The Volume Slicing Display: a tangible interface for slicing and annotation of volumetric data

Alvaro Cassinelli, Yoshihiro Watanabe, and Ishikawa Masatoshi

Ishikawa-Oku Laboratory, Dept.of Information Physics and Computing
The University of Tokyo
E-mail: Alvaro_Cassinelli@ipc.i.u-tokyo.ac.jp

The Volume Slicing Display (VSD) is a device enabling intuitive exploration and annotation of volumetric data (e.g. medical images) using a piece of Plexiglass that functions both as a passive, unthethered *projection screen* and as an *control interface* for free-hand annotation and slice selection.

1. Introduction

The motivation behind this project is to facilitate the analysis and annotation of volumetric image data. The method consists on generating 2d slices that appear *physically* at the right location in a three-dimensional reference frame associated to the virtual volume (Fig.1). This is done using a (passive) display that also functions as the volume slicing tool (Fig.2) and as a free-hand input surface (Fig.3). As a consequence, exploration of 3d data is rendered more intuitive thanks to the integration of *proprioceptive* information: the user intuitively "knows" the location and pose of the slice, as if he/she were in direct contact with the virtual object (the VSD is an example of a "Tangible User Interface" [1]). Even if this system does not generate stereoscopic views, radiologists would be able to retrieve a certain amount of three-dimensionality from an X-ray plate by freely orienting the (unthethered) screen in space, instead of relying on a negatoscope to form a mental representation of the volume, or using secondary hardware interfaces such as a mouse, knobs or keyboard to select a planar slice to be rendered on a standard monitor.
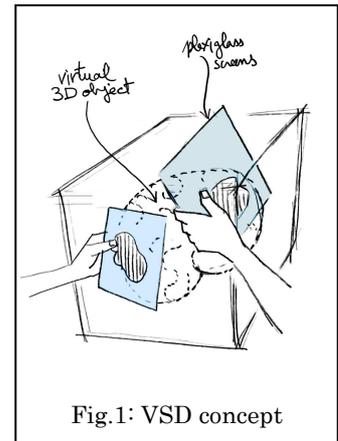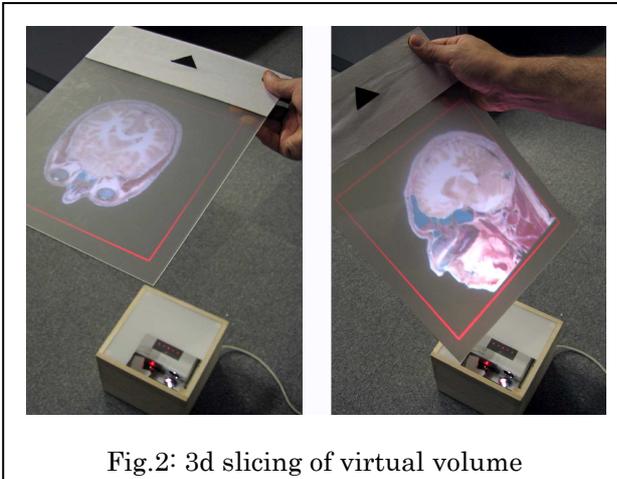

