HAL
open science

Creating 3D Interaction Technique Empirical Evaluation with the use of a Knowledge Database of Interaction Experiments

Christophe Domingues, Samir Otmane, Frédéric Davesne, Malik Mallem

To cite this version:

Christophe Domingues, Samir Otmane, Frédéric Davesne, Malik Mallem. Creating 3D Interaction Technique Empirical Evaluation with the use of a Knowledge Database of Interaction Experiments. IEEE Conference on Human System Interaction (HSI 2008), May 2008, Krakow, Poland. pp.170–175, 10.1109/HSI.2008.4581428 . hal-00339458

HAL Id: hal-00339458

https://hal.science/hal-00339458

Submitted on 30 Nov 2008

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.

Creating 3D Interaction Technique Empirical Evaluation with the use of a Knowledge Database of Interaction Experiments

C. Domingues†, S. Otmane†, F. Davesne† and M. Mallem†
†IBISC Laboratory, University of Evry – CNRS FRE-3190, France

Abstract — **Designing usable and effective 3D Interaction Technique (3D IT) is very challenging for system developers and human factors specialists. Indeed, time consuming empirical evaluation is necessary to have an idea about the goodness of 3D IT at the end of its development lifecycle.**

We have designed and implemented a new system called Empirical Evaluation Assistant to rapidly gather significant feedbacks about the usability of 3D IT during its development lifecycle. Thus, it may be possible to enhance iteratively the 3D IT before it would be evaluated by ergonomics experts at the end of the development.

EEA is based on a database where experiments results and evaluation tasks scenario are used to permit to give assistance to the experimenter during the creation of the evaluation and to automate some tasks (e.g. selecting useful questionnaires).

EEA has been used to gather results about a 3D IT developed in our research laboratory which is still under study. Results show that EEA has permitted to redesign some characteristics of this 3D IT.

Keywords — **3D Interaction, Database, Empirical evaluation, Evaluation tool, Human computer interaction, Virtual Reality.**

I. Introduction

In our laboratory we have been creating interaction models and techniques for our semi-immersive Virtual Reality/Augmented Reality platform, especially dedicated to robot teleoperation and collaborative telework. The most important constraint we are facing is the usability of our techniques. In the teleoperation case, IT must be user-friendly and usable to prevent from damages and to easily manipulate the distant robot.

However, seeking the best 3D Interaction Technique (3D IT) and Interaction Method (IT) requires a long period of time of usability testing.

Moreover, rapid changes in hardware capabilities, devices, and lack of mature methodology in interaction design [1] involve problems to design user-friendly and usable interfaces.

No true guidelines exist to remove these constraints when building and implementing an interface on a VR/AR platform. Thus, the only choice we had was to validate our IT by ergonomics experts at the end of their development lifecycle. But this validating phase takes a long time and if it appears at the end that the result is poor, the validation feedback comes too late.

Our idea is to build an Empirical Evaluation Assistant (EEA) in order to perform light evaluations of 3D IT without the need of being an ergonomics expert.

We want this tool to:

- Bring assistance to fasten the design of the validating experiments by using information collected in a dedicated database;
- Bring fast feedbacks about a tested 3D IT during the validation experiments;
- Collect data and enrich the database to increase the knowledge about 3D IT behaviors.

We hope the EEA system may improve the quality of our 3D IT, leading in most cases to positive final evaluation by ergonomics experts.

This paper is structured as follows. Section 2 will briefly review the classical techniques of ergonomic evaluations. The EEA system is described in section 3. And finally, section 4 gives some insights about the feedback given by EEA.

II. Related Work

Seeking the best 3D Interaction Technique for a given application requires numerous and long ergonomic evaluations. Indeed a lot of ITs for Virtual Environment (VE) have been developed without specifying requirements for specific applications, although it is a necessary step in order to create intuitive and transparent interaction for final users [2].

