

# Towards Affective-Psychophysiological Foundations for Music Production

António Pedro Oliveira<sup>1</sup> and Amílcar Cardoso<sup>1</sup>

Coimbra University, Department of Informatics Engineering, Portugal

**Abstract.** This paper describes affective and psychophysiological foundations used to help to control affective content in music production. Our work includes the proposal of a knowledge base grounded on the state of the art done in areas of Music Psychology. This knowledge base has relations between affective states (happiness, sadness, etc.) and high level music features (rhythm, melody, etc.) to assist in the production of affective music. A computer system uses this knowledge base to select and transform chunks of music. The methodology underlying this system is essentially founded on Affective Computing topics. Psychophysiology measures will be used to detect listener's affective state.

## 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.

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. The next section makes a review of some of the most relevant contributions from Affective

Computing and Music Therapy. Section 3 presents our methodology. Section 4 presents our emotional representation. Section 5 shows how our affective music database will be organized. Section 6 presents how will be done the validation of the emotional output, and finally section 7 makes some final remarks.

## 2 Background

Affective computing systems need to recognize and/or express emotions. These activities are vital to an appropriate stimulation and interaction with users. They can be even more important if the system uses multiple modes of interfacing with the user. Multimodal interaction is very important because it stimulates the user through various senses.

One thing that has to be cleared up is what we understand by emotions, moods and other types of affect. We accept Scherer's suggestion [2] that there are affective states that result from the combination of 5 types of affect: emotions, moods, interpersonal stances, preferences and affect dispositions. The main differences between these states are their duration and intensity. This research like most of the works reviewed in this section focus on emotions. The next paragraphs are dedicated to the presentation of some relevant research to our work that deal with recognition and expression of emotions, particularly by using music as a stimulus. Other relevant works of Music Therapy and affective music generation are also presented.

### 2.1 Emotions Recognition

Friberg [3] designed an algorithm to analyse emotional expression in music performance and body motion. This is done by extracting musical features (tempo, sound level and articulation) and motion features (quantity of motion and gesture speed), and then by mapping them to emotions (anger, sadness and happiness) using an expression mapper (fuzzy sets). Taylor et al. [4] also worked in the analysis of emotional expression in music performance. To be more specific, a virtual character was designed to respond in real-time to the musical input. Appropriate behaviors are defined in the character to reflect the perception of the musical input. These characters are developed through the 3-layer framework ANIMUS. The first layer (perception) is responsible for the extraction of musical features (e.g., pitch, amplitude, tone and chord) from live musical input. The second layer (cognition) uses the major findings of music perception and cognition (e.g., harmonic structural rules), and Gestalt theory to organize these features. The third layer (expression) is responsible for the character animation using musical data obtained from the previous layers.

Haag et al. [5] presented a method of recognizing emotions (arousal and valence) using bio-signals. The subsequent bio-signals were used: Electromyography (EMG), electrodermal activity, skin temperature, Blood Volume Pulse (BVP), Electrocardiogram (ECG) and Respiration. Vyzas [6] used pattern recognition techniques to recognize emotional and cognitive states from physiological data.

Like Haag et al., he used BVP, EMG, electrodermal activity and Respiration but also Heart Rate (HR). Each emotion was characterized by 24 features extracted from bio-signals. In the recognition process these features were subjected to dimensionality reduction techniques.

## 2.2 Emotions Expression

Leman and Camurri [7] developed the MEGA system for real-time communication of musical expressive gestures. Computational tools for non-verbal gesture-based communication of expressiveness were developed for this system: expressiveness extraction, synthesis of expressiveness and multimedia generation (e.g. audio and animation). This system envisioned the transmission of expressive parameters between several modalities: dance, music, audio, computer animation and robots. MEGASE (MEGA System Environment) is organized into 3 layers: low-level features (e.g. music signal and movements), non-verbal gesture representations and high-level symbolic descriptions.

Leman et al. [8] studied expressiveness in audio music. Expressiveness was recognized by extracting the following features: prominence, roughness, loudness, articulation, brightness, onset and tempo in musical audio. These features were mapped to a 3 dimensional affect space (valence, activity and interest). Castellano [9] studied the relationship between emotions induced by music and movement. This work was focused on the Component Process Model of emotion [10], namely motor activation as a component of emotional process. This work supports the idea that there is a close connection between emotions induced by music and motor activation (movement).

