International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

Generation of Indian Classical Music using Artificial Intelligence: A Survey

Bharavi U. Desai[1], Kunjal I. Tandel[2], Rahul M. Patel[3]

[1]Master's Student, Dr. S & S S Gandhy Engineering College, Gujarat, India

[2]Professor, Government Engineering College, Gujarat, India

[3]Professor, Dr. S & S S Gandhy Engineering College, Gujarat, India

Abstract: ***This paper is a survey and an examination of the various ways in which classical music content is produced. For our survey, we propose a framework focused on five dimensions: Objective, Representation, Architecture, Challenge and strategy. For each dimension, we conduct a comparative analysis of various models and techniques and we propose some tentative multidimensional typology. This typology is bottom-up, based on the analysis of many existing deep-learning based systems for music generation selected from the relevant literature.***

Keywords: LSTM, GRU, RNN, Arohana, Avrohana, FSM, Gamakas, HMM

1. Introduction

When Music is an artistic process that includes the emotion and the concept of a composer. The first computer-composed music appeared in 1957[1], and since then more and more music has been produced using computer technology. There have been a variety of algorithmic methods for creating music, such as grammatical based, Markov models, neural networks and other deep learning methods. As shown by the survey [2], the benefit of the use of deep learning (including machine learning) to produce musical content is its generality. Compared to other approaches, a machine learning based system can learn a pattern from an undefined corpus of music. In particular, several researchers have found the task of producing music to be a probabilistic model of monophonic music, represent music as a series of notes, and attempt to model music as a probability distribution, where the next note was allocated based on the probabilities of the previous note sequence and some background, such as chord, beat. Specifically, compared to a rule-based composition, we can train a particular model based on a huge number of musical corpus and allow it to explore patterns automatically.

During the generation process, we sample from a qualified probability distribution to generate new music pieces. Recurrent neural network (RNN), especially long-term memory networks (LSTMs) [3], has been shown to significantly predict time series data. In reality, multiple researchers have used LSTM to produce music, which has also provided good results. In Indian classical music, Raaga is one of the melodic styles of music. This is identical to a scale but also distinguished by a number of ornamentations, such as slides, vibrato, pitch bends, etc. and consists of five or more swaras arranged in a particular manner. The exact frequency of swaras used in some raagas that differ slightly from the reference frequency. Selected swaras are sung / played in ascending and descending order, depending on the music criteria, along with the appropriate ornamentation. In certain raagas, the notes may not be in ascending or descending order.

There are many applications by the Hindustani Classical Music Artificial Composer. The use of a generative method will be of great benefit to performers who want to pursue new forms of improvisation in a specific Raga of interest. A computer-generated composition may have drastically different qualities than a human-created one, in the same Raga. There might be unusual combinations of notes in vistaras or tanas, which will certainly be difficult for a human being to think about. Therefore, such a program will help to uncover new possibilities in the Raga that have not been explored by performers and learners. There are many applications by the Hindustani Classical Music Artificial Composer.

The use of a generative method will be of great benefit to performers who want to pursue new forms of improvisation in a specific Raga of interest. A computer-generated composition may have drastically different qualities than a human-created one, in the same Raga. There might be unusual combinations of notes in vistaras or taans, which will certainly be difficult for a human being to think about. Therefore, such a program will help to uncover new possibilities in the Raga that have not been explored by performers and learners.

2. Basic Characteristics

Indian classical music is focused on raga and its various characteristics. Raga is similar to melody, but more complex than melody in western music. Raga is a set of various specific notes that have certain unique properties (e.g. aroha, avaroha, taal, pakad, etc.).

2.1 Raga:

EachRaga comprises of a specific hierarchical structure of swaras or notes. Notes are termed swaras in classical Indian music. The fundamental seven notes or swaras or symbols in classical music are S (Sa), R (Re or Ri), G (Ga), M(Ma), P (Pa), D (Dha), N (Ni) that can be perceived analogous to C, D, E, F, G, B, A[1,2]. In other words, we have 12 swaras or shrutis in Carnatic music, S, r, R, g, G, rn, M, p, d, D, n, and N.

2.2 Arohana and Avrohana:

Raga is a set of swaras or documents. Its specific identity is created by raga, depends on the set of notes or the fusion of swara and arohana and avarohana. Arohana is a subset of the series of raga notes structured in an ascending sequence. Avarohana is a subset of the series of raga notes structured in a descending sequence.

