AN ARCHITECTURE FOR INTERACTIVE MULTIMODAL VISUALIZATION SYSTEM

ISABEL H. MANSSOUR[#], SÉRGIO S. FURUIE[+], LUCIANA P. NEDEL[*] AND CARLA M. DAL SASSO FREITAS[*]*

[*]UFRGS - Universidade Federal do Rio Grande do Sul, Instituto de Informática/PGCC
Caixa Postal 15064, 91501-970 Porto Alegre, RS, Brazil
Tel: +55 (51) 316-6842, 316-7037 or 316-6811 Fax: +55 (51) 319-1576
e-mail: {manssour, nedel, carla}@inf.ufrgs.br

[#]PUCRS - Pontifícia Universidade Católica do Rio Grande do Sul, Faculdade de Informática
Av. Ipiranga 6681, Prédio 30, Bloco 4, 90619-900 Porto Alegre, RS, Brazil
Tel: +55 (51) 320-3558 Fax: +55 (51) 320-3758

[+]InCor / Instituto do Coração HC-FMUSP, Unidade de Pesquisa e Desenvolvimento, Serviço de Informática
Av. Dr. Enéas de Carvalho Aguiar, 44, 05403-000 São Paulo, SP, Brazil
Tel: +55 (11) 3069-5544 Fax: +55 (11) 3069-5311
e-mail: Sergio.Furuie@incor.usp.br

ABSTRACT
Although the use of information acquired from multiple sources to help in medical diagnosis is increasing, the integration of multivariate data into a unique 3D representation is non-trivial. To overcome this problem, many researchers have been trying to develop suitable strategies to integrate important characteristics of multiple data sets into a single visual representation. By means of multimodal visualization techniques these researchers aim to provide better insight about data coming from different imaging modalities. This paper briefly describes multimodal visualization, emphasizing the requirements and open issues for the development of such systems. The architecture modeling of a multimodal interactive visualization system is also presented.
Keywords: Multimodal Visualization, Registration, Interaction Tools, and Medical Imaging.

1 INTRODUCTION

With the evolution of image acquisition technology in terms of resolution and tissue distinctiveness, the capacity and fidelity of image diagnosis were further extended. Several image acquisition modalities have been used for years to facilitate the medical diagnosis, e. g. Computed Tomography, Magnetic Resonance Imaging, and Positron Emission Tomography. These medical images can be categorized as anatomical (CT, MRI) or functional images (PET). The first ones depict primarily morphology, and the second ones, show functional and metabolic information. In fact, these modalities show different, complementary and/or partially overlapping aspects of the examined anatomy and function [1, 2].

Currently, the tendency of information acquisition using multiple sources to help the medical diagnosis in critical diseases is increasing. However, registration and fusion of the multivariate data into a 3D representation of the patient are extremely difficult, time-consuming and error-prone tasks. So, the integration of images from multiple modalities has rapidly evolved into an important area of research: multimodal visualization, which is concerned about the proper integration of data obtained from separate patients or from separate scanners. Such systems combine functional and metabolic information with anatomical data, and increase the confidence of the observers in the location of a functional abnormality in relation to the anatomy [3].

Other good examples of applications that might benefit from multimodal visualization are interactive simulation of neurosurgery, radiotherapy treatment planning and suitable comparison with a reference atlas [2, 3, 4, 5, 6,]. The motivation for developing new system architectures for interactive multimodal visualization comes from the usefulness of the integrated display of functional and anatomical images in several medical applications. Multimodal visualization systems can be developed for this specific purpose, or can be built up as modules of generic scientific visualization systems.

The main goal of this paper is to present the architecture of a multimodal visualization system. This work is part of the development of a framework for medical applications, which guarantees software reuse and integrates existing tools [7]. This paper is organized as follows. Section 2 presents the needs and issues concerning the use of multimodal visualization techniques. The problems still found for the development of such applications are discussed in Section 3. Section 4 presents our proposal for the architecture of a multimodal interactive visualization system that integrates several tools and tries to overcome some of the problems reported. Final comments are drawn in the last section.

