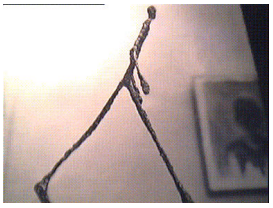
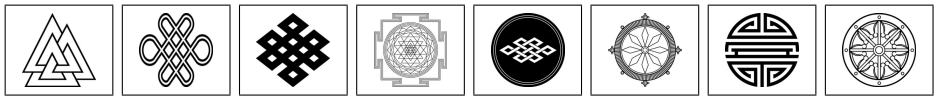
part iv. applications & tools

a journey of a thousand miles begins with the first step
chinese/japanese proverb



8. virtual environments
9. digital content creation
10. application development



2

reading directives In this part we will look in more detail at virtual environments as an interface to complex multimedia information spaces. In chapters 9 and 10, we will consider the issues that come into focus when creating digital content, and more in general, when developing a multimedia application. In chapter 10, the final chapter, we will look at some examples of multimedia application development.

Essential sections are section 8.1, which argues how virtual reality interfaces may be relevant, and sections 9.2 and 10.2, which provide examples of multimedia application development. Dependent on your experience you may skip sections 9.1 and 10.1, which provide rules of thumb for respectively content creation and application development.

perspectives Even in a more practical sense there are many perspectives that may characterize your academic stance. Design and application development, obviously, does not only consist of aesthetic or technical issues. So, non-exhaustively, you may look at such issues from the following perspectives:

perspectives – multimedia applications

- technical – algorithmic effects
- sociological – stakeholders and teamwork
- tool selection – Maya vd 3DSMAX
- political – negotiating support
- scientific – experience design
- computer science – tools and technologies
- artistic – portfolio as a design product

For example, the issues you may come across in an actual project may have to do more with people then any thing else, in other words may be more of a political nature, than of an aesthetic nature.

essay topics Even when you are primarily interested in the practice of developing digital content, it might well pay of to reflect on more theoretical issues. For example, consider writing an essay about:

- 2D vs 3D aesthetics animation – stills, sequences and stories
- elements of style – diversity and confluence
- models of creativity – a critical evaluation

In writing about such issues you should always beware of the risk of abstract speculation. So, look for examples in the domain of art, design or street culture to demonstrate your point.

the artwork

1. walking figure – sculpture by Alberto Giacometti, Giacometti.
2. signs – meteorological symbols, Signs, p. 214, 215.
3. photographs – Jaap Stahlie[1], commissioned work, using traditional non-digital techniques.

[1] www.jaapstahlie.com

8. virtual environments

augmented virtuality acts as an intelligent looking glass

learning objectives

After reading this chapter you should be able to characterize the notion of virtual context, discuss the issue of information retrieval in virtual environments, explain what is meant about intelligent multimedia and discuss the potential role of intelligent agents in multimedia applications.

From a user perspective, virtual environments offer the most advanced interface to multimedia information systems. Virtual environments involve the use of (high resolution) 3D graphics, intuitive interaction facilities and possibly support for multiple users. In this chapter, we will explore the use of (desktop) virtual environments as an interface to (multimedia) information systems. We will discuss a number of prototype implementations illustrating, respectively, how paintings can be related to their context, how navigation may be seen as a suitable answer to a query, and how we can define intelligent agents that can interact with the information space. Take good notice, the use of virtual environments as an interface to information systems represents a major challenge for future research!

1

8.1 virtual context

Imagine that you walk in a museum. You see a painting that you like. It depicts the Dam square in 17th century Amsterdam. Now, take a step forwards and suddenly you are in the middle of the scene you previously watched from some distance. These things happen in movies.

Now imagine that you are walking on the Dam square, some sunday afternoon in May 2001, looking at the Royal Palace, asking yourself is this where Willem-Alexander and Maxima will get married. And you wonder, what did this building and the Dam square look like three centuries ago. To satisfy your curiosity you go to the Royal Museum, which is only a half hour walk from there, and you go to the room where the 17th century city-scape paintings are. The rest is history.

We can improve on the latter scenario I think. So let's explore the options. First of all, we may establish that the Dam square represents a rich information space. Well, the Dam Square is a 'real world' environment, with it has 700 years of (recorded) history. It has a fair amount of historical buildings, and both buildings and street life have changed significantly over time.

So, we can rephrase our problem as

how can we give access to the 'Dam square' information space

But now we forget one thing. The idea underlying the last scenario is that we somehow realize a seamless transition from the real life experience to the information space. Well, of course, we cannot do that. So what did we do?

2

Look at the screenshot from our *virtual context* prototype. You can also start the VRML demo version that is online, by clicking on the screenshot. What you see is (a model of) the Dam square, more or less as it was in 2001. In the lower part, you see a panel with paintings. When you click on one of these painting,

your viewpoint is changed so that you observe the real building from the point of view from which the painting was made. Then using the controls to the right of the panel, you can overlay the real building with a more or less transparent rendering of the painting. You can modify the degree of transparency by turning the dial control. You may also make the panel of paintings invisible, so that it does not disrupt your view of the Dam and the chosen overlay.

In other words, we have a VR model of Dam square and a selection of related paintings from the Royal Museum, that are presented in aa panel from which the user can choose a painting. We deploy viewpoint adjustment, to match the selected painting, and we use overlay of paintings over buildings, in varying degrees of transparancy, to give the user an impression of how the differences between the scene depicted in the painting and the actual scene in (the virtual) reality.

We have chosen for the phrase *virtual context* to characterize this prototype, since it does express how virtual reality technology enables us to relate an information space to its original context.

From the perspective of virtual reality, however, we could also have characterized our prototype as an application of *augmented virtual reality*, since what we have is a virtual reality model of a reallife location that is augmented with information that is related to it, (almost) without disrupting the virtual reality experience. In summary, we may characterize our approach as follows.

augmented virtual reality

- give user sense of geographic placement of buildings
- show how multiple objects in a museum relate to eachother
- show what paintings convey about their subject, and how

Considering the fact that many city-scape paintings of Amsterdam have been made, many of which are in the Royal Museum, and that paintings may say many things about their subject, we believe that our approach is viable for this particular instance. The augmented virtual reality approach would also qualify as a possible approach to cultural heritage projects, provided that sufficient pictorial material is available or can be reconstructed.

Although we were quite satisfied with what we accomplished, there are still many things that can be done and also a number of open problems. Guided tours are a wellknown phenomenon. But how to place them in our virtual context is not entirely clear. As another problem, our approach does not seem suited to account for buildings that do no longer exist. Another thing we have to study is how to change the temporal context, that is for example change from a model of the dam in 2001 to a model of the Dam in 1850. We would then also like to have 'viewpoint transitions' over space and time!

Finally, to give better access to the underlying information space we must also provide for textual user queries, and find an adequate response to those queries.

VRML To realize our prototype we used VRML, which limits us to medium quality desktop VR. At this stage, VRML is a good option, since it is a relatively stable format with a reasonable programmatic model. In short, what VRML offers is

VRML

- declarative means for defining geometry and appearance
- prototype abstraction mechanism
- powerful event model
- relatively strong programmatic capabilities

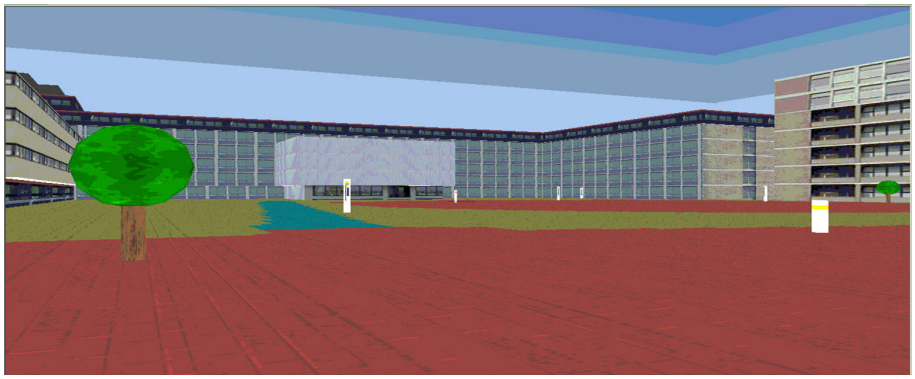
Although VRML allows for writing models (including geometry and appearance) using a plain text editor, many tools support export to VRML. As a consequence, often tools are used to create more complex models.

In addition, VRML allows for defining prototype abstractions, so reuse of models and behavior can be easily realized.

Defining dynamic behavior involves the routing of events that may come from a variety of built-in sensors (for example a TimeSensor for animations) to scripts or so-called interpolators, that allow for the manipulation of geometry and appearance parameters of the model.

In particular, the use of scripts or the *External Authoring Interface* (EAI), that allows for defining behavior in Java, is essential for realizing complex behavior.

Summarizing, VRML is a sufficiently rich declarative language for defining 3D scenes, with a relatively powerful programming model for realizing complex behavior. Some may think that VRML is dead. It isn't. The underlying model is endorsed in both the X3D and RM3D standards, simply since it has proven its worth.



3

research directions– *augmented virtuality*

Given an information space, there is a duality between information and presentation. For an audience or user to be able to digest a presentation, the amount of information must be limited. Effective presentation, moreover, requires the use of proper rethorics (which may be transcoded as *ways of presenting*) that belong to the medium. Using VR, which is (even in its desktop format) a powerful presentation vehicle, one should always beware of the question *what is it good*

for? Generally one may ask, what is the added value of using VR? In an abstract fashion the answer should be, to bridge the gap between information content and presentation. Or, in other words, to resolve the duality between information and presentation!

Let's look at an example, a site about archeology, announced as a site offering *Virtual Archeology*. Perhaps it is good to bring to your attention that the *virtual*, in Computer Science, means nothing but another level of indirection to allow for a (more) flexible usage of entities or objects. See Eliens (2000), section 1.2.

virtual archeology

- variety of archeological sites
- various paths through individual site
- reconstruction of 'lost' elements
- 'discovery' of new material
- glossary – general background knowledge

For a site about archeology, *virtual* means the ability to present the information in a number of ways, for example as paths through a particular site, with the possibility to explore the reconstruction of lost or perished material, and (for students) to discover new perspectives on the material. In addition, for didactic reasons there may also be a glossary to explain concepts from archeology.

Now, how would you construct such a site about virtual archeology? As a collection of HTML pages and links? It seems that we can do better, using VR and rich interaction mechanisms!

So, what is meant by *augmented virtuality*? Nothing that hasn't been expressed by the notion of *augmented virtual reality*, of which an example has been given in this section. The phrase *augmented virtuality* itself is just one of those potentially meaningless fancy phrases. It was introduced simply to draw your attention to the duality between information and presentation, and to invite you to think about possible ways to resolve this duality.

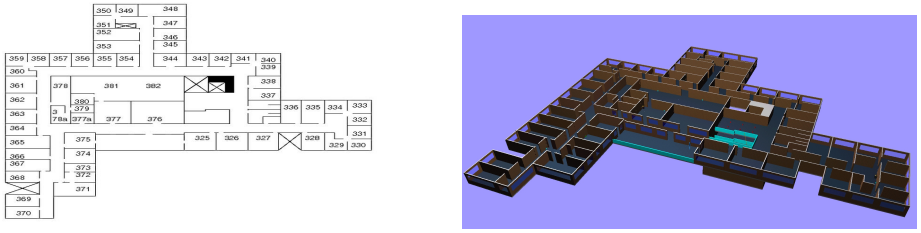
8.2 navigation by query

Virtual worlds form (in itself) a rich repository of multimedia information. So, when working on the musical feature detector, sketched in sectin 6.3, the thought occurred to ask funding for a research project on information retrieval in virtual worlds. This project is called RIF, which stands for

RIF

Retrieval of Information in Virtual Worlds using Feature Detectors

For the RIF project, we decided to develop a small multi-user community of our own, using the *blaxxun* Community Server. Then, during the development of our own virtual environment, the question came upof how to present the results of a query to the user. The concept we came up with was *navigation by query*, and in this section we will look at the prototype we developed to explore this concept.



On the left is the 2D map of the third floor of CWI, on the right the model generated from it.

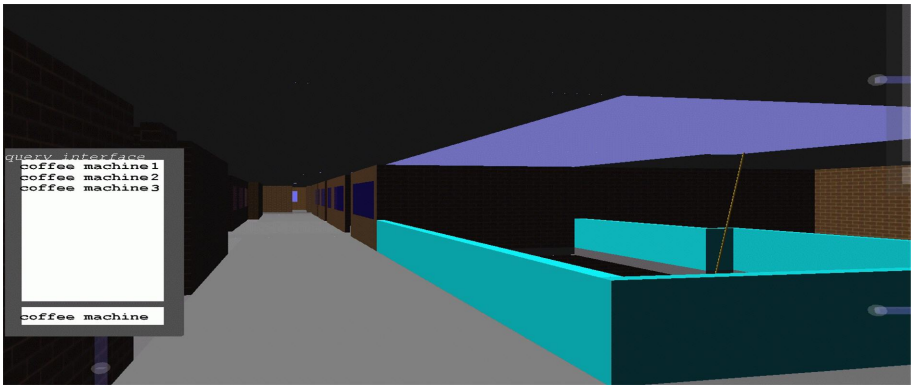
4

case study – CWI

For our prototype, we took one of the worlds of our virtual environment, the third floor of the CWI. The reason for this is that we were (at the time) doing our research there, and so there were plenty locations of interest, such as the rooms of our collegues, the printer room, and not to forget, the coffee corner.

We started out by taking a map of the third floor, and developed a model of it, using a tool developed by a student, who needed such a tool for realizing his game *Out of the Dark*.

When dwelling around in (this part of) our virtual environment, the user may pose (arbitrary) queries, for example *where is the coffee machine*.



5

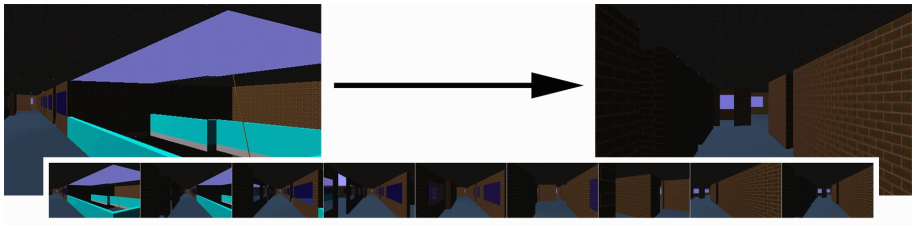
Remind, that after a few hours of research, coffee might be needed to get fresh ideas!

As a result, the user is then so to speak taken by the hand and led to one of the coffee machines that can be found on the third floor. In effect, with knowledge of the layout of the building a viewpoint transformation is executed, in a tempo that allows the user to

explore and discover

the (model of the) third floor of the CWI.

The idea is rather straightforward. Some have asked us why *navigation by query* might be useful. Well, simply, it seems to offer an interesting alternative to navigation by explicit interaction and navigation in the formof a guided tour. Our primary goal in developing the prototype, however, was to see whether navigation by query is feasible, and under what conditions.



6

information in virtual worlds

Developing the prototype has forced us to think more explicitly about what information is available in virtual worlds, and (perhaps more importantly) how to gain access to it. So the question we asked ourselves was

what are we searching for?

Now, in a virtual world, such as the ones built with VRML, we can distinguish between the following types of information: viewpoints, that is positions in the world from where interesting things can be looked at or accessed in any other way; areas of interest, where those intersting things are located; objects, that may provide information or offer particular kinds of functionality; persons, that is other users thatare visiting the world; and even text, which might be on billboards or slides.

Some of this information is, so to speak, hard-wired in the model and may be accessed anytime, in some cases even by scanning the VRML file. Other information, however, is of a more dynamic nature, since it might be due to the presence of multiple users, the execution of scripts, or events that happen in response to user interaction. Some information may even be explicitly hidden, such as for example the actions one should take in solving a puzzle or playing a game.

When the virtual world is loaded, all the information (or at least most of it) is present in the so-called scenegraph, the structure thatis built to render the world. Using the software interface to access the scengraph (which is usually browser-specific), we can look for annotations, node types and textual content to extract information from the world. This information may then be stored in

a database, and be reused later for other users and queries. In principle, more advanced techniques could be used to extract information from the materials used, and even from textures and geometry.

presentation issues

In our prototype, we aimed at solving the question how to present the results of a query, using navigation. First of all, we had to

choose a metaphor

for navigation. Dependent on the object of interest a viewpoint can be selected. For a viewpoint, it is just that viewpoint. For an area of interest, the viewpoint selected must enable the user to view the area, and when objects or persons are chosen, care must be taken not to block the users' view by some obstacle.

Now answering a query then comes down to planning a suitable route and apply a series of viewpoint transformations along that route.
Not surprisingly, the navigation metaphor we chose was

walking

as the preferred mode of viewpoint transformtions.

the prototype

The structure of the prototype is depicted in the figure below.

In realizing the prototype,we made the following (simplifying) assumptions.

We avoided a number of difficulties by choosing for explicit annotations (which indicate locations and areas of interest), and by avoiding the intricacies of route planning and advanced text processing.

The requirements laid down before hand just stated that we would have a database and that we would avoid superfluous user interface elements. Instead, we used control and input panels written in VRML, in order to provide a 3D(pseudo-immersive) interface.

Now, our assumptions may in principle be relaxed. For example, annotation might be done incrementally by users that visit the world or to some extent even automatically, by using feature extractors. Instead if explicit maps, we may dynamically create maps based on users' navigation patterns. And, instead of simple keyword matching, we may apply more advanced text retrieval techniques. But this is left as future work. Anyway, we were satisfied that we could state the following conclusions:

conclusions

- navigation by query is feasible and may help users to find locations and objects
- determining suitable navigation routes without an explicitly defined map is hard

As is often the result with good research, you solve one problem and a number of other problems come up. So, one of the questions that remains was: how can we improve on navigation? What additional navigation support can we provide?

research directions– *extended user interfaces*

Is desktop VR a suitable candidate as an interface technology for multimedia information systems? And if so, what needs to be done to apply this technology effectively?

