On the narrative structure of Virtual Reality walkthroughs

An analysis and proposed design

Pieter Jan Stappers, Daniel Saakes, and Jorrit Adriaanse[1]
Delft University of Technology
[1] SARA Computing and Networking Services

Key words: Virtual Reality, User-centered Design, Narrative theory

Abstract: Architectural walkthroughs have often been presented as prime examples of applications that can benefit from Virtual Reality (VR) technology, but still VR presentations can be disappointing. A main reason for this is that most VR applications have been developed on purely technical criteria, with an emphasis on geometrical precision rather than experiential quality. In this paper we present a human-centered analysis and propose design solutions, by focusing on the narrative aspects of the walkthrough, such as connecting transitions of the kind used in contemporary computer games. The solutions show how such narrative enhancements can improve the user's experience of presentations at modest technical expense.

1. WALKTHROUGHS IN VR PRESENTATIONS

Architectural walkthroughs have been among the most cited applications of Virtual Reality (VR), and are becoming a regular feature in the presentation of large architectural design projects. As always, such presentations are partly show, partly evaluation. Just as design sketches, VR presentations can elicit constructive criticism from clients and users who find it hard to read the formal language of plan views and construction drawings. The power of Virtual Reality (VR) presentations lies in its direct appeal to the senses.

Nevertheless, VR presentations can be disappointing as experiences. The audience is impressed by being immersed in a 3D animated environment, but notices that the graphic quality of VR is lesser than that of static renderings that can be made with a down-to-earth PC or workstation. Also, the simulated environment appears dead and empty, devoid of life which architects can readily suggest in conventional media, but which are hard to program in VR. The audience can see spectacular events, such as flying through space, but have no direct interaction with the environment: a guide controls the input device.

These criticisms are not a necessary feature of the technology. In another area of simulation, 3D computer games, the integration of technology and user-centered design has produced compelling experiences on relatively modest machines. Especially aspects of narrative are better represented, and some of these techniques can be fruitfully applied in VR walkthroughs as well. In this paper, we use narrative theory to analyse the structure of existing walkthroughs in the CAVE, compare this to recent computer games, and draw conclusions in the form of solutions for a user-centered design for architectural walkthrough presentations.

1.1 Role of the walkthrough in the design process

Figure 1 depicts an architectural presentation in the CAVE, a 3*3*3 meter cubic space in which small audiences can take part in a Virtual Reality experience. On its walls (in the Amsterdam CAVE three walls and the floor) stereo images are displayed whose perspective matches the position of a 3D sensor on one user's head. For that user, views in all directions correspond to a 3D environment. Other users close to this prime user, receive stereo images which appear a little distorted, but still give a compelling spatial impression. In this way, small groups of users can be led around and through simulated buildings.

A CAVE walkthrough is typically used for final presentations of large building projects. Participants in these presentations, e.g. staff of municipal councils or project developers, are often less trained in reading construction drawings, and benefit from experiencing the design in a 1:1 immersive simulation. The CAVE supports communication between the participants, also because the participants can see each other (which is not the case in helmet-mounted VR). Because multiple CAVE systems can be digitally coupled, the same environment can be shared by teams at separated geographical locations.

*Figure 1.*Participants in a CAVE presentation wear stereo goggles to see stereo images on four wall displays. In a typical architectural walkthrough, an architect and a VR specialist guide a group of people representing the client through parts of the proposed building.

The model of the proposed building that is viewed in the CAVE, is usually derived from a CAD model that had been made earlier in the design process. This derived model (which is often simplified to fit the needs of real-time visualisation) can be traversed by means of navigation software. It is important to note that the 3D session is not an isolated product, but plays a part in a larger presentation.

1.2 Worst-case scenario

We illustrate the shortcomings of many current presentations in a worst-case scenario. This scenario is exaggerated, but we have seen all its elements in various demos in various labs. The scenario is not meant to criticize the makers involved, but to indicate aspects which are easily overlooked because they are often thought as 'not a technical part of the simulation'. Figure 2 outlines the sequential structure of the scenario.

Before the simulation has started, the audience, filled with anticipation because they have been told to expect 'the wonders of VR', enter a heavily wired room in the basement of a computer center, where technicians are just calibrating the equipment or making last fixes to programs. Then they go through a confusing ritual of putting on stereo goggles, the light is turned off and the VR displays are filled with command line displays showing technical data.