ITs for VE are totally different from 2D IT that are typically used with a keyboard and a mouse to manipulate a graphical interface (WIMP paradigm). In this case, there are a lot of guidelines, principles or predictive models (e.g. Fitt's law or KLM model) to build an effective 2D IT.

However, it is not the case for 3D IT. The main reasons are rapid changes in hardware capabilities, devices, fewer experts and the lack of mature methodology in interaction design (no strong models) [1]-[3]. However, many authors as Wingrave [4] and Ray [5] started to propose specifications for building 3D IT.

Nevertheless, methodologies, 2D laws extension, guidelines or principles are emerging for VE such as usability engineering methodology [6] that focuses only on

VR application but not on 3D IT, principles and guidelines for VE [7] , however they are taken from experimenter experience rather than from empirical results [3] or Fitt's law extension [8] for VE. Moreover, there are evaluation tools using guidelines and principles such as MAUVE system [9] which provides a structured approach for achieving usability in VE system design and evaluation. Those evaluations methods are based on analytical approaches. They compare the behavior of the interaction to a reference model which describes the necessary conditions to obtain an efficient interaction.

Unfortunately, the VR domain is not as mature as the 2D domain due to a lack of norms and ergonomics experience, analytical approaches cannot be used to evaluate 3DIT [3] consequently, empirical approach must be used. However this approach is complex to perform due to many difficulties: large list of parameters such as subjects' profile, subjects' questionnaires or the design of tasks scenarios.

But is not all, there is huge list of performance metrics and outside factors pointed out in [2]-[3] where they are gathered in 4 categories. In [10]-[11], the authors have classified the elements especially for Selection/Manipulation tasks. More particularly in reference [12], the author focuses only on factors for vision or system command tasks [13].

In consequence, we propose to include all the knowledge of interaction experiments in a database and build a tool that allows creating empirical evaluation quickly and easily in order to iteratively test and improve the design of 3D IT.

III. The Empirical Evaluation Assistant

A. *Motivations and objectives*

Excluding problems pointed out in the related work we have encountered many problems (e.g. long period of time for testing experiments, software implementation and VE specification) during the development and tests of our 3D IT. So, we propose a tool that permits to:

- **Save evaluation time:**
  - By adjusting 3D IT outside problems such as Virtual Environment design or VR devices integration in order to get ideal conditions. This stage permits to eliminate factors in the system and only studies the 3D IT;
  - By extracting a limited list of outside factors that may influence selected performance metrics in order to find a compromise between the total duration of an experiment and its soundness.
- **Share experimental results:**
  - By permitting replicability of experiments;
  - By permitting collaborative work with 3DUI developers and evaluation experts.
- **Centralize 3D IT knowledge in a database.**

EEA is intended to be used during the development lifecycle of 3D IT by non experts of ergonomics. Typically, it is dedicated to 3D IT developers. It has no aim to bypass a complete evaluation process made by ergonomics experts. We propose to divide the evaluation process timeline in three stages (see figure 1) where stages part correspond to specific objectives.

Stage 1 is an iterative debug stage where the experimenter, may configure the hardware parameters, adjust the 3D IT software (e.g. technique internal parameters) and improve the VE specification (e.g. adjusting obstacles for a navigation task).

In stage 2, we are iteratively testing/improving the design of the 3D IT with voluntary subjects. This stage is split in three parts (before, during and after the experiment) and corresponds to different objectives.
Before the experiment, the system must provide assistance for the preparation of the experiment:

- Auto generation of qualitative questionnaires;
- Help to select useful metrics and factors to be used during the experiment;
- Help to design task scenarios.

During the experiment, the system must permit to build a trace of selected quantitative metrics during the whole experiment.
After the experiment, the experimenter can perform statistical analysis over qualitative and quantitative data.

Finally, **the stage 3** is a complete evaluation done by ergonomics experts.

During all stages, the system must permit collaborative work:

- To enrich the Knowledge Database;
- To share results with experts and 3D IT developers;
- To annotate and compare experiments.

