Preliminary Usability Evaluation of PolyMeCo:

A Visualization Based Tool for Mesh Analysis and Comparison

Beatriz Sousa Santos[1], Samuel Silva[1], Leonor Teixeira[2], Carlos Ferreira[2,3], Paulo Dias[1], Joaquim Madeira[1]

1-IEETA/DETI, University of Aveiro 2-DEGEI, University of Aveiro 3-CIO, University of Lisbon

{bss@det.ua.pt, sss@ieeta.pt, lteixeira@egi.ua.pt, carlosf@egi.ua.pt, paulo.dias@det.ua.pt, jmadeira@det.ua.pt}

Abstract

An overall description of the methods used and the results obtained in the on-going evaluation of PolyMeCo —a mesh analysis and comparison tool — is presented in this paper. We are trying to evaluate some aspects of both the user interface and the visualization techniques implemented. Heuristic evaluation, observation and querying techniques were used and produced encouraging preliminary results, which provided new ideas, as well as information that will inform the development of a more usable version o PolyMeCo, including new functionality.

Keywords: **Visualization based tool, Usability evaluation, Polygonal meshes.**

1. Introduction

Polygonal meshes assume an important role in several application areas ranging from Architecture to Computer Games. Often, mesh processing is necessary to reduce their complexity, for instance to produce models that can be easily manipulated or transmitted over a network. A wide variety of polygonal mesh processing operations exist and since each of these operations can be performed through several different methods, leading to different meshes depending on the used algorithm, it is important to evaluate the differences among the resulting meshes in order to choose the algorithm which better suits a specific situation. Such an assessment is essential not only for researchers developing new methods, as they have to fine tune them to provide better solutions than those already available, but also for developers applying these methods for particular purposes. The role of developers is also of paramount importance, since they often test or apply the mesh processing methods in unexpected conditions not considered by researchers, disclosing important results. In order to include an assessment stage in their workflow, it is important that researchers and developers have tools providing all the necessary features to allow a systematic assessment and enhanced insight into the obtained data. PolyMeCo is such a tool [1], providing an integrated environment for mesh analysis and comparison, where data can be explored using different visualization options.

In this paper we describe an on-going evaluation intended to inform the development of a new version of PolyMeCo. To help understanding the methods and results presented, we first give an overview of the philosophy, as well as the main functionalities, user interface features and visualization options available in PolyMeCo; then we present the evaluation methodology and the results obtained with the collaboration of two different types of users.

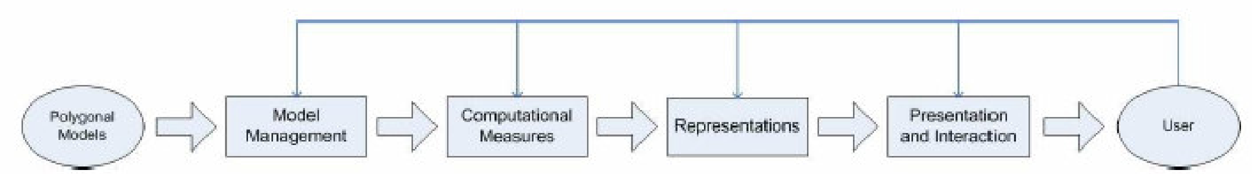

Figure 1 – Mesh analysis/comparison pipeline

2. Overview of PolyMeCo

PolyMeCo provides an integrated environment including several features not available in other mesh analysis tools [2][3][4]. These new features are important, in particular if one aims to provide a tool for systematic use by researchers and developers. The main functionalities include the possibility to:

- load several mesh models during the same work session in a clear and systematic way;
- compute different mesh features/properties (e.g., smoothness, curvature) and difference measures (e.g., geometric distance) between two meshes;
- visualize analysis/comparison data using several visualization alternatives, such as coloured models, histograms and box-plots;
- visualize the data obtained using several measures for the same mesh;
- visualize the data obtained using the same measure for different (and comparable) meshes;
- assess measure values associated with a specific mesh entity (e.g., a vertex);
- export the data obtained to allow further analysis with statistical tools.

