A Computer System to Control Affective Content in Music Production

António Pedro Oliveira[1] and Amílcar Cardoso[1]

Centre for Informatics and Systems, University of Coimbra, Portugal

Abstract. Music is a ubiquitous media in our lives, used in many contexts. The possibility to select appropriate affective music can be helpful to adapt music to our emotional interest. Our work intends to design a system to control affective content in music production. This is done by taking into account a knowledge base with mappings between affective states (happiness, sadness, etc.) and music features (rhythm, melody, etc.). The knowledge base is grounded on background knowledge from Music Psychology. Our system starts with the reception of an emotional description specified by the user. Next, mappings are selected from the knowledge base, according to the emotional description. Music is retrieved from a music base (recorded sound and MIDI files) according to similarity metrics between music features (of mappings and music base). Afterward, selected music can be subject to transforming, sequencing and remixing algorithms and then played. The inclusion of third party composition software is also envisaged. To assess the system, listener emotional state can be analysed using psychophysiological or self-report measures.

1 Introduction

Although emotions are essential to human life there is not yet a universal definition for them. According to Scherer [1] they may be conceived of as consisting of various components: cognitive appraisal, physiological activation, motor expression, behavior intentions, and subjective feeling. Emotional states can be described as particular configurations of these components. For a long time, Cognitive Sciences have been studying the foundations of emotions. More recently computational models have also been proposed. These contributions have been applied in several domains (e.g., robotics and entertainment). Music has been widely accepted as one of the languages of emotions. Nevertheless, only recently scientists have tried to quantify and explain how music influences our emotional states [9] [10].

In recent years many research areas have been working to reduce the semantic gap that exists between music features and human cognition. From this research some works [4] [2] established mappings between music features and emotions (e.g., figure 1). Our work intends to design a system to produce affective music by taking into account a knowledge base with mappings between emotions and music features. This work is being developed in two stages. Firstly, studies of the relations between emotions and musical features are examined, to try to select mappings useful in our computational context. Secondly, a computer system that uses these mappings to produce affective music is designed, implemented and assessed.

A possible application of this system will be music therapy as a way of emotional, cognitive and physical healing, as well as in the production of soundtracks for arts, movies, dance, deejaying, theater, virtual environments, computer games and other entertainment activities. The purpose of this system is the induction of an emotional experience in human listeners by using music. This work entails an interdisciplinary research involving Music Psychology, Affective Computing, Music Information Retrieval, Psychophysiology, Computer Music, among other fields. The next section makes a review of some of the most relevant contributions from Music Psychology. Section 3 presents Affective Computing works that are related with our system. Section 4 presents the proposed system, and finally section 5 makes some final remarks.

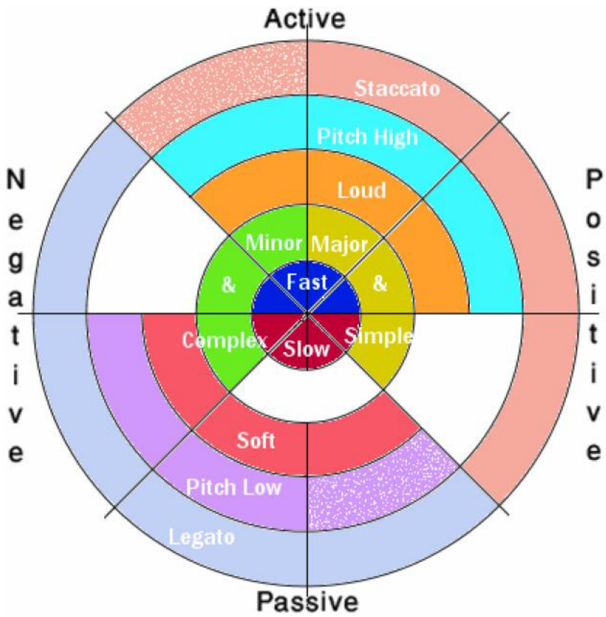

Fig. 1. Relation between musical and emotional space (from [2])

2 Music Psychology

Music Psychology is a sub-field of Psychology that intends to understand what are the musical processes involved in our brain [3]. One objective is to comprehend how music influences our behavior. In the following we review some of the most relevant works on musical content and emotion experiencing, including studies involving listener and environmental variables that apparently influence emotions induction. Relations between musical features and emotional responses are presented in this section, based on the study of works on Emotions and Music, and Music Semantics.

