KUNGL TEKNISKA HÖGSKOLAN
Royal Institute of Technology
Numerical Analysis and Computing Science

CID-79, KTH, Stockholm, Sweden 1998

Constructing and Manipulating the Virtual: Gesture Transformation, Soundscaping and Dynamic Environments for Extended Artistic Performance

John Bowers, Sten-Olof Hellström, Kai-Mikael Jää-Aro,
Jonas Söderberg, Bino, Pär Hansson and Lennart E Fahlén

CID
Centre for
User Oriented IT Design

Author: John Bowers, Sten-Olof Hellström,
Kai-Mikael Jää-Aro, Jonas Söderberg, Bino, Pär Hansson and Lennart E Fahlén

Title: Constructing and Manipulating the Virtual:
Gesture Transformation, Soundscaping and Dynamic Environments for Extended Artistic Performance Report number: CID-79

ISSN number: ISSN 1403-0721
Publication date: May 1998
E-mail of author: bowers@nada.kth.se
URL of author: http://www.nada.kth.se/erena

Reports can be ordered from:

CID, Centre for User Oriented IT Design
Nada, Dept. Computing Science
KTH, Royal Institute of Technology
S-100 44 Stockhom, Sweden
telephone: + 46 8 790 91 00
fax: + 46 8 790 90 99
e-mail: cid@nada.kth.se
URL: http://www.nada.kth.se/cid/

DELIVERABLE 2.2

Constructing and Manipulating the Virtual

Chapter 1

Introduction to Deliverable 2.2

John Bowers
Centre for User-Oriented IT-Design (CID), Royal Institute of Technology (KTH), Stockholm, Sweden
Department of Psychology, University of Manchester , U.K.

This is Deliverable 2.2 of the eRENA project. It is the outcome of work conducted in Task 2.3 of Workpackage 2. This Workpackage is concerned with various techniques for extending artistic performances so that an 'electronic arena' could become a viable performance space for the arts. It is this emphasis specifically on artistic performance, with its normal presupposition of an audience co-present with one or more performers, which makes the Workpackage distinct from others in the project. In a gallery space (the extension of which is the concern of Workpackage 1) or in broadcast television (Workpackage 3), there is no necessity for this kind of co-presence. What is at stake specifically in performance and its extensions is uniquely the concern of Workpackage 2.

Considerations such as these caused work in Task 2.3 to be rethought to some degree. In its original formulation, Task 2.3 was focused around a particular group of researchers-artists (those associated with the Membrane series of performances) and a particular kind of technology (tracking technology for the capture of performer movements). As initial work on eRENA began to make clear, this would not give the Task a unique focus or contribution. For example, tracking technologies are discussed *ubiquitously* in eRENA, not just in Workpackage 2. Indeed and furthermore, within Workpackage 2, it became clear that important work was being conducted, concerned with performance and based on tracking technologies, in Tasks 2.1 and 2.2. The duplication of this work or its pale imitation would be inefficient. In addition, historical work at the ZKM as well as their ongoing experience in staging an ambitious multimedia opera (both reported in Deliverable 2.3) was emerging from Task 2.4 indicating a series of important artistic, practical and technical challenges for contemporary work in extended performance.

For these reasons, the work at KTH sought constructive alternatives to the work within Tasks 2.1 and 2.2 while addressing some of the challenges for the notion of extended performance emerging within Task 2.4. In this way, we could maximise the perspectives on extended performance within eRENA and minimise the duplication of technological development, while rising to some of the artistic challenges raised in work emerging from the ZKM. Work within Task 2.3 was, then, first prosecuted through the development of a performance piece, *Lightwork*, which has been uniquely developed and performed within eRENA and presents a very different picture of possibilities for extended performance than does the work reported in Deliverable 2.1 (the combined outcome of

Tasks 2.1 and 2.2). Following this, a cooperation between KTH and the Swedish Institute of Computer Science (SICS) was engaged in to plan a mini-opera, *Mimoid*, one component of which, *The Mimetic Blob*, has been constructed and performed within this first year of eRENA. *Mimoid* also represents a cooperation between the EC's i3 project eSCAPE and eRENA. Again, *Mimoid* has a number of features which make it distinct from other work on performance in Workpackage 2. Members of the Membrane group of artists-researchers have been responsible, in various combinations, for the work on these two endeavours.

The comparison of the various approaches to extended performance in eRENA is properly the concern of Deliverable 2.3 but some anticipation is necessary here to introduce the two chapters which follow and to further legitimate the claim for variety in the Workpackage which led us to replan Task 2.3.

Lightwork is an improvisatory piece of performance art involving the combination of live experimental electroacoustic music with the real-time construction and controlled navigation of virtual environments. Its emphasis on improvisation makes it distinct from the *Cyberdance* performance, which forms the focus of Deliverable 2.1 and which was very carefully choreographed and rehearsed in advance. The performers involved in *Lightwork* see an essential tie between improvisation and interactivity. An interest in improvisation is one reason precisely why interactive technologies are used and why virtual environments are constructed and navigated on the fly and why pre-rendered animations or video material are not used by them. Indeed, this emphasis on improvisation cuts quite deep in *Lightwork* as whole virtual environments are constructed within the performance. This shifts the focus of interest in virtual reality and performance art away from a notion of 'virtual scenography' and, arguably, marks a potential point of departure in this work from a dramaturgical-theatrical tradition. Similarly, if virtual reality technology is not to be used to generate virtual scenography, then the status of a performer is not that of a real human-being acting before a scenographic illusion. Rather, as *Lightwork's* title suggests, performers become worker-constructors building a virtual environment from sound and light. The emphasis is on displaying their actions and their actions' effects rather than on displaying performer character or expression—again in stark contrast to the *Cyberdance* performance. Furthermore, as in *Lightwork* performers stand in relation to virtual environment (in both visual and sonic aspects) as builders to artefact, there is no necessity to represent the performer (or any other source of agency) as an embodiment within the virtual world. Thus, again in clear contrast to *Cyberdance*, no avatars, either humanoid or otherwise, appear in *Lightwork*

The concern for the real-time construction and navigation of virtual environments has brought up some fundamental issues in the design of interaction for complex technologies. For example, are methods of 'direct manipulation' (or DM as it is sometimes abbreviated), though 'classical' in contemporary graphical workstations, appropriate for supporting artistic performance in an electronic arena. The authors of *Lightwork* suggest an alternative which they refer to as 'algorithmically mediated manipulation' which brings into focus the transformations which occur between user-performer gesture and its effects. These transformations can be many and varied. A set of performer gestures could be 'compressed' into a compact representation (e.g. if many gestures were required to bring about some outcome in a virtual environment) or 'expanded' into a multiplicity of effects. In either case (and both can be found in how the interaction techniques have been designed in *Lightwork*), the algorithm connecting identified gestures to their effects in performance becomes a topic for research focus, and, indeed, a topic for aesthetic deliberation as some ways of transforming performer gesture might lead to 'ugly' results. An open and controversial question, which is returned to in the contributions to Deliverable 2.3, is whether such 'indirect' techniques make for a satisfactory audience experience as it may be less than obvious to an audience how the gestures of a performer in such an approach have the effects that they do. Again, a contrast exists with the work demonstrated in *Cyberdance* where rather closer couplings between performer gesture and effects were demonstrated—at times through the copying of a human performer's dance by an avatar.

Mimoid shares this concern with algorithmically mediated interaction and goes further in suggesting a 'smart layer in interaction control'. That is, they hint that the insistence on transforming performer gesture should be given potentially a system architectural recognition, i.e. 'taken out' of the application for explicit attention. If followed through, this could make the VR platforms of the future rather different from those of today in allowing richer support at the platform level for the (potentially radical) transformation of human input. The question of the nature of the audience experience when interactive effects are potentially complex is also considered in the work developing *Mimoid,* and a specific desideratum is suggested. Audiences should "experience *the surface*... and maybe *get a sense* of something more complex and involved lurking behind" (emphasis given here). In other words, these novel suggestions for interaction techniques are seen as having associated with them *new roles and experiences* for audiences. Perhaps, the challenge of 'deciphering' the logic behind the interaction techniques could be part of the point of a work and/or finding exactly the right kind of 'surface' which would give hints of the complexity that 'lurks' will turn into a major technical and aesthetic challenge. Both of these possibilities, suggested in the proposals for *Mimoid* and—it is hoped—exemplified in *The Mimetic Blob,* indicate different kinds of audience experience, artistic artifact and interaction techniques from those explored elsewhere in Workpackage 2.

Mimoid is also concerned with developing strong concepts for the interaction with and control of sound. Too often, music is subordinated to visual experience in interactive performance art. Perhaps a familiar piece of music will be used or, at any rate, one which has an existence independently from the work in question. The situation is very different in *Mimoid* as it has been conceived of as an opera and, furthermore, related interaction techniques are proposed for the manipulation of sound as are for the manipulation of visual features of the virtual environments in the piece. In other words, a thoroughly 'spatial' approach is given to sound—not merely to ensure an interesting diffusion in a multi-speaker sound system, but to 'navigate' multiple sound sources. In short, a principle of 'soundscaping' can be identified in the proposals for *Mimoid* and an example is given (the so-called Eurocentric sound compass) of symbolic associations which can be assigned to sounds of different types relating them to spatial directions. (Similar emphases on soundscapes can be found in *Lightwork,* where some attention has been also given to questions of 'synaesthesia', i.e., in this case, the perceptible correlations between seeing and hearing. A spatial approach is used in *Lightwork* to relate sounds to certain visuo-dynamical features of the virtual environments which are constructed and projected in the piece.)

The two chapters which follow discuss in turn *Lightwork* and *Mimoid.* Both chapters, in addition to describing the pieces and projects involved, contain their own suggestions for future work. We also provide several 'galleries' of images from these pieces as often these can make clear the results of the techniques discussed better than further description. Future possibilities within eRENA with these works, and others in Workpackage 2, is further discussed in Deliverable 2.3.

Note: Chapter 2 of this Deliverable is a much expanded and revised version of 'The *Lightwork* Performance: Algorithmically Mediated Interaction for Virtual Environments' published as a 'short paper' in the Proceedings of CHI98, Los Angeles, USA, New York: ACM Press.

Chapter 2

Making *Lightwork* : The Algorithmic Performance of Virtual Environments

John Bowers[1,2] Sten-Olof Hellström[1] Kai-Mikael Jää-Aro[1]

[1]Centre for User-Oriented IT-Design (CID), Royal Institute of Technology (KTH), Stockholm, Sweden

[2]Department of Psychology, University of Manchester , U.K.

Introduction

In this chapter of Deliverable 2.2, we describe an improvisatory performance art work called Lightwork.—a 15 minute long piece combining live electroacoustic music with the real-time construction and navigation of projected virtual environments (VEs). Lightwork embodies, we believe, a number of innovative human-computer interaction concepts for managing interaction within VEs, as well as exploring some novel uses of VEs for artistic purposes. Work on Lightwork has combined artistic, social and computer science skills, building on our existing experience of VEs and studies of them [1, 2].

Central to our interest in developing technologies for performance art is that this provides one of the most testing contexts for computer system development. Naturally, the highest standards are to be met for visual and sonic design but also reliable real-time system-performance is essential to an effective piece. The 'users' of such technology—the performers themselves and their audience—are also highly critical and demanding people, who are unlikely to be reticent if the interaction experience is unsatisfactory. All these features present challenges often met in only diluted form when research results are publicly appraised in a 'demo' format.

Just as we can benefit from performance art as a rigorous 'target domain', so—reciprocally—can performance and installation art profit from innovative interaction concepts. The prevalence of so much 'push-button' interaction in CD-ROM art, for example, suggests to us that new interaction principles need to be explored. Indeed, a core principle of our work is that interaction design can be an aesthetic matter and that techniques should be developed for their aesthetic value as well as for technical feasibility.

Perhaps this is most strongly felt in artistic applications of VR research. Many familiar VR interaction concepts and devices are not well suited for performance art applications. Performances tend to require a large public display interacted with by means of gestures which are themselves public to the audience. This means that both 'immersive' and 'desktop' VR techniques and devices are rarely appropriate—besides head-mounted displays and other VR accoutrements are rather clichéed in a performance art context. Finally, some of the debated interaction issues in VR are even more strongly experienced in performance contexts. For example, overshooting one's destination while navigating and getting 'lost in cyberspace' and losing fluency of interaction might be disastrous for all concerned.

As we shall discuss, our approach to interaction in *Lightwork* has been strongly influenced by the requirement to develop techniques which are relatively 'safe' in application, allowing performers latitude for spontaneous activity, (even) error and (yet) recovery, without some blatant discontinuity in performance being necessary. This is strongly necessitated by our dual focus on both real-time improvisation and complex VEs and, we would argue, leads us to conclude that some form of departure from classic Direct Manipulation (DM) human computer interaction techniques is required in this context.

Lightwork: Some Aesthetic Themes

In several respects, the aesthetic themes of *Lightwork* can be regarded as 'user-requirements' that our technical development work has to fulfil. It is worth spending a little while discussing these as they also offer some novel perspectives, we believe, one the potential uses of VEs within artistic work.

We wished to construct what might be called a *combinatorium*, a (virtual) place of combination, where image, sound, text and three dimensional virtual objects can be juxtaposed. In so doing, we wished to explore a principle of *infinite collage,* whereby multiple elements could be combined into new virtual forms on the fly and a mobile viewpoint onto these forms would permit an indefinitely large number of samplings of the result.

