Proceedings of ICADSS 2026 Symposium: Contemporary Visual Arts, AI Creativity, and Interactive Media: Theory and Practice
DOI: 10.54254/2753-7064/2026.BJ38455

A Study of Artificial Intelligence-Based Emotional Music Generation Systems

Linyang Lyu

Basis Bilingual School Guangming Shenzhen, Shenzhen, China
258012lvlinyang@gmail.com

Abstract. Artificial intelligence-based affective music generation (AI-AMG) systems generate music automatically with the purpose of expressing or evoking particular feelings. This review examines studies published between 1990 and 2022 and organizes existing systems according to their main components and generation methods. A typical AI-AMG system includes target emotion identification, emotional music generation, and emotional evaluation. Rule-based, data-driven, optimization-based, and hybrid are the four categories of generation methodologies. Rule-based systems include connections between musical features and emotions, whereas data-driven systems identify structural patterns using Markov models or neural networks. Optimization methods use techniques such as genetic algorithms and graph or tree traversal, and hybrid systems combine multiple approaches. Evaluation is commonly performed through computational measures or human listener studies that assess perceived or induced emotion. Current systems can produce music associated with target emotional states, but important limitations remain, including dependence on labeled datasets, inconsistent evaluation procedures, limited measurement of induced emotion, and lower creativity than human composition. The review highlights the need for more reliable evaluation, improved emotional control, and systems that generate natural, diverse, and context-sensitive music.

Keywords: Artificial intelligence, Affective music generation, Computational creativity, Emotion evaluation, Music generation

1. Introduction

Music can make people feel good or sad [1]. Music is often used to provoke and regulate the emotions and feelings of people [2, 3]. At the same time, with the development of the computer, people have begun to build artificial intelligence (AI) systems that can create music [4, 5]. "AI-based affective music generation (AI-AMG) systems" refers to the group of computational music synthesis systems mentioned above that concentrate on producing affective music. Such frameworks possess several benefits compared to manually composed songs, including the prevention of intellectual property disputes, algorithmic techniques for integrating melodic components and styles through innovative means, and, specifically for live sound synthesis platforms, the potential to modify generated tracks based upon shifts in the audience's physiological and affective conditions or surrounding surroundings. Moreover, AI-AMG systems may be able to create an infinite number of

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7064/2026.BJ38455

unique affective musical works and compose music without any time restrictions [6]. Due to the above advantages, AI-AMG systems have attracted the attention of researchers and companies such as Google and Apple in recent years, who are actively developing creative and interactive music generation methods.

Research on AI-based emotional music generation is presented in this paper, systems are classified according to their main generation methods, representative approaches are compared, current problems are pointed out, and future research directions are proposed.

The rest of this paper will address the problems mentioned above. Section 2 describes the review methods; Section 3 introduces the background and main content of AI-based emotional music generation; and Section 4 presents the conclusion.

2. Methodology

This review identifies relevant studies through searches of three academic databases: Google Scholar, Scopus, and IEEE Xplore. Four query groups were used. The first combined terms related to affect, emotion, or mood with terms related to synthetic or artificial music generation. The second used similar emotion and generation terms together with "lead sheet." The third combined valence, emotion, affect, or mood with synthesized or artificial melody and rhythm. The fourth combined affect, emotion, or mood with generated, synthesized, or composed music, melody, or score.

During the initial screening, studies that did not address the emotional aspects of music generation were ignored, including those on pleasant synthetic music, emotion recognition from music, emotion-based automatic playlist development, and related topics. The remaining papers were examined and shortlisted, including research on algorithms or computational emotional music production systems.

To find more papers published between 1990 and 2022 that satisfied the same requirements, reference lists from the initial search were also reviewed. These studies were critically reviewed and compared in this manuscript. The history of emotional music generation systems is presented in the next section.

