Evaluation of Volume Visualizations: Proposal of a Tool

Beatriz Sousa Santos

Resumo **- São apresentadas algumas questões relevantes para o projecto de uma ferramenta de apoio à avaliação da qualidade de visualizações em volume, de dados baseados em voxeis. Apresenta-se também uma proposta para a funcionalidade de uma tal ferramenta, que deve oferecer ao utilisador, entre outras, a possibilidade de importar e exportar dados reais ou sintéticos, gerar dados sintéticos a partir de parâmetros controláveis, usar várias técnicas de visualização em volume bem como aplicar um conjunto de métodos de avaliação a essas técnicas.**

Abstract **- Some issues relevant to the design of a tool for the evaluation of volume visualizations are presented, as well as a proposal for its general functionality. This tool should offer the user, among others, the possibility of importing and exporting real or synthetic test data, generating synthetic data with controlled parameters, using several volume visualization techniques and apply a set of evaluation methods to those techniques.**

I. INTRODUCTION

Visualization techniques and systems are actually just instances of technology in science for detecting, analysing and interpreting signals - albeit signals of a broad type, often rather different from the most usual. The need to provide a basis for the quantitative evaluation of similar techniques and systems is well recognised in many areas of science and technology; thus an analogous need exists for the evaluation of visualization techniques and systems; it is necessary to know how well a given visualization system or technique is helping a user to detect and interpret structure in his or her data. To accomplish this, standard sets of test data as well as evaluation methods are needed [1,2,3].

Evaluating the quality of visualizations of voxel based data obtained using volume visualization techniques is an example of the above mentioned problem. It is a complex, mostly still ill-defined process and, perhaps due to these reasons, not commonly performed. Although some tools exist which allow the implemention of several of those techniques [4,5], there are almost no tools which help visualization researchers on the complex task of evaluating the resulting quality. Possibly, off-the-shelf visualization packages could be used, but this would imply a great deal of work [3].

Starting the development of tools that could be used to provide support for developing and applying rigorous evaluation methodologies of volume visualization techniques is the aim of the present work. This will involve several fundamental issues as the definition of the testing data set and, most difficult, the definition of such complete methodologies. Since the later goal seems very ambitious, we shall begin by defining the kind of test data and a few types of methods for evaluating the quality of volume visualizations that could be used. After this is accomplished, it will be necessary in order to design our tool, to decide which choice of volume visualizations techniques and which functionality are to be offered to the researcher as well as how this functionality is going to be presented (i.e. the general organisation of the user interface).

Section II is concerned with two fundamental issues with which investigators are confronted when evaluating visualizations: the choice of the test data set and the kind of evaluation methods to be used. In section III an example of the funtionality that the referred tool could offer to the user is presented. Section IV addresses some issues concerning the user interface design and the choice of visualization techniques. Finally, section V, concludes this work.

II. EVALUATION OF VISUALIZATIONS

Evaluating visualizations in general and volume visualizations in particular, confronts researchers with many options; two general considerations are the type of data and the kind of methods to use in such studies. The kind of test data that should be used in the proposed evaluation tool is briefly described and three types of evaluation methods are presented, as well as a possible taxonomy of these methods based on some dimensions that seem adequate to their classification.

A. Test Data

Basically two fundamental choices can be done: synthetic or real data [6,7]. Ultimately, the evaluation of a technique should be done with real data, but it is reasonable to begin by using fully specified and systematically controlled data structures embedded in synthetic data. The use of computer generated data results flexible and allows the detection of errors and

inaccuracies of the technique to be evaluated in a way much easier than using real data. In some applications it may require a lot of modelation and may be only approximative, however it is perhaps the only method that allows a complete knowledge of the "ground truth" [8]. Thus, fundamental features of the proposed evaluation tool should be the possibility to easily create and edit different types of synthetic data and the possibility to import synthetic or real data already existing in other systems as well as exporting them.

B. Evaluation Methods

Several alternatives to evaluate the quality of volume visualizations exist. It seems reasonable to consider the possibility of including the following three types of evaluation methods in the proposed tool:

- Methods involving pannels of human observers: which rate sets of visualizations, allowing the computation of some quality measures, analogously to what is done to image quality evaluation [9,10]
- Quality indices: obtained directly from some kind of measures that seem relevant to the quality of the visualization, computed directly from the application of the visualization technique to the data [11]
- Digital observers: which use models of the Human Visual System (HVS) [12,13] to estimate ratings that human observers would attribute to visualizations.