De-sourcing/re-sourcing: much of our image, text and sonic material is sourced from everyday experience and then processed in various ways. The image material, for example, comprises of mostly digital photographs of found (i.e. not composed) scenes and objects collected and manipulated by John Bowers, classified into thematic image archives such as 'architecture', 'machine', 'plantlife' and so forth. These themes were also followed by Mark Jarman, who provided a series of short three-line texts, again composed of found textual material (dictionary, encyclopaedia or museum catalogue entries, fragments of 'lost literature', public information announcements, and so forth). Similar principles were also observed in the compilation of sonic material by Sten-Olof Hellström. In all cases, although image, sound and text may have a definite origin, further processing or manipulation makes them into *raw materials* for combination in *Lightwork*, disassociated from their original sources so that they can 're-source' virtual world construction.

Multiple coexisting perspectives: the combination of two dimensional pictorial material with virtual forms enables some experimentation with the perspective cues suggested by the images projected in performance. For example, a digital photograph containing its pictorial depth cues may be juxtaposed with those suggested by virtual objects and the whole scene may have an angular 'scaffolded form' (see below) within it as well. In this way, the perspective cues indigenous to standard 3D computer graphical rendering can coexist—not always consistently—with other forms of visualisation.

Improvising the world: we are concerned with the construction of a series of visual and sonic VEs *as* the performance. While the process of VE construction has been a theme of interactive art before (e.g. in Bill Seaman's 1996-1998 piece *The World Generator*, an eRENA Deliverable, 1.2, in its newest incarnations), to our knowledge, the notion of 'performing virtual worlds' is innovatory here. However, familiar techniques for VE construction hardly make for apt performance gestures. It is not engaging to watch someone edit VRML files or interact with a 3D modeller no matter how flamboyant their gestures are! Thus, in *Lightwork*, performers manipulate VE interaction *algorithms*. They do not directly manipulate VE content. Again, algorithms have been used to generate VEs in installations (e.g. Marcos Novak's 1995 *transTerraFirma*) but not yet, to our knowledge, applied in real-time in artistic performances.

Lightweight assembly: contrary to what one might intuitively suppose from what is suggested by the onerous image of 'world construction', we wished to make interaction with *Lightwork's* VEs depend on gestures of little effort. In performance, an electronic wind instrument player is responsible for interacting with the graphical world construction, animation and navigation algorithms. In a sense, then, the building of a world is achieved through the exercise of this performer's breath! Equally, it is our intention to explore a range of contact (e.g. enhanced joysticks) and non-contact (e.g. Theremin-style proximity sensing) controllers for interaction with the sonic elements of *Lightwork*. It is intended that movements against relatively little resistance (e.g. the 'resistance' of the air in the case of non-contact devices) might manipulate whole masses of sound. In this way, we wish to be true to the word play in the title *Lightwork*: a work made of light, done lightly.

Paradoxical devices: by using a musical instrument to determine VE interaction and devices usually employed for VEs (e.g. joysticks) to control music, we reverse conventional associations. The intention here is to explore the boundaries of what is 'intuitive' in gestural control, and of audience expectation, while opening up possibilities for 'synaesthesia' as the basis for interaction [3].

Narrative from within. Our aim was to produce an improvised work which would be most likely to be quite different on different occasions of performance. At each juncture, performers would have considerable choice in how to continue with graphical VE construction or manipulation of sound. However, we did not want to make this choice boundless (as if it ever could be—there are always limits to the free-est of free improvisations). As suggested already, image, sound and textual material is *prepared* and 'archived'—it is the combination of pre-existing material that is improvised. Equally, as we shall see, there is a finite and *pre-designed* set of algorithms which are interacted with, although algorithm choice is ad hoc—as are the exact moments when they are triggered. Finally, we wanted to design a performance in which a certain internal coherence could be achieved as the performance itself unfolded in time. In other words, we wanted to encourage a sense of narrative but without mandating a fixed, pre-decided structure, which would effectively be to give up on the improvisatory status of *Lightwork*. We formulated the phrase 'narrative from within' to capture what is intended here and below we describe how we attempted to support this.

Lightwork: Algorithmically Mediated Interaction

Lightwork is an exploration of *algorithmically mediated interaction*. Interaction is mediated by algorithms (rather than following a principle of the Direct Manipulation, DM, of world content), some *constructing* VE content, others governing the *animation* of virtual objects, yet others controlling the *navigation* of the viewpoint around the VE.

Some of these algorithms, in their abstract specification, were developed as part of the eSCAPE project by the first author of this Deliverable and represent an instance of co-operation between this other i3 Inhabited Information Spaces project and eRENA. Deliverable 2.1 of eSCAPE is, in part, concerned with the algorithmic generation of large-scale VEs, and eSCAPE project document escape-man-2 discusses some of the algorithms mentioned below (specifically `chamgen`, `immersiveForm/formModulator`, `scaffolder`, `plenumbulator` and `THALES`) in more detail (so q.v.). To make these algorithms suitable for application within *Lightwork*, numerous further developments have had to have been made in eRENA, including their calibration for interactive use, their implementation within the DIVE environments (see below), modifications to enable them to interwork with *Lightwork's* image, text and sonic material. In summary, while the general approach of algorithmically generating large-scale VEs has been discussed in eSCAPE, the implementation of algorithms in a specific artistic application and the investigations of 'algorithmically mediated interaction' as an orientation to human-computer interaction of general interest are specific to eRENA.

VE Construction and Animation Algorithms

A *Lightwork* VE is composed of several elements which can change independently with new material being created and deleted on the fly. With the exception of the VE's background (which is a series of images which change about every 80 seconds), each element is algorithmically generated, and—in the case of animated elements—has algorithmically controlled behaviour. For each algorithm, its parameters are intended to correspond to perceptually obvious features of the material generated by it.

For example, 'enclosures' which tend to surround the viewpoint can be generated by a chamber generation algorithm called `chamgen`. `chamgen` will create VE content resembling a room with various objects protruding from its walls. A performer can influence the size and regularity (e.g. the range of rotations and stretches) of the protuberances but the precise values which enter the VE world model are calculated algorithmically by sampling from probability distributions. On each occasion of its use, `chamgen` will map mages from an image archive onto the forms comprising the chamber. The 12 thematic image archives are drawn from in fixed rotation. At the centre of the chamber will be placed the three line text corresponding to the theme of the images.

Another algorithm is available for a rather different kind of enclosure (`immersiveForm`), where the viewpoint can be surrounded by a complex form repeatedly folded in on itself. Again an image is mapped to this form, selected from the next image archive in the ordering.

Various further 'forms' can be placed within enclosures. `scaffolder` creates strongly angular forms by aggregating 'pipes' onto each other. The overall size of the form and the parameters influencing how the aggregation takes place can be influenced in performance. These pipes are composed of 3D computer graphical material (essentially using a 'cylinder' primitive) without mapping on images, but rotating around blue, green and red as the form's emissive colour.

`formModulator` takes a sphere and distorts it non-linearly to generate complex 3D shapes. In fact an analogue of the equations for frequency modulation, FM, familiar from radio and sound synthesis applications, is used. The size of the form and degree of distortion are parameterised by performer gesture (the same algorithm is used for `immersiveForm` but parameters are scaled differently to give different sizes of objects and relations to the viewpoint in the two cases).

`plenumbulator` fills up the enclosure with image and text material according to a constrained random distribution where the amount and density of material are the main parameters. Finally, `THALES` creates a set of objects which orbit each other in a nested manner (orbits within orbits)—here the main parameters fix the distribution of orbit radii and cycle times. Textual material (e.g. the names of the thematic image archives) and image material are both used.

Navigation and Viewpoint Control

In *Lightwork* the VE projected to the rear of the stage-space is given by the view along a path which is computed in real-time by means of a selection from two algorithms. One employs a modulated sinusoidal function which generates periodic orbits. The function has been selected so that circular, elliptic and a family of 'looped' paths (e.g. figures of eight and three and four-leaf 'clovers') can be generated by different settings for it. The formula used is:

r = m + n*sin(t*Ø)

This formula should be interpreted so that r gives a radius from some notional origin and Ø corresponds to an angle, i.e. a polar coordinate representation of instantaneous position in a 2D plane is being used. m gives a notional radius which, when n is set to zero, defines the extent of a circular orbit. The term sin(t*Ø) modulates this notional circle with t related to the 'modulation frequency' and n giving the degree to which this modulation impacts upon the form of the path. In performance, any or all of these parameters could be under interactive control—as could be the rate at which the function is updated (i.e. the speed along its path).

It is through these features (rather than, say, pointing in a desired direction) that movement is controlled. We hope an easier task to manage in performance than the manipulation of a 6DOF interaction device. It is also intended that a bounded periodic function should be well suited to exploring enclosed VEs (e.g. m can be set in relation to the size of the enclosing form, room, chamber or whatever). However, as we shall discuss below, the calibration of such a function becomes a critical task.

(It is to be noted that we have used a formula which describes paths in a (r, Ø) 2D plane. In usage, we have subjected this plane a slow tilt, which—like the procession of background images—is not under interactive control. Naturally, formulae to extend the principle of navigation by means of parameterising a path equation into 3D could be explored. However, we wished to maintain a sense of 'up-down' by constraining motion. We also wanted to keep the number of parameters manageable.)

Another algorithm is available to approach/retreat from the centre of the VE. The viewpoint moves along a fixed radius from the centre of the VE. For both algorithms, the direction of gaze can be adjusted. In the case of the orbiting view, gaze can direction can be either in the direction of rotation or towards the centre. In the other case, gaze is along the radius, either inwards towards the centre or outwards away from the centre. One of the performers, in addition to selecting the viewpoint control function, 'toggles' the gaze direction.

Whatever function is selected for viewpoint control, this also influences the diffusion of some parts of *Lightwork's* specially composed electroacoustic music through a four speaker sound system. Sound sources are associated with static objects in the VE and the navigation function is used to compute their relative location in the soundspace.

Sound Generation

Sound material was prepared for *Lightwork* again following the principles of algorithmically mediated interaction. We employed a number of sound synthesis and processing techniques and a feature of most of them is that a very large number of parameters need to be specified for sound generation to take place. For example, many of the sounds employed in *Lightwork* are the product of sound synthesis by physical modelling where, in the Yamaha VL-1 synthesiser we employed, an string/air column model can be manipulated in real time in software. The model requires the specification of a vast number of parameters and, in order to find interesting and usable sounds, some methods of simultaneously controlling this large parameter set has to be found. An application named SO-2 was developed using the MAX multimedia programming language from Opcode Systems to manage, transform and combine the parameters. The composer-user defines a profile of dynamical tendencies for parameter change for four sounds. These profiles are represented as loci at the corners of a 2D square region within which the composer can make mouse movements. Within the square, linear interpolations between the four sets of dynamical tendencies are made and the result is realised as sound. Movements make for continuous sound transformations. Regions are also available outside of the square made by the four sets of dynamical tendencies where various kinds of further transformations can be defined (e.g. reflections, extrapolations or thresholding). A figure showing one of the screens from SO2 can be found in Appendix 2.

In addition to material generated using physical modelling techniques, a number of location recordings collected during a visit to Gottland, Sweden, Summer 1997 were also assembled. In total, 60 different 8 channel sound files were prepared in this way.

Performing The Interactive Narrative Machine

To date, *Lightwork* has been realised by two performers placed either side of a conventional stage-space with the VE projected between them. One performer, S, improvises a response to the projected VE by processing and mixing sonic elements further to those whose spatialisation is given by the viewpoint function just described. A number of interaction devices and techniques are being explored for their suitability for use by S. In Deliverable 6.1, we discuss various gesture processing techniques and contact and non-contact devices we have prototyped. However, in actual performance up to the time of writing, we have confined ourselves to the use of conventional MIDI-faders. Again, MIDI data does not directly control the realisation of the sound. Rather, it drives and interacts with a network of four chaotic algorithms which, in turn, generate streams of values which govern mixing, filtering, pitch modulation and spatialisation in *Lightwork's* multi-speaker environment.

The other performer, V, interacts with the algorithms which generate visual VE content by playing an electronic wind instrument—the Yamaha WX-11—and using footswitches on a Yamaha MFC-10 MIDI foot controller to trigger VE modification and select the algorithm to be used.

V's playing is analysed by a program called 'The Interactive Narrative Machine' (INM) which converts the WX-11's MIDI data stream into parameter values. The INM works with three 'time windows' which can compute level and variability statistics for the last 20, 100 and 500 notes. Three attributes of V's playing can be reported on: pitch, loudness (MIDI-velocity) and timings between notes. Selections from all these statistics are mapped onto the VE algorithms. For example, short-term (20 note window) timing values are used to control navigation (faster playing causes speeding up, syncopation yields increased 'loopiness' in the path). Some forecasting is employed to adjust for the 'inertia' which would occur in basing current values only past activity. Other mappings involve 'narrative rules' which define how past performance statistics get further transformed to generate future values. For example, one rule might specify that long-term loudness statistics define the size and regularity of `chamgen`'s protruding objects, such that if V has been playing loudly, the next chamber will contain small objects. In this way, the performer can *systematically* respond to existing and *predictably* generate new material. Playing loudly to big objects and quietly to small ones leads to an alternation between small and large protuberances. In this way, the INM is intended to enable the temporal unfolding of *Lightwork* to be improvised through performer activity within the piece itself—narrative from within.

