

An interaction abstraction model for seamless avatar exchange in CVET

Rozenn Bouville Berthelot, Thomas Lopez, Florian Nouviale, Valérie Gouranton, Bruno Arnaldi

▶ **To cite this version:**

Rozenn Bouville Berthelot, Thomas Lopez, Florian Nouviale, Valérie Gouranton, Bruno Arnaldi. An interaction abstraction model for seamless avatar exchange in CVET. 2014 International Workshop on Collaborative Virtual Environments (3DCVE), IEEE VR, Mar 2014, Minneapolis, United States. hal-01018825

HAL Id: hal-01018825
https://inria.hal.science/hal-01018825

Submitted on 7 Jul 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.

An interaction abstraction model for seamless avatar exchange in CVET

Rozenn BOUVILLE, Thomas LOPEZ, Florian NOUVIALE, Valérie GOURANTON and Bruno ARNALDI
IRISA/INSA de Rennes

ABSTRACT

Collaboration and interaction between users and virtual humans in virtual environments is a crucial challenge, notably for Collaborative Virtual Environments for Training (CVET). A training procedure, indeed, often involves several actors: trainees, teammates and many times a trainer. Yet, a major benefit of CVET is to propose to users to be trained even if the required number of person needed by the procedure is not available. Therefore, almost every CVET use autonomous virtual humans to replace the missing person. In this paper, we present the main results of our project that aims at improving the effective collaboration between users and virtual humans involved in a complex task within CVET. Using an entity called the "*Shell*", we are able to wrap the features common to both users and virtual humans. It gives us an abstraction level to pool the management of the main processes useful to control an avatar, interact with the environment and gather knowledge from a CVET. Besides, the Shell allows seamless exchange of avatars during a procedure. Thanks to the Shell, the exchange can be carried out at any time during a task while preserving all the data associated to a role in a procedure.

Index Terms: [Computing methodologies]: Virtual reality—
[Software and its engineering]: Virtual worlds training simulations—

1 INTRODUCTION

Training in real life hardly ever involve a single person. Indeed, a trainee is helped by a trainer and assisted by teammates in case of a teamwork training. Therefore, virtual humans are an essential feature for CVET. The challenging point with virtual humans in CVET is how to make users and virtual humans collaborate and interact as naturally as possible. Our objective is to provide a CVET that proposes intuitive and seamless ways to collaborate and interact with autonomous virtual humans. Besides we intend to enhance the collaboration between trainers and trainees to make training sessions more efficient and more interactive. Finally, depending on the requirements of the training, a variety of hardware devices is available for a user to interact with a virtual environment. Thus, we think that the solution we propose must run on many set-ups.

Hereafter, the term *actor* refers to both a user and a virtual human. During a CVET session, users are learning a team procedure. To complete the collaborative task, the actors involved in the training must be aware of the role of each teammate and understand what they are doing at any moment of the procedure. Thus, it is essential to them to have access to the knowledge acquired in the course of the procedure. Moreover, in order to synchronize their tasks, they need to exchange information. Besides, it can also be interesting for actors to exchange the control of their embodiments in order to change their role and access to the knowledge associated to this new role. Based on these observations, we believe that CVET must provide a common interface to both type of actors in order to ease the interaction, the knowledge management and enable exchange.



Figure 1: A sample CVET to learn how to operate a car.

To further illustrate this claim let us consider a scenario in the industrial training field. For example, trainees that are learning how to operate a car as shown in Fig. 1. During collaborative tasks, replacing the engine from the car for instance, at least two teammates must synchronize their tasks to achieve the collaborative goal. In real-life situation, the teammates would use dialog to share information and advices. By providing a common interface for both type of actors, the communication of information and even dialog is possible between actors whatever they are controlled by a user or by a virtual human. Otherwise, during real-life training, the trainer is able to replace one trainee to help him/her or to show him/her the correct manner to proceed. By allowing actors to exchange their roles, this situation can be simulated in the CVET.

In the remainder of this paper, we present in section 2, a literature review on collaboration between virtual humans and users in CVET as well as related work on exchange of avatars. In section 3, we introduce our interaction abstraction model allowing avatar exchange called the Shell then in section 4, we explain our exchange protocol based on the Shell. Section 5 depicts our implementation of the model of the Shell, the different set-ups used to run our sample application and some observations on the distraction of the user. The following section, section 6, explains the improvement for collaboration induced by the *Shell*. Finally, the section 7 concludes our contribution.

2 RELATED WORK

In this section, we will first present how virtual humans for CVET are represented in the literature and how users interact and collaborate with them. We will then discuss the exchange of avatars in CVET and how it improves the collaboration and the interaction between users and virtual humans.

2.1 Virtual Humans in CVET

In every CVET, the purpose of virtual humans is to interact and collaborate with users or with other virtual humans in order to complete a task. Nevertheless, they do not always have the same role. In the CVET literature, three primary roles assumed by autonomous humans have been identified [31] :

- as a *personal assistant* assigned to a single trainee to take care of basic tasks and help him,
- as a *team assistant* assuring the communication between users of the simulation, helping them to coordinate their actions and to better focus on the tasks requirements,

- as an *equal team member* operating as an autonomous entity that performs a collaborative procedure alongside users and other virtual humans.

