IEEE Computational Intelligence Society

International Conference on
Emerging Techniques in Computational Intelligence
22–24 August 2024

IEEE

Self-Organizing Map for Automated Quality Assessment in Algorithmic Music Computation

Prafulla Kalapatapu, **Ventrapragada Sai Srikar**, Aakarsh Reddy, Vineeth Mallavarapu, Bhargav Karra and Arya Kumar Bhattacharya

Mahindra University
Global Thinkers. Engaged Leaders.

Problem Statement

- This research is a part of work that proposes a unique algorithmic music composition system, using Differential Evolution (DE) as an Evolutionary Algorithm (EA) to synthesize musical creations.

- The entirety of the work maps musical representations into chromosomes using fundamental low level frequency-amplitude-phase representations of sound, aiming for maximal alignment with universal and pristine human emotional states (Rasas).

- Self Organizing Maps (SOM) are employed as a Fitness Function and discriminator across emotional states, guiding the evolutionary process toward targeted emotional directions.

- In this work we delve into the challenges in modelling a Fitness Function (FF) for the given problem statement. The Fitness Function (FF) must accurately assesses composition quality based on defined metrics and generate continuous quality scores that capture the nuanced improvements in the music generated.

Introduction

- **Music Emotion Modeling** involves understanding and mapping human emotions to musical elements, ensuring that the compositions evoke the desired emotional responses.

- **Evolutionary Algorithms** are employed to optimize the music generation process, leveraging principles of natural selection, mutation and crossover to evolve musical pieces that meet the emotional criteria.

- **The Objective Function** is critical for evaluating how well a given musical piece aligns with the intended emotions, guiding the evolutionary process towards producing emotionally resonant music.

- **Generative AI** models often rely on patterns from existing data, potentially limiting their ability to create truly original and emotionally resonant music. Whereas Evolutionary Algorithms creatively explore a vast musical search space and create novel combinations that align with target emotions

Music Emotional Modelling - Rasas

- According to ancient Indian philosophy all human emotions can be categorized into nine pure form classes. Rasas refer to these emotional flavors or essences that can be evoked through art, including music.
- The four pure form Rasas or emotions amenable to musical representation are Karuna (sorrow), Shanta (peacefulness), Shringar (excitement and romanticism) and Veera (thrill for action).
- Audio example of Shringar
- Audio example of Veera
- The decision to use only these four Rasas, ensures that the music we create is emotionally resonant, universally comprehensible, culturally neutral and optimized for effective representation. These Rasas provide a balanced and inclusive foundation for our music composition process, allowing us to create emotionally rich and universally appealing musical pieces.

Evolutionary Algorithms – Differential Evolution

- Evolutionary Algorithms (EAs) mimic natural evolution to explore complex solution spaces effectively. They generate an initial population of musical compositions and iteratively apply mutation, crossover and selection to improve their alignment with the intended Rasas

- The most profound springs of creativity lie within human nature itself. Unlike other methods that interpolate within a space of existing examples, Evolutionary Algorithms (EAs) venture into uncharted territories, enabling the discovery of truly original and innovative solutions.

- Here we employ Evolutionary Algorithms (EAs) and specifically Differential Evolution (DE) due to its innate capacity to transition focus from exploration to exploitation as evolution advances through generations.

The Objective Function – Fitness Function

- The Objective Function serves as the cornerstone of the optimization process in Differential Evolution (DE), providing the criteria for evaluating solutions, driving selection pressure, guiding convergence and shaping the exploration of the solution space. Its careful design and implementation are essential for the effectiveness and efficiency of the Evolutionary Algorithm (EAs).

- The prime challenge is to synthesize a Fitness Function (FF) that can accurately detect the "degree of adherence" of the composition to the targeted Rasa.

- The fitness of generated musical compositions is assessed using Self Organizing Maps (SOM), learning clusters aligned with Rasas classified based on our compiled dataset. Each composition targets a specific Rasa, with fitness determined by adherence to the target Rasa; minimizing distance signifies higher fitness.

Self Organizing Map (SOM) – Fitness Function

- A training sample of 62 songs, evenly distributed among four Rasas was used. Each song's mid-level features were extracted into 102 variables using **jAudio**. The model must learn from these 62 samples to function effectively as a **Fitness Function (FF)** .

- For our work, we used a **SOM** architected as 12x12, i.e. it had 12 rows and 12 columns of output nodes. The initial neighbourhood was set to 4X4 while the maximum number of iterations were limited to 30,000.