2.1 Emotions and Music

Schubert [4] studied the relations between emotions and musical features (melodic pitch, tempo, loudness, timbral sharpness and texture) using a 2 Dimensional Emotion Space. This study was focused on how to measure emotions expressed by music and what musical features have an effect on arousal and activation of emotions. Likewise, Korhonen [5] tried to model people perception of emotion in music. Models to estimate emotional appraisals to musical stimuli were reviewed [4] [6]. System identification techniques were applied to make linear models of the emotional appraisals. Livingstone and Brown [2] also proposed relations between music features and emotions in a 2 Dimensional Emotion Space (figure 1).

2.2 Music Semantics

Meyer [7] studied the meaning of emotions in music by bringing some gestalt principles into music. He defended that music is meaningless when it is in a style unfamiliar to us. Some principles of pattern perception in music are presented in this book. The role of music structure and shape were studied and some ideas were drawn: sound stimuli is conceived as part of a structure; a weak/bad shape leads to tension; one way to weaken shapes is the excessive similarity of them; pitch uniformity is characterized by equidistant series of tones; harmony uniformity is characterized by equal vertical intervals or unchanging harmony or repetitive progressions; formal completeness is characterized by texture. From this study expectation seems to be a relevant musical characteristic to induce emotions. Expressive variations in pitch, tempo, rhythm (rubato, vibrato, etc.), ornamentation and tonality are also important musical characteristics that are related to emotions induction. Meyer analyzed structural characteristics of important music and its relation with emotional meaning in music, which can be very useful in the process of emotions induction with music. Some ideas of Meyer were applied in Lindstrom work [8], where Lindstrom varied some musical features (melody, but also rhythm and harmony) to know how music expresses emotions to listeners.

Gabrielsson and Lindstrom [9] studied relations between happiness and sadness, and musical features. They found that: major mode is associated with grace, serenity and solemnity states; minor mode with tension, disgust and anger; staccato articulation with gaiety, energy, activity, fear and anger; legato articulation with sadness, tenderness, solemnity and softness; high loudness with joy, intensity, power, tension and anger; low loudness with sadness, softness, tenderness, solemnity and fear; high register (pitch level) with happiness, grace, serenity, dreaminess, excitement, surprise, potency, anger, fear and activity; low register (pitch level) with sadness, dignity, solemnity, boredom and pleasantness. Scherer

[10] also proposed rules to map between musical features and emotions. According to Scherer, the experienced emotion depends on different features: structural, performance, listener and contextual.

3 Affective Computing

Affective Computing is a discipline that deals with the design of devices used to process (recognize, embed and express) emotions [11]. The aim of this section is to discern common approaches and techniques used to generate music with appropriate affective content.

Starting from results of previous work [4], Livingstone and Brown [2] established relations between music features and emotions. Both emotions and a set of music-emotion structural rules were represented in a 2 Dimensional Emotion Space with an octal form (figure 1). A rule-based architecture was designed to affect the perceived emotions of music, by modifying the musical structure. Livingstone et al. made a symbolic music performance engine [12] that could adapt score reproduction with the audience emotions. This engine was grounded on a list of performative and structural features, that could be changed, and their emotional effect [13]. For instance tempo, mode and loudness are structural parameters, and tempo variation, melody accent and note accent are performative parameters.