2 REQUIREMENTS FOR MULTIMODAL VISUALIZATION

A first and fundamental step to generate images of multimodal volume (as the ones obtained from RenderVox system shown on Figure 1 [8]) consists of bringing the involved modalities into spatial alignment; a procedure referred to as *registration*. After registration, a *fusion* step is required for the simultaneous display of the two data sets. As reported by Maintz [1], it is important to emphasize that the terms *registration* and *fusion*, as well as *matching* and *integration* appear with different meanings in the literature, either referring to a single step or to the whole integrated process.

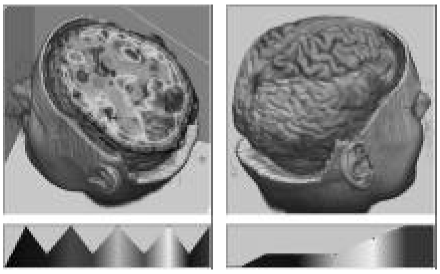

Figure 1 – Examples of multimodal visualization generated from MRI and PET data

Registration is an important task used to match two or more images (2D or 3D), obtained at different times, from different sensors and scanners, or from different viewpoints [1, 9]. Several common characteristics can be distinguished in registration algorithms. For example, whether it is based on artificial objects introduced into the image or not, whether it is based on rigid or affine transformations, or whether it is monomodal (images from the same modality) or multimodal (images from different modalities) [1].

In the last few years, researchers have focused on *mutual information* as the image registration technique for multimodal applications. It is based on information theory and works directly with image data [5, 6, 10, 11, 12]. Registration is achieved by adjusting the relative position and orientation of the images until the mutual information between the images is maximized [13, 14, 15].

Just after the registration step, an **interactive volume visualization** algorithm can be used for the integrated display. Besides the integrated display, it is important to provide ways for: manipulating the data; extracting measurement and functional information from the final image (quantification); and allowing different types of visualization and navigation inside the structures.

To allow interactive data manipulation, several volume rendering acceleration techniques have been proposed in the last few years [8, 16, 17, 18, 19, 20]. Moreover, since the calculation of each ray in volume rendering algorithms is independent, another strategy used for speed improvement is to take advantage from computers with parallel architecture, or execute the algorithm in several computers or workstations forcing a distributed parallel program implementation [21].

Often, depending on the application it is necessary to implement a **segmentation** step to isolate a structure from the rest of the data set. In multimodal visualization systems, a segmentation step may be needed for visualization, registration and measurement extraction. This last application is very useful in **quantification** and volume exploration techniques when the user points some structure in the image and obtains several data about it (e.g., dimension and functional information). Considering user actions during segmentation processes, we can distinguish *manual*, *automatic* and *interactive segmentation* methods [22, 23]. So, the development of a reliable segmentation method can be a decisive step during multimodal visualization process, since it will affect all the remaining steps. This led us to conclude that building a multimodal visualization system involves the development and integration of tools like registration, segmentation, and optimized interactive visualization techniques.

3 OPEN ISSUES

In spite of the technological advance in image acquisition and the multimodal visualization systems already developed, such type of systems still have a lot of shortcomings for effective clinical use. In this section some of these challenging problems will be discussed.

Platform Independence. One big challenge in the design of such a system is the achievement of a platform-independent system capable to solve a large range of the visualization problems, in order to be really used indistinctly from medical workstation to the personal computers of the physicians. ANALYZE^{TM}, for example, is a complete system that runs only in Unix systems, which in several circumstances may not be available during a diagnosis procedure in critical diseases. The system developed by Hastreiter and Ertl [24] is based on OpenGL and OpenInventor and tries to take advantage of procedures supported by hardware. VROOM achieves portability by adopting object-oriented programming; it was implemented using C++ programming language.

User-Friendly Interface. An important requirement for clinical acceptance of visualization systems is the design and implementation of a user-friendly interface. Manufacturers of medical image equipment often adopt the WIMP (*Window, Icon, Menu, and Pointer*) style of interaction. Unfortunately, 2D windows and mouse interface are less suited for direct 3D interaction with volume data. So, suitable input and output devices, which include virtual reality and haptic devices, should be developed to support 3D interaction [23]. However, these kinds of devices have to be pretty improved for real utilization, since they do not provide adequate precision and it is difficult to achieve real-time interaction with a low cost. Thus, the design of new graphical user-friendly interfaces is a promising research area, where physician participation will be indispensable.

