3DUI-EF: Towards a Framework for Easy Empirical Evaluation of 3D User Interfaces and Interaction Techniques

Christophe Domingues, Samir Otmane, Malik Mallem

To cite this version:


HAL Id: hal-00450311
https://hal.science/hal-00450311v2
Submitted on 14 Feb 2010

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.
3DUI-EF: Towards a Framework for Easy Empirical Evaluation of 3D User Interfaces and Interaction Techniques

Christophe Domingues, Samir Otmane and Malik Mallem

IBISC Laboratory – University of Evry, France.

Abstract— Designing usable and effective 3D User Interfaces and 3D Interaction Techniques is very challenging for Virtual Reality system developers and human factors specialists. Indeed, time consuming empirical evaluation is necessary to have an idea about the goodness of the 3D User Interface (3DUI) and the 3D Interaction Technique (3DIT) at the end of their development lifecycle. This may induce a huge loss of time if the result appears not to be satisfying in the end. Moreover, 3DUI evaluation is much more complex than 2D User Interfaces evaluation which is due to heterogeneous Virtual Reality (VR) devices and 3DIT.

The aim of this work is to provide a framework allowing developers and experimenters to quickly evaluate 3DUIs and 3DITs during the design and the development lifecycle. The proposed framework is divided into two tools. The first one enables to create an evaluation protocol based on a knowledge database using two data mining algorithms, the “C4.5” to avoid from impossible combinations between devices and indicators and the “Spv Assoc Tree” to build a decision tree between indicators and factors. The second tool of the framework is an Evaluation Virtual Environment (EVE) to perform the evaluation according the protocol created with the first tool.

Index Terms— Human computer interaction; Empirical evaluation; Evaluation tool; 3D User Interface; 3D Interaction Technique Database; Virtual Reality; Augmented Reality.

I. INTRODUCTION

As the technology of display and graphics systems has developed, Virtual Environments (VEs) applications have come into common use outside the research laboratory [1]. The technology of VEs offers to the user new interfaces which enable him to interact easily and naturally with the VE. Currently, interaction is one of the main problems of the majority of applications in this field. There are a lot of varieties of existing 3DITs, which attempt to solve the problem of grabbing and manipulating objects in VEs (for example, see [2, 3, 4, 5, 6]). However their exists numerous problems for designing user-friendly and efficient 3DUIs (see Rizzo et al. [7] and Wingrave [8]). It may be explained by the rapid changes in hardware capabilities, the multiplicity and heterogeneity of VR devices, and the lack of mature methodology in interaction design. Indeed, no established guidelines can guaranty the soundness of the building and implementation of a 3DIT for Virtual Reality (VR) or Augmented Reality (AR) environments. Thus, until now, the only choice has been to validate 3DUIs and 3DITs 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.

For all these reasons we propose in this paper a framework called 3DUI-EF (3D User Interface Evaluation Framework). This framework has no aim to bypass a complete evaluation process made by ergonomics experts but it aims to:
- Bring assistance to fasten the preparation of the validating experiments and to analyze data collected during the experiment;
- Bring fast feedbacks about a tested 3DUI and 3DIT during the validation experiments;
- Use past collected data to enrich the knowledge about 3DUI and 3DIT behaviors;

This paper is structured as follows. Section 2 will briefly review the classical kinds of ergonomic evaluations. The presentation of the 3DUI-EF is developed in section 3. Section 4 will present a case study applying the proposed Framework to evaluate an interaction technique. Conclusion and future work are given in section 5.

II. RELATED WORK

Looking for the adequate 3D User Interface for a given application requires time consuming ergonomic evaluations. Indeed most of 3DUI for Virtual Environment (VE) have been developed without meeting requirements for specific applications, although it is a necessary step in order to create intuitive 3DUI for final users (Bowman et al. [9]).

3DUI for VE are totally different from 2DUI which 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) that help in building an effective 2DUI.
But it is not the case for 3DUI. The main reasons are: 1) rapid changes in hardware capabilities; 2) many heterogeneous devices; 3) few experts; 4) lack of mature methodology in interaction design (no strong models) [7,10]. Moreover, some 3D interaction techniques have been designed only for a specific VR device as a navigation technique using pinch gloves (Bowman et al. [11]). Nevertheless, methodologies, 2D laws extension, guidelines or principles are emerging for VEs like:

- Usability engineering methodology (Gabbard et al. [12]), which only focuses on VR application without dealing with 3D IT;
- Principles and guidelines for VEs (Kaur et al. [13]), which are taken from the experimenter knowledge rather than from empirical results (10);
- Fitt’s law extension (Mackenzie et al. [14]) for VE.