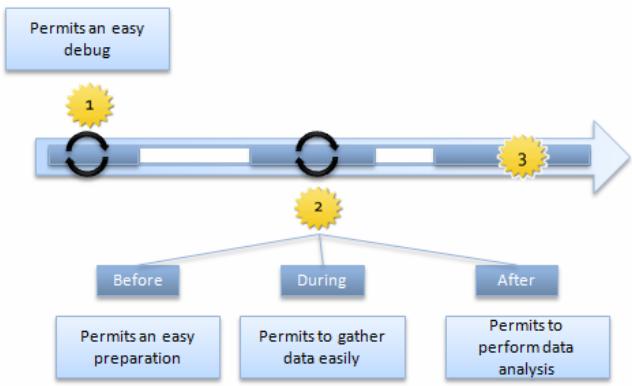

Fig.1. The three stages to obtain a finalized and usable 3D IT using EEA system. Stage 1 is an iterative debug stage. Stage 2 is an iterative testing/improving stage.

B. *System Features*

1) *The Experimental Protocol Conception (EPC) Tool*

In order to achieve the specifications of the EEA system, we have built two distinct tools. The first tool is dedicated to Experimental Protocol Conception, which we call EPC tool. EPC is composed of four main modules and a Knowledge Database (see figure 2 for the KD components).

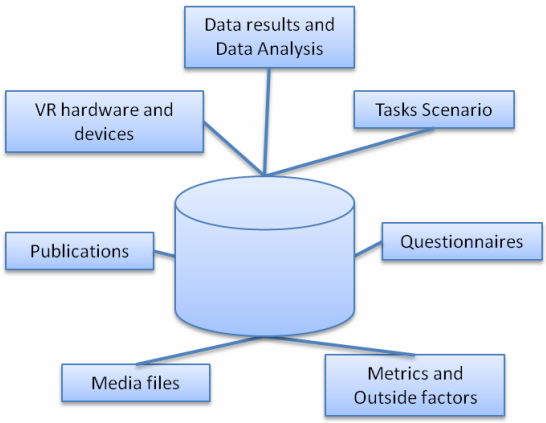

Fig.2. Composition of the Knowledge Database

To design tasks scenario, the experimenter uses the Scenario Conception Module (SCM) that allows:

- An assistance to minimize meaningful outside factors that may influence selected performance metrics. The system automatically excludes factors and metrics according to the VR hardware selected by the experimenter (e.g. no selection of stereo glasses excludes Stereo/Mono vision outside factor). To do that, we have linked hardware to many metrics and factors according to Bowman or Wingrave work ([2] [3] and [9]);
- Next, we try to link factors and metrics together to get an auto-selection of factors according to selected metrics. For the moment, this step remains semi-automated.
  Moreover the system gives the experimenter information about metrics or factors grouped by categories. This is possible due to links with KD (e.g. using publications or former experiments).
- Auto-setting of software resources in order to run the experiment. Questionnaires for qualitative results and MD tool configuration file are created according to selected elements;
- The task scenario is automatically created, stored and will be attached at the future results (see figure 3 for an illustration).

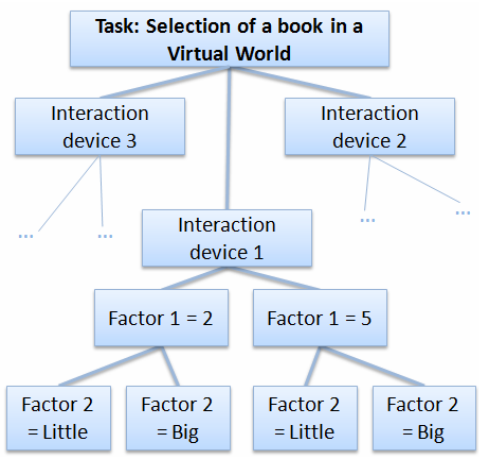

Fig.3. A task scenario where 3 factors are tested: Book size (factor 2), book-avatar distance (Factor 1) and three interaction devices.

