



# A Survey of Communication and Awareness in Collaborative Virtual Environments

Thi Thuong Huyen Nguyen, Thierry Duval

## ► To cite this version:

Thi Thuong Huyen Nguyen, Thierry Duval. A Survey of Communication and Awareness in Collaborative Virtual Environments. 2014 International Workshop on Collaborative Virtual Environments (3DCVE), IEEE VR, Mar 2014, Minneapolis, United States. hal-00974022

**HAL Id: hal-00974022**

**<https://inria.hal.science/hal-00974022>**

Submitted on 4 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Survey of Communication and Awareness in Collaborative Virtual Environments

Huyen T.T. Nguyen\*

INRIA Rennes Bretagne-Atlantique, IRISA, UMR CNRS 6074, France

Thierry Duval†

Université de Rennes 1, UEB, IRISA, UMR CNRS 6074, France

## ABSTRACT

In the domain of Collaborative Virtual Environments (CVEs), many virtual worlds, frameworks and techniques are built based on a specific and direct purpose. There is not a general and still good and efficient enough solution for all the collaborative systems. Depending on the purpose of the collaborative work, the techniques of interaction and of manipulation change from one application to another. Despite this difference between interaction techniques, they always benefit greatly from awareness features that help in explicating implicit knowledge related to one's own and others' working activities as well as to virtual workspace. In addition, people in CVEs also use communication channels to negotiate shared understandings of task goals, of task decomposition and of task progress. Therefore, awareness and communication are usually considered as "instruments" to complete collaborative tasks in the environment. However, few research work have been devoted to improving the awareness and the communication channels in CVEs for a better collaboration between users. In this paper, we will study the importance of awareness and communication in collaborative virtual environments. We will investigate different kinds of awareness which need to be carefully designed. We will also discuss different communication means and how to cope with this diversity, so we can benefit from the availability of different peripheral devices and can find an effective communication means to work together. Finally, we will make some propositions to overcome these actual limitations of CVEs.

**Index Terms:** H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: User Interfaces—User-centered design; H.5.3 [Information Interfaces and Presentation (e.g., HCI)]: Group and Organization Interfaces—Computer-supported cooperative work; I.3.6 [Computer Graphics]: Methodology and Techniques—Ergonomics, Interaction techniques; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

## 1 INTRODUCTION

Collaborative Virtual Environments (CVEs) allow people to work together on shared 3D artifacts and datasets or exchange information. A collaborative working cycle in CVEs includes the following steps: perception and awareness, communication and negotiation, coordination and performing. This cycle can be considered as an extension of the Model Human Processor cycle [7] including Perceptual Processor, Cognitive Processor and the Motor Processor which is interpreted as the Perception - Decision - Action Process. A collaborative task begins by the perception and awareness steps when users get to know the activities happening within the temporal and physical bounds of the task in a visual workspace [17]. The next step is to communicate with other people in the same virtual world to enhance the awareness knowledge as well as to make an

agreement about how to perform the task. Once the agreement has been achieved, the performing step takes place while the communication and the awareness steps still go on.

In the collaborative working process, the two important steps including coordination and performing, will highly benefit from awareness features that help in explicating implicit knowledge related to one's own or others' working activities. In addition, people in CVEs also use communication channel to negotiate shared understandings of task goals, of task decomposition and of task progress. Therefore, awareness and communication are usually considered as "instruments" to complete collaborative tasks in the environment.

In this paper, we will study the importance of awareness and communication in CVEs. We will investigate different kinds of awareness which need to be carefully designed. We will also discuss different communication means and how to cope with this diversity, so we can benefit greatly from the availability of different peripheral devices and can find an effective communication means to work together.

The remainder of this paper is structured as follows. The next section describes in more detail several kinds of awareness, followed by a small discussion on how to design a CVE for improving the awareness. Different means of communication are briefly presented in the third section. The paper concludes with a discussion on the actual limitations of existing awareness and communication work and some suggestions of different aspects to develop in order to obtain a better collaboration in CVEs.

## 2 AWARENESS IN COLLABORATIVE VIRTUAL ENVIRONMENTS

The awareness concept in a CVE as it is defined in [8] mainly concerns the activities of other users and their presence. This concept summarizes the idea of the spatial model of interaction [3] where Benford et al. define some properties of space such as medium, aura, focus, and nimbus representations as the basis for mediating interaction. This model can determine when, where and how the interactions happen. Therefore, this model can help a user to manage his/her awareness about the presence, whereabouts, attention and activities of others. Moreover, he/she also can manage his/her own level of giving attention so he/she can predict at which level the others can be aware of him/her. However, the awareness concept need to broaden in order to help a group of users be able to think and work together in a more complex CVE.

The collaboration in virtual environments does not simply start when users begin working together but it includes early activities such as using environmental cues to establish a common understanding, knowing who is around and what they are doing, knowing the state of artifacts in the shared workspace, observing other people's gestures and what they are referring to, etc. Therefore, awareness is defined as having complete knowledge of the environment within which a user is working and of other people he/she is working with. Awareness can help users reduce effort, increase efficiency, and reduce errors for the activities of collaboration. Gutwin et al. [17] states that awareness is knowledge about the state of a particular environment and because environments change over

\*e-mail: thi-thuong-huyen.nguyen@inria.fr

†e-mail: thierry.duval@irisa.fr

time, people must maintain and update their awareness by interacting with the environment.

We will describe in more detail in this section several kinds of awareness which need to be carefully designed including the awareness of the virtual environment and its artifacts, of people in it, of the ongoing collaborative task, of the limitations and constraints of the physical environment, and of network delays.

## 2.1 Awareness of Others

The first and most important factor of the user's awareness in a CVE is whom he/she is working with. All the information about others' presence, their location and their actions and activities needs to be well considered and represented. This awareness involves identifying and locating other users at a particular moment, so it concerns mainly synchronous collaboration.