A MIDI-based software named REMUPP was designed to study aspects of musical experience [14]. This system allows the real-time manipulation of musical parameters like tonality, mode, tempo and instrumentation. For instance, articulation is changed by altering the length of notes and register by altering the pitch of notes. This system has a music player that receives music examples and musical parameters. Music examples are composed by MIDI data and a set of properties. Musical parameters can be used to control the sequencer, to control synthesizers or to employ filters and effects on MIDI stream. Winter [15] built a real-time application to control structural factors of a composition. Models of musical communication of emotions were reviewed to get an insight of what musical features are relevant to express emotions. Pre-composed musical scores were manipulated through the application of rules. These rules have some control values for different musical features: mode, instrumentation, rhythm and harmony.

Hybrid approaches can be used to generate music based on examples of human performances [16]. This approach combined Case-Based Reasoning (CBR) and fuzzy techniques. Musical knowledge was stored in cases, which had score concepts (melody with a sequence of notes and harmony with a sequence of chords), musical analysis of the score (a tree describing metrical, tensing and relaxing relations among notes) and information about expressive performances (affective expressivity of sequences of notes). The system used fuzzy techniques in the reuse step of CBR.

DInca and Mion [17] varied a set of musical parameters to synthesize music with different affective content. They established associations between affective and sensorial categories by using Principal Component Analysis. Tempo parameters (attack, duration, note per second), intensity parameters (peak sound level, sound level range) and perception parameters (roughness and centroid) were manipulated. Tempo and intensity parameters were controlled with ADSR (Attack, Decay, Sustain, Release) envelope values and perception parameters were controlled by changing frequency and amplitude of partials. Mappings were established between an expressive control space and rendering model parameters.

Kim and Andre [18] made a music composition system to manipulate user's emotional state. To validate the emotional impact of the system on user they correlate user's impression with physiological responses (Electrocardiogram, Electromyogram (EMG), Galvanic Skin Response (GSR) and Respiration). They found that GSR is a useful indicator of unsettling-relaxing music and similarly EMG for positive-negative valence of music. A 2 Dimensional Emotion Space (valence and arousal) was used to characterize music, for instance, stimulant music is characterized with high arousal and valence.

MAgentA [19] is an agent that automatically produces real-time background music for a virtual environment. This environment has an emotional state, which is given to the agent through the perception module (cameras, etc.). The emotional state is used to select an algorithm from a database of affective composition algorithms. These algorithms try to match the emotional state to musical features (harmony, melody, tempo, etc.).

4 Proposed System

In this section an overview of the system is presented with the aid of figure 2. The input for the system is the description of the emotional experience that the system is intended to induce in the listener. Some information about the listener profile can also be given.

A knowledge base with mappings between emotions (calm, anger, sadness, happiness, fear, among others) and musical features (harmony, melody, rhythm, dynamics, tempo, texture, loudness, among others) allows the system to retrieve the more appropriate music from the music base. These mappings are to be built using background knowledge from Music Psychology and Music Information Retrieval.

The music base contains labeled music chunks that may origin from two sources. The first one comprises pre-composed music; the second (optional) is an automatic music composition module. Both can feed a segmentation module [20] [21]. The resulting music chunks are then labeled by feature extraction mechanisms. Several already developed frameworks, tools and/or algorithms for features extraction will be equated for this task [22]. Concerning the implementation of the automatic composition module, several approaches may be considered: e.g., fractals, L-Systems, statistical models and genetic algorithms. We also consider recurring to third-party solutions.