The level and variability statistics for the 20 note window are automatically updated every 10 notes. This means that the parameter values determining viewpoint control are the most frequently updated and updated without V have to command them to be so. The statistics for the other two, longer time windows are updated on demand. For example, if—as described above—V wishes to insert new material into the VE generated by `chamgen`, he will press a designated foot pedal. This action will cause loudness statistics to be computed for the last 500 notes. These, after appropriate transformation defined by the INM's narrative rules and some scaling, will then be passed as parameter values to `chamgen`. A new chamber will then be inserted into the VE. This makes for some efficiency savings as statistics on performer activity are only calculated when an algorithm that needs them is triggered.

To avoid the VE becoming overly cluttered, the use of the foot switches is interpreted to toggle between VE content insertion and deletion. For example, if there already is a chamber in existence in the VE, pressing the chamber foot switch will cause its deletion, otherwise it will initiate the calculations just described. Some further dependencies also enable cluttering (and hence overall system performance) to be controlled and to enable V to perform a number of related actions in one go. For example, the insertion of material using the `chamgen` algorithm will cause any existing material in the VE generated by the `immersiveForm` algorithm to be deleted. In other words, at most one kind of 'enclosure' can exist at any one time. In summary, V's selections of algorithms are interpreted in a *context sensitive* way in the light of the material that already exists in the VE. These context sensitivities, and the dependencies within the 'architecture' of the VE which they correspond to, all facilitate V under time-critical conditions and help avoid excessive clutter and potential errors on V's part in trying to eliminate it.

Technical Implementation Details

The algorithms for construction, animation and navigation are run as an application within the DIVE [4] VR system (see also `http://www.sics.se/dive/`) on an SGI Octane. Parts of the object generation and viewpoint animation is performed by a C program receiving input from the INM while other parts involve Tcl scripts contained in pre-existing objects in the VE reacting to events raised by critical changes in the environment. The INM has been authored in Opcode's MAX programming language which gives language-level support for handling MIDI data. Two Apple Power Macintoshes are used for sound processing and running the INM. Communication between machines is quite lightweight as, for example, parameters are passed to algorithms only intermittently. RS232 serial and MIDI communications have been explored to date.

Experience in Performance

Lightwork was performed in its first version, after just three months of highly concentrated development time, on 16th December 1997 in Stockholm with the first two authors as respectively V and S. Development work was being undertaken right up to the moment of performance and some decisions had to be rapidly taken, on occasion, it must be admitted, with a degree of arbitrariness.

For example, although aesthetically we had the principle of 'narrative from within' to guide us and a number of features of the INM application had been developed to support this, we were still unsure exactly how to start and finish the performance! We decided on the idea of 'laying out' the materials as a way of introducing the work for the first minute or two, with the main body of the performance consisting of their improvised combination as described, before finishing the piece with the materials being laid out again. This idea was implemented by flying through selections from the image archives together with some textual material against a black background. A straight trajectory was followed deterministically (i.e. non-interactively) for 90 seconds before entering a sphere upon which the first background image was texture mapped. A small animation sequence then turned the viewpoint from this straight trajectory to one controlled by initial default values for the viewpoint path function above. Once this turn was complete, the algorithms became available and the performance proper could commence. After a fixed period of some 12 minutes 'within the combinatorium', the viewpoint was turned again using a 'reverse' of the above mentioned animation sequence and then withdrawn from the sphere, backing off through the images and text in an exact reversal of the opening. The whole performance could be deemed complete when the viewpoint had retreated so far that the projection screen was devoid of any perceptible object. While we still favour the idea of beginning and ending the performance by having the materials 'laid out', the flythrough was far too clichéed an implementation of it.

Some other details of this performance also reflected our rapid development time. For example, we did not have the opportunity to try and 'break' the *Lightwork* application in a kind of 'stress-testing' rehearsal and some errors occurred during execution in performance. Luckily none of these caused a crash or any other serious interruption. However, two anomalies were detected. The closing 'withdrawal from the combinatorium' sequence took place along an incorrect path, probably due to an accumulation of rounding errors leading to a false estimate of where the viewpoint was before initiating the final animation sequence. Also, the `immersiveForm` algorithm placed a form within the VE when first triggered but the form was not removed when the `immersiveForm` foot switch was depressed the second time. This proved to be an error in the code for interpreting interaction with the foot switch, though for some time, we thought that this might have been due to a serial message between the INM and DIVE having been lost. Although serial communications provides no guarantee of message reception, we have not experienced any critical failures of this sort.

Almost certainly the biggest negative impact our rapid development time had upon this first performance concerns the calibration of the algorithms. Finding sensible and safe ranges for parameter values requires an empirical approach. The INM also requires further re-scaling at a number of junctures as, for example, 7-bit MIDI values (encoding pitch and velocity) have to be scaled to be consistent with the results of analysing timing data. Each data transformation presents the opportunity for further scaling and, in the INM, we conduct some forecasting on past values as well as allow narrative rule based transformations of various sorts. It is important to realise that calibration is not just a technical matter as decisions sometimes need to be taken on whether to request the performer-user to adjust their behaviour in lieu of recalibration or whether to keep legitimate ranges for values broad so as to accommodate a variety of performance styles. As it was, in performance, V became worried that the INM was overly sensitive to changes in his playing, leading to a lurching viewpoint for example. To counteract this, V adopted a very regular playing style which was effective in bringing the INM into line but did not allow the expressivity he would otherwise have desired. Even so, the ranges of, for example, viewpoint radii did not permit the variety in views that is intended given our aesthetic emphasis on 'infinite collage'.

Performance of *Lightwork* has brought out some other matters worthy of discussion here. Currently, we do not feel the relationship between the sonic and visual elements of *Lightwork* have been quite thought through enough. Aesthetically, we are critical of attempts to very tightly couple changes in vision with sound. We wish to pursue more subtle relationships than those exemplified by simple couplings like 'the greater the z coodinate the higher the pitch' or even those which tightly relate a sense of 3D visual space to a sense of 3D sonic space. Equally, the strong coupling of movement in virtual space to the movement of virtual sound sources may seem appropriate for computer gaming but not necessarily for the kind of more experimental artistic activity we are engaged in. It is for these reasons that we relate the location of sound sources in a virtual soundspace to graphical objects only partially.

Our experience has been that this is successful and that an *impression* of strong dependence can be given even if, say, only the approximate overall speeds of virtual sound sources are correlated with apparent viewpoint speed. More involved calculations of the *exact location* of a virtual object followed by a detailed auditory rendering do not need to be undertaken. Anyway, it is unlikely that a solution based on these principles would work for all listening locations in a real physical space. Nevertheless, we feel that there is still much scope for experimentation within our emphasis on the loose coupling between sound and vision. There are many other aspects of sound in addition to spatialisation which we are yet to consider. For example, the selection of sounds could be made more systematically in relationship to the image content (e.g. machine-images could be combined with like—or unlike—sounds algorithmically, i.e. the 'sound palette' could be algorithmically configured too).

Finally, it must be admitted that informal discussions with audience members revealed that few people were aware of the interaction principles involved in the performance—at least as regards V's wind instrument playing and S's MIDI fader manipulation. In some respects, we wouldn't expect this—as if audience members could decode all of the detail it has taken several pages to describe above!—and we are not sure whether, in the context of artistic performances, the complete transparency of interaction techniques is desirable. From our aesthetic position, part of the point of our work is to create combinations of material which people are invited to invest with sense. We do not proscribe the sense that is to be made of things, nor do we create pieces which have an 'inner mystery' to be deciphered. We do not expect audience members to recover the source of our de-sourced images, texts and sounds. So, equally, we are uncertain as to whether it would be right to make our interaction techniques more transparent as if an audience should be able to recover the source of, say, a form in a VE from the gestures of the performers. On the other hand, we are certainly not arguing that we should wilfully ignore what it is possible for the audience to glean from performer activity, nor that we should take some kind of perverse pleasure in making things obscure.

Perhaps an avenue of approach with respect to these issues is revealed above in our discussion of V's experience of feeling overly constrained in his expressivity in performance with the INM in its current form and with its current, perhaps too coarse, calibration. It is our hope that with further iterations of design and testing that this can be remedied. Indeed, in principle, this is the whole point of how we have designed the INM. It is not the case that every single gesture has a direct effect upon the VE. We have been deliberately designing indirect interaction and manipulation techniques to make the relationship between performer gesture and effect more fluid. Again in principle, this should allow more latitude for expressivity (e.g. the additional emphasis given in performance over and above what is 'technically' required) as the performer need not have to attend to the precise design of every gesture. On the other hand, other forms of expressive emphasis are well served by more direct interaction principles. For example, V found himself exaggerating the gesture of stepping on the foot switches to bring out the relationship this had to imminent changes in VE content and, indeed, a number of audience members reported being aware of this, if not of the role of the wind instrument playing. Again, this brings to centre stage the importance of calibration issues when indirect, algorithmically mediated interaction techniques like those we have been exploring are used. It also suggests that some hybrid of direct and indirect techniques might allow a richer repertoire of expressive performance gestures than either alone.

Conclusion

The experiences above are principally documented from our point of view—that is, from the perspective of the developers/performers. In this regard, we have a number of self-critical comments. However, our preliminary performance of *Lightwork* was largely received successfully by those in attendance. Indeed, it received a very flattering published review. All this in spite of the exact principles of interaction being opaque to 'third parties'.

While there is clear scope for greater 'user-evaluation' and more systematic assessment of our work, for the time being, we have concentrated on the interaction *concepts* which *Lightwork* embodies. Most important in this is the principle of *algorithmically mediated interaction*, which is intended as an alternative or complement to *direct manipulation* (DM) for the construction, animation and navigation of VEs. Our preliminary experience is that real-time interaction with algorithms should work well provided that (1) algorithms are selected so their parameters have obvious *perceptual effects* for the features that performers will respond to (something which is not be true for many graphical algorithms studied in the literature on fractals or for other methods of form generation with strong 'non-linearities') and (2) performer activity can have a *loosely coupled* relationship to parameter values so that performers need not feel themselves in a 'straitjacket' where the slightest infelicity could have disastrous effects (in principle, a well calibrated INM would achieve this (i) by basing its results on *sets of gestures* so that 'errors' can be compensated for within the time window, and (ii) leaving several of the calculated statistics *unmapped* to algorithm-parameters—in the short term time window, for example, pitch and loudness have no effect). In these ways, we are exploring interaction techniques which can give performers control, do not overly restrict their latitude for action, allow error correction, and still enable computations in complex virtual worlds—a combination of features rarely considered possible in current debates on, for example, DM *versus* software agents [5].

In current and future work, we shall largely concentrate on the following matters:

- strategies for the calibration of techniques for interacting with algorithms (in particular, for viewpoint control);
- more systematic techniques for the control of sound and the interrelation of sound to visual materials (including experimentation with interaction devices for these purposes);
- revising the overall structure of the piece (e.g. experimenting with new ideas for the beginning and end);
- strategies for the combination of direct and indirect interaction techniques to allow for the richest repertoire of performance gesture.

Much of this work is already underway . Some of our current prototypes for gesture processing are described in Deliverable 6.1. Other techniques we are working on will be reported early in Year 2 of eRENA when results from them have stabilised.

Our future work will be also complemented by a consideration of whether the interaction techniques we are investigating can have further application, in particular related to whole body interaction and intuitive interfaces developed within Workpackage 6.1, but also in domains outside of artistic performance. Naturally, further performances with *Lightwork* will take place throughout 1998.

References

1. Bowers, J. et al. Talk and embodiment in collaborative virtual environments. Proceedings of CHI’96. 1996.

2. Jää-Aro, K.-M. & Snowdon, D. Retaining body language in subjective virtual environments using artifact-centred coordinate systems. Proceedings of HICSS 30. 1997.

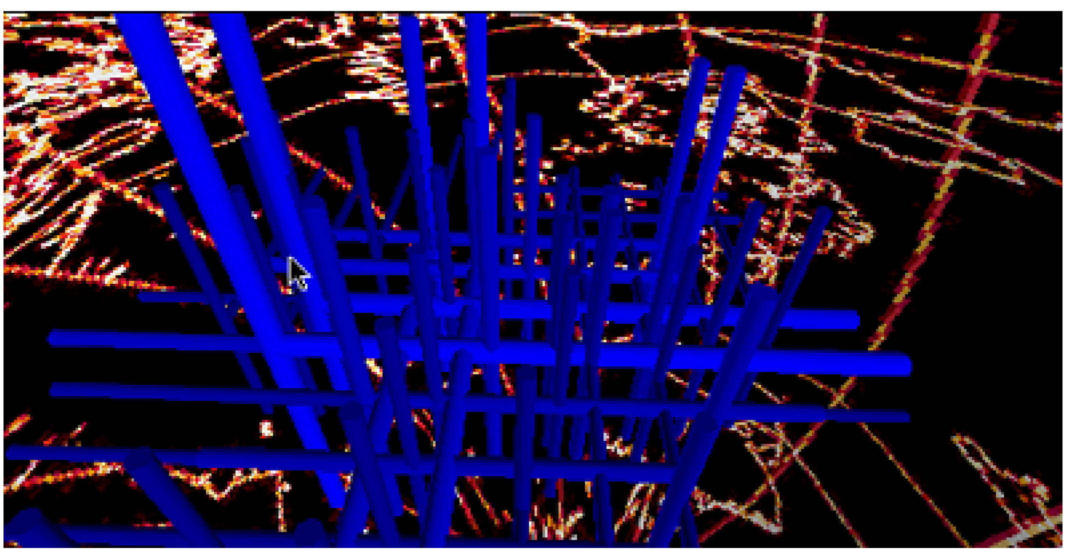
3. Smets, G. et al. Form giving: expressing the non-obvious. Proceedings of CHI’94. 1994.