At first sight, our vision of applying VR as an interface to multimedia systems seems to be doomed to fail. As Ben Schneiderman, in a keynote for the Web3D Symposium 2002, observes:

3D GUI

> *Wishful thinking about the widespread adoption of three-dimensional interfaces has not helped spawn winning applications. Success stories with three-dimensional games do not translate into broad acceptance of head-tracking immersive virtual reality. To accelerate adoption of advanced interfaces, designers must understand their appeal and performance benefits as well as honestly identify their deficits. We need to separate out the features that make 3D useful and understand how they help overcome the challenges of dis-orientation during navigation and distraction from occlusion.*
>
> Ben Shneiderman

So, even if advanced (3D) user interfaces might be useful, there are a number of questions to raise. Again, following Ben Schneiderman:

> *Does spatial memory improve with 3D layouts? Is it true that 3D is more natural and easier to learn? Careful empirical studies clarify why modest aspects of 3D, such as shading for buttons and overlapping of windows are helpful, but 3D bar charts and directory structures are not. 3D sometimes pays off for medical imagery, chemical molecules, and architecture, but has yet to prove beneficial for performance measures in shopping or operating systems.*
>
> Ben Shneiderman

In particular, according to Schneiderman, we must beware of *tacky 3D*, gadgets in 3D space that are superfluous and only hindering the user to perform a task. Well-spoken and based on adequate observations! Nevertheless, at this stage, we should (in my opinion) adopt a slightly more liberal attitude and explore in what ways the presentation of (multimedia) information could be augmented by using (desktop) VR. But enough about *augmentation*. Let's discuss technology, and investigate what is required for the effective deployment of VR from the point of view of intelligent agents!

7

8.3 intelligent agents

Visitors in virtual environments are often represented by so-called avatars. Wouldn't it be nice to have intelligent avatars that can show you around, and tell you more about the (virtual) world you're in.

Now, this is how the idea came up to merge the RIF project, which was about information retrieval, with the WASP project, another acronym, which stands for:

WASP

Web Agent Support Program

The WASP project aims at realizing intelligent services using both client-side and server-side agents, and possibly multiple agents. The technical vehicle for realizing agents is the language DLP, which stands for

DLP

Distributed Logic Programming

Merging the two projects required providing the full VRML EAI API in DLP, so that DLP could be used for programming the dynamic aspects of VRML worlds.

background Historically, the WASP project precedes the RIF project, but we started working on it after the RIF project had already started. Merging these two projects had more consequences than we could predict at the time. The major consequence is that we shifted focus with respect to programming the dynamcis of virtual environments. Instead of scripts (in Javascript), Java (through the EAI), and even C++ (to program *blaxxun* Community Server extensions), we introduced the distributed logic programming language DLP as a uniform computational platform. In particular, for programming inteligent agents a logic programming language is much more suited than any other language. All we had to do was

merge DLP with VRML, which we did by lifting the Java EAI to DLP, so that function calls are available as built-ins in the logic programming language.

When experimenting with agents, and in particular communication between agents, we found that communication between agents may be used to maintain a shared state between multiple users. The idea is simple, for each user there is an agent that observes the world using its 'sensors' and that may change the world using its 'effectors'. When it is notified by some other agent (that is co-located with some other user) it can update its world, according to the notification. Enough background and ideas. Let's look at the prototypes that we developed.

8

multi-user soccer game

To demonstrate the viability of our approach we developed a multi-user soccer game, using the DLP+VRML platform.

We chose for this particular application because it offers us a range of challenges.

multi-user soccer game

- *multiple (human) users* – may join during the game
- *multiple agents* – to participate in the game (e.g. as goalkeeper)
- *reactivity* – players (users and agents) have to react quickly
- *cooperation/competition* – requires 'intelligent' communication
- *dynamic behavior* – sufficiently complex 3D scenes, including the dynamic behavior of the ball

Without going into detail, just imagine that you and some others wish to participate in a game, but there are no other players thatwant to join. No problem, we

just add some intelligent agent football players. And they might as well be taken out when other (human) players announce themselves.

For each agent player, dependent on its role (which might be *goal-keeper*, *defender*, *mid-fielder* and *forward*), a simple coginitive loop is defined: sensing, thinking, acting. Based on the information the agent gets, which includes the agent's position, the location of the ball, and the location of the goal, the agents decides which action to take. This can be expressed rather succinctly as rules in the logic programming formalism, and also the actions can be effected using the built-in VRML functionality of DLP.

Basically, the VRML-related built-ins allow for obtaining and modifying the values of *control points* in the VRML world.

control points

- get/set – position, rotation, viewpoint

These control points are in fact the identifiable nodes in the scenegraph (that is, technically, the nodes that have been given a name using the DEF construct).

This approach allows us to take an arbitrarily complex VRML world and manipulate it using the control points. On the other hand, there are also built-ins that allow for the creation of objects in VRML. In that case, we have much finer control from the logic programming language.

All in all we estimate that, in comparison with other approaches, programming such a game in DLP takes far less time than it would have taken using the basic programming capabilities of VRML.

agents in virtual environments

Let us analyse in somewhat more detail why agents in virtual environments may be useful. First of all, observe that the phrase *agents in virtual environments* has two shades of meaning:

agents in virtual environments

- virtual environments with embedded autonomous agents
- virtual environments supported by ACL communication

where ACL stands for *Agent Communication Language*. Our idea, basically is to use an ACL for realizing shared objects, such as for example the ball in the soccer game.

The general concept op multi-user virtual environments (in VRML) has been studied by the *Living Worlds Working Group*. Let's look at some definitions provided by this working group first. A *scene* is defined as a geometrically bounded, continuously navigable part of the world. Then, more specifically a *world* is defined as a collection of (linked) scenes.

Now, multi-user virtual environments distinguish themselves from single-user virtual environments by allowing for so-called *Shared Objects* in scenes, that is objects that can be seen and interacted with by multiple independent users, simultaneously. This requires synchronization among multiple clients, which may either be realized through a server or through client-to-client communication.

Commonly, a distinction is made between a *pilot* object and a *drone* object.

Shared Object

- *pilot* – instance that will be replicated
- *drone* – instance that replicates pilot

So, generally speaking, pilot objects control drone objects. There are many ways to realize a pilot-drone replication scheme. We have chosen to use agent technology, and correspondingly we make a distinction between *pilot agents*, that control the state of a shared object, and *drone agents*, that merely replicate the state of a shared object.

Since we have (for example in the soccer game) different types of shared objects, we make a further distinction between agents (for each of which there is a pilot and a drone version). So, we have *object agents*, which control a single shared object (like the soccerball). For these agents the pilot is at the server, and the drone is at the client. We further have agents that control the users' avatars, for which the pilot at user/client side, and the drone either at the server or the client. Finally, we have autonomous agents, like football players, with their own avatar. For those agents, the pilot is at the server, and the drones at the clients.

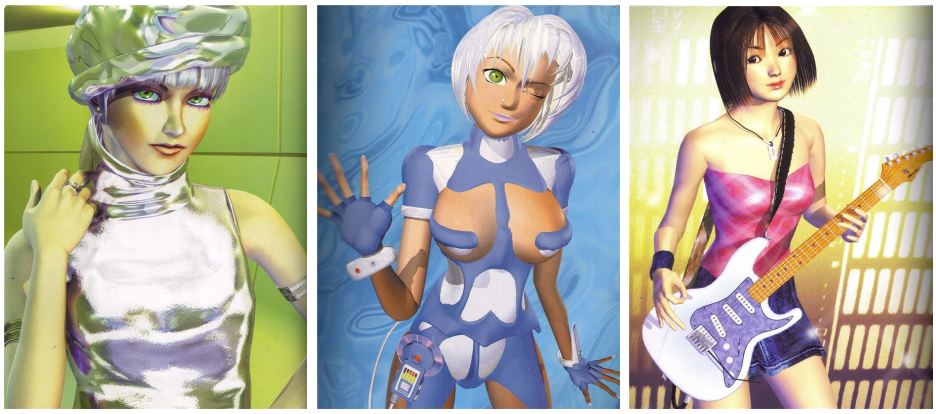
Now, this classification of agents gives us a setup that allows for the realization of shared objects in virtual environments in an efficient manner. See Huang et al. (2002) for details.

The programming platform needed to implement our proposal must satisfy the following requirements.

programming platform

- VRML EAI support
- distributed communication capabilities (TCP/IP)
- multiple threads of control – for multiple shared objects
- declarative language – for agent support

So, we adapted the distributed logic programming language DLP (which in its own right may be called an agent-oriented language *avant la lettre*), to include VRML capabilities. See the online reference to the AVID project for a further elaboration of these concepts.



9

PAMELA The WASP project's chief focus is to develop architectural support for web-aware (multi) agent systems. So, when we (finally) got started with the project we developed a taxonomy along the following dimensions:

taxonomy of agents

- 2D/3D – to distinguish between text-based and avatar embodied agents
- client/server – to indicate where agents reside
- single/multi – as a measure of complexity

A classification along these dimensions results in a lattice, with as the most complex category a *3D-server-multi-agent system*, of which the distributed soccer game is an example. See Huang et al. (2000).

When we restrict ourselves to *3D-client-single-agent systems*, we may think of,for example, navigation or presentation agents, that may help the user to roam around in the world, or that provide support for presenting the results of a query as objects in a 3D scene.

Our original demonstrator for the WASP project was an agent of the latter kind, with the nickname *PAMELA*, which is an acronym for:

PAMELA

Personal Assistent for Multimedia Electronic Archives

The PAMELA functional requirements included: autonomous and on-demand search capabilities, (user and system) modifiablepreferences, and multimedia presentation facilities. It was, however, only later that we added the requirement that PAMELA should be able to live in 3D space.

In a similar way as the soccer players, PAMELA has control over objects in 3D space. PAMELA now also provides animation facilities for its avatar embodiment.

To realize the PAMELA representative, we studied how to effect facial animations and body movements following the *Humanoid Animation Working Group* proposal.

H-Anim

- control points – joints, limbs and facial features

The H-Anim proposal lists a number of control points for (the representation of the) human body and face, that may be manipulated upto six degrees of freedom. Six degrees of freedom allows for movement and rotation along any of the X,Y,Z axes. In practice, movement and rotation for body and face control points will be constrained though.

presentation agent Now, just imagine how such an assistant could be of help in multimedia information retrieval.

presentation agent

> Given any collection of results, PAMELA could design some spatial layout and select suitableobject types, including for example color-based relevance cues, to present the results in a scene. PAMELA could then navigate you through the scene, indicating the possible relevance of particular results.

persuasion games But we could go one step further than this and, taking inspiration fromthe research field of *persuasive technology*, think about possible persuasion games we could play, using the (facial and body) animation facilities of PAMELA:

persuasion games

- single avatar persuasive argumentation
- multiple avatar dialog games

Just think of a news readerpresenting a hot news item. or a news reader trying to provoke a comment on some hot issue. Playing another trick on the PAMELA acronym, we could think of

PAMELA

Persuasive Agent with Multimedia Enlightened Arguments

I agree, this sounds too flashy for my taste as well. But, what this finale is meant to express is, simply, that I see it as a challenge to create such synthetic actors using the DLP+VRML platform.

10

research directions– *embodied conversational agents*

A variety of applications may benefit from deploying embodied conversational agents, either in the form of animated humanoid avatars or, more simply, as a 'talking head'. An interesting example is provided by *Signing Avatar*, a system that allows for translating arbitrary text in both spoken language and sign language for the deaf, presented by animated humanoid avatars. Here the use of animated avatars is essential to communicate with a particular group of users, using the sign language for the deaf.

Other applications of embodied conversational agents include e-commerce and social marketing, although in these cases it may not always be evident that animated avatars or faces actually do provide added value.

Another usage of embodied conversational agents may be observed in virtual environments such as Active Worlds, *blaxxun* Community and Adobe Atmosphere. Despite the rich literary background of such environments, including Neil Stephenson's Snow Crash, the functionality of such agents is usually rather shallow, due to the poor repertoire of gestures and movements on the one hand and the restricted computational model underlying these agents on the other hand. In effect, the definition of agent avatars in virtual environments generally relies on a proprietary scripting language which, as in the case of *blaxxun* Agents, offers only limited pattern matching and a fixed repertoire of built-in actions.

In contrast, the scripting language for *Signing Avatar* is based on the H-Anim standard and allows for a precise definition of a complex repertoire of gestures, as exemplified by the sign language for the deaf. Nevertheless, also this scripting language is of a proprietary nature and does not allow for higher-order abstractions of semantically meaningful behavior.

scripting behavior In this section we introduced a software platform for agants. This platform not only offers powerful computational capabilities but also an expressive scripting language (STEP) for defining gestures and driving the behavior of our humanoid agent avatars.

The design of the scripting language was motivated by the requirements listed below.

STEP

- *convenience* – for non-professional authors
- *compositional semantics* – combining operations
- *re-definability* – for high-level specification of actions
- *parametrization* – for the adaptation of actions
- *interaction* – with a (virtual) environment

Our scripting language STEP meets these requirements. STEP is based on dynamic logic Harel (1984) and allows for arbitrary abstractions using the primitives and composition operators provided by our logic. STEP is implemented on top of DLP,

As a last bit of propaganda:

DLP+X3D

> The DLP+X3D platform provides together with the STEP scripting language the computational facilities for defining semantically meaningful behaviors and allows for a rich presentational environment, in particular 3D virtual environments that may include streaming video, text and speech.

See appendix D for more details.

evaluation criteria The primary criterium against which to evaluate applications that involve embodied conversational agents is whether the application becomes more effective by using such agents. Effective, in terms of communication with the user. Evidently, for the *Signing Avatar* application this seems to be

quite obvious. For other applications, for example negotiation in e-commerce, this question might be more difficult to answer.

As concerns the embedding of conversationl agents in VR, we might make a distinction between *presentational VR*, *instructional VR* and *educational VR*. An example of educational VR is described in Johnson et al. (2002). No mention of agents was made in the latter reference though. In instructional VR, explaining for example the use of a machine, the appearance of a conversational agent seems to be quite natural. In presentational VR, however, the appearance of such agents might be considered as no more than a gimmick.

Considering the use of agents in applications in general, we must make a distinction between *information agents*, *presentation agents* and *conversational agents*. Although the boundaries between these categories are not clearcut, there seems to be an increasing degree of interactivity with the user.

From a system perspective, we might be interested in what range of agent categories the system covers. Does it provide support for managing information and possibly information retrieval? Another issue in this regard could be whether the system is built around open standards, such as XML and X3D, to allow for the incorporation of a variety of content.

Last but not least, from a user perspective, what seems to matter most is the naturalness of the (conversational) agents. This is determined by the graphical quality, as well as contextual parameters, that is how well the agent is embedded in its environment. More important even are emotive parameters, that is the mood and style (in gestures and possibly speech) with which the agents manifest themselves. In other words, the properties that determine whether an agent is (really) convincing.

8.4 development(s) – the metaverse revisited

When creating presence in Second Life, as discussed in section 1.4, our initial targets were

initial target(s)

- build initial (throwaway) prototype
- explore content creation technology
- create tutorial(s) for content contribution
- analyse technological requirements

After this first meeting, we put an announcement on some student mailinglists, and two and a half months later we were online, with a virtual campus, that contains a lecture room, a telehub from which teleports are possible to other places in the building, billboards containing snapshots of our university's website from which the visitors can access the actual website, as well as a botanical garden mimicking the VU Hortus, and even a white-walled experimentation room suggesting a 'real' scientific laboratory. All building and scripting were done by a group of four students, from all faculties involved, with a weekly walkthrough in our 'builders-meeting' to re-assess our goals and solve technical and design issues.

(a) outside view (b) inside view

11

As can be seen in the figure above, the overall style is realistic, although not in all detail. Most important was to create a visual impression of resemblance and to offer the opportunity to present relevant infomation in easily accessible, yet immersive, ways. Cf. Bolter and Grusin (2000). Our virtual campus, as depicted above, is meant to serve as an *information portal* and as a *meeting ground*, where students, staff and visitors can meet and communicate, as well as a place were teachers and researchers can conduct experiments aimed at discovering new ways of teaching and doing research.

In Eliens et al. (2007), we looked at the shift of culture that made the growth of Second Life possible, and the background against which (the phenomenon of) Second Life could be understood. In particular, we asked ourselves why earlier attempts at introcing (3D) virtual environments failed, and observed that: in less than a decade after the publication of William Gibson's novel *Neuromancer*, the *metaverse* was realized, albeit in a primitive way, through the introduction of VRML[2], introduced at the Int. Web Conference of 1992. Cf. Anders (1999). The German company *blaxxun*[3], named after the virtual environment in Neil Stephenson's *Snowcrash*, was one of the first to offer a 3D community platform, soon to be followed by *AlphaWorld*[4], already mentioned in the introduction, which offered a more rich repertoire of avatar gestures as well as limited in-game building facilities. However, somehow 3D virtual communities never seemed to realize their initial promises. Furthermore the adoption of VRML as a 3D interface to the Web never really took off.

The history of Second Life is extensively descibed in the official Second Life guide, Rymaszweski et al. (2007). Beginning 2004, almost out of the blue, Second Life appeared with a high adoption and low churn rate, now counting, March 2007, over 4 million inhabitants. Considering the cost of ownership of land, which easily amounts to 200 euro per month rent after an initial investment of 1500 euro for a single piece of land measuring 65,536 square meters, the adoption of Second Life

[2] www.web3d.org
[3] www.blaxxun.com
[4] www.activeworlds.com/worlds/alphaworld

by individuals as well as companies such as ABN-AMRO, Philips and institutions such as Harvard is surprising.

What is the secret of the success of Second Life?, we asked in Eliens et al. (2007), and we immediately confessed: *We don't know!* But in comparison to other platforms for immersive worlds, including MMORPGs such as *World of Warcraft*[5] and *Everquest*[6], Second Life seems to offer an optimal combination of avatar modification options, gesture animations, in-game construction tools, and facilities for communication and social networking, such as chatting and instant messaging. Incorporating elements of community formation, commonly denoted as Web 2.0, and exemplified in MySpace, YouTube and Flickr, the immersive appearance, perhaps also the built-in physics and the inclusion of elementary economic principles, seem to be the prime distinguishing factors responsible for the success of Second Life. In addition, the possibility of recording collaborative enacted stories, Stories, using built-in *machinima*[7] certainly also contributes to its appeal.