Awareness of presence is the knowledge about whether there are other users in the same virtual environment and who they are. This awareness of presence can be easily done by representing each user as a recognizable embodiment or avatar. The DIVE system [4] uses simple graphical 3D-bodies to show where is everyone in the virtual world. However, the awareness of presence and then later of identity and of actions can only be revealed when there is a collision between auras which are implemented as a volume around graphical 3D-bodies. These aura representations allow the system to calculate whether there is a collision happening and if there is, the identity as well as position and angles of the related entities are sent to other processes within the environment. However, the calculation time can increase and become a burden when there are many objects and users in the environments.

Awareness of location covers the knowledge of gaze, view and reach area of other users. Awareness of gaze and view involve understanding the area where a person is looking at and what he/she can see. Awareness of reach characterizes the knowledge about the area within a person can interact or manipulate artifacts [17]. Awareness of presence also implies the awareness of location if these users are working in the same limited "space": "I know you are there because I can see you". Awareness of location in a narrow space can be improved if the working conditions of each user are explicitly represented. For example, by attaching a "wire-framed" view frustum [15] to a user's avatar, others can see the orientation of his/her face and his/her field of view. They can also see where he/she is looking at and which artifacts are possibly seen by him/her. The same idea can be applied to the representation of interaction or manipulation area to improve the awareness of reach. Based on the limitations and the constraints of control devices (e.g., mouse, flystick, 6 DOF controller, etc.), the interaction area of each user is represented, so the others can understand what he/she can possibly do. However, a problem may arise when using view frustums as well as explicitly specifying interaction areas to improve the awareness of location. When there are many users working in the same narrow space, the confusion of these "wire-framed" representations can happen. When users work together in a large-scale CVE, the design to improve the awareness of presence and of location is different. The characteristic of this kind of collaborative task does not require the permanent presence of users in the same place or in the field of view. A user can observe others over a map and their small 3D representations in a world in miniature [28] or in a supplementary view called "bird's-eye view" [9]. This view can provide information about the presence, the position and may be about the activities of the others. However, this information may be incomplete, especially for the current actions of the others, because detail is lost and it makes the perception particularly difficult.

Awareness of actions and activities implies the understanding of what others are doing, either in detail or at a general level. It also includes the knowledge about their intentions and what artifact they are working on [17]. In a limited working space, a simple idea of

how to represent the others' actions and activities is to show them on the avatar of each user. The more detail in pseudo-humanoid representations there is, the easier it is for others to understand what a user is doing. Depending on the goal of collaborative work, if the users are working together to manipulate or interact with artifacts in the virtual world, a humanoid representing postures or gestures can perfectly show the current actions of each user and so the others can predict his/her intentions and which artifact he/she is working on. For other applications such as social meetings or teleconferences, the emotional expression and gestures may be important. If they are working in a large-scale CVE, in order to understand the actions of others, another solution can be found in [31]. In [22], Nguyen et al. provided "what you see is what I see" or "over-the-shoulder" perspective views so each user can have his/her independent view and also other shared perspectives. By using this technique, one can observe and know what the others are actually doing and can predict their actions and activities. A user can directly access to the current viewpoint of others and can share his/her own viewpoint. This technique is very useful to improve the awareness of actions and activities for collaborative tasks in large CVEs. The drawback of this approach is that multiple viewpoints can destroy the immersion and the continuity in the collaboration.

To summarize, awareness of other users is the first important step to increase the efficiency of collaborative tasks and reduce errors for collaboration activities. This kind of awareness can be obtained from different perception sources such as visual, auditory, embodiment ones. Therefore, a good design of different factors in the virtual world to improve the awareness of users is important to get closer to an effective collaborative work.

## 2.2 Awareness of the Virtual Environment

Being aware of the virtual environment in which people are working together is also an important step toward being able to efficiently work with others. In this paper, we define a virtual environment as an environment that includes all the 3D models, data, information, artifacts and the tools that users use to interact with the environment and with the others. In the literature, many research work have been devoted to improving the awareness of people and their presence in CVEs, but few were interested in improving the representations of virtual entities for facilitating the perception process and the awareness of the virtual world. When designing 3D models and representing the datasets or information of the virtual world, some features, representations or metaphors of these entities need to be taken into account to allow users to gain a great understanding of abstract concepts and of all possibilities to manipulate these representations [20].

The spatial model of interaction [3] is one of the earliest work which treated the behaviors of artifacts in the virtual world in the same way as behaviors of 3D embodiments. However, the behavior of artifacts and of 3D embodiments have different properties. A 3D embodiment which can possibly support the use of gesture and non-verbal cues combined with real-time conversation channels such as auditory and visual ones can represent the identity of a user and the information about his/her actual whereabouts, actions, activities, etc. [8]. In the virtual world, 3D embodiments can be considered as an active factor while artifact representations are passive ones. Artifact representations may be interpreted as a means to reflect the effect of users' actions onto the environment. By using many advantages of virtual reality, developers have many options to represent this effect on artifacts in the virtual world so the interaction of users with artifacts and also the collaboration between different users may become easier and much more effective. An example can be found in [15] where Fraser et al. used a 'wire-framed' representation of objects to show their current state of "being moved" differently than the state of "being pointed". This representation can help other users distinguish the difference between the grasp-

ing and the pointing actions of a user because both grasping and pointing actions are usually done by a similar gesture: extension of the embodiment “arms”.