The process of analysing and comparing meshes can be described by the pipeline presented in Fig. 1. After loading polygonal meshes, some of their features/differences are described by computational measures; then the obtained results are visually mapped to suitable representations. After this mapping has been performed, the chosen representations are presented to the user who may interact with them and change parameters along the pipeline in order to, for example, choose another computational measure, or change the way data is being mapped onto a representation. Notice that this pipeline is analogous to the Visualization pipeline presented in Card et al. [5]: the *Computational Measures* stage involves obtaining data and transforming them; the *Representations* stage corresponds to building visual structures and mapping data onto them; finally, *Presentation and Interaction* comprises all tasks regarding the user interaction with the visualization.

PolyMeCo includes three main functional modules implementing these pipeline stages, as well as a *Mesh Management* module. A detailed description of PolyMeCo and a comparison between two mesh simplification methods (performed using it) can be found in [6]; an executable version for test purposes can be downloaded from http:// www.ieeta.pt/polymeco.

Several representations (visualization techniques) are used to allow the user to visualize the measure data obtained for a given mesh; however, the fundamental technique consists in showing the measure values colour coded on the surface of the mesh (coloured model), thus allowing the user to interact with and analyse the mesh, observing the features that correspond, for instance, to maximum and minimum

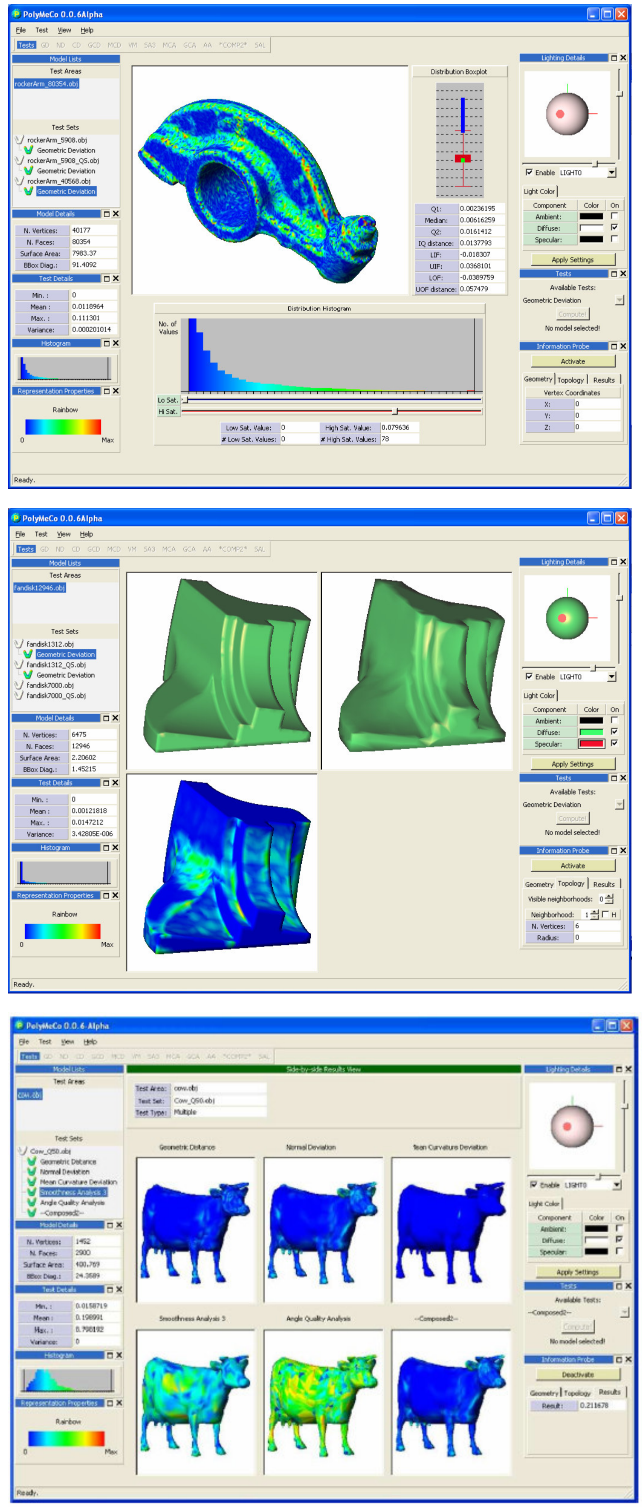

Figure 2 – Visualization modes provided by PolyMeCo to visualize mesh quality measures: a) extended results view, b) original *vs* processed *vs* coloured, c) simultaneous visualization of different measures.

values of the geometric distance as shown in Fig. 2a. Since the selection of an appropriate colour scale is a difficult issue [7], as a first step to a more sophisticated approach, several alternative scales are offered to the user (rainbow colour scale, grey scale, linearized optimal, heated-object, etc.). Although a coloured model conveys a good idea of the distribution of the computational measure values along the mesh surface, a global idea of the distribution characteristics can be difficult to obtain: to tackle this difficulty, statistical representations are used. In fact, histograms and box-plots allow the user to get, at a glance, an idea of the characteristics of the distribution, analyse ranges and the symmetries of the data and detect the presence of outliers. Even though they are useful in their own right, these representations augment the colour coded visualization of a given measure, since they allow to better adequate the colour mapping to the current data, by selecting the saturation values to be coded using the first and last colours of the colour scale and, thus, ignoring severe outliers (as shown in Fig. 2a).