What has been characterized as a shift of culture, from a media consumer culture to a participatory culture, Jenkins (2006), where users also actively contribute content, (*was*) for our institution one of the decisive reasons to create a presence in Second Life, to build a virtual platform that may embody our so-called *community of learners*, where both staff and students cooperate in contributing content, content related to our sciences, that is. Basically following up on companies like Nike, ING and ABN-AMRO, from which we, incidentally, borrowed the island on which we built our virtual campus.

The 1st of March 2007, we went live. In the evening there was a news item on national televison, RTL4 news, featuring the students showing the virtual campus and our project leader explaining the reasoning behind our presence in Second Life and how to give a course in the virtual classroom. A similar item appeared at AT5, local Amsterdam television, and various newspapers, among which Parool, Telegraaf and Volkskrant, spent a multiple-column article to report on our efforts. As a note, not surprisingly, all items focused on what we have characterized as the naive interpretation of our efforts, exemplifying the old credo *the medium is the message*. To be clear, as will be discussed below, our intention was not to provide a virtual replica, nor to provide an analogon of the Open University, in Second Life. After the news broadcasts, the number of visitors increased dramatically, having stayed at a modest below 100 during the day. In the evening, however, just after the news items on the national television, the number of visitors increased rapidly. Since at the time we did have only one island, it appeared to be very difficult to separate internal experimental activities from visitors just asking for additional information, and to exclude potentially malicious visitors. In that evening, we were even surprised by the invasion of an army of Mario Brothers. Hilarious and non-harmful. But enough reason to sit back and limit access to our campus for students and staff only the day after our open day. A few days later, after the first turbulent days following the TV broadcasts, we re-opened our virtual campus

[5] www.worldofwarcraft.com
[6] everquest.station.sony.com
[7] www.machinima.org

to allow visitors to walk/fly around, and enjoy our news items and informative videos.

The first idea that comes to mind, naturally, is to use Second Life to offer courses online. But, although we did have plans to give lectures (college) on law, probably including the enactment of a particular case, we did consider this approach as rather naive, and frankly I see no reason to include what may be considered an outdated paradigm of learning in our virtual campus, where there might be more appealing alternatives. Similarly, using the virtual laboratory for experiments might not be the best way to offer courses, although, again, we do intend to provide a model of a living cell, allowing students to study the structure, functionality and behavior of organic cells in virtual space.

Considering the success of our multi-disciplinary building team, it seems more worthwhile to take the cooperative effort of building as a model, and switch to a paradigm of learning in which in-game exploration and building plays an important role. As we observed in section 3.4, gaming may provide a form of *active learning*, that is allowing the gamer to

active learning

- experience the world in new ways
- form new affiliations
- preparate for future learning

This is due to intense involvement or immersion in the game environment, which even encourages *critical learning* or as we characterized it, following Gee (2003), *situated cognition in a semiotic domain*, that is a *world of meaning*. What is this *world of meaning* that a game exemplifies, and how is it related to the more general notion(s) of *immersion* and *flow*?

We explored the use of 3D desktop VR for presenting artworks in context, using 3D not to contruct a replica of physical space, but as a means to allow immersive access to both (representations of) artworks and information about these artworks. In Dossier, we wrote: *the* abramovic *dossier presents itself as a digital archive in 3D space, containing information about the artworks of the performance artist Marina Abramovic by presenting media content and relational structures, serving as an information source for museum curators to conserve and install the artworks.* As a follow-up on the *abramovic* dossier, the 2005 Casus group developed a digital dossier for the artist Jeffrey Shaw[8]. One interesting aspect of the dossier for Shaw is the availability of a tool environment to learn about the construction and de-construction of the Revolution[9] installation and to experiment with the exhibition space parameters of the artwork, such as the lighting conditions, and the color and texture of the walls and the floor. In Dossier we further observed that with the Casus 2005 group there was, interestingly, a strong resistance against using 3D for the concept graph navigation mechanism. So we explored a mixed approach, using 2D for the concept graph, and 3D only for the representation of the Revolution installation. Nevertheless, although the dossier for Shaw does realize many of the goals set for the next generation

[8] www.few.vu.nl/∼casus05

[9] www.medienkunstnetz.de/works/revolution

dossier, see section 10.2, it did fail in providing an immersive application. It did not achieve a natural transition between browsing the concept space and inspecting/experiencing the media recordings of the artwork, thus disrupting the natural flow of attention ...

12

questions

virtual environments

1. (*) Discuss how *virtual environments* may be used for giving access to *(multimedia) information*. Give a brief characterization of *virtual environments*, and indicate how *information (hyper) spaces* may be projected in a virtual environment.

concepts

2. What is meant by *virtual context*?
3. Give an example of *navigation by query*, and indicate its possible advantages.
4. Discuss the deployment of *(intelligente) navigation agents*.

technology

5. Give a brief characterization of: VRML.
6. What is a *viewpoint transformation*?
7. What kinds of navigation can you think of?
8. How may intelligent avatars be realized? Give an example.

projects & further reading As a project, I suggest the implementation of storytelling in virtual environments, with (possibly) an embodied agent as the narrator. You may further explore or evaluate the role of agents in multimedia applications and virtual environments.

For further reading in (real) VR, I advice Sherman and Craig (2003), and for gaining an understanding in story telling and applications you may try to get hold of the proceedins, of TIDSE 2003[10], and TIDSE 2004[11].

the artwork

1. another series of *dutch light*[12].
2. *virtual context* – Dam Square, Amsterdam, see 8.1.
3. VU Campus in VRML – student project.

[10] www.zgdv.de/TIDSE03
[11] www.zgdv.de/TIDSE04
[12] www.dutchlight.nl

4. CWI 3th floor, floormap and model, see 8.2..
5. query – on 3th floor of CWI.
6. navigation – on 3th floor of CWI.
7. soccer game – image from WASP project, see section 8.3.
8. *digital beauties* – taken from Wiedermann (2002).
9. *digital beauties* – taken from Wiedermann (2002).
10. VU @ Second Life – screenshots.
11. signs – sports, van Rooijen (2003), p. 276, 277.

Another sequence of *dutch light*, opening this chapter, is meant to make you wonder about *realism*. Is virtual reality less 'real'? With a reference to section 2.3, where I quoted Bolter and Grusin (2000) on *re-mediation*, I may remark that the graphic style chosen for presenting the virtual environment strongly determines whether the environment is experienced as 'realistic'. In our culture this is generally a *photorealistic style*, as for example in the *Mission Rehearsal Exercise* discussed in the next chapter, section 9.2. The *digital beauties* are not only a pleasure to look at, but do also display a wide range of postures and styles.

9. digital content creation

post-modern design allows for sampling

learning objectives

After reading this chapter you should be able to mention some basic rules of digital content creation, discuss what criteria your portfolio should meet, describe how you would approach the design of a logo, explain the notion of user-centered design, and characterize the issues that play a role in dveloping multimedia for theatre.

Whether your ambition is to become a professional designer or not, also for students of information science and computer science, a course in visual design is a must, I think.

In this chapter, we will treat various aspects of digital content creation. The first section discusses how to approach visual design and gives a number of basic design assignments, that can be used to get experience with visual design. Section 2 discusses the issue of workflow and tools, and investigates how design fits in with the process of developing multimedia applications. In the final section, I will elaborate on a theatre project I was involved in, for which I had to develop an augmented reality application.

1

9.1 visual design

When you are trained as a visual artist, as I once was (in the pred-digital era), you must do many basic exercises with form and color, learn the skills of drawing

and painting, including sill lifes and portraits. And, to graduate you must make an exhibition of your work over the years.

In the digital era, things have changed. The skills that you need to learn, which include the use of tools for digital content creation, as well as the artistic goals, I would say. What has not changed, however, is the need for basic exercises and the presentation of your work, that is the creation of a portfolio.

In this section, I will give an outline of the course *visual design*, that I started in februari 2005. This outline is meant to give you a general idea of how to approach design, and to give you some hints on how to acquire the skills needed to act as a designer. If you have the opportunity to take a course in visual design, then don't hesitate and do it!

2

perspective(s)

The overall goal of the *visual design* course is to establish some basic *aesthetic awareness*, by providing suitable exercises and assignments. In addition, the student is supposed to become familiar with the craft of design, which necessarily, but not exclusively, involves the use of tools and techniques.

To accomodate for the various interests and backgrounds of the students, we distinguish between several *tracks* or perspectives, as summarized below:

track(s) – perspective

- styling – concept and presentation
- digital content – material, animation
- tech track – special effects

To illustrate the various perspectives, I invited guest speakers who showed their work and talked about their approach.[13] The process of design is very complex, ranging from conceptual explorations and sketching up to the stage of finalizing delivery. It is also very personal. However, as each of the speakers testified to, a significant portion of the time and effort goes into negotiating with the client. Whether it involves taking photographs for an advertisement or setting up a campaign, it takes a lot of back and forth to get an idea of what the client wants. In the agency of one of the speakers Mark Veldhuijzen van Zanten, they have created roles to help eachother come up with the right ideas.

www.178aardigeontwerpers.nl

[13] A selection of their works is hown in this book.

- *e-motionist* – make emotion, rhythm and movement flow together
- *chaoticus* – who sees chaos within order
- *formologist* – who approaches the fabrication of forms as an art
- *infonaut* – who moves in the twilight zone of information and meaning
- *transformator* – who transforms images and concepts into new matter

Although slightly ironic, and in practice not so clear-cut, these roles give you an idea what cognitive modes are involved in bringing an original concept to a stage of finalization.

Coming up with an idea and sketching require a more reflective cognitive mode, whereas finalizing a design requires a more experience-related cognitive mode. It is important to choose the right tools to work with, dependent on which phase you're in, for example paper and pencil when you are still sketching and Photoshop or maya when you put your ideas into production.

3

deliverable(s)

However interesting the process may be, design is not about process but about product(s). Such a product may be, dependent on what you ar good at, one of the following:

products

- web site – e.g. conference, campaign (browse)
- 2D/3D animation – promotion/ad (temporal sequence)
- virtual space – game/infotainment (navigate)
- ebook – story (sequential experience)

As indicated between brackets, each of the products favors a particular mode of interaction. Although *interaction* is not an aspect of visual design as such, it is an important aspect to take into account. In section 9.2, we will look in more detail at the issue of interaction and usability in general.

One easily overlooked issue in a design project, is the creation of a portfolio. There may easily be some confusion here with regard to what should be considered the product of design, one of the items in the list above, or your portfolio. The answer is simple. Both! As a record of the process of design, the portfolio is itself a product of design.

portfolio – design as a product

- concept(s)
- sketches & explorations

- finalized products
- evaluation & reflection

What criteria should a portfolio meet? Well, nowadays it is not only common to have your portfolio on the web, I have been told that you cannot do without it. So, first of all, your portfolio should be web-friendly. And your work should not be too many clicks away! And, secondly, it should give sufficient insight in what you have to offer, so that a potential client can decide whether it is worth the effort to contact you.

In our visual design course, I require that the portfolio contains a description of the concept(s) underlying the design, sketches and explorations as well as the finalized products. It is also required, after all it is an academic course, to provide an evaluation and some reflection in the form of an essay on a topic such as *2D versus 3D aesthetics*, *animation techniques*, or (more theoretical), *elements of style* or *theories of creativity.*

www.jaapstahlie.com

> In my perception a portfolio is about the
> past and I feel much more related to the
> present especially in my work as a
> photographer. To me the relation with
> the present and the subject/assignment directs
> my creativity, the experiences over the
> past draw my skills. My challenge is to be
> truly inspirred, to be present in the present.

However, as illustrated by the motivation Jaap Stahlie gives with his portfolio (above), it is perhaps wise not to overdo it!

There are basic exercises, obligatory for all students, and a final assignment, where you have a choice between three productions, each with a different supervisor. In addition, as explained in the guidelines, all students must write an essay, and give a presentation in class. For deadlines, see the schedule. There will be periodic checks on the status of your work. Each year there will be recommended themes.

assignment(s)

In the *visual design* course there are basic exercises, obligatory for all students, and a final assignment, where the student has a choice between three productions.

basic exercises

1. develop a logo
2. create a sign
3. design a collage
4. write a story

For the final assignment, there is a choice between the following assignments: developing a house style, creating a non-linear visual story, and designing a suitable game environment For this assignment) the students are allowed to work in groups. However, the contribution of each individual must be reflected in his/her own portfolio.

4

regulation(s)

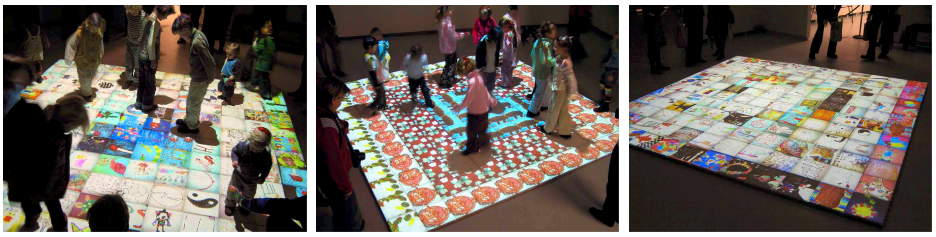
The first requirement when working in an area such as visual design is that you acquire sufficient self-discipline to find the challenge in the assignments and to complete the tasks involved. Since a department of computer science is not the natural habitat for a course in visual design, I have laid down some strict rules:

rules

- be present – 2 omissions max.
- be in time – hard deadlines
- be online – have your portfolio available
- be creative – don't steal without a reason/mentioning
- be smart – there is no 2nd chance

These rules may well apply when you work, after you graduated, as an individual designer/developer or as a member of a team in some agency. To finish this section, I may remark that design is an interesting field, full of implicit (not

always so obvious) wisdom and apparent paradoxes. Whatever you do, deliver! Silence is lethal. And, as another item of colloquial wisdom, be authentic, but only if possible!



5

research directions– *on creativity*

One of the assignments in the *visual design* course is an essay. As one of the recommended topics we have a reflection on theories of creativity. As such it is not a training in creativity[14]. This section contains some random thoughts on the processes and products of design, and ends with the provoking statement *there is no theory of creativity*. You may, however, try to find some counter-arguments, for example in the line of Hewlett and Selfridge-Field (1998).

multimediocrity Multimedia is a promising technology, and (nowadays) affordable. So we see that multimedia (which includes 3D-graphics, video and sound) is increasingly being used, also in information visualisation. But what is it good for? To quote Klabbers (2006):

> *multimedia's promise is terribly generalized, it simply lets you do anything.*

As with any new technology, the early multimedia productions (in particular CDROM and CD-I) were not optimal with respect to (aesthetic) quality. To quote Klabbers (2006), again:

shovelware – *multimediocrity*

> ... far from making a killing, it looked as if the big boys ... had killed the industry by glutting the market with inferior products.

Perhaps the industry in the late eighties did not have the right business model. But, then again, what are the chances of multimedia in our time. One more quote from Klabbers (2006):

> if multimedia is comparable to print then yes, we'd be crazy to expect it to mature in a mere ten years.

[14] www.goshen.edu/%7Emarvinpb/arted/tc.html

eliminating complexity So now, in the new millenium, we are (sadder and wiser) in a position to approach the effective deployment of multimedia afresh. What we look for is aesthetic quality. How do we find it? Easy enough, just be authentic.

> *"Learning how to not fool ourselves is, I'm sorry to say, something that we haven't specifically included in any particular course that I know of. We just hope you've caught it by osmosis."* Richard Feynman

Authentic in creating multimedia means, apart from not fooling yourselves, that you must become aware of the message orinformation you want to convey and learn to master the technology to a sufficient degree. But beware, an effective multimedia presentation is not the same as scientific argumentation:

the media equation

> *We regularly exploit the media equation for enjoyment by the willing suspension of our critical faculties. Theatre is the projection of a story through the window of a stage, and typically the audience gets immersed in the story as if it was real.*

These quotes, as well as the following one have been taken from an online essay on *eliminating complexity* which provides an argument against inessential gadgets and spurious complexity and bells and whistles in whatever you can think of, including user interfaces and scientific theories. Back to the subject, what does *master the technology to a sufficient degree* mean? Just remember that what you do is a form of engineering.

> *"engineering is the art of moulding materials we do not wholly understand ... in such a way that the community at large has no reason to suspect the extent of our ignorance."* A. R. Dykes.

In other words, learn the tool(s) that you are using to a degree that you master the basics and easily cut through its apparent magic.

theories of creativity Producing multimedia, in whatever form, has an element of craftsmanship. But, given the need for aesthetic quality, whatever way you approach it, there is an element of creativity. That means, you're in for a challenge. And, to quote Klabbers (2006),

> *The best thing is to empower yourself. But before you can do that, you need to understand what you are doing – which is a surprisingly novel thing to do.*

Now it is tempting to look for a set of guidelines and rules that give you a key to creativity. So let me be straight with you:

there is no theory of creativity

On the other hand, there are techniques for producing ideas. And some recommend a sequence of steps, such as:

steps

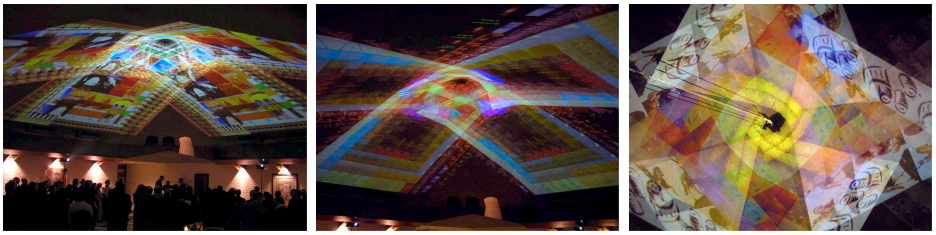
> browse, explore; chew it over; incubation, let it rest; illumination (YES); verification,*does it work?*

And in addition, still following Klabbers (2006), there are some general rules:

general rules

- *if you aim to please everybody, you will please nobody*
- *constraints come with the territory, you must learn to love them*
- *emotional charge is the key to success*

Now, if you'd ask me, I would say, just
make your virtual hands dirty.