Quantitative or qualitative, objective or subjective methods can be considered in these three types of evaluation methods. It is also possible to have methods that take into consideration the Human Visual System (HVS) and methods that do not. Based on these "dimensions" which seem adequate to the classification of evaluation methods, it is possible to propose a taxonomy for the referred methods (figure 1).

With the purpose of clarifying what is each of these dimensions let us define them as:

- Qualitative/ Quantitative- related to the type of result yielded by the evaluation method: quantitative (having a magnitude that can be and is denoted by a numerical expression), versus qualitative (which cannot or is not)
- Subjective/ Objective- related to the way the result is obtained by the evaluation method: subjective (through the judgement of human observers), versus objective (solely from the nature of the data and visualization method, without the intervention of the observer's judgement)
- Filtered by the Human Visual System/ Not Filtered by the Human Visual System- taking into account the response of the Human Visual System or not.

Note that the second and third dimensions must not be mixed up, they correspond to different characteristics; an evaluation method may not be subjective in spite of taking into consideration the HVS response [7]; for instance the methods we have called "digital observers" should use models of HSV to compute some results (possibly quantitative albeit it should not be impossible to produce a qualitative result) without the intervention of the judgement of any human observer and thus they are objective methods.

Another dimension which could, perhaps, be considered in a taxonomy of the proposed methods is related to the type of observer's perceptual/cognitive processes (of different levels and nature) used in the quality evaluation of visualizations. For instance, if an observer is asked to choose from a set of visualizations which one has less noise, jaggies or blur, the type of perceptual/cognitive processes used by the user seems to be of a different level and nature than when he/she is asked to give an interpretation of the visualization. These experiments could integrate methods meant to evaluate what we could, perhaps, call "visual quality" and "cognitive quality", respectively. According to this nomenclature, visual quality would be concerned only with the quality of the image and cognitive quality would be concerned with the information conveyed to the observer by the visualization. The referred dimension was not used in the taxonomy shown in figure 1 for the sake of brevity and also since that, due to the complexity of the involved processes, it seems currently exceedingly difficult to develop digital observers to compute predictions about human observer ratings on "cognitive quality" and thus this dimension would only apply to the methods involving panels of human observers.

	Qualitative / Quantitative	Subjective / Objective	HVS fitered / not filtered
Pannels of Observers	X / X	X	X
Quality Indices	X	X	X
Digital Observers	X	X	X

Fig. 1- Taxonomy of the proposed evaluating methods

Evaluation methods, preferably of all three described types, should integrate the proposed evaluation tool; however the development of such methods is expected to involve much research work since no ready to use methods exist (at least to the best of our knowledge).

III. GENERAL FUNCIONALITY OF THE TOOL

Considering all issues addressed in the previous section and some examples of evaluation environments in the scope of visualization in general [6,14], the following main functionality is proposed for the tool:

- importing and exporting external test data in some standard formats
- generating and editing synthetic test data with characteristics controllable by the researcher
- archiving test data (imported or generated)

- creating visualizations from the test data, using several relevant volume visualization techniques
- archiving visualizations
- importing and exporting visualizations in some standard formats
- computing several quality indices
- implementing digital observers
- administrating test procedures to observers
- keeping track of the results and scores of test procedures
- generating evaluation reports.

The interest of importing and exporting synthetic or real data as well as visualizations is easily justified by the great advantage it means to be able to exchange information with other systems; obviously this will be facilitated by the use of several standard and/or widely used formats [4]. The needs for the capacity of archiving both data and visualizations as well as providing several volume visualization techniques goes without justification in a tool such as the proposed one.

Finally, and in spite of not being as fundamental as the previous functionality items, the capabilities of supporting the administration of tests to human observers, keeping track of the results and generating repports, would be a great help to the researchers intending to apply evaluation methods involving large pannels of human observers and/or large visualization sets.

IV. OTHER CONSIDERATIONS ABOUT THE TOOL

Among other relevant matters that should also be taken into consideration when designing the proposed evaluation tool, are the ones related to the user interface and the choice of volume visualization techniques. For the matters related to the user interface, the guidance provided by the general literature on Human-Computer Interaction [15,16] as well as by the one more specific to visualization systems [17], should be used. Concerning the volume visualization techniques that should be chosen, and taking into account that many different techniques exist [18], possibly the best recommendations is that the tool should allow, as easily as possible, the integration of any visualization technique the researcher might choose to use.

