Journal Article

Affective Audio

Weinel, J., Cunningham, S., Griffiths, D., Roberts, S. and Picking, R.

Recommended citation:

Weinel, J., Cunningham, S., Griffiths, D., Roberts, S. and Picking, R. (2014), 'Affective audio', *Leonardo Music Journal*, Vol.24, pp.17-20. published online 19.11.2014. doi: 10.1162/LMJ_a_00189

Affective Audio

Jonathan Weinel, Stuart Cunningham, Darryl Griffiths, Shaun Roberts and Richard Picking

ABSTRACT

The authors discuss their interdisciplinary research, which investigates the use of affective computing technologies in the context of music, audiovisual artworks and video games. One current project involves the expansion of mobile sound walk apps through incorporating environmental and emotional factors, forming new sonic landscapes. What type of music could reflect driving through a hot desert landscape at midday or walking through a snowy cityscape at dawn? Through a discussion of their collective work in this area, the authors aim to elicit a vision of the computer-based musical experiences of the future.

In this article we discuss our work in Creative and Applied Research for the Digital Society (CARDS) at Glyndŵr University (North Wales) that explores the application of approaches from affective computing within the context of audio and audiovisual projects. "Affective computing" [1] is an area of research that focuses on the design of computer systems that respond to and exhibit human affectations such as mood or emotion. Such computer systems can make use of a variety of available sensor technologies and biofeedback equipment, which are becoming increasingly affordable. They can also be used in mobile scenarios, as many of these sensors are already available on modern smartphones. When used with appropriate interpretive algorithms, a "context-aware" system [2] that informs us about an individual can be developed. We may consider as "context" a variety of factors pertinent to the user, including environmental, physical, emotional and social factors. The challenge for research in this area is to provide systems having the necessary sensors and inputs, suitably interpreted to give meaningful contextual information that can then be utilized for the intended purpose.

Such systems can be used for a variety of purposes, ranging from the automation of systems for informational retrieval to advertising. Among these uses, affective computer systems can be employed to control the delivery of sound, music and visualizations. These can be designed for informative purposes or be artistically driven to provide new sonic experiences. For example, we may conceive of new types of creative experience that enable the composition of visual music using gestural control alongside automatic reflections of emotion. Imagine the new types of creative social activity that could be facilitated on a global scale through computer networks. Alternatively, mobile "sound walk" apps can be expanded to incorporate environmental factors and user emotion, forming new sonic landscapes. What type of music would reflect driving through a hot desert landscape at midday, or walking through a snowy cityscape at dawn? "Affective audio" is the term we have adopted to describe our collective work in this area, which seeks to address the technical, aesthetic, social and philosophical questions this research area presents.

Jonathan Weinel (researcher), Creative & Applied Research for the Digital Society (CARDS), Glyndŵr University, Plas Coch Campus, Mold Road, Wrexham, LL11 2AW, North Wales, U.K. Email: <j.weinel@glyndwr.ac.uk>.

Stuart Cunningham (educator). Email: <s.cunningham@glyndwr.ac.uk>.

Darryl Griffiths (student). Email: <griffiths.d@glyndwr.ac.uk>.

Shaun Roberts (student). Email: <roberts.s@glyndwr.ac.uk>.

Richard Picking (educator). Email: <r.picking@glyndwr.ac.uk>.

Affective Technologies

The principal requirements of affective computing are sensor technologies—or other data input sources such as logs of online activity—that provide the raw materials that may be used in determining a user's context. These can broadly be divided into categories that deal with different aspects of the latter, such as environmental, biophysical or social factors. For example, a GPS sensor may tell us about the environmental context of an individual, while a heart-rate sensor indicates one particular aspect of biophysical context. However, there are codependencies that emerge—biophysical data alone may not necessarily give a clear indication of what a user is doing or might be feeling until it is seen in relation to other factors, such as environmental data and online activity, that help to build a clearer picture.

While many sensors, such as GPS, accelerometers and altitude meters, are available on modern smartphones, there are others that have not yet been provided that are relatively inexpensive. Co-author Darryl Griffiths has constructed an Arduino-based "Sensor Belt" that houses a comprehensive range of input sources, including temperature, humidity, light, speed, altitude, acceleration, latitude, longitude, date and time [3] (Fig. 1). Consisting of several small boxes mounted on a belt and powered by long-life batteries, it can be worn while performing most typical daily activities. Sensors for temperature, humidity and light primarily give information that is useful in discerning environmental context, while the accelerometer indicates the physical movement of the user. Through suitable interpretation and cross-examination of signals, and using noise filtering, signal analysis and fuzzy logic, Griffiths intends to determine complex contextual information regarding a user's activities, such as could be provided by smartphones in the near future.

