Arne Eigenfeldt.

Generating Affect: Applying Valence and Arousal values to a unified video, music, and sound generation system

Multimedia Generative Art

Authors:
J. Bizzocchi[1], A. Eigenfeldt[2], M. Thorogood[1] [1] School of Interactive Art and Technology, Simon Fraser University, Surrey, Canada
[2] School for the Contemporary Arts, Simon Fraser University, Vancouver, Canada

Abstract:
Art can clearly affect viewers and listeners in very emotional ways; however, artists will often reject the claim that the emotion is *in* the artwork itself, and instead insist that emotions felt are solely within the viewer/listener. How does one reconcile these seemingly opposing views? Efforts have been made to discover the relationship between emotion and music [1] as well as moving image [2]; however, these studies have limited direct application for the generative artist.

Russel's circumplex model [3] introduced two very significant parameters for describing features that may produce emotional responses in listeners: *valence* (pleasant/unpleasant) and *arousal* (eventful/uneventful). These objectives measures can be used both analytically as well as for generative purposes.

Within our multimedia installation, *Seasons*[1], we are using these measures to drive the music and soundscape generating systems based upon an analysis of the video system's current output. The artwork uses visually evocative nature shots, and its goal is to support an ambient user experience that is calming and contemplative; as such, the values for valence and arousal in the video are relatively moderate.

Seasons combines three very different generative systems: video, music, and soundscape. In order to maximize aesthetic coherence and flow, the artwork relies on a chain of valence/arousal assessments and communications. The chain starts with the database of video clips. The video sequencing system uses a set of content tags to select and order the stream of clips. Each clip has also been assessed and tagged by the artists for its valence/arousal values. This assessment is based on the artists' subjective evaluation of each of the shots.

The valence/arousal values for each selected shot are then sent to the two audio systems. The music system uses artificial agents (called "Musebots") to compose and create an original generative music track that reflects the valence-arousal values of the images. The soundscape system uses both content tags and the valence-arousal values from the video stream to select and mix a sound effects soundscape that is consistent with the video and the music. Future modifications of our system will include the development of software that can generate valence-arousal values based on computational feature extraction and analysis.

1. https://vimeo.com/136362499

Main References:
[1] Hevner, K. (1937). "The affective value of pitch and tempo in music." The American Journal of Psychology, 49(4), 621-630.
[2] Cohen, A. (2001). "Music as a source of emotion in film." Music and emotion: Theory and research. 249-272.
[3] Russell, J. (1980). "A circumplex model of affect." Journal of Personality and Social Psychology, 39(6), 1161-1178.

Contact:
Jim Bizzocchi, jimbiz@sfu.ca

Keywords:
affect, emotion, generative video, generative soundscape, generative music

Applying Valence and Arousal Values to a Unified Video, Music, and Sound Generative Multimedia Work

Arne Eigenfeldt

School for the Contemporary Arts, Simon Fraser University, Vancouver, Canada

arne_e@sfu.ca

Jim Bizzocchi, Miles Thorogood, Justine Bizzocchi

School of Interactive Art and Technology, Simon Fraser University, Surrey, Canada

jimbiz@sfu.ca, mthorogo@sfu.ca, justine@sfu.ca

Abstract

We describe our research-creation across multiple generative systems, using the parameters of valence and arousal as unifying parameters. A variety of methods have been explored that translate emotional responses in viewers into objective measures; however, most of these are not useful for artists, especially generative artists. However, valence (pleasantness) and arousal (eventfulness) are two parameters that do suggest generative potential. We describe three generative systems – a recombinant video system, a soundscape generation system, and a multiagent music system – and how they individually use valence and arousal for generative purposes. Finally, we describe an artwork in which these three generative systems operate collaboratively to produce a multimedia installation.

