Towards an engineering approach for advanced interaction techniques in 3D environments

Louis-Pierre Bergé, Emmanuel Dubois, Minica Houry-Panchetti, Mathieu Raynal, Cédric Sanza
University of Toulouse; IRIT; Elipse / Vortex
F – 31 062 Toulouse France
Louis-Pierre.Berge@irit.fr; Emmanuel.Dubois@irit.fr

ABSTRACT

In recent years, Virtual Environments have appeared in new areas such as mass-market, web or mobile situations. In parallel, advanced forms of interactions are emerging such as tactile, mixed, tangible or spatial user interfaces, promoting ease of learning and use. To contribute to the democratization of 3D Virtual Environments (3DVE) and their use by persons who are not experts in 3D and occasional users, simultaneously considering Computer Graphics and Human Computer Interaction design considerations is required. In this position paper, we first provide an overview of a new analytical framework for the design of advanced interaction techniques for 3D Virtual Environment. It consists in identifying links that support the interaction and connect user's tasks to be performed in a 3DVE with the targeted scene graph. We relate our work to existing modeling approaches and discuss about our expectations with regards to the engineering of advanced interaction technique.

Author Keywords

3D User Interaction, Design, Analytical Framework, 3D Interaction Techniques, Multimodal Interaction.

ACM Classification Keywords

D.2.2 [**Software engineering**]: Design Tools and Techniques - *User interfaces*; H.5.2 [**Information interfaces and presentation**]: User Interfaces - *Theory and methods, User-centered design*; I.3.6 **[Computer Graphics**]: Methodology and Techniques - *Interaction techniques.*

INTRODUCTION

With the evolution of technologies, computing capabilities and rendering techniques, the use of 3D Virtual Environments (3DVE) is becoming popular. 3DVE are no longer restricted to industrial uses and they are now available to the mass-market in various situations: for leisure in video games, to explore a city in Google Earth or in public displays [26], to design house furniture [17] or to explore cultural heritage sites in a museum [6]. However, in these mass market contexts, the user's attention must be focused on the content of the message and not distracted by any difficulties caused by the use of a complex or inappropriate interaction technique. This is especially true in a museum where the maximization of the knowledge transfer is the primary goal of an interactive 3D experience. Common devices, such as keyboard and mouse [21] or joystick [30] are therefore widely used in museums. To increase the immersion of the user, solutions combining multiple screens or cave-like devices [6] also exist. However, these solutions are cumbersome and expensive.

Meanwhile, the Human-Computer Interaction (HCI) research domain is rapidly evolving and growing in complexity with new advanced forms of interaction such as mobile [16], ambient computing [13], spatial interfaces [15] and tangible user interface [29]. A common feature to these advanced forms of interaction is the attempt to involve and combine the use of multiple objects and entities taken in the physical and digital environments: interactive solutions are smoothly integrated in the user's activity and have been proved to be easier to apprehend by newcomers [24]. Successful uses of such advanced interaction have been recently demonstrated in mass-market applications involving 3DVE for museums [11][12].

It thus appears that 3D interactive applications are more and more widespread, from professional context to public spaces and from expert users to very occasional users. In addition, advanced forms of interaction techniques offer new potentials such as being based on personal belongings (devices or artefacts), integrated in the physical environment, easy to apprehend. But developing 3D interactive applications on one hand and advanced interaction techniques on the other hand are two preoccupations that are mostly considered through separated approaches, leading to compartmentalized progresses. There is therefore a need for understanding and supporting the engineering of advanced interaction techniques for exploring and taking advantage of 3DVE.

In this position paper, we first provide an overview on a new analytical framework for helping and guiding the design of advanced interaction techniques for 3DVE. We motivate and illustrate the choice of its grounding elements and then discuss a number of existing modeling approaches potentially useful to complement or refine this framework. We finally discuss our expectations from the workshop, with regard to the proposed framework and more widely with regards to the engineering of advanced interaction technique.