3. Generation Techniques

In this section, we are going to present a survey of existing system dealing with the Classical music generation with their different approaches, implementations and issues regarding these systems.

3.1 Using State Space Model

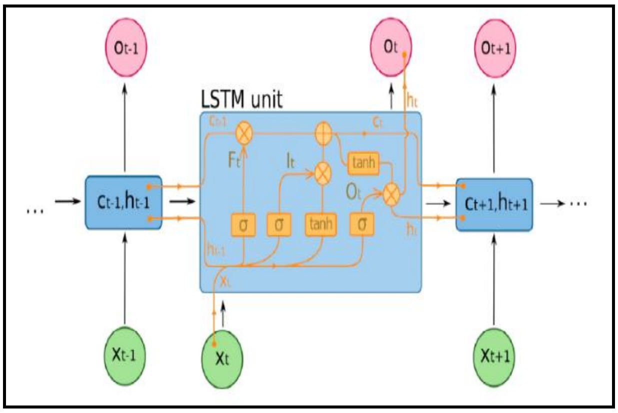
The main purpose and contribution of this paper is to use classical probabilistic time series models to produce piano pieces from the Romantic period that sound like human compositions, and to establish metrics to compare the compositions produced with the originals in terms of originality, musicality and temporal structure [4]. The models and datasets used here are, the 14 time series models we used to generate compositions, and the 10 piano pieces from the Romantic era we used as training data. . The TVAR model for a sequence (X1; :::;XT) is specified as:

$$X_t = \sum_{j=1}^{d} \theta_j\left(\frac{t-1}{T}\right) X_{k-j} + \sigma\left(\frac{t}{T}\right)\varepsilon_t, \quad t = 1,\dots,T \qquad (1)$$

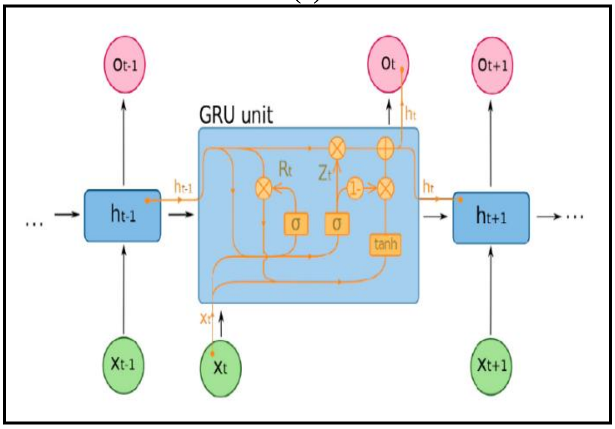
Time-Varying Auto-regression (TVAR) models have been tested to model the non-stationarity and periodicity of the note pitches for the original pieces. TVAR models are capable of modeling non-stationarity and longer dependency over time between note pitches. The biggest shortcoming was the lack of global structure or long-term melodic progression of the produced sections.

3.2 Using Recurrent Neural Network

A recurrent neural network (RNN) is a category of artificial neural network where correlations between nodes form a graph along a series. This allows time dependent behavior to be seen for a time series [5]. Initially, they used 3 GRU layers stacked one on top of each other, with a hidden neuron count of 512, but then they also chose to play with LSTM layers and varied the number of hidden neurons and the number of layers in order to obtain a deeper understanding of what is best achieved.



(a)



(b)

Figure 3.2: (a) LSTM unit, (b) GRU unit

It comprises 35 manually transposed to Note Worthy J. S. Bach works, with an average length of 50 musical measures. They have been normalized to a measure length of 6/8, in the B minor scale, with a tempo of 100 [5]. The GRU out performs the LSTM, reaching a training accuracy of 95% and a validation accuracy well above75%.

3.3 Using Hidden Markov Model

Statistical techniques along with time series analysis of periodograms for different instruments, structural models for melody and harmony decompositions, Markov chains for determining the probability of fluctuations between tween pitch intervals and discriminant analysis for the identification of pitch and historic period compositions have all been used to study classical music. The original piano pieces were downloaded from mfiles.co.uk (2017), Krueger (2016) and MIDIworld (2009) in Musical Instrument Digital Interface (MIDI) format. The RMSE was mainly used to evaluate the pieces and to pick the most highly produced pieces for assessment by human listeners and to gain insight into some of the general patterns found in the produced pieces [6].