Another aspect that needs to be studied is the way object representations or data may or may not remember the actions made on them and show them as a history metadata which can be used as an archive of past actions for later consultations. This aspect is mostly important in asynchronous collaborative work where not all the users connect and work at the same time. This metadata may contain the information about whether the artifact it represents can be manipulated or changed and how. It can also contain all the modifications and information about these modifications are made by whom, where and when. This capability of storing history actions of artifacts may be applied in educational applications of CVEs where course materials, learning contents or other auditory, visual, textual documents, etc. can be archived for all the learners as well as for each individual one.

Access control to artifact representations is also necessary. It defines the possibility of having many levels of access to artifacts depending on the authorization level of each user. Access control have been long studied in collaborative systems. In [29], Tolone et al. have summarized many different access control requirements for collaboration and these requirements can be applied in the same way in CVEs. In [8], Churchill et al. states that “one individual may require multiple representations to reflect different aspects of his/her task(s), whilst in other cases different individuals may require tailored representations to provide information specific to their tasks”. In the same view, developers need to determine which objects and which properties of them are available and visible to which kind of users. Accordingly, in order to improve the awareness of virtual world for users, the requirement of having different subjective views as well as alternative ones of objects can become necessary and indispensable depending on the nature of each collaborative task.

Information produced by artifacts, either by their internal state or by actions of users when they interact with them is one of the primary ways that people maintain workspace awareness [17]. If the interaction and manipulation process produces minimal feedback, the people’s ability to maintain awareness of virtual environment reduces. This reduction of awareness may become worse when many people try to interact with a same artifact or object. Getting distracted by people’s action, a user may not be aware of the change in the artifact’s state and so do not modify his/her actions accordingly. Therefore, effective feedback from artifacts is a key factor to improve the awareness of virtual world.

### 2.3 Awareness of Coordinating Actions

We want to discuss in this section a type of awareness that did not get much attention from the CVE developers even though it is very important for cooperative manipulation. It is the possibility to make users be aware of actual collaborative activities so they can work together with the coordinating actions occurring in the right order and at the right time to complete the task [17]. In [27], Robinson et al. states that there are two ways of coordinating actions in a shared workspace. The first solution is using explicit communication about how to do the task together. The second way is using shared materials used in the work process. The second way is more efficient but requires people to maintain the awareness of coordinating actions. This kind of awareness enhances the coordination of users because it informs them about the temporal and spatial boundaries of others’ actions, and it helps them fit the next action in the working process. The awareness of coordinating actions is particularly efficient in continuous actions when people are working with the same objects [17].

The cooperative manipulation is a type of collaborative work wherein users manipulate simultaneously an object. Because this

particular activity involves the actions of several users on an object at the same time, the awareness of the ongoing actions of the others’ takes an important role in the accomplishment of the collaborative work. Some approaches have been proposed for cooperative object manipulation but not many interaction techniques take into account this awareness aspect. The collaborative pointing technique for colocated multi-user interaction [26] allows users to pick up an object simultaneously using pick rays. By bending these pick rays which are normally straight, users can be aware of the change made by their actions on the object and also on the control metaphor (i.e. pick rays). Pinho et al. have proposed some visual metaphors to improve the awareness for cooperative manipulation [23]. In the selection phase, the selected object’s color changes when a user points to it. This feature allows another user to know what his/her partner is pointing to. The object’s color also changes when the users enter the other phases such as the attachment, positioning and release phases. Because this cooperative manipulation technique combines simultaneous user actions based on the separation of degrees of freedom between two users, the shape and color of the control metaphor (in this case, pointers) change accordingly to each particular manipulation situation which allows a user to predict the interactive capabilities of his/her partner.

In order to accomplish a successful cooperative manipulation task, each user has to maintain a continuous awareness of coordinating actions to track, predict and adapt his/her actions with other users. Therefore, some simple but effective metaphors need to be implemented to improve this kind of awareness.

### 2.4 Awareness of the Physical Environment

The awareness of the constraints and limitations of the physical world would be very important when a user is fully immersed within a CVE because his/her feel of presence of the real world can be totally replaced by the virtual world. This immersion aspect may cause serious collision with hardware parts of the system or more slightly the disturbance and misunderstanding between users during a collaborative work. We will specify how researchers cope with these constraints to make sure that users can be aware of their infrastructure limitations as well as of the others’.

As in most of the virtual reality applications, a physical environment of a collaborative system must be integrated into the virtual one so users can be aware of their own interaction capabilities as well as of the others’. Duval et al. [11] states that the representation of physical devices can describe the spatial relationships between these physical devices and model the users’ physical workspace associated to each device. A model for embedding the features of the physical world into the virtual world has been proposed in [14]. In this model, all the possible workspaces have been described including a motion workspace (the area where a user can move his/her body, e.g., in a tracking zone of a CAVE space, in a zone of maximum wire length of a HMD), a visual workspace (what the user can see through and around a display device), and an interaction workspace (the area where a user can interact by a control or an input device). Even though these workspaces may not be always visible to users, they help developers to implement their collaborative framework without worrying about the changes it could make to the virtual world when there is a new physical device added to the system. These workspaces also help precisely define the parameters of physical devices so they can appear in the virtual world and become a part of it. Two examples of how to use the physical representations in a virtual world to enhance the users’ awareness of their physical surroundings and to prevent the collision between users and the front display screen can be found in [11] and are illustrated in Figure 1. These representations are easy and simple to implement if the physical world limitations are taken into account when designing CVEs. Besides the visual representations, auditory signal can also be used as a warning of the collision between users



Figure 1: From left to right: In the tracking zone of a CAVE-like system, in order to warn a user not to hit the real wall, the motion workspace [14] is represented in the virtual world by a 3D grid which becomes clearer and sharper when the user goes closer to the display screen or his/her hand reaches out close to it.

and the physical world.