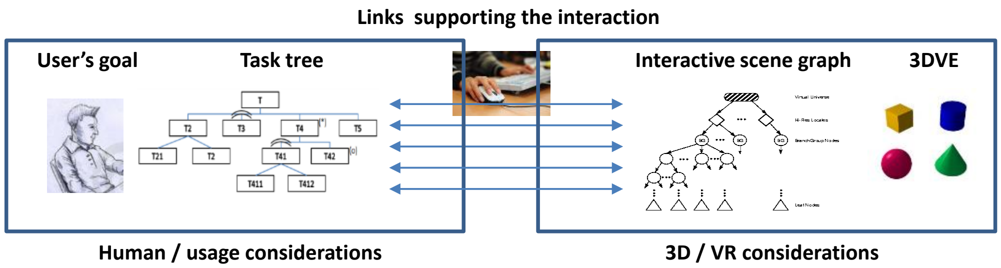

Figure 1: Overview of our analytical framework

OVERVIEW OF A NEW ANALYTICAL FRAMEWORK

We first present the two pillars of our modeling approach, tasks analysis and interactive scene graph. We then introduce the notion of links between these two pillars.

Tasks tree

The first pillar of the analytical framework is the result of a task analysis: a task tree. Task analysis (Figure 1 – left) consists in a decomposition of user's goal into tasks and sub-tasks which must be achieved to reach the user's goal. Task analysis is widely accepted in the HCI community as a starting point to the design of interactive techniques because it thus provides an algorithm of the user's activity, the logic and the dynamic of tasks accomplishment. But a task analysis does not express how the task is concretely performed with the system: no information related to the interaction technique is provided. Task analysis is particularly useful to understand and structure the user's activity, define functional specifications, identify data requirements, etc. [2].

Different formalisms exist to represent the result of a task analysis. We choose to rely on the familiar Hierarchical Task Analysis (HTA) formalism [2]. With HTA, the decomposition of the task is presented as a hierarchical tree of tasks and sub-tasks, enriched with some attributes (iteration, optional, parallel, etc.). Choosing a light approach ensures that non-experts are able to use it as a support for designing advanced multimodal interaction for 3DVE.

When used to describe tasks with an interactive 3DVE, most of the identified tasks are specific to the application domain and the leaves of the tasks tree are close to Bowman's tasks [5] (navigation, selection, manipulation, system control). But none of the 3D elements impacted by the tasks can be specified. This is where the interactive scene graph comes to action.

Interactive scene graph

The second pillar of the analytical framework is the interactive scene graph which provides a structured description of the 3DVE to be used. This scene graph is therefore specific to each 3DVE. A 3DVE generally consists of 3D objects (meshes, widgets or basic geometrical elements such as cone, cube, cylinder, etc.), lights and virtual cameras. At a finer grain 3D objects are described as a set of vertices (geometry), faces and edges (topology). Manipulations of 3D objects (translations, rotations or scaling) must therefore take into consideration the underlying topology and geometry. To assist this process, the concept of scene graph [25] has been developed to organize the 3D elements and provide for developers a structure for the assembly of a 3D scene.

Scene graph is a widely accepted method used in the Computer Graphic (CG) community to describe the essential components of a 3DVE. Scene graphs are also relevant to our context because we need to understand and take into account the structure of the 3D scene to design the interaction with it. But, we are not interested in elements related to the implementation of the scene graph by the 3D API in charge of the rendering. We are also not interested in the way the scene graph may impact the use of 3D engines for solving issues like texture management or collision.

However, with the scene graph description, only geometric and topologic aspects are expressed. It is not clearly identified which parts of which components of the 3D scene are likely to be impacted by user's interaction. To this end, we propose to define the "interactive scene graph" (Figure 1 – right). Its aim is to highlight and characterize **handled** and **not-handled** objects, i.e. objects impacted or not by one of the user's tasks identified during the tasks analysis.