3. Context: AI-powered creation of sentimental music

Since artificial intelligence-based affective music generation (AI-AMG) sits at the nexus of several disciplines, it is necessary to integrate knowledge from all of these subjects in order to build a foundation that takes emotions into account in music. The first places for interaction in producing AI-based affective music—that is to say, computationally generated music that aims to arouse or evoke a specific emotion in the listener. Algorithms for generating music automatically need to be based on artificial intelligence (AI). These algorithms need to understand the rules and structure of music based on music theory and composition to produce good-sounding, realistic-sounding music. Ideas from the science of emotion can also be used to understand the emotions in music and how people feel when listening to it. In short, the cross-disciplinary knowledge that AI-AMG needs comes from all sides: computer science, mathematical modelling, signal processing, cognitive science and psychology, music theory and composition, and occasionally physiology and neuroscience.

Research on the current systems for creating AI-based emotive music examines the contributions made by the main areas that affect the fundamental Design and components of AI-AMG systems. An introduction to these systems and an outline of their components are presented in the following sections.

DOI: 10.54254/2753-7064/2026.BJ38455

Target Emotion Recognition (TER), Affective Assessment (EA), and Music Composition for Emotion (MCE). The Target Emotion Recognition (TER) module collects information from a user or a sensing device and converts it into a psychological domain format which the framework can employ. The MCE unit subsequently produces affective melodies based on the feeling data supplied by the TER module. Finally, the emotional music output (such as a musical fragment or excerpt) is assessed by the EE component, and both human and computational methods may be employed to determine how expressive it is. The next subsections display the specifics of the aforementioned components.

3.1. Recognition of target emotions

Finding or selecting a target emotion—that is, the feeling that the music seeks to arouse or convey in listeners—is the purpose of this module. Text, images, films, sensor information, sheet scores, or specific desired feelings (as illustrated by instances of emotion charts or distinct groups of moods) may be fed into the module by an individual or a recording tool. The target sentiment is subsequently produced, which serves to inspire the creation of affective melodies. The generated information from the TEI system can be displayed as coordinates on a two-dimensional or three-dimensional mood chart or as elements of a particular emotional framework.

Both [happy and sadness] and [happiness, sadness, serenity, and rage] are often used sets of distinct emotions. Emotions that differ greatly from one another are included in these distinct emotion sets. A ring model of emotions, such as the one put forth by James Russell to display target emotions in the valence-arousal plane, is most frequently utilized in emotion maps [7].

The application of affective composition affects the selection of both the input and output datasets for the TEI module. A special TEI module can be used to create a quiet soundscape for a gaming environment based on video materials from a game. Many other frameworks only need to add explicit emotions or textual emojis to the TEI module based on their own use cases. Different kinds of emotions will be produced by different applications. For example, the TEI component can set a specific set of desired output emotions, such as happiness and relaxation, to create emotional music for people who are sad.

TEI modules generally use several kinds of feature extraction and classification methods to map the input data (text, video, etc.) into an emotional space. Many algorithms are used to project visual or video content in an emotional way for various applications. For example, a Convolutional Neural Network (CNN) structure has been used to map the input image into a seven-emotion discrete space by extracting emotional information from the visual content [8]. In some cases, the TEI module does not perform any algorithm processing because the user or creator manually provides one or more emotions as input. To create emotional sound, the TEI module can serve as an intermediary and directly pass the received emotions to the emotional music generation module.

3.2. Techniques and procedures for creating emotional music

The first goal of the AMG system is to create moving music that can touch the hearts and minds of the people. The specific goal emotion is used as the primary input for this processor, and the resulting product is an audio track, snippet or persistent musical flow that can express the desired sentiment. Musical content is generally produced by the AMG module as a sequence of sonic events, and each instance contains multiple auditory features at different structural levels (such as rhythm, duration, chord progression, pitch contour, harmony, tone colour, etc.) [9]. That is to say, the basic musical units of pitch and interval serve as building blocks for more abstract entities such as motifs

DOI: 10.54254/2753-7064/2026.BJ38455

and sections in complex harmonic structures, and these entities are somewhat interdependent [9]. In short, the AMG module uses certain techniques and algorithms to modify these features and their relationships at different scales to create affective soundscapes when given a particular emotional state as a trigger.