In addition to the recognition of their own physical world, the representation of physical devices can also help users to be aware of the working condition of others. They can predict the others' possible actions based on these representations. View frustum representation [15] is a simple way to show to other users about the limited field-of-view of a user and so the capability of his/her display screen or HMD (e.g., the field-of-view of a HMD is about 50 to 60 degrees and the computer display screen's one is about 90 degrees). In order to visually display a haptic device with a limited workspace in a virtual world in [10], a semi-transparent sphere which surrounds the manipulated cursor is used to control the force feedback based on the sphere position relatively to the cursor.

The ultimate goal is how to benefit greatly from the availability of different infrastructures of users to build an abstract representation of the virtual world. The semantic metaphors and the flexibility of the system to each user make sure that in spite of the difference, the metaphors can make up for the lack of devices or can make users be aware of this difference so they can find an effective way to work together. The constraints and limitations of the physical world normally stay invisible so the immersion of users can be guaranteed. The features of the physical world only need to be visible in case of collision or of explicit representation of the real world besides the virtual one.

## 2.5 Awareness of Inconsistency due to Network Delays

The problem from network delays is one of the main reasons of ineffective collaborations in CVEs. In [30], Vaghi et al. states that "the deployment of CVEs over wide area networks as well as different communication bandwidths increase typical network delays, potentially breaking the consistency between the replicated versions of an environment at the participants' sites". All the information about the activities of all the actual participants, their whereabouts and all the changes that are made to the data and objects in the virtual world need to be transferred over the network. Therefore, the more complicated the shared environment is and the more people there are in the same world, the more present the increasing network delays are and so is the incoherence of the participants' awareness and actions.

In order to solve this problem, the most evident solution is to reduce the amount of circulating updates and messages on the network [21]. However, if the network troubles still arise, the collaboration manager system has to make all users aware of this potential problem. Lamboray et al. [21] have proposed some system features that need to be applied so the network delay effect could be coped by the understanding of users in the virtual world. Depending on the level of trust that each system requires, each feature needs to be considered, so determining which one(s) will be implemented is a compromise between the required level of trust and the system performance. The features such as the prediction capabilities and

indication of level of trust associated with the prediction or the calculation of expression and preservation of user's expectations might not be necessary if the system's requirement of the consistency is not too strict.

We want to discuss here some features that may be easy to be implemented and applied in CVEs. These features include the capability of explicit indication of delays with all their characteristics and their effect on the influenced objects or data; and the capability of explicit display of corrections due to discontinuities. Duval et al. present an example of explicit indication of network delays that can be found in a collaborative system [12]. All the possible delays or disconnection due to low-level network problems are detected by sending synchronization messages between different sites in order to coordinate these parallel processes to similarly evolve in each system. If there is a breaking of the real time concept, they choose to let the collaborative task continue by freezing only the parts of the world whose state is uncertain for consistency considerations. The remotely shared objects lose their interactivity as long as the disconnection remains. They use echo objects that represent the state of their associated distant objects and a marker system to inform users about the out-of-date state of shared objects. The problem of this approach is that when a large number of sites are participating to the same collaborative session, there will be many echo objects which maybe "pollute" the visualization. The solution is then to make sure that there is no more than significant metaphors which are necessary for indicating the participants about the state of the world and of the other problems concerning network delays. One metaphor can also be used for different indications in different contexts so the redundancy will reduce. Another approach is to create dynamically echoes after the detection of a network problem because the echo objects are not necessarily explicit or available when there is no network problem to be shown. The users are informed about technical problems only when they obviously influence the users' activities. In addition, these network problems can be represented in form of visual, auditory metaphors or other metaphors in the virtual world so their immersion is not interrupted.

Developers of CVEs need to find a compromise between the need of representing the network problems because of the unacceptable inconsistency in a collaborative work and the discontinuity and the perturbation it could cause when the goal of almost all VR applications is to totally immerse users in the virtual world.

## 2.6 Summary

In collaborative virtual systems, the awareness of a user is the up-to-the-moment understanding of the collaborative task to achieve, of the people he/she is working with, of the shared environment, of the system infrastructure [16]. Because acquiring awareness information by perception process is obvious, simple and often happens inside users' head, the design of a CVE for improving the awareness is often not considered enough. However, so much information

of the virtual world as well as of the physical one does not really help the user to collaborate successfully with others due to information overload. Additionally, the perturbation of information can happen and obscure the main goal of collaborative work, reducing its efficiency and so compromising the collaboration.

### 3 COMMUNICATION IN COLLABORATIVE VIRTUAL ENVIRONMENTS

As mentioned in section 1, the communication in CVEs is an important “instrument” to complete the collaborative work. Communication means provide for a user a possibility of keeping in contact with other people, either communicating in real time in synchronous CVEs or leaving notes for others wherever they are needed. Users need to negotiate shared understandings of task goals, of task decomposition and sub-task allocation, and of task progress. It is important that users can be aware of what is currently being done and what has been done in context of the task goals. In addition, the communication between users can be explicit (e.g., speaking, writing) and implicit (e.g., emotional expressions, gestures, postures, visual metaphors).

However, the more there are users who participate in the collaborative work, the more various is the peripheral infrastructure. Each user’s site may have auditory devices (e.g., headphones, speakers), visual systems (e.g., display screens, head-mounted displays, CAVE-like systems, handheld devices), and input devices (e.g., mouse, keyboard, joystick, flystick, motion detectors). The MASSIVE system [3] is an example of collaborative virtual systems which supports interaction between users whose equipment has different capabilities. MASSIVE uses graphics embodiment which have different representations: an audio user has ears, a non-immersive user has a single eye and a text user has the letter “T” embossed on his/her head. These simple visual metaphors allow users to know how to communicate with others without wasting time to try each communication channel until finding one in common.