The definition of the "interactive scene graph" is based on the most relevant features used to support 3DVE user's interaction in 3D engines like Unity 3D [32], Irrlicht [34]. It is also derived from the standard description language X3D [33]. X3D supports the description of ***animated*** 3D scenes: behaviors among 3D nodes are expressed in script nodes or simple links among 3D nodes. The role of the "interactive scene graph" goes beyond the description of animations in the 3DVE: it emphasizes which are the elements and attributes of the 3D scene with which external elements may ***interact***. Based on these existing approaches, we distinguish two types of handled objects: **components** and **renderers**.

Components are 3D objects composing the 3DVE (mesh, geometrical element, widgets). User's interaction may impact Components through the modification of two classes of attributes: ***state*** and ***manipulation***. State attributes refer to the color, the texture or the visibility (display or not in the 3DVE) of the object. Manipulation attributes are more complex. Three levels of manipulation attributes coexist:

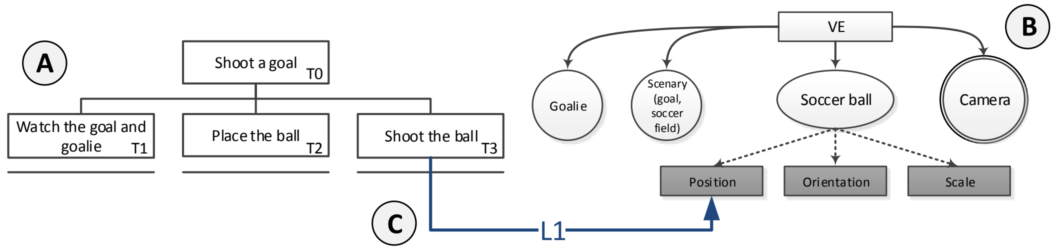

Figure 2: Task tree (A), scene graph (B) and link (C) of the Penalty Shootout interaction technique with Portico [4]

the *object*, its *faces* and its *points*. First, one may modify position, orientation or scale of an *object* as a whole. Second, an object is made of a set of *faces*: depending on the object structure, modifying a face can be limited in terms of degrees of freedom in orientations and scales. Third and finest level, a face of an object is made of a set of *points*: at that level one can only acts on the 3D position of each point. Faces and points levels are thus useful to refine and characterize the deformation of a 3D object.

Renderers are objects such as lights or camera, taking part in the rendering of the 3D scene. User's interaction may impact Renderers through the modification of two classes of attributes: ***state*** and ***manipulation***. States include attribute such as enabled/disabled and color. Manipulation attributes correspond to position and orientation, with orientation depicting the definition of the point of view of the camera.

The next step consists in identifying existing links between elements of the 3DVE and user's sub-tasks of the task tree that are affecting them.

Linking user's task with elements of the 3D scene

It consists in identifying for a set of sub-tasks of the task tree, the attribute(s) of the interactive scene graph affected by the realization of each of these specific sub-tasks (Figure 1 – middle). As a result pairs of user's sub-task and attributes of the 3DVE are clearly highlighted. This set of links provides a complete view on the user's interaction that will be performed in the 3DVE to perform the user's activities required to reach his/her goal. Each link depicts the use of one interaction modality [20]: on the one hand each link may involve a different interaction modality, on the other hand every links is using the same and unique interaction modality. In addition, we anticipate that the operators (sequence, alternative, etc.) present in the task tree at higher-levels will influence the design of the links.

Highlighting the pairs of user's sub-task and attributes of the interactive scene graph therefore constitutes the description of the overall user's interaction with the 3DVE. Furthermore, it establishes a link between user's activities (task analysis) and the 3DVE content and behavior (the interactive scene graph). Therefore, this description constitutes a support to reason about the design of the overall system and takes advantage of HCI and CG specificities. It overcomes the description of one specific task or one specific technique out of the context of use and manipulation of the 3DVE. In the next section, we illustrate the overview of our analytical framework with an example of the literature.