After a warning by the guide, a 3D world suddenly pops up around the audience, and the guide flies them through the world. Because interaction is still a cumbersome part of most VR applications, the audience can only look around, while the guide controls movement through using a spatial joystick. Although the 3D view is impressive, the image quality is less than that of the renderings of the building that they saw before. The audience quickly finds out that they cannot easily indicate parts of the scene by pointing with their hands. Perspective distortions in the CAVE make it impossible for one person to judge what direction another person is aiming at (CAVE perspectives are only geometrically correct for the single user who is wearing the tracking sensor). Moreover, the simulated building appears static, dead, clean, straight from the CAD package, and participants find it hard to orient themselves in the building. After the guide has shown the different parts of the building, the world disappears as suddenly as it came up.

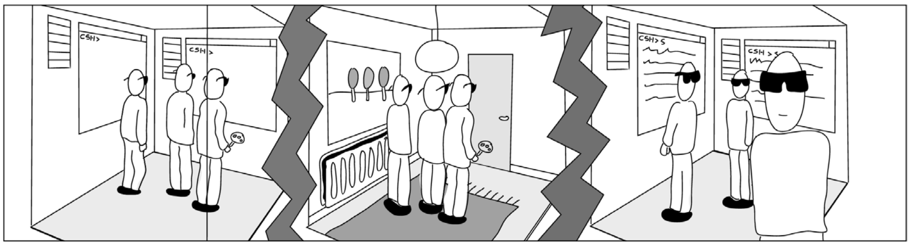

Figure 2. In the worst-case scenario, the simulation is sandwiched abruptly between screens with technical data which lessen the quality of the VR experience.

After the simulation ends, the audience is slowly ushered out of the room. The CAVE contains nothing to remind them of their walkthrough. Although the audience has seen the geometry, their experience might have been much more compelling.

Some VR sessions continue with a few spectacular demos in order to impress and amuse the guests. These demos can provide an unintended counterpoint to the walkthrough itself, because they contain graphics and interaction that was designed and tuned to impress the audience, and therefore are experienced as 'much better than the walkthrough', creating a sense of disappointment about the architectural presentation itself.

1.3 Reality check

The scenario sketched above does not do justice to the current state of the art, but all of the elements can be found in current practice around the world. In our own practice we found

– people were not able to recognise the VR model of the existing building in which they worked. The simulation lacked environmental detail and spaces filled with people.
– people were disappointed by the graphics and interaction. Their expectations had been raised by special effects in games and movies.
– a lack of support for the architect giving the actual presentation.
– the walkthrough not adding enough value to the presentation to justify the high costs involved.

In general the VR walkthrough lacks the interesting interactive techniques of VR and also lacks the richness and expression of traditional presenting methods. Some of the techniques which we will discuss below are used in some demonstrations at SARA, but are not standard practice.

In this section we have outlined some limitations of current VR practice: although the graphics can be impressive, the user experience is often found lacking. This should not be surprising, since VR originated in computer graphics laboratories, with the addition of richer input devices, but only more recently has it received interdisciplinary user-centered study and design.

2. THEORY: NARRATIVE STRUCTURE

Narrative theory was popularized in the field of computer-human interface design by Brenda Laurel with her book ‘Computers as Theatre’ (1993). The theory has widely been praised, but real applications of the theory are hard to find. Virtual Reality walkthroughs can be an excellent case study, exactly because the presentation of a building is a form of storytelling, of engaging an audience into an experience.

Here we distinguish three narrative elements. First is the role of time. Many VR simulations consist of placing the user in a piece of simulated geometry, and allowing him or her to fly through it for the duration of the experience. But people's interaction with products is not an iteration loop. Just like a book or a movie, it has a beginning, a middle, and an end, and sometimes an intricate structure in-between.

Second, the role of space. The world in which we live is littered with traces of history, little details which help us distinguish where we are, what we and others have done, and what we can do in the future. Most VR simulations are empty, and present a geometrical world which is experienced as 'dead', 'expressionless', 'too abstracted to be real'.

Third, the role of action. In the natural world, and in a play, we act upon the world by using our body, not by issuing commands. For example, we open a drawer by pulling it with our hand, not by speaking to it, either by voice or through a remote-control device.

In this section we illustrate how these roles have been used to heighten experience and engagement by the makers of 3D computer games. These games, which emerged in a highly competitive and experience-oriented market, can teach us 'tricks of the trade' which can be used to improve the experiential quality of VR simulations.