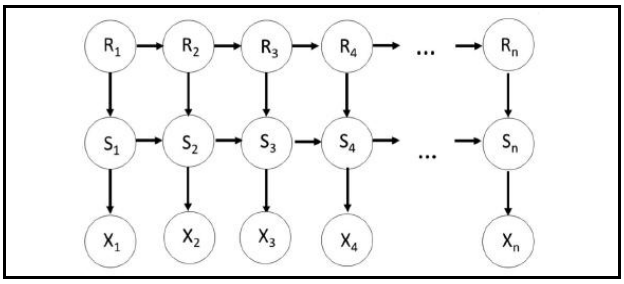


Figure 3.3: Directed graph of the HMM with two hidden states. Both the R1: n and the S1:n are hidden states.

The major limitation in using Hidden Markov Models to generate music is that, the generated pieces suffered most notably from a lack of global structure or long-term melodic progression. Generating musical pieces that have a global structure will be a key challenge to overcome in order for statistical models such as HMMs to be able to compose music at the same level as humans.

3.4 Using Finite State Machine

The probabilistic finite state machine (FSM) for each rising and decaying movement for an unique Raga is built for generation purposes. In the current implementation of the FSM, the probabilities p1, p2 and p3 for the three distinct edges e1, e2 and e3 for the individual node n are set for all Ragas [7]. The probability used is p1 = 0.70, p2 = 0.20 and p3 = 0.10.

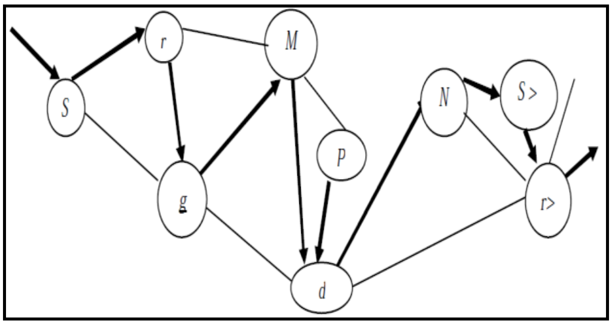


Figure 3.4: The part of an example transition diagram for the Arohana for Raga Miyan Ki Todi.

The total of 15 compositions are mixed up and presented to a subject. This study was carried out on 10 different people, five of whom had formal education in Indian Classical Music, and the others with little or no experience but with an interest in music. Since the mathematical model behind the system is universal for all Ragas and is just a large-scale representation of the notes that the Raga makes, an expert in Hindustani Classical music will clearly understand the inability of the system to achieve the structure of complex Ragas.

3.5 Using Bidirectional RNN:

In time direction, we allocate a single instance of the network to every note. In order to assure "translation invariance property," each network instance will have bound weights and each note will have its output in the same calculation process [8]. In addition, the use of coupled weights can greatly reduce the number of parameters in the system and minimize the chance of over-fitting.

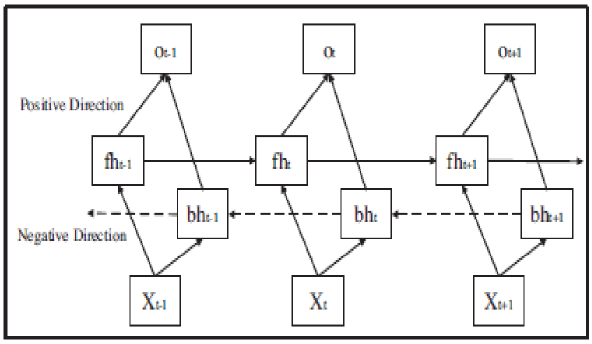


Figure 3.5: General structure of the BRNN shown unfolded in time for three time steps

The model trained on Classical Piano Dataset includes 295 MIDI files, which are included. MIDI files have a standard pitch range which ranges from 0 to 127, but in the train MIDI file, the vast majority of note pitch ranges from 48 to 96. Music generated by this model has lack of long term structure same as the most existing systems.