In addition to these sensors, we are also using biofeedback sensors, such as electroencephalographs (EEG), electrocardiograms (ECG) and electro-dermal activity (EDA) sensors (Fig. 2). In principle these sensors can be used to provide information about physical activity, affect [4] or conscious state [5]. There are some difficulties in providing accurate information about user emotion; "arousal" is indicated by EDA and can reflect the intensity of an emotion but does not necessarily tell us if that emotion is pleasant or unpleasant. However, comparing this data with that provided by other sensors may provide an answer.

Information from these sensors can also be used in conjunction with third-party sources. For example, online sources such as Google Maps can be used to enrich GPS data; a location can then be identified as rural or urban (for example). Similarly, user profiles or logs of online activity may provide useful information about a user's preferences or habits. This can be compared with sensor data to improve the quality of predictions. (However, one must naturally be aware of the inherent ethical and data-protection issues arising from the use of this data.)

Utilizing the sensors and input sources discussed, we are in principle able to acquire information about the context of a user in real time (or close to it). Many of the individual "low-level" components of a contextual scenario are relatively easy to acquire; environmental temperature or humidity can be directly measured using sensors. However, others—such as user emotion or desire—that relate to ephemeral, conscious experience are inevitably much more difficult to gauge. These "high-level" features require more complex interpretative algorithms that take in the range of data sources discussed—these features are essential to provide truly affective computing systems.

This research must therefore inevitably consider the tensions that exist between the commonality of individuals and their habits, and their unique, spontaneous decisions and whims. Through his Ph.D. work, co-author Shaun Roberts is exploring the dimensionality of decisions and preferences across a variety of scenarios, with the intention of determining a universal framework for context-aware systems. In general, we believe that sufficient commonality does exist to facilitate useful affective computing systems.

Affective Musical Playlists

Once context information has been acquired, we must then decide how it can be used and mapped to sound. We shall first consider this with regard to Griffiths's project, which aims to use context information to improve the automatic generation of musical playlists. Griffiths's project proposes that it is possible to improve the automatic selection of music from a user's library of songs, based on knowledge of their context.

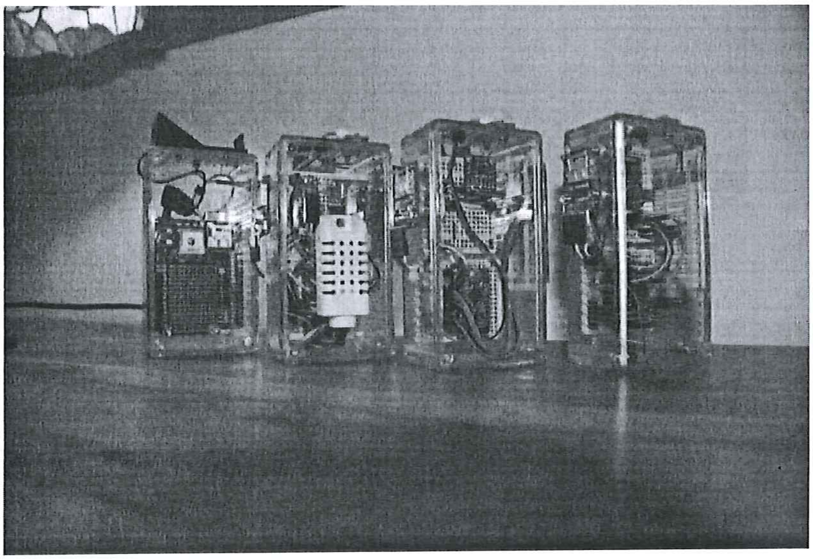
Using audio analysis, we may identify various low-level and high-level audio features [6] of musical compositions. For example, we may use tempo analysis to identify the speed of a composition or pitch analysis to identify scales. Further information can also be obtained using metadata or third-party sources to build up a complex set of descriptors for the qualities of a piece of music, including aspects such as genre or thematic content.

We may then consider what these descriptive qualities might mean for a user. For example, using our audio-analysis, can we identify a song as happy or sad, energetic or relaxed? The typical Western understanding of minor key music would suggest that "sad" music can be identified by this sign in some cases. Similarly, perhaps due to anthropomorphic signification, music involving rapid rhythms tends to be seen as energetic. DeNora cautions us against ascribing absolute interpretations to music, yet nonetheless comments that common meaning in music can be established and then reinforced through patterns of use [7]. Although there may be cultural differences regarding interpretations, we may argue that there is some commonality that can be exploited by our system.