## 2.3 Music Therapy

Chiu and Kumar [11] reviewed works done on music therapy, in particular to find what physiologic mechanisms are affected by music. Effects on body of 4 music parameters were analysed. Tempos around 60 bpm (similar to heart rate) have a sooth effect, faster tempos evoke tension and slower tempos evoke suspense. High pitches induce tension and low pitches elicit tension. High volume can cause pain and low volume contributes to relaxation. Rhythm is seen as very important to assist the regulation of our body on: circadian rhythms of temperature and sleep, ultradian rhythms of autonomic system, metabolic processes, cerebral dominance, rhythms of respiration, peristalsis and heart rate. This study supported the idea that music listening is widely accepted as being helpful for music therapy.

Hilliard [12] reviewed 11 empirical studies of music therapy (MT) and its emergent role in hospice and palliative care. From this review it was found that the following variables are positively influenced by MT: pain, physical comfort, fatigue and energy, anxiety and relaxation, time and duration of treatment, mood, spirituality and quality of life.

Erkkila et al. [13] made a music analysis system to analyse improvisations in clinical music therapy. This work intends to develop models (statistical or

neural nets) to estimate perceived musical qualities (activity, valence, tension, salience, congruence, etc.) given extracted musical features (texture, articulation, melody, rhythm, harmony, tonality and interaction between features). This tool and similar tools can help the musical interaction and experience shared by the client and therapist. The clinical relevance of the extracted features is still unknown.

Liljedahl et al. [14] developed 2 interactive music systems, which used REMUPP [15] to generate music to improve user's physical health. The Digiwall system uses music to help the user in the climbing activity and to foster the user to make music by physical motions. The Body Rest is a biofeedback system that encourages the control of users' physiology through visualization, muscle relaxation and other techniques. Biofeedback techniques are used to reduce stress, anxiety and harmful physiologic responses. Sensors are used to monitor and control heart rate, blood pressure, brainwave activity and respiration. The music generator used in this system uses mappings between musical parameters (tempo, instrumentation and rhythmic complexity) and physiological parameters (heart rate and heart rate variability). Body Rest system is presented with more detail by Fagerlönn [16]. It is a prototype used to induce relaxation by controlling heart rate using relaxing music. Results from this system supported the idea that heart rate is a stress indicator and that music of user's preference easily induces states of relaxation.

## 2.4 Affective Music Generation

Nakra [17] developed a system to synthesize music after analysing physiological and expressive gesture data using a wearable device. Sensors were used to measure muscle tension, respiration, heart rate, galvanic skin response (GSR) and temperature. Gestural and physiological data is used by musical software to map performer's gestures and breathing signals to real-time expressive effects by defining musical features (beats, tempo, articulation, dynamics and note length) in a musical score. In the same direction, Nakra and Picard [18] used this wearable system to monitor physiological data subjacent to emotional expressions of music performances.

Healey et al. [19] developed an interface of a wearable computer that perceives and responds to the user's affective state. This is done by the recognition and response to signals with emotional information. An algorithm is used to help in music selection with the objective to change user's current affective state to the intended state. This algorithm compares GSR of last 30 seconds of previous song with GSR of first 30 seconds of current song. Current affective state is predicted based on user preferences and physiological variables. These variables are measured based on Electromyogram, Photoplethysmograph (heart rate and vasoconstriction) and Galvanic Skin Response.

MAgentA [20] is an agent that automatically produces real-time background music (virtual DJ) for a virtual environment. Music is produced using rule-based algorithmic composition. The environment has an emotional state which is used by an affective composition algorithm. This algorithm is selected from

a database of affective composition algorithms, which try to match elements of music with emotions. The emotional state is obtained using perception modules of the environment (cameras, etc.).

Chung and Vercoe [21] developed a system to generate real-time music based on intended listener's affective cues. Personal expression was analysed when listening to music, like head nodding, hand tapping, foot tapping, hand clapping, mock performing, mock conducting, dancing and other gestures. The goal of this system was to correlate musical parameters with changes in affective state. Music files are generated in real-time by music composition/production, segmentation and re-assembly of music. There is also a model to select music segments and remix them to induce an appropriate emotion. 4 Affective states (engaging, soothing, boring and annoying) are represented in a 2 Dimensional Emotion Space (valence and arousal). Musical parameters are also organized in a 2 Dimensional Space (rhythm and harmony). The analysis of listener's affective state was based on physiological data (skin arousal), physical data (foot tapping) and a questionnaire. Listener data is used to develop a probabilistic state transition model to infer the probability of changing from one affective state to another. Results from this system supports the ideas that: engaged and annoyed listeners tend to stay in the same affective state, soothed listeners tend to stay soothed but can become easily bored and engaged, and annoyed listeners tend to become engaged if induced to boredom.