In addition to the visualization modes illustrated in Fig. 2, PolyMeCo offers a feature comparison mode allowing the visualization of coloured models, histograms or box plots of data obtained with the same computational measure for several meshes, using exactly the same colour mapping. This can be very useful to study different processing algorithms, since it allows direct comparison among the meshes obtained using each different algorithm (as illustrated in Fig. 3).

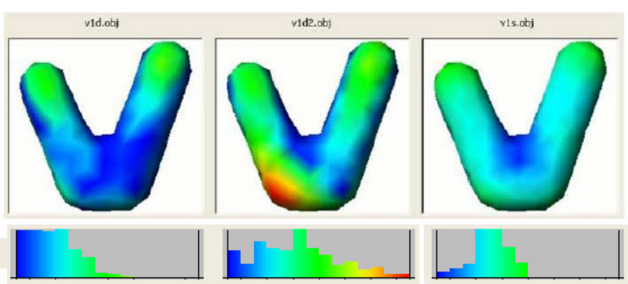

Figure 3 – Feature comparison visualization mode: it allows direct comparison among meshes obtained with different processing algorithms.

PolyMeCo provides users with additional information about what they are visualizing through *Information Windows*. These windows can either allow scene parameter configuration (such as light sources) or show detailed information (as in a probe) and can be moved, docked inside the main window or hidden.

3. Evaluation Method

Although there are several visualization techniques and many systems that use them to visualize large amounts of information, there are comparatively few studies on the evaluation of those techniques and systems. This is perhaps due to the inherent complexity of this evaluation. While there is not yet a body of knowledge on visualization evaluation, the literature presents some examples explicitly using evaluation in the process of designing a visualization system, evaluating specific systems and visualization techniques, as well as comparing alternative visualizations (e.g., [8][9][10][11][12][13]). Moreover, there are also some authors recognizing the importance of making an effort to develop more systematic approaches to this complex problem [14][15][16]. Nevertheless, those who need to evaluate visualization techniques and systems still struggle with a lack of specific methodologies and techniques to conduct the evaluation. Currently, it seems that a reasonable approach could be to adapt the well known and already widely used methods of Usability Engineering [17], taking into account the specificities of the techniques and systems being evaluated.

We considered the ideas in [14] to structure our work, adapted usability testing techniques to obtain feedback from users and that information is now being used to improve our tool.