What about other descriptive terms that describe nonmusical properties, but which a piece of music might signify through thematic content or some other aspects of its compositional design? For example, can a piece of music evoke environmental characteristics—a desert scene or a jungle? Can music be "hot" or "cold"? "Wet" or "dry"? "Night" music or "day" music? [8] Nonmusical descriptors that suggest such signifying properties in music can be found almost everywhere; this is a pervasive feature of music journalism, for example. Even titles of musical genres can be suggestive of nonmusical properties such as an environment or location: "country", "desert-rock" (a type of rock music originating from Palm Desert, California) or "jungle" (a type of electronic dance music) are three examples. These descriptors may be suggested by lyrics, song titles, packaging or the place from which an artist originates. They may also be indicated by purely sonic qualities such as timbre, instrumentation or musical arrangement; think of the "wet" sound of surf music, suggested by the use of reverb, tremolo and tape delay effects. These signifying factors can also be designed through musical composition and presentation of the cultural artifact (records) in order to create a particular impression. A country artist need not be from rural North America in order to make country music or evoke a similar impression for audiences. Hence there is a certain fluidity in the interpretation of such factors, yet arguably the fact that such signifying properties can be fabricated would suggest that there is significant commonality that we might be able to analyze.

Unpacking and quantifying these signifying properties may require much research. Which signifying descriptors can we meaningfully attach to music, and for which groups of people are they valid? In order to investigate this, we are currently designing a system that will allow us to query many listeners from different regions regarding their interpreta-

Fig. 1. The main components of Griffiths's wearable sensor system; each box contains a variety of sensors and can be mounted on a belt. (Photo © Darryl Griffiths)



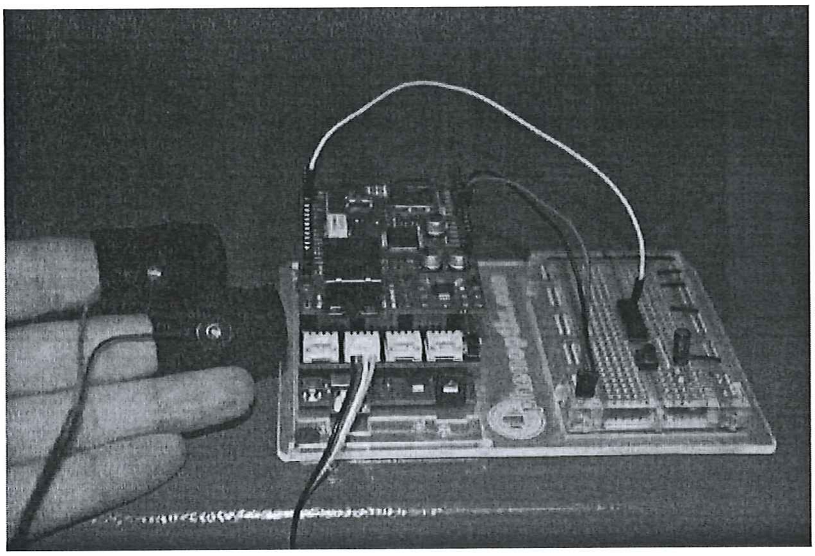


Fig. 2. An electro-dermal activity (EDA) sensor from Griffiths's system. (Photo © Darryl Griffiths)

The generation of these sounds could be entirely automated/generative, or perhaps the role of specialist composers will be to isolate certain contextual features, thereby giving us their own sonic schemas from which listeners can then choose.

As with musical playlists, we can also view these context-dependent ambient sonic experiences as a means to engineer our moods. For example, Volkswagen and Underworld's recent Play the Road app [15] plays sections of music by Underworld based on the speed of the vehicle. They propose that "the car becomes the instrument, the driver the musician." However, if driving faster causes an increase in tempo, does this encourage faster driving? As exciting as these new paradigms of sonic experience are, there are far-reaching ethical and philosophical implications developing from these systems that need to be investigated.

tion of music. Our approach will utilize a repertory grid, a technique based on personal construct psychology [9], which we have used previously to explore user interpretations and preference in single- and mixed-media environments [10,11]. This will enable us to explore the boundaries of such interpretations on both macro and micro levels by investigating the terms listeners associate with whole genres, artists, albums, songs and individual instruments or sounds.

Once this data regarding the signifying characteristics of music has been acquired, we can use it to analyze a musical library, assigning these properties to music tracks. The properties discussed relate directly to environmental and emotional characteristics, which we propose can also be identified using context data. Real-time contextual data can therefore be mapped to the signifying properties of music; music can be selected that matches a given context. Different environmental or emotional contexts may be used to affect the choice of music. This paradigm of music listening suggests a form of automated "life soundtracking"—something many of us arguably already do by picking music that "matches" our experience when in a car, on a train or having dinner.

