Audio-visual hybrids – between immersion and detachment

Sabine Fabo

Close interrelations between sound and image are not a mere phenomenon of today's multimedia technology. The idea of the synthesis of different media lies at the core of the concept of the Gesamtkunstwerk in the second half of the 19th century and it can also be traced back to the synaesthesia debate at the beginning of the 20th century. At that time the development of colour pianos and other colour instruments reached its climax, with musicians such as Alexander Skrjabin, Alexander Laszlo, Vladimimir Baranoff Rossine, Alexander Wallace Rimington, Thomas Wilfred, or Ludwig Hirschfeld-Mack and Kurt Schwerdtfeger at the Bauhaus.

Yet, the wish to combine sound, image, movement and language was accompanied by a strong romantic idealism that envisioned a transfer from the union of sensory experiences to nothing less than the spiritual enlightenment and the unification of mankind itself. Religious expectations were shifted to an aesthetic level that relied on an advanced media technology. Skrjabins *Mystère,* 1914*,* aimed at the redemption of mankind achieved by the experience of synthesized media. Thirty years later Ivan Wyschnegradsky designed a light temple where colour fields were activated by music. He developed the multimedia project as a means to evoke the sensibilities of his time:

> "The final aim lies in the idea of a work of art that achieves the unification of all arts and that at the same time is capable to provoke a healing choque and thus evokes the powers of a cosmic consciousness in man, that lie at the heart of the deepest unconscious."[1]

In the beginning of the 20th century the project of the fusion of media was not restricted to mere experiments in art and technology. Against the background of the depressing experience of World War I the early days of multimedia fulfilled the role of a substitute for religion. Only in the synthesis of media could the utopia of a universal language prosper. Kandinskys "The Spiritual in Art" paralleled the artists' endeavour to reveal common structures between colour and sound on a theoretical level.

The development of experimental instruments was accompanied by public presentations and concerts. Both, the spiritual elevation of man and the entertainment of an audience constituted parallel intentions.

---

[1] *Vom Klang der Bilder. Die Musik in der Kunst des 20.Jhdts*, Karin von Maur (Ed.), München 1985, p.282

The emergence of digital technologies has brought the multimedia experiments of the beginning of the 20th century into focus again. The different media can now be processed, combined and synchronized by the same binary abstract code. Images and sound are interpreted as “material” that can easily be accessed and shaped. Thus the concept of a unification of all arts seems to have found its ultimate tool.

Digital media syntheses of the nineties share the search for a common structural language of all media. On the basis of the digital code auditive and visual data undergo smooth transformations that keep media in a state of constant flux. The early nineties concentrated their acitivities on the experimental synthesis of sound, image and performance , appealing visual results were often neglected in favour of an intermedia discourse. Michael Saup made experiments with *Interfacing Audio and Images*, 1990-1993, where digitalized video images were manipulated by an acoustic input. The level of image manipulation was made transparent. Instead of achieving a complete work of art, the concept of the media work in progress was prominent, with all its clumsiness of testing and probing.

In today’s work with audio-visual hybrids two attitudes might be observed as central working philosophies:

The immersive approach that endeavours to entertain and embrace the public in an audio-visual stream of fusing media. The continuing development of digital technologies has intensified the smooth multimedia transformations, which is accompanied by a tendency to overcome the difference between the diverse media involved. Where the merging of media has reached the field of popular entertainment the technology behind the synthetical magic is deliberately held out of focus. Here the synthesis of media functions as a means to evoke multisensory sensations and to create a state of well-being.

The approach of detachment demonstrates a certain scepticism towards a multimedia enthusiasm. Here the discussion on intermedia has continued to be prominent. Media characteristics are paralleled, the in-between of media is investigated and rendered visible, but difference is not regarded as an obstacle to the experience of wholeness. On the contrary, the aspect of media difference plays a vital role, whereas holistic philosophies and their implication of addressing the spectator totally are viewed with a certain reservation. Instead, the process of combining media is part of the work, inconsistencies within the “material”, ruptures and the irritation of the audience are emphasized to heighten the critical awareness of the spectator.