During the development of PolyMeCo we tried to apply usability principles and guidelines and had some feedback from several people. However, no user centred design approach was followed nor formal usability evaluation was applied, since the development of PolyMeCo started in the scope of a research on mesh processing algorithms, as a tool merely to support a more systematic approach to the analysis of processed meshes. But, we have now a version that has already been used by others with useful results and we wish to improve its usability. Thus, we are performing a first formal evaluation cycle having two major goals: 1) evaluate the main features of the user interface, and 2) evaluate the visualization techniques. Visualization techniques encompass the visualization representations and the interaction mechanisms provided to users, which allow interactive data analysis aiming at greater insight into the visualized data.

We first had heuristic evaluation [18] performed by 3 people having some knowledge in usability, in order to find a list of usability problems, and then performed evaluation sessions with the collaboration of students.

In informal evaluation sessions we asked users to freely use the tool and give suggestions on how to improve its usability and utility. In the formal sessions we used observation and query based techniques to evaluate some user interface features, potentially more confusing, as well as some aspects of the visualizations. These usability techniques are commonly used in user interface evaluation, but they have also been considered appropriate to evaluate some aspects of visualizations [19]. In the next sections we describe the methods used at the evaluation sessions.

3.1. Users

We asked for the collaboration of two different types of users: 3rd year Computer Engineering students attending an introductory course on Human-Computer Interaction (HCI), and 5th year Electrical Eng. students attending a Computer Graphics (CG) course. While the students are not researchers/developers working with polygonal meshes, the intended users of PolyMeCo, they have a profile that makes them reasonably suitable as subjects for the evaluation of our user interface: all of them have high computer literacy and the CG students have a background allowing them to easily understand the issues in mesh processing. Therefore, we conducted an evaluation more focused on the user interface aspects with the HCI students and on the visualization techniques, as well, with the CG students, in formal and informal evaluation sessions, respectively.

3.2. Formal evaluation (HCI students)

Given that the HCI students had been practicing user interface evaluation through different methods, we decided to profit from their capabilities and ask them to act as observers, as well as users. Such procedure would allow us to obtain observation data from a larger number of users. We considered this would provide an interesting practice for the students and that having all students acting both as users and as observers would increase their motivation. We had already used a similar approach in a previous usability evaluation with interesting results [20].

Therefore, while half of the students would perform some predefined tasks, the other half would observe them and register times, task completeness, as well as other relevant information. After some predetermined time they would change roles. Obviously, students that would act first as observers, and later as users, would have a greater acquaintance with the interface, a different level of awareness, and could be considered as more experienced users. We should notice that this evaluation cannot be considered as a controlled experiment, since there was no hypothesis to accept or reject: its main purpose was to gather ideas that could help improve the usability and utility of PolyMeCo.

Two observation sessions were performed with these students; one (OS1) with 18 students and another (OS2) with 17 students. All students were organized, during each session, in groups of two (User#1 and User#2): User#1 would act first as user and later as observer; User#2 would do the opposite. Before asking the participants to perform a predefined set of tasks, they were given an overall description of the tool (30 minutes) and asked to freely use the tool for about 15 minutes. Then, they were informed about the tasks they were supposed to perform and observe, the questionnaires they were supposed to answer, and the used scales. Subsequently, participants were asked to start the test and perform or observe tasks during 30 minutes, fill a questionnaire and then change roles.

The questionnaire was prepared to collect some basic demographic data, aiming also to assess users' reactions to PolyMeCo and their opinion on several issues.

The tasks and the questionnaire were adjusted in a pilot evaluation, conducted with two users, in order to assess the difficulty of the tasks, their duration and the clarity of the questions. As a result, some modifications had to be made both in the tasks and questions.

3.3. Tasks

We defined two sets of tasks for the users to perform during the sessions: a simpler one for Users #1 and a more complex set for Users #2. These tasks were relatively simple, yet they were regarded as representative of typical operations researchers would perform in mesh analysis.