Besides, whatever their role, they shared similar attributes: a knowledge base containing information on the procedure, a control module to interact within the world and, in some cases, a decision making module.

In order to represent virtual humans, different models have been proposed. Some proposes to split their representations in two parts: the *body* and the *mind* [10, 11]. The *body* can act within the world and perceives events whereas the *mind* can take decisions and influences the emotive state of the actor. Here, the *mind* part does not handle the knowledge base associated to the actor. These kinds of models were exclusively designed for virtual humans, they are not shared by users.

On the contrary, in the Interactive Storytelling field, a unified interface for both real and virtual actors is often proposed [14, 15]. An actor can thus be controlled by either a human or by the system. Here, a memory concerning the actor's part and its perceptions is attached to the actor. It allows directing the actor or to help a user to better fit its part. In this model, actors are considered omniscient and the memory associated to an actor is only used when the actor is controlled by a virtual human. In case a user is controlling the actor, this memory is not accessible but used by the system to supply the pedagogy, giving pieces of advice to better stick to his/her role. However, even if the interface is shared by both real and virtual actors, this method does not propose to exchange the control of a character during the simulation.

For the purpose of our project, we focus on the situation where virtual humans interact with other actors in CVET as full-fledged team members and take part in the procedure i.e they are equal team members. This is the most complex and challenging case. They are able to perform tasks, interact with the objects in the environment and, to some extent, communicate with other teammates. In most CVET, virtual humans are able to perform their task independently [1, 24]. They are generally able to play different roles such as collaborators, instructors or assistants that help the trainee. They can also replace missing team members needed for a training. In the collaborative version of STEVE [25, 26], virtual humans play the double role of collaborator and instructor. They play their part in the procedure, but a specific human is assigned as tutor to help the trainee. In the fire training simulation of Dugdale [7], a group of trainees deals with a fire situation in a hotel. Various characters interact in the training and each one of them can be controlled by either a real user or an autonomous human. Unfortunately, interactions between team members and particularly between virtual humans and users are limited. They perform parallel tasks, working towards the team's shared goal but cannot neither interact collaboratively on a same object nor exchange their roles during the simulation. Nevertheless, some platforms allow collaborative interactions between teammates. This is the case of the Generic Virtual Training (GVT) [9, 19] in its collaborative version. The MASCARET model [3] also potentially handle collaborative interaction [5].

2.2 Exchange of Avatars in CVET

The exchange of avatars, also termed role switching or role swapping, is a powerful feature to improve interaction and collaboration in CVET. It gives to users, the ability of taking control over another avatar to see and act in the environment from another actor's perspective. In the context of a team training, this feature has a real pedagogical interest but it also increases the training sessions efficiency by giving the possibility to learn several roles with less training. For instance, a trainer can switch his/her role with a trainee so that he/she can show a specific manipulation or gesture as he/she would have done in real life. In the Intelligent Robots field, it has been studied for several years [16, 20] . It is used to improve the achievement of collaborative tasks by giving the possibility of switching the role assigned to robots between leader and follower to speed up a task and to resolve conflicts. In the Virtual Reality domain, Evrard et al. [8] also take advantage of the possibility of switching between a leader and a follower role to improve the collaboration of users involved in a dyad haptic interaction for a physical collaborative task. More recently, role exchanges have also taken place in the context of human computer haptic collaboration [21]. The aforementioned work provides haptic feedback to the user through a force negotiation mechanism, in order to help him/her achieve a dynamic task. Furthermore, the system is able to detect when the user needs help and, if necessary, switches from an assistive role to an equal-control role to complete the task.

In the context of CVET, a few works proposes an exchange of avatars. For example, Raij et al. [23] present after-action reviewing tools using role exchange to enhance experiences between users and virtual humans. Yet, the exchange is a promising and challenging trend for this area because it allows new usages for both designers and users of CVET. However, outside the scope of CVET, it is interesting to note that several studies have demonstrated the importance of the link between users and their avatar [18, 12, 22]. For instance, Slater et al. [30] conducted a perception study about the experience of body transfer in a virtual environment. In this study, the user point of view is modified during the experimentation allowing the user to watch the same scene sequentially from different perspectives, from an external viewpoint to a view embodied in the main character's eyes. This study shows that there are differences in the users' reactions depending on their point of view, in other word, their role in the virtual environment. These works demonstrate that an exchange of avatars is not a simple exchange of embodiment but that it induces users to act differently.

2.3 Synthesis

To sum up, the interactions between virtual humans and users are usually restricted to their participation in parallel tasks working towards the shared goal of the team. Their contribution mainly consists in advancing the procedure on their own, with limited interactions with others. We feel these are strong limitations to collaboration between autonomous humans and users in CVET. Besides, having an effective collaboration with virtual humans, is not only a promising feature for the realism of CVET but it will also increase performance and cognition. Our interaction abstraction model, by answering this issue, also enables an interesting feature: the exchange of avatars. The exchange of avatars is a promising feature for CVET even though only a few studies exists for this domain. Yet, it has been used in several fields for many years and it is still commonly used, even in video games like in the last version of Grand Theft Auto (GTA 5)[1]. Furthermore, it has been demonstrated that its usage induces interesting mechanism in users' reaction for the purpose of CVET.