Fig.1: VSD concept
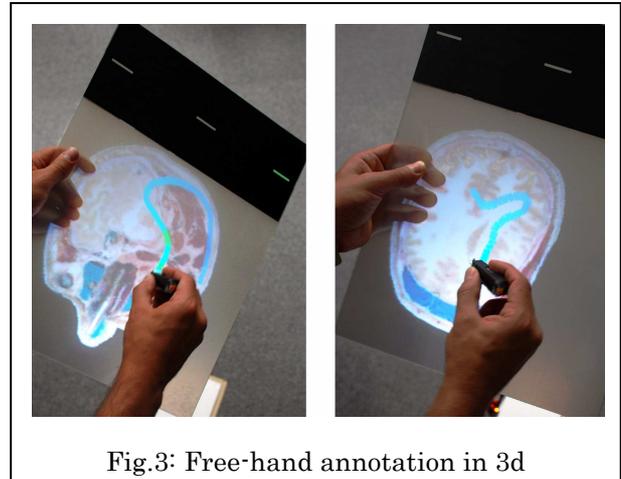

Fig.2: 3d slicing of virtual volume

Fig.3: Free-hand annotation in 3d

2. Prototype

This paper presents a second prototype of the VSD conceived with constraints of portability and cost in mind, and deployed over commercially available off-the-shelf hardware. A first prototype used a custom-made vision chip to continuously track four infrared LEDs at the corners of the screen (corners were discriminated by their different blinking sequences), and a closed solution for the P4P problem was used to estimate the pose in real time [2]. The new prototype presented here uses instead custom retro-reflective ARToolkit markers [3] to compute the pose of a piece of translucent Plexiglass (Fig.4). The computer then calculates the corresponding slice from a virtual volume 'hovering' in mid-air, and prepares the content to be displayed by applying
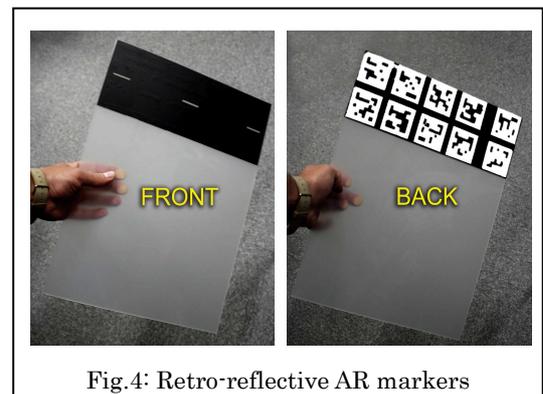

Fig.4: Retro-reflective AR markers

the proper perspective correction. All the hardware (a small LED-based projector, an infrared camera, infrared LED sources and a Mac-Mini with an Nvidia GeForce 9400M graphic card having 256 MB for storing the volumetric data as a *3d texture*) is contained in a small box measuring only 35x35x35cm (see Fig.2). The interaction space is a cone that extends from this box (placed on the floor) towards the ceiling, so that at adult hand level, the interaction *area* represents about 1 square meter. Camera/projector calibration is performed off-line, using a real and a projected checkerboard and the Matlab Stereo Calibration Toolbox [4].

3. Functions and Modes of Operation

Markers are simultaneously used as fiducials for pose estimation as well as discrete switches for toggling between different *operation* and *rendering* modes. A switch is activated by briefly covering the corresponding marker with one hand, from below the screen (this does not affect the pose estimation process because we use a redundant configuration of ARToolkit markers). Secondary uses of the markers are for position reset, zooming, or taking a snapshot of the current slice for printing purposes. Finally, the present prototype enables free-hand annotation and drawing (in 3d) using an IR pen over the Plexiglass screen. We detail these functions in the following paragraphs:

- **Modes of operation**. There are two basic modes of operation: *volume lock* and *volume drag*. In the first mode, the virtual volume being explored is locked to the world reference frame, and the screen acts then as a "slicing surface" tool. In the second mode, the virtual volume follows the screen motion (the displayed slice remains unchanged during this operation). This is useful when showing others a particularly interesting slice for scrutiny, or when there is a need to rotate or translate the volume in space (this is certainly needed in the case of a single projector/camera system, because a slicing plane perpendicular to the projector axis cannot be appropriately rendered). Therefore, using the *volume drag* mode, one can easily orient the volume to produce sagital, coronal or horizontal slices, without exaggerated tilting of the screen (see videos in [5]).
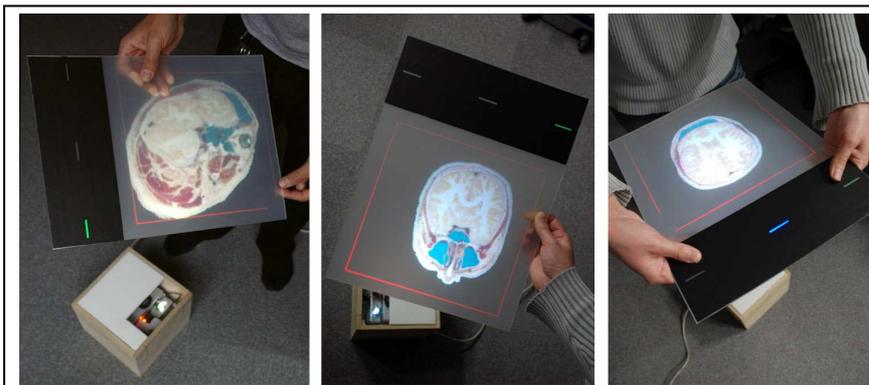

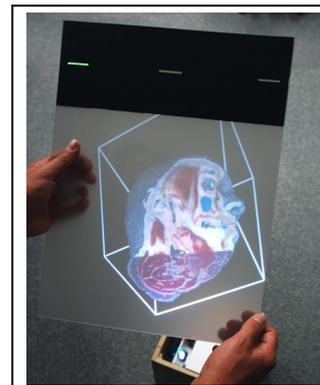Fig.5: slice / drag / reset (a simulated indicator is shown on the screen).

Fig.6: Volumetric rending

- **Rendering Modes**. It is also possible to change the type of visualization, from a unique *2d slice* to a *rendered volume* (Fig.6). Volume rendering presently assumes that the head of the user is right above the screen, at a distance of about 50 cm, so that the projection looks right for the user (in the future we plan to produce a real "fish-tank" VR effect by tracking the user head).

- **Other uses of markers**. Markers can function as independent switches, or as "chorded" switches. For instance, covering at the same time the upper left and right markers (which are independently used to select the mode of operation) produce a reset of the virtual volume position, in particular enabling the adjustment of its height with respect to the user. Covering the lower left and right markers, produce a gradual zoom of the current slice. Finally, separate markers can be brought over the Plexiglass screen and function as physical widgets (as done in [6] and [7]). We have demonstrated for instance *continuous zooming* by rotating a special maker over the screen (see Fig.7).
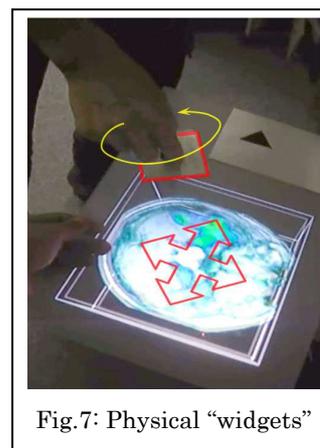

Fig.7: Physical "widgets"

- **3d Annotation**. It is possible to "write" freely on the screen using an IR pen. This generates 3d trajectories that remain attached to the reference frame of the virtual volume. This is useful to highlight regions of interest, produce alphanumeric annotations, supervise image segmentation, or indicate a surgical path to be checked in (volume-registered) tomographic samples of a different nature (bones, soft tissue, etc.).

Additional functionality is possible using a (wireless) keyboard to toggle and display angles or 3d grids. Most importantly, it is possible to change the *content* being displayed altogether. For instance, it is possible to first explore a volumetric CT scan, create a trajectory over it, and then switch to soft-tissue rendering (from MRI for instance - see Fig.8) or other source of data (PET scan, etc). This could be useful for modifying a surgical path that would appear compatible with, say, bone structure, but that intersects vascular structures better visualized in an angiographic scan.