Keeping tasks simple makes it easier to analyse user performance; however tasks should not be so simple that their ecological relevance is unclear (i.e., we have to ask how frequently do those tasks actually occur in real-world tasks, and how significant they are in the overall task solution process).

Each user had to complete 14 tasks within a given time window. The following are some of the tasks:

- Load a given mesh;
- Set the scene illumination parameters;
- Inspect the mesh and locate a particular detail;
- Obtain a coloured model using a given measure;
- Select a given colour scale;
- Find the value of the measure at a given vertex;
- Obtain a new colour mapping for the same colour scale;
- Compare results for meshes obtained with different processing methods using a given measure.

The first five tasks are simple and directed to the evaluation of some user interface features (e.g., use functionality through menu options, change viewing conditions, manipulate meshes); the two following tasks correspond to using functionality not as readily accessible, and the last task was a more complex task focused on evaluating the performance of the user using the visualizations to extract some qualitative or quantitative information through interaction with the data to attain a complex goal. "Find the value of the measure at a given vertex" is oriented to evaluate the interrogation features of the visualization (*Information Probe* tool). "Obtain a new colour mapping for the same colour scale", intends to evaluate an interaction mechanism of a visualization (according to the evaluation criteria in [14]).

3.4. Informal evaluation (CG students)

A less structured, approach was used for the evaluation performed with 20 CG students. In an initial two hour session a presentation was given about the goals and features of PolyMeCo. Some details of the user interface and visualizations modes were shown. Then, a period of questions and answers was allowed so that the participants could better understand the overall purpose of the tool and the type of evaluation we intended to perform. After this, they installed the prototype on their computers and were given some time to use it on their own. During this session we asked users to freely use the tool and think how they would use it if they had to conduct a research concerning mesh processing methods. We also asked for suggestions on how to improve the usability of the tool and on which new useful functionalities should be included.

A week later, in another session, the students were asked to further comment on the tool and to answer the same questionnaire as the HCI students.

3.5. Collected data

Throughout the evaluation sessions performed with the HCI students, each observer had to register the following data concerning the performance of the user, for each task:

- time spent performing the task;
- completion of the task;
- answers to questions (for some of the tasks);
- the difficulty felt by the user as judged by the user and by the observer;
- any additional observation considered relevant.

After completing all the tasks, the overall user reaction to PolyMeCo was collected using a set of bipolar semantically anchored items [21] (e.g., 1-frustrating *versus* 7-gratifying) using a qualitative scale from 1 to 7:

Frustrating Gratifying	Boring Stimulating
1 2 3 4 5 6 7	1 2 3 4 5 6 7
Rigid Flexible	Difficult Easy
1 2 3 4 5 6 7	1 2 3 4 5 6 7

Users were also asked to evaluate several features (listed in Table 1) of PolyMeCo using a qualitative scale:

| 1 | 2 | 3 | 4 || | N |

where 1 is complete disagreement, 4 is complete agreement and N corresponds to not having an opinion or not wanting to express it.

Opportunity to give suggestions or make any comments was given through open questions, as well as during informal conversation at the end of the sessions. Demographic data concerning age, gender and experience with 3D computer graphics were also collected.

Exploratory Data Analysis was used to analyze the collected data and results (obtained using STATISTICA [22]) are presented next.

4. Results

In this section we present some results obtained with the HCI students, more focused on the user interface, as well as results obtained with the CG students more focused on the evaluation of the visualization modes and overall interest of the tool.

4.1. Results obtained with the HCI students

We had the collaboration of 35 HCI students (5 women and 30 men); the median age of this group was 21, one student had colour perception problems and most of them were frequent users of several applications that gave them some familiarity with 3D computer graphics (such as MatLab, Blender and computer games).

The median value of the overall users' reactions to PolyMeCo in the above mentioned items (e.g., 1-frustrating *versus* 7-gratifying) was 5 for all items. Every user expressed an opinion on all items, except one user that didn't give any opinion on one item, and only five users classified items with less then 4. These results convey a positive reaction to the tool. Moreover, the results concerning the users opinion on the features listed in Table 1 convey a positive opinion, as well.

