

# The concept of Carrying in Collaborative Virtual Environments

*Martijn Stellingwerff*  
*Faculty of Architecture, Sector MEDIA*  
*Delft University of Technology*

## 1. Introduction

Collaborative Architectural Design can take place within a virtual environment with a team of remote but virtually present people. However, in most virtual environments, the ability to perform actions is still limited to the availability of some interactive objects and a set of tools for the specific purposes of the system. As the interface of most systems is designed for unshared use, the graphic feedback signals are limited to local information about the state of objects and tools. If multi-user interaction is added to such Virtual Environments, many new possibilities and problems emerge.

Users of shared applications should not only be informed about the state of local objects, tools and their own actions, they should also be made aware of what the other users undertake. Aspects, which are in daily life so obvious, should be re-studied thoroughly for the application within Virtual Environments for Collaborative Design. Much research has to be undertaken in order to make such virtual places as intuitively interactive as ordinary shared working places.

The 'concept of carrying', which is proposed and explained in this paper, is expected to become a useful metaphoric mechanism for solving several issues related to Spatial User Interfaces (SUI's) and Collaborative Virtual Environments (CVE's). The visual feedback from 'carrying-events' should provide more mutual understanding about ongoing processes in shared applications and it should add a more 'natural' interface for processes concerning people, tools and content in virtual and digitally augmented environments.

At the start of this paper some basic human action patterns for tasks on a 2D-desktop are compared to tasks in a 3D-environment. These action patterns are checked for their implementation in Windows Icons Menus and Pointer (WIMP) interfaces and Virtual Reality systems. Carrying is focused upon as an important interactive event in Virtual Environments. Three carrying actions related to Collaborative Architectural Design are explained by means of prototypes in Virtual Reality Modeling Language (VRML). Finally the usefulness of a general carrying concept as part of a new Visual Language is considered. The research at hand is in its first exploring phases and draws from a running PhD research about SUI's for Context Related Architectural Design and from recent experiences in CVE's.

## 2. Human action patterns

A commonly accepted method for the consideration of human factors in interface design is the observation and analyses of human behavior in simulated or real-life situations. In order to find patterns in the behavioural systems, a researcher can describe each single step within a chain of actions. The test person can be instructed to 'think aloud' in order to clarify the motivations and considerations for certain actions. By means of protocol analyses, every action is interpreted and coded in schemes. Generalisation of the schemes can enlighten certain action patterns, which can be used to improve and optimize systems, processes and machines.

The development of the desktop metaphor for computer operating systems is the result of interpreted human action patterns in real offices, conference rooms, archives, etc. While the computer desktop has been developed for over three decades, the development of a visual language for Virtual Environments (VE's) has just started. The development of experimental VE's is a challenging subject for research in the field of human action patterns and further development of visual language. With the continuing and enormous increase of rendering speeds and the possibility to share information over the Internet, many complex interactive processes should be guided by accurate visual interfaces. Many aspects of the Graphic User Interface (GUI) can be transformed into SUI-aspects, but other aspects still have to be invented. The description and comparison of real life actions for documents on a desktop and actions within a city can be helpful for the design of new visual language aspects in VE's. The further investigation of human interaction and the setup of experiments can elaborate on system aspects for Collaborative Virtual Environments.

### Actions at a desk

Desktops, whether part of real furniture or metaphors on computer screens, are surfaces where documents are processed. The order, structure and design of a desktop can be understood without considering what tasks, such as writing and drawing, are actually performed. People sitting behind a desk are in general just busy with documents. They read, create, change and store documents. The media for documents are pieces of paper, reports, books and document-windows on a computer screen. Near the desktop, there is a filing cabinet. This cabinet can be compared to a server computer or disk storage. All documents can be seen as the objects of actions. The measurements and divisions of the filing cabinet, the sorts of paper and the size of the desk are all fine-tuned for the average working methods of people in an office.

The real behaviours of the person behind the desk are specific for each person. Those behaviours can range from very orderly to extreme chaotic habits. Nevertheless, most people do not get lost in their own piles of documents. Each person uses small sub-actions to arrange their daily tasks. Some people make diversified piles of new, unfinished or 'in' and 'out' documents, others stick one or more small yellow notes to each document. Computer desktop interfaces provide

the people with similar mechanisms such as: shortcuts, aliases, virtual yellow notes, bookmarks, cut-copy-paste operations, drag-&-drop operations and different states of document windows. About all traditional actions have their metaphorical or novel counterparts in the computer interface. The traditional work at a desk is successfully transformed into work at a computer screen.

### **Actions in urban or interior space**

Most actions in urban space or in the interior of a building are related to motion. Meeting someone at a given place (occasionally or by chance), exploring a building or a city, shopping, even going on a pub-crawl, it all has to do with moving from one place to another.