Regarding technical implications, both approaches rely their work on the same digital basis, the computer is the prominent instrument to synthesize sound, image, text and movement. The crucial difference lies in addressing the technical process as part of the content of the work or in dismissing it completely in favour of creating pure sensations devoid of their technical background.

To distinguish between these attitudes does not mean to simplify the diversity of multimedia art or to make evaluations within the range of multimedia spectacle on one side and critical art experiment on the other side. But the concepts of immersion and detachment help to clarify the actual role and understanding of digital hybrids.

Immersive approaches

In 1997 Toshio Iwai and Ryuichi Sakamoto resumed the idea of the colour piano in a collaborative performance which received the Ars Electronica Golden Nica . The title of the work itself *Music plays images x images play music* alludes to the reversibility of digital sound and image. Based on Iwais *Piano as image media* (1995), a grand piano that is controlled by MIDI signals, image objects were closely connected to the keys of a piano. 88 image objects, mainly abstract graphic images, that corresponded to the 88 keys of the keyboard, were transformed and projected onto large semi-transparent screens. The computer as the technological tool behind all these synaesthetic revels remained mostly unseen. The illusionism was reinforced by generation of sound-image parallels in real-time. Graphical light effects were rendered visible according to duration and loudness of the notes, live images were added to the range of visual initiators/impulse of music. Though the synthesizing technology demonstrated an advanced state of the art, the message of the installation calls to mind the idealistic views of the early days of the colour piano: According to Iwai, the media devices should “raise people’s consciousness, sensibilities and creativity to even higher levels.”[2]

The artist Golan Levin has focused on the design of systems for the creation and performance of simultaneous image and sound, within the more general framework of communication and social dialogue. Levin has developed several multimedia installations where the process of sound-image interaction can be generated by the audience. Levins *Audiovisual Environment Suite* (2000) consisted of five interactive systems where a correspondence between an abstract animation and synthetic sound could be created in real

[2] ZKM (Ed.), Hardware, Software, Artware. Die Konvergenz von Kunst und Technologie. Kunstpraktiken am ZKM Institut für Bildmedien 1992-1997, Karlsruhe, Ostfildern-Ruit 1997, p. 90.

time. Visually, organic metaphors prevailed: The application *Aurora* produced diffusive cloudlike figurations of colour whereas *Floo* showed entwining plants that grew accordingly to the movements of the user.

> "Ideally, these systems permit their interactants to engage in a flow state of pure experience. The AVES systems are built around the metaphor of an inexhaustible and dynamic audiovisual »substance», which is freely deposited and controlled by the user's gestures. Each instrument situates this substance in a context whose free-form structure inherits from the visual language of abstract painting and animation. The use of low-level synthesis techniques permits the sound and image to be tightly linked, commensurately malleable, and deeply plastic."[3]

The idea of synthesis is achieved on the level of media and is further emphasized by making use of organic imagery. Media and nature are visually interwoven to create the experience of an integral whole. Levins latest work *Messa di Voce*, 2003, is an installation " in which the speech,shouts and songs produced by two vocalists are radically augmented in real-time by custom interactive visualization and sonification software".[4] Using speech analysis algorithms and head-tracking systems, the installation establishes close connections between speech, sound and graphical images that move and change according to the singer's loudness or the quality of the vowel (which has an influence on the colour, see the application *Fluid*). The installations' main interest lies in "phonesthesia, or phonetic symbolism". Close relations to the visual arts are established by the Messa segment "Rothko", whereas *Pitchpaint* offers a direct control of the visualization of the voice, descending or rising pitches qualify the direction of graphic lines whereas colour is generated by the quality of vowels.[5]

Organic metaphors are often employed at the point of synthesis where different media meet. The media involved are often described within the terms of transformation, fluidity or "liquid material", whereas the field of technical connotations is deliberately avoided. Despite a complicated set of apparatus the product of these machines, the hybridization of media, is understood in terms of a parallel to human life. Real-time is a necessary requirement to overcome the time gap during the process of data transfer. The illusionary integration into the sphere of colour and sound depends upon a dense and effective connecting with the digital machines, a considerable delay within the man-machine interaction might interfere with the quality of immersion. That seems to be one of the main reasons why multimedia artists like *Station Rose* refer to real-time as ""the precious moments of real time & how holy they are" [6], whereas the German art critic Bazon Brock, regards the electronic recording of

---

3 Golan Levin, "An Audiovisual Environment Suite", http://acg-media.mit.edu/people/golan/aves/

4 tmema, blonk, la barbara, "messa di voce", PDF, http://www.tmema.org/messa.