In order to obtain an effective common communication channel, developers of CVE frameworks need to find a general representation of these peripheral devices as well as their limitations, so users can be aware of their difference of communication means and can choose the right channel to use for communication. We will discuss in this section some major means of communication to understand their capabilities and their limitations to be used in CVEs. They are the audio communication, embodiment and nonverbal communication, text and 3D annotation, and some visual metaphors using for communication and exchange information.

#### 3.1 Audio Communication

Talk is one of the first and most important means of communication in virtual world besides the visual channel [6], especially for social meetings and teleconferences. But is talk solely enough for an effective communication in CVEs? What happens when there is a network delay that influences the quality of the audio channel? If the network delays happen, silence in the audio channel could be interpreted differently by each user because of some confusion over whether the origin of the silence is a technical failure or simply the intended silence in talk [6].

The simple approach to deal with this problem is to add a visual channel and visual metaphors to cope with some problems of consistency in the audio communication of CVEs. This approach allows users to get informed about who is talking or when there is a turn-talking, especially in social meeting or educational applications where the number of users in a shared virtual world becomes important. The whiteboard, conference table, podium entities [4] are some typical visual metaphors which allow users to organize the turn-talking, to get attention to the one who is talking as well as to determine who are listeners.

Audio communication becomes a fast and effective means of communication if it is combined with another communication means or with virtual representations. An avatar for each user with the capabilities of emotional expression and/or gestures and/or postures can be a valuable factor which helps users have a natural communication as it is in the real life.

In addition, the audio media can improve the interactions in the virtual world. When the 3D audio signal is rendered according on the position of users in the virtual world by a tracking system, their feeling of presence and immersion in the environment become more real. This possibility of 3D audio rendering can be done locally and it depends on the apparatus of each user.

Another advantage of the audio channel is that we can use this means not only to communicate to each other, but also to control or manipulate virtual elements or artifacts in the virtual world by using a speech recognition system. This advantage can change the user interfaces designs for CVEs deeply if the speech recognition system is efficient.

Another aspect that needs to be considered is to make the audio channel become a representation or an entity in CVEs. Until now, audio communication via Skype or telephone is normally used as a complementary means of communication, and it does not truly become an element or a factor in the virtual world that can affect the other means of communication or can be replaced by them when some problems happen. Therefore, the developers of CVE platforms mainly do not take into account the representation of this communication means. In [6], the virtual representation of talk presence is a “mouth” which opens when a user’s speech exceeds a certain amplitude threshold. As far as we know, there is still no richer and more advantaged representation of audio media in the virtual world. So we can imagine that the different nuances in speaking could be dynamically and automatically interpreted in gestures or emotional expressions or even other visual metaphors so the one without audio channel can globally understand what others are talking about.

#### 3.2 Embodiment and Nonverbal Communication

Natural human communication is based on speech, facial expressions, body postures and gestures. In order to obtain a virtual and yet still effectual and natural representation of users in the virtual world, humanoid embodiments or avatars are used as a medium to improve social interactions and communications in non-verbal aspect. Nonverbal communication signals include gaze, gestures and postures, facial expressions, touch, etc. as well as paralinguage cues such as variations in intonation and voice quality. These nonverbal cues help interlocutors express more feelings or thoughts through the use of their bodies and augment spoken messages [13, 18]. Moreover, besides improving the bodily communication and the natural interaction, a visual embodiment is a means to self represent, to interact with the world and to experience various properties in it. According to [13], the avatar can provide direct representation as well as feedback about the actions, the direction and degree of attention, and the interactive abilities of one specific user to the others at all times. By gaining the understanding and knowledge of their whereabouts and activities, users can develop a strong mutual awareness of the others. However, Fabri et al. [13] argue that the requirement for the rich representation of users in the virtual world does not necessarily imply the using of realistic avatar. The graphical representation of users in the DIVE system [3] is very simple and allows the system to detect the possible interactions between users by using the aura, focus and nimbus representations. This representation plays a role as a “placeholder” of users in space so it does not require a complicated model. The same thing goes for “blockie” models in the MASSIVE system [6] when Bowers et al. only wanted to represent some simple gestures such as “sleeping” state when the user it represents has left his/her machine, “flapping

ears” or “opening mouth” when the user is listening or speaking. These gestures only inform the presence or the actual action of the user and not the emotional expressions or the other complicated behaviors. In order to make the avatars become more realistically emotional, Guye et al. [18] have proposed an interface of different built-in gestures, postures and expressions. The problem of this approach is that users need more time to choose the adequate emotional expressions by panel and this interferes with the continuity of actions in VE just for choosing the right expressions to use. This is also not a natural interaction in CVEs, specifically from the immersion perspective. On the contrary, in [2], a CVE can render a chosen subset of nonverbal behaviors, filter or amplify that subset of behaviors, or even render nonverbal ones that the user may not have performed. In order to obtain a complicated subset of nonverbal behaviors, nonverbal signals (e.g., eye gaze, facial gestures, body gestures) need to be tracked and then rendered via realistic avatars. The choice of using which kinds of avatar with all the possibilities to express the emotions and feelings is made based on the required social interaction levels in the CVE. However, gestures and postures don’t need to be too complicated: for example the smiley icons on a chat messenger are simple but effective to express feelings without slowing down the system and the network transmission for complex graphical avatar and movements. However, if the interaction focuses on social behaviors, the complex graphical avatars can improve the sense of reality and immersion. And on the other hand, if the collaborative work focuses on “task-focus” interaction or on a specific collaborative task, the feeling and emotional expression will become less important. It also needs to allow users to plan on doing things, a means to show their actions, etc.

