A HUMAN-MACHINE INTERFACE FOR MEDICAL IMAGE ANALYSIS AND VISUALIZATION IN VIRTUAL ENVIRONMENTS

Christian Krapichler, Michael Haubner, Andreas Lösch, Karl-Hans Englmeier

GSF - National Research Center for Environment and Health
Institute of Medical Informatics and Health Services Research
Ingolstädter Landstraße 1, D-85764 Neuherberg, Germany
krapichler@gsf.de

ABSTRACT

Virtual worlds open new dimensions in human-machine and even human-human communication. Medicine is predestined to benefit from this new technology in many ways. For the field of visualization and analysis of tomography data, an application is introduced which expedites identification of spatial coherencies and exploration of pathological regions. To facilitate work in such an environment and to avoid long periods of accustoming, a human-oriented interface is required allowing physicians to interact as close to the real world as possible. Hand gesture recognition (with a data glove) and eye tracking (using biosignals) are essential parts to fulfil this demand. Their integration into the virtual environment as two components of the human-machine interface is presented.

1. INTRODUCTION

Evaluation of three-dimensional anatomical data, as they are obtained by techniques like Magnetic Resonance Imaging (MRI) or Computed Tomography (CT), is an elaborate and time-consuming process. Though there is remarkable success in accelerating this task by automation [1], it still demands human interaction [2].

Specialists can be supported in this work through modern means of man-machine communications. Virtual environments offer new ways of image analysis and visualization [3]. Three-dimensional image representation empowers radiologists to recognize spatial coherencies in a much easier and faster way than today‘s arduous analysis of two-dimensional sequences of tomographical image sets. Immersion into a virtual environment facilitates exploration of a scene of three-dimensional medical objects or tomograms.

Besides of high-performance hardware environments and visualization applications, a new and efficient human-machine interface is required to enable analysis of and interaction with the 3D data [4]. Such an interface should support control of image processing and segmentation algorithms as well as offering numerous tools for analyzing, manipulating and representing both volume- and surface-oriented structures. And in spite of the desired variety of features, usage has to be intuitive, natural and convenient. Therefore, interaction devices should recognize and interpret human actions and gestures and offer, in opposite to standard devices like keyboard or 2D-mouse, operations in six degrees of freedom. In an ideal case users can interact in virtual worlds in the same way as they would do in reality.

In this paper a combined tracking of two major organs of human interaction – hands and eyes – and integration into a medical imaging environment is shown.

2. HARDWARE CONFIGURATION

The entire interface is based on the following equipment:

- *Computer System:*
  SGI Onyx Infinite Reality (dual pipe, 12 R4400 CPUs, 2GB RAM, 4 RMs);
- *Display systems:*
  n-Vision head-mounted display (HMD);
  Fakespace BOOM 3C (interactive stereo viewer);
  TAN stereo projection system (polarized filtering);
- *Interaction devices:*
  Fifth Dimension Technologies 5th Glove ‘95;
  Ascension 3D-mouse;
  DLR Space Mouse;
- *Tracking systems:*
  Ascension Flock of Birds (normal and extended range);
  BioControl BioMuse eye tracking system.

3. HAND GESTURE RECOGNITION

Hands play a preeminent role in natural human interaction [5]. Recognizing and interpreting hand gestures and actions aims at improving intuitivity and efficiency of human-machine interfaces. Hence, a multiplicity of tools and actions is offered for image analysis and representation, controlled via hand gestures.

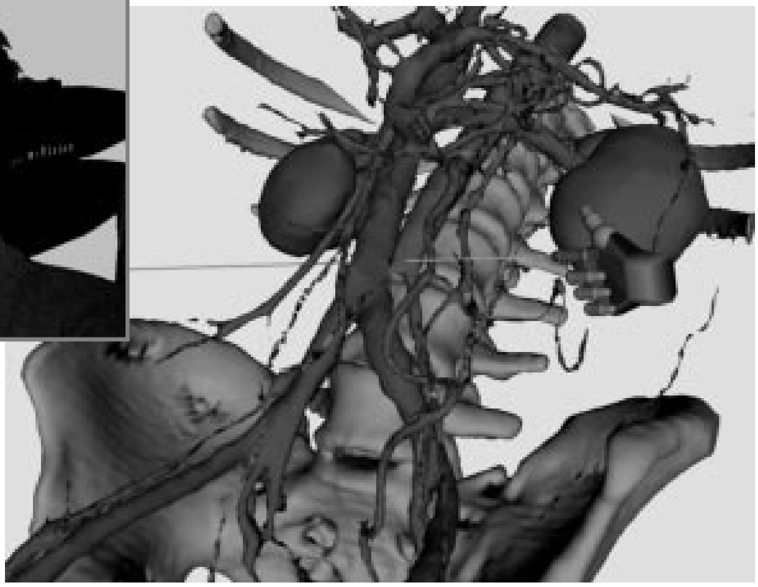

Figure 1: Performing a *pointing* gesture with the data glove (left picture) enables selection of segmented objects within a VR scene, e.g. an arterial occlusion inside the abdomen (right picture).