ILLUSTRATIVE EXAMPLE

Portico [4] is an interactive system for enabling tangible interaction on and around tablet computers. We focus our illustration on one of their application example named penalty shootout. Let us describe the interactive setting with the use of our analytical framework. The aim of the user is to shoot a goal (Figure 2 – A - task T0). For this, the user has to watch the goal and goalie (T1), place the ball (T2) and shoot the ball (T3). The system does not support multiple balls thereby there is a unique ball available for the user. Now, the interactive scene graph (Figure 2 – B) is composed by a **non-handled** camera (*renderer*), a **handled** soccer ball (*component*) and some **non-handled** objects like the goal, the goalie and the soccer field. Finally, the designer has adopted a tangible interaction to establish the link between the task tree and the scene graph (Figure 2 – C). Concretely, the soccer ball is a physical object manipulated by the user and thus, the sole task (L1) impacting the interactive scene graph is the task T3. The link L1 of this example connects the task T3 to the *manipulation* attribute position of the 3D ball in the scene graph. The task T2 does not impact the virtual ball position because the virtual ball position change only when the physical ball touch the tablet.

From this description, we can identify that the interactive situation supports the manipulation task in a 3DVE through the link L1 and an implicit selection task through the manipulation of the physical ball. The remaining Bowman's 3D tasks (navigation, system control) are not supported by this interactive technique. The description of the link L1 also highlights a direct connection between a physical object and a virtual 3D element (the ball): physical behavior and representation are directly mapped to the behavior of the corresponding digital object. Although it is a rather simple example, it shows that the framework can help in visualizing when the user's focus has to be on the 3D scene with regards to the user's task realization. If several links are present it may also help identify inappropriate sequence of interaction, such as switching modalities while focusing on the same 3D parts.

To better structure this kind of reasoning, designing the link L1 is subject to a set of design aspects that we extracted from the literature and summarize in the following section.

DESIGN ASPECTS OF THE INTERACTION WITH 3D

The HCI and CG communities have already been working on the design and implementation of advanced interaction techniques with 3D. Different points of view have been adopted thus revealing multiple design aspects. We summarize these considerations along the three parts of our analytical framework: user's interaction, 3D system and links between them.

Regarding the **user's interaction**, relevant considerations include the specification of the users expertise with the manipulation of the application as defined in Rasmussen work [22]. The definition of the application type (AR, VR, desktop application) is also considered as crucial in the 3-DIC model [10]. Given the use of advanced interaction technique, it is also required to specify which parts of the physical world are involved. It includes a description of the objects used and their constraints [27]. It also requires to specify how information are transferred from the user to the 3D system and vice-versa, as partly addressed in the ASUR [8] and MIM [7] models. It goes even to more precise description of the gestures and characterization of movements that will be performed by the users [19] as well as perceptual properties of interest such as visual, tactile and auditory properties.

Regarding the **3D system**, as already mentioned X3D [33] is the standard to describe the 3D nodes, structure and internal animation or behavior. Most of works however focus on the description of virtual reality interaction techniques such as ray casting and are focusing on their behavior or implementation. Among them the reusable library of 3D interaction technique [9], the Petri Net model [31], ontology model [14], the 3-DIC model [10] or IFFI [23] and Viargo [28] library are complementary alternatives.

Finally, in the literature, **the link between our two pillars** is often limited to the analysis of the required input device. Simple taxonomies offer an overview of the possibilities such as the Mackinlay taxonomy [18]. More elaborated models like RVDT [1] or InTml [9] deal with a particular aspect of input device, e.g. the type of data (float, integer, boolean) and the number of sensed DOF.

DISCUSSION

Obviously, developing advanced interaction techniques for 3DVE requires to confront multiple design considerations and to pay attention to both communities' preoccupations. To do so, offering a structured and refined set of design attributes that reconcile these multiple aspects will lead to a better understanding of the links between a user's goal and attribute of a 3D scene. For example metrics might be extracted to efficiently compare techniques; properties might be defined to clearly express how design choices in the user's part impact design choices in the 3D system parts and conversely.

