Valery Vermeulen

(Paper) **The EMO-Synth, an intelligent music and image generator directed by human emotion-**

Abstract:
This paper gives an overview of the EMO-Synth project, a project focusing on generative music and image generation, biofeedback, artificial intelligence, affective computing, advanced statistical modelling and creative evolutionary systems.
The EMO-Synth is a new interactive multimedia system capable of automatically generating and manipulating sound and image to bring the user in certain predefined emotional states. During performances the emotional responses of the user are measured using biosensors that register certain psychofysiological parameters such as heartrate (ECG signals) and stress level (GSR signals).

The paper is organised in three parts. In the first part we give description of the EMO-Synth, its functioning and background. This part includes an elaboration of the functional diagram as well as a detailed description of the organisation of performances with the EMO-Synth. The second part of the paper focuses on the music generating engine that lies at the heart of the EMO-Synth. We start with a description of the genetic algorithm and other computational techniques that have been used. Subsequently an overview is provided of the three sources of music generation; digital audio streams, music generated using the MIDI protocol and finally sound generation by virtual scores directing live musicians. The third part of the paper deals with a general discussion on how generative arts can extend the human creative process as well as how the integration of human (creative) interaction can provide humanised generative systems. In this discussion the EMO-Synth will be used as a key example.

EMO-Synth performance at Center for Digital Cultures & Technologie (iMAL, Brussels, Belgium)

The EMO-Synth project was realised with the support of the Flemish Audiovisual Fund (www.vaf.be), Flanders Image (www.flandersimage.com) and Center for Digital Cultures and Technology (www.imal.org)

Topic: interactive multimedia & generative systems

Authors:
Valery Vermeulen Phd
Independent researcher
Phd in Pure Mathematics (University of Ghent, Belgium)

References:
[1] Valery Vermeulen, "*The EMO-Synth, an emotion drive music generator*", eContact 14.2!, Canadian Electroacoustic Community, 2012 (http://cec.sonus.ca/econtact/14_2/vermeulen_emosynth.html)
[2] "EMO-Synth", Artefact Magazine, 2011
[3] Marc Leman, Valery Vermeulen et al.,"Correlation of gestural musical audio cues and perceived expressive qualities ", LNAI 2915, Springer-Verlag, 2004
[4] www.emo-synth.com

Contact:
Officetamuraj@gmail.com

Keywords:
affective computing, generative music and image generation, creative evolutionary systems, genetic programming, artificial intelligence, statistical modelling and biofeedback

The EMO-Synth, an intelligent music and image generator directed by human emotion

Dr. Valery Vermeulen, PhD
www.emo-synth.com
e-mail: officetamuraj@gmail.com

Abstract

This paper gives an overview of the EMO-Synth project, a project focusing on generative music and image generation, biofeedback, artificial intelligence, affective computing, advanced statistical modelling and creative evolutionary systems.

The EMO-Synth is a new interactive multimedia system capable of automatically generating and manipulating sound and image to bring the user in certain predefined emotional states. During performances with the EMO-Synth, emotional responses of users are measured using biosensors that register certain psychophysiological responses such as heart rate (ECG signals) and stress level (GSR signals).

The paper is organised in three parts. In the first part we give a description of the EMO-Synth, its functioning and background. This part includes an elaboration of the functional diagram as well as a detailed description of performing with the EMO-Synth. A second part of the paper focuses on the music generating engine that lies at the heart of the EMO-Synth. We start with the description of three sources of music generation within the system: digital audio streams, music generated using the MIDI protocol and finally sound generation by means of virtual scores directing live musicians. Subsequently an overview is given of music generation during the training and performance phase. The used genetic algorithm and other computational techniques are briefly described as well. The third part of the paper deals with a general discussion on how generative arts can extend the human creative process as well as how the integration of human (creative) interaction can provide humanised generative systems. In this discussion the EMO-Synth will be used as a key example.

1. The EMO-Synth

1.1 Short description

Figure 1: EMO-Synth performance at iMAL (Center For Digital Cultures and Technologies, Brussels, BE, 2011, picture by Tom Van Laere)

The EMO-Synth is a new interactive multimedia system capable of automatically generating and manipulating sound and/or image to bring the user in certain predefined emotional states. (for a thorough overview and discussion on interactivity see [1]). During performances the emotional responses of the user are measured using biosensors that register certain psychophysiological parameters such as heart rate (electrocardiogram or ECG) and stress level (galvanic skin response or GSR).