### 3.3 Visual Metaphors

Visual metaphors for communication are the most important tool of CVEs. 3D virtual environments are first and foremost the world where different abstract or realistic data and models can be created and visualized. By using simple visual metaphors as a direct means of communication in CVEs, users can establish an implicit channel of communication which is easy understand. Visual metaphors also help users get an immediate and direct visual feedback. In addition, not only visual metaphors are used for communication purpose, changing parameters of objects or data in the environments such as color, position, orientation, scale, etc. also provides opportunities for users to express their actions and activities to others. The examples in Figure 2 and [1, 22] illustrate the approach of using some simple visual metaphors for a mutual awareness and also a collaborative exploration without using verbal communication. A user can also use some colorful signals to inform others of his/her current states while navigating together (e.g., orange signal for “waiting”, red for “stopping”, green for “going”).

The drawback of this communication approach is that all the users need to get a common understanding about the regulations or the signification of visual metaphors before they start working together.

### 3.4 Text and 3D Annotation

Although the importance of providing text input and annotation capabilities in virtual interfaces has been recognized, not many research work have been devoted to integrating and manipulating textual data in immersive VEs [5, 19]. Text and 3D annotation can become powerful tools for communication in a CVE, especially for data visualization and collaborative analysis applications, if the ergonomics aspect for writing or taking notes is improved and if these tools do not degrade the immersion of users. Another advantage of text and 3D annotations is that they are easy to store offline so next time users can access to the same data. This aspect can facilitate asynchronous collaborations between distant users.

Ribarsky et al. [25] studied how to use text to annotate graphic

models in VEs. Notes might take the form of digitized speech or of written text, depending upon the user’s needs and preferences. Some textual metaphors such as text or annotation which are texture-mapped onto the walls of a room or of the billboard can become a communication tool to represent information to many users at the same time like in a meeting room. In CVEs, these notes should exist in a visualization space and be directly associated with the part of the visualization to which they apply so the other users can see, access and modify them. Poupyrev et al. [24] have proposed a collection of interface tools that allow users to take notes, annotate documents and input text using a pressured-sensitive graphics tablet, a pen and a handwriting recognition software while still immersed in VEs. This handwriting approach may become a new modality for interaction and communication in immersive CVEs. But using too much devices just for taking notes would “spoil” the immersion of users and the flexibility of VR interfaces. However, with the development of hand-held devices, this tool can be integrated in a tablet or a smart-phone so users can use it as a supplementary device to explore and work in CVEs. Additionally, the complexity of input devices to produce text in the virtual world can be avoidable by using some graphic metaphors such as a 3D virtual paintbrush or a virtual pen controlled by normal input devices to draw and write without any constraints.

A virtual annotation system by voice input has also been studied as a possible method for the immersive input of text [24]. Voice annotations are represented as a small marker attached to objects, and can be selected later for playback. However, a speech recognition application need to be used to recognize the audio input and then translate it into text for manipulation and editing purpose. This is the main reason why the voice annotations are easy to create but difficult to use.

Some problems arise when using text and 3D annotation in a virtual environment concerning the limitations of display devices and the integration of text and annotations in the virtual world. A low resolution and a small field-of-view of head-mounted displays normally make it difficult for users to read text in the environment. In a CAVE-like system, the field-of-view of the users is expanded as well as the physical motion space. This is the reason why if the system wants to get users to notice the text and annotations in the environment, they need to be attached to the users’ position or in their working area. So the integration of these elements need to be designed so they are not disturbing the visualization of users and yet not lost from their sight zone.

Besides the utilization of text and 3D annotations for annotating data, models in the virtual world as well as for communicating between users, synchronously or asynchronously, many CVE platforms integrate the text on menus or pop-up menus to provide for users more controls or options to choose. These menus normally float in the virtual world and can be visually customized if users have control over the level of detail in the visualization and over the kinds of tools employed. This customization process can help users limit graphical structure and means of interaction to retain immersion while still looking on details they think are important. But the problem with pop-up menus in a CVE is the access and display of these menus need to be local so participants do not access to the menus of others and vice versa. On the other hand, we need some metaphors to represent the actual action of selection and controlling over the menus of one participant to the others so they can know what is really going on, especially for the synchronous collaboration in the same shared space.

### 3.5 Summary

We have studied different means of communication, which CVEs normally support and exploit. Besides these kinds of communication means, there are still other communication media such as videoconferencing, tactile or haptic feedback, etc. which could

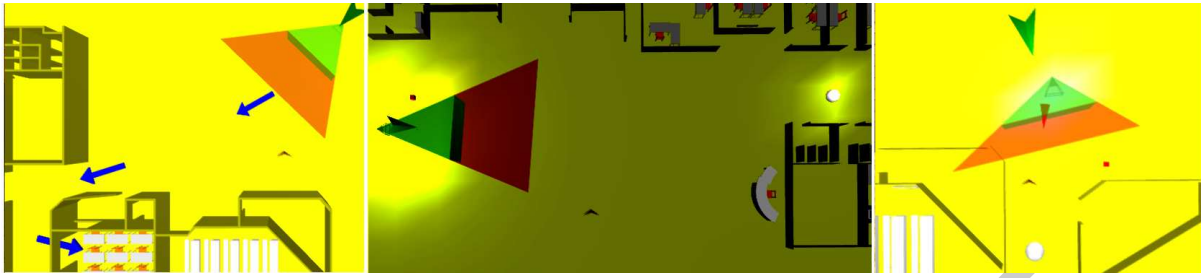


Figure 2: Directional arrows (left), beacon (middle) and compass (right) as navigation metaphors in a collaborative exploration task [22].

help users send some intuitive and effective messages or feedback. In order to cope with the diversity in the peripheral infrastructures of different users' sites, CVE frameworks should at least support and integrate these different kinds of communication means into the system by different representations. The choice of which communication channel to be used would be automatically proposed by frameworks. This choice should be based on the common available means of communication at different interlocutors' sites. It also should be based on the range of communication: it is either a face-to-face communication or a situation where many interlocutors are involved in the same communication process. The problem of data overloading on the network due to the communication signals such as video or audio data might happen. In this case, an alternative communication channel would be proposed. Therefore, a communication manager in CVEs would be necessary for managing and recognizing these possible means of communication, for proposing an appropriate channel of communication for interlocutors without degrading their immersion and their workflow continuity.