2.1 Time: Sequential structure in computer games

Figure 3. Advertisements, packaging, cinematics and gameplay together enhance the game experience in computer games such as Core Design's 'Tomb Raider'.

The experience of a computer game player starts long before the game titles appear on the computer screen. In the case of ‘Tomb Raider’, the main figure was billed in the way of a movie star, long before the game could be bought. People who bought it first heard about it from an orchestrated press campaign, then went through the ritual of opening the (carefully designed) box.

When the game commenced, the actual gameplay was preceded and interspersed with cinematics, prerendered animations which are of much higher graphic quality than the interactive part can be. An introduction provided a background (date and place of birth, social class, etc.) to the main figure and storyline and atmosphere that extended far beyond the jump, search, and shoot actions of the game itself. Interestingly, players often remember the visual quality of the gameplay as that of the cinematic scenes.

At the end of the game, the player has to combat the dragon in the final confrontation. When he or she wins, a special cinematic 'reward' fragment follows. In the past years, several papers in the professional press have addressed the structure of the storylines in this kind of games (e.g., Costikyan, 2000).

2.2 Space: environmental narrative

In many respects, it’s the physical space that does much of the work of conveying the story the designers are trying to tell. Color, lighting and even the texture can fill an audience with excitement or dread (Carson, 2000). Objects in the environment provide feedforward: keys lying about and footstep trails of footsteps on a floor inform the player of interesting new locations, and how to get there.

2.3 Action: enactment versus indication

Despite the limitations and diversity of the input devices that the games must support (keyboard, mouse, joystick, and game paddle must all work), the player is tied to the environment through actions tuned to the tasks in the game. He or she acts upon objects instead of giving disembodied verbal or key-press commands according to an arbitrary convention.

2.4 Conclusion

Many of the narrative elements do not affect the technical performance of a simulation. Many of the elements take place *outside* the simulation itself, but carry a psychological enhancement into it. The same elements can be used to improve VR simulations. This does not mean that architectural walkthroughs should be made into computer games. But it does suggest that some narrative devices may be used to enhance the user's experience of the proposed architecture. In the final parts of the paper we discuss how simple solutions in the setting, the hardware, and the software can improve the look and feel of existing walkthroughs, and better integrate the Virtual Reality simulation within the larger structure of the architectural presentation.

3. APPLICATION: NARRATIVE ELEMENTS IN WALKTHROUGHS

In this section we discuss the architectural simulation within the framework of narrative theory.

3.1 Sequential structure

The transition from the real world into the virtual environment is often severe. As the guests enter the hall which contains the CAVE, the CAVE's walls displays double images in distorted perspective, because the sensor and stereo goggles are hanging from the ceiling. Once the tracked user wears his goggles, the 3D environment can be experienced.

The second transition is the one that brings the user in the surroundings of the building, or into its interior. The route in the environment reflects the story that the architect wants to present. In our own experience we found that some architects want to start with an overview, while others start with details gradually exposing more and more, to create tension and to set expectations.

The route shows the rhythm of the building, the contrasts between small and large spaces, and is the key benefit of VR compared to the traditional presenting methods. The walk through the building conveys a much richer experience of it, than can be had from viewing a number of vistas. A carefully planned route, can make spaces look bigger, lighter etc. Currently, the route and its transitions are all implicit, 'in the head' of the guide, *enabled* but not *supported* by the simulation. In section 5 we propose a few solutions to support the guide in telling the story.

3.2 Narrative in the environment

The task of the guests in a presentation is to look around and understand the building, and we want to make this limited task interesting by creating a rich, self-explaining environment. Let the users discover the function and atmosphere of spaces themselves, rather than having the guide tell them what to see or feel. If the simulated building is empty of furniture, it is hard to imagine it as rooms that are used. Simple furniture gives cues about the scale, but what's more important it explains by whom and how the room is used. Props and lighting direct the guests' attention, and allow the guests to indicate places (to which they cannot point) by naming them, e.g., 'the door next to the painting'.

All the elements that define a room, the ambience and cues must be added to the visualisation. In immersive VR we need to explicitly show things which in sketchy renderings can often be left to the imagination.