The four main categories of algorithms/techniques for classifying AMGs are rule-oriented strategies, data-centric approaches, optimisation procedures and integrated frameworks. Many Algorithms and methods have been proposed for generating good music. The above methods are presented below.

3.3. Rules are used for

To organise the connections among all parts of sound in music synthesis, rule-based methods provide musical rules. The AMG module shows musical concepts using mathematical formulas, logical propositions or heuristic logic (if-then statements). The AMG module is a framework that can generate emotional music to make people feel or express emotions. The two main kinds of rules in this system are based on whether they correlate sonic features with emotions or regulate one sonic feature based on another. The mapping of musical parameters to emotional coordinates in the emotion space is presented by the feature-to-emotion rule subset. Wallis and others have built a system to explore how timbre, clarity and roughness are related under different arousal conditions (emotional dimensions) [10, 11]. Recently, researchers have developed more general ways to express the principles of valence and arousal (emotional dimensions) through parametric expressions by quantifying musical elements [12]. Regulations of features also show how different clusters of attributes are connected in the musical hierarchy. A typical case is "A protocol for selecting chords from a specific set that match a given melody". In an AI-AMG architecture, such ideas are expressed as mathematical expressions, conditional logic or heuristic rules. Many of these guidelines are combined to build knowledge-based AMG modules that, in turn, generate emotionally moving music based on that information. The above rule-based methods for music creation are governed by a set of mathematical formulas and are relatively simple in terms of calculation and time.

3.4. A way of using data

Data-driven AMG techniques use data to learn the structure of music directly. Train and save the patterns as model parameters, then use the model to generate new music. One is to create a desired emotion and then write expressive music that evokes it. The two types of data-driven methods are, according to the kind of model, (1) neural network-based approaches and (2) Markov model-based techniques. Syntactic and statistical information on the distribution of musical features in a model are obtained to predict feature values that are suitable for a particular piece of music using a Markov model-based algorithm. Researchers have used this way to predict features of a particular piece of music, such as note values, chord progressions and octave shifts/register changes [13]. Several high-performing deep neural network models have also been used to discover the intrinsic relationships in the music data. These models will be used later in the generation stage to create new, pleasant and moving music effectively. Emotional music is also created by some other popular neural network designs, such as seq2seq, Variational Autoencoders (VAE), Long Short-Term Memory (LSTM), etc. Training models needs a large amount of labelled data; however, these data-driven architectures are very attractive for creating emotional music due to their unified design, ease of access and operation, and high accuracy.

DOI: 10.54254/2753-7064/2026.BJ38455

3.5. Method of optimization

Some optimisation strategies for generating emotional music automatically have also been proposed. The two general categories of the optimisation methods are (1) tree/graph traversal optimisation methods and (2) evolutionary algorithms. AMG is a type of genetic optimisation that generates moving music based on cues from nature. The above way can be used to optimise cost functions with and without constraints. Articulating musical norms as constraints in the creation of affective sound involves viewing it as a multi-target restricted optimisation problem. Using a genetic algorithm, find the optimal set of auditory features that meet the above requirements (musical principles). Scirea and others have used genetic algorithms to generate melodies that meet some of the musical theory requirements for harmonic sequences [14]. As auxiliary refinement strategies for path exploration, AI-AMG frameworks also employ (a) tree-based navigation and (b) graph-based exploration. The above methods find trajectories that reduce a specific cost function in a graph or hierarchy. A graph/tree framework is used to generate affective sound by having diverse musical state configurations as vertices and assigning weights to the edges based on their importance or closeness to a particular context or sentiment. Kuo and others have built a tree structure for harmonic choice [15], where the distance between nodes is the degree of emotion, and each vertex is a particular chord. Select the closest chord in the progression according to a given valence or mood, and then minimise the total distance of the trajectory. Chord selection for emotional music has also been achieved by graph traversal-based optimisation methods [14].