Segmentation. Despite the widespread acceptance of interactive segmentation methods, the majority of multimodal visualization systems implemented automatic segmentation (e.g. ANALYZE™ [4] and Stokking [3]). So the merging of interactive segmentation with a multimodal interactive visualization system is a challenging issue.

Registration. Among some existent problems in this subject, it is possible to emphasize the large processing time required by registration. For example, several authors [24, 25] had improved the registration technique of mutual information. But, despite the optimized algorithms developed so far, depending on the volumes used, it is not possible yet to obtain real-time results. Another problem is how to validate the reached results .

Multimodal Visualization. There are few works dedicated to the development of a complete multimodal system that integrates several tools in an optimized way. Analyzing these works, one notice that several new tools could be devised to allow the interactive visualization of multimodal medical images. These systems also have to increase their visualization capabilities, as for example, to allow the visualization of local features like high intensities and curvature [19].

Quantification. After the visualization and segmentation steps, the user could explore and extract useful information, as measurements. Three steps need to be considered to allow the extraction of functional information and measurement when, for example, a user selects one region of the volume. The first one is exploration, when the recorded image is visually inspected; the second is *measurement*, when geometric data are extracted from the recorded image; and the last one is *statistical*, when the reliability of the measurement is determined [19]. So, since measurements are obtained from geometric data, and the functional information is associated with specific structures, the quantification accuracy is extremely dependent of the segmentation process. Moreover, to work with volumes from different modalities, the registration step also influences the final result.

Interaction. During visualization, the user needs to explore and interact with the volumes. Several kinds of interaction tools could be used [8]: determination of cross sections and selection of different regions and structures; cut volumes and cut planes; quantification of selected structures; and obtainment of structure information in each modality. It is also important to provide tools for the specification of classification tables to assign different colors and transparency levels for each volume, or structure inside the volume. To simulate these complex interactions a computer system has to provide fine details, realistic behavior during interaction and real-time images generation. To perform all these requirements, the processing becomes very heavy, and sometimes, it is not possible to generate the results in real-time. Noordmans [19], as well as Leventon [26] and Grimson *et al.* [27] that reported on image-guided surgery focused their works in the development of interactive tools.

Real-Time Image Generation. In accordance with the open issues discussed in this section, we could say that one of the central problems of multimodal visualization is still the development of real-time algorithms for registration, segmentation and visualization, in order to allow a real-time interaction with the volumes. This interaction is very useful, especially in CAS (Computed Assisted Surgery) applications. Several acceleration techniques were already developed for integrated representation [21, 19], as described in Section 2, but this problem is not completely solved.

Validation. Validation and accuracy represent a general problem in medical imaging, not only for registration, as extensively reported in the literature [28, 5, 1], but also for segmentation and visualization. Segmentation is one of the hardest problems in medical image analysis. It is very difficult to automatically isolate a structure of interest, because regions often do not have continuous borders or homogeneous interior. Because of this, interactive methods are gaining more acceptance in the last few years [22, 23].

For validating new registration techniques, or measuring the accuracy of algorithms, usually a *gold standard* registration, i. e. a bone marker match is done. Moreover, a phantom could also be used as a source for image acquisition. However, one of the most difficult tasks related to this still is the quantification of how accurate the results are. To minimize the validation problem, items like precision, accuracy, robustness, and algorithm complexity are usually analyzed [1, 12, 28, 29]. Consequently, an obvious problem is to quantify the fidelity of rendered images, since image acquisition, segmentation and/or registration are likely to introduce errors affecting the quality of rendering [21].