- **SOM's** create a juxtaposition of *Competitive Learning* and *Co-operative Learning*

- Initially let us consider for any given **i^{th}** song **F_i** to be the transpose of the feature vector of size 102 for that particular song.

$$F_i = [f_1, f_2, \ldots, f_{102}]^T \qquad (1)$$

Self Organizing Map (SOM) – Fitness Function

- Given that we randomly initialize weights between 0 and 1 for each node in the **SOM.** Here each node connects to all the input features. For example for any given **j^{th}** node let **W_j** be the transpose of weights of size 102 for that particular node.

$$W_j = [w_1, w_2, \ldots, w_{102}]^T \qquad (2)$$

- We then find the **Best Matching Unit(BMU)** for the respective input song and update the weights of the nodes.

$$d_j = \sum_{i=1}^{n} (x_i - w_{ij})^2 \qquad (3)$$

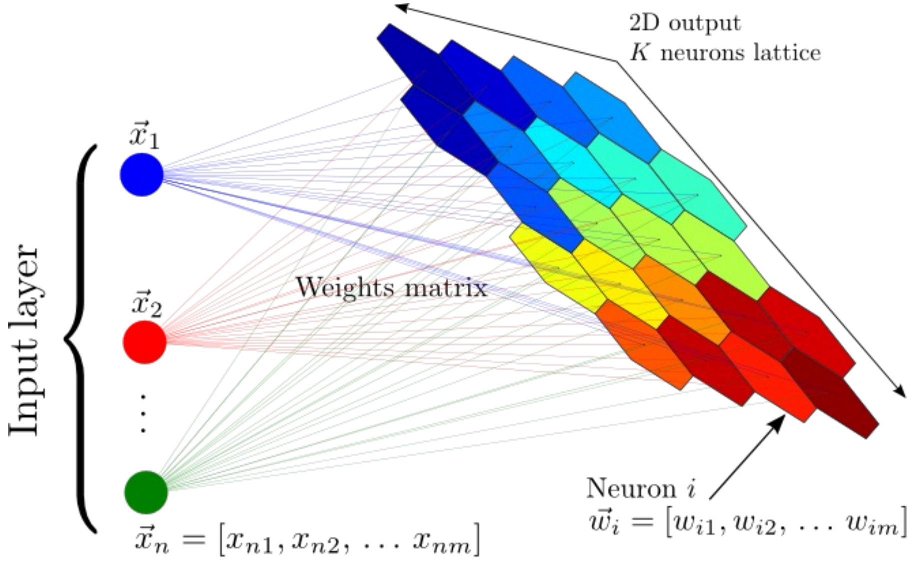
Self Organizing Map (SOM) – Fitness Function

$$i(F_i) = \arg\min_j \|F_i - W_j\| \quad \forall j \in \{1, \ldots, l\} \tag{4}$$

$$w_j(k+1) = w_j(k) + \eta(k) h_{ij(x)}(k)(x - w_j(k)) \tag{5}$$

$$h_{ij(x)} = \exp\left(-\frac{d_{\text{node}}}{2 \times \text{radius}^2}\right) \tag{6}$$

$$d_{node} = (row_{win} - p)^2 + (col_{win} - q)^2 \tag{7}$$



- Essentially different regions of the **SOM** learn different pattern present in the higher dimensional input space.

Self Organizing Map (SOM) – Fitness Function

- By defining the degree of belonging '**z**' to a specific Rasa based on the Euclidean distance '**d**' from the created composition to the Rasa's representative point in the SOM. The degree of belonging is defined as:

$$z = \frac{1}{1+d} \quad (8)$$

- Two versions of the FF were defined:
  - Version 1: $$FF_{SOM,1} = -\ln(z^2) \quad (9)$$
  - Version 2: $$FF_{SOM,2} = d^2 \quad (10)$$
- Both versions reach their minimum value as '**d**' approaches zero and increase nonlinearly with increasing distance.

Results

- We had tested out different well known clustering and classification algorithms, namely **DBSCAN**, **Fuzzy C-Means** and **SOM**. Except the latter, the others failed to achieve even rudimentary levels of accuracy under these challenging conditions.