1. Introduction

Art can clearly affect viewers and listeners in very emotional ways; however, artists will often reject the claim that the emotion is *in* the artwork itself, and instead insist that emotions felt are solely within the viewer/listener. How does one reconcile these seemingly opposing views? Efforts have been made to discover the relationship between emotion and music [1] as well as moving image [2]; however, these studies have limited direct application for the generative artist.

For example, many psychology studies have focused upon a stimulus response model, in which subjects are asked to rate musical excerpts using a set of adjectives, such as "cheerful, gay, happy; fanciful, light; delicate, graceful; dreamy, leisurely; etc." [3]. Other studies have used fMRI analysis to determine subject's neural responses to "pleasant" and "unpleasant" music [4]. While such studies may produce interesting data regarding listener experience, they leave few cues for artists: Dvorák's *Slavonic Dance No. 8* in G Minor may be considered "happy", but how can that inform a compositional practice?

Russell's circumplex model [5] introduced two very significant parameters for describing features that may produce emotional responses in listeners: *valence* (pleasant/unpleasant) and *arousal* (eventful/uneventful). These objective measures can be used both analytically as well as for generative purposes, primarily because such objective measures can be considered during the creative process. Artists can readily translate these

measures within their medium: in music, for example, eventfulness can be translated as activity, and pleasantness can be translated as tension.

The circumplex model overlays emotional states on the two-dimensional scale (see Figure 1); significantly, these emotions result from the relationship between the two measures of arousal (or eventfulness, also called activation/deactivation) and valence (or pleasantness). Therefore, an artist, generative or otherwise, can create an artwork that has low arousal and high valence – two objective measures – and be confident that it will be perceived as "calm and relaxed", two subjective responses.

