

AVIEM
LIBRARY OF EMOTIONS

SİNAN BÜYÜKBAŞ

SABANCI UNIVERSITY
Fall 2011

AVIEM
LIBRARY OF EMOTIONS

by
SİNAN BÜYÜKBAŞ

Submitted to the Graduate School of Visual Arts and Visual Communication Design
in partial fulfillment of the requirements for the degree of Master of Arts

SABANCI UNIVERSITY
Fall 2011-2012

AVIEM
LIBRARY OF EMOTIONS

APPROVED BY:

Elif Ayiter
(Dissertation Supervisor)

Yoong Wah Alex Wong

Lanfranco Aceti

DATE OF APPROVAL:

© Sinan Büyükbaş 2011

All Rights Reserved

Table of Contents

ABSTRACT.....	6
1. INTRODUCTION.....	8
2. THE PROJECT.....	9
2.1 The Methodology.....	9
2.2 The Creation of the AVIEM Library.....	11
2.2.1 How human emotions work.....	11
2.2.2 Emotive perception process of the audiovisual signals.....	12
2.2.3 Audiovisual perception.....	14
2.2.4 Audiovisual Synaesthesia.....	14
2.2.5. Difference in the Speed of Perception.....	16
2.2.6. How to design emotional signals in audiovisual media:.....	16
2.2.6.1 Theoretical aspects of emotional design.....	16
2.2.6.2 Audiovisual Metaphor.....	18
2.2.6.3 The audiovisual interaction of the AVIEM components of video [v6] and audio [a16].....	18
2.2.7 Cross-sensory Correlations (Parameter Mapping).....	21
2.3. The Interface.....	23
3. CONCLUSION.....	28
4. BIBLIOGRAPHY.....	30

AVIEM LIBRARY OF EMOTIONS

Sinan Büyükbaş

Istanbul, 05.01.2012

Abstract: The project AVIEM runs a comprehensive study on the relationship between audio, visual and emotion by applying the principals of cognitive emotion theory into digital creation. It focuses on the perception and creation processes to design an audiovisual emotion library and develops an interactive interface for experimentation and evaluation.

AVIEM primarily consist of separate audio and visual libraries and provides user a wide range of experimentation possibilities. AVIEM Library is formed by digitally created abstract virtual environments and soundscapes that are designed to elicit target emotions at a preconscious level. The interface allows users to create audiovisual relations and logs their emotional responses. AVIEM Library grows with user contribution as users explore different combinations between audio and visual libraries. Besides being a resourceful tool of experimentation, AVIEM Library aims to become a source of inspiration to build genuine audiovisual relations that would engage the viewer on a strong emotional level. Consequently, the project proposes various information visualizations in order to improve navigation and designate the trends and dependencies among audiovisual relations.

Keywords: *audiovisual, emotion library, abstraction, cognitive emotion theory, digital creation*

AVIEM
LIBRARY OF EMOTIONS

Sinan B y kbaŐ

Istanbul, 05.01.2012

 zet: AVIEM projesi, kognitif duygu teorisi prensiplerini dijital yaratım s recine uygulayarak ses, g rsel ve duygu arasındaki iliŐkiyi inceler. G rsel-iŐitsel bir duygu k t phanesi tasarlamak amacıyla algı ve yaratım s re lerine odaklanır; deney ve deęerlendirme ama lı bir kullanıcı aray z  geliŐtirir.

Keywords: *g rsel-iŐitsel, duygu k t phanesi, kognitif duygu teorisi, dijital yaratım*

1. INTRODUCTION

Judging the emotional stimuli of a situation is a daily action that we even do unknowingly. Emotional evaluation of these moments handled mainly by our auditory and visual perception, since they are the essential senses that we use to communicate. Zapping fast between TV channels, we make instant decisions on what these audiovisual signals make us feel (Fahlenbrach, 2002). Within a split second, our emotional perception evaluates the subjective attraction of the received multi-sensory messages, which differentiate between staying or keep zapping until we find an “emotional hook”.

This mysterious equation between audio, visual and emotion finds new platforms to be investigated while technology gives artists new mediums for creative expression. At the latest century, the production of affordable hardware for sound and moving image recording carried the earlier experimentations such as *Color Organs* at the 18th century one step further (Peacock, 1991). Avant-garde animators such as Oskar Fischinger¹ was among the first ones who experimented on the audiovisual synaesthesia with the advantage of analog editing, which enabled artists to make temporal connections between sound and moving image.

Not long before, computers set the new rules and gave users the ability to digitally create and edit audio and visual data. A new generation of artists learns to use audio-vision as a tool of artistic expression. Ranging from 3d projection mappings to virtual environments, recent audiovisual works cover a wide variety of genres. Music videos by directors such as Edouard Salier² and Alex Rutherford³ explore different approaches on how to meaningfully connect virtual reality with computer generated music to engage us into unique emotional experiences. A new form of underground art, VJing appeared on the club scenes, focusing on live performance by projecting computer generated visual materials to reflect the interactive emotional nature of the venue (Michael Faulkner, 2006, p.9).

¹ See Motion Painting No. 1, 1947

² “Massive Attack - Splitting the Atom” Music video. Accessed on 11.10.2011. <<http://www.edouardsalier.fr/#/fr/films/music-videos>>

³ “Autechre – Gantz Graf” Music Video. Accessed on 05.11.2011. <<http://warp.net/records/autechre/player/video/gantz-graf>>

The project AVIEM (Audio Visual Emotion), takes its route idea from VJing tools: experimenting on the audiovisual relation with an interface that enables the user to compare, arrange and create multisensory compositions. With the integration of a tag database, AVIEM forms an audiovisual library where the user can sort all single and multisensory components according to emotion tags added by users. All the visual elements are 3d generated abstract animations describing virtual environments along with digitally created soundscapes, which in the end form all together diverse and engaging emotionscapes. AVIEM puts emphasis on the creation process and strategic use of sound and image in order to catalyze strong emotional response on the viewer.

AVIEM Library aims to employ immersion and interactivity with its audio and visual elements designed to engage participant on an emotional level, thereby enhances the overall experience of the virtual environment. Participants are able to choose, watch and tag the previously created audiovisual compositions by selecting any desired emotion from the interface menu; thus AVIEM Library grows and evolves with user contribution.

As the technology evolves, we take one more step further to express ourselves free from creative limitation in order to find better ways to express how we feel, aiming at the true transmission and sharing of an experience one to another. In the end, the question comes to what Howard Rheingold insightfully stated, "If our technology ever allows us to create any experience we might want, what kinds of experience should we create?" [Rheingold, 1991, p.116].

2. THE PROJECT

2.1 The Methodology

AVIEM project began by focusing on the creation of audio and visual components in order to investigate what makes visual or aural stimuli intense and effective. The goal was to construct an effective and rich audiovisual library that would provide user a wide range of experimentation through its possible matching combinations among the visual and audio libraries.