## 4 DISCUSSION

We will discuss in this section some existing techniques in CVEs which have been developed to deal with the awareness improvement and the efficiency of communication means in collaboration from a global perspective.

### 4.1 Awareness

We have detailed in section 2 different kinds of awareness that users need to obtain when they are working together in CVEs. We have also developed some requirements for each kind of awareness and how they have been applied in existing CVEs. Up until now, to our best knowledge, there are few developers of CVEs who have given much thoughts to awareness aspect while designing collaborative virtual worlds.

The awareness cues do not need to be too complicated to give a "hint" about the state of the workspace or of people in it. Awareness of others is a key factor of the collaboration between users. That is the reason why many developers have integrated embodiments or avatars into their systems to represent the identity, presence, location and activities of each user in the virtual world. These representations vary from simple "blockies" [6], view frustums [15], to complex 3D avatars for a high level of realism [18]. Artifacts in the virtual world can have some special behaviors such as changing the color or shape [23] to inform users about their change of state. Physical constraints and technical problems can be informed to users when some problems might happen in forms of visual metaphors [11] or audio signal for warning.

All these awareness cues can be simply integrated in CVEs to improve the perception process and so to easily get a common understanding between users. However, it is not necessary to make these awareness cues visible all the time to the users due to the perturbation of too much information to process. The awareness knowledge of the artifacts and people in the virtual world must be kept up to

date. The awareness of the physical world and network delays can only be revealed to users when there is a potential collision between users and the physical world or when the network delays may influence the consistency of the virtual world. Most of the time, this kind of awareness should remain invisible. In addition, depending on the nature of collaborative work, some awareness cues may not need to be present. For example, when users are working in a large-scale virtual environment, the awareness of gesture, of posture, or of gaze might not be necessary.

### 4.2 Communication

As mentioned above in section 3.5, with the development of communication media technology, CVE frameworks need to support a multi-media infrastructure so these communication means can be integrated into the virtual world, becoming virtual entities in the environment. Different communication means can be represented on the same virtual entity. For example, a humanoid avatar can "open its mouth" to speak, to move its head or arms to express some gestures or move its body to show different postures. This representation can improve the natural communication in the virtual world and at the same time guarantee the immersion of users.

The flexibility and the dynamic aspects of communication channels need also to be investigated. For example, a user A can only communicate with a user B by talk while a user C can only use the nonverbal means to communicate with user B because of the difference in communication means at users' site. In this case, to make all the three users have a common understanding and agreement, normally user B has to become a communication "bridge" between A and C. The problem of how to transfer information on different communication channels in CVEs still remains unsolved.

Additionally, to our best knowledge, the verbal communication is the sole means to get agreements for all the users in the collaboration in order to make sure that they have a common understanding about the task. This kind of communication channel might block all the other steps of a collaborative task if the system can not find a common verbal communication channel for all the participants. That is the reason why we should investigate an alternative communication means which is effective, simple, direct, quick. This communication means might not require too much control because the communication step normally goes on with the performing step.

## 5 CONCLUSION AND PERSPECTIVES

We have studied different features and requirements of awareness and communication in Collaborative Virtual Environments (CVEs). This features and requirements could help CVE developers to design an effective collaborative framework with a natural interaction; to cope with the diversity of peripheral infrastructures which causes the inconsistency of the collaboration in terms of no common communication means can be found, or network delays; to integrate all the physical entities and human representations into the virtual world so the collaborative task becomes independent of the

infrastructure layers of the system; to surpass the misunderstanding caused by a lack of awareness cues in the environment, by the inefficiency of communication channel; and so on.

In order to be able to meet these requirements, further investigations about the improvement of awareness (e.g., integrating all the important awareness cues into CVEs) and of communication (e.g., supporting a diversity of peripheral infrastructures, integrating communication means into the virtual world, investigating the way of communication when there is an asymmetric communication means of different users, etc.) are still necessary and urgent.