3 THE SHELL

This section presents our interaction abstraction model called the *Shell*. The *Shell* aims at simplifying the collaboration between virtual humans and users by providing them a common interface.

3.1 Motivation

Collaboration between virtual humans and users is of great interest for CVET. This type of cooperation not only requires compatible means of representing the world and the tasks for virtual humans and users, but also that these representations can be interpreted and used by actors of both nature [2]. Thus, humans need to be able to understand the collaborative task from the virtual human's perspective [6], and vice versa. In order to meet these requirements, we create the concept of the *Shell*[27], an entity containing a knowledge

[1] http://www.rockstargames.com/V/

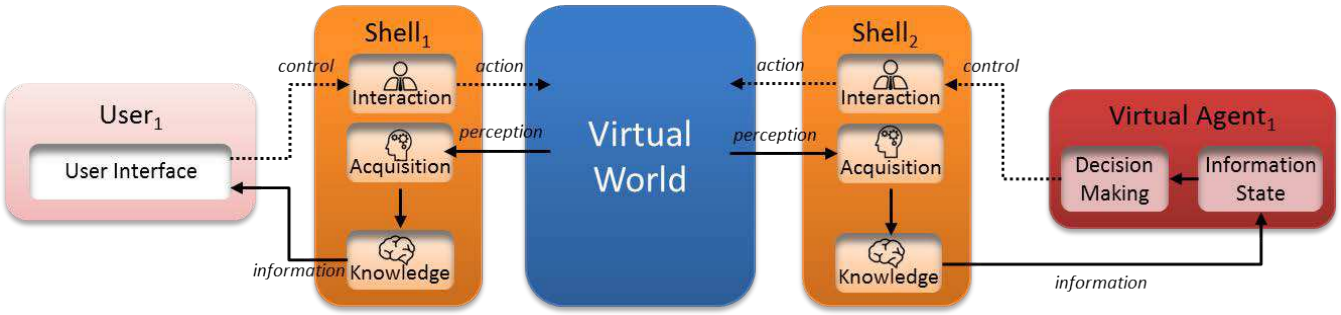


Figure 2: The shell: an abstraction level for the control of an avatar that is generic whether the avatar is controlled by a user or by a virtual agent.

base and which controls can be exchanged between virtual humans and real users. The *Shell* acts as a gateway unifying the actions and knowledge of both type of actors through different modules.

3.2 Concept

The *Shell* is an abstraction level for the design of avatars that is generic whether the avatar is controlled by a user or by a virtual agent. It wraps common features for both users and virtual humans to allow seamless interaction between a user and his/her avatar as well as between a virtual agent and its avatar. The *Shell* holds three main modules as shown in Fig.2 : the **interaction module**, the **acquisition module** and the **knowledge module**.

The **interaction module** handles interaction capabilities by dealing with the inputs from both virtual humans and users and by relaying it to the avatar which acts consequently in the virtual world. It allows collaborative interactions between several avatars such as collaborative manipulation of an object.

The **acquisition module** gathered information from the virtual world during the training session. Information is acquired through a perception process and is transmitted to the knowledge module. The perception modes used to acquire new knowledge could be of different nature: vision or dialogue.

The **knowledge module** contains at once basic information that is known a priori and information transmitted by the acquisition module.

The information managed by the knowledge module is of four types: *task*, *ego*, *team* and *world*. We refer to *task* information, knowledge that is known a priori, such as pieces of knowledge concerning the procedure to be trained or objects in the world. In the context of CVET, this information is crucial. The actor needs to know the series of tasks required to complete the procedure (i.e. the scenario) and whether a task is individual or collaborative in order to synchronize its actions with the team. The other three types of information (*ego*, *team* and *world*) are managed jointly by the knowledge module with the acquisition module. In fact, the acquisition module fills the knowledge module with these types of information. *Ego* refers to information about the avatar: internal resource states, task-specific skills and knowledge, interaction skills, internal states or collaborative profile. It is filled using an introspection process. *Team* refers to information about the collaborators and their actions like the role of each teammates or the completion of their current task. *World* refers to knowledge concerning the surroundings and the virtual world in order to locate objects or tools needed by the procedure. For instance, the last known position, the type, the properties or the interactive capacities of an object. These last two types of knowledge are filled using multi-modal perception i.e. through vision and dialogue.

The *Shell* concept is especially relevant in CVET with mixed agent/human teams. In this context, each actor has a role to play in the collaborative task whatever its real or virtual nature. The *Shell* wraps all the features needed to the completion of the task without considering whether the actor is controlled by a user or a virtual agent. The *Shell*, indeed, allows to pool the management of all the processes that are common to both type of actors. Moreover, using this representation, the knowledge base of a *Shell* can be accessed by both type of actors to help their decision process or update its own knowledge base.

4 The Exchange

This section presents how the model of the *Shell* enables a seamless role exchange between actors in CVET. Moreover, we discuss here how to make exchange of avatar efficient. Users must indeed understand it and, therefore, an appropriate representation of the exchange must be provided.