Creation of such library in the context of academic research purposes requires a few ground rules. My first step was to decide that all moving image and sound should be abstract. The reason was to avoid cultural stereotypes. The ideology of a “non-narrative, non-discursive mode of expression” dated back to first color organ creations, eliminates “subjective interaction of all sensory perceptions” that is directly related to participant’s cognitive experience (Zilczer, 1987, p.101). A child's smile refers to hope or a red rose refers to love for the most of us. AVIEM aimed to convey emotions without repeating such "clichés" to seize accurate emotional responses. In the end, the first challenge was to elicit target emotions with the strategic use of very primary notions such as geometry, color, movement, timbre and pitch.

AVIEM profits from today’s cutting-edge CG tools that allow creating virtual environments that can be perceived almost as real. In contemporary computer simulations, immersion is mostly depending on the accuracy of lighting and shading. With the exploration of Global Illumination engine in 3d software, artists and designers have become able to create the most physically accurate virtual environments up to date. Immersion is how much we believe in what we see and hear in this case. Thereby, successful immersion of sound and moving image enhances the impact of the audiovisual stimuli, thus the emotive response of the participant is highly intense and lasting.

The duration of the audio and visual components are created no longer than 40 seconds in order to preserve the impact of the first impression, but also to give the user enough time to experience an emotive range of experience. Unlike most of the VJ samples and audio loops, AVIEM Library components are not designed to be synched to one another; knowing that temporally synched audiovisual relations increase the impact of the emotive response. However, this necessary sacrifice is crucial to obtain the diversity and interactivity of the library: numerous cross-modal matching combinations that ensure a fruitful experimentation for the user. In order to overcome lack of synchronization, each cross-modal component is designed with several dynamic and textural structures to encourage happy accidents of temporal synchronization. For example, when a visual sequence which have several jump cuts from different camera angles (referring to cinematic montage techniques), it is highly possible to witness numerous synched moments if it interacts with a sound of a similar nature that has ups and downs in timbre and pitch with an arrhythmic tempo.

2.2 The Creation of the AVIEM Library

2.2.1 How human emotions work

After setting the ground rules of the project, I began thinking on how human emotion works in order to create successful emotional signals. Emotions play highly important role on our mental state as they represent a synthesis of subjective experience, expressive behavior and neurochemical activity to help uniquely define our experience of reality. In his work *Rhetoric*, Aristotle suggests that an emotional response must be triggered by a certain pleasure or pain in order to qualify it as an “emotion”. However, “that is not to say that every person will feel the same pleasure or the same pain with any particular emotion, but if a feeling is to qualify as an emotion it must be attended by some physiological sensation of pleasure or pain” (Worth, 1998). One could think of the emotional response as a fingerprint of individual’s unique emotional perception. However we could not deny that there are several similarities in human’s emotional response when it comes to aggregate analysis of behavioral trends. We all shared a group laugh in a cinema or grief in a funeral; that clearly points out the emotive trends that we do respond similarly. Cognitive psychologists emphasize the role of comparison, matching, appraisal, memory, and attribution in the forming of emotions. Mark Johnson (1987), in his theory of cognitive metaphors, points out how much of human perception and thinking is affected by repeatable patterns of experience that include: motion, directness of action, degree of intensity and structure of causal interaction; what refers to namely “image schemata” (p.44). Johnson states that the meaning of balance –“the bodily experience in which we orient our selves within our environment”- creates a pre-conceptual data in our memory. This bodily experienced knowledge of balance also refers to the feeling of harmony (p. 74). In contrary, the state of “out of balance” corresponds to “disharmony”; that we know from our bodily experiences we feel fear when we lose balance. Thus, the theory of cognitive metaphors demonstrates that we share subconsciously emotional responses that are rooted in our physically experienced knowledge on reality. As I will example later, these mentally rooted visual and aural experiences are effectively used by the audiovisual media in order to create strong emotional stimulus on the viewer.

2.2.2 Emotive perception process of the audiovisual signals

But, how our sensory and cognitive emotional system works during the reception of the fast running visual and audio signals? They are evaluated mostly on three dimensions:

- Sensorial (Intermodal) processing of the audio-visual stimuli
- Cognitive evaluation
- Emotional experience

Recent theories on cognition and emotion points out that sensorial, cognitive and emotional process of the stimuli are strongly related to each other (Fahlenbrach, 2002, p.3). In this case, sensorial processing is directly related to audiovisual perception (that we will look in more depth later). Thus, it consists of parallel and simultaneous synthesis of two different sensorial data, in other words, cross-modal or intermodal processing (Stern, 1993, Marks, 1978). Intermodal processing provides the essential data for cognitive and emotional processing by interpreting diverse stimuli from audio and visual channels. Therefore it plays a crucial role on the emotional response emphasizing that the successful reception might rely on the harmonic perception of the moving image and sound which we will talk in more depth in cross-sensory correlations (Fahlenbrach, 2002, p.4).

Cognitive evaluation relies on the interpreted stimuli from diverse channels. Recent theories point out that the evaluation process starts with our cognitive attention; meanwhile our divided attention tries to use attentional resources on two or more stimuli. However, attention has limited capacity; the instant processing of the data from the two different sensory channels is specifically reinforced by the capacity of the brain (Fahlenbrach, 2002). On the other hand, attention is also related to individual's physical and mental well being at the moments of intermodal reception and processing. For example, tiredness, sickness or mental and environmental distractions can cause a state of emotional numbness and weaken the impact of the emotional response.

Grodal (2002) states that cognitive evaluation instantly processes the audiovisual data and starts to build "a web-like structure of associations" by producing semantic relations (p.64-65). He categorizes three mechanisms in this process:

- Establishing connections

- Chunking, grouping (making gestalts)
- Labeling

During the reception of audiovisual signals, the brain starts to mentally construct the meaning of what we see and hear; some elements are recognized consciously and create links to previous events of memory; at the same time the viewing and hearing activate many associative connections at a preconscious level. Therefore, the *meaning* that is processed as a result of this complex cognitive evaluation seems to be highly flexible and fluid as it depends on subjective and unconsciously predefined relations (Fahlenbrach, 2005, p.4).

On the level of emotional experience, the functional relationships between cognition and emotion are bidirectional (Lazarus, 1991). The processed meaning of the audiovisual signal generates an emotional response. “It is always a response to cognitive activity, which generates meaning regardless of how this meaning is achieved” says Richard Lazarus. Triggering an emotion can cause the activation of a subsequent thought or meaning, which would trigger other emotional responses. According to Carroll Izard’s differential emotion theory; “Emotion included cognitions may trigger complex memory clusters (attitudes, moods, beliefs) and cognitions may act as a positive feedback loop and amplify the ongoing emotional state” (Izard, 1977, p.100); as in anger, shame and fear mostly triggers each other.