Physicians Participation. An efficient and easy to use interface is an essential requirement for interaction tools, in such a way that the physician can determine, manipulate and analyze specific regions. Moreover, interactive segmentation techniques depend on the participation of the physicians and needs a user-friendly interface. So, it is important to notice that their participation is essential, not just in interface development, but in the whole process, from diagnosis to surgery planning and simulation. Since they will be the users, they have to identify important features of the interface, to enumerate appropriate visualization and interaction tools, and to analyze the results.

4 PROPOSED ARCHITECTURE

Considering the several requirements stated before, and the possibility of using some of the optimized algorithms presented in the literature, we developed a conceptual model of a system for interactive multimodal visualization. This model is based on the MVC (*Model-View-Controller*) pattern [30] and is described here using UML (*Unified Modeling Language*) [31]. MVC is extensively used in interactive systems to keep the functional core independent from the user interface.

As a first step, the use cases diagram was settled (Figure 2). According to this diagram, the system has two actors: the physician (radiologist, specialist, etc.), who will manipulate the available tools and data, and the data acquisition system, which will supply the system with volume images. The use cases modeled correspond to some of the functionality listed in Section 3.

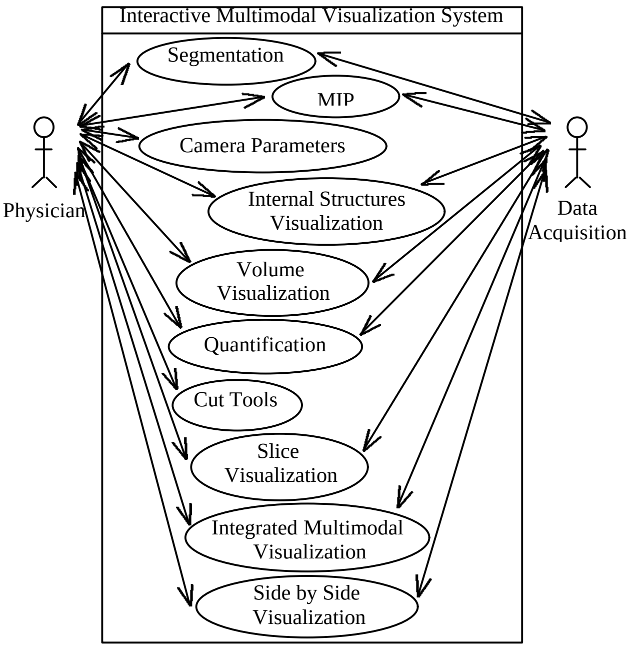

Figure 2 – Use cases diagram

The proposed architecture integrates registration, segmentation, and interactive visualization of multimodal data sets. Figure 3 shows a simplified UML description of this conceptual model, where it is possible to see that the use of the object-oriented paradigm allows easy integration of existing tools as well as system extension.

In accordance with Figure 3, the *UserInterface* class is the system controller, responsible for event management, while the *View* class is responsible for data presentation. In this way, the model, i. e. the *Scene* class composed by *Camera*, *Light* and *GraphicObject* lists, is totally independent of a specific platform. Interaction tools, such as *CutTools*, *SegmentationTool* and *QuantizationTools*, are associated with the *View*, since they need a graphical representation at the interface and invoke image generation and model changes. The detailed behavior of these relationships is still being defined. The *RayCasting* class has a set of methods that implement the different visualization algorithms (e. g. intermixing, side by side, MIP) and internal structure visualization. Registration is just a method of the *Volume* or *Image2D* class, which is also responsible for data loading. To simplify the diagram presented here, basic classes such as *Point*, *Vector* and *Color* are not included.