4.1 Motivation

Based on the *Shell* architecture, we have defined an exchange protocol that supports role switching between two actors. The exchange protocol gives to users the possibility of taking control over another avatar to see and act in the environment from a new perspective. This exchange can result from different reasons: control of multiple avatars at the same time, making an avatar perform a specific action, or accessing some knowledge it possesses but also resolving conflicts and deadlock situations or assisting another actor. In fact, the exchange of avatar enables new usages for both designers and users of CVET. Designers can create more complex scenarios than those generally used to enrich user experience. Training scenarios could indeed integrate mandatory exchanges to force users to assimilate various perspectives of the same procedure. Conversely, embodied in a new avatar, the trainee can feel and act as if he had a different role in a procedure at a crucial step.

4.2 Concept

We define an exchange protocol that allows two actors to exchange the avatar they are controlling. It basically consists in a *Shell* exchange between the actors. As the *Shell* provides a similar abstraction for the interactions and the knowledge, taking control of a new *Shell* gives users and virtual humans access to all the elements needed to resume the procedure started by the previous owner of the *Shell*.

Figure 3 shows the sequence of messages exchanged between two users and their shell in order to perform an exchange of avatars. First, $User_A$ ask for an exchange with the avatar of $User_B$ The *shell* associated to $User_A$ ($Shell_1$) transmits the message to the *shell* associated to $User_B$ ($Shell_2$). $User_B$ receives and accepts the demand. Finally, the *shells* proceed to the exchange and notify their new controller: $Shell_1$ notifies $User_B$ and $Shell_2 2$ notifies $User_A$.

This exchange protocol presents various assets. First, the two actors involved in the role switching do not have to physically move

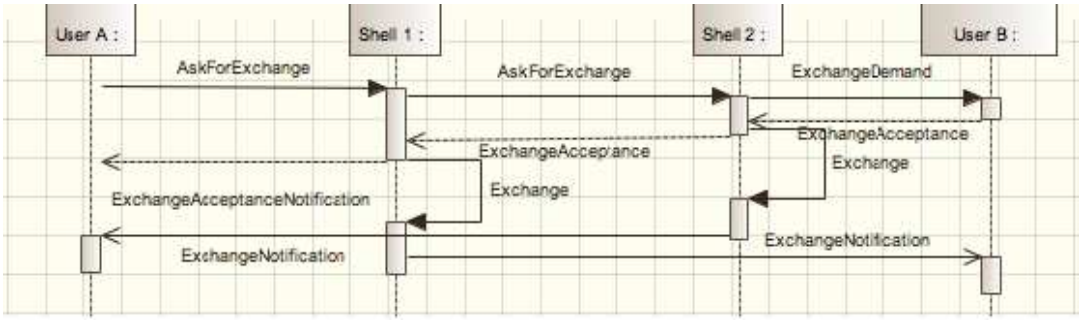


Figure 3: Thanks to the shell abstraction, a user can exchange seamlessly his/her avatar on-the-fly while preserving the data collected by its new avatar.

or even to navigate in the virtual world. They directly take control of the avatar previously controlled by another actor. Second, this protocol handles a critical key element in the role switching as it keeps consistent the knowledge representation associated to an actor. Therefore, the exchange can be carried out at any time of a procedure without disrupting the training. After an exchange, a user has access to the knowledge associated to his/her new embodiment. It may differ from its previous knowledge so that it improves his/her own knowledge.

4.3 Evaluation of the Exchange of Avatars

While working on the exchange protocol, a question emerges:

> "How to represent an exchange to make it understandable by users ?"

We highlight three criteria that an exchange feature has to meet in order to be usable by users. First, users must comprehend they are involved in an exchange process. Second, they must understand they are now embodied in a new avatar. Third, they must be able to locate themselves in the virtual environment before and after the exchange happened. Otherwise, they also need to deal with various metadata associated with their new avatars such as new capabilities, different knowledge on the environment or on simulation history and so on. To meet these requirements, we have worked on the visual representation of an exchange of avatars. We call *exchange metaphor*, a specific configuration of parameters used to represent an exchange (animation, color, duration, ...). Although exchanges of avatars have often been used in the literature (see section 2.2), the method of representing such an exchange has never been evaluated.

Therefore, we have conducted a perception study to evaluate different exchange metaphors in order to assess the relevant criteria that need to be taken into account in the design of exchange metaphors. Our study revealed several results that have been published in [13]. First, the study indicates that different parameters should be taken into account to define an exchange metaphor depending on whether the user is involved or not in the exchange. If a user performs an exchange, the other users indeed needs to be aware of it. Nevertheless, the exchange metaphor used for a user witnessing the exchange should not be the same as for a user triggering it. Second, different metaphors should be proposed depending on the level of familiarity of the end-user with Virtual Environment. The study indeed shows that users that are familiar with VR want efficient and quick metaphors whereas novice users need explicit and comprehensible metaphors. Eventually, metaphors must be adapted to the targeted application. User-friendly metaphors, despite being well-appreciated by users, often lead to a longer task accomplishment duration, even though no difference has been found in terms of either ease of use or exchange awareness. This study answered our initial question and enabled us to improve the representation of exchange of avatars in our sample application.