Another important aspect of the emotional processing is the experienced *density of the stimuli*. It is primarily dependant on the interpreted data by *intermodal processing* referring to “how effective the audiovisual signal is designed”. It is also dependant on the individual’s coping potential: how much he or she can handle an emotional experience. Like in the TV channel zapping example, we make instant decisions according to subjectively experienced density of the stimuli. If the individual is getting too much information that he or she can process instantly at that moment, it is highly possible that he or she feels exhausted or uninterested in the audiovisual after a while.

As we see, the construction of the audiovisual meaning and the emotional response as a result are highly subjective matters depending on the individual’s personal luggage. However, researches on human behavior have indicated that a target emotive response can be produced through strategic employment of sound and imagery (Thompson, 1988; Watson and Rayner, 1920). For

the purpose of the project, next, I will focus on the audiovisual perception and examine the techniques for connecting sound and image in more abstract terms.

2.2.3 Audiovisual perception

The audiovisual language first sparkled within theatre and opera; Wagner's concept of *Gesamtkunstwerken* suggested a "total" multimedia experience which would create a multi-sensory media art form in order to change our understanding of what unifies sound and image in audiovisual works. With the developments in digital image and sound recording, computers literally accomplished Wagner's insightful vision; the digital media eliminated traditional distinctions between individual media through digitization of data from audio and visual channels into standardized series of numbers. Any medium can be translated into another with a total media link on a digital base that will erase the very concept of medium (Kittler, 1999). In cinematic experience, we acknowledge movies as functioning wholes, without making distinction between image and sound. They are designed to be perceived in the most immersive way to make viewer believe what is happening, by building meaningful audiovisual connections into intriguing stories. Thus, over years of practice and experience, the movie industry has developed its own audiovisual language by discovering many approaches on the audiovisual perception.

2.2.4 Audiovisual Synaesthesia

Russian filmmaker Sergey Eisenstein reports in his studies on film montage, how he interprets the phenomenon of "synaesthesia" to link image and sound to his shots in order to generate the exact emotional effect he wanted the audience to experience which he vividly imagined. He describes it as being "the ability to unite in one whole a variety of feelings gathered from different sources through different sense organs." (Robertson, 2009, p.143). Here, he refers to the audiovisual synaesthesia which, decades later, Michel Chion calls "synchresis" and describes as "forging of an immediate and necessary relationship between something one sees and something one hears" (Chion, 1994).

Even though, both definitions describe a simultaneous perception of the audiovisual, they reflect different approaches on the audiovisual creation. In the former, while filmmaker Eisenstein emphasizes on the impact of emotional design, in the latter, composer Michel Chion points out on the necessity of synchronization in order to forge and fortify what is perceived from two senses.

The word *synchresis*, formed by the combination of *synchronism* and *synthesis*, symbolizes how Chion uses synchronization as a tool for merging image and sound into an audiovisual unit. However, he stresses that this bond can also occur independently and without reflecting any logic. Early experimental audiovisual works verify Chion's finding; Oskar Fischinger combines his abstract animations with classical music and creates moments of synchronization between his colorful, geometric shapes and melodic lines with a metaphorical approach. Chion adds;

Play a stream of random audio and visual events, and you will find that certain ones will come together through *synchresis* and other combinations will not. The sequence takes on its phrasing all on its own, getting caught up in patterns of mutual reinforcement and phenomena of "good form" that do not operate by any simple rules.(Chion, 1994, p.63).

Later on, Bailey, Fells and Moody (2006) also suggest that *synchresis* may still function with abstract visuals without any contextual information. They suggest that the appearance of abstract elements might be perceived in synchrony if there is a temporal synchronization between events. Thus, the harmonic motion among the intermodal channels may allow the formation of consistent audiovisual relations regardless of cultural background.

The statement of Chion verifies a primary intention behind AVIEM; experimentation on the audiovisuals with random sound and imagery, and letting possible happy accidents of *synchresis* in order to observe a diverse array of mutual reinforcements of the "good form"; while Bailey, Fells and Moody support the use of abstraction in audiovisuals so as to generate emotions free from subjective associations with individual's cultural background. The point is that the sound makes a spontaneous and irresistible weld and enhances the perception of the image (or vice versa). Thus, viewer's brain intuitively creates a connection that makes the perceived more real. Chion calls this enhancement the "added value" (Chion, 1994, p.63). For example, consider the visuals and the sound of an explosion; we can separately recognize the signals from both senses thanks to our cognitive memory. However they never create individually the strong impact of an

explosion; because the brain needs more information about the nature of the event; if not, the immersion is incomplete and the emotional response is weak.

From this point of view, AVIEM library forms an efficient tool to investigate the effect of “added value” by simply allowing the user to choose between sounds that will adhere to a particular visual better than the others in order to elicit a target emotion. As a result, the most immersive combinations would be chosen and tagged by the users, hoping that further analysis on these successful results would give hints for new approaches on the emotional design of audiovisuals.

2.2.5. Difference in the Speed of Perception

The way how our auditory and visual senses process the signals individually also affects the audiovisual perception. By their nature, sound and visual have different influence speeds; meaning that “the ear analyzes, processes, and synthesizes faster than the eye” (Chion, 1994, p.10). This condition is all natural as a consequence that our primary communication sense is auditory since we are faster and better to communicate with spoken language in daily life. In *ordinary* listening sound is always used as a vehicle of meaning. However, human ear is able to isolate and focus on a detail of its auditory field. In example, we are able to select a sound source in a multi-instrumental song and follow the sound source changes in time by stripping our auditory perception so as to ignore the other sound sources. The important point is that we tend to perceive and follow the sound temporally unlike our visual sensory does. The eye perceives slower because; in order to follow a visual object our brain needs to relate its movement within a space; meaning that the eye has to process more information at once. Thus, the audio perception processes the data temporally (faster) while the visual perception does spatially (more information). This condition have a major influence on the design of the audiovisual signal; since we know that the sound signal is evaluated faster than the image and perceived temporally.

2.2.6. How to design emotional signals in audiovisual media:

2.2.6.1 Theoretical aspects of emotional design

In order to create audio and visual components of AVIEM library, I looked for previous approaches on emotional design in the audiovisual media that are applicable to the ground rules of AVIEM project. Although the early cinematic approaches are still in prevail, new theories that are mostly derived from VJ culture and music videos offer convenient solutions on notions such as short duration, intense emotional signal and the use of abstraction.

