Real-Time Emotion-Driven Music Engine

Alex Rodríguez López, António Pedro Oliveira, and Amílcar Cardoso

Centre for Informatics and Systems, University of Coimbra, Portugal
lopez@student.dei.uc.pt, apsimoes@student.dei.uc.pt, amilcar@dei.uc.pt

Abstract. Emotion-Driven Music Engine (EDME) is a computer system that intends to produce music expressing a desired emotion. This paper presents a real-time version of EDME, which turns it into a standalone application. A real-time music production engine, governed by a multi-agent system, responds to changes of emotions and selects the more suitable pieces from an existing music base to form song-like structures, through transformations and sequencing of music fragments. The music base is composed by fragments classified in two emotional dimensions: valence and arousal. The system has a graphic interface that provides a front-end that makes it usable in experimental contexts of different scientific disciplines. Alternatively, it can be used as an autonomous source of music for emotion-aware systems.

1 Introduction

Adequate expression of emotions is a key factor in the efficacy of creative activities [16]. A system capable of producing music expressing a desired emotion can be used to influence the emotional experience of the target audience. Emotion-Driven Music Engine (EDME) was developed with the objective of having such a capability. The high modularity and parameterization of EDME allows it to be customized for different scenarios and integrated into other systems.

EDME can be controlled by the user or used in an autonomous way, depending on the origin of the input source (an emotional description). A musician can use our system as a tool to assist the process of composition. Automatic soundtracks can be generated for other systems capable of making an emotional evaluation of the current context (i.e., computer-games and interactive media, where the music needs to change quickly to adapt to an ever-changing context). The input can be fed from ambient intelligence systems. Sensing the environment allows the use in installations where music reacts to the public. In a healthcare context, self-report measures or physiological sensors can be used to generate music that reacts to the state of the patient.

The next section makes a review of related work. Section 3 presents our computer system. Section 4 draws some conclusions and highlights directions for further work.

2 Related Work

The developed system is grounded on research made in the areas of computer science and music psychology.

Systems that control the emotional impact of musical features usually work through the segmentation, selection, transformation and sequencing of musical pieces. These systems modify emotionally-relevant structural and performative aspects of music [4, 11, 22], by using pre-composed musical scores [11] or by making musical compositions [3, 10, 21].

Most of these systems are grounded on empirical data obtained from works of psychology [8, 19]. Scherer and Zentner [18] established parameters of influence for the experienced emotion. Meyer [13] analyzed structural characteristics of music and its relation with emotional meaning in music. Some works have tried to measure emotions expressed by music and to identify the effect of musical features on emotions [8, 19]. From these, relations can be established between emotions and musical features [11].

3 System

EDME works by combining short MIDI segments into a seamless music stream that expresses the emotion given as input. When the input changes, the system reacts and smoothly fades to music expressing the new emotion.

There are two stages (Fig. 1). At the off-line stage, pre-composed music is segmented and classified to build a music base (Section 3.1); this makes system ready for the real-time stage, which deals with selection, transformation, sequencing and synthesis (Section 3.2). The user interface lets the user select in different ways the emotion to be expressed by music. Integration with other systems is possible by using different sources as the input (Section 3.3).

3.1 Off-line stage

Pre-composed MIDI music (composed on purpose, or compiled as needed) is input to a segmentation module. An adaptation of LBDM [2] is used to attribute weights according to the importance and degree of proximity and change of five features: pitch, rhythm, silence, loudness and instrumentation. Segmentation consists in a process of discovery of fragments, by looking to each note onset with the higher weights. Fragments that result are input to a feature extraction module. These musical features are used by a classification module that grades the fragments in two emotional dimensions: valence and arousal (pleasure and activation). Classification is done with the help of a knowledge base implemented as two regression models that consist of weighted relations between each emotional dimension and music features [14]. Regression models are used to calculate the values of each emotional dimension through a weighted sum of the features obtained by the module of features extraction. MIDI music emotionally classified is then stored in a music base.