Besides, there are tools like the MAUVE system (Stanney et al. [15]) which are based on analytical approaches. They provide a structured approach to achieve usability in VE system design and evaluation. They compare the behavior of the 3DUI with a reference model, which describes the conditions to obtain an efficient User Interface. Unfortunately, the VR domain is not as mature as 2D desktop domain. So, analytical approaches cannot be used to evaluate 3DUI because of a lack of norms and ergonomics experience.

Hence empirical approach must be used to measure the skill of different users using the IT in the VE [10]. However, empirical evaluations are easy to perform due to many difficulties: large list of parameters like users’ profile, users’ questionnaires, conception of scenario and particularly the huge list of performance metrics and outside factors pointed out in [9,10,16,17] for 3D Interaction tasks. Dünser et al. [18] have written a technical report in which they classify publications which include an AR evaluation, sorted by evaluation method. For them, the main reason for the lack of user evaluations in AR might be that evaluation is too complex and there are no tools to assist experimenters during this process. Moreover, collaborative interaction brings more problems for evaluating techniques [30, 31].

The results of these different reports confirm that empirical measurements methods (objective and subjective) are most commonly used. Objective methods are studies that include objective measurements such as: completion times, accuracy/error rates and generally, statistical analyses are made on the measured variables. Subjective methods are studies in which questionnaires are used to perform the analysis. Usually, we found these measurements in many evaluations [21, 22, 23, 24, 25, 26, 27].

Fig. 2 shows the number of novel 3D interaction techniques by year. This graphic shows that this number is decreasing since 1995. One of the possible explanations may be that evaluation is too complex and there are no tools to assist experimenters during this process. Moreover, collaborative interaction brings more problems for evaluating techniques [30, 31].

### III. 3DUI Evaluation Framework

#### 3.1 3DUI-Evaluation Framework objectives

3DUI-EF is designed to give experimenters a complete evaluation platform. It provides:

- An assistance to select useful elements to design task scenarios with automated process;
- An automatic generation of qualitative questionnaires;
- An assistance to select adequate statistical analysis;
- A collection of Evaluation Virtual Environment (EVE);
- A collection of existing 3D Interaction Techniques to draw comparisons between the evaluated technique and classical ones;
- Tools to gather data and to control the evaluation.

3DUI-EF is dedicated to two kinds of tests which are included in the V development cycle of the 3DIT (see Fig. 3): System Testing which is an iterative debug stage where the experimenter may configure the hardware parameters, adjust the 3D IT software (e.g. technique internal

---

**Fig. 1. Classification of publications by evaluation method**

**Fig. 2. Number of novel 3D interaction technique publications cited by year in 3D User Interface.**

---

*Extracted from [18]*
parameters) and improve the VE specifications (e.g., adjusting obstacles for a navigation task). Hence, 3DUI-EF will permit the experimenter/developer of the 3DUI to check the initial 3DUI specifications and the actual developed 3DUI. The second test is an acceptance testing where the experimenter will perform an empirical evaluation with many users.

3.2 3DUI-EF overview

3DUI-EF is divided into two distinct tools (see Fig. 4). The first tool is dedicated to Experimental Protocol Conception (EPC). The second tool is the Measurement and Debug tool (MD). MD tool permits to gather data and manage the VE environment and devices through modules. Consequently the experimenter has a complete evaluation platform. MD tool needs the XML document to be created with EPC to initialize the evaluation.

3.3 The Experimental Protocol Conception tool

EPC is intended to design tasks scenario. At the beginning, the experimenter has the possibility to choose logical implications between elements such as devices, metrics or factors (e.g., using a selection VR device will imply the use of time selection as metric).

We suggest using data mining algorithms to generalize the logical implications over successive experiments to refine the design of tasks scenarios. After multiple experiments the system can provide:

- An assistance to minimize meaningful outside factors that may influence selected performance metrics. The system automatically excludes factors and metrics according to the Virtual Reality hardware selected by the experimenter (e.g., no selection of stereo glasses excludes Stereo/Mono vision factor);
- An assistance to link factors and metrics to get a selection of factors according to selected metrics. Moreover the system gives the experimenter information about metrics or factors grouped by categories;
- Auto setting of software resources in order to run the experiment. Questionnaires for qualitative results and MD tool configuration file are automatically generated according to selected elements.

1) The Knowledge Database

The main component of the EPC is the Knowledge Database. Fig. 5 illustrates the elements which are stored in the database. Metrics, Factors and Devices are automatically linked together with the use of data mining algorithms. Table 1 illustrates how achieved evaluations are stored in the database.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Metrics</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Speed, Time, ...</td>
<td>Size, Color, ...</td>
</tr>
<tr>
<td>2</td>
<td>Time, Errors, ...</td>
<td>Size, Initial Position, ...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Fig. 5. Composition of the database

TABLE 1: EXAMPLE OF STORED EVALUATION METRICS AND FACTORS IN KD
2) An assistance to select useful parameters

The idea is to use evaluation knowledge to assist users during the conception of their own protocol. The data mining [31] is a set of algorithms and methods for the exploration and analysis of databases, to detect data.