When we observe the people in urban space, the first thing that catches the eye is who it is and that he or she is going somewhere. If you ask these people what they do, they will tell you where they go, in order to do something. What they actually do is moving and taking things to other places. When they leave, they collect things to be carried. When they arrive, they get rid of the things; they deliver, store, drop, use or give away. They carry many different objects such as money, presents, documents, tools and maps. The places in which they keep the things while they carry them are their pockets, hands, cars and bags. Carrying is a method to have things available while you are not 'at the desktop'. While 'drag-&-drop' is a useful mechanism for WIMP interface events, 'carrying' might be a good metaphorical mechanism for multi-user events in simulated spatial environments.

### **Carrying is a medium**

What and how people carry things can inform other people about their intentions. While real life situations are rich of signals by which we fine-tune our actions and social behaviors, a meeting in a virtual space is often radically simplified and can lead to blunt actions and unrefined messages.

In real life, carrying is a medium and what we carry is the message. Imagine someone carrying a gun, a nice odor, a rotten tomato, a photo camera, a badge, a hat, a backpack, a child, an adult, a pencil, a mask ... It gives much information. The message is expressive and high in participation or completion by observers. Carrying is, if I understand McLuhan's media theory right, a cool medium. The message is attractive but it is not yet articulated in a way that we know how the gun, the pencil or the camera will be used.

The rather informal way in which carrying informs a person in an environment makes it an interesting concept for Collaborative Virtual Environments. It informs about the intentions of each participant and can be the starting point for a more articulated virtual meeting.

## **3. Prototypes for carrying**

An avatar is the virtual embodiment of someone in a CVE. Avatars can be found in many forms on the Internet and they become increasingly detailed and capable.

Software, like Avatar Studio, works as a factory for avatars. Body type, gender, dimensions, skin color, clothes, coiffures and paraphernalia can be adjusted in order to mimic or contrast to the physique of the real person who uses the avatar. Another interesting technique is the capturing of avatar data from real persons in an enhanced 'passport photo booth'. ([www.avatarme.com](http://www.avatarme.com)). The still developing VR standard for humanoid animation, 'H-ANIM', makes it possible to give the avatar a set of behaviors, such as: nod, walk, wave, shake hands, clap, etcetera.

Current avatars are still very limited; most of the advancements are related to social aspects and the visual profile of the avatars. The relation to, and the interaction with the Virtual Environment are in most applications limited to wandering around, chatting and being seen. In order to be useful for collaborative design and design education, actions of avatars should be developed further and the superficial physique should be considered in a more critical way.

The following prototypes attempt to broaden the above-mentioned developments and techniques. Three different carrying concepts are described in order to explore the identity, the ability to orientate and the mechanisms for interaction of 'next generation avatars'.

### Carrying an identity

Essentially an avatar carries all information related to a single user in a multi-user environment. Carrying an identity is a main feature for an avatar. Identity enables to distinguish between different participants. The identity of avatars should be unique, as their owners are unique. Avatars of programmed agents could be carrying a not-unique, clone-like, identity.

At the moment, users of most CVE's do not have a personal/unique avatar. They get a general avatar body or they can choose from a limited list of available avatars. This leads to confusion of other participants in the CVE. Future CVE's should get a mechanism that prevents against misuse of identities and vagueness of redundant avatar types. Carrying your own identity is a clear and polite attitude towards the others. As a solution for the problem of fake identities, the use of a specific avatar could be linked to the username and password.

Identity can be fixed by the shape of the avatar and by the name it carries. In CVE's for architectural design and other applications with a distinct goal, the complexity and amount of realistic details of the identity-shape is not a first priority. Therefore the here presented prototype for 'carrying identity' is kept sober. This initial

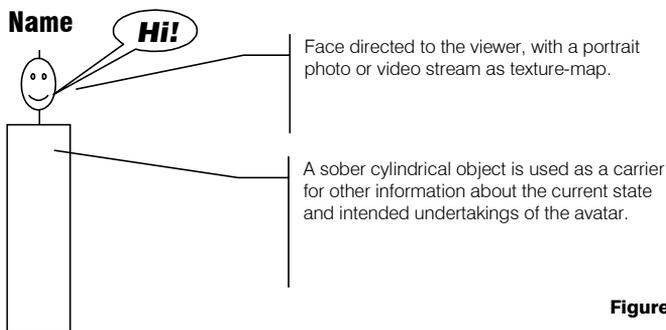


Figure 1.

soberness helps in image-clarity when other carrying objects such as tools, maps and dialog texts are added to the avatar.

Below is a prototype for an avatar that carries the identity of a participant. For a better understanding, the personal chat-texts of each participant in a CVE should be carried by their avatars as well as it should be displayed in a chat-window.