3. Discussion

This system represents a logical evolution of our previous research on the 'Khronos Projector' [8] and the 'Deformable Workspace' [9], both of which explored the paradigm of a deformable screen through which the user is capable of 'sculpting' or 'slicing' a virtual volume of data (in particular stacked video frames). Although the screen was deformable, its borders were fixed and it was impossible to rotate the screen with respect to the volume being sliced. The Volume Slicing Display enables this, but the screen is rigid. Although we have been working towards a screen that is both orientable *and* deformable using structured laser light to compute the deformation in real time [10], we feel that this complex mode of interaction is only necessary in a limited number of applications. Since this project was started similar systems were built by other researchers (see in particular [11] and [12]); however, the prototype proposed here is particularly compact, can be built with off-the-shelf hardware, and presents the advantage that the screen is a passive Plexiglass screen (or even a simple translucent sheet of paper) with a set of printed markers that serve for pose estimation and control, but that can also be used
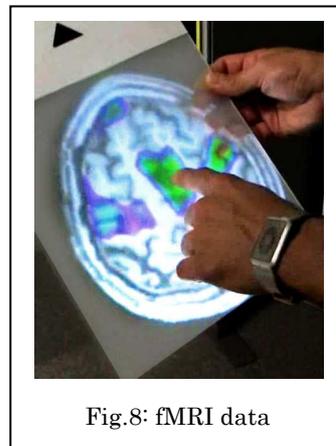

Fig.8: fMRI data

to *identify the patient in a confidential manner*. Annotation in 3d space is another singular feature of our system; the recorded 3d trajectories are associated to the virtual volume and can be saved for latter analysis using more classic volume rendering software and hardware. The present prototype does not have enough resolution or stability to be of immediate use in clinical medicine; however, it has been pointed out by professionals that it can come handy for teaching anatomy, or facilitate preliminary analysis of fMRI data in the field of neurosciences.

4. Conclusion and further research

This experimental visualization interface may one day enable a team of experts (surgeons, geologists, designers or architects) to explore 3D virtual objects as if these co-existed in real physical space, as well as to explore them interactively using a simple piece of paper that functions both as a control interface (through changes in its shape or by drawing on it) and as a passive, untethered projection screen. We are now pursuing usability tests comparing the efficiency of this interface with more classic volume rendering software and hardware (for the discovery of hidden features, help in supervised segmentation, etc). On the other hand, we are contemplating several attractive improvements and evolutions of the system, the most important being the possibility of enlarging the interaction space. Indeed, the present prototype uses a single projector; this reduces the working space as well as the maximum possible inclination of the screen. We are now considering the use of multiple projectors to make the workable space as large as a

Fig.9: multiple screens

whole room. Related to this is the need for a simple auto-calibration procedure for multiple projectors and cameras. This could be done by active marker tracking as in [13], using stable markers or feature points on the projection screen. Alternatively, one can imagine a system using multiple passive screens, each showing slices from different (but registered) tomographic sources (see Fig.9). This could accelerate the analysis and understanding of particularly complex structures. Another goal is to make the markers invisible to the naked eye (we have tried using infrared ink, but the contrast is not sufficient). This will allow for a clear and large projection surface. Since the present prototype uses an infrared source collinear with the projector, an interesting possibility is to use a sanded plastic semi-transparent sheet which will couple light into the screen. Markers can then be engraved on the surface, and will appear brighter than the rest of the screen. Interestingly, this configuration will also permit FTIR-based tracking of fingers over the surface, as done in [14]. This will enable zoom and rotation of images using *bare fingers*, as well as annotate and draw trajectories (such as surgical paths) in space without the need of an IR pen. Among other interesting directions that this project highlights, is the development of an "origami"-like user interface, in which different shapes and folds of a flexible screen are interpreted by the machine as specific displaying commands (this could be done using polarized light to detect pressure and bending of plastics or silicone surfaces as in [15]). As noted previously, "fish-tank" VR would be feasible if we could track the user's head - this would be relatively simple in the "room" configuration using several cameras (these could be recruited to perform robust head and gaze –direction tracking). In this situation, the screen would function as a "window" to a 3d space (as has been done in the

"Deskrama" project [11] and in our own previous research on the 'Parallax Augmented Desktop' [16]). In this situation, using a pair of polarized projectors it could even be possible to render stereoscopic imagery. (Interestingly, a very cost effective implementation of this fish-tank capable VSD could be implemented using the front and rear cameras in the second generation iPad: the rear camera could be used for pose estimation (using SLAM), and the front camera could be used to track the user head.)