Features	median
Is easy to learn	4
Organization is understandable	3
Response time is reasonable	4
Information layout is adequate	3
Terminology is consistent	3
Text is easy to read	4
Messages are clear	3
Feedback is adequate	3

Table 1 – Users opinion on general and specific aspects of PolyMeCo (in a scale of 1- complete disagreement to 4 - complete agreement).

Table 2 shows the median time values collected for all the tasks performed by both types of users: Users #1 that carried out the tasks (1 to 14) first and then observed their colleagues, and Users #2 that acted as observers first and then performed the tasks (15 to 28). Analysing this table, we notice that the tasks performed by Users #2 took generally longer than the tasks performed by Users #1, which might result from the fact that the latter tasks were designed to be more complex, since Users #2 would have already some more knowledge about the tool just by observing their colleagues performing tasks 1 to 14. Table 3 shows, for all the tasks that had questions, the number of users that completed and did not complete such tasks, as well as the number of users that gave totally correct, partially correct and totally wrong answers (in order to assess if the task had been correctly performed). Figure 4

Task	Time (s)	Task	Time (s)	Task	Time (s)	Task	Time (s)
1	10	**8**	10	**15**	10	**22**	17
2	10	**9**	20	**16**	15	**23**	40
3	8	**10**	90	**17**	12	**24**	70
4	4	**11**	15	**18**	15	**25**	37
5	46	**12**	20	**19**	20	**26**	30
6	29	**13**	11	**20**	30	**27**	25
7	10	**14**	15	**21**	36	**28**	45

Table 2 – Median times for all tasks: Users #1 – tasks 1-14; Users #2 – tasks 15 - 28

Task	Comp. task	Total. corr.	Part. corr.	Total. wrong	Didn't comp.
5	14	14	0	0	4
7	18	18	0	0	0
8	18	18	0	0	0
10	13	12	1	0	5
12	18	18	0	0	0
22	17	16	1	0	0
24	15	12	2	1	2
26	17	17	0	0	0
27	17	17	0	0	0
28	17	16	0	1	0

Table 3 – Number of users that completed and didn't complete the tasks having questions, and gave totally correct, partially correct or totally wrong answers.

shows the box-plots of the times taken for tasks 1-14 and 15-28. Observing Tables 2 and 3, as well as Fig. 4, we notice that the tasks that took longer (10 and 24) also correspond to a larger number of users that didn't complete the task or didn't answer correctly the questions. Analysing these tasks we concluded that both involved the use of the probe in order to assess the value of a computed measure at a given vertex. This result was not expected since the probe seemed reasonably easy to use; however, examining the way this functionality is offered and considering some observations performed throughout the evaluation sessions, we believe these longer times might be due to Fitts Law [21]: given that the probe only returns a value for the measure when the user is pointing at a vertex (and vertices are rather small), an accurate gesture is required to select a vertex, which takes longer.

The student with colour perception problems could not perform a task asking to locate a maximum error point on a mesh surface (maximum error appeared in red, as he was using a rainbow colour scale). After a short explanation, he was capable of selecting another colour scale (from those available in PolyMeCo), which allowed him to complete the task.

The median time and the number of users that did not complete task 5 are also high. However, analysing the task, we concluded that it was not correctly positioned in the task sequence; thus, the poor performance of users does not