### 3 Methodology

This section presents the methodology of our work: objectives, scope, proposed computational approach and architecture, and validation.

#### 3.1 Objectives

The principal objective of this research is to build a computational model of music production that may express and induce intended emotions. This objective can be split in the following two objectives:

1. Examine studies of the relations between emotions and musical features, to select mappings useful in our computational context;
2. Design, implement and assess a computer system that uses these mappings to control affective content in music production through the manipulation of structural and performance features.

Generally speaking, models of music perception and expression founded on research works from Music Perception, Music Cognition, Music Performance, Music Theory and Music Therapy were examined. With this knowledge, relations are established between musical features and emotions. Through a multi-disciplinary review we intend to employ a holistic approach that bring together scientific, technological and artistic background into a computational model. Music manipulation algorithms will be used to create arrangements of pre-composed

music with appropriate affective content. The plan of our work also includes the analysis of induced emotions.

There are also other objectives to be achieved: build a music base, compose music (instrumental mainly), extract music features (high level features from both audio and symbolic music) and segment music. Now we will present some premises used to guide our objectives.

**Music has emotional information [22]** Music is characterized by distinct music features (e.g., rhythm and melody). The systematic variation of these features is closely related to variations in listeners' emotional states. So, music affective content can be changed, through the transformation of music features, to evoke intended emotions. Relations between emotions and music features are grounded on a state of the art done in areas of Music Psychology.

**Music induces emotional states** In situations, where we are playing computer games, specifically the ones which have action, when a sudden sound/music is perceived, an emotion can be induced. Roughly speaking, after being perceived, this sound/music is transmitted to our brain nerve cells, which will communicate with our limbic system that will interact with our motor system to send signals to specific parts of our body. There are some patterns that can be detected by using psychophysiological techniques, like Electromyography, Galvanic Skin Response and Heart Rate allowing us to predict induced emotion.

**Music can be a way of healing** As it was said before, music can be therapeutic by healing your mind (cognition and emotion) and body (physical). For instance, music can be used to stimulate our body to dance, and as a result cause the motion of our body in ways that can heal it. This premise can also be supported through the use psychophysiological techniques.

Our work, when compared with works described in the previous section, intends to bring together the following ideas:

1. Build a knowledge base with relations between emotions and musical features that evolve in such a way that can be used by Music Psychologists/Music Therapists;
2. Create an autonomous DJ-like application;
3. Produce music with appropriate emotional content;
4. Establish a music production framework that can be used by musicians;
5. Use music to psychomotor therapy through the induction of appropriate emotional states.

### 3.2 Scope

Both emotions and music are of multidimensional nature. This work is focused on music content, so other dimensions are not studied. Social variables like context, human listener experience and other sociopsychological variables are not considered. Moreover, editorial and cultural metadata and songs' lyrics are not meant to be analysed. The central attention of this work is on the role of high level features (acoustic metadata) in induction of emotions.