3.6. The hybrid approach

Several scholars have combined the above approaches to create expressive melodies rather than using them individually. Many methods have been used to build the AMG module for predicting or selecting specific sound effects based on emotion, such as traversal techniques, genetic algorithms, rules and neural networks. A piece is subsequently generated utilizing these identified attributes. Kuo and his group have integrated graph-based traversal, evolutionary algorithms and rule-based methods to select the next chord in a sequence, and have produced tracks with suitable tonality and texture for a particular atmosphere or emotion [15].

The above AI-AMG techniques are used to create music that can express the desired emotions. Determine whether the artificial music can achieve the desired emotional effect on the people. The third is carried out by the emotion-assessment module.

3.7. Emotional assessment

The aim of the following section is to determine if the artificial music can inspire the desired emotions. Some ways researchers have been used to study emotions. The two types of above methods are divided into algorithm-based evaluation and human-based evaluation. Algorithm-based methods compare some features of the generated music with the reference template and output quantitative indicators. The above indicators are used to evaluate the degree to which the system produces music with a specific emotion. One is to show the chord position and where it falls in the "Circle of Fifth Emotions", and thus also express emotions. The created music's emotional characteristics are ascertained using the resulting index. Algorithm-based evaluation methods are quicker and simpler to use than human evaluation, but they cannot directly detect human emotions. Additionally, they are typically unable to assess the creativity of the music—that is, the system cannot demonstrate emotions in a novel way through repetition—or determine whether the music

DOI: 10.54254/2753-7064/2026.BJ38455

has a "good structure" and sounds natural. At the same time, the other type of evaluation is based on how listeners perceive the emotions in music. Emotion is a feature of people; thus, a person can directly feel or be moved by emotions. The listeners' emotions are also relatively unstable. As people have different musical knowledge and tastes, many scholars have used this method. A group of human listeners are shown a few examples of music produced by the system, and they are then invited to share their thoughts or feelings on the emotions in the music. The purpose of a QuestionarQues is to collect views and self-assessment data. In general, the aforementioned two methods seek to either measure the emotions that individuals express or feel as a result of music (perceived emotions and induced emotions). Formally speaking, induced emotion is an introspective experience of psychophysiological changes that often take place during emotional self-regulation, and perceived emotion is the cognitive processing of the perception of intended or stated emotions. The guidelines for emotion evaluation state that there are two types of evaluation that can be achieved from human study: perceived emotion assessment and induced emotion assessment.

In the past, the degree and intensity of emotions experienced by people have been recorded through self-reports of pleasure and strength, or classified into categories. Induced-emotion-assessment methods also seek to identify the emotions experienced by listeners after hearing a work of art. A separate set of questionnaires will be used to investigate whether listeners have been moved or affected emotionally by the work. Other Ways to Observe and Quantify the Resulting Emotions are Physiological and Sensor-Based. To identify and measure listeners' emotional states, additional physiological signals including electroencephalography (EEG), face electromyography (EMG), and photoplethysmography (PPG) can be gathered and examined. Because it employs impartial, involuntary indications that accurately represent generated emotions, the type of physiology-based evaluation approach described above is generally regarded as being reasonably reliable. The real implementations of AI-AMG systems should guide the selection of emotion-assessment techniques, and it is important to take into account whether the system's goal is to elicit or convey emotions in listeners.

Even tho the existing AI-AMG systems can produce musical works that express the desired target emotions, there is still much work to be done. The emotions that listeners feel, however, have not been studied by the majority of researchers, and their computational creativity is still below that of human ingenuity. The two are as follows: It begins by outlining the existing SOTA techniques for AI-based emotive music production. Secondly, it can assist the reader in locating pertinent research and guiding them to further reading.