### Carrying viewpoint lists, maps and each other

When it comes to orientation in a CVE, we easily get lost. Most VR environments provide only visual feedback at eye-level (there are no 'multiple view-ports' like in most CAD programs). Fast turns and uncontrolled movements can easily lead to disorientation. As a first remedy, avatars can carry a list of places where they have been. The list should be extendable and each item should contain data about the viewpoint, the view direction, accompanied by a small view image or a description. Another 'carry-able' item for navigation is a map. As a prototype a carry-able map was programmed (see figure 4). The specific map can be picked up when a specific area is entered. The map moves or scrolls according to the movements of the avatar. If the avatar turns, the map is turned accordingly. When the area is left or the user becomes familiar to the environment, the avatar can drop the map or keep it for later use.

The third mechanism for orientation in CVE's is 'carrying each other'. This sounds a bit strange, but it is no more than sharing the same view with someone else. Avatars should be able to perform a guiding role and invite others to be carried by them. They can advise each other in order to orientate and explore a new environment.

Carrying implies that objects can be picked up, dropped and exchanged. This also applies for carried viewpoint lists. With this kind of 'list-keeper' mechanisms, avatars can keep track of where they have been and where they want to go. They can guide each other by sharing and exchanging viewpoints and points of view. A group of carried people could be represented by a group-avatar, e.g. a multiple headed creature. A virtual bus tour can be used as a metaphor for shared, animated, viewpoints in a CVE. The bus can be seen as an agent for navigation.

If the person in a CVE wants to see his or her own avatar, they can have an 'Out Of Body Experience' (OOBE). The OOBE makes the viewpoint shift slightly away from the own avatar's viewpoint. This can also help to understand the own position in relation to other avatars and related to the Virtual Environment.

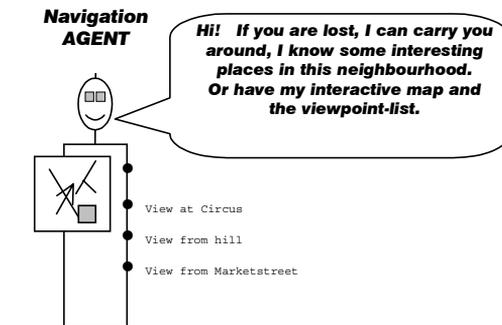


Figure 2.

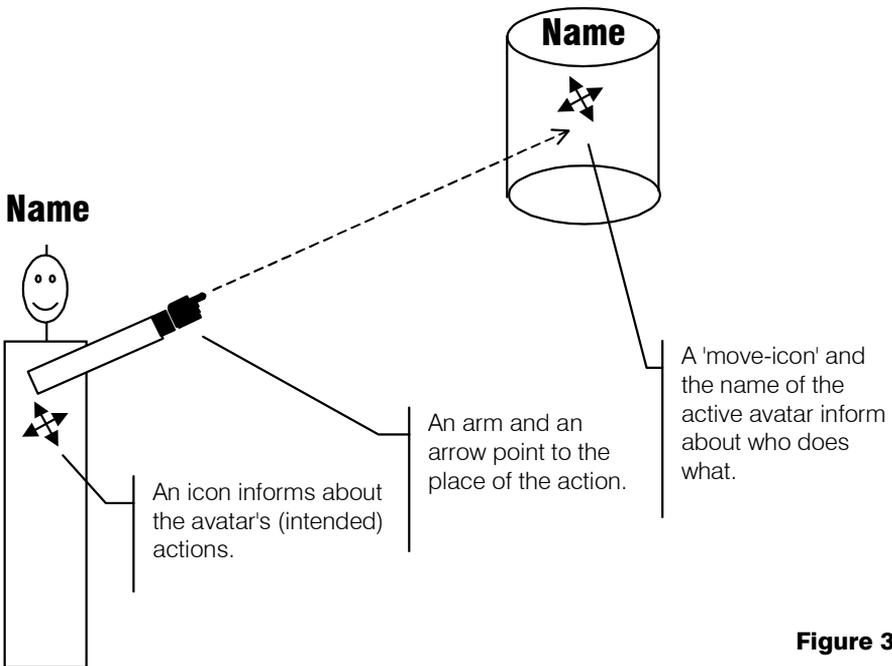
### Carrying tools for pointing, painting and dragging.

In a single user environment, an arrow on the screen represents the movements of the mouse. If a pointing action of another user in a 3D multi-user environment should be represented, this cannot be a similar arrow, as it would give conflicts with the arrows of other users and with the many different points of view.

The pointer of a user should be perceived in relation to his or her avatar. A direct connection between the avatar, the pointer and the object or place that is pointed at, should be visible.