A special case is adding people, for scale, to denote a room's function, to depict its atmosphere. In drawing classes all architects have learned quick and sketchy methods to add people and furniture to visualisations. In VR there is a difference between images of furniture and people, in that the VR users are part of the visualised scene, and therefore expect people to react to their presence. In the Alladin VR-ride, Pausch et al. (1996) found that users did not mind that simulated agents were rendered in cartoon-style, but were disturbed by the fact that the agents did not react to the presence of the users. When we populate our architectural simulation with simulated people, we must make sure these react at least in some way to the proximity of the guide and guests.

3.3 Enactment

The example just given of simulated actors was already an example of the role of action. In current presentations, the guests have no direct interaction with the environment. All control over the presentation is in the hands of the guide. An experienced guide in the demonstrations at SARA can give the user the needed feedforward by announcing that interesting things are about to appear around a corner. The anticipation thus created enhances the experience of the audience.

One way to enhance the experience is to let the guide open a door by grasping its handle rather than pushing a button on the remote control. The grasping action is visible to all participants, and is experienced as a natural action in the context of going through a door. Although the guests do not act themselves, this action heightens the suggested connection between the virtual and real environments, and therefore heightens their sense of presence in the virtual environment.

3.4 Proposed solution: a better scenario

We conclude this section by presenting solutions in the form of a scenario. The added value of Virtual Reality systems is that the proposed design is not just told by the architect, but also that the guests can themselves experience, discover, and explore the design. The simulated spaces must tell their story, and enhance the limited task of the guests: looking around, pointing out interesting places or problems. Solutions to this problem involve enhancing the simulation in the direction of expressiveness, which need not be photorealistic modelling. In this section we present a number of solutions, none of which involve much technical effort, but try to improve the fit between the simulation and the users' needs on the basis of existing technology.

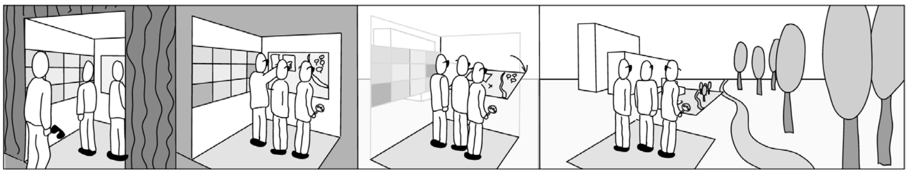

Figure 4. At the start of the proposed scenario, the CAVE displays introductory 2D presentation material about the proposed building project. When the 3D goggles are active, a smooth transition is made to 3D presentation, and the building is approached.

3.4.1 Before the simulation starts

The users enter the large space with the CAVE, the hall is dark, and the CAVE is light and attractive. The walls of the CAVE are filled with 2D images, animations and maybe music. These images provide the background story and the lit CAVE attracts the users to enter it. The imagery is displayed in mono, on the walls, so there’s no perspective distortion or double images. While standing in the CAVE the users put on the goggles.

The architect starts presenting using the CAVE as a large, conventional slide display. This preparation ends with the plan view, the map of the building, depicted on one of the CAVE's walls. While the guide is talking and pointing at parts of the map, the simulation is loading.

3.4.2 Starting the walkthrough

The 3D experience begins when the guide 'pulls' the map of the building down from the wall, and it turns into a tabletop-size scale model, and the building can be discussed on model scale. The walls of the CAVE gradually disappear to create a smooth transition to the virtual environment. The model helps to smoothen the passage of the guests from the real environment into the virtual environment. It connects these environments, preventing the guests from becoming disoriented.

3.4.3 During the walkthrough

In the virtual environment there are simulated actors which support the route. The guide can 'hook' the navigation software to these actors, so the group follows the actor rather than flying through space. The actor provides feedforward by leading the way, strengthens the tie between the route and the virtual environment, and gives the movements a sense of purpose related to the story. So in our story the guests follow the actor to the building. The building is not shown in isolation, but other buildings surrounding it, are suggested by some form of outlines. The environmental context helps to set the mood. The simulated street environment contains the suggestion of life and traffic: cyclist and cars pass by, and react to the presence of the guests. A cyclist can look at the users, a car stops while the users cross the street.