6

9.2 designing the user experience

In a time in which so much information is available as in ours, we may make statements like

postmodern design

> ... postmodern design is of a highly reflective nature ... appropriating design of the past ... in other words, sampling is allowed but no plagianarism

One interpretation might be that it is impossible to be original. But another interpretation might be that it is undesirable to be (totally) original. In other words, as discussed in section 2.3, it is necessary that your design contains references to not only some real situation but also to other work, so that it can be understood and experienced by the user/spectator/client whatever you like to call the people that look at your work. As observed in Sherman and Craig (2003), designing for multimedia does not take into account only the technological or aesthetic issues, but also constraints on what people can perceive and what the experiential context is in which the work is presented, which may be re-phrased more plainly as what expectations the user has.

game design

Let us consider how these observations affect one of the project assignments for our *visual design* course. Also for *game design*, there are several options, dependent on the track the student is most comfortable with.

game design

- style – develop concept, plot and visual assets for *a game of choice*
- content – develop environments, models and animations for *a game of choice*
- effects – develop models, textures and special effects (shaders) for *a game of choice*

To explain, *style* may be considered to involve the whole gamut of concepts, plot and genre, as well as the visual assets or props of the game, those things by which the game differentiates itself from other games. Since the reader of this book will probably be more familiar with games than the author, there is no need to expand on these issues. *Content* is concerned with the actual game environment, including the models and animations. Finally, *effects*, to simplify things, is everything else, those things that are visual but does not belong to the story line or game environment.

Games, perhaps more than any other multimedia application, are appealing, not because they are useful, although they might be, but because the user gets emotionally involved, not to say addicted. Now, following Norman (2004),

> did you ever wonder why cheap wine tastes better in fancy glasses?

Exactly, because cheap glasses do not give us the same emotion. It is, indeed, a matter of style!

Obviously, games are played for fun. As applications, games may be classified as seductive, which is, see section 2.3, stronger than persuasive. Norman (2004) distinguishes between four categories of pleasure.

seduction

- physio-pleasure – of the body
- socio-pleasure – by interaction with others
- psycho-pleasure – due to use of the product
- ideo-pleasure – reflecting on the experience

In other words, games are seductive, or fun to play, because they arouse any combination of pleasure from the categories above. Which combination depends on the kind or genre of game. Quoted from Norman (2004), but originally from Wolf we can list, not exhaustively, the following genres of video game:

genre(s)[15]

> Abstract, Adaptation, Adventure, Artificial Life, Board Games, Capturing, Card Games, Catching, Chase, Collecting, Combat, Demo, Diagnostic, Dodging, Driving, Educational, Escape, Fighting, Flying, Gambling, Interactive Movie, Management Simulation, Maze, Obstacle Course, Pencil-and-Paper Games, Pinball, Platform, Programming Games, Puzzle, Quiz, Racing, Role-Playing, Rhythm and Dance, Shoot Em Up, Simulation, Sports,

[15] www.robinlionheart.com/gamedev/genres.xhtml

Strategy, Table-Top Games, Target, Text Adventure, Training Simulation, and Utility.

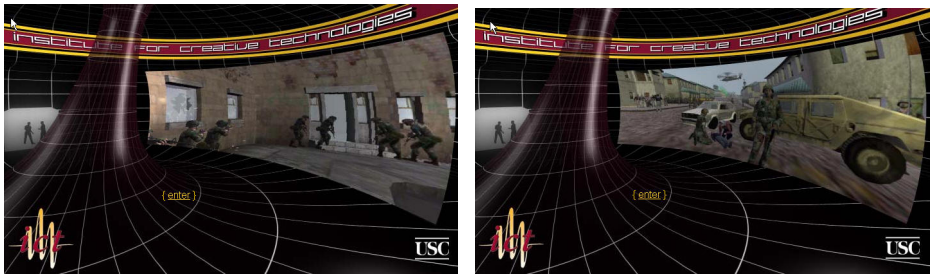
When you develop a game it is good to reflect on what genre your game belongs to, because that will directly affect the user's expectation when playing the game, due to the conventions and rules that exist within a particular genre. For video games, which can be characterized as *a mixture of interactive fiction with entertainment*, interaction is evidently another distinguishing factor in determining the success of the game.

Corresponding to the kind of pleasure a user may experience, Norman (2004) distuinguishes between three levels of design:

levels of design

- visceral – what appeals to our intuition (*affordance*)
- behavioral – is all about use (*performance*)
- reflective – its all about message, culture and meaning

Of these, the latter two should be rather obvious, although we will elaborate on the notion of *usability* later on. But what does *affordance* mean, and how is it related to our intuition.



7

affordance – ecology of behavior

The notion of affordance[16] has a long history. According to Don Norman, the word "affordance" was invented by the perceptual psychologist J. J. Gibson, to refer to the properties of the world that 'invite' actions, as for example a char invites one to sit. Originally, however, the notion of affordance dates back to the beginning of the 20th century, when it was used in phenomenologist philosophy to describe how the world around us presents itself as meaningful. Affordance, in other words, is a concept that explains why it seems natural for us to behave in a particular way, either because it is innate, as the reflex to close one's eyes by sudden exposure to light, or because we have learned that behavior, as for example surfing the web by clicking on links. In game or product design, thinking about 'affordance' may help us to find the most natural way to perform certain

[16] www.jnd.org/dn.mss/affordances-and-design.html

actions. Natural is in this context perhaps not the best phrase. What we must take into account is what is perceived as an affordance, and how actions fit in with what we may call an exology of behavior (with the system).

How does this diversion in abstract philosophy help us design better games? To answer this question, I'd like to recount my visit at the Virtual Humans Workshop, held in october 2004 at the Institute of Creative Technologies[17], in Los Angeles. Of interest in particular is the ICT Games Project:

ICT Games Project

> The goal of the ICT games project is to develop immersive, interactive, real time training simulations to help the Army create a new generation of decision-making and leadership-development tools.

As further explained on the website:*with the cooperation of the U.S. Army Research, Development and Engineering Command, Simulation Technology Center (RDECOM STC), Training and Doctrine Command (TRADOC), and commercial game development companies, ICT is creating two training simulations that are intended to have the same holding power and repeat value as mainstream entertainment software.*

The two training applications developed by ICT are:

- Mission Rehearsal Exercise – to solve a potential conflict after a car accident
- Language Training Simulation – to learn how to contact local leaders in arabic

The *mission rehearsal exercise* is situated in former Yugoslavia. The trainee is confronted with the situation after a car accident in which a boy got injured. The mother of the boy is furious, and a potentially hostile crowd is waiting. An army convoy is on its way to a nearby airport, and needs to pass the crossing where the accident took place. The trainee must decide what to do and give appropriate orders, knowing that the wrong decision may lead to serious trouble.

The *language training simulation* is situated in the Middle-East, and is meant to teach the trainee not only some basic arabic but also proper ways of conduct, in conformance with local customs to gain confidence.

Both applications are highly realistic, with impressive graphics.[18] The both support speech input. The challenge in both simulation games was to come up with a natural way to indicate to the trainee what options for actions were available. Natural means, in this context, that it should fit within the simulation or game environment. Obviously, a menu or a row of pushbuttons does not fit naturally within such an environment, and would break what we have previously, in section 2.3 called 'immersion'.

I was invited at ICT for the Virtual Humans Workshop because of my involvement with embodied conversational agents (ECAs), as discussed in section 8.3. The topic of the workshop was, among others, to investigate whether the notion of affordance could help in analyzing and evaluating the interaction of a user/trainee with the simulation game. These were the questions we tackled:

[17] www.ict.usc.edu

[18] A lot of attention has been devoted to creating the models and environments. Both simulations are implemented using the Unreal game engine.

Virtual Humans Workshop[19]

- Is it more appropriate to construct a frame of analysis that encompasses both user and ECA in a single interaction graph?
- Is it fitting to think in terms of a fixed graph that the user comes to recognize, or is the graph itself a dynamic structure?
- Is it even appropriate to focus on "affordances to act," or is it more fitting to consider cues that influence the mental interpretations that lead to action (e.g., affordances of control, affordances of valence of potential outcomes, etc.)? How does this relate to intrinsic motivation?

This workshop was a follow-up on a seminar in Dagstuhl on Evaluating Embodied Conversational Agents[20], where we discussed the tipics of interaction and affordance in a special interest group. In the *research directions*, I will further discuss an evaluation study that we did on agent-supported navigation in a virtual environment.

Back to our question, how can affordance help us in designing a game? In the *mission rehearsal exercise*, described above, it would be much more easy to have a menu with all available options listed. However, such amenu would defeat the purpose of the simulation, since such menus will not likely occur in real life. Immersion is, in other words, necessary to maintain the emotional involvement with the application, and affordance is the key to immersion. But, although it sounds like an answer, it does rather lead to another question, *how can we define the usability of a game*?

8

usability and fun

In interaction design there is a clear, yet unresolved, tension between usability and fun. Usability is, despite the many disputes, a well-defined notion:

usability (ISO DIS 9241-11)

> ... the effectiveness, efficiency and satisfaction with which specified users can achieve particular goals in particular environments ...

[19] www.ict.usc.edu/∼vhumans/2004/
[20] wwwhome.cs.utwente.nl/∼zsofi//eeca

This is the ISO DIS 9241-11 definition, cited from Faulkner (2000). In sectiona 10.3 we will further investigate usability as a means to evaluate systems from an interaction perspective. Now, I wish to focus on why artefacts or games might be appealing even if these same aspects may compromise usability in the traditional interpretation.

In describing a fancy juice squeezer, designed by Philip Starck Norman (2004) observes, following KS, that is:

emotional involvement

- entices by diverting attention – unlike the common
- delivers surprising novelty – not identifiable to its function
- goes beyond obvious needs and expectations – it becomes something else
- creates an instinctive response – curiosity and confusion

The phrase *satisfaction* in the definition of usability above seems somewhat meagre to explain the emotional involvement with games, and even inappropriate as one realizes that, in the *mission rehearsal exercise*, frustration might actually be beneficial for the learning experience.

9

example(s) – *visual sensations*

The dutch *visual sensations*[21] festival is an annual contest for VJs. In 2005, in cooperation with the festival, a parallel seminar seminar was held discussing the topic of the history of VJ-ing, a aplenary discussion of the relation between club-VJs and the established art circuit. In addition there were two guest speakers, Geert Mul and Micha Klein[22], both visual artists who also have a ten-years experience as VJ.

[21] www.visualsensations.nl
[22] www.michaklein.com

10

Above is another work of Geert Mul, in cooperation with DJ Speedy J. It was shown a dance event in cooperation with Rotterdam Maritime Museum. On the right, the cranes are swinging on the rhythm of the music.

The portfolio of Geert Mul[23] starts with a quote from Arnheim (1957):

form and content

> Very often people assume that "form" as a concept is the opposite of something called "content". This assumption implies that a poem or a musical piece or a film is like a jug. An external shape, the jug, contains something that could just as easily be held in a cup or pail. Under this assumption, form becomes less important than whatever it is presumed to contain.
>
> We do not accept this assumption. If form is the total system, which the viewer attributes to the film, there is no inside or outside. Every component functions as part of the overall pattern that is perceived. Thus we shall treat as formal elements many things that some people consider content. From our standpoint, subject matter and abstract ideas all enter into the total system of the artwork (....)

I totally agree with this. And perhaps this is why I have a preference for artworks that are slightly out of the main stream of tradional art.

research directions– *engaging with fictional characters*

What do you need to evaluate your game or multimedia application? There are many ways to gain insight in how your system is being used, see section 10.3. But if you want to establish functional properties of a multimedia application, for example the effectiveness of using an agent in navigating a virtual environment, in a scientifically more rigorous way, you need to have:

experimental validation

- a theory – in our case: PEFiC
- a test scenario – for example, memory tasks in a digital dossier
- the technology – to realize applications

[23] e.mac.com/geertmul2

In this section, I will briefly describe our efforts in experimentally validating the use of ECAs in virtual environments. As technology, we use our *intelligent multimedia technolgy*, described in sectiona 8.3 and appendix E. So what must be explained is the theory we adopt and the test scenarios we use.

PEFiC is a theory developed by Johan Hoorn and Elly Konijn, to explain *Perceiving and Experiencing Fictional Characters*, see PEFIC. The PEFiC theory may serve as the basis for the experimental evaluation of user responses to embodied agents. In summary, PEFiC distinguishes between three phases, encoding, comparison and response, in analyzing the user's behaviour towards an agent. Encoding involves positioning the agent (or fictitious character) on the dimensions of ethics (good vs bad), aesthetics (beauty vs ugliness) and epistemics (realistic vs unrealistic). Comparison entails establishing personal relevance and valence towards the agent. Response, finally, determines the tendency to approach or avoid the character, in other words involvement versus distance.

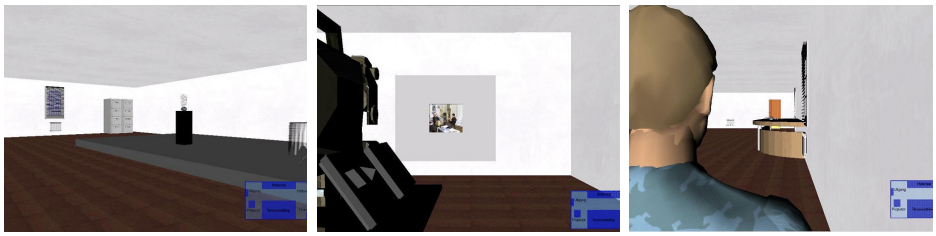
In general, having a virtual environment, there is, for developing test scenarios, a choice between:

validation scenario(s)

- navigation – pure interactivity
- guided tours – using some narrative structure
- agent-mediated – navigation and guided tours

For our application, a virtual environment of an artist's atelier, we have three experimental conditions, navigation without an agent, navigation with a realistic agent and navigation with a cartoon-like (unrealistic) agent. To ensure that these conditions can be compared, the acrtual information encountered when using the application is in all conditions the same.

The independent variable in our experiment, the degree of realism of the agent, corresponds with the epistemic and to some extent the aesthetic dimension of appraisal in the PEFiC theory. As dependent variables we have, among others, user satisfaction, believability, that is estimated usefulness of the agent, and also the extent to which the relevant information is retained.



11

The application is a *digital dossier* for the Dutch artist Marinus Boezem. The spatial metaphor we used for the dossier is the artist's atelier. We created a virtual environment containing a display of the artworks, in 3D, a file cabinet with textual information, a workbench for inspecting the artist's material, and a

video projector, with which the user can display a video-recorded interview with the artist.

The actual task to be performed by the user is to learn what constraints do apply to the installation of one of the artworks, *Stone and Feather*:

Stone and Feather

- feather: 70 cm, from ostrich, curved
- stone: 13.5 cm, white marble
- position: alignment with pedestal, no glue
- environment: 50 lux of light max.

The items mentioned in this list must be reproduced by the user in a subsequent memory test, and in another experiment the user must be able to choose the right materials and reconstruct the artwork.

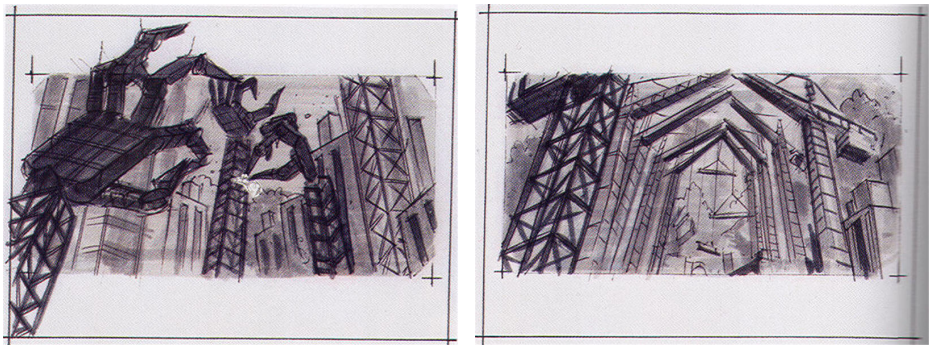
Our assumption in designing this test scenario was that the gestural nature of positioning the artwork will be favorable for the condition with a gesturing agent, whereas believability will be positively affected by the degree of realism of the agent.

12

9.3 multimedia augmented theatre

In june 2003, I was asked to advise on the use of VR in a theatre production of the Odyssee. Lacking experience in this field, I accepted the invitation to participate with some reluctance, since at the time I didn't have any clue what

the VR for the theatre production should look like. Nevertheless, I took the invitation as a challenge and started looking for appropriate hardware, bothering collegues for information on mixed reality art productions, and downloading code to explore software technologies. Many hurdles ware to be taken. We had to deal with organizational issues, such as finding the money for financing the actual production (which unfortunately never came through), finding the right people (students, in our case) to select material and contribute to the code; aesthetic issues, in particular to determine which approach to take to reach at an effective solution; and not in the least technical issues, to realize the production on a sufficiently efficient low-cost platform.



13

context – the Odyssee theatre production

The Odyssee[24]. theatre production was initiated by Ground Control[25], as a successor of previously succesful theatrical spectacles, including an open air performance of Faust[26]. In effect, two performances of the Odyssee were planned, an out-door (external) version, involving real ships at the shore of a lake, and an in-door (internal) version, to be played in temporarily empty office buildings. The in-door version is meant to give a more psychological rendering of the Odyssee Entanaclaz (2003), where the travels of Ulysses are experienced by the audience as a confrontation with themselves. Our contribution was asked for the in-door version, to enhance the experience of the audience with additional VR.