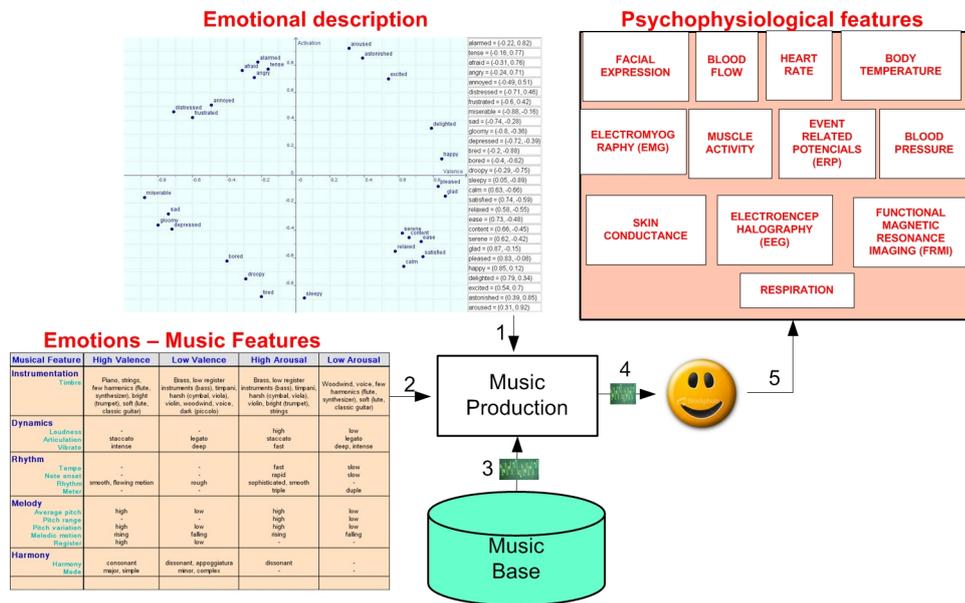


Fig. 1. Computational Approach

### 3.3 Computational Approach

Our computational approach deals with the problem of inducing emotions with music. A brief overview of our approach is presented with the aid of figure 1.

The input is the description of the emotional experience that the system is intended to induce in the listener. 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. Then, music is played and the emotional state of the listener can be analysed using psychophysiological and/or self-report measures.

**Models** A computational model used to control affective in music production will be designed. There are some studies of computational models in Music Cognition and Music Performance that can be helpful in the development of our model. Honing [23] proposed 3 criteria for model selection (goodness-of-fit, model simplicity and surprisingness) when comparing kinematic and perception-based models in a Music Cognition case study. Widmer and Goebel [24] studied computational models for the emotional expression in Music Performance.

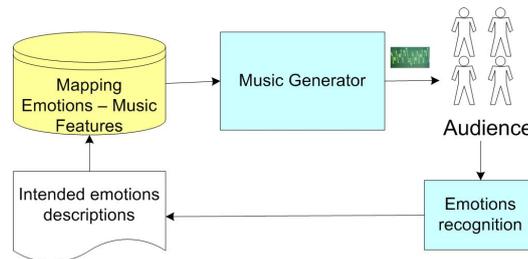
**Techniques and Algorithms** The Knowledge Base (KB) is like a white box module with mappings between emotions and musical features. This means that

all represented mappings are always visible. This option derives from the fact that this KB can be used in the future by Music Psychologists and Therapists. Both Case-Based and Rule-Based techniques are known to be adequate for this kind of representation. Semantic networks and frame/analogy-based systems can also be helpful.

Some algorithms will be developed: music transformation, music sequencing and music remixing. Others will be adapted from third party software/algorithms: music composition, music selection (energy, timbre, etc.), music mosaicing (automatic remixing) and music synthesis. All these algorithms take into account the control of music affective content.

### 3.4 Validation

Figure 2 presents the assessment methodology to be used in this system. Music with specific features is selected according to the intended emotions descriptions and then emotional playlists are generated. Musical features are previously selected from the mappings of the intended emotion. After the play of the emotional playlist, emotions are identified using psychophysiological and self-report measures. Then, there will be a comparisons between recognized and intended emotions. These comparisons will be used to refine the mappings in the KB. Techniques that were previously referred are known to be adequate for this kind of operation.



**Fig. 2.** Validation

**Sessions planning** 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 one affective state (e.g., happiness). Later, other musical stimuli are used to induce other affective states (sadness, tenderness, etc.) sequentially. 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 stage of human development. However, it will be easier to do some tests with people in stages of adolescence or early adulthood, which are almost non-musicians.

## 4 Emotions representation

A 2 dimensional emotion space (valence and arousal) is being considered in this work like in [25] and [26]. 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 rises 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 types of affect [2]. These types 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). Currently there is not a unifying model of affective states. Nevertheless, EARL [27] was recently proposed as an XML-based representation language for emotions. This proposal takes into account limitations of similar works.

Figure 3 illustrates a possible 2 dimensional affective space that can be used in our work. This space has the representation of 28 affective states. For each of these states we can see on the right side of this figure their approximate coordinates.