The central process of the main module (CGAM) starts with the reception of an emotional description in the form of an emotional spectrum specified by

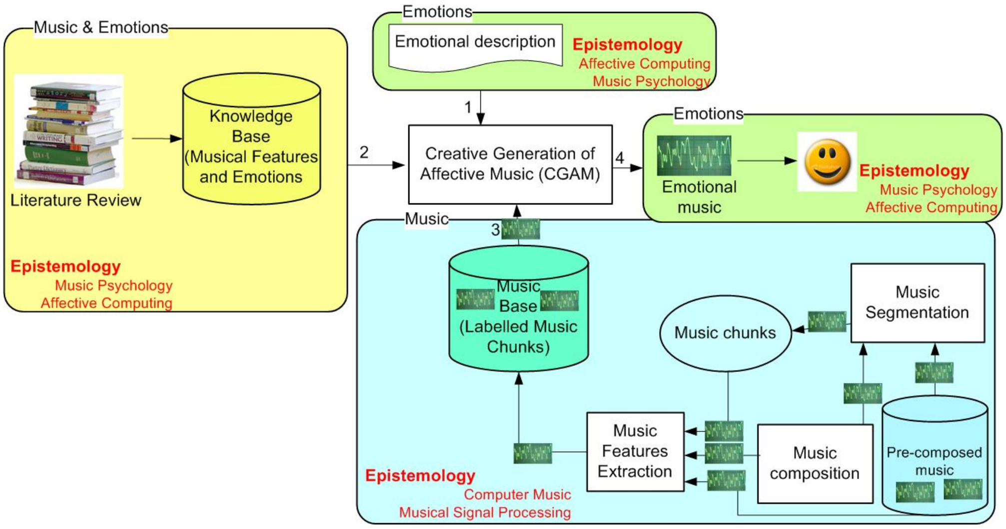

Fig. 2. System architecture

the user. Next, mappings are selected from the knowledge base, according to the emotional spectrum. Music is retrieved from the music base according to similarity metrics between music features (of mappings and music base). Case-Based and or Rule-Based Reasoning can be used to help in the retrieval process. Afterward, selected music can be subject to sequencing, remixing and transforming algorithms and then played. To assess the system, listener emotional state can be analysed using psychophysiological or self-report measures.

4.1 Architecture: a closer look

This sub-section is devoted to the description of the 4 main stages depicted in figure 2. They are based on the subsequent epistemological research topics: Music Psychology, Musical Signal Processing, Affective Computing and Computer Music. Special attention is also devoted to the description of the main module.

Emotion Description A 2 dimensional emotion space (valence and arousal) is being considered in this work like in [4] [23] [2]. The user will describe the intended emotion by selecting a region in the emotional plane. Another useful dimension that can be considered is the emotional intensity. This dimensional approach seems to be intuitive for the user, but raises some problems in the mapping of the user input to the discrete descriptions of emotions in the KB. Actually, it is not clear whether there is an obvious matching between some regions in the cartesian plane and descriptions of emotions.

Another open possibility is to let the user represent the input by a set of lexical terms that represent affective states. This last approach can be aided by an ontology of affective states [24]. These states can be emotions (e.g., sad), moods (e.g., depressed), interpersonal stances (e.g., cold), preferences (e.g., loving) and affect dispositions (e.g., anxious).

Mapping Emotions - Musical Features The Knowledge Base (KB) is constructed based on a literature review of works that studied ways to establish mappings between emotions and musical features [4] [5]. Literature usually uses lexical terms to represent musical features. Feature extraction algorithms usually represent the result as a value. Therefore, literature terms need to be mapped to (fuzzy) values in the knowledge base. For instance, a slow tempo is about 60 Beats Per Minute (bpm), while a rapid tempo can be 120 bpm.