The Odyssee is a wellknown account of the travels of Ulysses leaving Troje, in 24 episodes ending in his return to Ithaca and his reunion with Penelope. The actual theatre production takes 12 parts which are played in 12 successive rooms through which the audience, subdivided in small groups, is guided one room after another for about five minutes per room. Our initial idea was to add information in the form of text and images, to direct the interpretation of the audience towards a particular perspective. In that beginning stage, somewhat optimistically, we

[24] www.odyssee2004.nl
[25] www.ground-control.org
[26] www.faust2002.nl

planned to offer multiple perspectives to each participant, in an individualized manner, dependent on the actual focus of attention of the individual participant.

initial ideas – VR and augmented reality: Our first problem was to find suitable hardware, that is see-through goggles. Searching the Internet gave us the name of a relatively nearby company, Cyber Mind NL[27], that specialized in entertainment VR solutions. Both price-wise and in terms of functionality semi-transparent see-through glasses appeared to be no option, so instead we chose for simple LCD-projection goggles with a (head-mounted) low-resolution camera. This solution also meant that we did not need expensive head orientation tracking equipment, since we could, in principle, determine focus using captured image analysis solutions such as provided by the AR Toolkit[28]. Moreover, captured video feed ensured the continuity and reactiveness needed for a true (first-person perspective) VR experience.

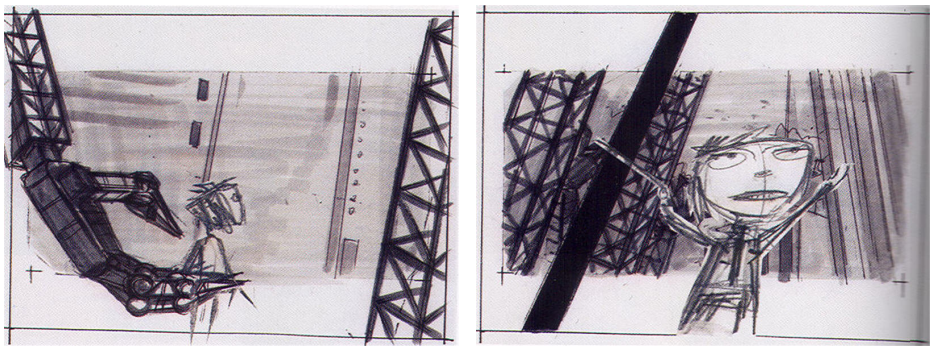
Augmented or mixed reality[29] is an interesting area of research with many potential applications. However, in the course of the project we dropped our ambition to develop personalized presentations using image analysis, since we felt that the technology for doing this in a mixed reality theatre setting is simply not ripe, and instead we concentrated on using the captured video feed as the driver for text and image presentation. In addition, we developed image manipulation techniques to transform the (projection of the) captured video, to obtain more implicit effects, as to avoid the explicit semantic overload resulting from the exclusive use of text and images.

technological constraints – the DirectX platform: After a few experiments with the AR Toolkit, it soon appeared that the frame rate would not be sufficient, on the type of machines our budget would allow for. Moreover, reading the AR Toolkit mailing list, marker tracking in a theatrical context seemed to be more or less unfeasible. So, we shifted focus to the DirectX SDK 9[30], both for video capture and projection in 3D. The DirectX toolkit is a surprisingly functional, and very rich technology for multimedia applications, supporting streamed video, including live capture, 3D object rendering and precise synchronisation between multimedia content-related events. At that time, and still at the time of writing, our own *intelligent multimedia technology*[31] was no option, since it does not allow for using live video capture and is also lacking in down-to-the-millisecond synchronisation.

After exploring texture mapping images copied from the incoming captured video stream, we decided to use the VMR-9 *video mixing renderer* introduced in DirectX 9, that allows for allocating 3D objects as its rendering surface, thus avoiding the overhead of explicit copies taken from a video processing stream running in a separate thread. See section 4.3. Although flexible and efficient,

[27] www.cybermind.nl
[28] www.hitl.washington.edu/research/shared_space/download
[29] www.se.rit.edu/~jrv/research/ar
[30] www.microsoft.com/windows/directx
[31] www.intelligent-multimedia.net

DirectX is a low-level toolkit, which means that we had to create our own facilities for processing a scenegraph, world and viewpoint transformations, and, even more importantly, structuring our mixed reality presentations in time.

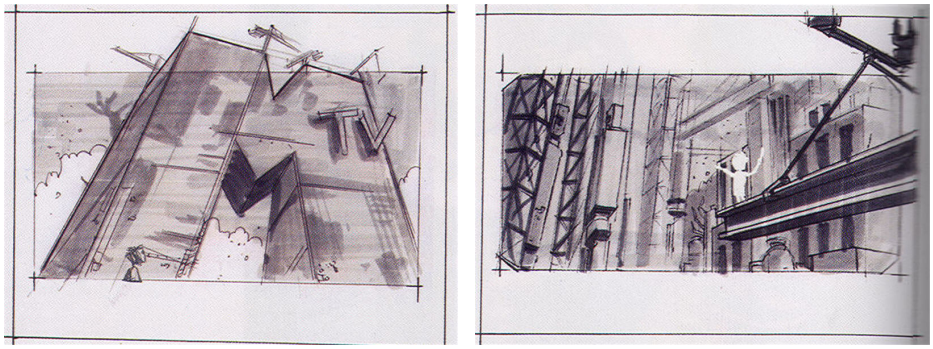


14

structuring time – maintaining 'see-through' aesthetics

One of the problems we encountered in discussing what we conveniently may call the VR with the producer of the Odyssee theatre performance was the high expectancy people have of VR, no doubt inspired by movies as the Matrix and the like. In mixed reality applications, manipulating persons, warps in space, and basically any intensive image analysis or image manipulation is simply not possible in real time. Moreover, there is a disturbing tendency with the layman to strive for semantic overload by overlaying the scene with multiple images and lines of text, thus obscuring the reality captured by the camera and literally blocking the participants view and awareness of the scene. Basically, as a guideline, we tend to strive for 70% visibility of the scene, 20% image or projection transformations and only 10% of information in the form of text and images.

The total duration of our presentation is only 2 minutes, or 118 seconds to be precise. We made a subdivision in 4 scenes, with transitions inbetween, hierarchically ordered in a tree-like structure. Initially, we abstracted from the actual duration, by taking only the fraction of the time passed (in relation to the total duration) as an indication for which scene to display. However, when the development reached its final stages, we introduced actual durations that allowed us to time the sequence of scenes to the tenth of a second. In addition, we used multiple layers of presentation, roughly subdivided in background captured image, the transformed captured image projected on 3D objects, and, finally, pictures and text. These layers are rendered on top of eachother, triggered in a time-based fashion, semi-independent of one another. The frame rate varies between 20 and 30, dependent on the number of images simultaneously used for texturing. Our final mixed reality theatre application may be considered a prototype, awaiting to be put to the test by the audience.



15

lessons learned – our explorations revisited: Altogether, the development of the mixed reality theatre application has been quite an experience, in multiple ways.

Not in the least it has been (and still is) a challenge to explain the possibilities of mixed reality applications to the layman, that do not take the abstractions we use in our daily academic life for granted.

To be frank, it also has opened my eyes to what some consider 'politically incorrect' technology, in other words Microsoft DirectX, originally developed as game technology, and no doubt a rich toolbox for real life multimedia applications.

Reinventing the wheel is not as simple as it seems. Nevertheless, developing scenegraph processing facilities and the appropriate timing mechanisms for controlling the mixed reality presentation was, apart from being a rekindling of basic skills, a learnful experience.

Finally, even before putting the application to the test, considering the aesthetics of mixed reality theatre productions, it may be called an eye-opener to realize how important reality is, and how meaningless explicit semantics (in the form of text and images) may become. Rather, our explorations were an incentive to further explore more implicit (graphical) modifications of the captured scene to convey meaning.

16

example(s) – *pizza boy*

Did you like to go to a theatre play, when you were sixteen? Peter van Kessel, one of the speakers in our *visual design* develeped a game *pizza boy* to overcome the resistance of secondary school students to theatre. The game, developed by Peter's agency Headland[32], introduced elements of the theatre play in a game setting. The player must deliver pizzas, riding a scooter, and finds him/herself in a situations where he/she must enact a role or perform some actions, related to the theatre play. Visiting the theatre, there is a MIDI-enabled scooter with which the students can play the game, in preparation for the actual play (image on the right in the figure above). Peter reported that the game did indeed help to overcome the scepticism these yougsters had with theatre.

The game was implemented by the Little Chicken Game Company[33], with Virtools. It was available for download at the website of the governmental department that initiated the development of the game, but became such a huge success, also with other people than the original target group, that it had to be taken offline.

research directions– *computational art*

I started studying computer music in 1979. For about four years, I worked on a system for real-time sound synthesis and algorithmic composition. Some years later, I wrote an essay that I presented in 1986 at the First Symposium for Electronic Arts, in Utrecht. The essay was the published in Leonardo, Eliens (1988). Now, almost 20 years later, I reread the abstract, which I OCRed from the original hardcopy:

Computational Art[34]

> The author conducts a simple thought experiment investigating the existence and scope of 'computational art': the utilization of the computer in the visual arts and music. In the experiment he sets the task of constructing an artifact that is capable of producing works of art. Since it appears that the artifact needs at least the capability of imagination, he queries the nature of images and imagery and argues that imagination is strongly intentional. Next he introduces the concept of notational systems, since they seem to govern the artistic activity of (not exclusively) machines. Confronted with the question of whether we are able to develop a computational analogue for taste, he finds that notational systems prove to be necessary for mediating the method of production of an artwork and the appraisal of its artistic value. Furthermore, the author shows that there are certain epistemological limits to the creativity of an imaginative device. Although the outcome of this hypothetical construction task clearly denies the possibility of an autonomously creative artifact, there seems to be no reason to worry about the opportunities for computational art: the computer appears to be a

[32] www.headland.nl
[33] www.littlechicken.nl/en
[34] www.cs.vu.nl/∼eliens/archive/documents/computational-art.pdf

> unique tool in exploring the possibilities of artistic production, guided by artists.

Today, despite the immense increase in computational power and the overwhelming amount of multimedia peripherals, the computer is still not able to produce art autonomously. But as a tool it is about to take over the entire industry, ranging from traditional publishing to film production and distribution. And, perhaps more tangible, as a tool for the creation of media content it is becoming ever better!

9.4 development(s) – dialectic(s) of awareness

There is an overload of information. Usually we don't pay much attention to this. Increasingly, we rely on search (engines) to find the information, when needed. But often our search is a bit disappointing, simply because there are too many results. If only there was more focus in search.

To assist you in proving focussed search, using google, that limits search to your site, look at the following fragement of code.

google(s)

```
<form action="http://www.google.com/search?" method="GET">
<input name="q" value="+site:www.cs.vu.nl/~eliens/media ">
<input type="text" name="q" size=40>
<input type="submit" value=" ">
</form>
```

Adapt the url, and put it in your web page, and, voila, your site is search-enabled.

On a deeper level, we may wonder why we are so impatient, in a hurry to search for information, and often not taking the time to properly digest it. In Eliens & Vyas (2007) we wrote: *In the course of our field study for the* PANORAMA *system, we tried to establish what relation users would have to the system, not only in the way they interact with it, but also in terms of what role the system plays in their lives, and when and how they would be aware of the systema.* Due to the intrinsic properties of the *PANORAMA* system, as a system meant to support social awareness in a work environment, we could not assume direct focussed attention. Instead, we must take the various forms of awareness or attention into account.

Our thoughts in this direction were triggered by a lecture of Linda Stone (former vice-president of Microsoft) at the Crossmedia Week[35] September 2006 in Amsterdam, entitled *Attention – the Real Aphrodisiac.* In that lecture Linda Stone made a distinction between applications popular before 1985, applications which were in general meant for self-improvement, for example language-learning, applications that were popular between 1985 and 2005, applications that she characterized as supporting *continuous partial awareness*, such as email and newsfeeds, and applications of the period thereafter, from now into the future, which

[35] www.picnic06.org

may be characterized as applications that allow the user to be creative, take part in a community, and are in other words more focussed and less dependent on the external environment.

Admittedly, it takes a few more steps to formulate a theory of the *dialectics of awareness*. However, with the function of the *PANORAMA* system in mind, we may make, following Benjamin (1936), some interesting distinctions between the experience of art and architecture. Where art is usually experienced in a delimited time span, and is similarly delimited in space, that is the position of the observer, architecture is everywhere and always there. As a consequence, art receives focussed attention and may be appreciated with reflective distance, whereas architecture is often not perceived consciously, but merely present and subject to an almost sub-conscious sensibility, which is only brought to the focus of attention when it is either aesthetisized, for example when taking photographs, or when something surprising is sensed, for example in the change of skyline in New York.

As argued in Hallnäss and Redström (2002), many of the new interactive systems, whether in the category of *ambient media*, *ubiquitous computing* or *calm technology*, will fall somewhere inbetween the spectrum spanned by art and architecture, or more likely even alternate between the forms of awareness associated with respectively art and architecture.

In designing the new interactive systems and games, we need to be explicitly concerned with the actual phases of awareness that occur, simply because it is not clear what role these systems play in our life. When introducing a new system or artefact, we may distinguish between the following phases:

- *initiation* – appeal to curiosity
- *promotion* – raising interest
- *progression* – prolonged involvement

As designers we must ask ourselves the following questions. *How do we appeal to the users' curiosity, so that our system is noticed? How do we get a more sustained interest? How de we get the user to interact with or contribute to the system? And, how do we obtain prolonged involvement, and avoid boredom?* These questions are not simple to answer, and require also an understanding of the actual context in which the system is deployed as well as an understanding of the level of (aesthetic) literacy of the user(s).

Aesthetic awareness is common to us all, Saw (1971). Having an understanding of aesthetic awareness, can we isolate the relevant design parameters and formulate rules of composition that may help us in developing interactive applications? According to our philosophical credo, Eliens (1979), no! However, the history of art clearly shows the impact of discoveries, such as the discovery of perspective, as well as conventions in the interpretation of art, as for example in the iconic representation of narrative context in 17th century Dutch painting. Moreover, the analysis of the visual culture of mass media may also give us better understanding of the implied meaning of compositional structures.

The notion of *perspective*, described in Alberti (1435), is an interesting notion in itself, since it describes both the organisation of the image as well as the optimal

point of view of the viewer. The normal perspective as we know it is the central perspective. However, there are variants of perspective that force the viewer in an abnormal point of view, as for example with anamorphisms.

Perspective had an enormous impact on (western) art and visual culture. It defines our notion of naturalist realism, and allowed for the development of the panorama as a mass medium of the 19th century, Grau (2003). Art that deviated from central perspective, such as cubism or art from other cultures, was often considered naive. Photography and its pre-cursors had a great impact on the perfection of perspectivist naturalism, and what is called *photorealism* became the touchstone of perfection for early computer graphics, Bolter and Grusin (2000).

Apart from perspective, other conventions regulate the composition of the 2D image, in particular, following Kress and van Leeuwen (1996), the *information value* related to where an object is placed in the image, and the *salience* of the object, determined by its relative size, being foreground or background, and visual contrast. Also *framing* is used to emphasize meaning, as for example in the close-up in a movie shot. In analysing a large collection of image material, Kress and van Leeuwen (1996), somewhat surprisingly found that *lef/right* positioning usually meant *given* versus *new*, *top/bottom* positioning *ideal* versus *real*, and *centre/margin* positioning *important* versus *marginal*. It is doubtful whether these meaning relationships hold in all cultures, but as a visual convention it is apparently well-rooted in western visual culture.

17

questions

digital content creation

1. (*) What are the considerations in developing digital content?

concepts

2. What guidelines can you give for the process of design?
3. What is a portfolio? And, what criteria should it meet?
4. What is to be understood by affordance? And, why is affordance important?

technology

5. How would you characterize the following items: logo, sign, collage, story?
6. Characterize the elements of game design.
7. What factors play a role in emotional involvement?
8. Explain how time constraints may be incorporated in the scene graph.

projects & further reading As a project, you may develop a dialog engine for non-linear interactive story telling or a *collage* generator, that produces artworks from a collection of images.

You may further explore the various presentation platforms, and assess the tradeoffs with respect to the support they offer for authoring.

For further reading, I suggest to study interaction design patterns[36]. It is also worthwhile to get some books on modern art, to gain some knowledge about art and design.

the artwork

1. street logos – images from Manco (2004).
2. photograph of oilpaint box.
3. Mark Veldhuijzen van Zanten – the six roles in their *agency*[37].
4. Mark Veldhuijzen van Zanten – to design for the *salon*, periodic lounge evenings in musea and art institutes.
5. Geert Mul – interactive multimedia installation.
6. Geert Mul – multimedia installation in dutch consulate in India.
7. website of Institute of Creative Technologies[38], showing scenes from *Mission rehearsal Exercise* (MRE).
8. street logos – more images from Manco (2004).
9. website for Visual Sensations[39], a yearly VJ contest in the Netherlands, developed by the agency of mark Veldhuijzen van Zanten.
10. Geert Mul – *Harbour Sound & Vision*, 1999
11. screenshots from virtual atelier of Marinus Boezem.
12. left: *don't spit*, a chines poster against spitting during the SARS period, taken from dutch newspaper; right: filmteckarna, Wiedermann (2004).
13. sketches – from filmteckarna, Wiedermann (2004).
14. sketches – from filmteckarna, Wiedermann (2004).
15. sketches – from filmteckarna, Wiedermann (2004).
16. game – *pizza boy*, developed by Headland[40], see 9.3.
17. signs – health and safety, van Rooijen (2003), p. 258, 259

The artwork for this chapter is meant to emphasize *context*. The *street logos* opening this chapter, as well as the work of Mark veldhuijzen van Zanten and Geert Mul, must be experienced in a context to fully appreciate their meaning.

Also for the MRE application, it is the context, in this case the stress and anxiety of a war situation, that determines the impact. The photorealistic graphic style of MRE, wellknown by the trainees from other games, is meant to strengthen the experience of *immersion*. Notice that the street logos assume an almost iconic character.

[36] www.visi.com/∼snowfall/InteractionPatterns.html
[37] 167aardigeontwerpers.nl
[38] www.ict.usc.edu
[39] www.visualsensations.nl
[40] www.headland.nl

10. application development

learn the craft, break through the magic of engineering

learning objectives

After reading this chapter you should be able to discuss the multimedia development process, to indicate the need for information system support in the cultural heritage domain, to characterize the notion of digital dossier, to provide solutions for navigating complex information spaces, and to discuss the data representation issues involved.