Statistical analysis is done by the Data Analysis Module (DAM). It allows the experimenter to perform different kinds of statistical analysis. During step 1 (stage 1), the experimenter may perform a quantitative analysis or linear regression in real time. Consequently, he may perform quick tests to adjust VE and 3D IT internal parameters.

During stage 2, the experimenter may perform analysis on the data traced during the experiment relatively to selected metrics and factors (e.g. inferential analysis as ANOVA, Student's t-test) and the evaluation with pre-selected people An EPC module which is called Collaborative Module (CM) permits to enrich the KD. It allows two main possibilities: share KD between different 3D IT developers to Consult/Add/Modify/ data (e.g. to add metrics) and utilize SCM with other experimenters to design efficient tasks scenario.

2) *The Measurement and Debugging Tool*

The second tool is dedicated to Measurements and Debugging, which we call MD tool. MD tool includes a debug module that allows experimenter to view in real time all quantitative metrics available on EPC tool (during stage 1). However, all available statistical tools cannot be used (e.g. inferential or qualitative analysis) and results are not stored in KD. During stage 2, all the statistical analysis tools are available due to the pre-selected metrics and results are stored in the KD. This stage permits to evaluate 3D IT with voluntary subjects. Storing the results permits to get evaluation traces and share results with others 3DUI developers.

C. *Implementation of EEA System*

1) *Hardware Context*

In our laboratory, we own a semi-immersive VR/AR platform which enables stereoscopic display, wireless hand/head/fingers tracking and force feedback. Each device is associated to a specific server which is accessed by clients via the VRPN library. The interactivity between the user and the VE is done by using Virtools as a front-end. We have chosen Virtools and VRPN library for two main reasons. First, our aim is to provide a tool that permits to share experimental results over the community. These tools are considered as standards and the experiments have a better reproducibility. Using these tools permits to create a standardized collection of interaction experiments to search for replicability.

Second, Virtools is suitable software for prototyping and testing our 3D IT because it offers a fast and graphical way to compute and link them with devices and VEs by connecting specific building blocks to each other.

2) *Software architecture*

The MD tool has been implemented by making specific Virtools blocks that we have called Probes and a master block called MD main module. Probes objectives are to retrieve quantitative data from VR devices, VE and subjects' tasks (e.g. navigation time, selection errors or system frame rate). Probes may be connected to building

blocks where the output has to be measured, traced or displayed in real time.

MD main module permits to initialize the measurement schema of all pre-selected quantitative metrics by using a configuration file (a XML document) created by the EPC. It sends synchronization signals to the probes (e.g. start/stop/pause signals) in order to gather data. This is done by the "Synchronization" and "Wait" modules.

When a probe is enabled and the synchronization done, the appropriate module (e.g. Distance Module) is started/stopped/paused. At the end of evaluation, MD main module transfers data to DAM of EPC to perform statistical analysis and store results into KD. The Input/Output module permits to retrieve data from different devices (Optical tracking, DataGloves, Spidar and Wireless accelerometers) and can set stereoscopy mode and audio outputs.

Moreover, the debug module (used on the debug stage) allows the experimenter to view metrics in real time (e.g. position, speed or time to reach a position in the 3D space) by using the probes. This stage is generally done before the empirical evaluation (stage 2). In the debug tool case, MD main module does not require a configuration file, all probes are enabled. However, results are actually not stored in KD.

There are two links between EPC (XML document) and MD (Data results) as we can see on the global architecture and software implementation (figure 4).