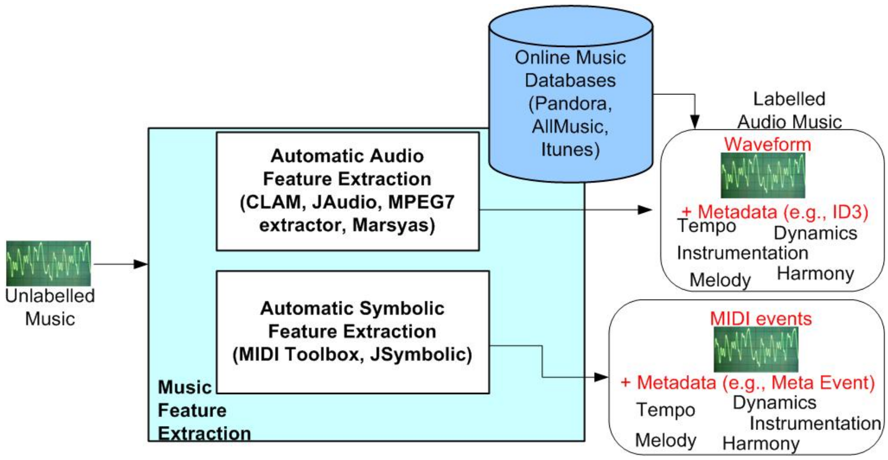

Fig. 3. Music features extraction module

Music base construction Musical Analysis and Musical Signal Processing research represent the most important sources of solutions for extracting features from music. The music features extraction module will comprise algorithms to mine features like melody, harmony, dynamics and tempo from either the music produced from the music composition module and from the database of precomposed music chunks. Music metadata that results from this module is then used to label music. Figure 3 presents this module with more detail.

A distinction between features from audio and symbolic music is established. There are some applications that can be used to extract music features from audio. Nevertheless, the automatic approach is not reliable for some music features and the analysis of large music databases is very slow. So, manual approaches like the ones used in Pandora[1] and AllMusic[2] which were done with the intervention of music experts are frequently more reliable and faster than automatic approaches.

[1] www.pandora.com/
[2] http://allmusic.com/

Figure 4 presents audio and MIDI features that can be used in our system. These features can be useful to generate playlists by using simple rules, similarly to what can be done with music metadata (artist, song, album, genre, year, etc.) in ITunes [3]

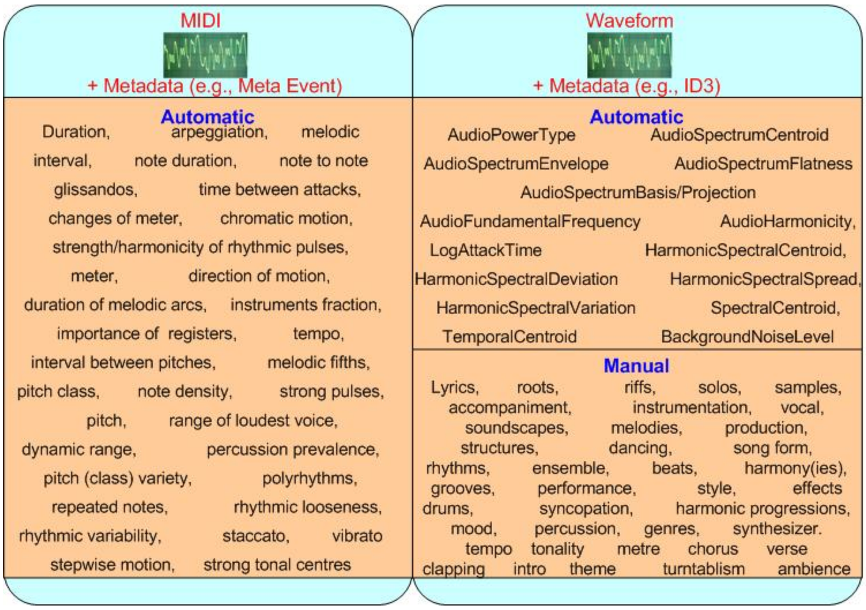

Fig. 4. Music features

Production of Affective Music The CGAM module interacts in 4 different phases (figure 5) with:

1. The input (emotion description);
2. The Knowledge Base (KB) to select the appropriate mappings to the emotion description;
3. The Music base to select the music according to the mappings selected from the KB;
4. The output to generate the music.