As a prototype for pointing and performing actions in CVE's, an avatar with a pointing arm and hand was built (see figure 5, next page). By means of several goniometric equations, the arm is rotated in the direction of the pointed point. From the pointer's point of view, the pointed point is coherent with the perspective, so there the normal pointer suffices. All other users in the CVE see the avatar 'carrying an arm', which points in the good direction.

In order to enrich the pointing actions, the arm is extended with a line to the pointed object. Additionally, the name of the acting avatar is displayed near the object. According to the specific action, e.g. move or paint the object, an icon is shown on the avatar and on the object.



**Figure 3.**



Figure 4: While walking with the carry-able map, it becomes a 'scrolling' element on the screen. The map rotates according to the viewing direction.

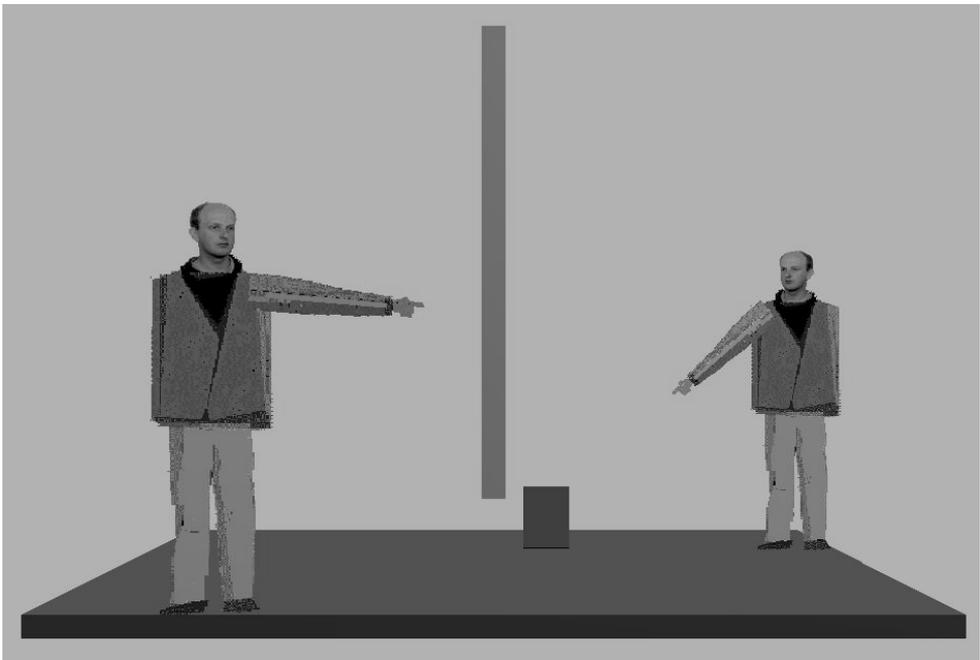


Figure 5: The author's avatar 'carries' his portrait photo and a pointing arm.

#### **4. New visual language aspects**

In the recent development of computer interfaces, we saw successively: punch cards, printers as output device, text-screens, vector-screens and dot-matrix screens with Windows, Icons, Menus and Pointers (WIMP interfaces). At the moment we get acquainted with interfaces in Virtual Environments. Especially for architects, and researchers at Faculties of Architecture, this is most interesting. Spatial interfaces are still underdeveloped and the further developments can benefit from architectural knowledge.

Each time when a new medium is developed, a new language emerges as well. The avant-garde artists explore the medium in all its aspects. They seek out all expressive means that are possible with the new media, and they add new inventions to the general diversity of human languages. These new inventions even have their repercussion on older media. For instance the brutal 'cut-off' way of framing a situation in photography can be found in several impressionists paintings as well. On the other hand, many new media are used in an old fashioned way. Again in the early days of photography many studios used the type of attributes and compositions that were custom in painting.

Personally I feel tempted to find new visual language aspects that become available with Virtual Reality. At the same time I am aware of the limits and biases that come from my TV-background which, although it is rich in visual language characteristics, is NOT interactive and VERY linear and FLAT in its narrative language. The next generation of researchers and artists has a game-background. They are brought up with non-linearity and many novel ways of game interaction.

Not surprisingly it is the Games Industry where the avatar comes from. In Multi User Dungeons, the avatars are trained for fights and (other) 'social' activities. The concept of carrying should free the avatar from those dungeons. The avatar is supposed to become increasingly important as part of a new visual language for CVE's. Carrying is suggested as an extension to avatars, in order to give them more 'body language' and to relate their behaviors to the represented context.

A general carrying concept as part of a new Visual Language is considered to be fruitful for application in Virtual Reality and Digitally Augmented Reality systems. A more capable type of avatars could solve the communicational aspects that are needed in a multi user environment. The carrying avatar becomes part of the interface and informs implicitly about what the others undertake.