5 Implementation and Results



Figure 4: Our test scenario running in an immersive room.

We present in this section our implementation of the Shell concept and explain how we handle the exchange of avatar. Our contributions has been implemented in a sample CVET application that is available on several platforms.

5.1 Implementation

According to our model, the implementation of the *Shell* gathers the control, the acquisition and the knowledge attached to an avatar in one single entity. Thus, its architecture, illustrated in Fig.2, is composed of three parts following the conceptual analysis of section 3.2 :

1. the interaction module that allows the *Shell* to interact with the environment,
2. the acquisition module that allows the entity to assimilate new knowledge,
3. the knowledge database that stores and classifies the knowledge acquired by the *Shell*.

At runtime, the acquisition module captures and transmits new information in order to update the knowledge database. Knowledge is classified in the database depending on its type as defined in the previous section (3.2): *ego*, *task*, *team* and *world*. Our implementation enables new knowledge to be acquired through both vision and dialogue but it can easily be extended to other perception channels. Various data is attached to the objects of the virtual world such as their types, their locations and their properties. Each time the object is perceived by the avatar (i.e each time the object is in its field of view), the data associated to the object is transmitted to the



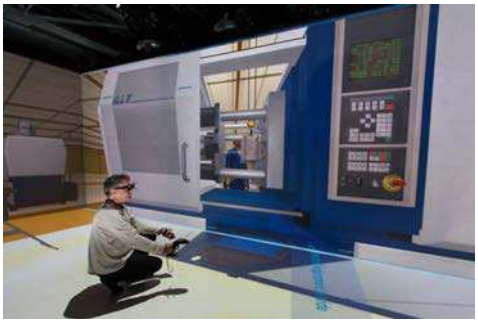
(a) Desktop computer



(b) Screen and projector



(c) HMD



(d) Immersive room

Figure 5: The different platforms running our sample CVET.

acquisition module. This data is updated each time the object is perceived. Besides, an inference system allows to acquire information about the other *Shell* when perceived by an avatar. For example, a *Shell* can retrieve information about the action performed by another *Shell*. Moreover, our application includes speech recognition and synthesis to enable users and virtual humans to exchange information about the current task. Information exchanges through dialog is also transmitted to the acquisition module.

Both virtual humans and users have access the knowledge database of their associated *Shell*. On the one side, autonomous virtual humans are based on a BDI architecture [17, 4] and uses the knowledge of their *Shell* for decision-making and for conversation using an Information-State based approach. On the other side, users have access to the knowledge of the *Shell* through a graphical user interface. The knowledge displayed by the user interface complete their expertise during the training.

The Shell has been implemented in a CVET using the GVT platform [2] and Unity[3]. We use a test scenario resulting from a concrete CVET application defined by an industrial partner. This scenario describes a maintenance procedure in a plastics manufacture that consists in the replacement of a mold in a plastic injection molding machine (see fig. 4). This specific intervention requires a precise coordination of tasks between two operators : a setter and a machine operator. The use of an autonomous virtual human allows the learners to execute the learning procedure even if there is no other user available for the training session. To begin, each learner chooses one of the two roles. When ready, he/she is given control of the *Shell* corresponding to his/her role and executes his/her part of the training scenario. If the other role has not been provided, a virtual human takes it and executes its part of the procedure. At the beginning of the training, each actor has the same knowledge but their role differs. While running the procedure, they acquire new knowledge using perception channel i.e. dialogue and vision. For example, an actor can ask its teammate what is the next action he/she is going to perform and improve his/her knowledge.

We have also implemented our exchange protocol in this application. In order to perform an exchange, the user selects an actor in its surrounding. If the selected actor is controlled by a virtual human, the exchange is automatically accepted. Then, immediately, the user is moved in its new embodiment, changing its point of view. Otherwise, if the selected actor is controlled by another user, the later must accept the exchange to make it effective as explain in section 4.2. If the actor he wants to exchange with is not in his/her field of view, a graphical user interface allows him/hers to select one of the role required by the procedure.

5.2 Immersion of the User in CVET

Most serious games and virtual environments for training are designed as classical computer softwares. We do think, however, that the immersion of the user is a crucial element that needs to be taken into account in the assimilation of a scenario. Indeed, the more a training situation will be close to real-life condition, the easier it will be for a trainee to adapt the acquired knowledge in the real world. Based on this idea, our CVET application has been designed for multiple immersive platforms. As a result, a trainee can now practice the example scenario presented in 5.1 on the following set-ups:

1. a classical desktop computer (Fig. 5(a)),
2. a portable screen combined with a projector and the tracking of the user's motions (Fig. 5(b)),
3. an HMD set-up using the Occulus Rift [4] combined with a

[2]Generic Virtual Training – http://www.gvt-nexter.fr/

[3]http://unity3d.com/

[4]http://www.oculusvr.com/

	Desktop station	Screen + movements tracking	HMD + Hands movements	Immersive room
Immersion degree	4	3	2	1
Affordability degree	1	3	2	4
Freedom of movement degree	4	2	3	1

Table 1: Set-ups tested for our CVET application and ranking from 4 (less adequate) to 1 (more appropriate) regarding three criteria: the immersion, the affordability and the freedom of movement.