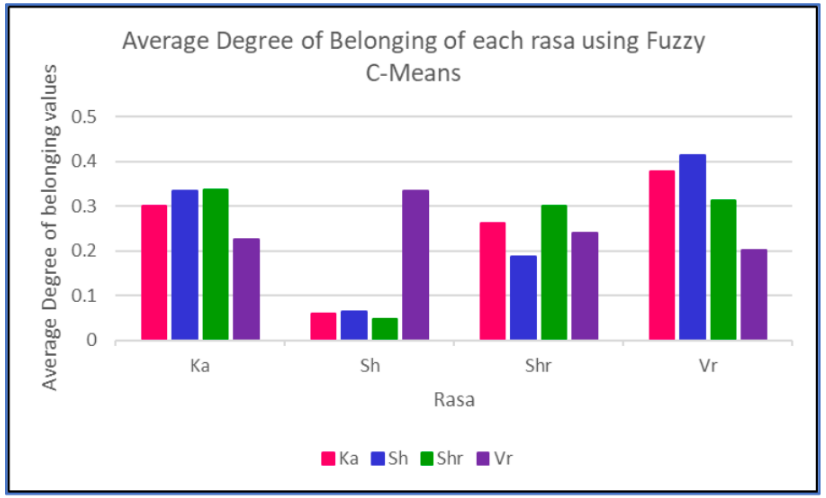


Fig. 8. Average degree of belonging of each Rasa using FCM; Ka implies *Karuna*, Sh *Shanta*, Shr *Shringar* and Vr implies *Veera*

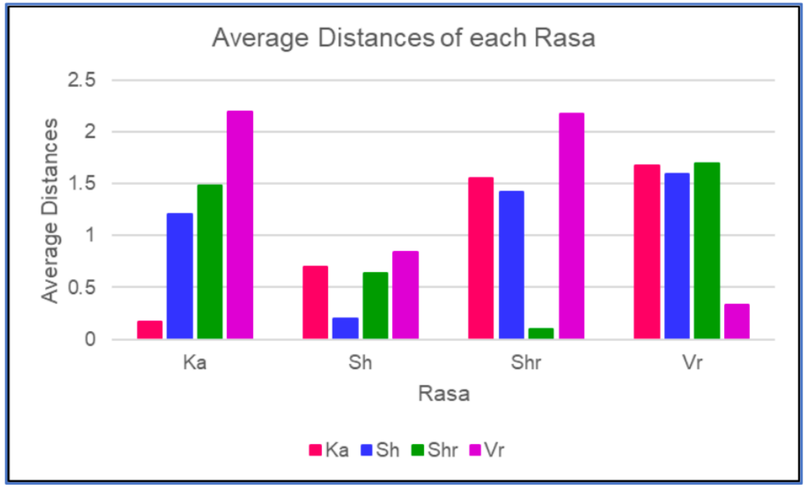


Fig. 3. Average distances of songs of each Rasa from targeted Rasa's node (shown on x-axis), using SOM; note Ka implies *Karuna*, Sh *Shanta*, Shr *Shringar* and Vr implies *Veera*.

Results

- The trained **SOM** was able to classify the given songs with an accuracy of 83.33 %
- In Fig. 4, x-axis represents the index numbers of the songs in each Rada and y-axis represents the distance of a song in a **Rasa** to the **BMU** for **Karuna**. The 4 curves represent the distances of songs in each of the 4 Rasas to that BMU.

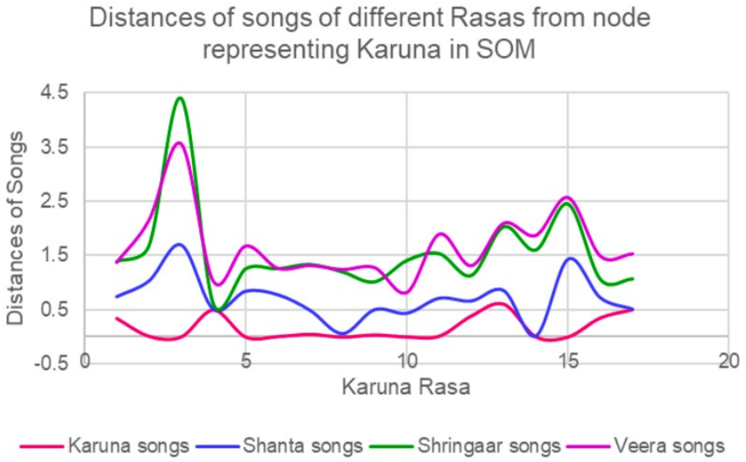


Fig. 4. Distances of songs of different Rasas from node representing Karuna in SOM.