As you gather from reading this book, the field of multimedia is widely divergent. However, when you develop a multimedia application, you will find that all topics treated so far will become relevant. There will be a need to mix multiple media formats. You will have to find suitable codecs for your video. You will be asked whether search is possible. And, not the least important, you will have to balance navigation and presentation.

This chapter is based on the work we, that is my students, have been doing in the domain of cultural heritage. In the first section, we will introduce the notion of *digital dossier* and outline our general approach. We will then in section 2 look at some examples, and describe how we deploy concept graphs as a universal navigation tool for complex information spaces. Finally, in section 3, we will explore the options for presenting multimedia material and discuss the design issues as well as the technical issues that have arisen in the course of our work.

1

10.1 multimedia casus

You can learn a great deal about technology, but there is no meaning to that unless the technology is applied to produce something worthwhile. In this final

chapter, the outline of a *multimedia casus* will be presented, that is a course in which students face the challenge of creating a veritable (intelligent) multimedia information system.

In the studyguide, the course is described as follows.

multimedia casus

> *The assignment in the* multimedia casus *is to develop a virtual environment for some cultural or governemental institute or company. The practicum takes the form a stage, in which external supervision plays an important role.*
>
> *In the* multimedia casus*, techniques learned in previous courses will be applied to create the application. At the start of the course the actual assignment will be determined.*
>
> *Examples of possible assignments are: the development of a virtual exposition hall for the Dutch Royal Museum of the Arts, a virtual city square, which gives information about both the present and the past, a virtual shop, with online buying facilities, or an online broker, which offers facilities for inspecting houses.*

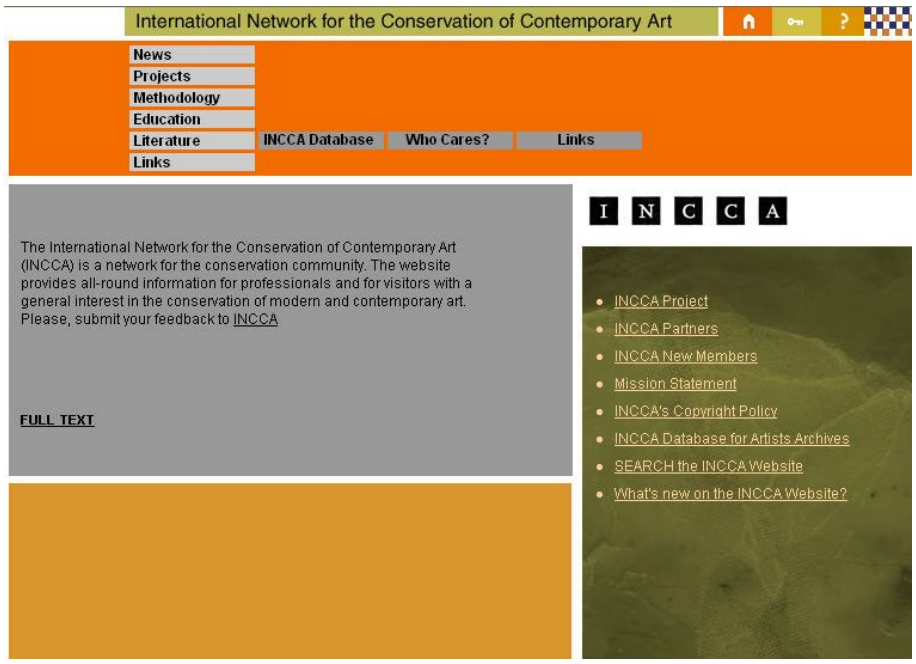
In effect, the availability of a representative of a cultural institute, industry, or governmental department is crucial, otherwise the assignment might easily degrade to the type of toy assignments so common in academia. Now, what is the challenge in such an assignment?

augmented information In the *research directions* of section 8.1 the notion of *augmented virtuality* was introduced to clarify the duality between *information* and *presentation*. More in particular, it was argued that the use of VR makes no sense unless there is some added value, that is by using the rich presentation and interaction facilities that come with this technology.

In an abstract fashion, we may rephrase the assignment as follows:

> Given an information space, create a VR that resolves the duality between information and presentation, using *intelligent multimedia* technology. The VR must offer access to all relevant information entities, organized in a suitable spatial layout, and must allow for presentations from a variety of perspectives, making full use of graphical and rich media facilities.

Below, we will see how this may work out for a concrete assignment.



front page of the INCCA website

2

project assignment – *present a complex information space*

Art is an interesting and complex phenomenon. No art, no culture! Hence, the preservation of collections of artworks is of crucial importance. The ICN (Netherlands Institute for Cultural Heritage) is a government-funded institute for the preservation of (dutch) cultural heritage. ICN gives advice, organises courses, does research, etcetera.

ICN is actively involved in the preservation of modern art, being project leader for INCCA (International Network for the Conservation of Contemporary Art), in the person of Tatja Scholte.

INCCA

> *In 1999, a group of eleven international modern art museums and related institutions applied to the European Commission (Raphael Programme) under the umbrella International Network for the Conservation of Contemporary Art (INCCA). The INCCA project was accepted and work started in January 2000 led by the organiser, the ICN (Netherlands Institute for Cultural Heritage) and the co-organiser, Tate, London.*

The objectives of INCCA are phrased as follows.

objectives

> *INCCA's most important set of objectives, which are closely interlinked, focuses on the building of a website with underlying databases that will facilitate the exchange of professional knowledge and information. Furthermore, INCCA partners are involved in a collective effort to gather information directly from artists.*

The INCCA web site contains a wealth of information about contemporary artists, as well as links to virtual collections of the works of a variety of artists, as for example Mondriaan. The way the virtual Mondriaan collection is presented is interesting in itself. It is a running display with iconic representations of his paintings. The speed of the display varies with the user's mouse movement, and at any time the user may select a painting to obtain more information about it. This particular site suggests where our *intelligent multimedia* approach may fit in.

Returning to the INCCA project once more, as its mission statement we read:

mission

> *INCCA's guiding mission is to collect, share and preserve knowledge needed for the conservation of modern and contemporary art.*

By now, the outlines of our assignment should become clear. Our information space is information about modern and contemporary artists, in the form of digital representations of there work, photographs, audio recordings from interviews and written text. The project assignment is to organize (part of) this material in a virtual environment and to include interaction facilities that highlight particular aspects of this information.

At this stage it would be too ambitious to cover all the material in the INCCA database, so we should restrict ourselves to one or more smaller case studies. The challenge, obviously, is to create presentations with a solid narrative structure and to augment the presented material in a suitable manner, using *intelligent multimedia* technology. What is *suitable*, is part of the challenge!

project management – *roles*

Can the challenge, stated above, be met? Well, there are many ways the project may loose its focus, or fail alltogether. Students should be aware of the fact that the challenge is real and that failure would bring about shame.

Since there are no golden rules for project management, the students themselves are responsible for keeping the project on track. In other words, project management is part of the experience. Here is a checklist.

checklist

- *roles* – create a team
- *project goal* – develop a vision
- *production* – construct the assets
- *quality assesment* – test and control
- *delivery* – present and archive

- *manage* – all along
- *document* – track project's history

The rule of the supervisor should be minimal, as a critical third party. The students work as a group, and they should take responsibility as a group, including the management of the project, assigning roles, and keeping track of progress. In such an approach *intervision* (students supervise one another) is a necessary mechanism in judging the final result of the project.

judgement

- *group* – (2) effort, 5 (product), 3 (documentation)
- *individual* – (4) responsibility, (3) productivity, (3) quality

On a scale of 0-10, both the group result and the individual efforts may be assigned a mark with proper weights, as indicated above. In addition, target deliverables should be defined to assure that the project meets its deadlines and to inspect the nature and quality of the students' work.

deliverables

- *group* – project plan, design, project report, product
- *individual* – detailed weekly account of activities

Dependent on the time available a schedule should be defined indicating when the deliverables should be ... delivered.

schedule

1. project organisation
2. project definition
3. planning and design
4. construction and development
5. integration and delivery
6. presentation and archiving

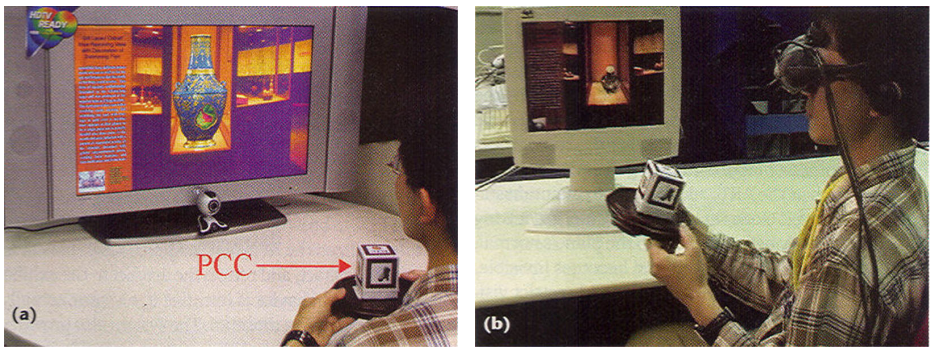
Is this a realistic setup? It should be. Besides, it is not the supervisor's responsibility, is it? It is first of all the responsibility of the students themselves!

peer reviews – *to control group dynamics* Whether you are a student or responsible for supervising projects, this account of how our *multimedia casus* is organized should give you some indications of what you may encounter in a team project. But apart from the organizational issues, you should be aware of the group dynamics, that is the individual relations and clusters of persons that emerge during the process of development. In general it takes some time before the various roles are established, that is who takes leadership, who takes notes during meetings, and who does most of the technical work. Also, sometimes some of the more creative members of the tem are overshadowed by some of the more outspoken ones. Some people simply need to learn to assert themselves!

For a supervisor, it is often quite difficult to assess the contribution of each individual. I remedy this by having both informal peer reviews, in the group, as well as formal peer reviews, where marks must be given for *responsibility*,

productivity and *quality of the work*. The order in which the students to be reviewed are presented to the reviewer is random, to avoid any bias due to presentation order.

I started using peer reviews about ten years ago, then informally, because I noticed that students could be much more direct in their criticisms that I dared to be. Recently, following suggestions from my collegue Johan Hoorn, who is an experimental psychologist, we formalized the procedure and introduced peer reviews for other courses as well, including the *visual design* course.



physical control cube head-mounted display

3

example(s) – *tangible virtual museum*

Don't touch that! Keep your hands off! This is what you often hear parents shouting at their children in a museum or gallery exhibit. More often, however, precious artifacts, ceramics, porcelain or bronzes, are stored away in glass showcases, precluding any kind of physical interaction, and many times a proper look as well.

To remedy this situation, researchers from the Academia Sinica and the National Cheng Kung University have developed a *tangible photorealistic virtual museum*, a system for real-time interaction with photorealistic museum artifacts, which allows for an immersive experience using tangible interfaces, in the form of a 3D control cube (image left), Rosenblum and Macedonia (2005). The display is a kiosk-like system showing a panoramic view of the exhibit, augmented with a collection of perspective photographs for each object. The user may examine any of the objects by using a handheld control cube (PCC) to control size and rotation of the object.

As indicated, the system is not 3D mesh-based but image-based, which allows for high resolutions on mid-range platforms, which would not be feasible according to the authors, when using 3D modelling techniques.

research directions– *metaphors and interaction style*

Given a problem statement as the one above, to present information about contemporary artists, how would you proceed? You might start by asking potential users, or stakeholders, how they would like the system to be. The answer you will get this way is likely to be disappointing. They will probably tell you that it must be like something they already know. So it might be better to rely on your own intuition and find a creative solution by choosing a fitting metaphor.

Let me give an example. In creating the digital dossier, a notion that will be explained in the next section, for the artist Marinus Boezem, as presented in the *research directions* of section 9.2, we choose the artist's atelier as a metaphor, and we used the spatial layout of the atelier as an organizational principle for presenting the information. In this, indeed very naturalistic, approach, we used pedestals to present the artworks, a file cabinet to present the textual information and a video projector to present the video recorded interview with the artist. The extent to which the virtual atelier does represent the artist's atelier faithfully is not important, in this context. What is important is whether the spatial metaphor did function as a valid organizational principle for presenting the information.

Instead of arguing whether this is the case or not, or whether the graphics chosen were right, etcetera, I would rather like to refer you to the literature, so that you can investigate the issues involved yourself.

In Preece et al. (1994), it is observed that interface metaphors act as conceptual models to support particular tasks. For office tasks, for example, we have the wellknown *desktop metaphor*. Preece et al. (1994) lists a number of such metaphors, for a variety of application domains:

application area	metaphor	familiar knowledge
operating environment	desktop	office tasks
spreadsheets	ledger sheet	columnar table
object-oriented environment	physical world	real world
hypertext	notecards	organization of text
learning environment	travel	tours, guides, movement
file storage	piles	categorizing
multimedia environments	rooms	spatial structures
cooperative work	multi-agents	travel agents, servants

In the most right column it is indicated why the metaphors should work, assuming real world situations that we are familiar with.

In some cases it is necessary to speak of a *composite metaphor*. For example, scrollbars are not easily to be found on your natural desktop. Form a cognitive perspective then, we may speak of multiple mental models.

When we look at what interaction styles are supported from a more technical perspective, we have following Preece et al. (1994), the following options:

interaction styles

- command entry
- menus and navigation
- forms fills and spreadsheets

- natural language dialog
- direct manipulation

However, each of these interaction styles may somehow be incorporated in the representation that we adopt for our metaphor.

2D vs 3D Surprisingly, each year that I start with another *multimedia casus* group, there is a discussion whether the application should be in 2D, using traditional web technology or *flash*, or 3D, using VRML or any other suitable 3D technology. My answer to the students objections, which can partly be explained by the fact that they fear the complexity of 3D, is flatly that anything that can be done in 2D can be done in 3D. But looking at the list of interaction styles above, I am tempted to add that a 3D representation allows for a more rich repertoire of interaction styles, such as spatial navigation. It would be interesting to investigate to what extent the interaction styles used in game playing can be incorporated in 'more serious' applications.

10.2 digital dossier(s)

After a first round of the *multimedia casus*, in which the students produced an application giving an overview of the INCCA information archive, the participants, but only incidental information about the artists and their artworks, we decided to focus on case studies of individual artists, and we introduced the notion of *digital dossier*:

digital dossier

> *Create a VR that realizes a* digital dossier *for a work of a particular artist. A* digital dossier *represents the information that is available for a particular work of art, or a collection of works, of a particular artist. The* digital dossier *should be* multimedia-enhanced*, that is include photographs, audio and other multimedia material in a compelling manner.*

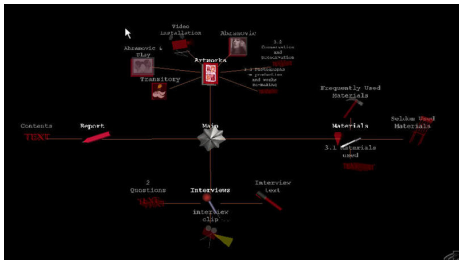
Like a medical dossier, the *digital dossier* was meant to give the information about the artist and the works of art readily at hand, so that it could effectively be used for the task of conservation and the re-installation of the artworks.

Since we were in doubt whether the phrase *dossier* actually existed in the English language, we looked it up in a dictionary:

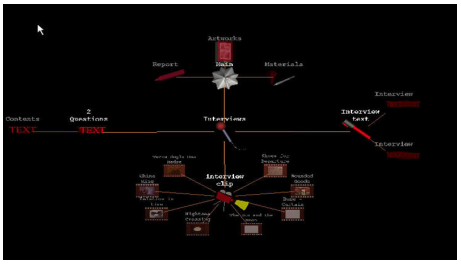
Webster New World Dictionary

- dossier (dos-si-er) [Fr < dos (back); so named because labeled on the back] a collection of documents concerning a particular person or matter
- archive – 1) a place where public records are kept ... 2) the records, material itself ...

We chose for the phrase *digital dossier*, and not for archive or library, to stress that our focus lies on presentational aspects. Although issues of data representation and content management are clearly important, our primary interest was with issues of presentation and navigation.



main node



interviews

4

the *abramovic dossier*

For the 2004 autumn group, we decided to take the work of Marina Abramovic, a serbian-dutch artist who became wellknown in the seventies with performances with her partner Ulay, and has since then produced numerous installations, videos and performances with what I would like to call 'high existential impact'. The directive with which the students where set to work was, quoting Ted Nelson:

> *everything must be highly intertwinkled*

Since virtual museums are by now a common phenomenon, and the virtual atelier for Marinus Boezem may be considered to be just a variant of this, the 2004 autumn group decided to explore alternative ways of presentation and navigation.

As material for the *abramovic dossier* there was an interview with Marina Abramovic from ICN, made in cooperation with the Dutch Foundation for the Visual Arts, and a great collection of videos from Montevideo. In addition, a transcription of the contents of the interview made by Michela Negrini, a student of media art at the University of Amsterdam, who also provided an interpretation as well as a categorization of the works of art. Given the material and the categories along which this material was classified, the students decided to explore the use of concept graphs as an instrument for navigating the information space.

navigation – *concept graphs*

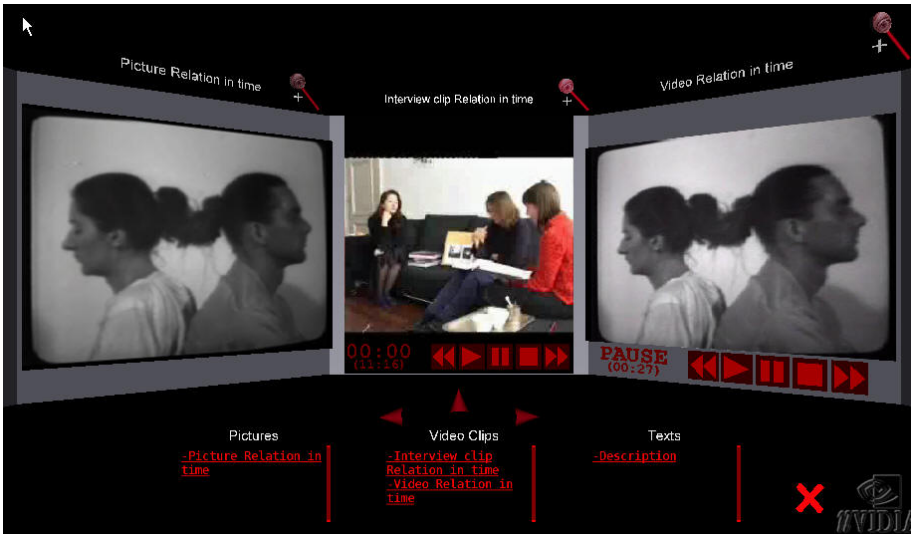
The reader has already encountered concept graphs in chapter 1, when the notions of multimedia, medium, television and communication were explained by indicating their relations to other concepts.