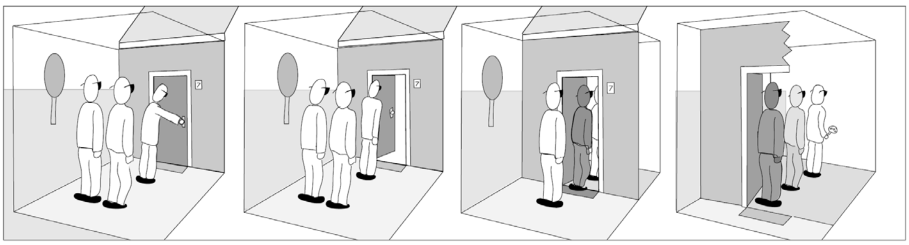

Figure 5. Opening ritual and light effects enhance the experience of entering the building. The building's wall is aligned with the physical display wall, so optical distortions are minimized.

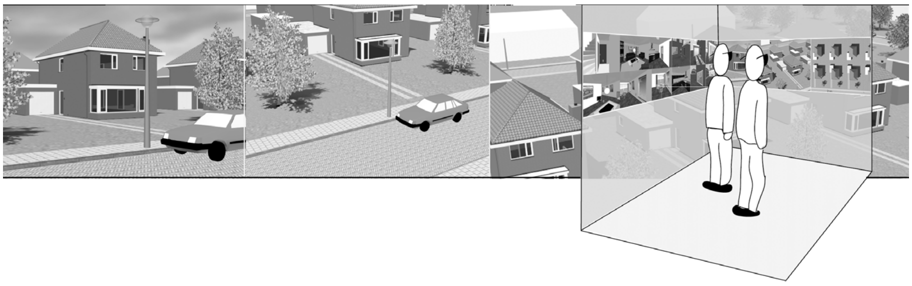

Figure 6. After the walkthrough, the viewpoint rises to an overview, and a connecting transition is made to 2D imagery showing highlights of the tour. and images of the places they visited. (Model courtesy of Zegelaar Onnekes BV)

When the group halts at the door of the building, a new transition occurs. The guide spends some time talking, then grasps to open the door and the users enter the building. We don't promote to always open all doors in such a ritual, but this ritual is important to stop and experience the entering of a building for the first time.

Two technical problems need to be overcome in this transition. First, in order to eliminate perspective distortions during the talking and grasping of the doorknob, navigation stops not just anywhere in space, but in such a location and orientation that the door and wall are aligned with the display walls of the CAVE (in that case, distortion is absent). Second, because the lighting model used in the CAVE's graphics software, typically indoor rooms are brighter than outside scenes, whereas in the real world a reverse relation holds. We can still give people the sensation of entering a building by explicitly modulating the overall light intensity at the moment they move into or out of the building. As the group enters, the overall light intensity is dropped quickly and noticeably, after that it increases again slowly and less noticeably. By tuning these transitions, the experience of entering a building is given both dramatic impact and a heightened degree of realism.

What we just described involves not only graphics but also sound. Environmental sounds can convey a strong sense of place, distance, and atmosphere (Mountford & Gaver, 1990). Sounds of traffic can easily be muffled or attenuated when the group enters a simulated building.

A building consist of important places to visit but also of less interesting spaces. We lead the users to the more attractive places by placing *attractors* to guide them, i.e., objects like furniture that draw the attention. Localized sound can also be used to draw the attention. In the less interesting spaces, movements can be increased a bit. Note that these additions do not take away functionality, but rather invite optimal behaviour. The guide can still move toward boring spaces or leave the route when he or she needs to.

The rooms are filled with simple cues which make the box-like simple geometry into rooms, cues like skirting boards and window frames, and dirt on the floor leading up to much-used doors. The aim in all these additions, however, is not to increase the supposed match between the simulation and a physical reality, but to increase the match between the simulation and the presentation's message (Stappers et al, 2001).

3.4.4 After the simulation

The walkthrough ends with the users rising on an elevator platform and thus get an overview of the building. Representative views of parts of the tour appear on the walls of the CAVE. Just as holiday snapshots, they remind the audience of the things they have seen, and help to reconnect the virtual environment to the real environment. These images draw the users out of the virtual environment, and the discussion continues, but now with an overview of the building and all the important images and details compressed in one space. Again the imagery is 2D, and the users can take off their goggles. When the talk is over, the users leave the CAVE and continue their talk in the conference room. There they may find the pictures summarizing the walkthrough in a printed form.

3.5 Partial conclusions

During this research we found that solutions which enhance the quality of the narrative (pauses, breaks, anticipation, and climax) can at the same time solve problems at the more mundane technical level of presentation. For instance, when presenting to more than one person in a CAVE environment, confusion often arises when someone points with a finger, because apparent distortions of perspective stereo lets different viewers perceive different directions of pointing. A solution is to place non-obtrusive, natural objects near points that are likely to be discussed, so they can be named. Similarly, objects placed beyond open doorways can be used to draw the viewer's attention, in the way some artists entice the viewer's eyes to move over their painting in a certain way, or the way in which objects lead the way in some maze-solving computer games.