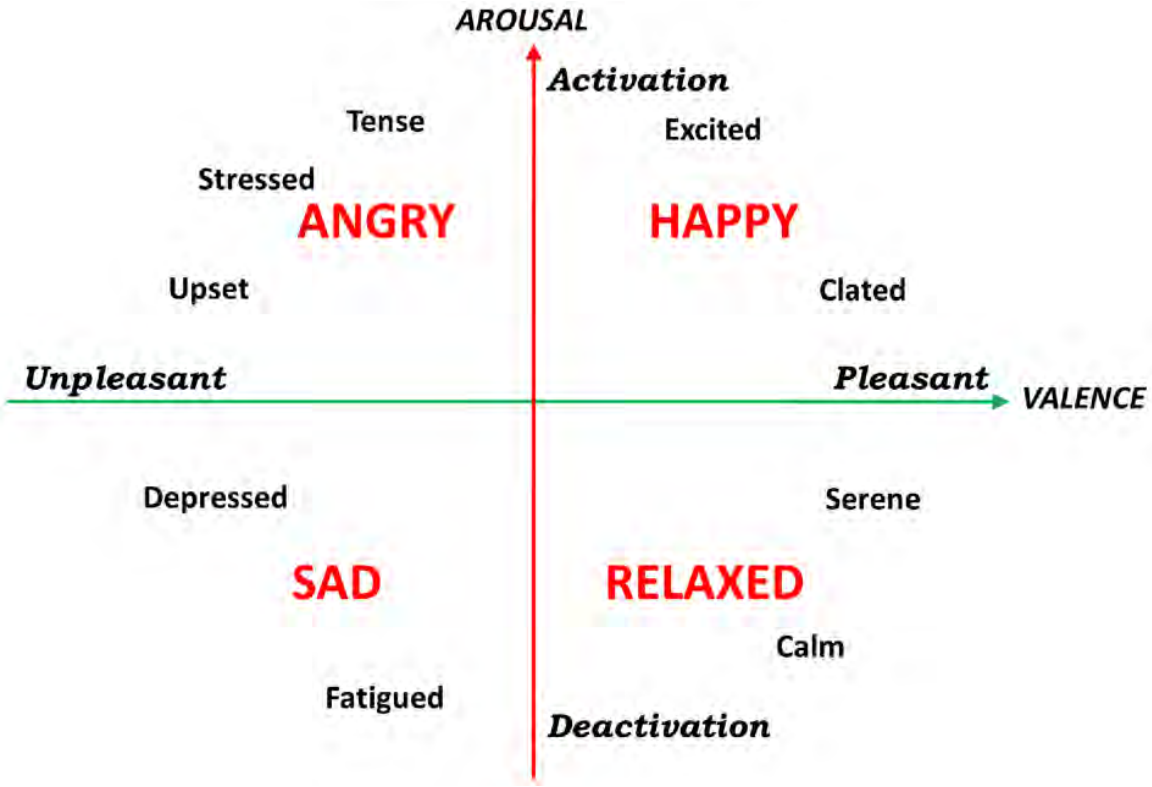

Figure 1. Valence / Arousal model

1.1 Generative Systems

Generative engines are systems that provide various, expected or unexpected patterns by using a series of rules. The use of systems, rather than intuition, for artistic creation has a long tradition, and for this reason, Galanter suggests that generative art can be considered as old as art itself [6]. A contemporary approach to generative art has arisen: metacreation. Using tools and techniques from artificial intelligence, artificial life, and machine learning, metacreation develops software that is creative on its own. In other words, software is a metacreation if it exhibits behaviors that would be considered creative if performed by humans [7].

While the potential of codifying artistic decisions may be alluring to many artists, the challenges are many: for example, can the notion of creativity be extended to machines, or can they (should they?) only remain as tools for the creative artist? The growing field of metacreation explores these questions, and is populated by psychologists, art theorists, cognitive scientists, artificial intelligence researchers, machine learning specialists, and, perhaps most importantly, artists. As such, it is not merely an academic pursuit: it is, and has been, a fertile creative domain for artists exploring new avenues of production.

Many examples of metacreative works in sound, music or video can be found, but the authors have not been able to discover any that combine these into a system where the elements interact and generate a blended audio-visual work in real-time. Some generative and metacreative systems that were inspirational to our work are described in a previous paper [8].

Some production systems, particularly in music, have investigated the potential for affective generation, including the use of valence and arousal. Wallis *et al.* created an Emotional Music Synthesis system in which valence/arousal were mapped to various musical parameters [9]; however, such heuristic and ad hoc mappings will always depend upon the desired musical output, and are seldom generalizable. For example, Wallis *et al.* decided that the phygrian mode constitutes a minimal valence, whereas the lydian mode constitutes a maximal valence. Such straightforward mappings limit the harmonic potential of the generated music to modal melodic material, despite any experimental user studies that may suggest correlations between a generated melody's intended and perceived valence/arousal.

Oliveira and Cardoso [10] use machine learning to discover these mappings objectively. 80 participants rated 96 musical excerpts from film music, albeit MIDI versions of the music, for valence and arousal. Classifying this data using 146 musical features, the more salient musical parameters for affect were derived. Of note, the experiments suggest temporal and density features are superior for predicting valence, countering the more intuitive notion that harmony is its main determinant; this might result from their study's limited harmonic features, as it suggests that the most important harmonic feature for valence was minor versus major scales, a somewhat simplistic notion.

As artists, generative or otherwise, have known for centuries, the affective power within an artwork is dependent upon its context; an image, melody, or sound will be perceived by viewers and listeners differently, given its circumstance. While it may be possible to determine that certain features relate to valence and arousal percepts, the application of this information is still reliant upon artists for its effective use.

2. Seasons

Seasons is an audio-visual journey through our natural environment across the span of a year. The work is situated within the genre of "Ambient Video" - artworks that provide ongoing visual and emotional pleasure without requiring our attention in any particular moment. The primary goal of the work's creative team is to work within this tradition to produce a successful generative artwork that fully satisfies our artistic sensibilities. Its generative underpinnings are integral to its appreciation, while its metacreative aspects are clearly important, yet secondary to its artistic intentions.