1.2 Practical use

Figure 2: the EEG Trainer by Mindmedia B.V.

Using the EMO-Synth involves two phases: a learning phase and a performance phase. In the learning phase the EMO-Synth will generate auditory artefacts for the user being attached to biosensors (for the project these are the EEG Trainer and bio-trace software by Mindmedia B.V.). Using the measurements of the bodily reactions the system subsequently analyses the resulting emotional impact of the artefacts onto the user. By means of machine learning techniques and statistical modelling the EMO-Synth will learn in an adaptive way to generate those sounds and music that bring the user in certain pre-chosen emotional states. During the learning phase the user has his/her emotional feedback matched to generated sound using artificial intelligence models constructed by the EMO-Synth. After the learning phase has passed, the artificial intelligence models contain an emotional response profile for the same user. The EMO-Synth is then ready to be used as a realtime responsive multimedia performance tool in the second phase that we will call the performance phase. During these performances the models which were constructed in the learning phase are used to produce realtime personalised soundtracks for live visuals. The soundtracks involve not only digital audio but also live musicians. Visual material is hereby partially generated by the EMO-Synth and partially controlled by the same person as in the learning phase placed in front of an audience. The live audiovisual concerts that result from this experience tend to be unique and entirely based on the emotional feedback and profile of the user. During every performance the EMO-Synth is programmed to maximise the emotional impact of generated sound and image on the user.

1.3 Functional diagram

To provide the reader with a mind map we include a theoretical functional diagram of EMO-Synth and its functioning which we shortly discuss.

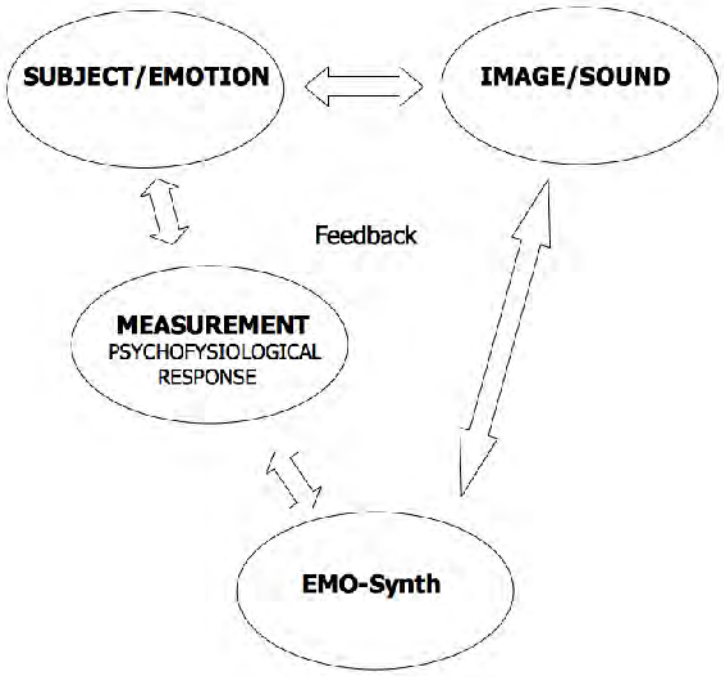

Figure 3: functional diagram of EMO-Synth

The subject is attached to biosensors that measure his or her psychophysiological reactions. Subsequently the EMO-Synth generates sound and image evoking emotional reactions with the subject. The reactions are fed into the EMO-Synth. The system learns from these reactions (in the present case this concerns an offline learning process) and starts to generate new sound and image and the cycle starts all over again. During this process a feedback loop is established between the subjects' emotional reactions and the images and sounds that are being generated. As the reader will notice in the further description, up to now image generation should be considered as a realtime manipulation of automatically generated video material. For the next generation of the EMO-Synth this image generation process will be extended to include generative algorithms as well.