We have chosen two rules for data exploration: the explanation and the association rules. The explanation will allow us to link the devices in the platform for virtual reality and performance indicators. The association rules will allow us to find correspondences between performance indicators and external factors to vary. For the development and testing of algorithms, we used Tanagra ¹, an open source application. We will use this to permit the user to be assisted during the metric and factors selection. This is a semi-automatic stage.

Explanation rules

The aim is to highlight the performance indicators based on devices selected by the evaluator. This step allows deleting unnecessary or impossible combinations such as the choice of indicator “force feedback measurement” if the experimenter selects stereo glasses. The rule explanation that we have created is based on the supervised learning algorithm called C4.5" [32]. The result of the C4.5 algorithm is a decision tree, which is used to create a tree for 3D Evaluation knowledge.

Association rules

The objective here is to link the performance indicators with external factors after the devices selection. We choose this because we agree the principle that people who assess with a certain parameter automatically choose a factor that will influence this indicator. The algorithm used for the construction of rules is the algorithm “Spv Assoc Tree” [33].

3) Adding new knowledge into the Database

To add new elements (metric, factor or device) in 3DUI-EF, the experimenter must write a form with the element information. At the beginning, the element will not be proposed by the system because there is no knowledge about it. After many evaluations using this element, the data mining algorithms selected for our system is able to automatically suggest this element to the experimenters. However, adding a new metric or a new factor implies the need to add a new probe or controller module to gather data or to control the factor during the evaluation. 3DUI-EF API is used to realize it. When new knowledge is added, a new decision tree and new association rules are created using the algorithms.

4) Parameterized building of a VE dedicated to evaluation

We have added in the process of creation of the evaluation, the possibility for generating a VE adapted to the evaluation based on experimenter’s choices. The experimenter has the option to choose a simple or more complex VE available in the system designed for the Interaction tasks (see Fig. 6). The system may give the possibility to specify the number of objects, their position and initial orientation and their predefined movement if needed (see Fig. 7). All the informations will be stored in the XML file, which will generate the test environment. For the 3DUI that the experimenter wants to evaluate, the experimenter will have to develop a module using the 3DUI-EF AP. The EPC will permit the experimenter to choose the elements of the test environment. The core of our system will then launch the evaluation and controller.

Fig. 6. A basic EVE to use for selection task. Objects are static.

Fig. 7. Evaluation controllers can modify the VE. Here he object controller has change colors and size of the objects because the experimenter has selected size and color as factors.

5) Data Analysis

The experimenter can perform different kinds of statistical analysis. During stage 1 (System Testing), the experimenter may perform a quantitative analysis or linear regression in real time. Consequently, he may perform tests to adjust the 3DUI or 3DIT. During stage 2 (Acceptance Testing), 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 voluntary people. At the end of the evaluation, results and analysis performed are stored in KD.

¹ http://eric.univ-lyon2.fr/~ricco/tanagra/fr/tanagra.html
3.4 Measurements 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 (System Testing).

During stage 2 (Acceptance Testing), the experimenter must use the EPC tool to design an evaluation scenario and to initialize the measurement schema. Results are stored in the KD to get evaluation traces and to share results.

MD tool is divided into five parts: Core, Classic Modules, EVE, Controllers and Probes (see Fig. 8).

![Fig. 8. Zoom on the Measurements and Debugging architecture](image)

1) Measurements and Debugging tool Core

The MD tool has been implemented by making specific Virtools blocks that we have called Probes, Controllers and a master block called MD Core. MD Core 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. MD Core also communicates with Controllers, which are designed to modify the VE (objects, colors, initials conditions). Finally, the MD Core communicates with modules designed for specific functions as the Devices Module which is designed to manage VR devices on platform during the evaluation.

2) Probes

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 Virtools building blocks for which the output has to be measured, traced or displayed in real time as an electrician using a voltmeter in an electric circuit. Each probe is a process, which communicates with the MD Core.

3) Controllers

When the experimenter is creating a new experimental protocol, the system may suggest him what factor should be used. However when the experimenter will use this factor during the evaluation, he doesn’t have a tool which helps him to change the state of the factor. So, he must adjust the factor by himself.

A controller is a module, which controls factors that will vary during the evaluation. Controllers are adjusted during the evaluation prototyping with EPC tool. For example, if an experimenter selects “Size” as a factor, the system will ask him to configure the factor with EPC (e.g. big or little). During the evaluation the controller will modify VE and objects with values set by the experimenter.

4) Classic Modules

Others Modules are elements of the system, which can be added or deleted, dedicated to perform a function during the evaluation or containing 3D Objects. For example, when an experimenter is prototyping his evaluation with EPC, he is going to choose VR devices, VE, etc. Modules will be added into the XML file and going to manage Devices or Virtual Environment. Finally, the statistics module manages the connection with Matlab to perform analysis in real time.