3.6 Incorporating narrative elements in the development process of architectural presentations

The scenario sketched in the previous subsection relied on technically simple tricks to enhance the support for the narrative elements of the presentation. In themselves, these tricks are not difficult. The difficulty lies in finding out what effects the hardware and software are able to produce (knowledge of the VR specialist), finding out what the presentation storyline needs (knowledge of the architect), and merging these knowledge into a well-tuned set of support effects. This requires communication between experts from different backgrounds, knowledge, and skills. In the current project, this merging was performed by an industrial design engineer. In the future, a formal structure is needed to support and promote the required communication between user-centered and technology-centered expertises.

A structured approach goes beyond a choice for a certain software package or library. No current software packages help the architect to see what narrative-sustaining effects the CAVE can perform, or what it costs to prepare them or to run them. And the technology changes rapidly. Neither can or should we turn VR specialists into architects. What is needed is finding a joined language for both parties.

One such language has been found in '(software) patterns', structured expositions of a problem, its importance, example solutions, and observations. This technique of describing, pioneered by Christopher Alexander (1977), and taken up enthusiastically by the software engineering community (e.g., Gebriel, 1996) and recently by the game developer community (Hecker & Simpson, 2000) can provide a common platform. In patterns, a concrete prototype plays the central role. All the discussants understand the prototype and can discuss it from their own perspective. Through the concrete nature of prototypes, a catalogue of techniques and parts can be grown. These may not be as formally unified as a software library, but are also more generally applicable.

The resulting effects can be used as a standard, shared, bag of tricks, which is enhanced and added to over time, and which helps in the specification of elements for new presentations.

4. CONCLUSIONS

In this paper we have argued how high technology VR systems in themselves do not guarantee highly convincing experiences for users. The level of experience and engagement in some commercial 3D games goes well beyond those of most academic and commercial VR presentations, for a large part because the former develop their applications from a user-centered criterion. We have shown how narrative theory, originally developed for theatre, but in the past decade also applied to product experiences, can be applied to enhance the experience users get from architectural VR simulations. All together, the results show how a high-end

VR architectural presentation can be improved by learning from the trade secrets of the game industry.

The proposed design solutions in this paper address only some aspects of a simulation, but indicate a practical way, through patterns and prototypes, through which narrative elements can be brought into the development process of a VR developer. Rigorous testing is still needed, and a suite of examples will have to be grown and tuned, before the solutions we present settle in the working practice of VR centers.

5. REFERENCES

Alexander, C. 1977, *A pattern language*, Oxford University Press, New York.

Carson, D. , 2000, Environmental storytelling *,Gamasutra web archive,* http://www.gamasutra.com/features/20000301/carson_01.htm

Costikyan, G. (2000) Where stories end and games begin. *Game Developer*, Sept., 44-53.

Cruz-Neira, C., Sandin, D.J., & DeFanti, T.A. (1993) Surround-screen projection-based Virtual Reality: the design and implementation of the CAVE. *Computer Graphics: Proceedings of SIGGRAPH '93*, August 1993.

Gebriel, R.P. 1996, Patterns of Software, Oxford University Press, New York.

Hecker, C., & Simpson, Z.B. (2000) Game programming patterns & idioms. Game Developer, Sept., 12-13.

Laurel, B. , 1993, *Computers as theatre*, Addison Wesley Longman Inc. , City.

Mountford, S.J., & Gaver, W.W. (1990) Talking and listening to computers. In: Laurel, B. (Ed), *The Art of Human-Computer Interface Design*. Addison-Wesley, 319-334.

Pausch, R. , Snoddy, J, Taylor R. , Watson S. and Haseltine, E. , 1996, "Disney's Aladdin: First Steps Toward Storytelling in Virtual Reality", *Computer Graphics Proceedings, Annual Conference Series*, 23(4), p. 193-203.

Stappers, P.J., Gaver, W.W. , & Overbeeke, C.J., (2001) Beyond the limits of real-time realism. In: Hettinger, L.J., & Haas, M.W., *Psychological Issues in the Design and Use of Adaptive, Virtual Interfaces*. Erlbaum, New York