5. Acknowledgements

References

[1] Ishii H.: *The tangible user interface and its evolution*. Communications of the ACM , Volume 51 Issue 6 (2008)

[2] Ito T., Cassinelli A., Komuro T. and Ishikawa M.: *3D Object Representation Using a Tangible Screen*. Proc. of the SICE2006 conference (2006)

[3] Kato H., Billinghurst M.: *Marker tracking and HMD calibration for a video-based augmented* reality conferencing system. In Proc. 2nd IEEE and ACM International Workshop on Augmented Reality IWAR99 (1999)

[4] Bouguet J.Y.: *Camera Calibration Toolbox for Matlab*, (Computational Vision at Caltech), web: http://www.vision.caltech.edu/bouguetj/calib_doc/ (2004)

[5] Images and video of the VSD in action can be found at:
             www.k2.t.u-tokyo.ac.jp/perception/VolumeSlicingDisplay/

[6] Jordà S., Kaltenbrunner M., Geiger G., Bencina R.: *The reacTable*, in ACM SIGGRAPH Sketches (2005).

[7] Weiss M., Wagner J., Jansen Y., Jennings R., Khoshabeh R., Hollan J.D., and Borchers J.: *SLAP Widgets: Bridging the Gap Between Virtual and Physical Controls on Tabletops*. In CHI '09: Proc. of the 27 conference on Human factors in computing systems (ACM SIGCHI), New York, NY, USA, (2009)

[8] Cassinelli A. and Ishikawa M.: *Khronos Projector*. ACM SIGGRAPH Emerging Technologies, LA (2005).

[9] Watanabe Y., Cassinelli A., Komuro T. and Ishikawa M.: *The Deformable Workspace: a Membrane between Real and Virtual Space*, IEEE Tabletops and Interactive Surfaces, Oct. 1-3, Amsterdam, pp:145 - 152, (2008)

[10] Watanabe Y., Komuro T., and Ishikawa M.: *955-Fps Real-Time Shape Measurement of a Moving/Deforming Object Using High-Speed Vision for Numerous-Point Analysis*, IEEE International Conference on Robotics and Automation (ICRA'07), Roma, 2007.4.11, pp: 3192-3197. (2007)

[11] Nagakura T. : *Deskrama*, ACM SIGGRAPH, Emerging Technologies (2006)

[12] Saeki Y. and Hirota K.: *Cross-Section Projection Display using Movable Screen*, 11th Japan Virtual Reality Conference, Sept. (2006)

[13] Gupta Sh. and Jaynes Ch.: *Active Pursuit Tracking in a Projector-Camera System with Application to Augmented Reality*, Proc. of the 2005 IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'05), Vol. 3, pp: 111, (2005)

[14] Han, J. Y.: *Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection*. In Proc. of the 18th Annual ACM Symp. on User Interface Software and Technology (UIST), (2005)

[15] Sato T., Mamiya H., Koike H., Fukuchi K.: *PhotoelasticTouch: Transparent Rubbery Tangible Interface using an LCD and Photoelasticity*, in ACM Science And Technology, pp: 43-50 (2009)

[16] Reynolds C., Cassinelli A., and Ishikawa M.: *Meta-perception: reflexes and bodies as part of the interface*, Conference on Human Factors in Computing Systems (CHI 2008), Florence, Italy, April (2008)