Concept-relation graphs are a familiar tool in linguistics and have also been used for a long time in Artificial Intelligence to describe the semantic relationships in complex domains. As a navigation instrument it is, to my knowledge only used in a kanji learning tool[41] and the Visual Thesaurus[42].[43]

[41] www.rikai.com/perl/KanjiMap.pl?

[42] ualthesaurus.com

[43] The Visual Thesaurus allows also for invoking Google image or document search from any of the elements of the concept graph.



presentation of video clips from Marina Abramovic

5

After the initial idea was there, one of the students of the group, Olaf van Zon, an AI student, managed to get a first version of a 3D concept graph working in VRML. This prototype implementation demonstrated the potential of the concept graph as a navigation instrument in the *abramovic dossier*.

presentation – *gadgets*

The original idea of presenting information, that is the actual interview, the videos and images of the works of art, as well as the textual information, was to use *rooms*, where the information could be projected on the walls. The *room* metaphor, which obviously stems from the virtual museum approach, did however not seem appropriate since it conflicted with the concept graph used for navigation. After some discussion, information rooms were abandoned in favor of *information gadgets*, that could be expanded from and collapsed into the concept graph.

In the original *abramovic dossier*, the presentation gadget consists of three panes that can simultaneously show a video of the work, the interview, that is the fragment in which Abramovic speaks about that particular work, and the textual information related to the work and the interview. However, it appeared that in some cases there was not enough information, because the work was not spoken about in the interview, and in other cases there was too much information, for example multiple recordings or text documents. It was then decided to extend the presentation gadget with lists of alternative material that the user could select from and direct to one of the panes for further inspection.

To enable the user to focus on one of the panes, for example to get a better view of the video material a zoom in/out button was provided. All these enhancements,

however, did complicate the interaction, as became clear when the *abramovic dossier* was presented at Montevideo.

In the course of the project, another interesting presentation feature was added, namely the reconstruction of one of the video installations in 3D, incidentally demonstrating the advantages of using 3D.

reconstruction – *recreating the installation*

In discussing the *abramovic dossier* with Bart Rutten from Montevideo, who provided us with all the video material, another project was mentioned which was concerned with 3D-recordings/models of existing installations. Having full confidence in the technical capabilities of my students, I promised to show that such a reconstruction of an installation would naturally fit within our approach.

Reconstruction of Terra della Dea Madre *in VRML.*

6

The installation for which the reconstruction was made is *Terra dea degli madre*, and installation with two chairs and a television, which was exhibited in the Stedelijk Museum of Amsterdam, in 1986. As a starting point, we took a video produced at the time of the exhibition, which shows the installation in an exposition room in the Stedelijk Museum, and which contains, apart from comments from Abramovic, also the video shown on the televison in the installation.

At this point, we can only speculate how useful such a reconstruction can be as a tool for the conservator responsible for the re-installation, to play around with the presentation parameters, the positioning in space, the overall size, light and ambient effects.

style issues – *how to improve the dossier*

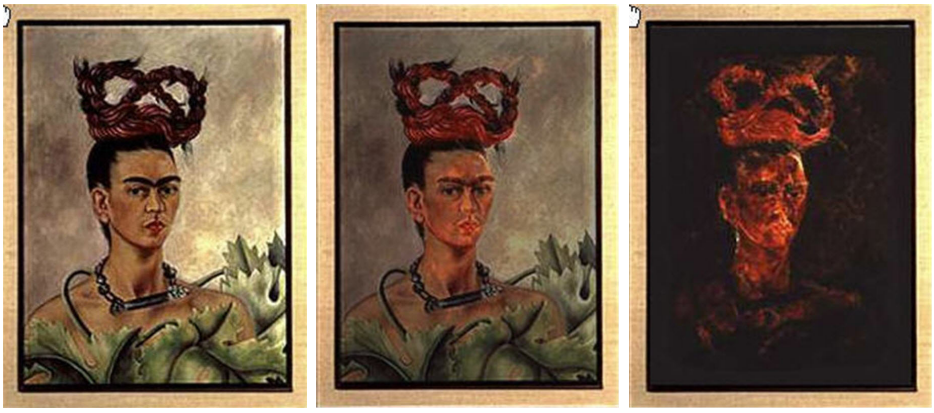
The *abramovic dossier* does also provide a facility for search, as well as online help. However, as already mentioned, when demonstrating the application to the interested parties, that is ICN and Montevideo, a number of issues came along, that I will here summarize as a list of questions:

style issues

- what icons should be used to identify the elements of the concept graph?
- what categories and relationships are most appropriate?
- how should the information be displayed, simultaneously or more focussed?

- how do we allow the user to choose between multiple information items?
- how do we avoid visually disturbing elements?

Obviously, although the *abramovic dossier* was very positively received, these issues must be dealt with to make it a success. Having a first prototype, we need to rethink our application, not only with regard to its style of presentation, but as we will discuss in section 10.3, also in terms of its underlying data representation.



no light half light full light

7

example(s) – *conservator studio*

Ever thought of becoming a conservator? Seattle Artmuseum's Conservator Studio[44] gives you the opportunity to explore this career options:

> Explore four paintings from the Mexican Modernism exhibition through the eyes of a conservator (what's a conservator? you'll find that out too!). You'll have a new perspective on the paintings as well as how they are handled and prepared for display.

The illustrations above show what occurs when manipulating *transmitted light* on the painting *Self-Protrait with Braid*, oil on canvas, from the Mexican painter Frida Kahlo. As explained in the accompanying text: *when a light is shone through this painting one can see that the hair and the flesh areas are painted with thin layers of paint.*

These series of images are part of an interactive *flash* application developed by the Seattle Artmuseum to engage the general audience in the conservation of art, and to arouse an interest in art in general. The application allows the user to experiment with the various techniques used for the analysis and conservation of oil paintings.

[44] www.seattleartmuseum.org/exhibit/interactives/mexicanModernism/enter.asp

research directions– *establishing usability*

In the March 2005 volume of CACM, an assessment is given of the current state of *user-centered design* practice. User-centered design is, quoting UCD, *a multi-disciplinary design approach based on an active involvement of users to improve the understanding of user and task-requirements, iterative design and evaluation.* In the article, which is based on a survey among user-centered design practitioners, user-centered design is claimed to have been beneficial for, among others, customer satisfaction and enhanced ease of use. Other measures mentioned are mostly relevant for e-commerce applications, which, as the authors observe, *have greatly bolstered the the appeal of usability and user-centered design, as users can take their business elsewhere with just one mouse click.*

In our case, the competition is fortunately less threatening. Nevertheless, usability issues such as legibility of text, ease in navigation and adequate task support are equally relevant. As a first step after completing the *abramovic dossier*, we have developed a test-plan and a sample task, and (the students) executed two test-sessions with participants from ICN and Montevideo, who where asked to work with the system thinking aloud. The test-sessions were recorded on video, and the participants were requested to complete a questionnaire.

In UCD, a list of approaches is given, which were reported to have been used by the respondents of the survey:

user-centered design methods

> field studies, user requirement analysis, iterative design, usability evaluation, task analysis, focus groups, formal/heuristic analysis, user interviews, prototype (without user testing), surveys, informal expert review, card sorting, participatory design

The three most frequently used methods in this list are, respectively, iterative design, usability evaluation and task analysis. These three methods were also considered to be important by the respondents. Frequently used, but not considered to be as important, were informal expert reviews. And less frequently used, but considered important, were field studies. This distinction can, according to UCD, attributed to cost-benefit trade-offs, since clearly field studies are much more costly.

Usability evaluation looks, according to Preece et al. (1994) to issues such as:

usability evaluation

- *learnability* – time and effort to reach level of performance
- *throughput* – the amount of work done
- *flexibility* – accomodating changes in the task
- *attitude* – of users to the system

To conclude this section, let's take a closer look at task analysis.

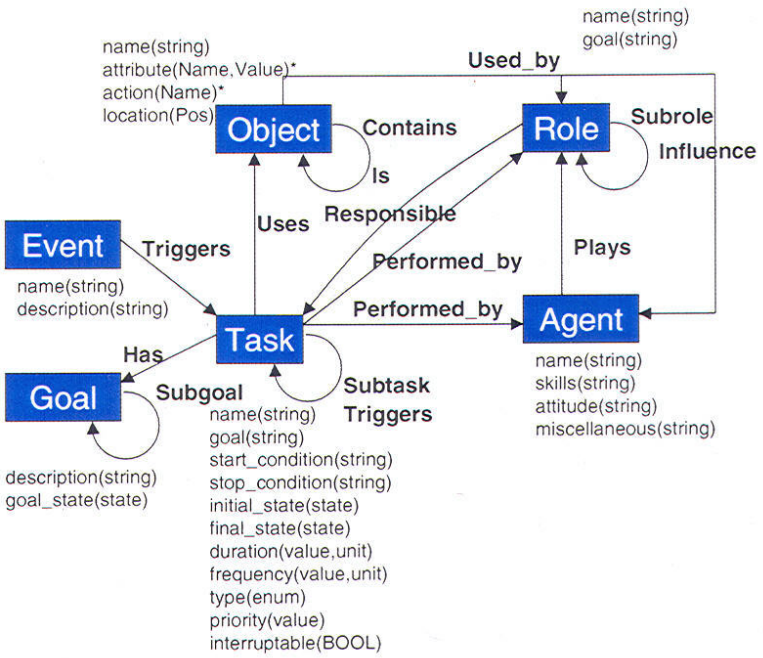
task analysis Task analysis may be characterized as the decomposition of a task into subtasks or steps, to arrive at a sufficiently detailed description of the task and its relation to the environment.

In Welie et al. (1998), a description is given of what might be understood as the task world ontology, the concepts and relations that play a role in performing a task analysis. The main concepts figuring in the task world ontology are, following Welie et al. (1998):

task world ontology

- *task* – activity performed by an agent to reach a certain goal
- *goal* – a desired state in the task world or system
- *role* – a meaningful collection of tasks
- *object* – refers to a physical or non-physical entity
- *agent* – an entity that is considered active
- *event* – a change in the state of the task world

As indicated in the diagram above, these concepts are related in various ways. Example relations include *uses*, *triggers*, *plays*, *performed_by*, *has*, etcetera.

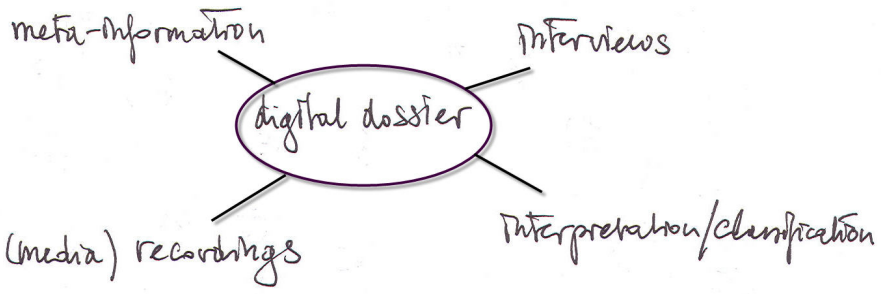


8

Creating a task model based on this, or a similar, ontology may help us understand what a user needs to accomplish and how this may be supported by an information system. As such, creating a task model should be considered to be an essential ingredient of the software engineering life cycle, Eliens (2000).

10.3 representation & interaction

In re-thinking the *abramovic dossier*, we first needed to re-establish what are our goals in developing this application and what are our primary data sources. The goal, first of all, is to support conservators in their task of preserving contemporary art, and to assist them with the re-installation of such artworks.



Our primary data sources are *meta-information*, coming from the INCCA initiative, and video-recorded artist *interviews*, which were initiated by ICN as a means to record information about contemporary art that would otherwise be lost. In addition we media-material, including images and video, that may be regarded as *recordings* of the works of art, as well as the textual *interpretations* and classifications that exist, or may be constructed from this material.

At this point, I may remark that one of the pitfalls in creating a dossier is to get trapped in the visually salient features of the dossier, the presentation of the artworks themselves, and forget about the primary focus of the dossier, to make all information accessible in an adequate manner.[45]

For our next generation of digital dossiers, we decided to take the following steps:

next generation dossier(s)

1. adaptation of representation to Dublin Core (+ annotation needed for presentation)
2. XML-based content management, with php forms (extending what we have now)
3. there should also be a possibility to present the information and material in a 'plain' web format
4. as well as in (a new version of) 3D dossiers
5. we should think about the proper presentation parameters.

Dublin Core is the standard used in the INCCA initiative, to record meta-information about existing information sources. See section 3.3 for a description of the Dublin

[45] For many cultural heritage applications, which aim to present art to the layman, presenting the artwork is the primary focus, and giving access to the information context generally comes second.

Core element set and the Resource Description Framework (RDF) on which it is based.

For the *abramovic dossier*, a collection of record-like structures was developed, together with a simple content-management tool, written in PHP. This content-management system must be adapted to be compatible with the Dublin Core-based resource descriptions.

Further, we decided that, along with the 3D presentation of the dossier, it would be worthwhile to develop a conversion tool that produces standard web-technology based presentations as well. This approach allows us to assess the usability merits of the 3D dossiers in a comparative way.

Finally, as I indicated before, an important issue that must be resolved concerns the proper presentation parameters. What do we present to the user? And how do we allow the user to interact with the material presented?

content management and data representation

For developing the *abramovic dossier*, we have a fixed number of record-like structures:

structures

- Video – to display video fragment, including interviews
- Picture – to present pictures of the artwork
- Artwork – contains all information connected to a work of art
- TextItem – to present text, from the interview or any other source
- MaterialItem – to present information about material(s) used
- GroupNode – to combine nodes in the concpet graph
- Information – acts as the outer container for all nodes

All these structures support a set of common attributes, including *shortName*, *longName*, *ID*, *connectedNodesIDs*, and *description*. In addition the *Video*, *Picture* and *Picture* have fields allowing to show a preview image. And the *Video*, *Picture* and *TextItem*, also have a *url* field giving access to the actual information item.

The *Information* and *GroupNode* structures are used for creating the top-levels of the concept graph, whereas the other structures, such as the *Video* and *TextItem* give access to for example a fragment of an interview and its transcription.

Below an example is given of the data underlying the concept graph of the *abramovic dossier*:

concept graph

```
Information {
 informationNodes [
  GroupNode {
       ID "MAIN"
       shortName "Main"
       longName "Main"
       urlModel "models/conceptGraph/main/modelMain.wrl"
```

```
        description [ "Central information node" ]
        connectedNodesIDs [ "ARTWORKS", "KEYWORDS",
                            "INTERVIEWS", "REPORT" ]
        }
    GroupNode {
        shortName "Artworks"
        longName "Artworks"
        description [ "Node that connects to all the artworks" ]
        ID "ARTWORKS"
        connectedNodesIDs [ "MAIN", "TRANSITORY",
                    "ULAY", "VIDEOINSTALLATION", "ABRAMOVIC" ]
        urlModel "models/conceptGraph/artworks/artworksGroup.wrl"
        }
        # # ...
  ]
  }
```

The *Information* node collects all available nodes, and takes care of connecting the individual nodes, based on the information specified for each node.

As an example of an *Artwork* node, that is an element of the list of nodes in the *Information* node, look at:

```
Artwork {
  shortName "Terra degla Dea Madre"
  longName "Terra degla Dea Madre"
  description ["15:40 min, colour, sound."]
  ID "AV24"
  connectedNodesIDs ["VIDEOINSTALLATION", "DTV24",
            "TTV24", "PV24", "CV24", "VV24", "G0"]
  urlPreviewImage "images/previewImages/AV24.jpg"
  widthPreviewImage 479
  heightPreviewImage 349
}
```

This node is connected to many other nodes, giving access to the information items that belong to it, such as the video clips of the interview, shown below.

```
Video {
  ID "CV24"
  shortName "Interview clip Terra degla Dea Madre"
  longName "Interview clip showing Terra degla Dea Madre"
  url "interviewclips/interview_terra_degla.avi"
  width 320
  height 360
  urlPreviewImage "images/previewImages/interview_terra_degla.jpg"
  widthPreviewImage 320
  heightPreviewImage 240
  description [""]
```

```
  connectedNodesIDs ["CLIP", "AV24"]
}
```

In the *url* field of this declaration, the actual video file is indicated, which should be displayed at a resolution of 320x360, as specified in the *width* and *height* fields.