Of course, it might not always be the case that users want music that matches their experiences, as people are known to use music to change their moods. In such cases, the listener might prefer to select the mood or environmental context (s)he would like to be in and have the system play this type of music. Hence, if the reality is a cold, wet, miserable day in the city, the user could tell the system to play the music that (s)he would hear if (s)he were happy and relaxed at a beach cafe in a warm and sunny location. The system of correspondences we have been discussing could easily be adapted to facilitate this. This usage suggests a form of semi-automated "emotioneering" [12]. One of the advantages of this, particularly when seen in the context of streaming music services, is that the listener could select a particular mood or idea and have the system find the music that matches it—I might not know exactly what album would give me that sunny beach cafe mood, but the system could find it for me.

Affective Ambient Sonic Experiences

While the system discussed could be used for musical playlists, we may also consider how contextual data might be employed in other ways to create ambient sonic experiences. For example, individual sound objects (in terms of spectromorphology [13]) could be used to signify environmental or emotional characteristics. Hence generative, ambient compositions can be produced according to the context of the user. This extends the concept demonstrated by "sound walk" apps such as Ignacio Percino's SonicMaps [14], which facilitate sonic experiences based on GPS location. Beyond GPS, we can use a wider range of other sensors, creating new musical experiences. In a sense, listeners become composers, since by navigating through different contexts they play a role in creating the music.

Affective Visual Music

Contextual data can also be used in other music-related contexts. The "Holophonor" is a fictional musical instrument from the science-fiction TV show *Futurama* that creates holographic images in response to the emotion of the performer. The Holophonor, in many ways the ideal portable "visual music" instrument, could become a reality through fusion of the artistic field of visual music with affective technologies, and we propose a possible design [16].

We have developed an early precursor to such an affective visual music system: the *Psych Dome* [17]. It utilizes a consumer-grade EEG headset in order to link brain activity to the generation of sounds and graphics based on psychedelically altered states of consciousness. This project was presented as an art installation in an immersive dome environment. Through the *Psych Dome* we began to investigate the user experience of inexpensive consumer-grade EEG in this type of context. The results of our pilot study showed that most of our participants did not experience a statistically meaningful connection between the EEG signals and the sounds and graphics produced. This underscores the need for further research regarding how to utilize consumer-grade EEG effectively.

Affective Video Games

The principles discussed are applicable to video games. Tom Garner has carried out substantial work regarding sound, biofeedback and fear perception in games [18]. In our ongoing Quake Delirium project [19], we have been exploring the adaptation of a first-person shooter video game to reflect altered states of consciousness. While the default approach used in video games of this type is to use a virtual camera to simulate the eyes of a game character, we can enhance this by simulating changes in visual or auditory perception by changing graphics or sound parameters. In principle, the visual and auditory components of an altered state of consciousness can be reflected through the use of a game engine. Utilizing Hobson's Activation, Input, Modulation model of consciousness, we can use EEG signals to control the conscious state of the virtual game character. Hence we can either reflect a state of consciousness that someone is experiencing or simulate one that (s)he is not.

In a related project, we are also exploring the use and display of biofeedback in games in social situations. While Garner's work and our Quake Delirium projects are oriented toward adrenaline gaming, systems of this kind may also have therapeutic uses in the future. As Tony Brooks's SensoramaLab demonstrates [20], such systems may be able to help the user obtain states of meditation or aid in relaxation or physical therapy.

Conclusions

In this article we have outlined the research area that we refer to as "affective audio," which merges the areas of affective computing and the sonic arts. In addition to sound-based projects, we are also exploring these approaches in the context of visual music and video games. However, audio remains an important unifying feature of our work, perhaps because of the powerful resonances that music has with places and emotions.

Affective audio may suggest a significantly new paradigm of interaction with sound and music. These interactions may give users ownership over their experiences of sound and music through automated reflections of context. The futuristic experience of navigating through sonic landscapes that relate to ephemeral qualities such as emotion may not be too distant. In many ways such experiences can be seen as a logical extension of the way in which we already choose music to suit particular situations or change our moods. It is because the underlying principles of affective audio are already pervasive features of how we listen to music that we believe this area of research holds such radical potential.

References and Notes

1. R. Picard, *Affective Computing* (Cambridge, MA: MIT Press, 2000).

2. B. Schilit, N. Adams and R. Want, "Context-Aware Computing Applications," *Workshop on Mobile Computing Systems and Applications 1994* (IEEE, 1994) pp. 85–90.