5) EVE Module

The EVE Module is a module, which contains a virtual environment and 3D objects to use during the evaluation. The experimenter can adjust these objects according the choice he made with EPC (objects trajectory, speed, size etc).

IV. 3DUI Evaluation Framework How To

5.1 User point of view

When a user wants to test his 3DUI, he must specify which devices he wants to use during the evaluation. The explanation rule will pre-select useful metrics to be applied to selected devices. The user can validate or select other metrics. In the second stage, the system will select factors according to selected metrics using association rules.

Third stage consists on selecting the VE and the 3DUI techniques, which the experimenter wants to compare. The experimenter can select no techniques if he only wants to test his technique. The experimenter must configure the objects according to the factors. For example if he chooses the size factor, he must specify the values of this factor.

At the end, the system will generate a questionnaire according objective metrics selected, factors and devices and a XML file where modules and EVE to load for performing the evaluation are stored. When EPC process is finished, the experimenter can launch his configured environment with XML and Modules. Figure 9 describes the process.

5.2 System point of view

1) EPC

In the first step, the system will propose the user to select VR devices System will “write” into the XML document the VR devices modules to load according selected devices validated by the experimenter. When this is done, the results of “C4.5” algorithm are used to find matches between the devices selected by the experimenter and metrics stored in
The “Spv Assoc Tree” algorithm results are used to find associations between metrics and factors. As probes, which are used to gather quantitative data, there are qualitative metrics. These metrics are gathered with questionnaires stored in the KD and manually linked with metrics and factors. If an experimenter wants to add some questions, the question will be added in Database. Next, the system will propose the experimenter to choose an Evaluation Virtual Environment and 3DUIs to compare to his 3DUI.

2) Measurements and debugging tool

The MD core will read the XML document to load adequate modules, probes, controllers and the Evaluation Virtual Environment. MD will manage the evaluation: send signals to probes; send signals to controllers to modify the factor. MD will store evaluation data and permit to display real time analysis or graphics using Matlab.

5.3 Example of a realized comparative evaluation to gather objectives measurements

During the debug stage, the experimenter will adjust the internal parameters of the 3DUI and the VE (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 to reproduce 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. Metrics selected for the evaluation are: selection and manipulation time, selection and manipulation mistakes, subjects age, skill in VR, gender, etc.. Selected factors are: book size, use of stereovision, 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 is enabled. Moreover, data results will be stored.

5.4 Study Case

We have performed a comparative evaluation of Follow-Me [34] and two other classical 3D IT (HOMER and Go-Go) 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 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

In this paper, we have presented a new framework called 3DUI-EF for quick evaluation and feedback during the V-cycle development of a 3D interaction technique. The aims of this tool are:
- to offer design and trace facilities;
- to bring fast design of an evaluation by assisted selection of useful metrics and factors;
- to easily draw comparisons between the tested technique and others techniques using developed Virtual
Environment and 3DUI;
- to perform statistical analysis;
- to recursively enrich a Knowledge Database that can be used for the future experiments.

We have proposed a system to create logical links between evaluation elements. This system uses data mining algorithms which are employed over multiple experiments. At the end of the process, the system is able to automatically suggest most probable design.

In order to accumulate knowledge about 3DUI, the whole experiments are 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.

We have used our framework in the V-cycle development of the Follow-Me technique. It has permitted us to point out some problems which have been corrected afterwards.

For future work, we are also investigating to store in the knowledge database users’ preference and skill to analyze if there is a long term learning of 3DUI skills for a particular user. This might permit to create flexible 3DUI, which could adapt the interaction according to user’s preference and skill.

VI. ACKNOWLEDGMENTS

We wish to thank “Le Conseil Général de l’Essonne”, french CNRS and the IBISC laboratory for funding this work.

REFERENCES


Christophe Domingues is a Ph.D. student in the IBISC lab at the University of Evry. Christophe received his BS and MS degrees in computer sciences from Evry University. His research focuses on collaborative work, 3d interaction, and virtual reality.

Samir Otmane received PhD degree in computer science from the University of Evry (France) in 2000. He is associate professor in computer science department of the University of Evry. He has been working on Virtual Reality and Augmented Reality applied to telerobotics (ARITI Project). He is working in the area of Virtual and Augmented Reality and CSCW. His current research efforts are focused on multimodal and collaborative 3D interaction in mixed reality environments.

Malik Mallem received a PhD degree in robotics and computer sciences from University of Paris XII, in 1990. His research deals with Augmented Reality (AR) and Collaborative Work applied to robotics and telerobotics at IBISC- Complex System Laboratory CNRS FRE 3190, Evry, France. Since 1999, he is Professor at the University of Evry and head of AR team which is implied in several French and European projects.