And finally, as an example of a *TextItem*, consider:

```
TextItem {
  shortName "Instruction"
  longName "Green Dragon Lying instructions for the public."
  description ["Text explaining the way the public has to interact with the
    artwork."]
  ID "ITO05"
  connectedNodesIDs ["AO05", "INTERACTION"]
  url "text/AO05_instruction.txt"
}
```

For constructing the *abramovic dossier*, Tim Verweij developed the content management tool, that allows the user to browse and edit existing nodes, and to insert new nodes into the graph.

integration with the Dublin Core

The Dublin Core is a general resource description formalism, that allows for specifying resources in a variety of domains. See section 3.3. For INCCA the Dublin Core was chosen, not because it is the most suitable formalism, but because it may serve as the least common denominator, and agreement on anything else simply seemed to be impossible. As a reminder, the Dublin Core provides the following elements:

Dublin Core[46]

- *title* – name given to the resource
- *creator* – entity primarily responsible for making the content of the resource
- *subject* – topic of the content of the resource
- *description* – an account of the content of the resource
- *publisher* – entity responsible for making the resource available
- *contributor* – entity responsible for making contributions to the content of the resource
- *date* – date of an event in the lifecycle of the resource
- *type* – nature or genre of the content of the resource
- *format* – physical or digital manifestation of the resource
- *identifier* – unambiguous reference to the resource within a given context
- *source* – reference to a resource from which the present resource is derived

[46] dublincore.org/documents/dces

- *language* – language of the intellectual content of the resource
- *relation* – reference to a related resource
- *coverage* – extent or scope of the content of the resource
- *rights* – information about rights held in and over the resource

Descriptions of items in the *digital dossier* should incorporate these elements, together with the attributes needed for the insertion of items in the concept graph and the presentation parameters, that are necessary for displaying the (media) material. Technically, the namespaces supported by RDF does allow for merging these different types of annotations. However, the challenge here is to derive the presentation attributes automatically, and to come up with a reasonable default for inserting these items in the concept graph.

location of *Tower of Babel* project

9

intelligent guidance – I-GUARD

Although digital archives or digital libraries[47] are by no means a new phenomenon, our concept of *digital dossiers* contains a number of innovative elements. A digital dossier provides a unified information and presentation space. In this sense it differs significantly from a digital archive with a traditional web interface, where navigation and presentation are distinct. Digital dossiers allow to a much greater extent for an immersive experience of the information related to works of art. As such it is reminiscient to explorations in *virtual archeology*[48], our to our notion of *virtual context*, presented in section 8.1.

Working out the issues indicated above, that is the integration with the Dublin Core and providing suitable content management, is a matter of diligent software engineering. But what can we further do to support the construction of digital dossiers and improve the usability of such dossiers? And what are the scientific issues, worth to be investigated?

To indicate the research issues, let me first expand the cope of our project and re-define the goal of our research:

I-GUARD

[47] www.ifla.org/II/etext.htm
[48] library.thinkquest.org/18261/?tqskip1=1

> Contemporary art is an intrinsic part of our cultural heritage. Installations, performances, video and other forms of media art, as for example *web art*, have the interest of a small group of adherents, but are in comparison with more traditional art forms, far more difficult to present to a general audience. Another problem presents itself, due to the type of materials used and the context-specific aspects of these art forms, in the conservation of the works.
>
> In our research we address the issue of providing access to these contemporary art forms from a wide variety of perspectives, ranging from the interested layman to the expert that has to deal with archiving, conserving and the possible re-installation of the art works.

The acronym I-GUARD stands for *Intelligent Guidance in Artist's Digital Dossiers*, and refers to a project the aim of which is to arrive at a general framework for artist's digital dossiers, that provide intelligent guidance to both the expert user, responsible for the future re-installation of the work(s), and the interested layman, that wishes to get acquainted with a particular work or collection of works. In general, there are two techniques that we can apply to provide such guidance:

intelligent guidance

- filtering the information space according to the user's perspective, and
- intelligent agents, that (pro) actively aid the user in searching the information space.

Filtering the information space may be done by using techniques from formal linguistics to restrict the concept graph that defines the navigation structure, that is by stating assumptions with respect to the relevance of particular (linguistic) categories or elements from a user's perspective. Intelligent agents is an approach stemming from artificial intelligence which allows for providing guidance in a variety of ways, possibly even in an embodied form using a face or humanoid figure to give suggestions to the user on what interactions to perform. With the latter type of guidance we have already experimented in the Marinus Boezem dossier, as described in section 9.2. So let's look at what natural language technology has to offer.

natural language: Having a concept graph as a generic navigation device, it still remains a problem how to fill the concept graph with meaningful content, and how to indicate meaningful relations between the concepts and aspects covered by the nodes of the concept graph. In the *abramovic dossier* this was done by hand, based on information derived from a transcripion of an interview with the artist. (provided to us by ICN). Interviews with artists is one of the means ICN deploys to gain knowledge needed for the conservation of contemporary artworks. Such interviews provide a rich source of textual information, that includes both general viewpoints on the artist's oeuvre as well as specific constraints that adhere to the (re) installation of the work(s) of art.

What we should strive for is to derive both structure and content of the concept graph for a particular dossier (semi) automatically. Using a basic lexicon of terms and phrases related to contemporary art we should be able to generate a

representation of the textual information that may serve as a basis for constructing the concept graph. This representation must contain an enumeration of the concepts, the relation between occurrences of concepts, as well as a reference to the work(s) of art to which the concepts apply.

Natural language processing technology may not only serve for the static analysis of the material, when the digital dossier is created, but also dynamically when the dossier is being used, to aid the user in finding relevant information. Research issues here are, on the one hand, the interpretation of user input (that is, loosely structured natural language), and on the other hand, filtering the concept graph representing the information space in such a way that it adequately reflects the user's interest or perspective.

In summary, from a research perspective, digital dossier(s) concern the following issues:

digital dossier(s)

- representation of information of one or multiple works of art,
- presentation of that information in a *rich media presentation environment*,
- intelligent navigation and interaction, and
- support for interaction with loosely-structured natural language.

And to conclude, *digital dossiers* will on the one hand contribute to making contemporary art forms accessible to a larger audience and on the other hand are explicitly meant to support the complex task of the conservation and re-installation of works of art in an effective manner.

outside view of *Tower of Babel* project

10

example(s) – *Tower of Babel*

In the *Tower of Babel*[49] project, shown above, multimedia material was projected from within buildings, on the windows, to the outside. Local citizens in a neighbourhood in Amsterdam where approached to submit material that expressed

[49] www.torenvanbabel.info

their emotions of daily life, with the question *what moves you*. The text and photograms could be submitted either by email or SMS. Also workshops were held, during which particpants could develop material. This material was then edited and prepared for projection, using 40 carroussel dia-projectors, taking about 2000 images, and six beamers projecting images and video. Also sound material, that was collected in the same manner, was being used during the projection.

Inside view of Tower of Babel *project.*

11

The centre of the location[50], a somewhat impoverished neighbourhood near the centre of Amsterdam, is a building dating from 1926, originally an antroposofic temple, that once served as a cinema for avant-garde movies, and is now being used as a library. The buildings surrounding it are, if not split up into apartments, being used as a local youth centre, a city archive and another library.

research directions– *media art*

In a recent symposium on the preservation of contemporary media art, a number of institutions presented their projects, ranging from more technical topics, such as the *conservation of videotapes*[51] and the *mass storage of digital material*[52] to the conceptual issues in *capturing new media*[53], the variety of *media formats*[54] and the need to record and maintain *meta data*[55] about the artworks and related information.

To get an idea what the phrase *media art* encompasses, have a look at the circumscription given in the WikiPedia[56]: *new media art* is a generic term used to describe art related to, or created with, technology invented or made widely available since the mid-20th Century, including technology stemming from telecommunications, mass media and digital modes of delivery the artworks Below, the disciplines that belong to this form of art are listed, together with their entries in the WikiPedia, in an abbreviated form:

(new) media art

[50] www.alphons.net/panos/tolstraat.html
[51] www.montevideo.nl/en/pdf/CONSERVERING_1tm80.pdf
[52] www.ichim.org/ichim03/PDF/128C.pdf
[53] www.v2.nl/Projects/capturing/summary.html
[54] www.variablemedia.net
[55] www.incca.org
[56] en.wikipedia.org/wiki/New_Media_art

- *audio art* – no definition available
- *computer art* – any art in which computers played a role in production or display of the artwork.
- *digital art* – art created on a computer in digital (that is, binary) form.
- *electronic art* – entry to game producer, should be Leonardo[57].
- *generative art* – art or design generated, composed, or constructed through computer software algorithms, or similar mathematical or mechanical autonomous processes
- *hacktivism* – the writing of code, or otherwise manipulating bits, to promote political ideology
- *interactive art* – a piece of art that involves the spectator in some way.
- *internet art* – art or, more precisely, cultural production which uses the Internet as its primary medium and, more importantly, its subject.
- *performance art* – art where the actions of an individual or a group at a particular place and in a particular time, constitute the work.
- *robotic art* – page does not exist
- *software art* – is an intersection of two almost non-overlapping realms: software and art.
- *video art* – is a subset of artistic works which relies on "moving pictures" and is comprised of video and/or audio data.
- *video game art* – involves the use of a computer game for the creation of a digital artwork.

By the nature of the WikiPedia, to which every user can contribute entries, this list nor the defining entries are by any means authorative. Nevertheless, it does provide an overview and may serve as a starting point for further research.

10.4 development(s) – hybrid multimedia

In chapter 1, we introduced the notion of *digital convergence* to explain the occurrence of the great variety of elements of multimedia applications, from a technical perspective. From an aesthetic perspective, this great variety of elements may easily lead to chaos, unrelatedness or divergence, where meaning gets lost in a multitude of perspective(s). To cut a long deliberation short, for simplicity, let's assume that *meaning* lies in the context, the story or *narrative structure*.

For 2D images, Kress and van Leeuwen (1996) identify narrative elements, that is relations between objects in the image that suggest a story, such as a diagonal line from a person to a door, or a relation of an object to the viewer, such as a gaze towards the viewer, a technique that has been used only since late renaissance painting.

More than paintings or 2D images, film is the medium for conveying narrative structures. The art of storytelling in film has been perfected in such a way that

[57] mitpress2.mit.edu/e-journals/Leonardo

Hollywood films may seem more real than life. However, as emphasized in Bolter and Grusin (2000), this is not due to any inherent form of naturalism, but to the fact that we have got accustomed to the conventions applied, that is the techniques of cutting, montage, camera movements, close-ups, etcetera. In a highly recommended book, Arnheim (1957), Rudolf Arnheim gives an extensive analysis of the principles of montage and film technique, and he explains why film is such an effective medium:

frame(s) of reference

> It is one of the most important formal qualities of film that every object that is reproduced appears simultaneously in two entirely different frames of reference, namely the two-dimensional and the three-dimensional, and that as one identical object it fulfills two different functions in the two contexts.

Due to the subtle play between these two *frames of reference* film may be considered an art form, and as such perhaps the dominant art form of the 20th century. As a mass medium, film may be characterized by what Arnheim, following Benjamin, called the *aesthetics of shock*, replacing reflective distance with immersive thrill. As an art form, however, it is the dominant paradigm for aesthetic awareness, lacking however still one dimension, *interactive dynamics*.

As observed in Bolter and Grusin (2000), interaction is what distinguishes video games from film. Current day technology allows for high-resolution photo-realist graphics, that make video games or virtual applications almost indistinguishable from film. Virtual reality technology as applied in video games adds arbitrary choice of perspective, as exemplified in first-person shooters or fly-overs, as well as an arbitrary mix of the imaginary and real, as in CG movies, in an interactive fashion.

Now, should we take the aesthetics of interactive video games as the standard for interactive applications? Not necessarily, since the naturalism strived for in most games may at best be characterized as naive realism, mostly photorealism. As observed in Kress and van Leeuwen (1996), realism is a social construct, and hence the program for developing an aesthetics for interactive applications should perhaps include the development of appropriate *realisms*. Again with an eye to the history of art, where we have for example *impressionism*, *cubism*, *expressionism*, as a guideline in the design of interactive systems, it might be even better to look for appropriate interaction-*isms*, styles of developing interactive systems and games from a particular perspective. Not excluding provocative perspectives! Cf. Burger (1981).

Where an arbitrary interactive system may differ from a game played for entertainment is obviously the actual outcome, the value attributed to that in the real world, and probably the effort required and the possible consequences. You would not like to run the risk to die a virtual death when answering your email, would you? However, when interactive systems replace task-bound functionality with fun, the difference becomes less clear.

As we indicate in Eliens & Chang (2007), one element not sufficiently captured by a classic game model, as introduced in Juul (2005), is the narrative aspect of the game play. To quote Juul (2005):

> *Game fiction is ambiguous, optional and imagined by the player in uncontrollable and unpredictable ways, but the emphasis on fictional worlds may be the strongest innovation of the video game.*

We may observe that many games already have a strong relation to reality in what narrative context they supply, or else in the realities of the media industry, in particular Hollywood. For *serious* interactive systems, we may assume an even stronger and in some sense more straightforward relation with reality, by the use of media content that is relevant for the life of the individual.

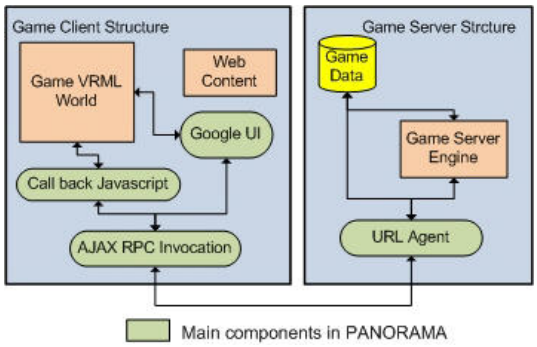
All these aspects of playing games are clearly relevant for the new interactive systems, which appeal more to *play* than *task-oriented* behavior. For example *rules* may be used to describe the visual characteristics of a system (e.g. the display of images as a flow in a particle system), *outcome* may be regarded as the benefits of the system (e.g. social awareness), *value* may include the risks of the system (e.g. a transgression of privacy), *efforts* is important when asking for contributions from the user (e.g. as image material to be displayed in the system), *attachment* may result when the system is installed (e.g. when people look forward to find new information), and finally *consequences* must be considered when a system is installed and used (e.g. interaction between people may actually change when they get to know eachother, for better or worse).

Given the large variety of games, including first person shooters, role-playing games, strategy games and decision-making simulation games, we can distinguish between a range of degrees of interaction, direct interaction, on the one hand, as for example in first person shooter and indirect interaction, on the other hand, as for example in simulation games, or role-playing games where the individual actions may contribute to a plot such that the effects will become visible at a later time. Where in game playing the variety of interaction modes seems to be well understood within each community of game players, for the development of more general interactive systems we will have to think seriously whether the target user will be able to learn the various modes of interaction, either by explicit instruction or during play. And as designers we must be concerned with the *rules of interaction* as well as issues of visualisation and interaction mappings, that is in other words which affordances the application offers for a particular group of users.

dynamic contribution(s) Another potential source of confusion lies in where the material comes from. Not in the sense of network transport or local storage of the platform of delivery, as discussed in the context of *convergence of delivery* in section 1.2, but in terms of *authorship*, which in our *participatory culture*, where users contribute content may result in a great variety of forms and formats. To develop multimedia applications and games that accomodate contributive authorship by a community of users is the great challenge for the next era (period).

In Eliens et al. (2007c) we wrote: *We explored the use of AJAX and web services in an X3D/VRML implementation of PANORAMA, a system meant to support social awareness in a work environment.* As explained in section 5.4, PANORAMA represents *casual encounters* in the work environment and displays

self reflections, that is postcards and other item contributed by employees, at a large display in a central public space.



12

The figure above, taken from Si & Eliens (2007), illustrates the architecture of an AJAX-based web implementation of PANORAMA, which includes facilities for game playing as *occasional battle(s)*, using a PHP server and the google GWT toolkit to allow users to contribute their image material, video's and whatever else.

13

questions

application development

1. (*) What information system support is needed in the domain of cultural heritage? Explain what considerations playa role in developing multimedia applications in this domain.

concepts

2. What phases do occur in the multimedia application development process?
3. Characterize the notion of digital dossier.
4. Explain the distinction between navigation and presentation, and discuss possible solutions for combining them.

technology

5. What elements do you include in your checklist when you set up a project?
6. What is a concept graph? And, how may it be used for navigation?
7. What structures do you need to represent the information in a cultural heritage application? Describe what descriptive features these structures must have.

8. What elements does the Dublin Core have? How can the these elements be integrated with for example the descriptive features of video?

projects & further reading As a project, develop a data format for text, images and video in XML, and implement stylesheets in XSLT to convert the format for display, for example in HTML frames or using SMIL.

You may further explore the formulation of criteria for selecting software and tool support for developing multimedia applications.

For further reading I suggest, apart from the manuals and learning materials that come with your tools, to study example projects and in particular the workflow, that is the dependencies between stages in the production, as for example explained in McCuskey (2002).

the artwork

1. website of Montevideo Collection[58] Catalogue[59]. To avoid being parochial here, I should also mention similar institutes abroad, such as Electronic Arts Intermix[60] from New York, USA, and LUX[61], from London, UK.
2. website of INCCA[62].
3. tangible virtual museum – from Rosenblum and Macedonia (2005), see section 10.1.
4. digital dossier – concept graph for *abramovic dossier*, see section 10.2.
5. digital dossier – presentation gadget in *abramovic dossier*, with video of *Relation in Time*, with Ulay.
6. digital dossier – installation *Terra dea degli madre*, as 3D model.
7. conservator's studio – *Self-Portrait with Braid*[63], see section 10.2
8. diagram – task world ontology, Welie et al. (1998).
9. tower of babel – location where the event took place, see below.
10. tower of babel – projection of *tower of babel* project, see section 10.3, submitted by Katelijne Arts. The project is a concept of Katelijne Arts, Tineke Goemans, Franka van de Goor, Leidi Haaijer en Bert Vogels.
11. tower of babel – a view from the inside of the building.
12. PANORAMA architecture – from Si & Eliens (2007).
13. signs – sports, van Rooijen (2003), p. 278, 279.

The artwork for this chapter is selected to emphasize *variety* and *experiment*. The collection of Montevideo contains a great number of works from the early history of video art, including the works of Nam June Paik and Bill Viola. Yet, despite the experimental flavor of these works, contemporary media art shows a strong

[58] catalogue.montevideo.nl
[59] catalogue.montevideo.nl
[60] www.eai.org
[61] www.lux.org.uk
[62] www.incca.org
[63] www.seattleartmuseum.org/exhibit/interactives/mexicanModernism/enter.asp

sense of *context*, *experience* and *communication*, as demonstrated for example in the *tower of babel* project. The issues of preservation we dealt with in this chapter, may now, to conclude this chapter, be summarized as: how do we preserve the *context of experience* of contemporary media art?