We believe that providing such a structured approach to describe the links between task and scene graph is a fruitful way to help reason about the design and implementation of advanced interaction techniques for 3DVE. The resulting model or notation will constitute a pseudo-formal description language of interaction techniques for 3DVE. From such description, a semi-automatic implementation of the described advanced interaction for 3DVE could then be built in a platform for rapid development of multimodal interaction such as the Dynamo framework [3].

CONCLUSION

Applying advanced forms of interaction to 3D applications is required to contribute to a more effective use of 3D interactive environment. As multiple types of user and context are potentially targeted a user centered approach to the design of interactive 3D application is particularly expected.

In this paper we proposed a way to narrow two communities by involving well established design resource of each domain as the two pillars of a dedicated approach. We then identified a set of existing design and implementation supports for bridging the gap between these two pillars.

During the workshop, we hope to find the opportunity to further illustrate the use of this framework on different prototypes we have implemented in our lab. We then expect a fruitful discussion with the other participants of the workshop to identify additional existing design approaches relevant to this context of interaction with 3DVE, or relevant metrics, properties or considerations. In particular, we are interested to discuss what could be the ways to tightly anchor 3D specificities in the design of interaction technique. We are also interested in refining the links between our two pillars with relevant approaches. Finally, we hope to hear about similar approaches in different contexts, i.e. a context in which advanced HCI and another domain are involved and in which engineering supports of the two communities have been brought together. From such situation we expect to hear about lessons learnt, benefits and limits of such approaches.

REFERENCES

1. Ajaj, R., Jacquemin, C., and Vernier, F. RVDT: a design space for multiple input devices, multipleviews and multiple display surfaces combination. *In Proc. of ICMI-MLMI'09*, ACM (2009), 269–276.
2. Annett, J. Hierarchical Task Analysis. In D. Diaper and N. A. Stanton, eds., *The Handbook of Task Analysis for Human-Computer Interaction*. 2004, 667.
3. Avouac, P.-A., Lalanda, P., and Nigay, L. Autonomic management of multimodal interaction: DynaMo in action. *In Proc. of EICS '12*, ACM Press (2012), 35–44.
4. Avrahami, D., Wobbrock, J.O., and Izadi, S. Portico: tangible interaction on and around a tablet. *In Proc. of UIST '11*, ACM Press (2011), 347–356.