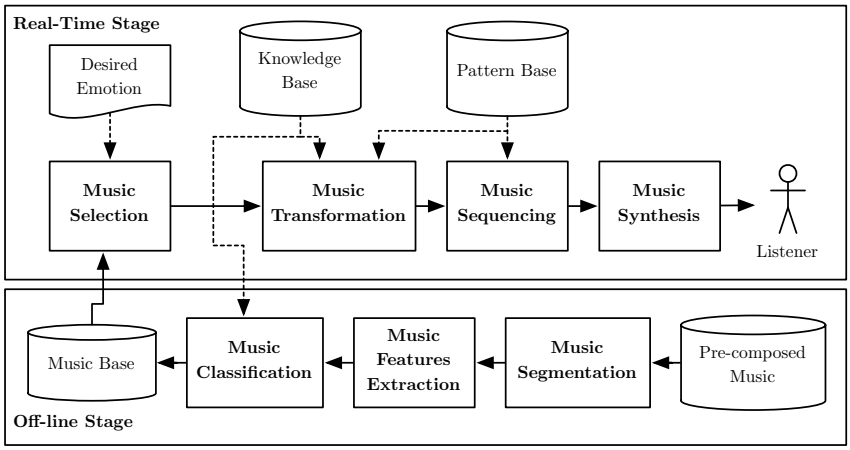

Fig. 1. The system works in two stages.

3.2 Real-Time Stage

Real-time operation is handled by a multi-agent system, where agents with different responsabilities cooperate in simultaneous tasks to achieve the goal of generating music expressing desired emotions. Three agents are used: an input agent, which handles commands between other agents and user interface; a sequencer agent, that selects and packs fragments to form songs; and a synthesizer agent, which deals with the selection of sounds to convert the MIDI output from the sequencer agent into audio.

In this stage, the sequencer agent has important responsabilities. This agent selects music fragments with the emotional content closer to the desired emotion. It uses a pattern-based approach to construct songs with the selected fragments. Each pattern defines a song structure and the harmonic relations between the parts of this structure (i.e., popular song patterns like AABA). Selected fragments are arranged to match the tempo and pitch of a selected musical pattern, through transformations and sequencing. The fragments are scheduled in order to make their perception as one continuous song during each complete pattern. This agent also crossfades between patterns and when there is a change in the emotional input, in order to allow a smooth listening experience.

3.3 Emotional Input

The system can be used under user control with an interface or act autonomously with other input. The input specifies values of valence and arousal.

User Interface. The user interface serves the purpose of letting the user choose in different ways the desired emotion for the generated music. It is possible for the user to directly type the values of valence and arousal the music should have.

Other way is through a list of discrete emotion the user can choose from. It is possible to load several lists of words denoting emotions to fit different uses of the system. For example, Ekman [6] has a list of generally accepted basic emotions. Russell [17] and Mehrabian [12] both have lists which map specific emotions to dimensional values (using 2 or 3 dimensions). Juslin and Laukka [9] propose a specific list for emotions expressed by music.

Another way to choose the affective state of the music is through a graphical representation of the valence-arousal affective space, based on FeelTrace [5]: a circular space with valence dimension is in the horizontal axis and the arousal dimension in the vertical axis. The coloring follows that of Plutchik's circumplex model [15].

Other Input. EDME can stand as an autonomous source of music for other systems by taking their output as emotional input

With the growing concern on computational models of emotions and affective systems, and a demand for interfaces and systems that behave in an affective way, it is becoming frequent to adapt systems to show or perceive emotions. EmoTag [7] is an approach to automatically mark up affective information in texts, marking sentences with emotional values. Our system can serve the musical needs of such systems by taking their emotional output as the input for real-time soundtrack generation.

Sensors can serve as input too. Francisco et al. [20] presents an installation that allows people to experience and influence the emotional behavior of their system. EDME is used in this interactive installation to provide music according to values of valence and arousal.