4. Hagsand, O. Interactive multiuser VEs in the DIVE system. IEEE Multimedia, 3 (1), 30-39. 1996.

5. Bradshaw, J. (ed) Software agents. MIT Press. 1997.

Appendix 1: Some Images of Virtual Environments Generated Using Some of the *Lightwork* Algorithms

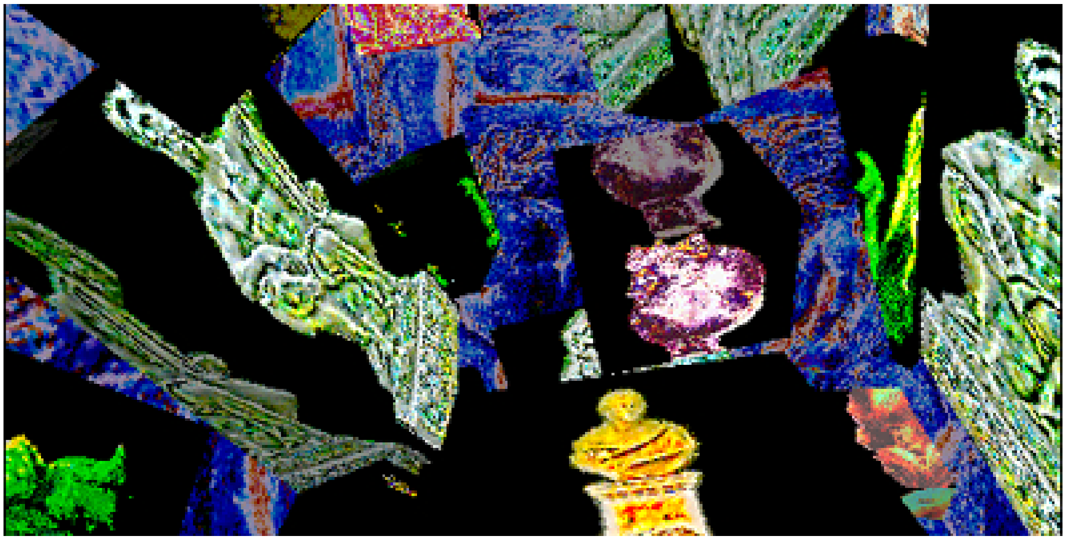
Chamber featuring images from the 'machine' archive (moderate magnitude and irregularity parameter values).



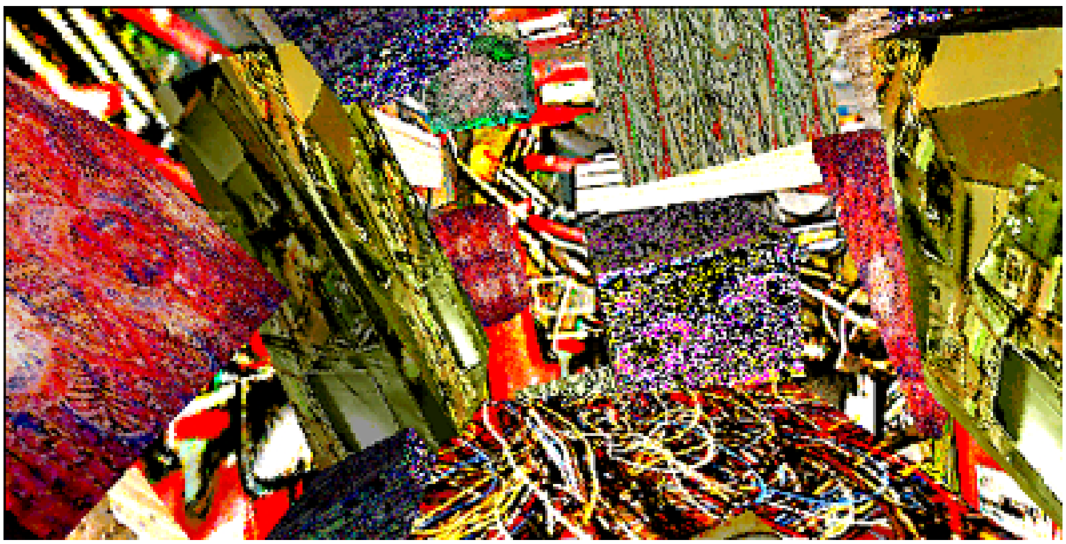
Scaffolding.

Scaffolding within a chamber featuring images from the 'myself' archive (composed of pictures from the homepages of the first 64 people one encounters when searching using alta vista with the string 'picture of me'—together with their filenames).

Chamber featuring images from the 'shadow' archive (low magnitude and irregularity parameter values).

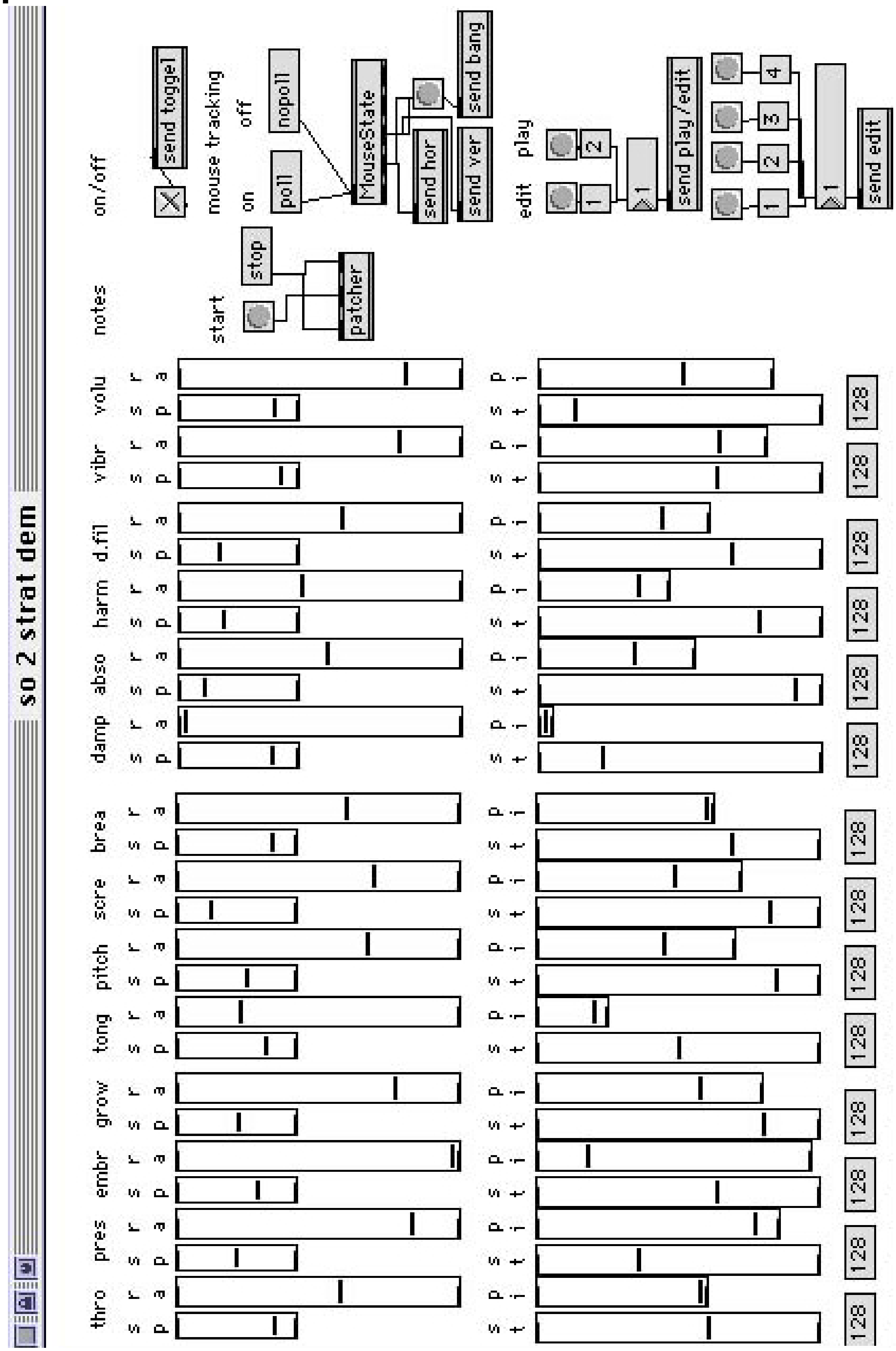


Chamber featuring images from the 'statue' archive (high magnitude and irregularity parameter values).



Chamber featuring images from the 'filament' archive (same virtual objects as previous VE, different images).

Appendix 2: A Screen Shot from SO2



Chapter 3

The *Mimoid* and *Mimetic Blob* Projects

texts by Jonas Söderberg, Bino, Pär Hansson and Lennart E Fahlén
Swedish Institute for Computer Science (SICS), Stockholm, Sweden

Mimoid

The *Mimoid* is an interactive opera based on ideas and story lines from among other things The Odyssey, Pygmalion, Kalevala and Stanislaw Lem. The term mimoid means a body that mimics. The 30 minute small scale opera is created for a multi-user mixed reality environment and focusing on new forms of artistic experiences and social interaction. In the opera singers and remotely controlled avatars are acting against each other accompanied by a small group of musicians and an electronically generated soundscape in a performance being simultaneously presented in a number of different CAVEs or similar kind of presentation and interaction environments.

Furthermore, *Mimoid* project contains some sophisticated ideas about how to involve audiences in a multi levelled way, live and electronically mediated.

In the *Mimoid* project a number of different issues will be addressed covering both "hard" computer science and "pure" artistic activities. Most importantly there is a focus on the hard to define and hard to do combinations of the two cultures.

Some of the objectives are the aesthetic deployment of both tracking- and behaviour controlled avatars, adopting the Web planetarium (as being in the developed within eSCAPE) style of interactive 3D visualization of the World Wide Web and spatialized soundscapes. A number of what on the surface seems to be pure artistic developments has to be technologically accommodated and mediated. Examples are the graphic designs of the virtual stage and environment, the real and virtual actors, and the music which is a mix of modern opera and vintage electronic music.

As part of the *Mimoid* efforts a highly reactive virtual "sculpture" has been developed. This artifact called the Mimetic Blob is an attempt to introduce very organic and in a sense "lifelike" chaotic behaviours into an interactive performance space.

The project was initiated in January 1998 and is being developed jointly within EC I3-project eSCAPE and eRENA.

In the following sections we will explore further some of the areas seen as being specifically of interest to the eRENA efforts. Especially there will be an extended treatise on the characteristics of the soundscape concepts being used in the *Mimoid* opera and to some extent in the Mimetic Blob as well as a discussion on the performance aspects of *The Blob*.

A fuller treatment of other aspects of *Mimoid* and *The Blob* can be found in the eSCAPE year one deliverables.

The *Mimoid* Opera

The *Mimoid* opera is very much concerned with issues of opposites: originals and replicas, real versus virtual, male versus female, real versus imagined, subject and object, inside and outside, being close and remote and scripted versus opportunistic.

The cast of the initial version *Mimoid* is a man (a baritone singer) who is acting against a virtual female avatar (a mezzo-soprano) either being pre-recorded or being digitally broadcast from some other space. The performance stage is an immersive multi walled 3D display and interaction space. The computer generated scenography is of course presented in this virtual stage but is also available in a more generally accessible electronic landscape.

The performers movements and actions are being tracked by various means and fed into controlling computers.

The music is performed live by a string quartet and a pianist as well as by an interactively controlled multi-channel pre recorded electronic music parallel sound track.

In the next section the interactive control paradigms for a spatialized soundscape will be further explored.

Acting in a Soundscape Interactive Performance in *Mimoid*

One aspect of the virtual stage in *Mimoid* is that it is an interactive sound landscape, a multidimensional sonic flow, where an actor or other entity can navigate and control using parts or the whole of his body. If a performer perceives something of extra interest or an especially attractive sonic process he can move closer to that point in the 3D soundscape and in that way change his perspective and focus in relation to that event. Furthermore, a movement like that can also change or influence the discourse of the whole sonic process. Not only the position of the body on the virtual stage can be used, the intensity and extent of gestures and other body movements can further modulate and transform the chosen sound flow.

Looking more specifically into the details of the first version of *Mimoid*, the components of the soundscape are all derived from recordings of voices of African women. These samples have been extensively modified through editing, filtering, pitch shifting and finally been mixed and arranged to generate the final tracks. This process is depicted in more detail in Figure 1.