Now, we will concentrate our attentions in the description of the music manipulator, which is responsible for the generation, sequencing, remixing and synthesis of music. This module uses algorithms to manipulate audio and MIDI. Sound effects usually used by DJs may be used to try to manipulate music to induce specific emotions. Equalization may be used to change the frequency envelope of a sound by applying filters. For instance, high-pass filters are used by DJs to reduce low frequencies (e.g., drums) to promote moments of relaxing.

[3] http://www.apple.com/itunes/

Other effects like 3D, chorus, diffusion, distortion, echo, flanging, phasing and reverbation are also used to manipulate audio. More information about the role of these effects on the emotional response of audience can be consulted on [25].

There are many algorithms to manipulate MIDI in ways that may be used to adapt music features to the intended emotional effects over the audience. Phrase expression, chord asynchrony, melody dynamics, performative originality, ADSR envelope, humanising functions, tonality, mode, harmonic and rhythmic complexity, register, instrumentation, articulation, tempo, intensity and perception parameters (roughness) are some of the features that can be manipulated in the MIDI plane. The following works can be consulted to obtain information about algorithms used to manipulate MIDI files [13] [14] [17] [15].

The manipulation and sequencing of MP3 files and/or MIDI files is done at this stage. The way this is done is similar to the approach followed in REMUPP [14]. Our work considers music examples represented in Standard MIDI Format (MID/MIDI) and MP3 format. These examples are selected from the music base in the previous stage. Musical parameters are selected from the knowledge base during the knowledge selection stage. Some algorithms for music manipulation of music properties (e.g., sound effects and synthesis) and the way music examples are sequenced will be implemented/adapted. Figure 6 presents the main processes involved in this module (it is similar to the architecture of REMUPPs music player).

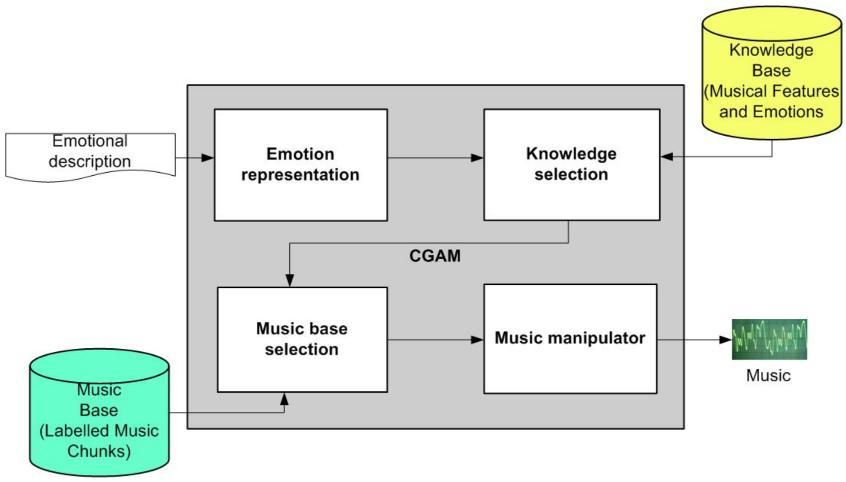

Fig. 5. Interactions of the main module

Emotion recognition Human listeners are stimulated with music that is appropriate to the emotional description given in stage 1. Emotions recognition will be done by using qualitative and quantitative methods. Quantitative methods will be based on acquisition and analysis of psychophysiological data from listeners, like the changes in heart rate, blood pressure, skin conductance, temperature and respiration. These methods are appropriate to verify induced emotions. Qualitative methods will be based on self-report measures (like in [4]), which are appropriate to verify perceived emotions. For more details about the expression, perception and induction of musical emotions, Juslin and Laukka article [26] shall be consulted.