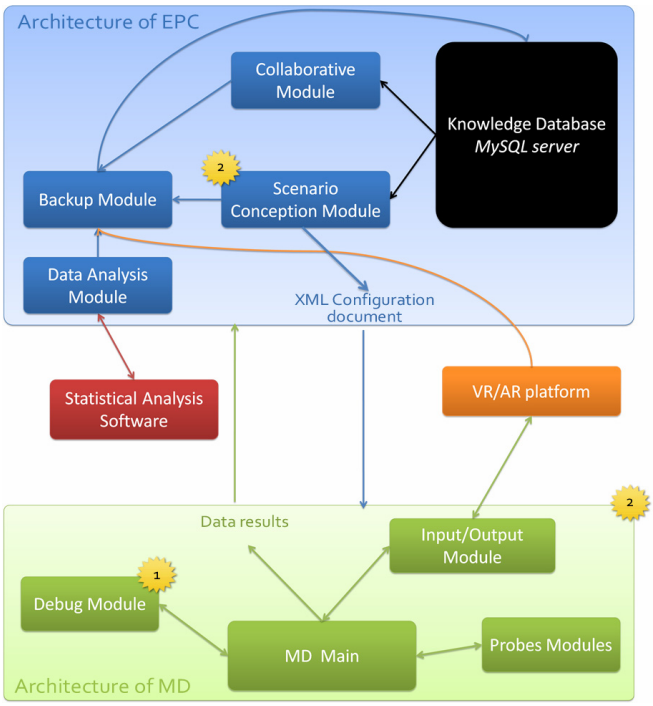

Fig.4. Software architecture of the EEA system. Debug tool is only used during stage 1. During stage 2, the experimenter uses SCM (EPC) and MD tools.

The EPC tool permits the access to the KD. We have chosen a WEB based architecture centered on an Apache 2 server. The database is implemented with a MySQL server which is accessed via SQL queries from the EPC tool written in PHP and AJAX. EPC creates the configuration file (XML document) that permits to MD main module to initialize the measurement schema.

We have chosen these tools because:

- They do not require plug-in like Java Applets or Macromedia Flash and work with all internet browsers. Moreover using PHP/MySQL permits an easy access over internet. However, Java EE is a good solution and similar to PHP solution but we prefer using the second;
- They are open source and used by a lot of web site.

3) The Knowledge Database

The Knowledge Database contains eight tables. Each table is composed of different attributes (e.g. name or type). The main table is called "Experimental Protocols" and is linked with the others table. The figure 5shows the relational model of our database. Tables are represented in blue. Relations between tables are symbolized in. Each relation is linked by one key.

Tables' composition:

- **Metrics/Factors table:** This table contains information about metrics and factors. This table is linked with "Questionnaire" because some metrics need questions to gather them, with publications to get the information of metric and with "VR devices" because some metrics are specific to VR devices. For example, if the experimenter selects "Age" as a metric (with EPC tool), the system will auto-selects the appropriate questions.

  New metrics/factors are added manually by experimenters.
- **Experimental protocols table:** This table contains the tasks scenario created with EPC tool. A task scenario incorporates different elements used for the evaluation (e.g. metrics, factors or devices used). This table is linked with all the other tables.
  A new task scenario is only added by the system during its creation with EPC tool.
- **Data Results:** This table is composed of two categories of data: the evaluation results and statistical analysis perform on the results. Results and analysis are linked with a unique protocol.
  Moreover, access permissions to results and data are stored in this table. The experimenter assigns access to selected users and groups to make his data public or private.
- **VR Devices:** This table is composed of information about VR devices and hardware in order to permit replicability of experiments and have comparison with others experiments. Moreover, it permits to prevent from impossible combinations due to its link with metrics table. When adding devices, the experimenter must create links between new device and metrics which are already in the database.