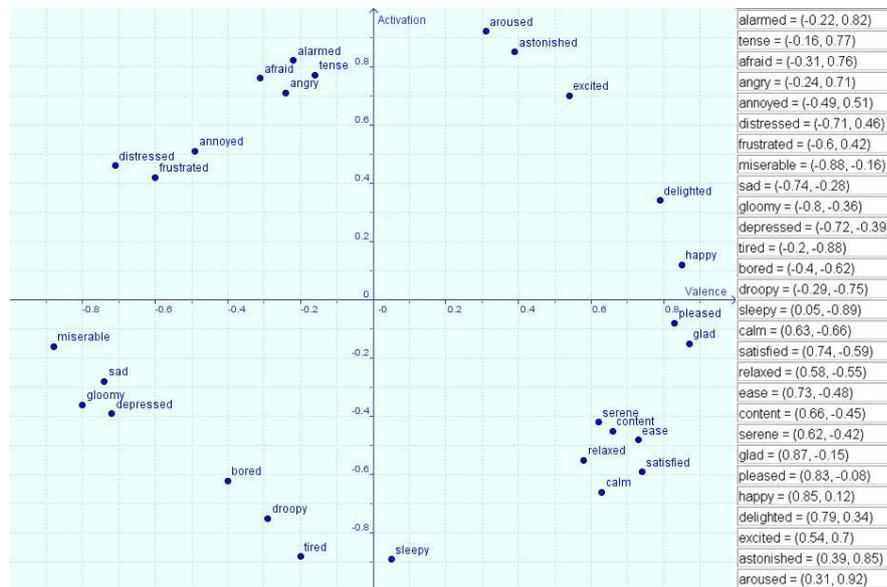
## 5 Affective Music Knowledge Base

The Knowledge Base (KB) mappings are grounded on reviewed Music Psychology research works that studied ways to establish mappings between emotions and musical features (e.g., [25] and [22]). 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.

Both audio and MIDI music will be used. Some mappings are only applied to audio or MIDI as there are some unshared features among them. Figure 4 illustrates a table with mappings established between affective states (valence and arousal) and high level musical features (instrumentation, dynamics, rhythm, melody and harmony). These mappings were established having a dimensional approach as a background. Nevertheless, we also established mappings with an approach based on categories where we used 10 affective states (sadness, happiness, anger, tenderness, fear, solemnity, depressed, dreamy, excitement and surprise).

## 6 Emotional Output Validation

Human listeners are stimulated with music that is appropriate to the emotional description given in stage 1. The system output is assessed using qualitative and



**Fig. 3.** Affective Space (adapted from Russel’s circumplex model)

quantitative criteria. To validate quantitatively the results we intend to analyse psychophysiological data from the listeners, like the changes in heart rate, blood pressure, skin conductance, temperature and respiration. Some classifier algorithms [28] can be used to help this analysis task. The qualitative validation can be done using self-report measures [25] and experimental methods from Psychology and Cognitive Science fields. Gestures and body motion are also important factors to detect emotional expression, however we do not meant to use them. Lastly, a comparison is made between the intended and the induced emotion.

## 7 Conclusion

This research reviewed affective and psychophysiological foundations relevant to a computational model that can foster the expansion of Affective Computing, namely in the domain of automatic music production according to an emotional description. We intend to foster 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). This can contribute to a systematized scientific research for psychologists and health scientists working in Music Psychology and Music Therapy.

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. The production

Musical Feature	High Valence	Low Valence	High Arousal	Low Arousal
<b>Instrumentation</b> Timbre	Piano, strings, few harmonics (flute, synthesizer), bright (trumpet), soft (lute, classic guitar)	Brass, low register instruments (bass), timpani, harsh (cymbal, viola), violin, woodwind, voice, dark (piccolo)	Brass, low register instruments (bass), timpani, harsh (cymbal, viola), violin, bright (trumpet), strings	Woodwind, voice, few harmonics (flute, synthesizer), soft (lute, classic guitar)
<b>Dynamics</b> Loudness Articulation Vibrato	- staccato intense	- legato deep	high staccato fast	low legato deep, intense
<b>Rhythm</b> Tempo Note onset Rhythm Meter	- - smooth, flowing motion -	- - rough -	fast rapid sophisticated, smooth triple	slow slow - duple
<b>Melody</b> Average pitch Pitch range Pitch variation Melodic motion Register	high - high rising high	low - low falling low	high high high rising -	low low low falling -
<b>Harmony</b> Harmony Mode	consonant major, simple	dissonant, appoggiatura minor, complex	dissonant -	- -