4. Our method

For each raagam, a finite state machine can be designed in such a way that the notes corresponding to that raagam may form states and multiple octaves can be protected by assigning separate states for notes in each octave as shown in the figure above. Raagam provides a structure that enables the use of Finite State Machines to generate Indian classical music. Raagam, can be considered to be a formula to compose a piece.

The transitions between states are defined by the probabilities that are learned by examining the transitions of notes in sections that are centered on that raagam. When such an FSM is designed, we could just seed the FSM with a random note, and this dynamic system will keep producing notes based on these transfer probabilities. For training purpose we have given one specified raagam with its MIDI files. There is a shortcoming of accurate gamakas in this model.

5. Conclusion

All the techniques have their own strategies and concepts. Even though having shortcomings they are very good with their ideas. Education of Classical music is very essential for the root knowledge of music. These models will help anybody to go deeper in to the musical education and with as easy efforts. We have proposed a multi-criteria conceptual framework based on five challenge and strategy.

We have analyzed and compared various systems and experiments proposed by various researchers in the literature. We hope that the conceptual framework provided in this book will help in understanding the issues and in comparing various approaches for using all machine learning techniques for music generation, and therefore contribute to this research agenda.

6. Acknowledgement

This paper and the survey behind it, would not have been possible without the exceptional support of my external guide as well as co-author, Prof. Kunjal I. Tandel. His enthusiasm, knowledge and exacting attention to detail have been an inspiration and kept my work on track from my first encounter with the Technical aspects of Music to the final draft of this paper.

Moreover, I also like to appreciate the efforts of my internal guide and co-author, Prof. Rahul M. Patel. Without their guidance I never have reached to this level. I also would like to thank my college, Dr. S & S S Gandhy government Engineering College and its staff and their technical support.

References

[1] Lejaren A. Hiller and Leonard M. Isaacson. Experimental Music: Composition with an Electronic Computer. McGraw-Hill, 1959.
[2] Jose D. Fern´andez and Francisco Vico. AI methods in algorithmic composition: A comprehensive survey. Journal of Artificial Intelligence Research, 2013.
[3] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural computation, 1997, 9(8): 1735-1780.
[4] Yanchenko, Anna & Mukherjee, Sayan. (2017). Classical Music Composition Using State Space Models.
[5] Marinescu, Alexandru-Ion. (2019). Bach 2.0 - generating classical music using recurrent neural etworks. Procedia Computer Science. 159. 117-124. 10.1016/j.procs.2019.09.166.
[6] Yanchenko, Anna. “Classical Music Composition Using Hidden Markov Models.” (2017).
[7] Das, D. and Choudhury, M. 2005. “Finite state models for generation of Hindustani classical music”. In Proceedings of the International Symposium on Frontiers of Research in Speech and music (FRSM - 2005) (pp. 59–64), Bhubaneswar, India [Google Scholar]
[8] T. Jiang, Q. Xiao and X. Yin, "Music Generation Using Bidirectional Recurrent Network," 2019 IEEE 2nd International Conference on Electronics Technology (ICET), Chengdu, China, 2019, pp. 564-569.
[9] Lejaren A. Hiller and Leonard M. Isaacson. Experimental Music: Composition with an Electronic Computer. McGraw-Hill, 1959.
[10] Schuster, Mike, and Kuldip K. Paliwal. "Bidirectional recurrent neural networks." IEEE Transactions on Signal Processing 45.11 (1997): 2673-2681.
[11] J. D.D., “Generating polyphonic music using tied parallel networks,”Correia J., Ciesielski V., Liapis A. (eds) Computational Intelligence in Music, Sound, Art and Design. EvoMUSART 2017., 2017.[Online].
[12] Choudhury M. and Ray P.R. (2003) Measuring Similarities Across Musical Compositions: An Approach Based on the Raga Paradigm. Proceedings of International Workshop on Frontiers of Research in Speech and Music, Kanpur, 2003.
[13] Franc¸ois Pachet, Pierre Roy, and Gabriele Barbieri. Finite-length markov processes with constraints. In Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011), pages 635–642, Barcelona, Spain, July 2011.
[14] Steedman, M. (1984). A generative grammar for jazz chord sequences. Music Perception, 2(1), 52–77.
[15] Bhatkhande V.N. (1993) Hindustani Sangeet Paddhati (Vol 1-4): Hatras: Sangeet Karyalaya, 1993.