V. CONCLUSIONS

Some fundamental issues to have in consideration when designing, specifying and implementing a tool meant to support researchers in the complex task of developing and applying evaluating methods of volume visualization quality were addressed. A tentative functionality for such a tool was proposed, as well as the type of data and methods that could be integrated in such a tool. A taxonomy for those types of methods, based on some dimensions which seem adequate for the classification of evaluation methods of volume visualization quality was also presented. The author expects this work will make easier the task of designing a first version of such a tool, which could become a very useful support for the researchers engaged is such an hard work as studing the problem of quality evaluation of volume visualizations. Moreover most issues addressed in this work can be generalised to other visualization applications.

REFERENCES

[1] Nielson, G., "Challenges in Visualization Research", *IEEE Transactions on Visualizations and Computer Graphics*, Vol. 2, N. 2, 1996, pp. 97-99

[2] Rushmeier, H., et al., "Panel on Metrics and Benchmarks for Visualization, Proceedings of *Visualization 95*, IEEE Computer Society Press, 1995, pp. 422- 426

[3] Keim D., D. Bergeron, R. Pickett, "Test Data Sets for Evaluating Data Visualization Techniques", Grinstein, G., H. Levkowitz (eds), *Perceptual Issues in Visualization*, Springer Verlag, 1995,pp. 8-21

[4] Brodlie, K., L. Carpenter, R. Earnshaw, J. Gallop, R. Hubbold, A. Mumford, C. Osland, P. Quarendon, *Scientific Visualization Techniques and Applications*, Springer Verlag, 1992

[5] Brown, J., R. Earnshaw, M. Jern, J. Vince, *Visualization, Using Computer Graphics to Explore Data and Present Information*, John Wiley, 1995

[6] Levkowitz, H., R. Pickett, R. Smith, M. Torpey, "An environment and Studies for Exploring Auditory Representations of Multi-dimensional Data", Grinstein, G., H. Levkowitz (eds), *Perceptual Issues in Visualization*, Springer Verlag, 1995 pp. 47-58

[7] Coatrieux, J.L., Dillenseger, J.L. Personal communication to the author , 1997

[8] Van Clynenbreugel J., Nuyts, R. Verbeek, M. H. Smet, G. Marchal, P. Suetens, "On the Problem of Validating Computer Assisted Medical Image Interpretation", *Contemporary Perspectives in Three-Dimensional Biomedical Imaging*, C. Roux, J.L. Coatrieux (eds.), IOS Press, 1997

[9] Kaukoranta, T., P. Franti, O. Nevalainen, "Empirical Study on Subjective Quality Evaluation of Compressed Images", Proceedings of SPIE Conference vol. 2663, *Very High Resolution and Quality Imaging*, 1996, pp. 88-96

[10] Avadhanam, N., R. Algazi, "Prediction and Measure of High Quality in Still Image Coding", Proceedings of SPIE Conference, vol. 2663, *Very High Resolution and Quality Imaging*, 1996, pp. 100-109

[11] Sousa Santos, B., J. Nunes, C. Coimbra, C. Ferreira, A. M.Tomé, N. Lau, " Assessing the Quality of Visualizations Produced Using Volume Rendering: Experiments with a Ray Caster", *Revista do DETUA* , Universidade de Aveiro, September,1997

[12] Daly, S., "The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity", *Digital Images and Human Vision*, A. Watson (ed.), MIT Press, 1993, pp. 179-206

[13] Lubin, J., "The Use of Psychophysical Data and Models in the Analysis of Display System Performance", *Digital Images and Human Vision*, A. Watson (ed.), MIT Press, 1993, pp. 163-178

[14] Wong., P., D. Bergeron, “A Multidimensional Multivariate Image Evaluation Tool”, Grinstein, G., H. Levkowitz (eds), *Perceptual Issues in Visualization*, Springer Verlag, 1995, pp.95-108

[15] Dix, A., J. Finley, G. Abowd, R. Beale, *Human Computer Interaction*, Prentice Hall, 1993

[16] Meyhew, D., *Principles and Guidelines in Software and User Interface Design*, Prentice Hall, 1992

[17] Hibbard, W., H. Levkowitz, J. Haswell, P. Rheingans, F. Schroeder, “Interaction in Perceptually Based Visualizations”, Grinstein, G., H. Levkowitz (eds), *Perceptual Issues in Visualization*, Springer Verlag, 1995, pp. 23-32

[18] Elvis, T., "A survey of Algorithms for Volume Visualization", *Computer Graphics*, Vol. 26, N. 3, August, 1992, pp. 194-201

AKNOWLEGMENT- The author wishes to thank Professor J.L Coatrieux and Dr. J. L. Dillenseger, from LTSI, University of Rennes II, for all the interesting discussions on the general subject of Visualization Evaluation which greatly contributed for the clarifications of some concepts used in this work.