The complex nature of an audiovisual relation triggers divided sensorial attention forms bindings, matchings, groups out of particles of knowledge; builds a giant web of semantic associations; arouses filaments of feelings that pulls strings of primitive emotions and finally evokes genuine emotional response(s) or waves of emotional reactions that are felt within the individuality of the observer. However, one can seize, analyze, decompose and apprehend the very components of this enigmatic phenomenon in order to sculpt unique emotional scripts, cues and signals. Once the influence of the signal is impactful and it is received by many, it blossoms new emotive experiences, feelings, moods; incepts ideas, thoughts that trigger decisions, actions and reactions; and creates a chaotic source of experience and inspiration for its precedents. Thus, the spread of the audiovisual message is inevitable when it is well designed.

Researchers on cognitive emotion theory assume that every basic emotion consists of a *script* that is derived from “behavioral and social events for the best or most typical case of the emotion, the essence of the category” (Fischer, Shaver, Carnochan 1990, p.92). Therefore emotional scripts are associative schemas - as Nigel J.T. Thomas defines *schema* as a data structure implemented in the brain that functions to manage perceptual exploration of the world (Thomas, 2002) - that are both related to schema of bodily experiences and cognitive components of the individual’s memory. The emotional script structure is formed by “typical antecedents and responses, including behavioral, expressive, experiential, and cognitive components” (Fischer, Shaver and Carnochan, 1990, 92). Therefore, it is a highly situational and flexible structure that does not form consistent meanings of primary emotions; however it consists of “dynamic and generic aspects” of emotional experiences. Emotional scripts thus cover both cognitive and physical elements that are perceived and processed simultaneously (Bartsch and Hübner, 2005).

Looking at the topics covered on the audiovisual perception, it is crucial to acknowledge to successfully design an emotional signal - in such as music videos- that “primary visual and

acoustic gestalt patterns are related to an audiovisual network that is experienced first of all physically” (Fahlenbrach, 2005, p.8). As I pointed out before, this condition deeply relies on our sensorial and bodily based patterns of experience. Since music videos have limited time to develop emotional plots, they do emphasis on these sensory and physical elements of emotional scripts, to create brief, intense emotional signals (cues). These instant signals which give clues about the audiovisual message, establish undirected moods that are the basis of target emotions (p.6). Therefore, short duration audiovisual compositions such as music and promotion videos, VJ loops -and AVIEM library in this case- tend to use emotional cues redundantly in order to gain control over the viewer’s perception first of all pre-consciously, to have access all over the flexible emotion system of the viewer.

2.2.6.2 Audiovisual Metaphor

In her study “the emotional design of music videos: approaches to audiovisual metaphors”, Kathrin Fahlenbrach (2005) introduces her concept of *audiovisual metaphors*, derived from the theoretical aspects covered above, in order to analyze emotional nature of music videos. She defines audiovisual metaphors as “a metaphorical mode of audiovisual synthesis within which acoustic and visual schemas are projected onto one other in a metaphorical way” (p.8). In light of this concept of designing emotional cues with metaphorical approach, I will analyze a possible audiovisual coupling from the AVIEM library in order to describe the creational approach of an emotional signal design (cue) that physically and pre-consciously triggers the primary emotion of fear.

2.2.6.3 The audiovisual interaction of the AVIEM components of video [v6] and audio [a16]

In the visual library component [v6], a black, displacing surface covered with random plain spaces and rough furrows is described in a dark and pale environment. The triangulated faces of the structure are highlighted with thin white lines on the edges, allowing the viewer to perceive

randomly and instantly displacing faces on the surface. Dust and blur blends the horizon with a diffusing pale grey light from the below.

During the analysis of the visual event, I will ignore culturally cognitive metaphors in the scene, except bodily experienced schemas– for the purpose of the demonstration and also that they are assumed to be perceived subsequently according to cognitive emotion theory-; and focus on physical and sensorial signals that are perceived instantly and pre-consciously. The instant and asynchronous movements of the surface aim to directly attack to our perception of balance. As I mentioned previously, disharmony -referring to the bodily experience of *loss of balance*- pre-consciously evokes the emotion of fear likewise in its typical antecedents and responses. The highlighted lines on the black surface emphasize on a clean contrast in order to ensure an instant and successful reception of the visual signal- by decreasing the amount of the visual data to be processed, to a level just enough to perceive the asynchronous motion, but also to keep it immersive through its highly segmented interesting structure. Therefore, the movement of this unstable surface is a visual node, which could be intermodally mapped onto an audio channel with a similar disharmonic nature, in order to design an emotional script of fear. Audio [a16] in this example, fits the criteria with its asynchronous structure; the instant fluctuations in the sound intersect with the textured motion of the visual material and create accidental moments of synchronization. Despite that these acoustic and visual components are not meant to be perfectly in synch, they form a harmonic audiovisual relation through synchresis likewise Michel Chion (1994) insightfully stated (p.63). All in all, one modality of the visual sense, the nature of the motion is projected onto an acoustic schema through intermodal processing; in which they form an audiovisual metaphor for disharmony and loss of balance all together (Fahlenbrach, 2005, p.9).



Excerpt from AVIEM component video [v6]

Another audiovisual metaphor formed by the interaction of the components [v6] and [a16] is the cross-modal equivalence in terms of intensity. Along the dark atmosphere of visual material [v6], we see random flash lights in various moments of the video. These flashes are initially designed to stimulate the visual attention and increase the emotional tension in the scene; beside their cognitive metaphor to thunder lights in the sky, they accidentally create synchronization with the instant and fluctuating loudness of the sound (sounds of shattering glasses). In that manner, the intensity of brightness from the visual channel creates a metaphor with the intensity of loudness in the acoustic channel. The emotional meaning of this audiovisual signal is again triggered by a preconscious knowledge from the sensorial and bodily based experiences. In his studies on the psychological development of infants, Daniel Stern (1993) indicates that mother and child develop “common gestalt patterns” without a verbal interaction to communicate their feelings and emotions. From our childhood, we know that mother raises her voice when we did something wrong; the instant, raising loudness of the sound symbolizes the anger of the mother (action) and results mostly in fear; thus, an instant loudness mostly evokes the emotion of fear.

This audiovisual relation is familiar to most of us for that is also a well known effect in the cinema industry.

Although there are other numerous defining factors that affect the emotional response, the impact of illustrated emotional signals is arguable. Even though they create an accidental synchresis all together, most of the modalities are not in perfect synchronization, thus, hypothetically results in an unsuccessful immersion. However, knowing that the emotional result is not comparable to the diligent effort and keen intelligence of an artist's creation, it clearly reveals an important point: the accidental interactions through AVIEM introduce a genuine formula for cross-modal mappings, the essence of a possible approach in audiovisual emotion design.