3.1 Methods

Finger flexions are measured with a data glove (*5th Glove ‘95* from Fifth Dimension Technologies) containing optical fibres, hand and head positions are tracked magnetically (*Flock of Birds* from Ascension).

All data glove positions and motions are reflected by a virtual hand, showing current finger bends. There is a set of predefined gestures (e.g. *pointing*, *grasping*, *navigation*, *release*), each of it assigned to an explicit action (→fig. 1).

Degrees of finger flexions are measured in a range from 0-255 units. Bending of finger joints is not determined separately. There is one measured value per finger, representing current finger flexion as a whole. Consequently, reference gestures and current flexion states are represented as five-dimensional vectors. Comparing angle and size of measured vectors versus those of all calibrated gestures, the current vector is assumed to represent that reference gesture which shows least deviation in angle and size and stays within a predefined tolerance range.

For both hand and head, position and orientation is measured. Accuracy of magnetical tracking is about 2.5mm or 0.5˚ RMS, respectively, with corresponding resolutions of 0.8mm/0.1˚.

Head tracking values specify the viewer’s location and orientation inside the virtual scene. Thus, peering and roaming in a scene is supported. Hand tracking values in respect of head tracking determine the virtual hand’s position and orientation. Hand tracking values are also applied to move and manipulate tools or objects while the corresponding gesture is recognized. For collision detection an algorithm is implemented which uses oriented bounding box hierarchies (OBB trees [6]) for intersection tests.

3.2 Applications

Medical image segmentation is an important step in computer aided analysis of tomography data. As this work still requires human interaction, it is supported and facilitated. One problem of segmentation algorithms emerges when two areas with similar Hounsfield values, but belonging to different objects, overlap in one or more layers of the tomogram. In such cases most volume growing or edge detection algorithms recognize those two objects as one. To correct such misinterpretations, it is necessary to segment those areas manually in the affected regions. Avoiding time-consuming layer-by-layer editing, selection of a three-dimensional barrier (plane, hemisphere, cube) is supported to place it at the according location and adapt it by scaling, rotation and translation. Segmentation algorithms are prohibited to exceed these boundaries, only the desired object is detected.

Navigation is possible by simply walking around, by additional devices like a space mouse, or with a specific hand gesture. While performing such a *navigation* gesture, hand orientation is transformed to a change of viewing position and orientation, evoking an impression of flying through the scene. Thereby, specialists can examine all relevant parts of the scene from different angles and positions, for example in virtual bronchoscopy (→fig. 2f).

By selecting and moving (or rotating) a medical object of interest, it can be examined without the need to walk around or fly through it. This is the natural way we are looking at things: take them and turn them in our hands to inspect them from every side. This means an improvement in diagnosis of complicated pathological mutations, e.g. tumours or compound bone fractures.