## REFERENCES

- [1] F. Bacim, E. Ragan, C. Stinson, S. Scerbo, and D. Bowman. Collaborative navigation in virtual search and rescue. In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI 2012, pages 187–188, Orange County, California, USA, 2012.
- [2] J. N. Bailenson, A. C. Beall, J. Loomis, J. Blascovich, and M. Turk. Transformed social interaction: decoupling representation from behavior and form in collaborative virtual environments. *PRESENCE: Teleoperators and Virtual Environments*, 13(4):428–441, Aug 2004.
- [3] S. Benford, J. Bowers, L. E. Fahlén, and C. Greenhalgh. Managing mutual awareness in collaborative virtual environments. In *Proceedings of the Conference on Virtual Reality Software and Technology*, VRST 1994, pages 223–236, Singapore, Singapore, 1994.
- [4] S. Benford and L. Fahlén. A spatial model of interaction in large virtual environments. In *Proceedings of the European Conference on Computer-Supported Cooperative Work*, ECSCW 1993, pages 109–124, Milan, Italy, 1993.
- [5] J. Bolter, L. Hodges, T. Meyer, and A. Nichols. Integrating perceptual and symbolic information in vr. *Computer Graphics and Applications*, IEEE, 15(4):8–11, Jul 1995.
- [6] J. Bowers, J. Pycock, and J. O'Brien. Talk and embodiment in collaborative virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI 1996, pages 58–65, Vancouver, British Columbia, Canada, 1996.
- [7] S. Card and T. Moran. The model human processor. *Ariel*, 192:50–50.
- [8] E. Churchill and D. Snowdon. Collaborative virtual environments: An introductory review of issues and systems. *Virtual Reality*, 3(1):3–15, 1998.
- [9] T. J. Dodds and R. A. Ruddle. Using teleporting, awareness and multiple views to improve teamwork in collaborative virtual environments. In *Proceedings of the Eurographics Conference on Virtual Environments*, EGVE 2008, pages 81–88, Eindhoven, The Netherlands, 2008.
- [10] L. Dominjon, A. Lecuyer, J.-M. Burkhardt, G. Andrade-Barroso, and S. Richir. The “bubble” technique: interacting with large virtual environments using haptic devices with limited workspace. In *Proceedings of the Eurohaptics Conference and the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, WorldHaptics 2005, pages 639–640, Pisa, Italy, 2005.
- [11] T. Duval, T. Nguyen, C. Fleury, A. Chauffaut, G. Dumont, and V. Gouranton. Improving awareness for 3d virtual collaboration by embedding the features of users physical environments and by augmenting interaction tools with cognitive feedback cues. *Journal on Multimodal User Interfaces*, pages 1–11, 2013.
- [12] T. Duval and C. Zammara. Managing network troubles while interacting within collaborative virtual environments. In *Proceedings of the International Workshop on Computer Supported Activity Coordination*, CSAC 2006, pages 85–94, Paphos, Chypre, 2006.
- [13] M. Fabri, D. Moore, and D. Hobbs. The emotional avatar: Non-verbal communication between inhabitants of collaborative virtual environments. In A. Braffort, R. Gherbi, S. Gibet, D. Teil, and J. Richardson, editors, *Gesture-Based Communication in Human-Computer Interaction*, volume 1739 of *Lecture Notes in Computer Science*, pages 269–273. Springer Berlin Heidelberg, 1999.
- [14] C. Fleury, A. Chauffaut, T. Duval, V. Gouranton, and B. Arnaldi. A generic model for embedding users’ physical workspaces into multi-scale collaborative virtual environments. In *Proceedings of the International Conference on Artificial Reality and Telexistence*, ICAT 2010, pages 1–8, Adelaide, Australia, 2010.
- [15] M. Fraser, S. Benford, J. Hindmarsh, and C. Heath. Supporting awareness and interaction through collaborative virtual interfaces. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, UIST 1999, pages 27–36, Asheville, North Carolina, USA, 1999.
- [16] C. Gutwin and S. Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):411–446, 2002.
- [17] C. Gutwin and S. Greenberg. The importance of awareness for team cognition in distributed collaboration. pages 177–201, 2004.
- [18] A. Guye-Vuillme, T. Capin, S. Pandzic, N. Thalmann, and D. Thalmann. Nonverbal communication interface for collaborative virtual environments. *Virtual Reality*, 4(1):49–59, 1999.
- [19] R. Harmon, W. Patterson, W. Ribarsky, and J. Bolter. The virtual annotation system. In *Proceedings of the IEEE Symposium on Virtual Reality*, VR 1996, pages 239–245, Santa Clara, CA, USA, 1996.
- [20] R. L. Jackson and E. Fagan. Collaboration and learning within immersive virtual reality. In *Proceedings of the ACM Conference on Collaborative Virtual Environments*, CVE 2000, pages 83–92, San Francisco, California, USA, 2000.
- [21] E. Lamboray, S. Wurmlin, and M. Gross. Data streaming in telepresence environments. *IEEE Transactions on Visualization and Computer Graphics*, 11(6):637–648, Nov. 2005.
- [22] T. T. H. Nguyen, T. Duval, and C. Fleury. Guiding techniques for collaborative exploration in multi-scale shared virtual environments. In *Proceedings of the International Conference on Computer Graphics Theory and Applications*, GRAPP 2013, pages 327–336, Barcelona, Spain, 2013.
- [23] M. Pinho, D. Bowman, and C. Freitas. Cooperative object manipulation in collaborative virtual environments. *Journal of the Brazilian Computer Society*, 14(2):53–67, 2008.
- [24] I. Poupyrev, N. Tomokazu, and S. Weghorst. Virtual notepad: handwriting in immersive vr. In *Proceedings of the IEEE Symposium on Virtual Reality*, VR 1998, pages 126–132, Atlanta, GA, USA, 1998.
- [25] W. Ribarsky, J. Bolter, A. Op den Bosch, and R. van Teylingen. Visualization and analysis using virtual reality. *Computer Graphics and Applications*, IEEE, 14(1):10–12, Jan 1994.
- [26] K. Riege, T. Holtkämper, G. Wesche, and B. Fröhlich. The bent pick ray: An extended pointing technique for multi-user interaction. In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI 2006, pages 62–65, Alexandria, Virginia, USA, 2006.
- [27] M. Robinson. Computer-supported cooperative work: Cases and concepts. In *Readings in GroupWare and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration*, pages 29–49. Morgan Kaufmann Publishers Inc., 1993.
- [28] R. Stoakley, M. J. Conway, and R. Pausch. Virtual reality on a wim: Interactive worlds in miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI 1995, pages 265–272, Denver, Colorado, USA, 1995.
- [29] W. Tolone, G.-J. Ahn, T. Pai, and S.-P. Hong. Access control in collaborative systems. *ACM Comput. Surv.*, 37(1):29–41, Mar. 2005.
- [30] I. Vaghi, C. Greenhalgh, and S. Benford. Coping with inconsistency due to network delays in collaborative virtual environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST 1999, pages 42–49, London, United Kingdom, 1999.
- [31] U. Wössner, J. P. Schulze, S. P. Walz, and U. Lang. Evaluation of a collaborative volume rendering application in a distributed virtual environment. In *Proceedings of the Eurographics Workshop on Virtual Environments*, EWVE 2002, pages 113–122, Barcelona, Spain, 2002.