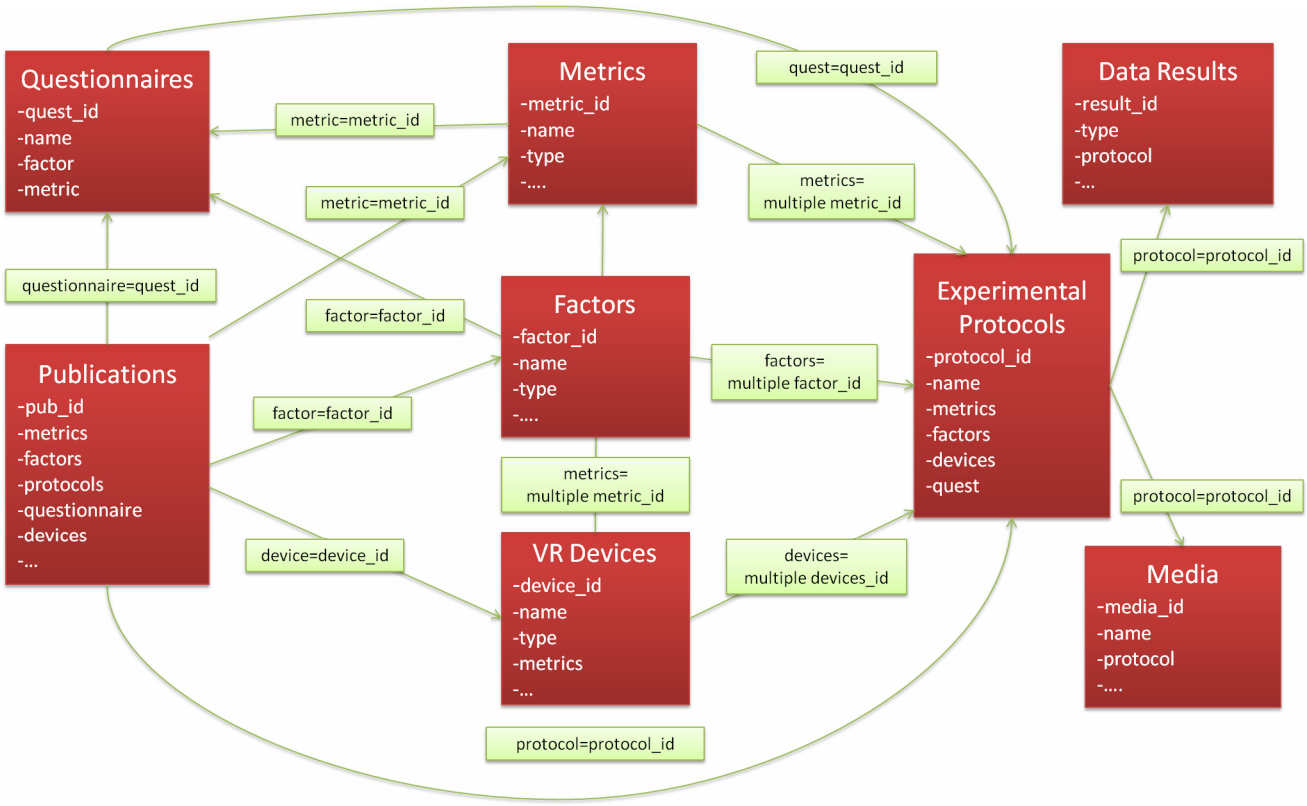

Fig.5. The relational model of the Knowledge Database.

IV. Using EEA System through an example

A. *Setting up the evaluation*

During the debug stage, the experimenter will adjusts the internal parameters of the Follow-Me technique (e.g. size or type of virtual guides) and the Virtual Environment (e.g. books position and size). To do that, the experimenter will only use the MD tool with connected probes (connections are done by the experimenter).

The second stage consists on creating the task scenario. Here, our task is to select and manipulate books as fast as possible.

To do that an experimenter must follow three steps to set up the evaluation using the EPC tool:

Firstly, the experimenter selects VR devices which evaluation task scenario we have selected: Active stereo capability, Flystick 1, the SPIDAR and the Data Gloves were not used, so they are not been selected).

This step is necessary in order to permit replicability of the experiment and to avoid impossible combinations according hardware.