The analysis demonstrates that, through synchresis, it is possible to create an audiovisual connection, - although it is not necessarily objective, will be perceived more or less the same by the audience- free from the majority of cultural and subjective associations (Moody, Fells and Bailey, 2006, p406). Inevitably, fundamental considerations of the human culture dominate the audiovisual meaning likewise Alves (2005) notes that the “literal mapping of pitch space to height [is] only intuitive because our culture has adopted that particular arbitrary metaphor of 'low' and 'high' to describe pitch.” (p.47); but even this knowledge, besides being cultural, is also a bodily based experience from our previous sensorial experiences. Thus, cross-sensory metaphors rooted on the physical world are more likely to success.

2.2.7 Cross-sensory Correlations (Parameter Mapping)

Apparently metaphorical mappings are not only related with gestalt patterns but also with other primary structures that affects our multi-sensory perception. The audiovisual interaction between the AVIEM components [v6] and [a16] demonstrates that certain qualities of the stimuli are perceived in both acoustic and visual senses. These qualities are commonly named as *amodal qualities* as they are present in all modalities of senses (Stern, 1993). According to the categories proposed by recent theories on neurology and developmental psychology, several qualities are as such;

- Intensity: strong- weak, loud –silent, bright- loud

- Rhythm / duration: fast-slow, rhythmic - arrhythmic
- Form / pattern recognition: moving – quiet, common – uncommon, harmonic – disharmonic, complex – simple, varied – redundant, contrasting – similar, symmetric – asymmetric

Moreover, there are many other qualities of image and sound that are not common in both senses, but effective in order to metaphorically connect image and sound. A well known example is the connection between color and timbre regarding its spiritualistic connection with emotions. All in all, the concept of audiovisual parameter mapping covers any metaphoric conversion between the two mediums as the metaphoric understanding is “the ability to perceive similarity among seemingly dissimilar objects” (Cytowic, 1993, p.207).

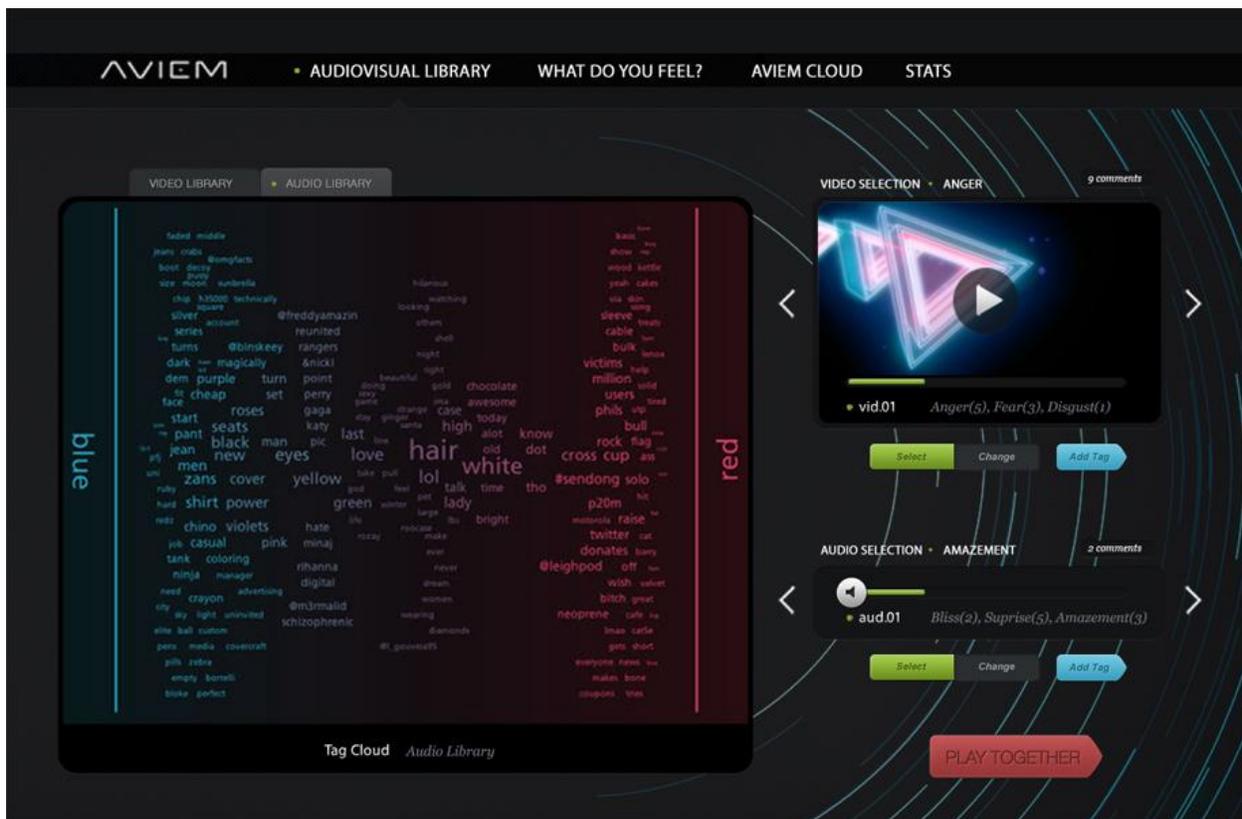
The success of the cross-sensory mapping is also due to its transparency. The interconnection between cross-sensory parameters is fully transparent when the receiver fully understands the nature of the audiovisual event without notification or explanation (Callear, 2010, p.5). In this respect, metaphorical connections that address specific cultural associations may guide to that level of mutual and transparent understanding. The bell sound that notices the opening doors of an elevator, the sounds that symbolize specific numbers on the phone pad etc. are such culturally formed audiovisual mappings that we come across in everyday life. For example, the elevator sound could be mapped onto a split-screen effect in the visual material, thus the visual schema of ‘opening doors’ would create a transparent connection to the movement of splitting screens. The transparency of the mapping clarifies the audiovisual meaning and provides a lucid reception; therefore the impact of the signal increases and may lead to a strong and memorable emotional response.