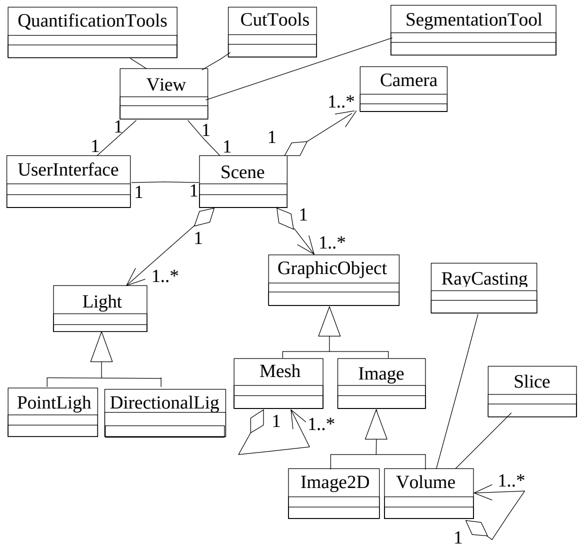

Figure 3 – Simplified conceptual model

It can be noticed that the big problem to keep the functional core independent of the user interface in interactive systems architecture specification is solved here by the MVC implementation. Since the core of interactive systems is based on its functional requirements, it usually remains stable. User interfaces, however, are often subject to change and adaptation. So, it is necessary to develop an architecture that support the adaptation of user interface without causing major effects to application-specific functions or the data model underlying the software [30]. Then, the class library was designed in such a way that the entire interface is concentrated at *UserInterface* and *View* classes, and the other tools and functions remain completely independent from the interface. With these classes, we are providing the fundamental structural organization for an interactive multimodal visualization system.

To illustrate objects interaction, a collaboration diagram is presented in Figure 4. In this example the user changes the visualization type to MIP. At first the event is detected; then one message is sent for the *GraphicObject*, i. e. *Volume*, notifying that it has to modify itself. After changes, a notification is sent to the *Scene*, which then notifies the changes to the *View* class that is responsible for image re-exhibition, and to the *UserInterface* class.

As mentioned before, optimized algorithms recently described in the literature can be easily included in this architecture. First, we can consider interactive segmentation, which is included in the framework as a tool to generate segmented volumes that can be used with visualization and quantification techniques. Regarding this, we are planning to adapted the 2D technique developed by Olabarriaga[22] to work with volume slices.

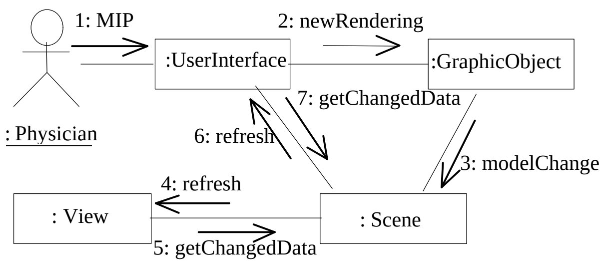

Figure 4 – Collaboration diagram

One indispensable pre-processing step is the registration algorithm implementation. As described in the previous section, the most used technique nowadays is mutual information. So, such algorithm to make automatic registration is being implemented. We are also studying some alternatives to improve its performance, as for example to ask for the user to interactively make an initial matching, in order to reduce the processing to maximize the mutual information [6, 5].

The multimodal visualization algorithm that will be implemented is based on the classical ray casting technique with many improvements, concerning to illumination, parallel and perspective projection, among other optimization techniques. At present, we are integrating the algorithms from RendexVox into the presented architecture. RenderVox is a very efficient in-house system developed as a Master Thesis [8]. It is also planned to develop an algorithm for the visualization of interior structures based on direct volume visualization. A good candidate for extension is the technique called *Confocal Volume Rendering*, presented by Mullick *et al.* [32] that enables the user to visualize interior structures in one data set by just controlling physically defined parameters, without performing segmentation.

Although the ray casting algorithm is still very computational intensive, this will not be considered as a limiting factor since this problem have been reduced with the development of parallel architectures and special purpose graphics hardware designed specially for fast manipulation of volume data [23]. In fact, since each ray is processed independently, we are planning to develop a new class to allow system execution in a parallel architecture. Thus it will probably be possible to generate the images in real-time. The only disadvantage in this case is that this kind of class will be totally dependent of an architecture that usually is not available in the environment work of the physicians.

Finally, it is important to point out that in order to develop a system as portable as possible, using free software, we choose to implement this framework using standard C++ language, OpenGL and FLTK toolkit [33]. C++ has been showed to be the most suitable programming language, since it is possible to use just the ANSI libraries and compile the same code in different platforms. Zuiderveld *et al.* [21] made an evaluation of its utilization. OpenGL is used for the 3D interface while FLTK is used for the GUI development. FLTK is portable, developed over the GNU Library General Public License and has an optimized code, considering performance and size.