Secondly, the experimenter selects remaining metrics in the list and desired factors according his environment and what he wants to test. All metrics and factors are given with applications of use and "help message". Metrics selected for the evaluation are: selection and manipulation time, selection and manipulation mistakes, subjects' age, experience, sex etc. Factors selected are: books size, use of stereo vision, distance between books and avatar initial position. The system will creates the required automated resources and saves all in the Knowledge Database. It produces questionnaires for qualitative results and a list of tasks to give to voluntary subjects.

Finally, the experimenter uses the MD tool to put probes in his Virtools Script to retrieve data. This step is similar to the debug stage. But here, only given probes, specified in the configuration file are enabled. Moreover, data results will be stored.

B. *Feedbacks after the preliminary evaluation*

We have performed a comparative evaluation of Follow-Me and two other classical 3D IT (HOMER [2] and Go-Go [16]) over 15 voluntary subjects.

Two days of work for one experimenter were necessary to:

- Adjust the internal parameters of the three 3D Interaction Technique and parameters of the virtual environment with the use of the debug module;
- Build and implement the experimental protocol depending on the questions we were asking [the EPC tool configures the probes and deliver questionnaires];
- Perform the experiment in itself with 15 voluntary subjects (an average of 30 minutes per user was necessary) [MD tool produces a dated trace of all probes];
- Analyze the collected data to produce a feedback (dated trace and qualitative data from questionnaires are submitted to DAM to perform an ANOVA).

EEA permitted us to know that Follow Me is favorably accepted by novices in VE and permits faster selection and manipulation that other 3D IT whereas experts are puzzled by *Follow-Me* and prefers classical 3D IT (Go-Go and HOMER). This feedback will be utilized in the future to refine the use of virtual guides in the Follow Me model. Moreover EEA will be reuse to see if the score of metrics is better.

V. Conclusion

The EEA system is dedicated to light evaluations of 3D Interaction Techniques during the development lifecycle. However, this tool is not limited to 3D IT and can be used for others kinds of systems. It may be used by non ergonomic experts.

The aims of this tool are:

- To offer design and trace facilities;
- A fast design of an evaluation by assisted selection of useful metrics and factors;
- To perform statistical analysis;
- To recursively enrich a Knowledge Database that can be used for the future experiments.

The core idea is to get fast feedbacks in order to improve the tested 3D IT. In order to accumulate knowledge about 3D ITs, the whole experiments may be stored in the Knowledge Database that may be accessed worldwide via a WEB interface, whereas the debug tool is connected to our VR/AR platform and it is implemented under Virtools. We have used EEA to test the 3D IT we have developed recently and the feedback we obtained in only two working days permitted us to upgrade the Follow-Me model.

However, the major problem of this approach is to populate the Knowledge Database (KD).

EEA system has been used to redesign conditions of use of 3D IT that we have developed, called Follow-Me [14]. The model had been roughly tested before but without the EEA system [15]. The characteristic of Follow-Me is, its use of virtual guides to reduce the number of degrees of freedom in VE when approaching an object to be selected and manipulated. Thus the system anticipates what the user may do to reduce subjects' concentration. We have performed a comparative evaluation of Follow-Me with two others 3D IT (HOMER and Go-Go) over 15 voluntary subjects. Each user had to select a book on a shelf with a Flystick and put it on another shelf as fast as possible. EEA permitted us to know that Follow-Me is favorably accepted by novices in VE and permits faster selection and manipulation that other 3D IT whereas experts are puzzled by Follow-Me and prefers classical 3D IT. This feedback will be utilized in the future to refine the use of virtual guides in the Follow-Me model.