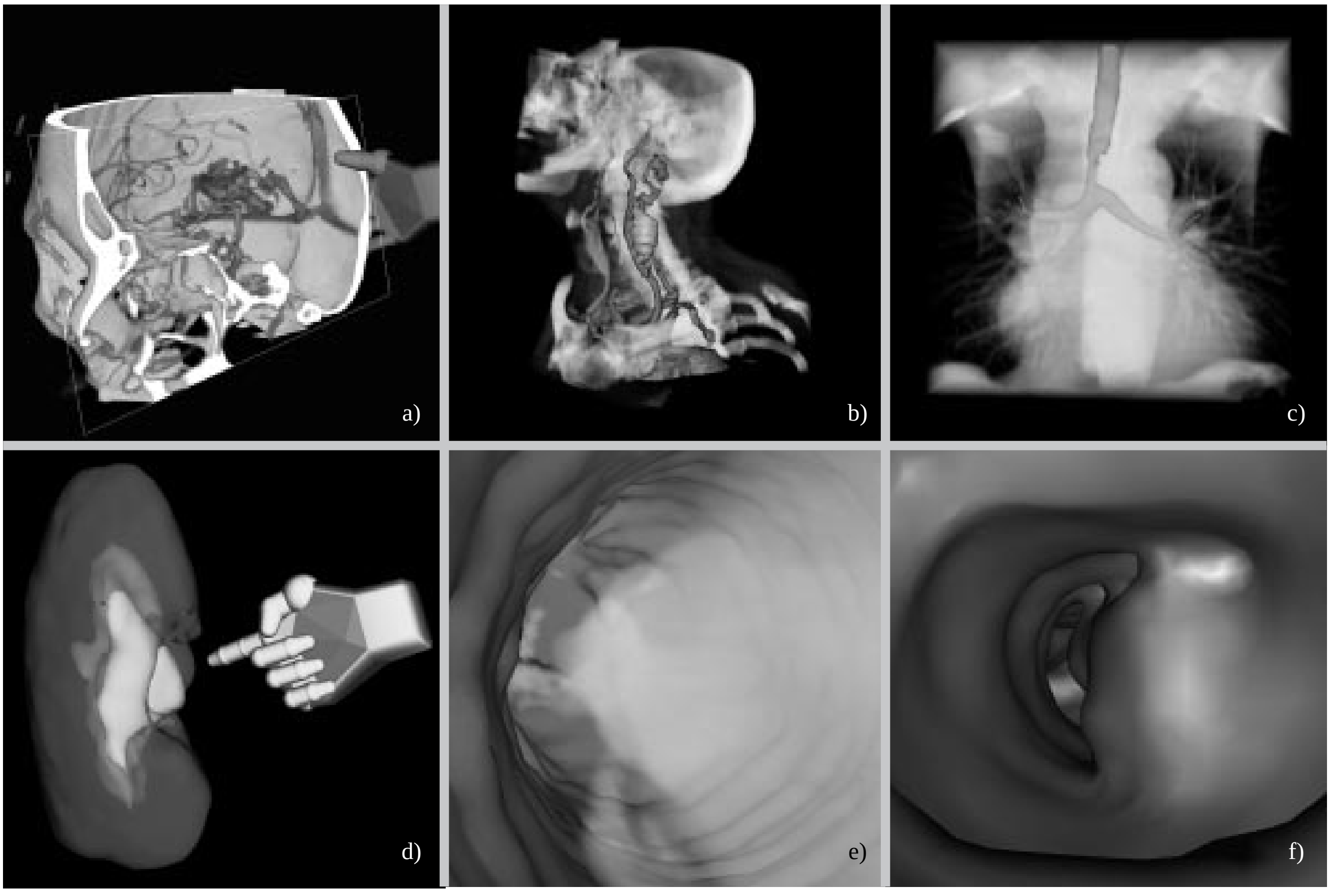

Figure 2: (a): Moving a clipping plane through the volume rendered skull opens view on a brain tumour; (b), (c): Hybrid visualization: vessels and bronchial tree are triangulated surfaces, rest is volume rendered; (d): a renal calculus, visible due to transparent representation of the kidney; (e): virtual angioscopy: analysis of blood flow (enhanced with contrast medium, volume rendered bright areas), inside the surface rendered aorta, restricted by a tumour (area without blood flow, switched to transparent state); (f): virtual bronchoscopy: view inside the trachea of figure 2c, from larynx to bifurcation, recognizing a tumour on the right.

Concave objects which hide other objects inside could be removed from the virtual scene. But to conserve information about location and orientation of the internal structure with respect to the hiding object and to improve orientation inside the scene, there are two possibilities: moving a clipping plane manually through the outer object (→fig. 2a), or switching the outer object with a single gesture to a transparent state (→fig. 2d).

Another advantage is the ability to show both original volumetric data and segmented objects (→fig. 2b,c). Such a hybrid approach is very helpful for diagnosis [7]. For example, it is necessary to examine the internal structure of a renal calculus to predict the result of an ultrasound treatment. Therefore, specialists can cut the calculus and analyze the original tomographical data inside the segmented object. As another example, blood flow can be explored in virtual angioscopy (→fig. 2e).

4. Eye Tracking

Knowledge about the user‘s current visual focus and its motion can be applied to control actions or tools, to adequately display VR scenes, or to log the eye path for later evaluation.

4.1 Methods

Eye positions are detected with an electro-oculography (EOG) system (*BioMuse* from BioControl [8][9]), measuring biosignals due to an electrostatic field inside each eye. A signal processor converts electrode signals to four data streams, representing horizontal and vertical positions of each eye.

Advantages of this method are the limited computing performance required for data processing, allowing real-time applications, and the possibility to wear the equipment in combination with a VR display, e.g. a HMD [10]. A major disadvantage of EOG is the sensitivity to some sources of interference, e.g. variing skin temperature or eye lid movements [11][12]. Besides of standard signal processing methods, errors are reduced by a combined evaluation of eye tracking and hand-controlled actions. As motion of selected objects requires attention on those objects, they will be focused most of the time. Hence, object coordinates are used to permanently recalibrate drifting EOG signals.