5. Bowman, D.A., Kruijff, E., LaViola, J.J., and Poupyrev, I. *3D User Interfaces: Theory and Practice*. 2004.
6. Christou, C., Angus, C., Loscos, C., Dettori, A., and Roussou, M. A versatile large-scale multimodal VR system for cultural heritage visualization. *In Proc. of VRST '06*, ACM Press (2006), 133–140.
7. Coutrix, C. and Nigay, L. Mixed Reality: A model of Mixed Interaction. *In Proc. of AVI '06*, ACM Press (2006), 43–50.
8. Dubois, E., Gray, P.D., and Nigay, L. ASUR++: A Design Notation for Mobile Mixed Systems. *In Proc. of MobileHCI'02*, Springer Verlag (2002), 123–139.
9. Figueroa, P. and Castro, D. A reusable library of 3D interaction techniques. *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, IEEE (2011), 3–10.
10. Figueroa, P., Dachselt, R., and Lindt, I. A Conceptual Model and Specification Language for Mixed Reality Interface Components. *VR 2006, In Proc. of the Workshop "Specification of Mixed Reality User Interfaces: Approaches, Languages, Standardization"*, 4–11.
11. Hornecker, E. Interactions around a contextually embedded system. *In Proc. of TEI '10*, ACM Press (2010), 169–176.
12. Huang, C.-R., Chen, C.-S., and Chung, P.-C. Tangible photorealistic virtual museum. *IEEE Computer Graphics and Applications 25*, 1 (2005), 15–17.
13. Ishii, H., Wisneski, C., Brave, S., et al. ambientROOM: integrating ambient media with architectural space. *CHI 98 conference summary on Human factors in computing systems - CHI '98*, ACM Press (1998), 173–174.
14. Latoschik, M.E. A user interface framework for multimodal VR interactions. *In Proc. of ICMI '05*, ACM Press (2005), 76–83.
15. LaViola, J.J. and Keefe, D.F. 3D spatial interaction. *ACM SIGGRAPH 2011 Courses on - SIGGRAPH '11*, ACM Press (2011), 1–75.
16. Lee, D., Hwang, J.-I., Kim, G.J., and Ahn, S.C. 3D interaction using mobile device on 3D environments with large screen. *In Proc. of MobileHCI '11*, ACM Press (2011), 575-580.
17. Lee, J.Y., Seo, D.W., and Rhee, G.W. Tangible authoring of 3D virtual scenes in dynamic augmented reality environment. *Computers in Industry 62*, 1 (2011), 107–119.
18. Mackinlay, J., Card, S., and Robertson, G. A Semantic Analysis of the Design Space of Input Devices. *Human-Computer Interaction 5*, 2 (1990), 145–190.
19. Manches, A., O'Malley, C., and Benford, S. Physical manipulation: evaluating the potential for tangible designs. *In Proc. of TEI '09*, ACM Press (2009), 77–84.
20. Nigay, L. and Coutaz, J. A design space for multimodal systems: concurrent processing and data fusion. ACM (1993), 172–178.
21. Pecchioli, L., Carrozzino, M., Mohamed, F., Bergamasco, M., and Kolbe, T.H. ISEE: Information access through the navigation of a 3D interactive environment. *Journal of Cultural Heritage 12*, 3 (2011), 287–294.
22. Rasmussen, J. Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering. (1986).
23. Ray, A. and Bowman, D.A. Towards a system for reusable 3D interaction techniques. *In Proc. of VRST '07*, ACM Press (2007), 187–190.
24. Shaer, O. and Hornecker, E. Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends® in Human–Computer Interaction 3*, 1-2 (2009), 1–137.
25. Strauss, P.S. and Carey, R. An object-oriented 3D graphics toolkit. *ACM SIGGRAPH'92 26*, 2 (1992), 341–349.
26. Tan, D.S., Gergle, D., Scupelli, P., and Pausch, R. Physically large displays improve performance on spatial tasks. *ACM Transactions on Computer-Human Interaction 13*, 1 (2006), 71–99.
27. Ullmer, B., Ishii, H., and Jacob, R.J.K. Token+constraint systems for tangible interaction with digital information. *ACM TOCHI 12*, 1 (2005), 81–118.
28. Valkov, D., Bolte, B., Bruder, G., and Steinicke, F. Viargo - A generic virtual reality interaction library. *2012 5th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS)*, Ieee (2012), 23–28.
29. Williams, C., et al., P. TZee: exploiting the lighting properties of multi-touch tabletops for tangible 3d interactions. *In Proc. of CHI '11*, ACM Press (2011), 1363–1372.
30. Wischgoll, T. and Meyer, J. An explorational exhibit of a pig's heart. *ACM SIGGRAPH 2005 Posters on - SIGGRAPH '05*, ACM Press (2005), 138.
31. Ying, J. An approach to Petri net based formal modeling of user interactions from X3D content. *In Proc. of Web3D '06*, ACM Press (2006), 153–157.
32. Unity 3D - Game Engine. 2012. http://unity3d.com/. Accessed: April 2014
33. Web3D Consortium : X3D for Developers. 2012. http://www.web3d.org/x3d/. Accessed: April 2014
34. Irrlicht Engine - A free open source 3D engine. 2014. http://irrlicht.sourceforge.net/. Accessed: April 2014