The system comprises video sequencing and transitions enriched through their interaction with music and soundscape. The full work is a real-time and ongoing cybernetic collaboration between three independent but communicating generative systems: video (*Re:Cycle*, see Section 3), soundscape (*AuMe*, see Section 4), and music (*Musebots*, see Section 5). The work runs continuously using a variety of computational processes to build the audio-visual output for a single large-screen display and multi-channel sound system.

Within this multimedia installation, we are using measures of valence and arousal to drive the music and soundscape generating systems based upon an analysis of the video system's current output. The artwork is intended to support an ambient user experience - one that is calming or contemplative. The visuals consist of a series of evocative nature shots that are consistent with this goal (see Figure 2). The values for valence and arousal in the video are therefore relatively moderate.

Figure 2. Example evocative nature shot from *Seasons*

In order to maximize aesthetic coherence and flow between the three very different generative systems, the artwork relies on a chain of valence/arousal assessments and communications (see Figure 3). The chain starts with the database of video clips. The video sequencing system uses a set of content tags to select and order the stream of clips. Each clip has also been assessed and tagged by the artists with text cues for sound as well as for its valence/arousal values, an assessment based on the video artists' subjective evaluation of each of the shots.

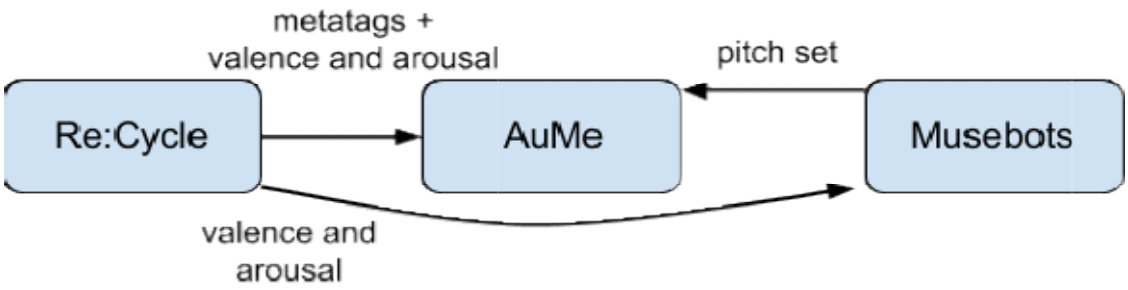

Figure 3. Metatag and Valence/Arousal pipeline between the video system (*Re:Cycle*), the soundscape system (*AuMe*), and the music systems (*Musebots*)

The valence/arousal values for each selected shot are then sent to the two audio systems. The music system uses artificial agents (musebots) [9] to compose and create an original generative music track that reflects the valence/arousal values of the images. The soundscape system uses both the text cues and the valence/arousal values from the video stream to select and mix a soundscape that complements the video and the music and will achieve the artists' overall aesthetic goals.

3. Video engine

The video system is an extension of Bizzocchi's computational video sequencing and presentation system entitled *Re:Cycle* [11]. This system relies on a recombinant process to combine and sequence shots and transitions drawn from the system's databases, currently consisting of over 250 shots. *Re:Cycle* uses metadata tags to nuance its selection process with an enhanced semantic coherence. As *Seasons* strives for an "Ambient Video" aesthetic [12], the content consists of imagery drawn from nature and landscape; therefore, the tags reflect the content of the individual shots: for example, "trees", "water", "mountain", "snow". Short sequences of shots are selected and presented based on these content tags. This very simple computational process significantly increased the visual flow and unity of the piece.

Creative use of tags determined by the artists for the video clips drives the selection and sequencing of the visuals, as well as providing triggers for the selection, processing and playing of the music and soundscapes. As shown in Figure 4, these commentaries form part of a pipeline that communicates video, mood, and tonal metadata between the subsystems of *Seasons* using the Open Sound Control (OSC) protocol [13].