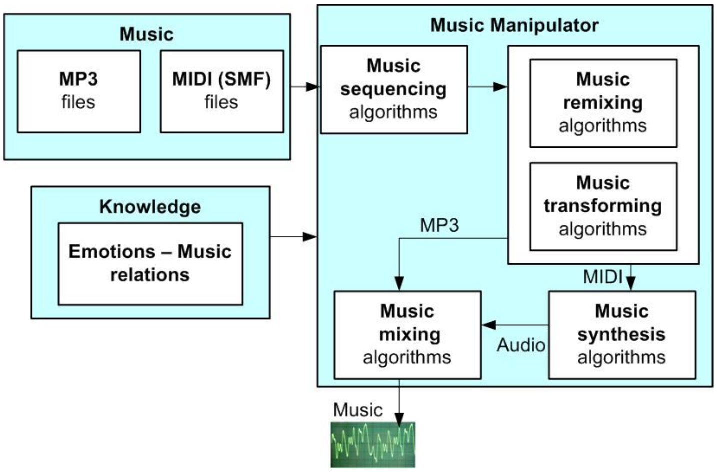

Fig. 6. Music manipulator

4.2 Assessment methodology

Figure 7 presents the assessment methodology to be used in this system. Emotional playlists are generated according to the emotional description input. Then, the audience listens to these playlists. Later, emotions are recognized through the analysis of psychophysiological and self-report data of the audience. Finally, comparisons are made between recognized and intended emotions. These comparisons will be used to refine the mappings in the KB. New mappings can be added, old mappings can be adapted or discarded. Both Case-Based and Rule-Based representations are known to be adequate for this kind of operation.

Several sessions are planned to assess our system. For each session we need to define the number of participants and what affective states we intend to evoke. To decrease the effect of exogenous variables each participant rests a period of time before the experiment. Then, each participant is subjected to a number of musical stimuli that induces a specific affective state (e.g., happiness). Later, other musical stimuli are used to induce (sequentially) other affective states (sadness, tenderness, etc.). For these stimuli, both audio and MIDI music will be used. Playlists that will be used can comprise music chunks that last more than 20 seconds and less than 10 minutes. These music chunks will be automatically taken from the music base. The audience is not limited to any age group. However, due to the existent receptiveness to the experiments, we are planning to do some tests with people in stages of adolescence or early adulthood.

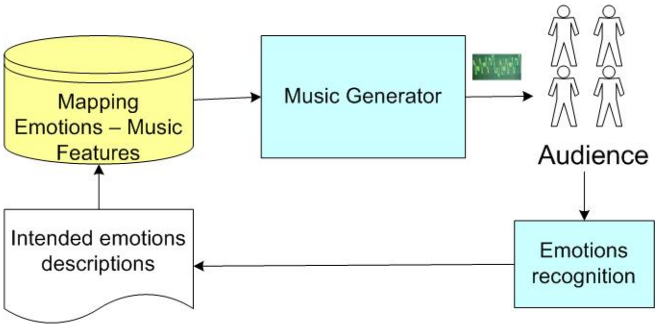

Fig. 7. Assessment methodology

5 Conclusion

This research reviews affective computing and music psychology works relevant to the design of a computational model of automatic music production according to an emotional description. We intend to develop a computational systematization of the relations between emotions and music, which can contribute to a high affective control in the selection and transformation of both structural (e.g., harmonic mode) and performing features (e.g., beat accent).

We intend to tune our system with users to promote a reliable induction and expression of emotions by using music. This way, our system can be applied in areas that intend to produce music given an emotional input. It may assume the role of a tunable support tool for scientific research in Music Psychology. It may be applied in the production of soundtracks for arts, movies, dance, theater, virtual environments, computer games and other entertainment activities. It may also be useful in Music Therapy as a way of emotional, cognitive and physical healing. Musicians can also benefit from this system as an affective music production tool or as an autonomous affective DJ-like application.