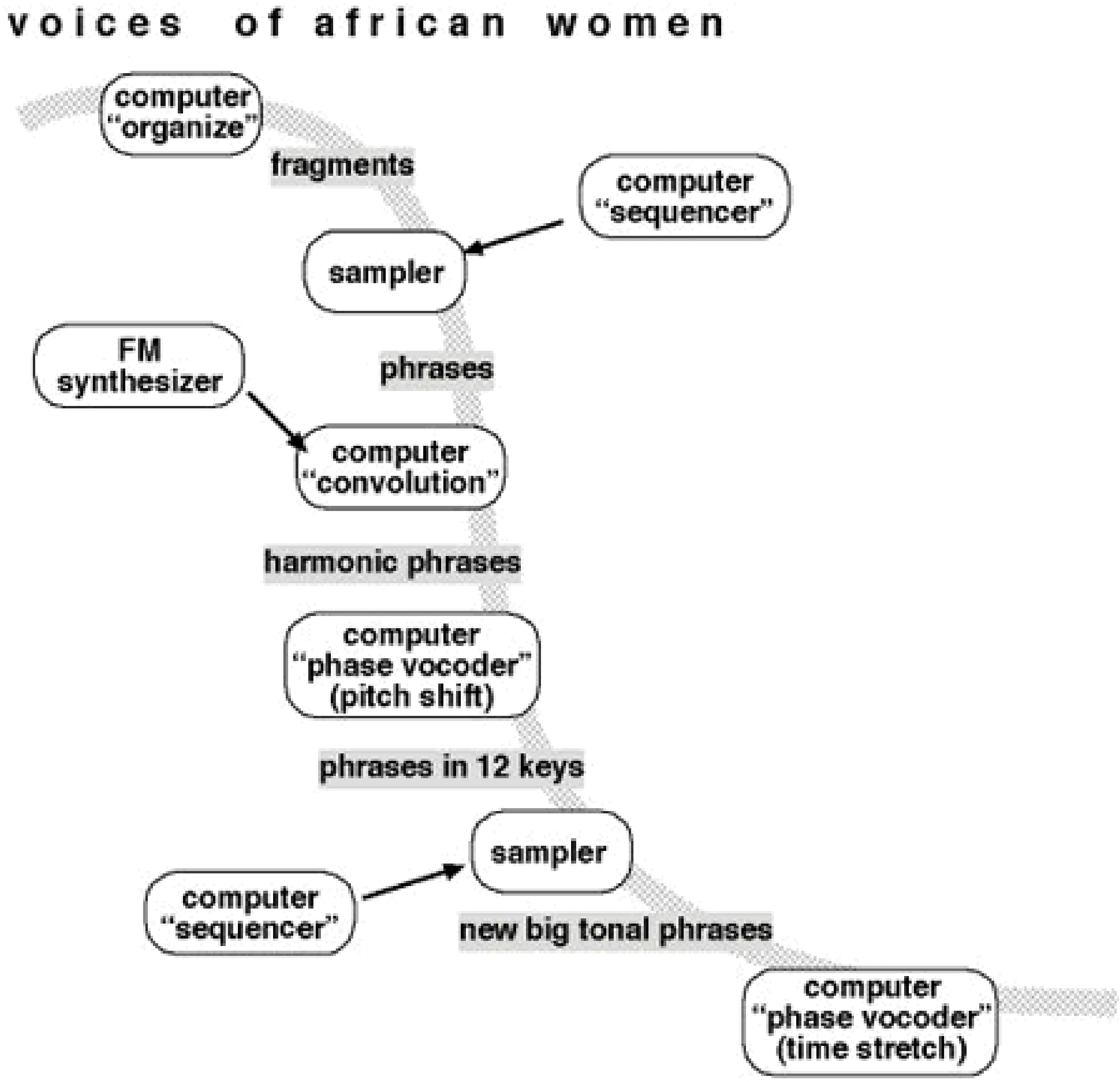


Figure 1: some of the components of *Mimoid's* soundscape.

The generated 8 soundtracks, arranged as 4 stereo pairs, corresponds to the four major bearings of North, East, South and West. of the virtual stage. This four bearings might or might not coincide with their real counterparts. In an actual performance the tracks are played in parallel with the musicians playing live off the score and subsequently the duration of the tracks corresponds to the length of the whole opera. There is always a strong correspondence between the resonant peaks (pitches) of the 8 tracks and the root note or some other dominant harmonic structure in the instrumental score. To assure that the parts played live stay in synchronization with the traversal through the soundscape there are special reference points composed into the tracks. These cues are used by the conductor (or the Primary in the case with the string quartet) who adjust tempo etc. accordingly.

Other common solutions to the problem of synchronization is for instance to have the Conductor listen via headphones to special "click track" or to have another person (typically the composer) manually and repeatedly cueing the sound tracks played from the computer.

The actors position on the stage has to be tracked and determined with a high degree of accuracy and the resulting data sent in the form of MIDI code to the computer controlling the mixing of the four stereo tracks. All "musical" communication in *Mimoid* is done via the MIDI (musical instrument digital interface) protocol. The MIDI standard might be old, slow and with too low resolution for many contemporary applications, but it is a very successful and useful standard in the sense that it is a protocol that nearly all sound (and light) equipment in the world uses and understands.

So let us suppose the following arrangement:

North - MIDI Channel 1 - Controller 7

East - MIDI Channel 3 - Controller 7

South - MIDI Channel 5 - Controller 7

West - MIDI Channel 7 - Controller 7

Controller nr. 7 (out of a total of 127) has been standardized to control volume or level. The reason is that the odd channel numbers in most computer based sound processing programs are directly controlled be different MIDI channels in an ascending order and what we are doing is to convert 8 monophonic channels to four stereo channels. This then will mean that MIDI Channel 1 will control mixer channel one and two, MIDI Channel 3 will control three and four etc. (see fig 2)

The gestures and behaviour of the performer must somehow be read and be input into the system. The idea here is not to have a direct mapping from a gesture to an immediate sonic response like some existing gesturally based instrument such as the vintage Theremin or the much more recently designed Buchla Lightning. Rather the intention is to use some higher order derived information such as gestural activity per time unit should be used as controlling data for the real time processing of the 8 parallel sound files. This treatment can be done either by a really fast computer or via a set commercially available digital processing units controlled via MIDI. The processing can consist of filtering, granularizations(whereby the sound file is chopped up into very small pieces, subsequently distributed in time and pitch according to some set of rules), frequency or amplitude modulation (a simple mapping would be to have the gesture density control the modulating pitch) or something completely different.

An important principle is that regardless of the treatment or processing being used, the resulting output (i.e. the controlling signal) should modulate some parameters in the generation the final output that spans a range of intensities.

The objective is to in one end generate a scale that is calm, soothing etc. and at the other extreme end very active, hectic, stressful. How linear or direct this mapping should be to the actual gestural activity is a subsequent issue.

Another possibility is to work with different types of treatments depending on where the gestural activity occurs. Movements high in the air could for example result in a different sonorous output than a gesture closer to the floor. A limitation in the context of performing together with an ensemble playing music with a harmonic character (i.e. not atonal sounds) is that the treatment of the material cannot be allowed to change the pitch nucleus (or "key") to a great extent.

Future Work on *Mimoid*

As part of the future work on *Mimoid* a multi site version will be developed. In this more complex arrangement the different actors, singers and musicians are all located remotely in relation to each other (as is of course the audience). This allows the interesting possibility that the audiences at the different performance spaces will see different versions of what is real and what is virtual. This can be made obvious to the people being present at one site by for example displaying video streams as part of the scenography from the other sites. Another even more intriguing idea is to make it possible for the audiences to make "virtual visits" or "avatar group tours" to other sites.

In the next section the use of intermediate layers in interactive control will be discussed.

Use of Smart Layer in Interaction Control

In the production of the *Mimoid* opera (and also to some extent *The Blob*) we will apply a whole range of different sensor and tracking systems. Examples are electromagnetic systems such as the Ascension boxes, pressure sensitive touch screens, inertial sensors, video tracking systems like the PFinder from MIT, sonar range finders, infrared light based devices such as the Buchla Lightning virtual drum system and finally electric field sensors. This set of technologies will be complemented with some more conventional interface devices such as computer keyboards and mouses and MIDI based keyboards, mixers and knob boxes.

One problem that all of the above input devices suffer from to a greater or lesser extent when used in higher level applications is that the data output is often thought to be directly used in their raw form (e.g. as absolute position values). The only pre-processing done is some simply low level filtering to, for example, remove deviating data.

However, for various reasons, there are several incentives to add a 'smart layer' on top of the low level filtering.

One could imagine a couple of examples of such situations:

- interpretation of gestures, as a sequence of movements, from the raw position data using some kind of matching with statistical models.

- adjustment of application parameters according to how a user interacts with an application (e.g. violent or slow movements) using trained neural networks.

- reduction of delays etc. by different predictive techniques.

These types of algorithms are also very interesting when it comes to art applications. Body expressions could for example be interpreted with any kind of abstract 'smart' layer to be used in interactive art to control objects, and colours, and sounds.

The Mimetic Blob

The Mimetic Blob is a vital part of the *Mimoid* opera work where the focus is on the development of a reflective, interactive, memory absorbing and memory fragment emitting "substance". This kind of artefact plays a central role in the *Mimoid* story and presents an "alien" counterpart to the real humans and the humanoid avatars. In the design work on *The Blob* the metaphors and elements of the ocean and life in water has been very prominent. It was especially important to try to capture in *The Blob*s behaviour the oceans constantly morphing form and organic shifting behaviour.

The birth history of *The Blob* contains a number of happy coincidences where technological development walked hand in hand with aesthetic and artistic considerations. It was developed jointly by technicians, artists, and musicians.

We tried to get the technology to meet the demands of art as well as getting the aesthetic input to be funnelled through the technology. Of special note is that a small graphics and display technology company (a SME) commissioned the first version of *The Blob* to show at a trade fair exhibition.

The Blob was first shown at Cebit 98 in Hannover, Germany in April 98 in the joint Realax/Graf booth. In this initial version *The Blob* was displayed in 3D in an interactive horizontally mounted virtual desktop.

The visitor stands in front of the virtual desktop wearing a pair of crystal shutter 3D glasses.

The desktop screen is pressure and position sensitive, so by touching the screen the "user" can control or rather influence the behaviour and sounds emitted by *The Blob* mass.

Typically *The Blob* will in an organic and squid like way track the finger but the place and the way it is touched can greatly change this and for instance trigger a very nervous and aggressive behaviour both in shape and in its visual appearance, colour characteristics, transparency etc. *The Blob* is constructed from tri-meshes, triangle formed polygons. In the first version there are 40 different textures being used, all of them inspired by the ocean and oceanic life forms. The textures are classified into three different groups according to their content, colour saturation and transparency etc.

A range of different responses can be setup by changing the parameters controlling the deformations of the polygons as well as the morphing between different shapes and textures.

Some of the surface textures contain images of concrete real world objects. These hidden pictures can be interpreted as fragments of long lost memories, unconscious thoughts and desires.

In the present version the images used as textures are all manually processed through Adobe Photoshop.

The mechanisms for controlling the sounds emitted by *The Blob* works in a similar way to the handling of the polygons and textures.

So in a truly *Mimoid* fashion *The Blob* can imitate and respond to touch and movements, change its appearance both in form, shape and colour, generate sounds and project memory fragments.

Anecdotal note: In one instance *The Blob* was seen stealing and projecting textures ("memories") from another virtual object sharing its virtual space....

Presently, moving round in Blob space is based on conventional navigational metaphors such as the ones currently found in driving and flight simulators/games, i.e. pointing and clicking with your finger, dragging in different directions in combination with various modes.

Future Work

Multi-User Multi-Site

There is a number of different ways this interactive virtual sculpture could be developed further.

The Blob currently discussed is single user. There is no reason for it to remain so and we currently have in our lab multi-user multi-site versions also built in the DIVE system (Distributed Interactive Virtual Environment).

Metaphoric Mappings

In an ancient theory known as Buddhist Finger Magic there is a system for the movements and positions of the fingers, where each finger represent one of the five elements. This could be explored as a possible set of mappings of behaviours and appearances of *The Blob*. A simple way of achieving this in terms of an interface is to instead of having a performer interacting via a touch sensitive surface, the performer would wear a sensor glove where the tracking data emanating from each of the fingers of the glove will be linked accordingly. This information could then be used to change *The Blob*s behaviour, parameters, textures, soundscapes, etc. in order to explore this Buddhist concept.

Another possible mapping could be the Chinese Akupressure technique. The Chinese discovered that pain in the body could be healed by akupressure through a theory about representation of the body. The idea being that if you for example wanted to relieve a pain in your back, you could apply pressure on your ear at appropriate points. The idea here would be to find *The Blob*'s akupressure points and map interactions to behaviour's based on this mapping theory.

Performance Application

Using video tracking technology images of the heads or torsos of members of the audience can be used as a surface texture material for *The Blob* rendering. An interesting extension of this idea is to merge this material with input from other sites and users. This *Blob* consisting of composites of material from many real time sources can of course also be deformed/transformed in the way described above, with the added ingredient of real time cooperative manipulation. In this way the audience becomes both artist, spectator, art artifact, provider of mimetic fragments and distorted mirror objects, collaborators in the same piece of art.

The WWW Connection

Some of the techniques currently being explored in the eSCAPE WebPlanetarium project for mapping objects or pages found on the World Wide Web into shared 3D space could be used to complement/replace the static imagery used as basis for *The Blob* textures. In this way the appearance of it could be directly influenced in real time by a specific or random set of web pages. This would further enforce *The Blob* as a memory projecting, association generating device.