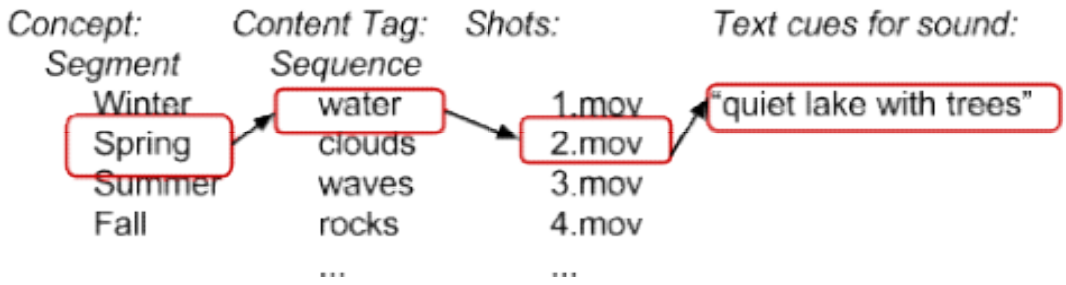

Figure 4. Video sequencing logic

There is a hierarchy to the tagging and sequencing logic. During performance, an initial *concept* is chosen, which gives coherence to an visual *segment*. In *Seasons*, concepts consist only of the four separate seasons (see Figure 4). These season segments are cycled through in the same order. Within each season, *content tags* are chosen from those available within that concept subset. The system then filters all shots in the seasonal subset of the database that contain that content. From this collection, the system uses random operations to select the shots for this *sequence*. Each shot, which lasts 55 seconds, has a unique description associated with it, which is used by the soundscape engine (see Section 4). For *Seasons*, the system has been set so that each segment contains four content-defined sequences, and that each content sequence has three video clips.

The system runs indefinitely, going through the seasonal segments as it proceeds. For this work, there is no beginning and no end; shots do repeat, but not in the same order or context. The effect of this method is the automatic generation of a series of coherent shot sequences nested within a larger thematically-based video segment. The artist's creative use of both the initial settings and the segment and content tags drives the resulting generated video thematic progression and viewer experience. There is still an element of randomness in the sequencing selections, which builds in an ongoing variability and helps to maintain viewer interest over multiple viewings. At the same time, the tagging and selection mechanisms maintain ongoing content coherence and visual flow. This unifying connection of sequencing decisions produces an experience that is often 'read' by the audience as a traditional linear video built upon human-produced visual and semantic integrity.

3.1 Valence/Arousal in video engine

The video sequencing drives the soundscape and music generative processes. At the moment, each shot in the database has been hand-tagged with a subjective valence/arousal measure. This in itself was a difficult process, and methods had to be determined as to how to apply these measures in as objective a way as possible. While there are a number of studies which are working to establish some objective measures of visual elements correlated to valence and arousal in photos or videos [14], they are not sufficiently advanced in their results to be able to be implemented by the artists for this work.

A method of scaling had to be determined: would the least "pleasant" of the shots be considered a low valence, and how would that be determined? The clips are assessed individually, although it is likely that affective perceptions by viewers of a particular shot would actually be skewed by its context in a sequence (the montage effect). Further, elements of the shot may change over the length of the clip. Finally, the videos for this piece were selected with the ambient video aesthetic in mind. They purposefully seek to be both calm and pleasant. This naturally limited the full range of valence and arousal values that might be found in randomly acquired images.

For *Seasons*, the assessments were done primarily as an intuitive and creative exercise, with an attempt to utilize some consistent guidelines. The subjective assessments combined both content references and more abstract visual elements. For example, close-ups of brightly coloured flowers were assessed as high on both valence and arousal. This reflects the intuitive sense that spring flowers are 'happy' and 'stimulating'. Fast-moving, rushing water was assessed as high in arousal (eventful), but could be low or high on valence depending on the surrounding elements (dark skies and rain compared to bright sunny day or flocking birds).