Razor Hydra [5] to follow the hands' movements of the user (Fig. 5(c)),

4. in a large 4-sided immersive room with stereoscopic rendering and motion tracking to follow the user's actions (Fig. 5(d)).

These set-ups, however, do not provide to the user the same feedbacks and are not accessible in the same conditions to the trainees. Table 1 provides an overview of these set-ups and proposes a ranking of them, from 1st to 4th, regarding three criteria : the immersion degree, the affordability for a single user (*i.e.* financial needs) and the freedom of movement for the user. On the one hand, the immersive room provides a better immersion as well as a better freedom to the user, however, it is a very expensive set-up which is not easily available for virtual training. On the other hand, desktop station are really common and a rather cheap equipment. Nonetheless, they do not provide an immersive training as the interaction using keyboard and mouse does not reproduce real movements and nor interactions. During our demonstrations, we observe that the use of a screen and a video-projector with a tracking system, or the use of an HMD with another tracking device might be appropriate solutions to obtain a good equilibrium between immersion, affordability and freedom of movements.

Besides, to reinforce the immersion of the user, and in addition to the use of different set-ups, we also explored two other features. First, we increase the interaction level of the user by adding a communication module to our CVET. This communication module handles natural communication between users and virtual humans. To do so, we are using speech recognition to capture the dialogue of the user. For virtual humans, we are using the architecture presented in section 5.1 which allows them to generate communicative behaviors toward the user. Second, we added a module to handle a physically-simulated environment. Indeed, most CVET are using precomputed animation during a training scenario. Using physics allows us to render real reactions of objects from the world instead of scripted and predefined reactions. Thus, when coupling this physics with a tracking of the user's gestures, he/she is able to interact easily and naturally with all the virtual training environment which enhances his/her user experience.

5.3 Exchange of Avatars and Attention of the User

It appears to us while working on our exchange protocol that this feature, despite being useful, could distract the user and make him/her less focused on the training. For that reason, we try to evaluate whether some metaphors are less distracting than another. Thus, the evaluation presented in section 4.3 has been interested in assessing the distraction of the user while an exchange of avatars happened.

The first experiment compares 3 exchange metaphors for users that are only witnessing the exchange, they do not take part in it. The 3 metaphors used are illustrated in fig.6:

- flickering avatar (see 6(a)): each time an exchange happened, the two involved avatars flickered from a color to white,
- ghost translation (see 6(b)): each time an exchange happened, two "ghosts" appear that translate from one avatar to another,
- popup notification (see 6(c)): each time an exchange happened, a popup with the picture of the involved avatar appears to notify the user that their control has been exchanged.

We had 54 participants that has been divided in 3 groups. Each one of them uses one of the 3 metaphors presented above. They were informed that several exchanges of avatar will happen during the session and they were asked to run a series of task while the exchange happened. To the question: "Does the metaphor distract you from the task ?", our analysis suggests that no metaphor was distracting the participant from the task to execute than others. This suggests that, concerning exchanges that are witnessed by a user, the visual representation is of few consequences on the distraction of the user.

The second experiment compares 5 metaphors for users that are triggering the exchange. The used metaphors combined two visual stimuli: camera moves and animated feedback. Camera moves can be whether a straight linear motion between the two involved avatars (camera move 1) or a camera motion that goes up smoothly from the first avatar to an external point of view then comes back down to the second avatar (camera move 2). The animated feedback is, in turn, a combination of the flickering avatar and the ghost translation described in the first experiment. Our five metaphors are defined this way:

- Metaphor 1: camera move 1 with animated feedback,
- Metaphor 2: camera move 1 without animated feedback,
- Metaphor 3: camera move 2 with animated feedback,
- Metaphor 4: camera move 2 without animated feedback,
- Metaphor 5: no camera move and no animated feedback i.e. direct teleportation.

We had 52 participants. They were asked to compare two of our 5 metaphors at a time and therefore performed 10 comparisons. At the end of the experiment, for each metaphor, they were asked to note the affirmation: "The metaphor allows remaining attentive to the environment." from 1 to 7 (1: not agree at all, 7: totally agree). The result is illustrated in Fig. 7. The results obtained suggest that the metaphor 1 was perceived as more distracting than the metaphors 2, 3 and 4. Nevertheless, this conclusion must be caught with caution for two main reasons. First, the conclusion is not really obvious if we refer to the Fig. 7. We can indeed notice that our statistical analysis could not compare the metaphor 5 with the others although its repartition is exactly the same as for the metaphor 3. Second, the experimental condition can not be applied to every virtual environment. The virtual environment used was indeed a small room and the user was always facing the avatar he/she exchanged with (see Fig. 8). Therefore, the conditions of our experiment are too specifics to be applied to every CVET. We believe that further studies must be conducted in order to answer whether the attention of the user is disturbed by the exchange he/she is triggering.

[5] http://www.razerzone.com/fr-fr/gaming-controllers/razer-hydra/



(a) Flickering avatar (b) Ghost translation (c) Popup notification

Figure 6: Illustration of the metaphors used in our first perception study.