Another defining factor on the success of the audiovisual design may be the complexity of the signal. Unless they create a strong metaphorical connection, one to one mappings are likely to be less successful compared to several transparent mappings that would aim to increase the sense of ‘reality’. In daily life, we do not question “reality”, since our brain continuously receives multi-sensory signals that create our understanding of it. The smell that comes through the wind, the sound of the moving branches, flitting leaves in the air etc. together form the essential data to understand our surroundings. Likewise, hypothetically, the cross-modal relations between different modalities could be metaphorically mapped onto audiovisual parameters in order to

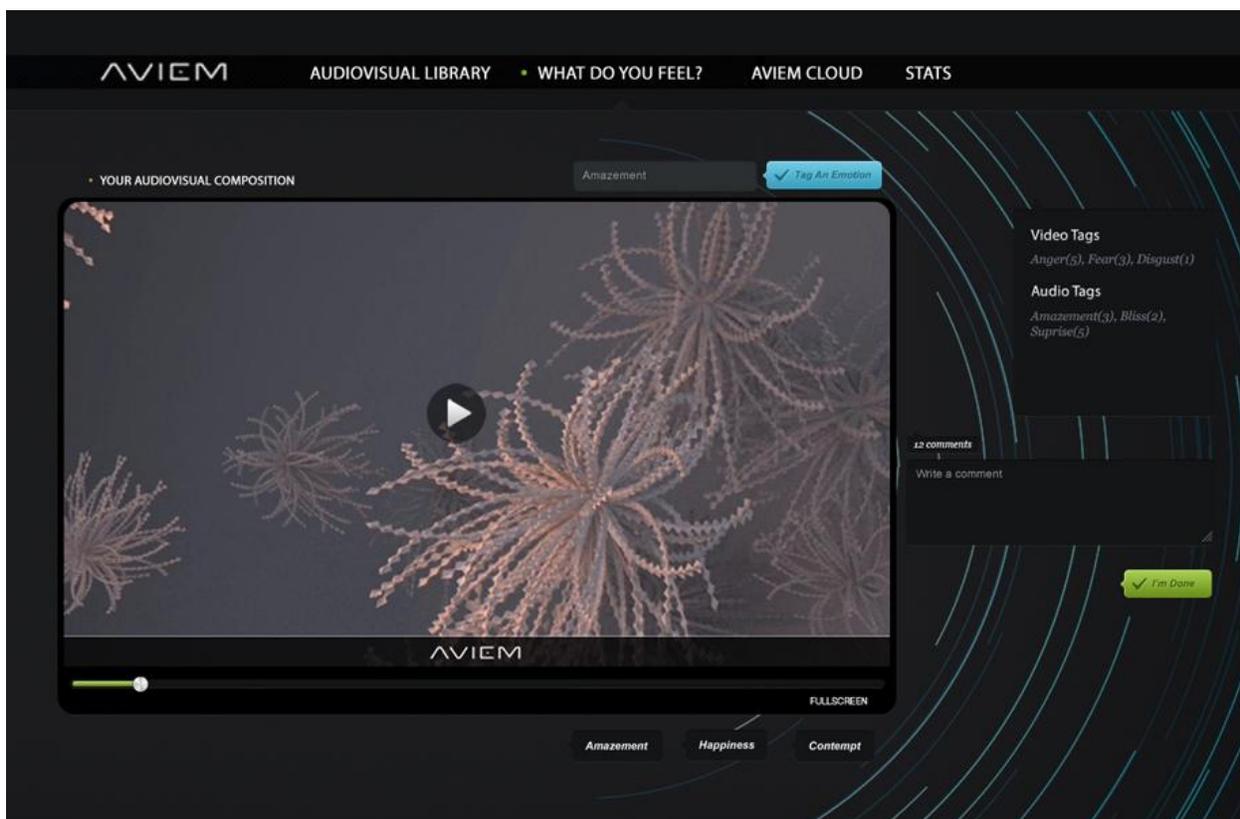
imitate what we perceive as real. However, the coping potential of the viewer must be always taken into account; too many and complex mappings will exhaust the viewer and break the level of attention.

2.3. The Interface

The interface design of AVIEM emphasizes on easy usage and accessibility in order let user to explore and experiment freely. Since AVIEM database (tag system) grows with user contribution it is crucial to secure the accurate evaluation by users. Thus, the interface aims to provide a comfortable navigation with its compact and lucid design elements. The interface consists of two main parts; “audiovisual library” and “what do you feel?” These two sections are the core of the interface where user can input data and provide statistical information for the other two sections which are “AVIEM Cloud” and “Stats”.



Audiovisual Library section consists of three parts. First, the tag cloud visualization, where the user can browse separately both audio and video library content according to emotion tags. Second, the video selection tab, in which the user chooses a visual component between the videos corresponding according to the emotion tag selected from video library tag cloud. In the third part, the audio selection tab works in the same manner as the video selection tab in order to define an audio component. The important point is that the user is able to sort and choose different emotions for audio and video components, which makes the final audiovisual evaluation interesting. The user may choose a visual component that addresses to “happiness” along with an audio component that is pre-tagged as “fear”. Hence, the opposing nature of selected audio-visual elements is subjected to user evaluation.



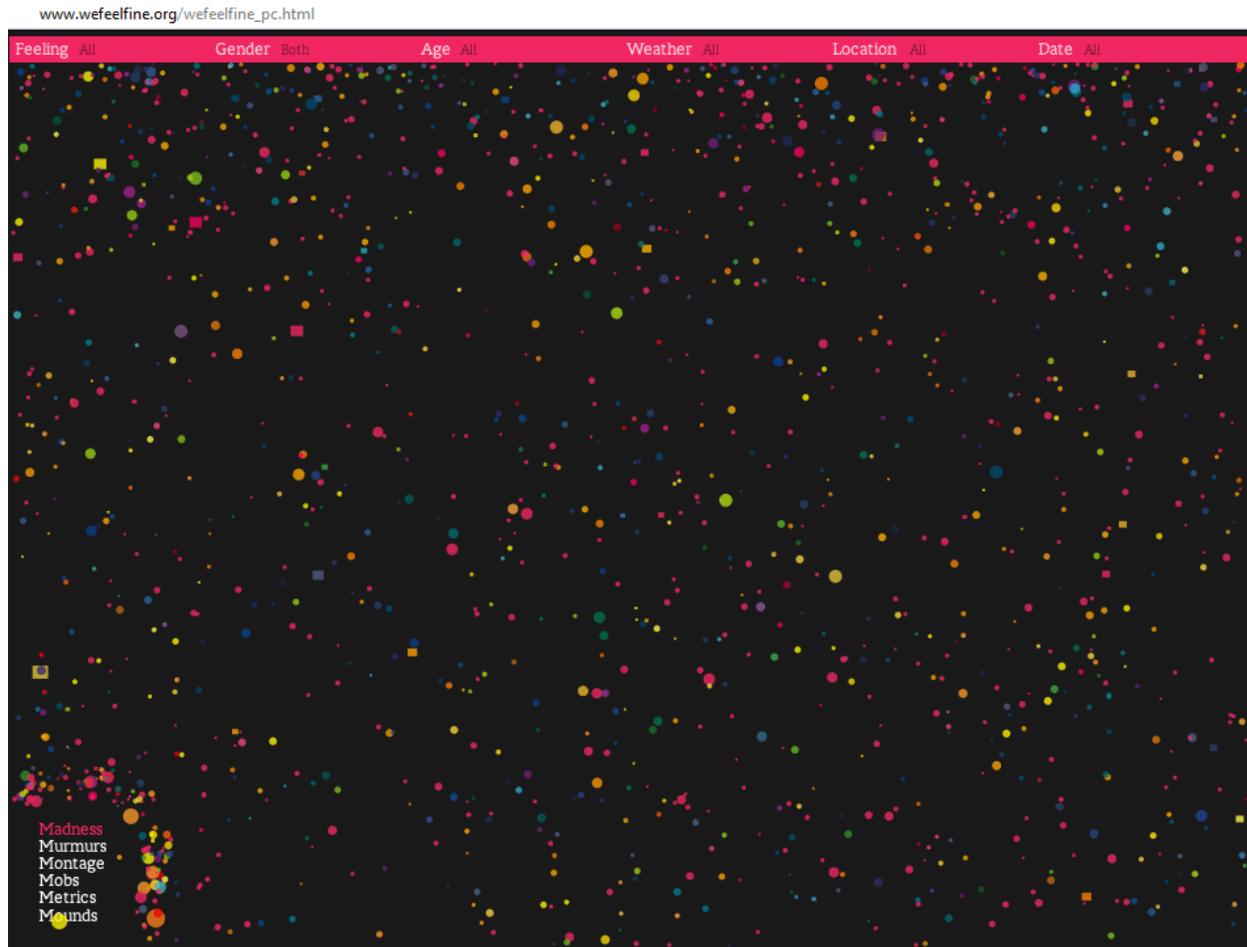
The tag cloud visualization and the emotion database are primarily designed to provide an easy navigation, but also not to affect the individual's subjective evaluation. The user also is able to add his/her preferable emotion tag in both video and audio selection tabs and contribute to the pre-tagged emotion databases. The pre-tag databases are formed during the testing process with the contribution of 30 individuals.