1.4 Background and musical diversity

Figure 4: EMO-Synth performance at art cinema OFFoff (Ghent, 2009, BE, picture by Tom Van Laere)

Due to its interdisciplinary character the development of the EMO-Synth relies on a broad range of scientific and artistic disciplines. These include affective computing, artificial intelligence, statistical modelling and algorithmic sound and image generation. The idea and subject behind the EMO-Synth project is not new and relies on a long history of use of biofeedback for the arts. Among the first to use biofeedback we can mention the visionary artist A. Lucier with his ground breaking *Music for Performance (1965),* the pioneering work of D.Rosenboom, author of the essential *Biofeedback fort he Arts, Results and Early Experiments* ([11]) and R. Teitelbaum with his *Spacecraft* installation in 1967. From the sixties biofeedback has

been used in various artistic settings and contexts. In a lot of cases though biofeedback data is used as input data in a non-responsive sonification or visualisation process. In this way the data is directly translated into sound or image. Keeping this in mind, the EMO-Synth project provides a new dimension to this paradigm. Music and sound generation in the EMO-Synth entails much more than a simple data translation. The system really tries to understand the emotional responses of the users to sound or music and stores this information in appropriate artificial intelligence models. These artificial intelligence models are at the very heart of the music generation algorithm in the EMO-Synth. In order to allow for maximal flexibility music generation incorporates different sources: MIDI based musical data-streams, sample based audio-streams and live musicians by virtually generated scores. Due to the implementation structure of the system the music and sound that is being generated by the EMO-Synth can be immensely diverse: from tonal jazz or pop music over more textual soundtracks to real experimental avant garde music. In this way the variety in personal musical taste is embedded as much as possible into the generative system of the EMO-Synth.

2. Music and image generation

In this section we will go into detail with regard to the sound and image generation engine that is used in the EMO-Synth.

2.1 Three sources for music and sound generation

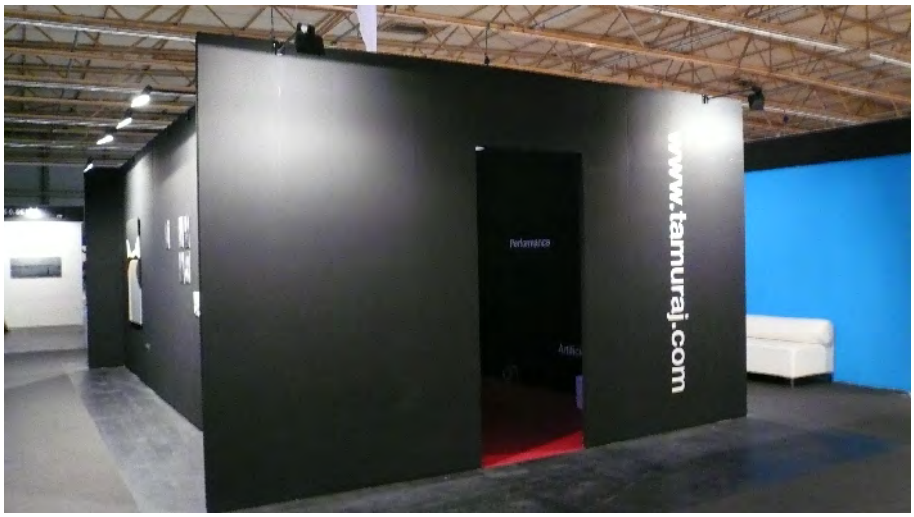

Figure 6: EMO-Synth performance at art fare Lineart (Ghent, BE, 2007)

Music generation during EMO-Synth performances includes three sources. All these sources can be combined according to the choice of the user. A first source consists of digital audio material which is organised in a database. For music generation this database can be consulted by he EMO-Synth. The audio material in this collection can completely be customised. Working with the database will moreover enable any artist to use his/her own audio clips during a performance. The second source of music/sound generation is implemented through the use of MIDI data streams. MIDI data is being generated by the EMO-Synth and sent to several MIDI channels. These

MIDI channels can comprise percussive lines as well as harmonic and tonal material. For sound generation the MIDI data is send to appropriate soft or hardware synthesisers. A third and final source of sound generation involves the use of live musicians. During performances, if desired, the EMO-Synth can generate realtime virtual scores. That way, live musicians can be directed by the system using these scores. Thus, the EMO-Synth becomes a realtime composer directing these live musicians using the scores.