5 Ibid., p 12.

6 Station Rose*, Private.//Public. Webcasting * netSTReams *realtime* , Wien 2000, p.9.

data as a "resurrection machine"[7]. "That is the deeper sense of real time – to reach out for timelessness and eternity".[8] It seems, as if the romantic concept of overcoming physical boundaries via an aesthetic experience can easily be reconciled with new technologies. The affinity to the religious interpretations of multimedia to the beginning of the 20th Century is obvious: "Electrons to us are what angels were to the medieval age. These angel-like creatures will allow us to concentrate on the implementation of the virtual."[9]

A critical, almost cynical comment on multimedia entertainment was delivered by the multimedia group *Bauhouse. Bauhouse* are part of the multimedia entertainment business and they are fully aware of it. Within the context of the German Preview for the European Song Contest *Bauhouse* was invited to combine the songs and presentations to a dense audiovisual synthesis. The *Grand Prix Ultimativ Mix,* 2001, did not really meet up with entertainment criteria. Instead the mix demonstrated bluntly the joy that lies in exerting power over the images and their transformations. The process of image-scratching revealed a cynical attitude towards the often smiled-at German popsong-mainstream. Parts of the songs were fragmented and endlessly repeated, until their modest narrative content was completely depreciated. Further the song performances were interrupted by images taken from commercials thus demonstrating the arbitrariness of the manipulation and syntheses of different media. Proteine capsules worked themselves into the song fragments whereas inserts of the cutting of meat alluded ironically to the structural cut of sound and images. The availabilty of all material, visual and audio data alike, was displayed with great relish, with the artist acting as a cynical master of show ceremonies. "Entertainment recycle", as *Bauhouse* has defined it.

Analytical attitudes towards audio-visual hybrids

The work of the multimedia group *Granular Synthesis* (Kurt Hentschläger, Ulf Langheinrich) concentrates on the research of sound- image configurations. The title of the group directly refers to the technical programme of fragmenting audio-visual data into minimal units and, in a second stage, to restructrure and synthesize them. The outcome of their explorations has been presented in installations and live performances since 1992. Prerecorded as well as live images are controlled by audio-sequencing software, the audio-video samples are generated by a specially developed software VARP 9.

---

7 „Station Rose in conversation with Bazon Brock", in: Station Rose, p.57.
8 Ibid., p.66.
9 Ibid., p.58

> "VARP 9 is proprietary audio-video re-synthesis software, developed by Granular Synthesis and Dirk Langheinrich in 1997-1999. It is a sample based MIDI software instrument that streams from RAM. It works in realtime in full PAL video resolution and frame rate. It allows accessing and allocating single frames at 25f/sec. At this video grain rate it is capable of audiovisual granular synthesis. The structure of the program resembles in many of its parameters an audio sampler. The specialty of VARP 9 is its flicker engine, which can mix 2 discrete video/sample streams frame by frame, as if 2 videos were intersecting each other."[10]

Though audiovisual material can be addressed and changed on a microlevel any idea of smooth transitions is denied. The images are disturbingly vibrating on a highspeed level (as in *Model 5*, 1994-1996), accompanied by a machine-like sound that works itself subsonically into the bodies of the audience, with approximately 20 000 watt, as in *POL*, 1998-2000. The impact of these "immersive works", as Granular Synthesis call it, lies in their almost aggressive formulation of synaesthesia, the audience is exposed to subliminal sensual irritations and physical strain, experiencing their physical boundaries. The popular notion of multimedia as the creation of overwhelmingly colourful moving images embedded in harmonious sounds is supplemented by a minimalistic environment where the changes of the images can hardly be perceived but the more physically felt.