With the integration of tag cloud visualization, the statistic data from each audio and video emotion databases provide visual information on the emotive trends in both audio and visual libraries. The user can observe the quantity of the tags according to their font size. Hence, the tag cloud visualization enables the user to see most the dominant emotions in both libraries at a first look.

Once the user chooses the audiovisual components and hits the "play together" button, the second section, "what do you feel?", is launched. Here, the user experiences and evaluates his/her audiovisual creation. With the "tag an emotion" button, more than one emotion tags can be added or the user can describe his/her experience in the comment tab. Once the evaluation is completed the user hits the "I'm done" button and views a pop-up screen that summarizes all the information on audio, video and audiovisual tags.

AVIEM Cloud and Stats sections are theoretical propositions and currently inactive due to their complex nature of advanced coding. The both sections are proposing various approaches in order to visualize the statistic data and strengthen the visual identity of a "library". AVIEM Cloud section is the visualization of an immense library that consists of all the audiovisual combinations created by users. It provides an easy navigation in order to observe previously created audiovisual combinations. Thus, the user is able to view all the previous creations as a whole library, add tags and write comments. Consequently, AVIEM Cloud section provides a wide range of observation and evaluation feature and encourages the user to contribute more. The proposed visualization is inspired by "We Feel Fine" emotional search engine, developed by Sepandar D. Kamvar and Jonathan Harris from Standford University.⁴

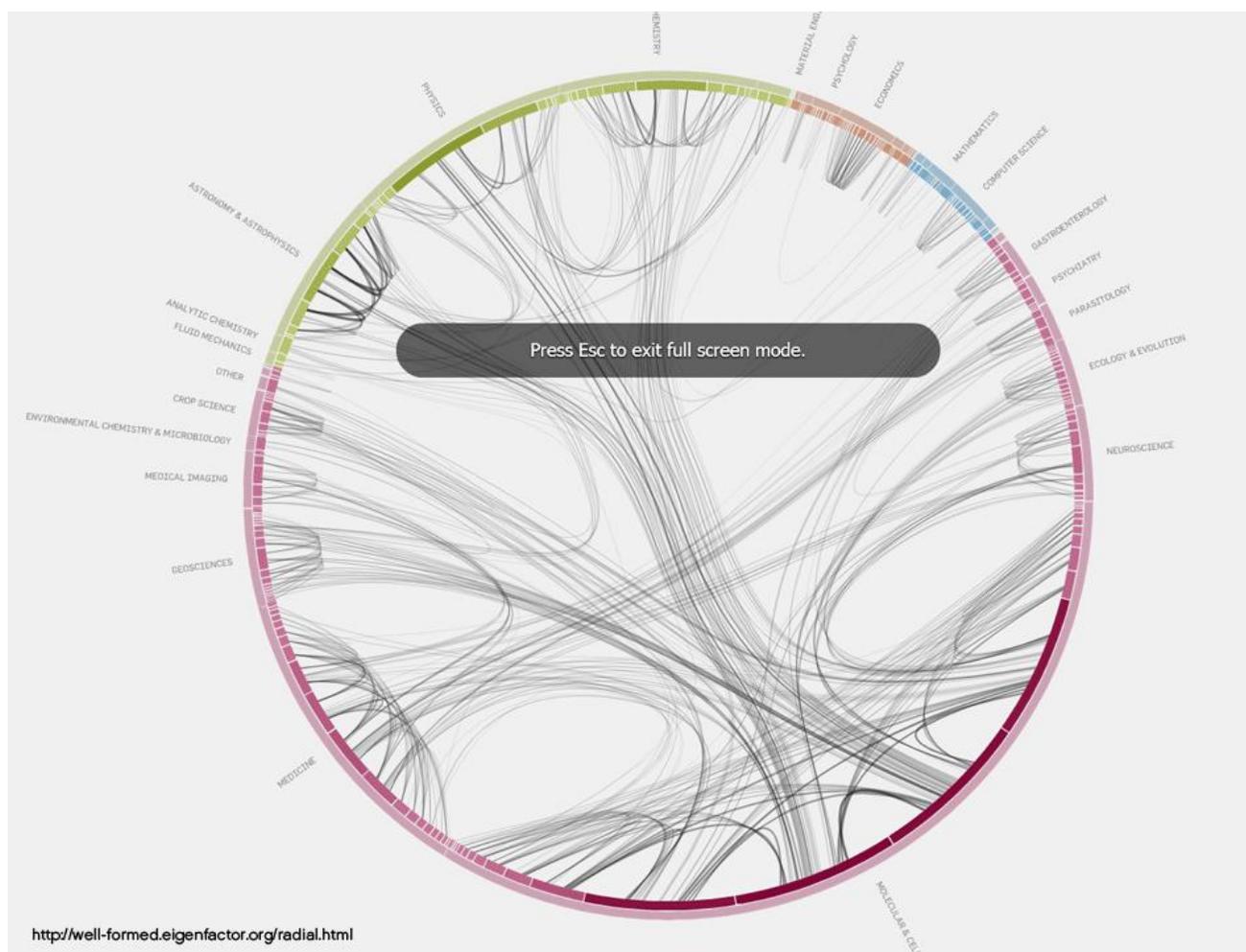
⁴ We Feel Fine emotional search engine, last accessed on 25.12.2011 from <<http://wefeelfine.org/>>



Excerpt from We Feel Fine emotional search engine

After creating and tagging his/her creation(s) the user will be able view his/her contribution to the massive pile of creation. The colored particles that symbolize audiovisual compositions would be selected randomly or sorted according to emotion tags. Here, the user is able to analyze and compare different approaches to create a target emotion. Therefore, AVIEM Library humbly hopes to provide a source of experimentation and inspiration for future creations in audiovisual media.