3. D. Griffiths, S. Cunningham and J. Weinel, "Automatic Music Playlist Generation Using Affective Technologies," *Proceedings of the Fifth International Conference on Internet Technologies & Applications 2013*, Glyndŵr University, Wales (2013).

4. J. Russell, "A Circumplex Model of Affect," *Journal of Personality and Social Psychology* 39, No. 6, 1161–1178 (1980).

5. A.J. Hobson, *The Dream Drugstore: Chemically Altered States of Consciousness* (Cambridge, MA: MIT Press, 2002) pp. 44–45.

6. D. Mitrovic, M. Zeppelzauer and C. Breitender, "Features for Content-Based Audio Retrieval," *Advances in Computers: Improving the Web* 78 (2010) pp. 71–150: <www.sciencedirect.com/science/article/pii/S0065245810780037>.

7. T. DeNora, *Music in Everyday Life* (Cambridge, U.K.: Cambridge Univ. Press, 2010) pp. 21–45.

8. This concept has been discussed with regard to videogame soundtracks by Karen Collins, in a keynote lecture presented at *Audio Mostly 2013*, Luleå University of Technology, Piteå, Sweden (18 September 2013).

9. A.G. Kelly, *The Psychology of Personal Constructs. Volume 1: A Theory of Personality* (W.W. Norton, 1955).

10. S. Cunningham, "Applying Personal Construct Psychology in Sound Design Using a Repertory Grid," *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound* (ACM, 2010).

11. S. Cunningham and R. Picking, "Sounds Relaxing—Looks Cool: Audio and Visual Selections for Computer Systems that Support Wellness," *International Journal of Ambient Computing and Intelligence* (IJACI) 4, No. 1, 40–53 (2012).

12. M. Grimshaw, "Computer Game Sound: From Diegesis to Immersion to Sonic Emotioneering," in Jens Gerrit Papenburg and Holger Schulze, eds., *Sound as Popular Culture: A Research Companion* (Cambridge, MA: MIT Press, forthcoming 2015).

13. Spectromorphology is an approach to sound materials and musical structure established by Denis Smalley, which focuses on the spectrum of available pitches and their shaping in time. See D. Smalley, "Spectro-Morphology and Structuring Processes," in S. Emmerson, ed., *The Language of Electroacoustic Music* (Basingstoke, U.K.: Macmillan, 1986).

14. I. Percino, SonicMaps, locative audio software for iOS and Android (2012), <sonicmaps.org/about.html>. Accessed 1 May 2013.

15. Volkswagen, "GTI + Underworld: Play the Road," web page regarding mobile application: <www.volkswagen.co.uk/playtheroad>. Accessed 1 May 2013).

16. We discuss the Holophonor in J. Weinel et al., "Holophonor: On The Future Technology of Visual Music," in K. Curran, ed., *Recent Advances in Ambient Intelligence and Context-Aware Computing*, (IGI Publishing, forthcoming 2014).

17. J. Weinel, *Psych Dome* (1:40), EEG-controlled interactive visual music installation in an immersive dome environment, first presented at William Aston Hall, Glyndŵr University, Wales (16 October 2013).

18. T. Garner, "Game Sound from behind the Sofa: An Exploration into the Fear Potential of Sound & Psychophysiological Approaches to Audio-centric, Adaptive Gameplay," Ph.D. thesis, University of Aalborg (2012).

19. J. Weinel, "Quake Delirium Revisited: Systems for Video Game ASC Simulations," *Proceedings of the Fifth International Conference on Internet Technologies & Applications 2013*, Glyndŵr University, Wales (2013).

20. T. Brooks, keynote lecture, *Fifth International Conference on Internet Technologies & Applications*, Glyndŵr University, Wales, 12 September 2013.

Manuscript received 2 January 2014.

Jonathan Weinel is a Postdoctoral Research Associate at Glyndŵr University; his work focuses on representations of altered states of consciousness (ASC) in sound and computer graphics.

Stuart Cunningham is Head of Creative Industries at Glyndŵr University. His research interests include audio compression techniques, human-computer interaction and sound design.

Darryl Griffiths is currently completing his Ph.D. at Glyndŵr University, studying automatic playlist generation using affective computing technologies.

Shaun Roberts is a Ph.D. candidate at Glyndŵr University. His research focuses on context-aware technologies in the organization of large digital media libraries.

Richard Picking is a Reader in Human-Computer Interaction and Head of the Centre for Creative and Applied Research for the Digital Society (CARDS) at Glyndŵr University.