Carsten Nicolais method of media work might be defined by the concepts of assembling and composing. Though Nicolai works as painter, installation artist, composer, DJ and editor of the raster-noton label, he rejects the notion of multimedia artist. Instead he speaks of a strategy of diffusion.

> "What I am in fact working on are only fractions and fragments, something which a laboratory situation makes possible. I start out with puzzles and different subject-matters, I sort them into groups, concentrate them and sample them. At the same time I am aware of the fact that it will never be possible for me to make complete sets out of them".[11]

Carsten Nicolai has been active in the field of music and art alike, with a strong interest in the dialogue between these media. Nicolai emphasizes the technical preconditions of his work. Electricity is one of his subjects functioning as aesthetic counterpart to entertainment electronics where the media impact has already been defined. Sound is sometimes produced by sine wave generators thus harking back to the early days of electronic sound. The installation *Wellenwanne* makes use of amplified CD-sounds that produce vibrations on the water of the wave baths, 2001. In *Telefunken*, an installation from 2000, audiosignals constitute the direct input to a TV thus creating austere images of electronic wave contractions. Nicolais visualization of sound is devoid of any graphical transformations of audio input, instead it follows the minimalist beauty of oscillographs, as can be seen in the

---

10 "VARP 9", http://www.granularsynthesis.info

11 Carsten Nicolai, *Polyfoto*, Nürnberg 1998, p. 76

installation *Model for visualizing sound through the effect of magnetic fields on an electronic beam*, 2001. Nicolai has composed several pieces of music that are collected on a CD that gave its electronic input to an electronic-beam tube. The dynamics of the beam corresponds to the intensity and pitch of the music. Nicolais working philosophy takes its point of departure from the idea of the fragment that despite all constructive efforts will never face integration into an all-complete whole. His profound scepticism about multimedia synergies formulates itself in the juxtaposition of media, that comprises analogue and digital media alike:

> "I have simply attempted to build up duality, different poles, which I then dissolve or split up again. I am interested in these small pieces, the residual fragments which are left when something is broken up. I tend to identify myself with conflict rather than unity, since conflict offers an open field where I can unite any number of elements in my artistic work: films, my feelings on architecture, music, painting."[12]

Actionist Respoke, an interactive musicclip, that now has changed its title to *Passivist Destroke,* is constructed on 30 small sound samples taken from the song *Actionist Respoke* by the experimental electronic group Mouse On Mars. At first glance *Passivist Destroke* , 2002, fulfills all requirements of an interactive interface where user actions are promptly rewarded with a feedback. But after a while acoustic troublemakers appear, according to the speed in which the user acts: the less the user does, the more intrusive become the disruptions of the interactive system. Any hesitation is answered by small microforms and letters that literally feed on the fruits of interaction and multiply. The music is slowed down, partly distorted, beats aggressively take over the interactive composition. The possibility of an interactive production of musical hybrids is questioned by the constraints of the programme, the claim for "user-friendliness" is confronted with a more complex understanding of user and machine. Simple 1:1-relations of user action and feedback, where sound and image fragments match harmoniously to a complete composition, are qualified against the idea of displacement and work in progress.

If audio-visual hybrids create "a moment of truth and revelation", as McLuhan had claimed, cannot be answered here.[13] Nevertheless, audio-visual hybrids constitute a background against which our understanding of sensory experience and its intensification via media can be discussed. Within the context of multimedia the synthesis of sound and image reveals high expectations from the side of artists and audience alike, digitally achieved immersion is almost endowed with a religoius impetus. At the same time the fusion of sound and image is

---

[12] Ibid, p.77

[13] „The hybrid ort he meeting of two media is a moment of truth and revelation from which new form is born. (..) The moment of the meeting of the media is a moment of freedom and release from the ordinary trance and numbness imposed by them on our senses" , Marshall McLuhan, *Understanding Media. The Extensions of man*, New York 1964, p.63.

seen as a critical demarcation line where romantic views of the whole are countered by a more analytical aesthetics that emphasizes media transparency and difference as constitutive qualities of our attitude towards modern technology.