Lastly, the Stats section aims to visualize the emotional relations and trends in the AVIEM Cloud. The dependency graph ⁵introduced in can be used to visualize dependencies among classes within the AVIEM Library. The outer and inner ring sections would symbolize the primary and subsequent emotions signifying the quality and the quantity of the class in the database structure. When selecting an emotion tag, the Bezier curves would highlight related emotion tags and creations. Thus, the user would be able to see which emotions are selected together at most. When AVIEM library gathers a considerable sample in its database with the contribution of users, in theory, the statistic interpretation of the dependency graph may provide key knowledge on emotional trends used in audiovisual media creation.



Dependency graph visualization at well-formed.eigenfactor.org

⁵ Dependency graph visualization, last accessed on 05.01.2012 from <http://well-formed.eigenfactor.org/radial.html>

3. CONCLUSION

All in all, the AVIEM project covers both the creation and evaluation processes and seeks the unrevealed possibilities of digital creation in audiovisual media design. It combines the very principals of cognitive emotion theory with the most up-to-date digital creation techniques in 3d animation and sound design. It provides experimental knowledge on audiovisual media design and aims to encourage exploration of new techniques emphasizing on the infinite possibilities of today's digital creation tools.

In the creation process, AVIEM investigates how to elicit a target emotion using digital abstraction in order to transmit audiovisual signals at a preconscious level. It gives examples on how to trigger an emotion by building associative audiovisual relations that are derived from repeatable patterns of experiences such as motion, directness of action, degree of intensity and so forth. In this respect, digital creation tools are essential for audiovisual design, since they are based on the very same principals. Current creation tools provide the user full control over the amodal qualities; various parameters such as intensity, rhythm, form can be fine tuned and mapped on to each other in order to design impactful audiovisual mappings. Moreover, they provide the most immersive digital renders up-to-date. Consequently, digital creation tools provide control over the key features for successful immersion and strong emotional response. Creative narration is also one of these key features, since the viewers gain an emotional numbness after consuming similar audiovisual media over and over in daily life. In this respect, AVIEM library offers a wide range of experimentation and aims to become a source of inspiration for designers and artists that are eager to explore new ways of conveying emotions.

In the evaluation process, AVIEM introduces an interactive audiovisual interface where the user is able evaluate audio, visual and audiovisual relations separately. The statistical data of the user evaluation are stored in each libraries database and visualized in order to reflect various aspects of AVIEM libraries. Such information visualizations are essential to analyze emotional trends and dependencies of the libraries. The interface is designed user friendly in order to reflect the comfortable venue of a library and also guide user through an uninfluenced evaluation.

In conclusion, AVIEM project and interface run a comprehensive investigation (creation, experimentation and evaluation) on the relationship between audio, visual and emotion in digital

creation. With further developing and user contribution, AVIEM Library aims to become an experimentation and inspiration source and a valuable data resource for cognitive emotion researches.

4. BIBLIOGRAPHY

Alves, B. (2005). "Digital Harmony of Sound and Light". *Computer Music Journal*, 29(4): pp. 45-54.

Bartsch, A. and Hübner, S. (2005). "Towards a Theory of Emotional Communication". *CLCWeb: Comparative Literature and Culture*, 2(4):
<http://docs.lib.purdue.edu/clcweb/vol7/iss4/2>

Callar, S. (2010). "Audiovisual Correspondence: An Overview". A modification of PhD progression assessment submission, retrieved from
<http://stephencallar.files.wordpress.com/2010/03/audiovisual_correspondence1.pdf> on 05.11.2011.

Cytowic, R. (1993). "The experience of metaphor". In: *The Man Who Tasted Shapes*. G. P. Putnam's Sons: New York, pp. 206-10.

Chion, M. (1994). *Audio-Vision: Sound on screen*. New York: Columbia UP.

Fahlenbrach, K. (2002). "Feeling sounds. Emotional aspects of music videos". *Proceedings of IGEL 2002 conference*, Pécs, Hungary.

Fahlenbrach, K. (2005). "The emotional design of music videos. Approaches to audiovisual metaphors". *Journal of Moving Image Studies*, 3(1): pp. 22-28.

Faulker, M. (ed). (2006). *VJ: Audio-Visual Art and VJ Culture*. London, UK: Laurence King Publishing.

Fischer, Kurt W., Shaver, Phillip R. and Carnochan, P. (1990). "How emotions develop and how they organize development". *Cognition and Emotion*, 4(2): pp. 81-127.

Izard, Carroll E. (1977). *Human Emotions*. New York: Plenum Press, p. 100.

Kittler, F. (1999). *Gramophone, Film, Typewriter*. Palo Alto, Calif.: Stanford University Press.

- Lazarus, R.S. (1991). "Cognition and motivation in emotion". *American Psychologist*, 46(4): pp. 352–367.
- Moody, N., Fells, N. and Bailey, N. (2006). "Motion as the Connection Between Audio and Visuals". Retrieved from <http://www.niallmoody.com/phdstuff/downloads/MotionAsTheConnectionLong.pdf> on 20. 11. 2011.
- Peacock, K. (1991). "Famous early color organs". *Experimental Musical Instruments*, 7(2): pp.1 and 17-20.
- Rheingold, H. (1991). *Virtual Reality*. New York: Summit Books, p. 116.
- Robertson, R. (2009). *Eisenstein on the audiovisual: The Montage of Music, Image and Sound in Cinema*. New York: Taurus Academic Studies.
- Stern, D. (1993). *The Interpersonal World of the Infant*. Stuttgart: Klett-Cotta.
- Thompson, J. G. (1988). *The Psychobiology of Emotions*. New York: Plenum Press.
- Thomas, Nigel J.T. (2002). A Note on "Schema" and "Image Schema". Retrieved from <http://www.imagery-imagination.com/schemata.htm> on 19.11.2011.
- Watson, J.B. and Rayner, W. (1920). "Conditioned Emotional Reactions". *Journal of Experimental Psychology*, 3(1): pp. 1-14.
- Worth, S. E. (1998). "Music, Emotion and Language: Using Music to Communicate". *Twentieth World Congress of Philosophy, in Boston, Massachusetts*.
- Zilczer, J. (1987). "Color Music: Synaesthesia and nineteenth-century sources for abstract art". *Artibus et Historiae*, 8(16): pp. 101-126.