Mimoid Integration

At some point the Mimetic Blob has to be integrated with the opera *Mimoid* for which it is originally designed before it took on a life of its own. In doing this it is anticipated that the characteristic of it has to be changed and adopted in several ways for it to become a subordinated part of the opera.

The Mimetic Blob has been accepted for presentation at the ACM Multimedia 98 conference in Bristol, UK.

The Artist's Dilemma

Having to describe details of an artwork not yet realized is a both unusual and awkward task for an artist.

"Will somebody rip off my ideas?" is an admittedly naive but still a very real gut reaction.

"Can this artwork really be realized" is a reaction that is even closer. In addition one can add the experience that tells you that even if you had really good ideas and clear concepts when starting to realize the artistic artefact you have by the time the work is finished left most of those behind and probably ended up somewhere completely different!

Furthermore, in the good old days composers saw themselves as just putting to paper music that had already been created by (their) God. Today, for many artists God does not exists (at least in the sense discussed here) but still some of the same sense prevails. The artist unveils a small piece of an artwork and the rest is somehow ready to uncover as well. The composer decides on a key, an interval, a timbre and the rest of the piece comes into existence by its own deterministic inertia.

Another drawback in describing the details of the ideas behind artworks is that the audience then is confronted with material not really targeted for them. The underlying theoretical structure being used by the artist(s) in the conceptual and creation phase of an art piece may have no importance to them who are going to experience it. The audience should experience the surface, the projected interface and maybe get a sense of something more complex and involved lurking behind. They should interpret according to their own experiences and abilities, generating their own thoughts and emotions.

To give an example from the *Mimoid* work described above: In the composers mind the design and content of the four stereo tracks deliberately reflects an euro-centric, colonial nineteenth century perspective according to following mappings:

North - Arctic - Calm, slow, apollonian, intellectual

East - Oriental - Sensual, mystical, chromatic, female

South - Tropical - Fast, rhythmical, dionysian, emotional

West - Occidental - Strict, functional, militaristic, macho

The performer/actor will view and command the sound landscape in a way similar to how europeans traditionally viewed the world as being centred around themselves. The information presented will automatically be sorted into a pre-existing belief system. New facts only seem to confirm and enforce a social and cultural based world view established for many generations.

The above section is one example of associations an intended audience does not need to know about.

There is no political and social motives for an artist to use these metaphors, they are just mental aids or stimulations helping the artist to compose, create, handle information and to generate interesting dialectic opposites and parallels.

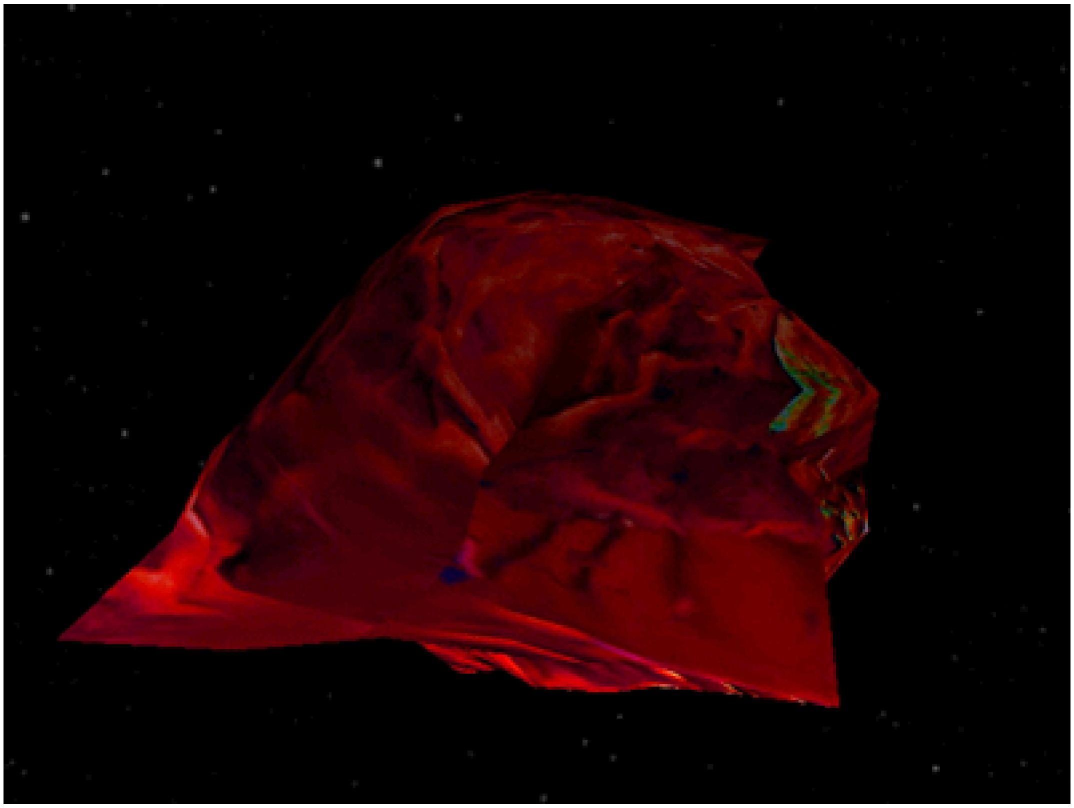
There is another side to this argument. In order for a performance piece to be engaging and exciting to an audience there needs to be a certain degree of understanding of the mapping between, let's say, a gesture and a resulting sound. Furthermore, there also needs to be an element of risk or effort or uncertainty involved in the perceived mappings. Otherwise, if these conditions are not met the performance element becomes superfluous and unnecessary and the audience might either be better off consuming a pre-recorded version or they might just experience the whole performance as chaotic and impossible to understand and to establish a relationship with. The relation between action and result might become too obvious and trivial (i.e. a conductor or virtuoso musician "going through the motions") or they could be just too obfuscated as exemplified by a person sitting more or less motionless on a darkened stage behind a computer clicking on a mouse!

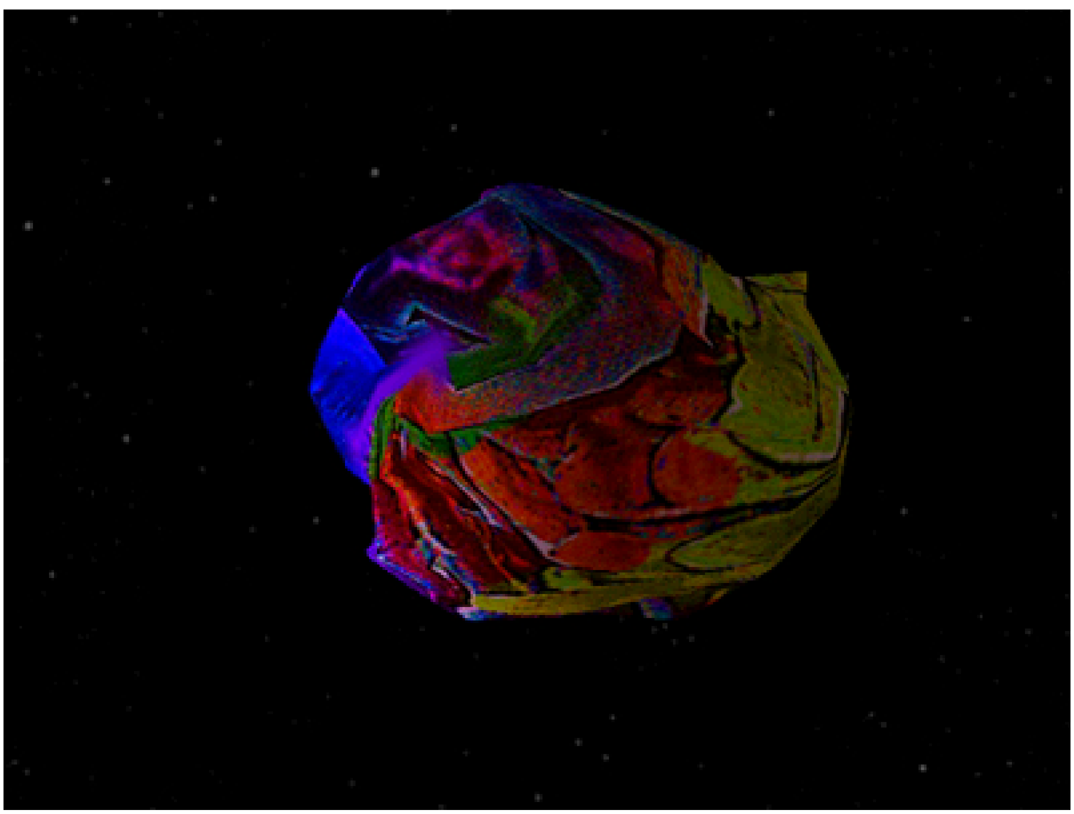
It is certainly up to the artist to strike the "correct" and exciting balance between these two extremes.....

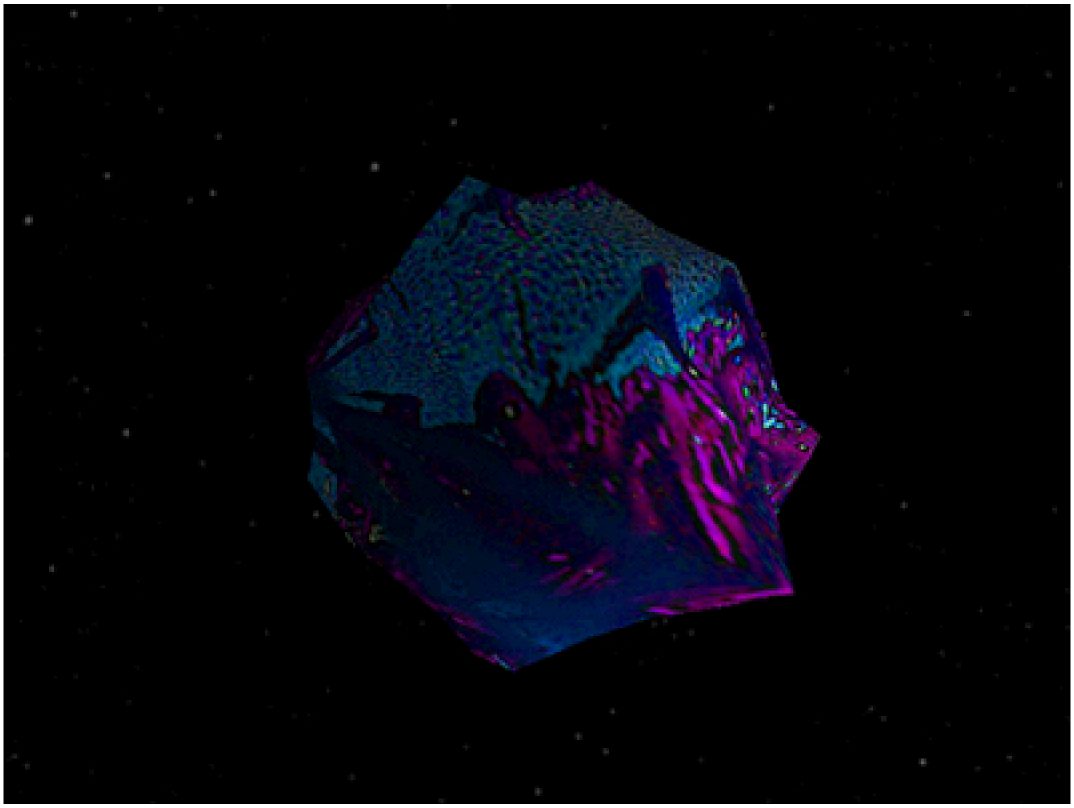
In the end it is left to the Art theoreticians and critics to in a crime detective like way trying to reconstruct the processes that led up to the crime, sorry, piece of art. The three way dialectic play between artists, critics and the audience is extremely important for the preservation of the vitality of art and its contextual relevance. Still, it is important to understand that the interpretations and analysis constructed in this way very seldom agrees with the real background and this does not matter and it does not have to!