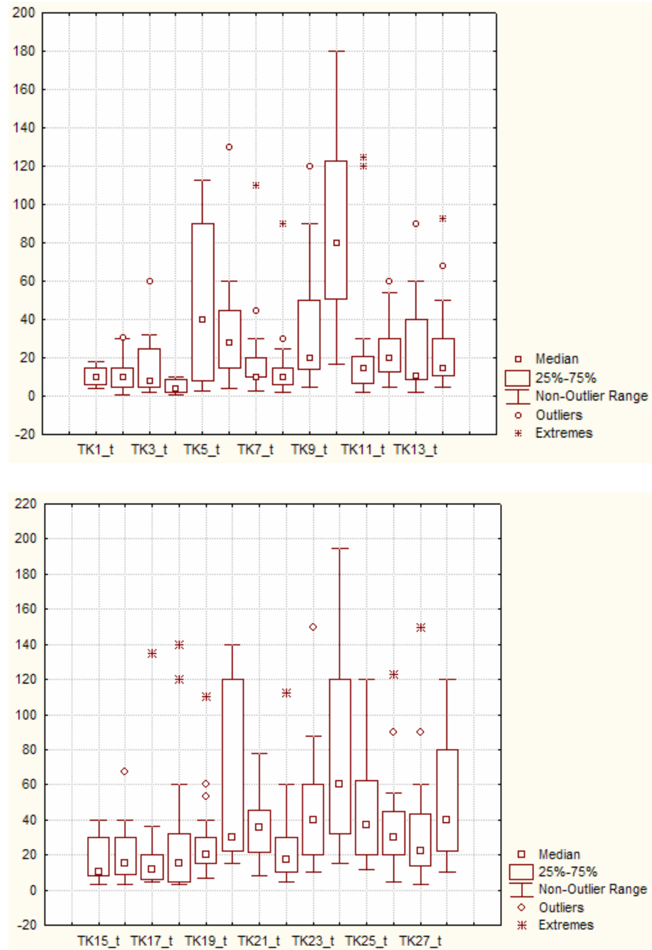

Figure 4 - Box-plots of time spent, task by task

seem to correspond to a usability problem of the user interface, but, instead, to a problem of the questionnaire.

4.2. Results obtained with the CG Students

CG students used PolyMeCo during the evaluation sessions in a more committed manner than we had previously anticipated, giving plenty of suggestions concerning interesting new functionality to include or ways to improve the usability of the tool.

The results concerning their reaction and opinion about PolyMeCo, collected through the questionnaire were very similar to the results obtained from the HCI students, conveying a positive opinion, as well.

5. Discussion and Conclusions

In this paper we describe the on-going evaluation of a visualization based tool under development, intended to offer researchers and developers an easy way to analyse and compare polygonal meshes. We applied several common usability evaluation methods (heuristic evaluation, observation and query techniques) in order to find usability problems and elicit new ideas that will, eventually, allow to improve the usability of the tool, as well as to produce a new version including additional functionality.

The main finding obtained through the evaluation conducted with observation techniques was the usability problem concerning the probe: user feedback describing their difficulties and some suggestions provided a few clues towards a solution.

Even though the users did not belong to the target audience of PolyMeCo (researchers/developers working with polygonal meshes), their performance was very satisfying and the suggested ideas were adequate, revealing their understanding of at least some of the features and goals of PolyMeCo.

Furthermore, this evaluation confirmed that an adequate selection of the tasks and their sequence is very important and, thus, carefully evaluating the questionnaire is essential.

On the other hand, the questionnaire gave us a general positive idea about the usability of PolyMeCo in the three dimensions of usability: ease of learning, ease of use and satisfaction.

We are encouraged by the results obtained; still, we intend to evaluate how PolyMeCo supports users as they become more experienced.

Finally, we consider that the used methodology could be adequate to test the usability of other visualization based tools.

Acknowledgements

The authors wish to thank all students of Human-Computer Interaction and Computer Graphics courses, of the University of Aveiro, that have graciously collaborated.

The second author would like to thank the research unit 127/94 IEETA, of the University of Aveiro, for the grant that supports his work.