Fig. 4. Knowledge Base Mappings

of soundtracks for arts, movies, dance, theater, virtual environments, computer games and other entertainment activities are some examples. Another one is 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. *Approaches to emotion* (1984) 293–317
2. Scherer, K.: Psychological models of emotion. *The Neuropsychology Of Emotion* (2000) 137–162
3. Friberg, A.: A fuzzy analyzer of emotional expression in music performance and body motion. *Music and Music Science* (October 2004)
4. Taylor, R., Boulanger, P., Torres, D.: Visualizing emotion in musical performance using a virtual character. *Lecture Notes in Computer Science* **3638** (2005) 13
5. Haag, A., Goronzy, S., Schaich, P., Williams, J.: Emotion recognition using biosensors: First steps towards an automatic system. *Lecture Notes in Computer Science* **3068** (2004) 36–48
6. Vyzas, E.: Recognition of Emotional and Cognitive States Using Physiological Data. PhD thesis, Massachusetts Institute Of Technology (1999)
7. Leman, M., Camurri, A.: Musical content processing for expressive gesture applications in interactive multimedia. *Conference on Interdisciplinary Musicology* (April 2004)
8. Leman, M., Vermeulen, V., De Voogdt, L., Taelman, J., Moelants, D., Lesaffre, M.: Correlation of gestural musical audio cues and perceived expressive qualities.

- Gesture-based communication in human-computer interaction. Berlin, Heidelberg, Springer-Verlag (2003) 40–54
9. Castellano, G.: Experiments, analysis, and models of motor activation as a component of an emotional process. Master's thesis, University of Genoa (2004)
  10. Scherer, K., Zentner, M.: Emotional effects of music: Production rules. *Music and emotion: Theory and research* (2001) 361–392
  11. Chiu, P., Kumar, A.: Music therapy: Loud noise or soothing notes? *International Pediatrics* **18**(4) (2003) 204–208
  12. Hilliard, R.: Music therapy in hospice and palliative care: a review of the empirical data. *Evidence-based Complementary and Alternative Medicine* **2**(2) (2005) 173–178
  13. Erkkilä, J., Lartillot, O., Luck, G., Riikkilä, K., Toiviainen, P.: Intelligent music systems in music therapy. *Music Therapy Today* **5** (2004)
  14. Liljedahl, M., Sjömark, C., Lefford, N.: Using music to promote physical well-being via computer-mediated interaction. *Musicnetwork Open Workshop* **5** (2005) 5
  15. Wingstedt, J., Liljedahl, M., Lindberg, S., Berg, J.: Remupp: An interactive tool for investigating musical properties and relations. *New Interfaces For Musical Expression* (2005) 232–235
  16. Fagerlönn, J.: A prototype using music responding to heart rate for stress reduction. Master's thesis, Luleå University of Technology (June 2005)
  17. Nakra, T.: Inside the Conductors Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture. PhD thesis, Massachusetts Institute of Technology (1999)
  18. Nakra, T.M., Picard, R.: Analysis of affective musical expression with the conductors jacket. *Col. Musical Informatics* **12** (1998)
  19. Healey, J., Picard, R., Dabek, F.: A new affect-perceiving interface and its application to personalized music selection. *Workshop Perceptual User Interfaces* (November 1998)
  20. 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
  21. Chung, J., Vercoe, G.: The affective remixer: Personalized music arranging. *Conference on Human Factors in Computing Systems* (2006) 393–398
  22. Gabrielsson, A., Lindström, E.: The influence of musical structure on emotional expression. *Music and emotion: Theory and research* (2001) 223–248
  23. Honing, H.: Computational modeling of music cognition: A case study on model selection. *Music Perception* **23**(5) (June 2006) 365–376
  24. Widmer, G., Goebel, W., Intelligence, A., Vienna, A.: Computational models of expressive music performance: The state of the art. *Journal of New Music Research* **33**(3) (2004) 203–216
  25. Schubert, E.: Measurement and Time Series Analysis of Emotion in Music. PhD thesis, University of New South Wales (1999)
  26. Friberg, A.: pdm: An expressive sequencer with real-time control of the kth music-performance rules. *Computer Music Journal* **30**(1) (2006) 37–48
  27. Schröder, M., Pirker, H., Lamolle, M.: First suggestions for an emotion annotation and representation language. *International Conference On Language Resources And Evaluation* (May 2006) 88–92
  28. Lim, T., Loh, W., Shih, Y.: A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning* **40**(3) (2000) 203–228