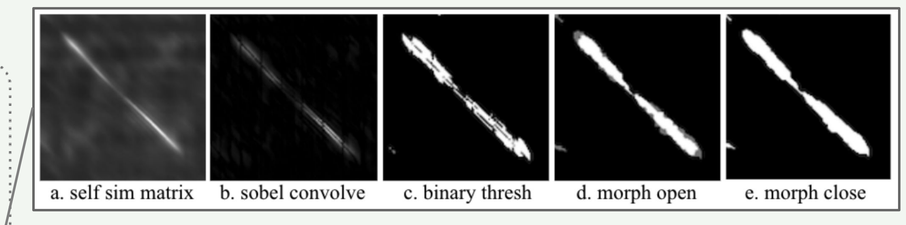
Cosine self-similarity matrix for non silent/stable regions 6

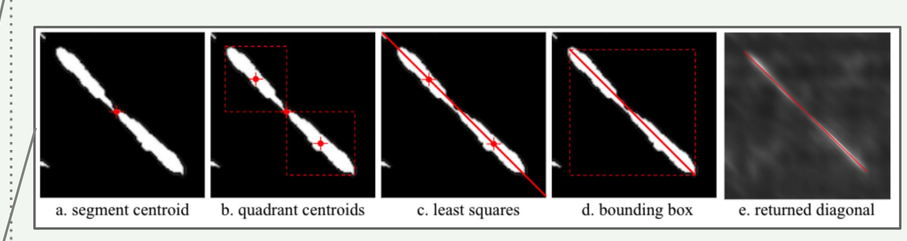
Regions that correspond to **silence** or **stability** are annotated:

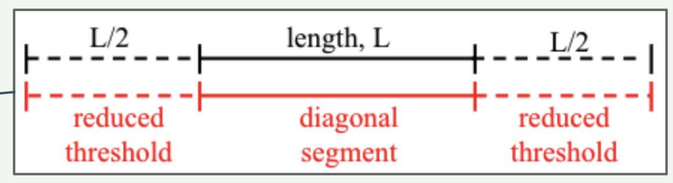


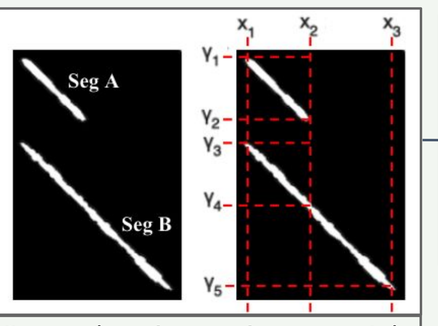
Predominant pitch track extracted using FTA-NET on bespoke carnatic dataset [2] 7