The result in the view of the artists' was remarkably successful. The music and soundscape work very well with the ongoing dynamic video sequences. However, there is much more work to be done in applying these affective measures to moving visuals and then understanding how best to utilize them to create a blended audio-visual work.

4. Soundscape engine

A generative system – Audio Metaphor (*AuMe*) – autonomously creates the soundscape by processing metadata communicated from *Re:Cycle*. *AuMe* takes a modular approach for analyzing the valence/arousal tags related to videos, retrieving and segmenting audio files, and processing that information to generate a representative soundscape. Each generated soundscape consists of multiple layers of background and foreground sounds related to the concepts found by semantic analysis. Described in greater detail elsewhere [15], the arrangement and processing of those sounds is controlled by a mixing engine, making decisions based on a set of soundscape composition rules. The duration of a soundscape correlates with the duration of a video clip, and the transition from one soundscape to the next happens by interweaving sounds as a video transition occurs.

4.1 Valence / Arousal in soundscape engine

AuMe uses a database of curated soundscape recordings that has been analysed for measures of valence and arousal [16]. Recordings are automatically annotated with these measures by a machine-learning algorithm trained with example recordings labeled by human listeners during a listening study. For the study, we were interested to observe if a correlation exists between low-level audio signal features and how listeners perceive a soundscape, in terms of valence and arousal. Having established that indeed such a correlation exists, we were then able to model a machine for classifying soundscape recordings based upon audio signal features.

Each soundscape generated by *AuMe* utilizes the machine-learning model to be sympathetic to the sentiment content of a video. For *Seasons*, *AuMe* selects segments of soundscape recordings using the valence and arousal tags associated with the current video clip. If a clip contains a high degree of arousal, then the soundscape will

be comprised of sounds that have a similar degree of arousal. Furthermore, the mixing engine will also respond to this measure by mixing the sounds to convey a more active scene. The same approach is taken for valence, where a higher degree of valence will result in a perceptually more pleasant soundscape. Similarly, if the video is tagged as having a low valence or arousal, then *AuMe* will generate a soundscape to convey those measures.

In its implementation for *Seasons*, *AuMe* generates a new soundscapes at every cycle of the year by creating slight variations in its search space, based on a Gaussian probability distribution. *AuMe* is influenced by this search space variation. As different sets of audio files are given for mixing, the behaviour of the *AuMe* system is altered. Thus there is a slight variation of the generated soundscapes even as the video clip tags remain constant.

6. Music engine

An agent-based production system creates the music through generative means based upon information received from *Re:Cycle*. The music engine is comprised of independent musical agents – musebots – that act autonomously, yet collaboratively with one another. Described in more detail elsewhere [17], each musebot generates a particular aspect and/or function within the music – for example, a bass line; however, it does so in relation to other active musebots. The musebots, of which there are currently dozens, function within curated ensembles, and communicate via messaging. Within *Seasons*, one musebot will generate a harmonic progression, and send this information to other musebots for their interpretation. A *Conductor* handles the tempo, as well as serving as an overall relay station for musebot messages and initiating each successive musebot ensemble. New ensembles are launched for each new season; because a season consists of four sequences of three clips, each of which lasts 55 seconds and contains a 22 second transition, an ensemble's duration is currently locked at 6 minutes 36 seconds. Tempo is consistent for a given ensemble: for *Seasons*, tempo varies between 40 and 60 beats per minute.

6.2 Musebot reaction to Valence/Arousal

Because each musebot is autonomous, yet sensitive and reactive to its ensemble partners, the music can vary substantially over the course of a season's duration. Each musebot has its own defining parameters, although several more general parameters overlap. For example, "density" is a common element which can be interpreted as the number of events over a given time period: the greater the number, the higher the perceived musical density. Incoming arousal (eventfulness) messages are mapped directly to a musebot's density parameter; however, each musebot may interpret density in a nonlinear way, depending upon its previous states, and other active musebots. In other words, a given valence will not produce a fixed response from a musebot ensemble.