5 FINAL COMMENTS

This paper presented a brief description of the multimodal visualization research area, which has the goal of finding suitable strategies to integrate important characteristics of multiple data sets into one image such that better insight can be provided [23]. We also focused on the necessary requirements to the development of an interactive multimodal visualization system. The architecture proposed in Section 4 was designed to fit those requirements.

Some systems were already developed to allow interactive multimodal visualization, but a big challenge still remains: the development of techniques to integrate registration, quantification, and interactive segmentation and visualization, including the visualization of local features in multimodal data sets. According to Johnson [34] interactive visualization systems need to be:

a) modular and easy to extend, as proposed in our object-oriented architecture design;

b) adaptable to hardware ranging from the largest computing systems to low-end workstations and PC's, implemented in the proposed architecture with the use of MVC pattern;

c) demonstrably usable in medical scientific research, what we expect to reach with the provided system functionality.

Previous work has focused on the development of optimized registration and visualization algorithms, in order to allow user interaction in real-time. With the proposed architecture, it will be possible to use newer optimized algorithms as soon as they come out, in such a way that the user can integrate them with several visualization and interaction tools without having to rebuild entire modules.

Acknowledgements – This research was partially supported by CAPES, CNPq, FAPERGS and UFRGS.

6 REFERENCES

[1] J. B. A. Maintz, M. A. Viergever, Survey of Medical Image Registration. *Medical Image Analysis*, 2(1), 1998, 1-36.

[2] J. P. W. Pluim, J. B. A. Maintz, *Registration of Multimodal medical images*. Available at

http://www.cv.ruu.nl/Research/Registration/registration-frame.html (March 2000).

[3] R. Stokking, *Integrated Visualization of Functional and Anatomical Brain Images* (Netherlands: Universiteit Utrecht, 1998, Ph.D. Thesis).

[4] R. A. Robb, Visualization Methods for Analysis of Multimodality Images. In: R. W. Thatcher, M. Hallet, T. Zeffiro *et al.* (eds). *Functional Neuroimaging: Technical Foundations*. Academic Press, 1994.

[5] F. Maes, *Segmentation and Registration of Multimodal Medical Images* (Belgium: Katholieke Universiteit Leuven, 1998, Ph.D. Thesis).

[6] A. Collignon, *Multi-Modality Medical Image Registration by Maximization of Mutual Information* (Belgium: Katholieke Universiteit Leuven, 1998, Ph.D. Thesis).

[7] Computer Graphics and Image Processing Group, *VPat - Visualization and interaction with Virtual Patients*. Available at http://www.inf.ufrgs.br/gpesquisa/cg/projects.html#vpat (July 2000).

[8] M. R. M. Silva, *Ray Casting Based Volume Visualization Techniques* (Porto Alegre: Instituto de Informática/UFRGS, 2000, Master Thesis).

[9] J. K. Udupa, G. T. Herman, *3D Imaging in Medicine* (CRC Press LLC, 2000, 2th ed.).

[10] J. B. A. Maintz, *Retrospective Registration of Tomographic Brain Images* (Netherlands: Universiteit Utrecht, 1996, Ph.D. Thesis).

[11] J. B. A. Maintz, E. H. W. Meijering, M. A. Viergever, General Multimodal Elastic Registration Based on Mutual Information. SPIE, 1998, San Diego, California. *Proceedings...* 1998, 144-154.

[12] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, Interpolation artefacts in mutual information based image registration. *Computer Vision and Image Understanding*, 77(2), 2000, 211-232.

[13] A. Collignon, F. Maes, D. Delaere, *et al.*, Automated multi-modality image registration based on information theory. *Information Processing in Medical Imaging*. Kluwer Academic Publishers, 1995, p. 263-274.

[14] P. Viola, W. M. Wells III, Alignment by maximization of mutual information. International Conference On Computer Vision, *Proceedings...* IEEE Computer Society Press, 1995, 16-23.