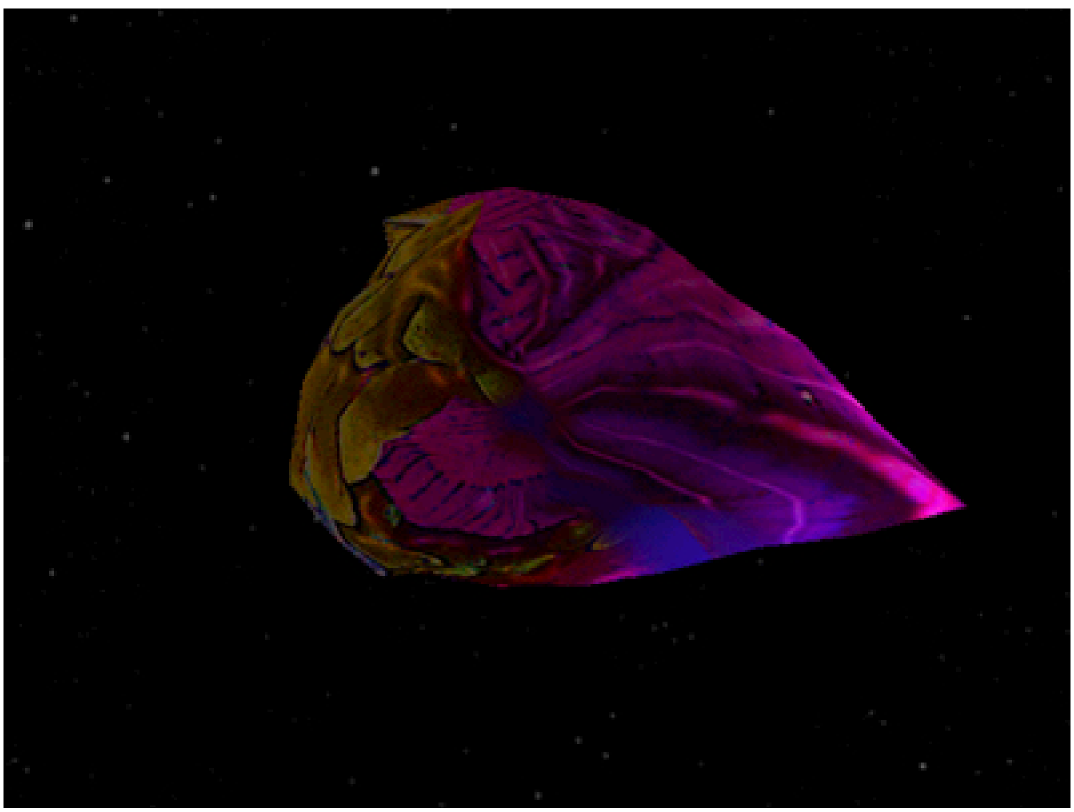
The intentions of an art piece is just not of interest to anybody but to the artist him/herself, with the possible exception of the researcher and the student.

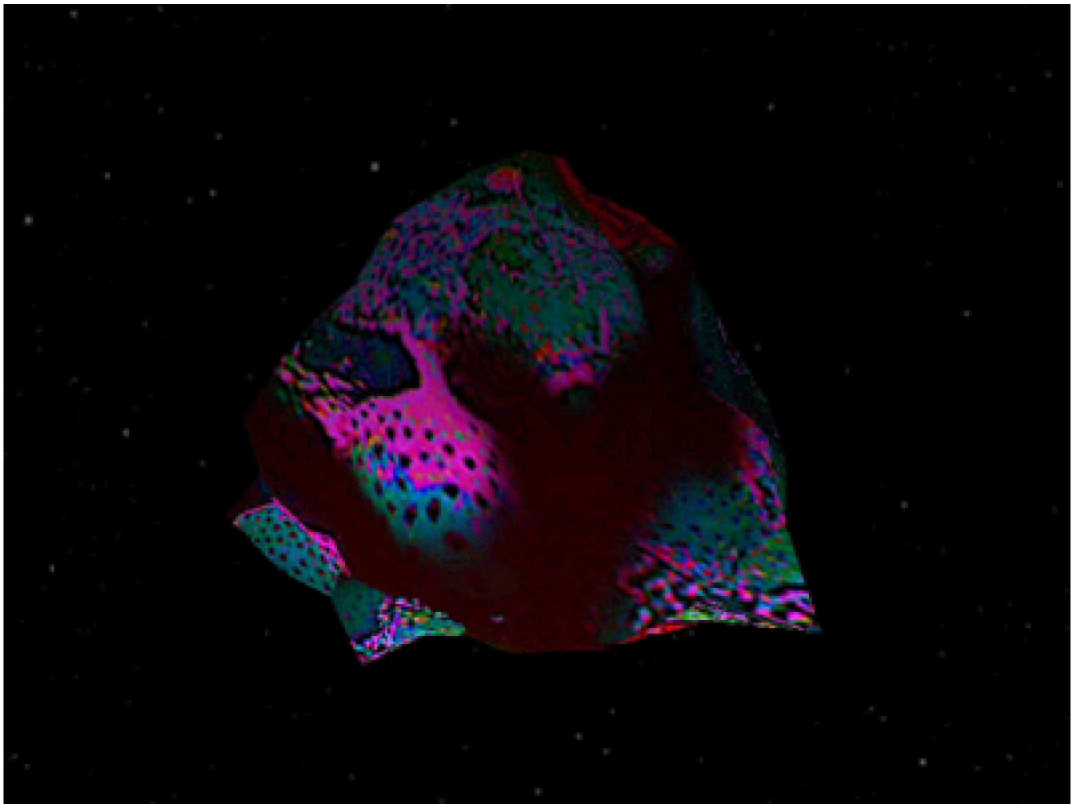
Images from *The Mimetic Blob* and *Mimoid*