Future work on the EEA will concentrate on:

- Continuing the development of our system;
- Removing the probes to get an autonomous system with pre-configured settings using EPC;
- Collaboration with ergonomic experts or others 3D IT developers to improve our system;
- Enrich the Knowledge Database to create a large collection of Interaction experiments;
- Related work on data mining techniques to find and extract useful information to enrich automatically the Knowledge Database;
- Allow creation and backup of VE using Virtools supported formats in order to have a standardized collection of VEs to replicate/compare 3D IT evaluations.

VI. References

[1] A. S. Rizzo, J. K. Gerard, S. Yeh, M. Thiebaux, J. Hwang. and J. G. Buckwalter, "Development of a Benchmarking Scenario for Testing 3D User Interface Devices and Interaction Methods," I*n Proc. of the 11th Int. Conference on Human Computer Interaction*, Japan, 2005.

[2] D. A. Bowman, E. Kruijff, J. Laviola and I Poupyrev "3D User Interfaces theory and practice", *Ed. Addison Wesley*, 2004.

[3] D. A. Bowman, D. Johnson and L. Hodges "Testbed Evaluation of VE Interaction Techniques," *In Proc. of the 1999 ACM Symposium on Virtual Reality Software and Technology, 1999*, pp. 26-33.

[4] C. Wingrave, "Understanding the Limits and Benefits of a 3D Interface Specification," *In Conf. IEEE VR Workshop on Specification of Mixed Reality User Interfaces,* 2006.

[5] A. Ray and D. A. Bowman, "Towards a System for Reusable 3D Interaction Techniques," *In Proc. of the 2007 ACM Symposium on Virtual Reality Software and Technology*, Newport Beach, California, 2007, pp. 187-190.

[6] J. L. Gabbard, D. Hix and J. E. II Swan "User-Centered Design and Evaluation of Virtual Environments," *In conf. IEEE Computer Graphics and Applications*, 1999, pp. 51-59.

[7] K. Kaur, "Designing virtual environments for usability," Ph. D. dissertation, University College London, 1998.

[8] S. Mackenzie and W. Buxton, "Extending Fitts' law to two-dimensional task," *In Proc. of the SIGCHI conference on Human factors in computing systems,* 1992, pp. 219-226

[9] K. M. Stanney, M. Mollaghasemi, L. Reeves, R. Breaux and D. A. Graeber, "Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design," *Int. J. Human-Computer Studies 58,* Elsevier Science, 2003, pp. 447–481.

[10] I. Poupyrev, S. Weghorst, M. Billinghurst and T. Ichikawa, "A Framework and Testbed for Studying Manipulation Techniques for Immersive VR," *In Proc. of ACM Virtual Reality Systems and Technology Conference*, 1997.

[11] C. Wingrave and D. A. Bowman, "Baseline Factors for Raycasting Selection," *In Proc of Virtual Reality Int.,* 2005, 10 pages (CD-ROM proceedings).

[12] T. Ni, D.A. Bowman and J. Chen, "Increased Display Size and Resolution Improve Task Performance in Information-Rich Virtual Environments," *In Proc. of Graphics Interface*, 2006.

[13] R.P. McMahan, and D. A. Bowman, "An Empirical Comparison of Task Sequences for Immersive Virtual Environments," *IEEE Symposium on 3D User Interfaces*, 2007.

[14] N. Ouramdane, F. Davesne , S. Otmane and M Mallem M."Follow-Me: A new 3D interaction technique based on virtual guides and granularity of interaction," *In Proc. of ACM International Conference on Virtual Reality Continuum and Its Applications,* Hong Kong, China, 2006, pp. 137-144.

[15] N. Ouramdane, F. Davesne, S. Otmane and M. Mallem, "Evaluation of the Follow-Me technique for grabbing virtual objects in semi-immersive virtual environment," *In Conf. IEEE 8th Int. Conference on Virtual Reality*, 26-28 April 2006, pp. 85–94.

[16] I. Poupyrev, M. Billinghurst, S. Weghorst and T. Ichikawa, "The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation," *In Proc. of the 9th annual ACM symposium on User interface software and technology*, 1996, pp. 7980.