[15] W. M. Wells III, P. Viola, H. Atsumi *et al.*, Multi-Modal Volume Registration by Maximization of Mutual Information. *Medical Image Analysis*, 1 (1), 1996, 35-51. Available by ftp.ai.mit.edu:/pub/sw/papers/mia-95-dist-color.ps.gz.

[16] M. Levoy, Volume Rendering - Display of Surfaces from Volume Data. *IEEE Computer Graphics & Applications*, 8(3), 1988, 29-37.

[17] M. Levoy, Efficient Ray Tracing of Volume Data, *ACM Transactions on Graphics*, 9(3),1990,245-261.

[18] M. Levoy, A Hybrid Ray Tracer for Rendering Polygon and Volume Data, *IEEE Computer Graphics and Applications*, March 1990, 245-261.

[19] H. J. Noordmans, *Interactive Analysis of 3D Microscope Images* (Netherlands: Universiteit Utrecht, 1997, Ph.D. Thesis).

[20] B. Lichtenbelt, R. Crane, S. Naqvi, *Introduction to Volume Rendering* (Prentice Hall, 1998).

[21] K. J. Zuiderveld; A. H. J. Koning, R. Stokking, *et al.*, Multimodality Visualization of Medical Volume Data. *Computers&Graphics*, 20(6), 1996, 775-791.

[22] S. D. Olabarriaga, *Human-Computer Interaction for the Segmentation of Medical Images*. (Netherlands: Universiteit Utrecht, 1999, Ph.D. Thesis).

[23] K. J. Zuiderveld, *Visualization of Multimodality Medical Volume Data using Object-Oriented Methods* (Netherlands: Universiteit Utrecht, 1995. Ph.D. Thesis).

[24] P. Hastreiter, T. Ertl, Integrated Registration and Visualization of Medical Image Data. CGI, 1998, Hannover, Germany. *Proceedings...* Available at http://www9.informatik.uni-erlangen.de/Persons/Hastreiter/deutsch.html.

[25] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, Mutual Information Matching in Multiresolution Contexts. F. Pernus, S. Kovacic, H. Stiehl, M. A. Viergever, (Eds.). *International Workshop on Biomedical Image Registration*, 1999, 46-60.

[26] M. E. Leventon, *A Registration, Tracking, and Visualization System for Image-Guided Surgery*. (MIT Master's Thesis, 1997). Available at http://www.ai.mit.edu/people/leventon/Research/SM-Thesis/thesis.pdf.

[27] W. E. L. Grimson, R. Kikinis, F. A. Jolesz, *et al.*, Image-Guided Surgery. *Scientific American*, June 1999, 54-61.

[28] J. M. Fitzpatrick, D. L. G. Hill, Y. Shyr, *et al.*, Visual Assessment of the Accuracy of Retrospective Registration of MR and CT Images of the Brain. *IEEE Transactions on Medical Imaging*, 17(4), 1998, 571-585.

[29] J. West, J. M. Fitzpatrick, M. Y. Wang *et al.*, Retrospective Intermodality Registration Techniques for Images of the Head: Surface-Based Versus Volume-Based. *IEEE Transactions on Medical Imaging*, 18(2), 1999, 144-150.

[30] F. Buschmann, R. Meunier, H. Rohnert *et al.*, *Pattern-Oriented Software Architecture, A System of Patterns* (John Wiley & Sons, 1996).

[31] C. Larman, *Applying UML and Patterns – An Introduction to Object-Oriented Analysis and Design* (Prentice-Hall, 1997).

[32] R. Mullick, N. Bryan, J. Butman, Confocal Volume Rendering: Fast Segmentation-Free Visualization of Internal Structures. SPIE, 2000, San Diego, California. *Proceedings...* 2000.

[33] B. Spitzak, *FLTK - The Fast Light Tool Kit Home Page*. Available at http://fltk.easysw.com (March 2000).

[34] C. Johnson, S. G. Parker, C. Hansen *et al.*, Interactive Simulation and Visualization. *Computer*, 32(12), 1999, 59-65.