References

[1] S. Silva, J. Madeira, B. Sousa Santos, PolyMeCo – A Polygonal Mesh Comparison Tool, *Proceedings 9th International Conference on Information Visualization (IV05)*, 842-847, 2005

[2] P. Cignoni, C. Rocchini, R. Scopigno, Metro: Measuring Error on Simplified Surfaces, *Computer Graphics Forum*, vol. 17, no. 2, 167-174, 1998

[3] L. Zhou, A. Pang, Metrics and Visualization Tools for Surface Mesh Comparison, *Proceedings SPIE 2001 vol. 4302. Visual Data Exploration and Analysis VIII*, 99-110, San José, California, 2001

[4] M. Roy, S. Foufou, F. Truchetet, Mesh Comparison using Attribute Deviation Metric, *Int. Journal of Image and Graphics*, vol. 4, no. 1, 1-14, 2004

[5] S. K. Card, J. Mackinlay, B. Shneiderman (eds.). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufman. 1999.

[6] S. Silva, F. Silva, J. Madeira. B. Sousa Santos, Evaluation of Mesh Simplification Algorithms using PolyMeCo. A Case Study, *Proceedings SPIE 2007 vol. 6495, Visualization and Data Analysis 2007*, 6495D, San José, California, 2007

[7] D. Borland, R. Taylor, Rainbow Color Map (Still Considered Harmful, *IEEE Computer Graphics and Applications*, vol. 27, no.2, 14-17, 2007

[8] D. Hix, J. E. Swan II, J. Gabbard, M. McGee, J. Durbin, T. King. User-Centered Design and Evaluation of a Real-Time Battlefield Visualization Virtual Environment. *Proceedings IEEE Virtual Reality 99*. 96-103. 1999.

[9] C. North, B. Schneiderman. Snap-Together: Can Users Construct and Operate Coordinated Views? *Int. Journal Human-Computer Studies*, 53(5). 715-739. 2000.

[10] M. Sebrechts, J. Vasilakis, M. Miller, J. Cugini, S. Laskowski. Visualization of Search Results: A Comparative Evaluation of Text, 2D, and 3D Interfaces. *ACM Conf. Research and Development in Information Retrieval*, ACMSIGIR 99, 3-10, California, 1999.

[11] U. Wiss, D. Carr, H. Jonsson. Evaluating Three-Dimensional Information Visualization Designs: a case Study of Three Designs. *Proceedings Information Visualization 98.* IEEE, 137-144. 1998.

[12] A. Kobsa. An Empirical Comparison of Three Commercial Information Visualization Systems. *IEEE Symposium on Information Visualization, InfoVis01*. 123-130. 2001.

[13] T. Barlow, P. Neville. A Comparison of 2D Visualizations of Hierarchies. *IEEE Symposium on Information Visualization, InfoVis01*. 131-138. 2001.

[14] C. S. Freitas, P. Luzzaerdi, R. Cava, M. Winckler, M. Pimenta, L. Nedel. Evaluating Usability of Information Visualization Techniques. *Proceedings 5th Symposium on Human Factors in Computer Systems* IHC2002, Fortaleza, Ceará, 2002.

[15] R. Brath. Concept Demonstration Metrics for Effective Information Visualization. *IEEE Symposium on Information Visualization, InfoVis97*. 108-111. 1997.

[16] G. Grinstein, P. Hoffman, S. Laskowski, R. Pickett. *Benchmark Development for the Evaluation of Visualization for Data Mining*. http://home.comcast.net/~patrick.hoffman/VIZ/benchmark.pdf (visited 3/2007)

[17] J. Nielsen, *Usability Engineering*, Academic Press, 1993

[18] J. M. Carlsmith, P. Ellsworth, E. Aronson. *Methods of Research in Social Psychology*, Addison-Wesley. 1976.

[19] C. Ware. *Information Visualization – Perception for Design*. Morgan Kaufmann, 2000.

[20] B. Sousa Santos, F. Zamfir, C. Ferreira, O. Mealha, J. Nunes, Visual Application for the Analysis of Web-Based Information Systems Usage: A Preliminary Usability Evaluation, *Proceedings 8th International Conference on Information Visualization (IV04)*, 295-301, 2004

[21] B. Shneiderman, *Designing the User Interface-Strategies for effective Human-Computer Interaction*, 3rd ed., Addison Wesley, 1998

[22] *STATISTICA for Windows*, version 6.0, StatSoft Inc., 2001.