Valence is even more subtle, and can be interpreted in a variety of ways. Pleasantness can be mapped to musical consonance, which in turn can be mapped to "simple" versus "complex". While each musebot can interpret complexity in a variety of ways, one of the most audible is by the harmonic generating musebot: simple chord progressions use simple chords (i.e. triads) and the root movement is predictable, whereas complex chord progressions use chords with extensions and alterations (for example, as found in jazz) and more unpredictable root movements. Similarly, melodic musebots can prefer stepwise melodic movements (high valence) or more disjunct melodic movement (low valence), and rhythmic musebots can emphasize simple metric divisions that enforce the beat (high valence) or complex subdivisions that subvert the beat (low valence).

Within the current implementation of *Seasons*, musebots will alter their generation subtlely with each new clip, as new valence/arousal measures are sent when clips initiate their incoming transitions. As such, the musebots behave very much "in the moment", without any sense of the future. For example, the given clip may have a high valence / low arousal tag, and the musebots will exhibit a "relaxed" emotion, but without knowledge of the next clip's tags, they will lack any real sense of movement or progress. Within the aesthetic of ambient art, such lack of musical motion is acceptable, and even preferred; however, as will be discussed in Section 8, a pre-

computed structure, complete with detailed vectors for time-varying valence/arousal, will be necessary for the musebots to anticipate changes in shot, segment, and concept.

7. Installations

Seasons has been shown in two different installation environments. In August 2015, as part of ISEA 2015, it was shown in an enclosed room in which the walls had thick black curtains that created an isolated viewing and listening environment. A high-quality four channel sound system was used, along with a large 60" LED monitor; a bench was provided for audiences, inviting them to experience the work in contemplative surroundings. Based upon our own experiences within the installation and watching audiences anecdotally, we found that viewers stayed within the space for anywhere from five to twenty minutes, with the latter time allowing them to experience multiple seasons. Responses were extremely positive, however anecdotal they might be.

A second installation was presented as part of a large exhibition, coexisting with several other generative installations. The event was held over four weeks, and audiences had docents present tours, explaining each work. In this case, *Seasons* was displayed in an open gallery context – replete with extraneous public noises and competing soundscapes and music from other installations. Although we did provide the possibility of headphones, as well as the soundscape and music being heard through the monitor's built in speakers, the lack of comfortable seating did little to motivate audiences to spend more than a few minutes with the work. While the audiences were more informed about the nature of the work and collaborative effort, due entirely to the docent's explanations, the tour format, which necessitated timely movement through all exhibits, sadly negated this benefit.

8. Conclusion and Future Work

We feel that *Seasons* is an effective installation, if presented in a situation where viewers can experience its ambient nature properly. The interaction between the three generative systems is not immediately obvious, and, somewhat paradoxically, requires some critical viewing and listening on the part of its audience. However, we prefer to consider this as rewarding the audience for active viewing/listening, as the ambient aesthetic of indirect interaction is still supported.

Each generative system operates successfully as an independent entity, while also following the unifying features of valence and arousal. The subjective, rather than purely objective, interpretation of these features allows for a considerable amount of variation; for example, a single video shot, with its set valence/arousal rating, will elicit different reactions depending upon its context.

Larger issues within generative art, specifically large-scale structure, are happily avoided within *Seasons*, as these are determined by the imposed structure of the seasons themselves. The approximate six minute duration of a single season can be treated as one coherent gesture; any longer, and sectional changes and more direct development of material would be required by the three systems.

Several improvements are still planned for *Seasons*. The first of which would be the selection of all shots by the video engine in advance, allowing for the season's complete valence/arousal measures to be communicated prior to commencing. This would allow the music and soundscape engines to sculpt their selections better, as it would provide a target shape for the longer section. Another modification will include the development of software that can provide feature extraction and analysis of video. This would lead to a computational assessment of valence/arousal, and could also provide additional cues on how to build certain shot sequences.