4. Conclusion

This paper provides an overview of AI-based emotional music generation (AI-AMG) systems. These systems consist of three parts: emotional evaluation, emotional music production, and target emotion identification. How to Build Each Part is Also Addressed. The existing generation methods fall into four categories: rule-based, data-driven, optimisation-based and hybrid, and pay attention to the musical features that control valence and arousal. Also in the review, it will compare computational and human-based evaluation methods for perceived and induced emotions. Although the current system can produce music according to the desired mood, many problems still need to be solved, such as the quality of the dataset, inconsistency in evaluation, lack of emotional reliability, musical unnaturalness and computational creativity. Through the above methods, components and shortcomings have been identified in a single system; thus, it provides a foundation for understanding current AI-AMG systems and for future research on more expressive, diverse, and reliable emotional music generation.

DOI: 10.54254/2753-7064/2026.BJ38455

References

[1] Patrik, N. J., & Sloboda, J. (2011). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press.

[2] Agres, K. R., Schaefer, R. S., Volk, A., van Hooren, S., Holzapfel, A., Dalla Bella, S., Müller, M., de Witte, M., Herremans, D., Ramirez Melendez, R., Neerincx, M., Ruiz, S., Meredith, D., Dimitriadis, T., & Magee, W. L. (2021). Music, computing, and health: A roadmap for the current and future roles of music technology for healthcare and well-being. *Music & Science, 4*, 2059204321997709.

[3] Clements-Cortés, A. (2004). The use of music in facilitating emotional expression in the terminally ill. *American Journal of Hospice and Palliative Medicine, 21*(4), 255–260.

[4] Briot, J.-P., & Pachet, F. (2020). Deep learning for music generation: Challenges and directions. *Neural Computing and Applications, 32*(4), 981–993.

[5] Herremans, D., Chuan, C.-H., & Chew, E. (2017). A functional taxonomy of music generation systems. *ACM Computing Surveys, 50*(5), 1–30.

[6] Williams, D., Kirke, A., Miranda, E., Daly, I., Hwang, F., Weaver, J., & Nasuto, S. (2017). Affective calibration of musical feature sets in an emotionally intelligent music composition system. *ACM Transactions on Applied Perception, 14*(3), 1–13.

[7] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161.

[8] Madhok, R., Goel, S., & Garg, S. (2018). SentiMozart: Music generation based on emotions. In *Proceedings of the International Conference on Agents and Artificial Intelligence (ICAART), Vol. 2* (pp. 501–506).

[9] McFee, B., Nieto, O., Farbood, M. M., & Bello, J. P. (2017). Evaluating hierarchical structure in music annotations. *Frontiers in Psychology, 8*, 1337.

[10] Wallis, I., Ingalls, T., Campana, E., & Goodman, J. (2011). A rule-based generative music system controlled by desired valence and arousal. In *Proceedings of the 8th International Sound and Music Computing Conference (SMC)* (pp. 156–157).

[11] Wallis, I., Ingalls, T., & Campana, E. (2008). Computer-generating emotional music: The design of an affective music algorithm. In *Proceedings of the International Conference on Digital Audio Effects (DAFx-08)*, Espoo, Finland, 7–12.

[12] Ehrlich, S. K., Agres, K. R., Guan, C., & Cheng, G. (2019). A closed-loop, music-based brain-computer interface for emotion mediation. *PLOS ONE, 14*(3), e0213516.

[13] Ramanto, A. S., & Maulidevi, N. U. (2017). Markov chain based procedural music generator with user chosen mood compatibility. *International Journal of Asia Digital Art and Design Association, 21*(1), 19–24.

[14] Scirea, M., Togelius, J., Eklund, P., & Risi, S. (2017). Affective evolutionary music composition with MetaCompose. *Genetic Programming and Evolvable Machines, 18*(4), 433–465.

[15] Kuo, P.-H., Li, T.-H. S., Ho, Y.-F., & Lin, C.-J. (2015). Development of an automatic emotional music accompaniment system by fuzzy logic and adaptive partition evolutionary genetic algorithm. *IEEE Access, 3*, 815–824.