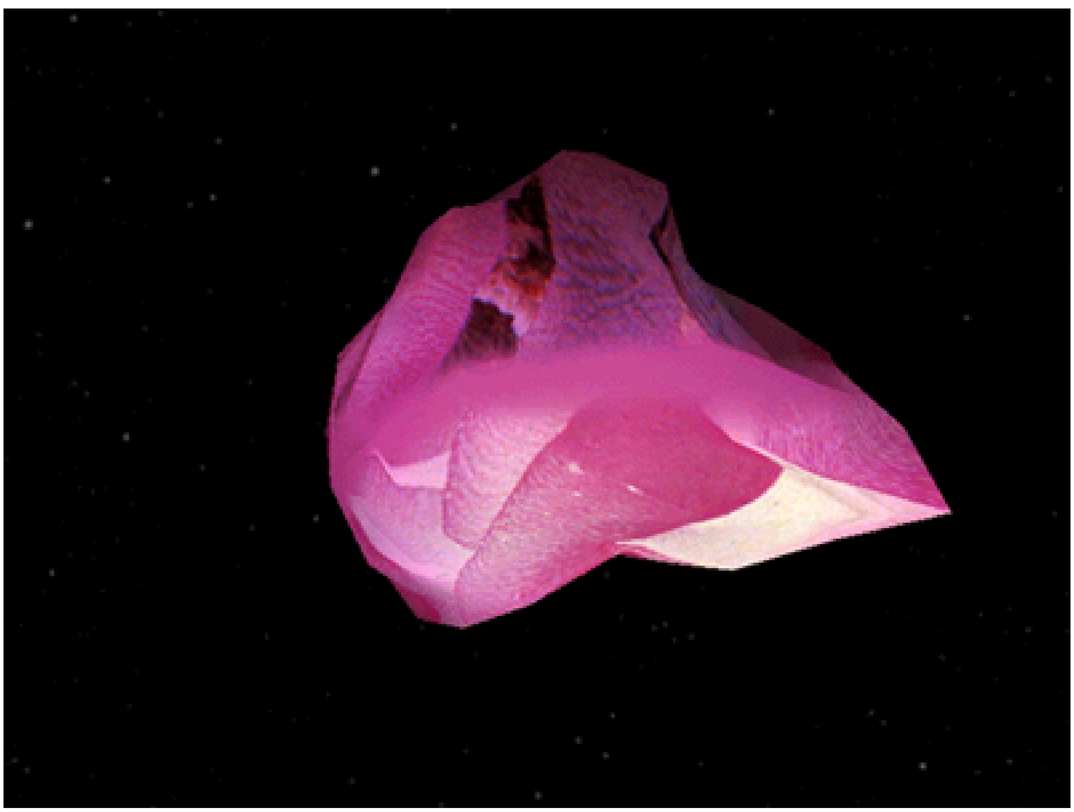


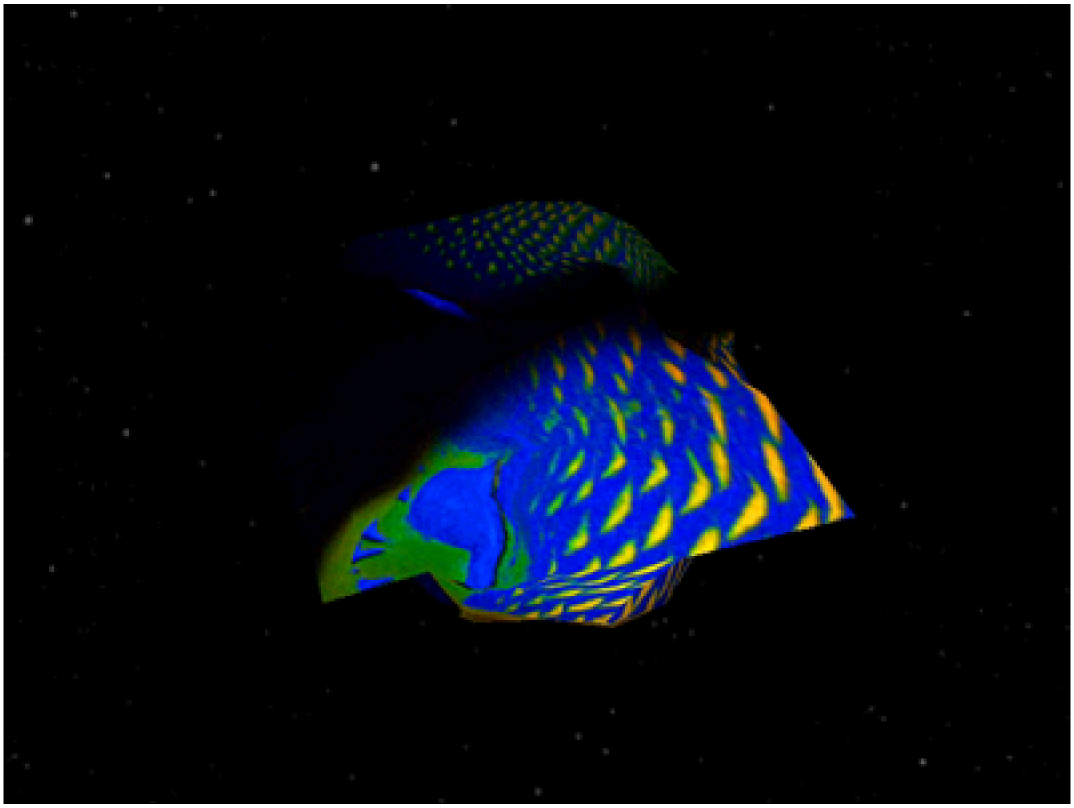


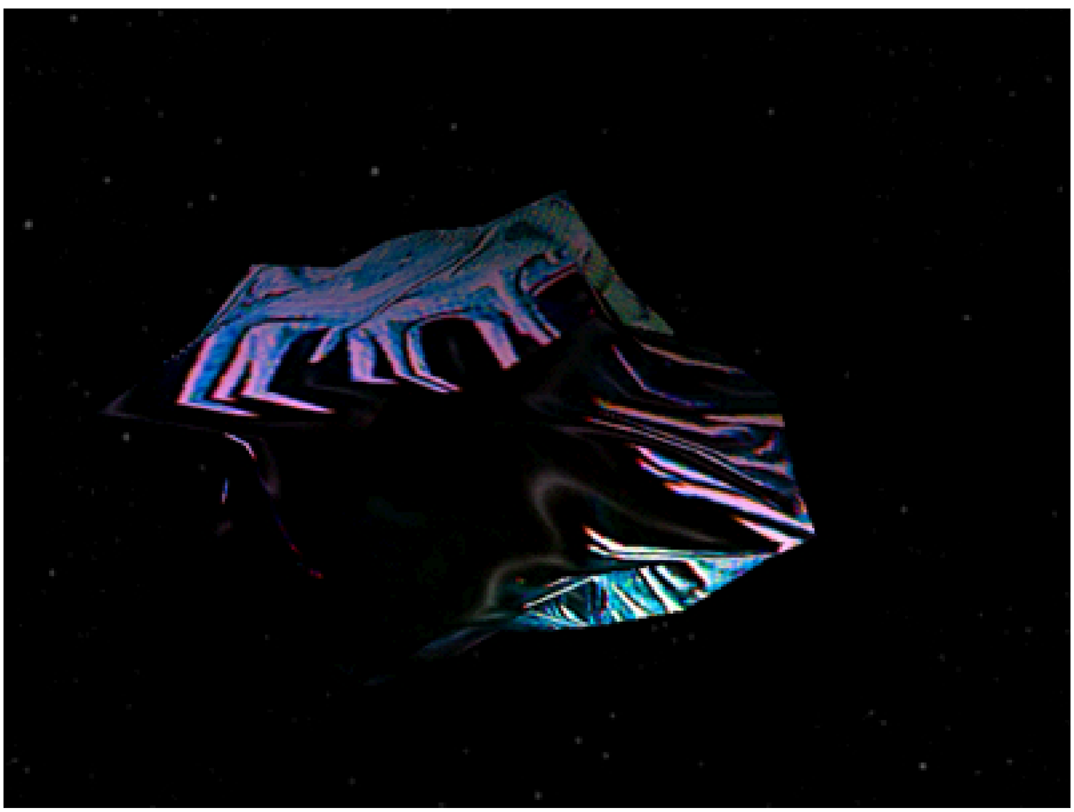


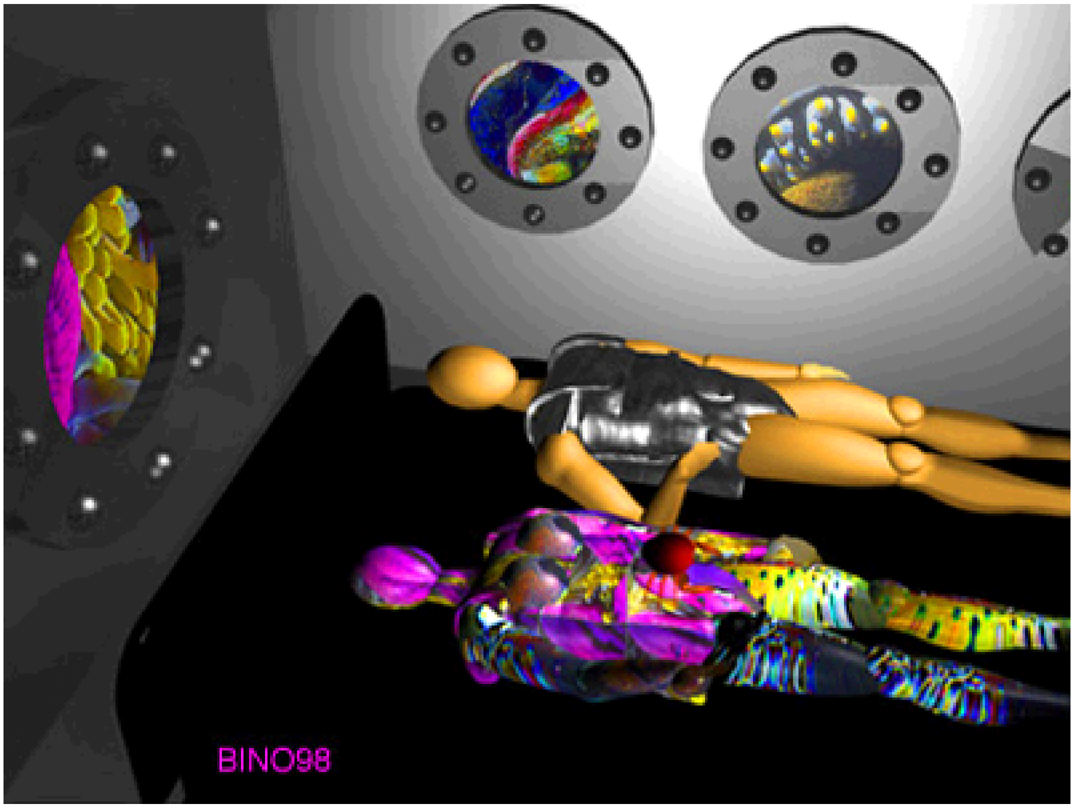


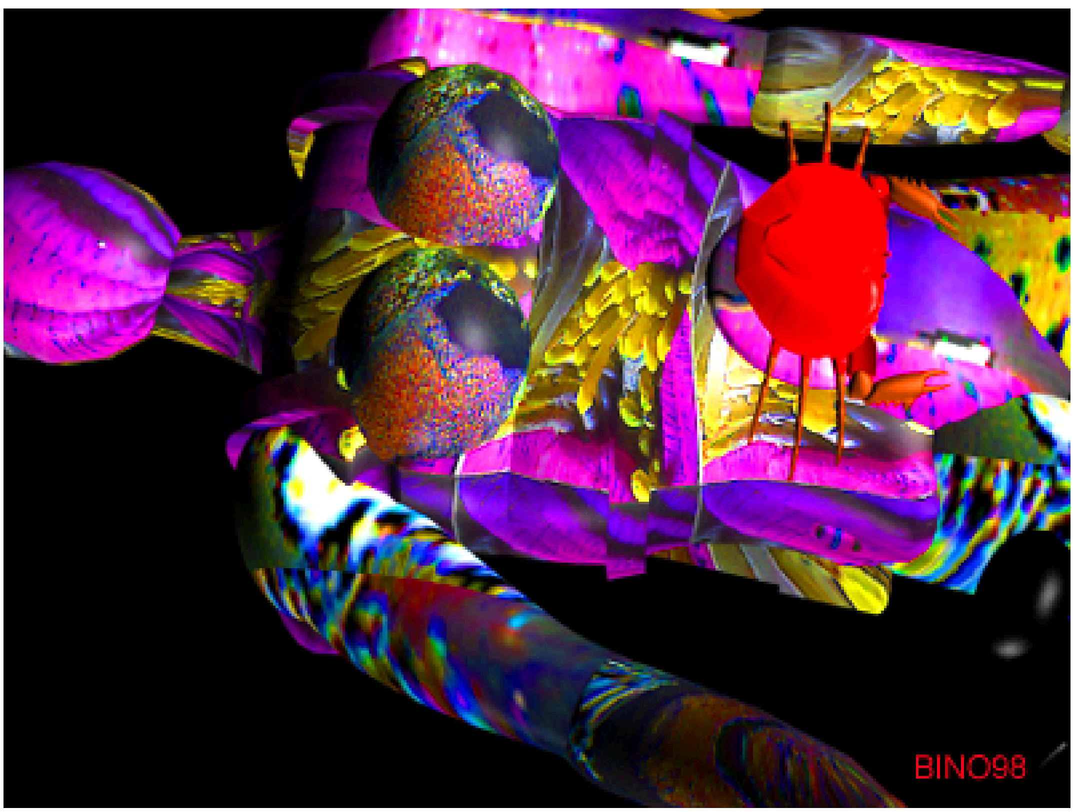


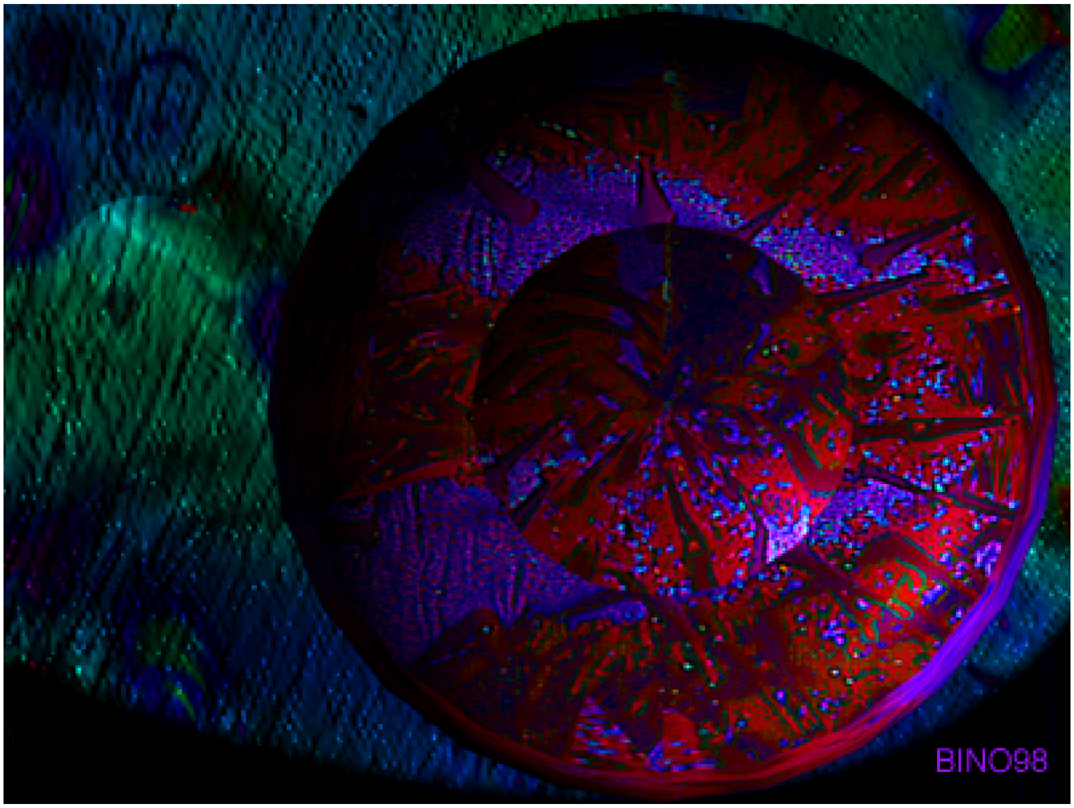


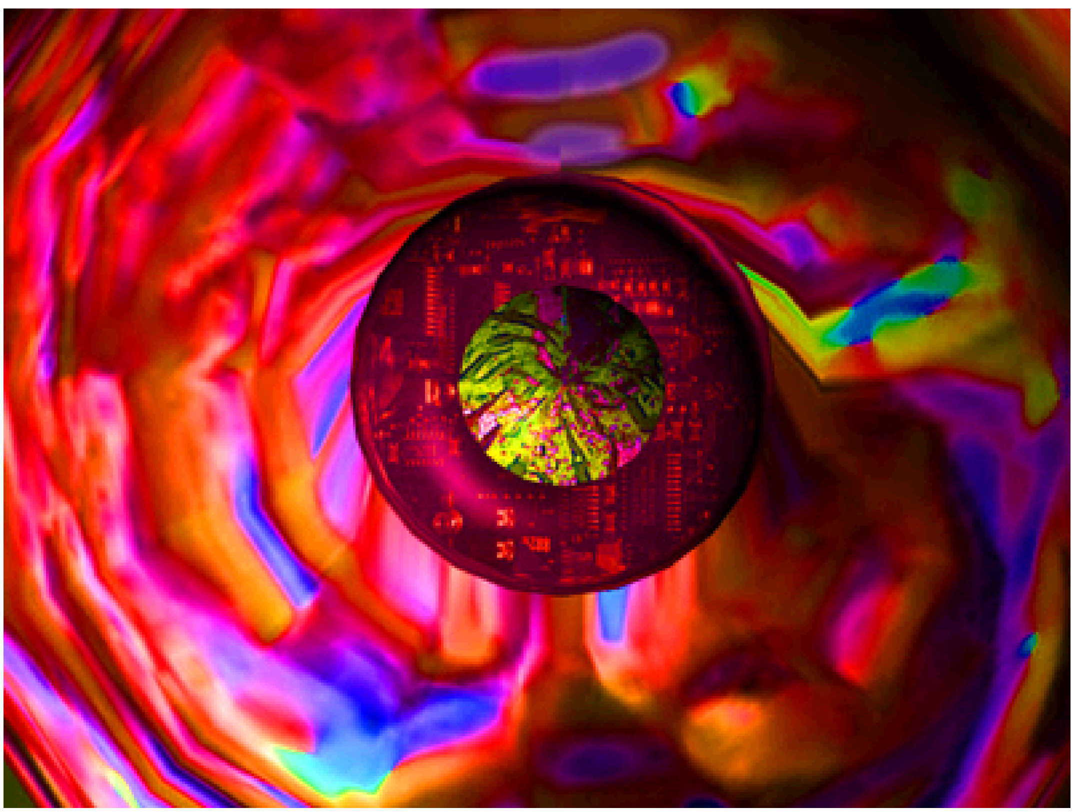


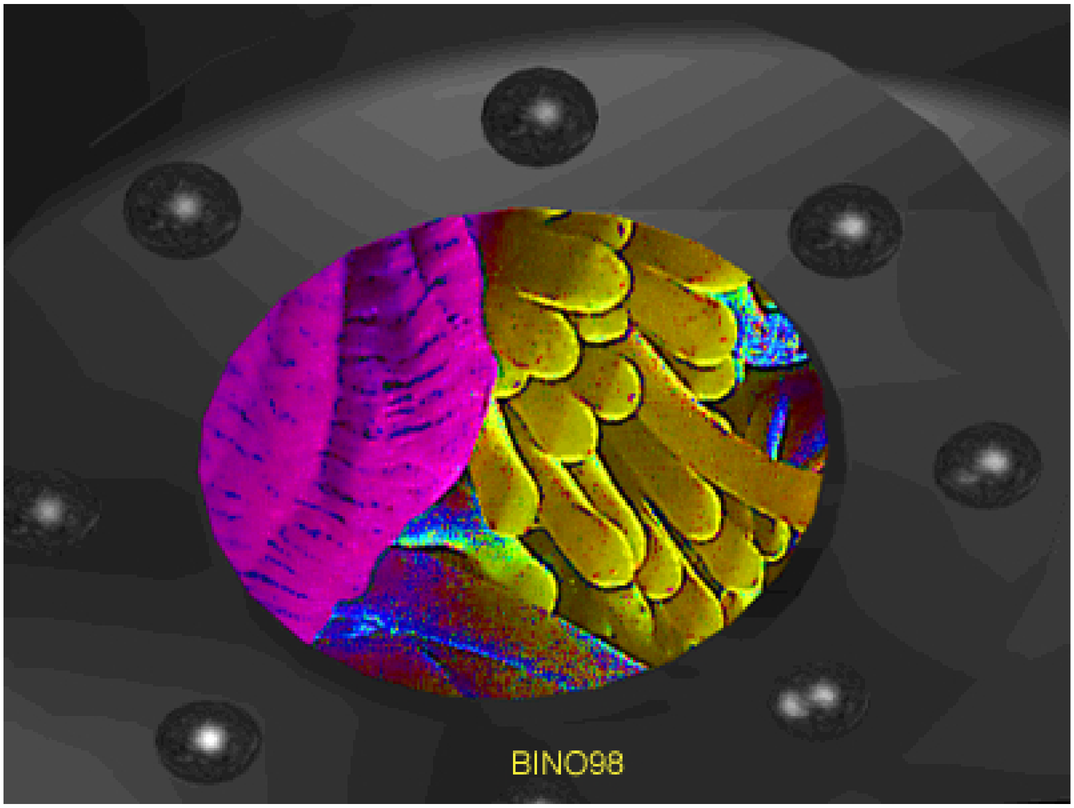



BINO98


BINO98


BINO98




BINO98