2.2 Music and sound generation during training an performance phase

As already briefly mentioned using the EMO-Synth involves two stages; a learning and performance phase. During the learning phase the EMO-Synth will learn how to bring the test person into four different states of arousal, in particular: a state of low arousal, a state of low average arousal, a state of high average arousal and a state of high arousal. The motivation to work with states of arousal can be found in general emotion psychology. According to most researchers in this field human emotions can be categorised using the two dimensions valence and arousal. Valence pertains to the positive or negative effect of an emotion and arousal to the intensity (for information on these dimensions we refer to [7, 8]). As the measurement of the valence component cannot be realised using classical biofeedback devices we chose to only work with the arousal component of the emotional state in the EMO-Synth project. For the practical implementation to build and store knowledge between music generation and emotional response, statistical modelling, genetic programming (cf.[4, 5, 6]) and the approach followed in the design of creative evolutionary systems (cf. [2]) are used. After the learning phase the knowledge is stored in so called artificial intelligence models. Every model in the same vein contains a statistical musical model that is central for music generation. Different artificial intelligence models will entail a different kind of musical genres. It are these artificial intelligence models that are trained by the EMO-Synth during the learning phase. The training hereby is implemented using genetic programming. In the genetic algorithm the artificial intelligence models are the individuals in the evolutionary pools that evolve under the darwinistic rules of survival, cross over and mutation. How survival, cross over and mutation affects the artificial intelligence models is hereby programmed into the EMO-Synth. Once the learning phase has passed the EMO-Synth is ready to be used in the performance phase. During the performance phase the EMO-Synth will use its knowledge on emotional reactions of the user captured in the trained artificial intelligence models to compose appropriate sound and music.

2.3 Image generation and manipulation

Figure 7: EMO-Synth performance at Logos Foundation (Ghent, Be, 2012)

The video generation during performances of the EMO-Synth depends on the following strategy. Initially source material is chosen and cut into consecutive video clips. The clips are cut in such a way that every clip expresses one outspoken emotional state. By doing this video clip can subsequently be annotated according to its arousal level or quality; i.e. video clips with low arousal quality, video clips with low average arousal quality, video clips with high average arousal quality and video clips with high arousal quality. Next, these video clips are designated to a database inside the EMO-Synth. Once a performance starts the EMO-Synth will compose a movie using the video clips from this same database. At the same time music and sound is generated by the system that has the same effect onto the user as directed by the annotations of the video clips. In order to monitor the actual effect the audiovisual stream has onto the spectator the visuals are partially manipulated by his or her stress level. The more stressful he or she gets the more distorted the visuals will be.

3. Generative arts, human creativity and the EMO-Synth project

As a generative sound and video generating engine the EMO-Synth is an example of how connections can be established between generative arts and human creativity. By using human input generative arts can be provided with a human factor. If one establishes this link by the use of human emotion as input source new links arise between human creativity and generative arts. Moreover if emotions are used as input or during the generating process one introduces a very interesting and highly non linear factor into the work of art. A factor that is on the one hand not completely random or randomly generated but on the other hand not completely predictable or programmable. It is this form of hybrid coexistence between software and emotion as human factor that also form the basis of the EMO-Synth project.

In order to establish the link between emotion and generative art works two domains are crucial. The first is the domain of psychophysiology and biofeedback. In this field researchers are looking for ways to measure human emotional states by means of biosensors that measure certain psychophysiological parameters such as ECG (electro cardiogram or heart beat), GSR (galvanic skin response or stress) and EMG (electro myogram or muscular tension). The second domain is the promising new field of affective computing founded by R. Picard at MIT (cf. [9, 10]). Originally build on the early work of M. Clynes (cf. [3]) researchers in affective computing are looking for new directions to establish emotional man-machine interactions. As affective computing heavily relies on accurate measurement of human emotions biofeedback and psychophysiology are widely used. During the last decade psychophysiological measurement devices became more and more affordable. As a result, research in affective computing is spreading from very specialised laboratories to worldwide teams or individuals looking for various applications.