To determine screen coordinates of the visual focus, the system has to be calibrated to get the eye position in relation to the display and the range of eye movements. As the HMD position is constant in relation to the viewer‘s eyes, it is not necessary to evaluate head tracking data for focus calculations. This is required for applications which allow users to move in front of a screen, e.g. a stereo projection system.

4.2 Applications

From a user‘s point of view there are two ways of applying eye movements to virtual environments: *active* control, e.g. moving an object or selecting a tool, or *passive* control, i.e. eye movements are evaluated in the background, without the user noticing any effects or reactions.

For active control, some gestures would be necessary. But besides of interpreting eye blinking e.g. as different button clicks (left eye, right eye, both eyes), which would be quite annoying in practice, eyes are not able to perform gestures. Eye tracking has to be combined with other input sources (e.g. speech recognition) for active control.

As active control of objects via eye movements is not the natural way of interacting with things in the real world, passive evaluation seems to be more reasonable. As a first step, eye movements influence image representation: Visual focus is used as parameter for a level-of-detail algorithm. It allows different levels of detail even for the same object. Only the region of interest is displayed accurately, peripheral regions show less details. Thus scene complexity is reduced and higher frame rates can be achieved, without perceivable loss of accuracy.

5. Conclusion

Virtual Reality offers a powerful environment for visualization and analysis [13]. Extended by an effective man-machine interface, it allows specialists to analyze and evaluate three-dimensional data in less time and in a less elaborate way, but with the required high precision.

Especially in medicine, where the basis of all data –the human beeing– exists in a three-dimensional world, virtual environments seem to be predetermined to support physicians in diagnosis, therapy, surgical planning, simulation or education. Therefore it is likely to become a standard tool for medicine of the future [14][15]. But acceptance of this new environment depends, among others, largely on the offered tools and features and on the intuitivity and convenience of the human-machine interface.

References

[1] N. Ayache et al., “*Steps toward the automatic interpretation of 3D images*”, 3D Imaging in Medicine, NATO ASI Series F60, pp. 107-120, 1990.

[2] M. W. Krueger, “*Automating Virtual Reality*”, IEEE Comp. Graph. & Applications, vol. 15(1), pp. 9-11, Jan. 1995.

[3] S. R. Ellis, “*What are Virtual Environments?*”, IEEE Comp. Graph. & Applications, vol. 14(1), pp. 17-22, Jan. 1994.

[4] K.-F. Kaltenborn, O. Rienhoff, “*Virtual Reality in medicine*”, Meth. Inform. Med., vol. 32, pp. 407-417, 1993.

[5] D. J. Sturman, D. Zeltzer, “*A survey of glove-based input*” IEEE Comp. Graph. & Applications, vol 14(1), pp. 30-39, Jan. 1994.

[6] S. Gottschalk, M. C. Lin, D. Manocha, “*OBB-Tree: A Hierarchical Structure for Rapid Interference Detection*”, Proc. ACM Siggraph ‘96, pp. 171-180, 1996.

[7] K.-H. Englmeier et al., “*A Method for combining fast surface and exact volume visualization technique in medicine*”, Proc. SPIE Medical Imaging, vol. 2164, pp. 22-33, 1994.

[8] R. B. Knapp, H. S. Lusted, “*Real-time digital signal processing system for bioelectric control of music*”, Proc. IEEE ICASSP ‘88, pp. 2556-2557, 1988.

[9] H. S. Lusted, R. B. Knapp, A. M. Lloyd, “*Biosignal processing and biocontrollers*”, Virtual Reality Systems, vol. 1(1), pp. 38-39, 1993.

[10] H. S. Lusted, R. B. Knapp, “*Medical applications for biocontroller technology*”, Proc. Medicine meets Virtual Reality II: Interaction Technologies & Healthcare, pp. 45-47, 1994.

[11] W. Barry, G. Melvill Jones “*Influence of eye lid movement upon electro-oculographic recording of vertical eye movements*”, Aerosp. Med., vol. 36, pp. 855-858, Sep. 1965.

[12] L. R. Young, D. Sheena, “*Survey of eye movement recording methods*”, Behav. Res. Methods Intrum., pp. 397-429, 1975.

[13] W. Ribarsky et al., “*Visualization and Analysis Using Virtual Reality*”, IEEE Comp. Graph. & Applications, vol. 14(1), pp. 10-12, Jan. 1994.

[14] R. M. Satava, “*The role of Virtual Reality in medicine of the 21st century*”, Virtual Reality Systems, vol. 1(2), pp. 43-52, Mar. 1993.

[15] S. Weghorst, “*Virtual Reality in medicine*”, AI in Medicine, vol. 6, pp. 277-279, Aug. 1994.