4 Conclusion

Real-time EDME is a tool that produces music expressing desired emotions that has application in theatre, films, video-games and healthcare contexts. Currently, we have applied our system in an affective installation [20]. The real-time usage of the system by professionals of music therapy and the integration of EDME with EmoTag [7] for emotional soundtrack generation are also being analysed. The extension of EDME to an agent-based system increased its scability, which makes easier its expansion and integration with external systems. Listening tests are needed to assess the fluentness of obtained songs.

References

1. Bresin, R., Friberg, A.: Emotional Coloring of Computer Controlled Music Performance. Computer Music Journal, 24(4), pp. 44–62 (2000)
2. Cambouropoulos, E.: The Local Boundary Detection Model (LBDM) and its Application in the Study of Expressive Timing. International Computer Music Conference (2001)

3. Casella, P., Paiva, A.: Magenta: An Architecture for Real Time Automatic Composition of Background Music. International Workshop on Intelligent Virtual Agents, pp. 224–232. Springer (2001)
4. Chung, J., Vercoe, G.: The Affective Remixer: Personalized Music Arranging. Conference on Human Factors in Computing Systems, pp. 393–398. ACM Press New York (2006)
5. Cowie, R.: Feeltrace: An Instrument for Recording Perceived Emotion in Real Time. SpeechEmotion, pp. 19–24 (2000)
6. Ekman, P.: Basic Emotions. In: Dalgleish, T., Power, M. (eds.) Handbook of Cognition and Emotion. Wiley, New York (1999)
7. Francisco, V., Hervás, R.: EmoTag: Automated Mark Up of Affective Information in Texts. EUROLAN 2007 Summer School Doctoral Consortium, pp. 5–12 (2007)
8. Gabrielsson, A., Lindstrom, E.: The Influence of Musical Structure on Emotional Expression. Music and Emotion: Theory and Research, pp. 223–248. Oxford University Press (2001)
9. Juslin, P., Laukka, P.: Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. Journal of New Music Research, 33(3), pp. 217–238 (2004)
10. Kim, S., André, E.: Composing Affective Music with a Generate and Sense Approach. Flairs - Special Track on AI and Music. AAAI Press (2004)
11. Livingstone, S.R., Muhlberger, R., Brown, A.R., Loch, A.: Controlling Musical Emotionality: An Affective Computational Architecture for Influencing Musical Emotions. Digital Creativity, 18. Taylor and Francis (2007)
12. Mehrabian, A.: Basic Dimensions for a General Psychological Theory. Cambridge: OG&H Publishers (1980)
13. Meyer, L.: Emotion and Meaning in Music. University of Chicago Press (1956)
14. Oliveira, A., Cardoso, A. Modeling Affective Content of Music: A Knowledge Base Approach. Sound and Music Computing Conference (2008)
15. Plutchik, R. A general psychoevolutionary theory of emotion. Emotion: Theory, research, and experience: Vol. 1. Theories of emotion, 3-33. New York: Academic (1980)
16. Russ, S. Affect and Creativity: The Role of Affect and Play in the Creative Process. Lawrence Erlbaum Associates Inc, US. (1993)
17. Russell, J.: Measures of Emotion. Emotion: Theory, Research, and Experience, 4, pp. 83–111 (1989)
18. Scherer, K., Zentner, M.: Emotional effects of music: Production rules. Music and emotion: Theory and research, pp. 361392 (2001)
19. Schubert, E.: Measurement and Time Series Analysis of Emotion in Music. PhD thesis. University of New South Wales (1999)
20. Ventura F., Oliveira A., Cardoso, A.: An Emotion-Driven Interactive System. 14th Portuguese Conference on Artificial Intelligence, pp- 167-178 (2009)
21. Wassermann, K., Eng, K., Verschure, P., Manzolli, J.: Live soundscape composition based on synthetic emotions. IEEE Multimedia, pp. 82–90 (2003)
22. Winter, R.: Interactive Music: Compositional Techniques for Communicating Different Emotional Qualities. Masters Thesis. University of York (2005)