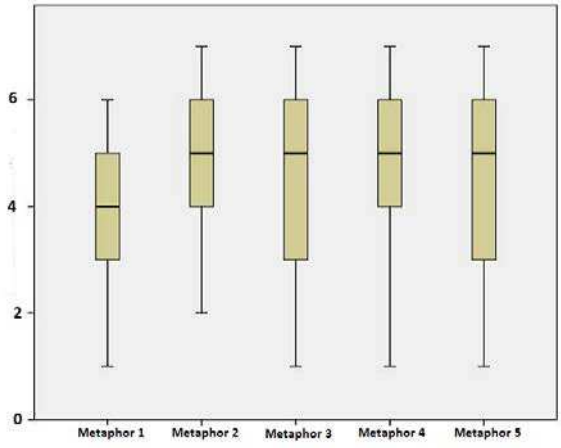


Figure 7: Our experiment shows that when the user is triggering the exchange, metaphors with multiple visual stimuli are a bit more distracting than more simple metaphors.



Figure 8: The environment used for our second perception study.

6 How Does the Shell Improve Collaboration in CVET ?

This section discusses how the contributions presented above raise new ways to collaborate in CVET. Most of the Shell improvements on collaboration between users and virtual humans are induced by the exchange feature it allowed.

6.1 Collaboration Between a Trainer and the Trainees

Using the exchange protocol, the trainer has several possible ways to collaborate and interact with trainees. He/she can :

- take control of any actor involved in the collaborative task with the trainee. In this way, he/she can interact with the trainee to help and guide him through dialogue. On the contrary, he/she is able to provoke unexpected situations by making mistakes or giving wrong information to the trainee on pedagogical purposes.
- take control of the trainee himself. In this way, he/she can induce the user to see the situation from a different perspective at a crucial step. He/she can also take the user's role to show him/her a specific step that the trainee failed to complete as he would have done in real life.
- let a trainee complete a procedure, started by the trainer, in order to evaluate the trainee's capabilities.

6.2 Collaboration Between Trainees and Real or Virtual Teammates

Thanks to our exchange protocol, a trainee can exchange his/her role in the procedure at any time and as often as desired. This can help him to collaborate more efficiently with his/her real or virtual teammates by:

- improving his/her virtual training experience. For instance, if he/she feels comfortable with one step of a procedure, then the trainee can choose another role. Thus, he/she remains focus and learns from this new perspective to better collaborate during another training session.
- improving his/her procedure comprehension. For example, during a surgical procedure, a nurse can exchange his/her role with another nurse at a specific step to integrate his/her point of view in order to improve her/his own reaction in real situation.
- reinforcing the team spirit. In fact, the exchange of avatars allows the trainee to develop more efficiently his/her cognitive empathy and thus, the team cohesion. It has indeed been demonstrated by Seers et al. in [28, 29] that the quality of lateral (within-group) relationships is as important as that of vertical relationships (with hierarchical supervisors).

7 Conclusion

This paper presents an interaction abstraction model for the control of avatars in CVET that is shared by both users and autonomous virtual humans. This abstraction level, called the *Shell*, allows pooling the main process of avatar control: the interaction response and the knowledge management. Moreover, the *Shell* enables us to define an exchange protocol that allows users of CVET to exchange their role seamlessly during a training session. This exchange feature preserved all the data associated to a role. It ensures that the new owner of a role has access to every data required to resume the task initiate by the previous owner. We show that the exchange is a

promising feature for CVET to improve collaboration and interaction. Nonetheless, it raises several questions that must be answered such as how much the exchange distracts the user and how to minimize its impact. We have conducted preliminary studies that tends to indicate that the representation of the exchange is of few consequences on the distraction of the user. Nevertheless, work remains to be done to clearly answer this question. Besides, several other questions must be studied to make the exchange feature more usable by users. For example, how to adapt the exchange metaphor to all kind of render devices from a portable device to a fully immersive set up ? Future works will also focus on proposing different uses of the *Shell*. For instance, it would be interesting to have a simultaneous access on a *Shell* by two or more actors. An actor could then follows a user as a passive controller, monitor his/her actions or take over a part of the procedure when needed. This control sharing could also allows a trainer to fine-tune a manipulation performed by a trainee.

ACKNOWLEDGEMENTS

This work was supported by the French Research National Agency project named CORVETTE (ANR-10-CONTINT-CORD-012) and the French Unique Interdepartmental Funds SIFORAS (FUI 11).

REFERENCES

[1] N. I. Badler, C. A. Erignac, and Y. Liu. Virtual humans for validating maintenance procedures. *Communication of the ACM*, 45(7):56–63, 2002.

[2] J. Bradshaw, P. Beautement, and A. Raj. Toward a deliberative and reactive agent architecture for augmented cognition. *DARPA Augmented Cognition Program White Paper*, 2002.

[3] C. Buche and R. Querrec. An expert system manipulating knowledge to help human learners into virtual environment. volume 38, pages 8446–8457. Elsevier, 2011.