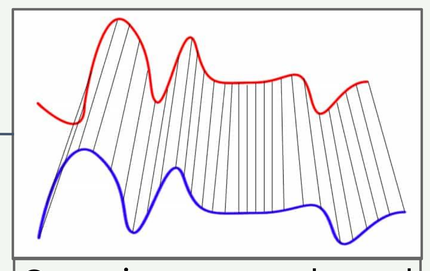






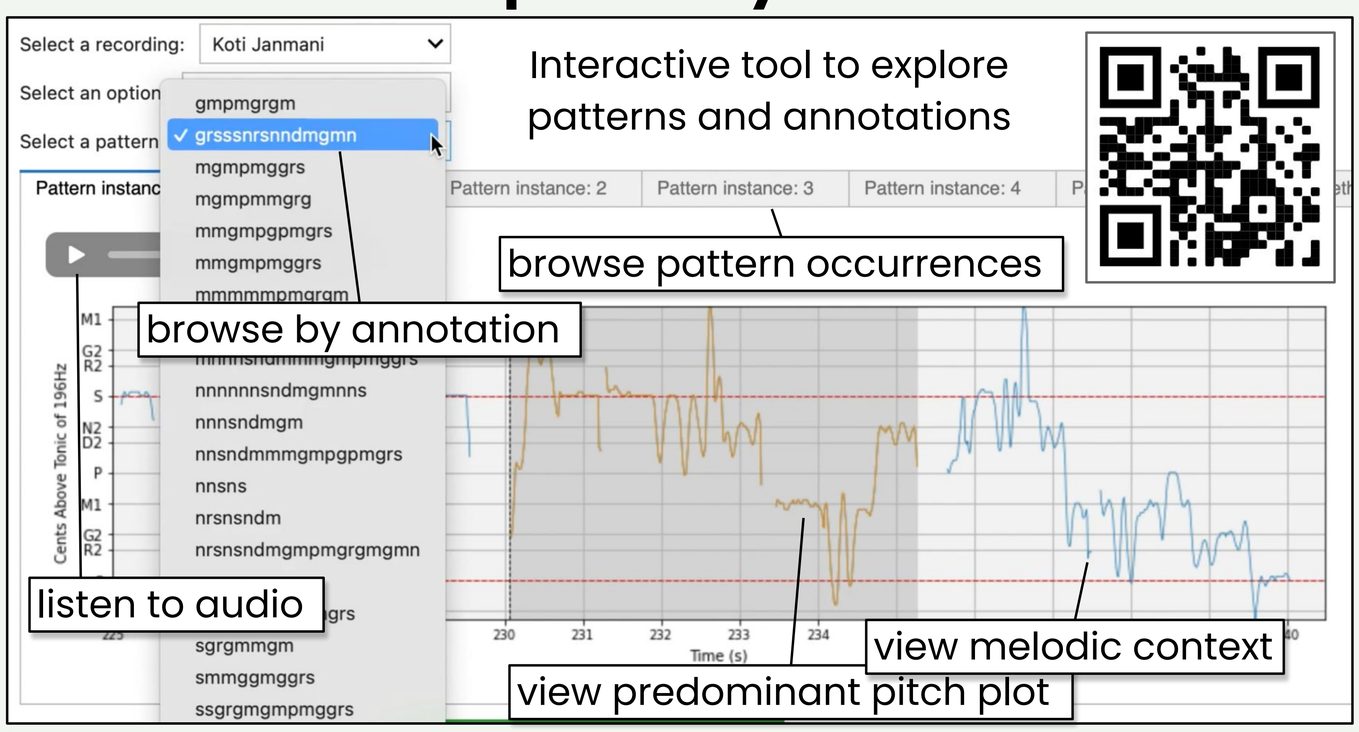


Grouping based on matrix geometry



Grouping groups based on mean DTW distance between them.

Exploratory Tool



Interactive tool to explore patterns and annotations

browse by annotation

browse pattern occurrences

listen to audio

view melodic context

view predominant pitch plot

Evaluation

Three carnatic performances annotated by professional Carnatic musician, Brindha Manickavasakan. A match is considered if there is 66% overlap between annotation and returned pattern. 2

Performance	N_r	N_a	Recall	Precision	F1
KJ	174	168	0.73	0.76	0.74
SJ	214	106	0.53	0.54	0.52
VNK	90	144	0.33	0.42	0.37
Overall	478	418	0.54	0.60	0.57

KJ - Koti Janmani - Akkarai Sisters (Ritigowla)
SJ - Sharanu Janakana - Salem Gayatri Venkatesan (Bilahari)
VNK - Vanajaksha Ninne Kori - Sumitra Vasudev (Ritigowla)

[1] S. Lattner, A. Arzt, and M. Dörfler, "Learning complex basis functions for invariant representations of audio," In Proc. of the 20th Int. Society for Music Information Retrieval Conf. (ISMIR), Delft, The Netherlands, 2019.
[2] G. Plaja-Roglans, T. Nuttall, L. Pearson, and X. Serra, "Repertoire-specific vocal pitch data generation for improved melodic analysis of carnatic music," 2022. [Online]. Available: 10.5281/zenodo.7036117

1 upf. MTG Music Technology Group
2 Max Planck Institute for Empirical Aesthetics

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
MUSICAL AI