References

1. Scherer, K.: On the nature and function of emotion: A component process approach. In: Approaches to emotion. Lawrence Erlbaum (1984) 293–317
2. Livingstone, S., Brown, A.: Dynamic response: real-time adaptation for music emotion. In: Australasian Conference On Interactive Entertainment. Volume 2., Creativity & Cognition Studios Press Sydney, Australia, Australia (2005) 105–111
3. Deutsch, D.: The Psychology Of Music. Academic Press (1982)
4. Schubert, E.: Measurement and Time Series Analysis of Emotion in Music. PhD thesis, University of New South Wales (1999)
5. Korhonen, M.: Modeling continuous emotional appraisals of music using system identification. Master's thesis, University of Waterloo (2004)
6. Li, T., Ogihara, M.: Detecting emotion in music. In: International Conference on Music Information Retrieval (ISMIR). Volume 4. (2003) 239–240

7. Meyer, L.: Emotion and Meaning in Music. University of Chicago Press (1956)
8. Lindstrom, e.: A Dynamic View of Melodic Organization and Performance. PhD thesis, Acta Universitatis Upsaliensis Uppsala (2004)
9. Gabrielsson, A., Lindström, E.: The Influence Of Musical Structure On Emotional Expression. In: Music and emotion: Theory and research. Oxford University Press (2001) 223–248
10. Scherer, K., Zentner, M.: Emotional effects of music: Production rules. In: Music and emotion: Theory and research. Oxford University Press (2001) 361–392
11. Picard, R.: Affective Computing. MIT Press Cambridge, MA, USA (1997)
12. Livingstone, S.R., Muhlberger, R., Brown, A.R.: Playing with affect: Music performance with awareness of score and audience. In: Australasian Computer Music Conference. (2005)
13. Livingstone, S.R., Muhlberger, R., Brown, A.R.: Influencing perceived musical emotions: The importance of performative and structural aspects in a rule system. In: Music as Human Communication: An HCSNet Workshop on the Science of Music Perception, Performance and Cognition. (2006)
14. Wingstedt, J., Liljedahl, M., Lindberg, S., Berg, J.: Remupp: An interactive tool for investigating musical properties and relations. In: New Interfaces For Musical Expression, National University of Singapore Singapore, Singapore (2005) 232–235
15. Winter, R.: Interactive music: Compositional techniques for communicating different emotional qualities. Master's thesis, University of York (2005)
16. Arcos, J., de Mantaras, R.: Combining ai techniques to perform expressive music by imitation. In: AAAI Workshop: Artificial Intelligence and Music. (2000) 41–47
17. DInca, G., Mion, L.: Expressive audio synthesis: From performances to sounds. In: International Conference on Auditory Display. (2006)
18. Kim, S., Andre, E.: Composing affective music with a generate and sense approach. In: Proceedings of Flairs 2004 - Special Track on AI and Music, AAAI Press (2004)
19. Casella, P., Paiva, A.: Magenta: An architecture for real time automatic composition of background music. In: International Workshop on Intelligent Virtual Agents (IVA '01), London, UK, Springer-Verlag (2001) 224–232
20. Tzanetakis, G., Cook, P.: Multifeature audio segmentation for browsing and annotation. In: IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. (1999) 103–106
21. Grilo, C., Cardoso, A.: Musical pattern extraction using genetic algorithms. In: Symposium on Computer Music Modeling and Retrieval, Springer (May 2003) 132–141
22. Tzanetakis, G., Cook, P.: Audio information retrieval (air) tools. In: International Symposium on Music Information Retrieval (ISMIR). Volume 1. (2000)
23. Gabrielsson, A.: Perceived emotion and felt emotion: Same or different. Musicae Scientiae, Special (2002)
24. Scherer, K.: Psychological models of emotion. In: The Neuropsychology Of Emotion. Oxford University Press (2000) 137–162
25. Barrington, L., Lyons, M.J., Diegmann, D., Abe, S.: Ambient display using musical effects. In: International Conference On Intelligent User Interfaces. Volume 11., New York, NY, USA, ACM Press (2006) 372–374
26. 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) (2004) 217–238