[4] H. Bunt. Multifunctionality in dialogue. *Computer Speech & Language*, 25(2):222–245, 2011.

[5] P. Chevaillier, T.-H. Trinh, M. Barange, P. De Loor, F. Devillers, J. Soler, and R. Querrec. *Semantic modeling of Virtual Environments using MASCARET*, pages 1–8. 2012.

[6] K. Christoffersen and D. Woods. How to make automated systems team players. *Advances in human performance and cognitive engineering research*, 2:1–12, 2002.

[7] J. Dugdale, B. Pavard, N. Pallamin, M. el Jed, and C. L. Maugan. Emergency fire incident training in a virtual world. In *Proceedings ISCRAM 2004*, volume 167, 2004.

[8] P. Evrard and A. Kheddar. Homotopy switching model for dyad haptic interaction in physical collaborative tasks. In *EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 45–50, 2009.

[9] S. Gerbaud, N. Mollet, F. Ganier, B. Arnaldi, and J. Tisseau. GVT: a platform to create virtual environments for procedural training. In *IEEE Virtual Reality*, pages 225–232, Reno United States, 2008.

[10] P. Kenny, A. Hartholt, J. Gratch, W. Swartout, D. Traum, S. Marsella, and D. Piepol. Building interactive virtual humans for training environments. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, volume 2007, 2007.

[11] J. Lee, D. DeVault, S. Marsella, and D. Traum. Thoughts on fml: Behavior generation in the virtual human communication architecture. *Proceedings of FML*, 2008.

[12] B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke. Video ergo sum: manipulating bodily self-consciousness. *Science*, 317:1096–1099, 2007.

[13] T. Lopez, R. Bouville, E. Loup-Escande, F. Nouviale, V. Gouranton, and B. Arnaldi. Exchange of avatars: Toward a better perception and understanding. *IEEE Transactions on Visualization and Computer Graphics (to appear)*, 2014.

[14] I. Machado, P. Brna, and A. Paiva. Tell me a story. *Virtual Reality*, 9(1):34–48, 2005.

[15] I. Machado, A. Paiva, and R. Prada. Is the wolf angry or... just hungry? In *Proceedings of the fifth international conference on Autonomous agents*, pages 370–376. ACM, 2001.

[16] E. Martinson and R. Arkin. Learning to role-switch in multi-robot systems. In *IEEE International Conference on Robotics and Automation, 2003. Proceedings. ICRA 03*, pages 2727–2734 vol.2, 2003.

[17] V. Mascardi, D. Demergasso, and D. Ancona. Languages for programming BDI-style agents: an overview. In *Proceedings of WOA'05*, 2005.

[18] A. Maselli and M. Slater. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience*, 7, 03 2013.

[19] N. Mollet and B. Arnaldi. Storytelling in virtual reality for training. In *Edutainment*, pages 334–347, Hangzhou Chine, 2006.

[20] E. Monacelli, C. Riman, R. Thieffry, I. Mougharbel, and S. Delaplace. A reactive assistive role switching for interaction management in cooperative tasks. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5118–5123, 2006.

[21] S. Oguz, A. Kucukyilmaz, T. Sezgin, and C. Basdogan. Haptic negotiation and role exchange for collaboration in virtual environments. In *2010 IEEE Haptics Symposium*, pages 371–378, 2010.

[22] V. I. Petkova and H. H. Ehrsson. If i were you: perceptual illusion of body swapping. *PloS one*, 3(12):e3832, 2008.

[23] A. B. Raij and B. C. Lok. Ipsviz: An after-action review tool for human-virtual human experiences. In *Virtual Reality Conference, 2008. VR'08. IEEE*, pages 91–98. IEEE, 2008.

[24] J. Rickel and W. L. Johnson. *Virtual Humans for Team Training in Virtual Reality*. 1999.

[25] J. Rickel and W. L. Johnson. Task-oriented collaboration with embodied agents in virtual worlds. *Embodied conversational agents*, page 95122, 2000.

[26] J. Rickel and W. L. Johnson. *Extending virtual humans to support team training in virtual reality, Exploring artificial intelligence in the new millennium*. Morgan Kaufmann Publishers Inc., San Francisco, CA, 2003.

[27] A. Saraos Luna, V. Gouranton, T. Lopez, and B. Arnaldi. The Perceptive Puppet: Seamless Embodiment Exchange Between Real and Virtual Humans in Virtual Environments for Training. In *International Conference on Computer Graphics Theory and Applications*, pages 1–6, Barcelone, Espagne, Feb. 2013. FUI11 SIFORAS.

[28] A. Seers. Team-member exchange quality: A new construct for role-making research. *Organizational behavior and human decision processes*, 43(1):118–135, 1989.

[29] A. Seers, M. M. Petty, and J. F. Cashman. Team-member exchange under team and traditional management: A naturally occurring quasi-experiment. *Group and Organization Management*, 20(1):18–38, 1995.

[30] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke. First person experience of body transfer in virtual reality. *PloS one*, 5(5), 2010.

[31] K. Sycara and G. Sukthankar. Literature review of teamwork models. Technical Report CMU-RI-TR-06-50, Robotics Institute, Pittsburgh, PA, 2006.