As to the EMO-Synth project one might say that affective computing has been crucial for its development. I personally treated it as the perfect way to practically develop my own methodology of connecting generative audiovisual systems and human creativity. The crux of this matter is my belief that every artwork has a deeply embedded ability or inability to communicate by means of human emotion. This belief is and was also motivated by my work both in mathematics and music. Far too often art and science are put in juxtaposition. It is taken for granted that every scientist, and mathematician in particular, acts in an utmost rational way. Just like the work of the scientist would be only related to a smart and complex network of programmed routines on a digital system. However, no matter how clever networks are implemented they are and remain subject to the world of the rigid binary first order logic. It is in the light of this naive approach that generative arts are too often conceived as products of logically ordered routines and subroutines living in a digital programmed system. Such a point of view leads more than desirable to a discussion on the artistic position generative arts need to take. To this end, a common crucial question relates to the intrinsic value of an art work as soon as everything is programmed in advance and artefacts are produced by generative systems. On the other hand, the methodology of the artist is always doing its part. It draws upon his/her personal creativity and imagination and forms the essential part of the process towards an artistic result or product. In classical terms every artistic creation process is supposed to be non rational, unpredictable and highly non linear. As if artists hold the key to a primal creative alter ego that has to be mastered.

Having worked in both mathematical research and as a musician, I experienced time and again that the boundary between art and science is nothing but artificial. Thanks to my work I have gained the insight that both disciplines are manifestations of one and the same mental activity or state. To my believe, this state is closely related to the core of what is being considered as artistic creativity. The EMO-Synth project presents new ways to integrate this way of thinking into a concrete generative multimedia system that stresses both the rational and irrational part of human behaviour and it hereby provides a new dimension to the paradigm of generative production.

Figure 8: EMO-Synth performance at the Royal Conservatory of Ghent (Ghent, BE, 2010, picture by Tom Van Laere)

Bearing the previous discussion in mind, the two following basic principles inspired by the EMO-Synth project and related to the creative artistic process can be formulated:

- The integration of biofeedback and emotions in generative arts offer new ways for creating virtual tools by which an artist can extend or complement his or her own creative process. Generative arts can accordingly function as inspirational source and moreover provide an understanding of the creative process on a whole new level both by artist and audience.

- By using systems like the EMO-Synth a new platform arises in which boundaries between artist, audience and the generated artefact are redefined. To what extend are the artefacts generated by the system creations of the artist, programmer, author or machine? To what extend can we speak of artistic input once a generative system like the EMO-Synth participates in the creative process. To what level are these artefacts a creation of the audience, the machine or the artist-programmer? Who is the author of what is being generated or is there even an author?

From this discussion it is clear that integrating emotions into generative artworks is a tremendous promising field when it comes to including human artistic creativity Creativity in this context can therefore be found on several levels: on the level of the audience, on the level of the machine or on the level of the artist.

3. References

[1] Beyls, P. (2003). *Interactivity in context/Interactiviteit in context*. KASK Cahiers 4.

[2] Corne, D. & W, Bentley P.J. (Eds.) (2002). *Creative Evolutionary Systems*. Morgan Kaufmann Publishers.

[3] Clynes, M. (1980). *The communication of emotion: the theory of scentics*. Academic Press: New York.

[4] Goldberg, D.E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.

[5] Koza, J.R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. The MIT Press.

[6] Koza, J.R. (1994) *Genetic Programming II: Automatic Discovery of Reusable Programs*. The MIT Press.

[7] Leman, M. & Vermeulen V. et al (2005). Prediction of Musical Affect Attribution Using a Combination of Structural Cues Extracted from Musical Audio. In *Journal of New Music Research, 34, no 1*, pp37-67.

[8] Leman, M., Vermeulen, V. et al (2004). Correlation of gestural musical audio cues and perceived expressive qualities. In Camurrin A. & Volpe G. (Eds.), *Gesture-Based Communiation in Human-Computer Interaction, Lecture Notes in Artificial Intelligence, 2915,* pp40-54.

[9] Picard, R.W. (1995). *Affective Computing*. MIT Technical Report 321.

[10] Picard, R.W. (1997). *Affective Computing*. The MIT Press.

[11] Rosenboom, D. (1974). *Biofeedback for the Arts, Results of Early Experiments.* Vancouver Aesthetic Research Center of Canada Publications.

[12] Vermeulen, V. (2012). The EMO-Synth, an emotion-driven music generator. In *eContact 14.2! Biotechnological Performance Practice,* .Consulted on 30th of october 2012 via http://cec.sonus.ca/econtact/14_2/vermeulen_emosynth.html.