Our next project will be to create a generative documentary in the tradition of the early "City Films" such as the Ruttmann 1927 classic silent film *Berlin: Symphony of a Great Metropolis*. We will deconstruct selected films in this genre in order to guide our plans for shot creation and for visual sequencing. We will then adapt our system of segment/sequence/shot selection to create a contemporary video montage in the spirit of these classic city films. This new project will allow us to more deeply explore the dynamics of valence/arousal values within the design and the experience of our work. A city documentary will have a much more varied emotional palette than the relatively constrained ambient aesthetic, so the variance in the valence/arousal values in the elements of this work will be much broader. The soundscape and music systems will therefore generate appropriate accompaniment material that reflects the more complex affective range of a complex urban environment.

References

[1] Hevner, K. (1937). "The affective value of pitch and tempo in music." The American Journal of Psychology, 49(4), 621–630.

[2] Cohen, A. (2001). "Music as a source of emotion in film." Music and emotion: Theory and research, 249–272.

[3] Li, T., and Ogihara, M. (2003). "Detecting emotion in music." International Society of Music Information Retrieval, Baltimore, 239–240.

[4] Koelsch, S., Fritz, T., Müller, K., and Friederici, A. D. (2006). "Investigating emotion with music: an fMRI study." Human brain mapping, 27(3), 239–250.

[5] Russell, J. (1980). "A circumplex model of affect." Journal of Personality and Social Psychology, 39(6), 1161–1178.

[6] Galanter, P., (2003). "What is Generative Art? Complexity theory as a context for art theory." GA2003–6th Generative Art Conference.

[7] Whitelaw, M. (2004). Metacreation. Art and Artificial Life. Cambridge: MIT Press.

[8] Eigenfeldt, A., Thorogood, M., Bizzocchi, J., and Pasquier, P. (2014). "MediaScape: Towards a Video, Music, and Sound Metacreation." Journal of Science and Technology of the Arts, 6(1), 61–73.

[9] Wallis, I., Ingalls, T., Campana, E., and Goodman, J. (2011). "A rule-based generative music system controlled by desired valence and arousal." Proceedings of 8th International Sound and Music Computing Conference, Padova.

[10] Oliveira, A. P., & Cardoso, A. (2010). "A musical system for emotional expression." Knowledge-Based Systems, 23(8), 901–913.

[11] Bizzocchi, J. (2011). "Re: Cycle – A Generative Ambient Video Engine." Entertainment Computing – ICEC, Springer, 354–357.

[12] Bizzocchi, J. (2008). "Ambient video: the transformation of the domestic cinematic experience." Small Tech: The Culture of Digital Tools, 22, 197–206.

[13] Wright, M., Freed, A. (1997). "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers." International Computer Music Conference, Thessaloniki, 101–104.

[14] Baveye, Y., Dellandréa, E., Chamaret, C., and Chen, L. (2015). "Deep Learning vs. Kernel Methods: Performance for Emotion Prediction in Videos." Affective Computing and Intelligent Interaction (ACII), 77–83.

[15] Thorogood, M., Pasquier, P. (2013). ”Computationally Generated Soundscapes with Audio Metaphor.” Proceedings of the Fourth International Conference on Computational Creativity, Sydney, 256–260.

[16] Thorogood, M., Pasquier, P. (2013). “Impress: A Machine Learning Approach to Soundscape Affect Classification for a Music Performance Environment.” Proceedings of New Interfaces for Musical Expression, Daejeon, 256–260.

[17] Eigenfeldt, A., Bown, O., and Carey, B. (2015). “Collaborative Composition with Creative Systems: Reflections on the First Musebot Ensemble.” Proceedings of the Sixth International Conference on Computational Creativity, Park City, 134–143.