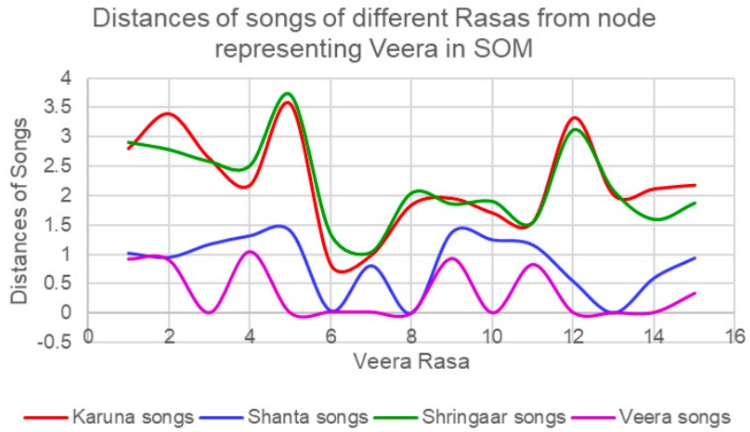


Fig. 7. Distances of songs of different Rasas from node representing Veera in SOM.

Self-Organizing Map for Automated Quality Assessment in Algorithmic Music Computation

Results

- Some sample songs that have were created with the proposed method are as follows:
- Veera:
- Shringar:

Veera Convergence history

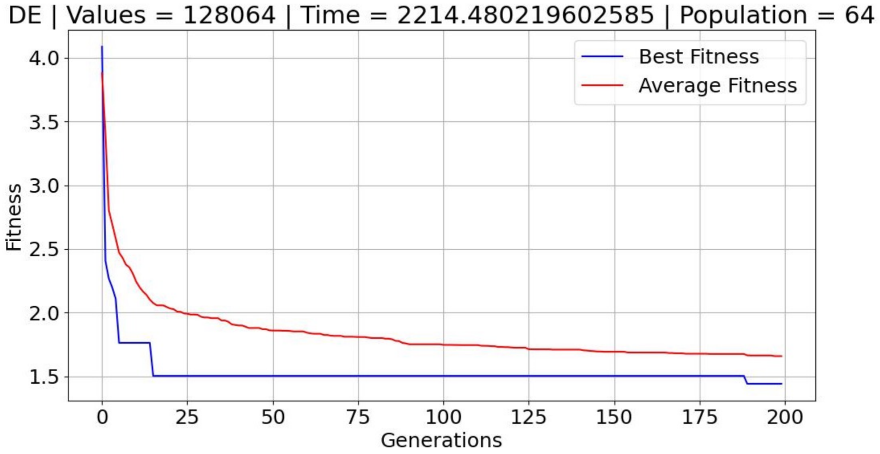


Table 2: Average Distances of every song for each Rasa, from the representative node for that Rasa in the SOM architecture

Rasa / Distances	Karuna	Shanta	Shringaar	Veera
Karuna	0.17	0.7	1.55	1.68
Shanta	1.21	0.2	1.42	1.6
Shringaar	1.48	0.64	0.1	1.7
Veera	2.2	0.84	2.18	0.33

Conclusion

- Our developed model has successfully demonstrated the efficacy of employing Self-Organizing Maps (SOM) for automated quality assessment in the realm of algorithmic music composition.
- In summary while both SOM and Fuzzy C-Means involve distance calculations, they serve different purposes and represent different aspects of data relationships. For our class of problem, SOM outperforms Fuzzy C-Means by clearly classifying the Rasas.
- This interdisciplinary approach, blending facets from machine learning, evolutionary computation, and music theory, holds promise for fostering a deeper understanding of the emotional impact of computationally generated musical compositions.

Future Work

- Expanding our original song data base, create longer songs and handle longer chromosomes. Generate variety by modulation of different parameters

- Having additional mechanisms for judging and validating adherence of the creation to the targeted Rasa – e.g. Fitness Function based on Information Entropy, Spectrogram-based classification, improving the SOM itself

- Generating songs reflecting mixture of Rasas (emotions) by using Multi-Objective Optimization having 2 or even 3 Rasas as objectives, and then a converged Pareto front from where the user may choose randomly based on her priorities and preferences

Important References

- P. Kalapatapu, I. C. Palli, R. T. Gangavarapu, A. K. Bhattacharya and N. K, "Emotion Aligned Music Composition from Sound Fundamentals Using Differential Evolution," *2024 IEEE Congress on Evolutionary Computation (CEC)*, Yokohama, Japan, 2024, pp. 01-08, doi: 10.1109/CEC60901.2024.10612152.

- T.Kohonen, T.Honkela, “Kohonen Network”, Scholarpedia, 2(1), http://www.scholarpedia.org/article/Kohonen_network, 2007. DOI: doi:10.4249/scholarpedia.1568

- R.Storn and K,Price, “Differential Evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces”, ICSI Technical Report TR-95-012, March 1995.

Thank You

prafulla.kalapatapu@mahindrauniversity.edu.in