Stereoscopic Volume Rendering

Roger J. Hubbold[1], David J. Hancock[1], and Christopher J. Moore[2]

[1] Advanced Interfaces Group,
Department of Computer Science,
University of Manchester,
Manchester M13 9PL
[2] Department of Medical Physics,
Christie Hospital,
Manchester M20 9BX

Abstract. In this paper we describe the extension of a parallel, distributed, direct volume renderer for use with a novel auto-stereoscopic display. We begin by briefly describing the target application of our research, radiation therapy planning, why we believe that stereoscopic viewing may be helpful for this, and the design of our DVR system. We then report on some of the problems we have encountered, and the results we have obtained in experiments. These demonstrate that stereoscopic viewing is beneficial for perceiving depth in transparent DVR images. We illustrate the application of the system to the visualisation of prostate cancer treatment plans. Finally, we describe the use of head-tracking to implement 3D stereo look-around.

Keywords: auto-stereoscopic display, volume rendering, visualisation, medical imaging, radiotherapy.

1 Introduction

For several years, direct volume rendering (DVR) by ray casting has been investigated as a tool for visualisation in medicine [11]. Here, we report on experiments relating to the application of DVR to radiation therapy (radiotherapy) planning. Radiotherapy planning is a safety-critical task – any mistakes could have serious consequences for the patient. It is also difficult, because it is a complex 3D problem involving multi-valued volume data. Our goal is to find ways to visualise treatment plans using 3D displays. These 3D views must unambiguously convey all of the information needed to ensure that a plan is viable and safe.

One approach to disambiguating depth in DVR images is to use a stereoscopic display. We are fortunate in having a collaboration with the Imaging Technology Group at Sharp Laboratories of Europe. They have developed a number of stereoscopic display technologies, including auto-stereo devices in which no special glasses, filters or shutters need be worn. The device which we have been testing presents a full-resolution image to each eye simultaneously [4] – an important consideration when high-quality images are needed for medical visualisation. Using this device, we have conducted experiments to evaluate whether stereoscopic visualisations offer advantages compared with monoscopic ones.

Direct volume rendering has been proposed previously for radiotherapy planning [12, 5]. However, we believe that our work is the first attempt to apply the method to conformal radiotherapy, and to investigate the use of auto-stereoscopic displays for this purpose.

The remainder of this paper is organised as follows. We begin with a brief overview of the target problem, radiation therapy planning. Next, we outline our parallel, distributed DVR system, and explain the techniques we use to obtain interactive rendering speeds. We then consider the problems of extending these ideas to work with an auto-stereoscopic display. We summarise results of experiments to explore whether stereoscopic display is helpful in interpreting DVR images. We show that although DVR is helpful, some of the techniques which can be used for monoscopic images do not work satisfactorily in stereo, because of aliasing problems. We illustrate the application of our method with some visualisations of a treatment plan. Finally, we describe the use of head tracking to control viewing on the stereo display.

2 DVR visualisations of radiation therapy plans

The purpose of radiation therapy planning is to devise a pattern of radiation beams to irradiate a cancer within a patient's body. In our study the cancers are in the prostate, which is surrounded by sensitive organs such as the bladder and rectum. The directions of the different beams can be adjusted relative to the patient's body in order to target the tumour. The cross-sectional profiles of the beams can also be varied using a series of programmable tungsten leaves – a process known as *conformal radiotherapy* [9, 2]. The goal is to plan multiple intersecting beams, which enclose the tumour. In this central region, where the beams intersect, a high radiation dose is delivered, but this must be achieved without damaging other organs. Current clinical practice is to use a series of 2D slices for visualising plans, but this has the serious drawback that a 3D model of the plan is only formed in the consultant's head, rather than visualised directly.

In order to determine whether a given plan is satisfactory, several different data fields must be combined: CT or MRI scans of the patient, the tumour volume, the geometry and magnitude of the radiation dose field (computed using a procedure which takes account of propagation and scattering), and particular organs which may be sensitive to radiation damage, such as the rectum and bladder. In our experiments the tumour and organs are segmented during a preprocessing step, so our task is to produce a coherent view in which different features are clearly visible.

The dose field has a complex 3D geometry, which makes it difficult to visualise. The field strength varies continuously through the volume, adding to the difficulty. The dose field is only really useful when viewed in combination with other features, such as the tumour volume. For example, we must be able to visualise the radiation dose inside a tumour to ensure that an adequately high level is delivered to the target region. Too low a dose results in a *cold spot*. Conversely, if the dose is too high inside critical regions, such as the rectum, then irreparable damage can result; such occurrences are termed *hot spots*. Because these hot and cold spots occur *inside* organs, some means is needed for seeing through the data. DVR offers a number of potential advantages for this, as volumes can be made partially transparent. However, our experience with monoscopic, transparent DVR images shows that depth relationships – important in our application – can be very difficult to discern. This provides the motivation for exploring whether stereoscopic renderings are helpful in interpreting the data correctly.

3 The KVR direct volume renderer

We use a purpose-designed, parallel, distributed volume renderer called KVR. We have previously reported a number of design issues and experimental results relating to this. Here, we restrict our discussion to features which impact – in some cases indirectly – on the effectiveness of stereoscopic display.

Direct volume rendering is notoriously time consuming, but when implemented carefully is capable of generating high-quality images with minimal artifacts. To address the computation speed we use a parallel implementation running on a remote supercomputer [6]. In order to access this from low-cost workstations and PCs we developed a distributed rendering protocol [7]. Its design is integrated with the parallelisation strategy of the renderer, so that images are compressed, transmitted, and displayed using a pipeline design to hide network latencies. Furthermore, the implementation includes a progressive refinement strategy so that initial low-resolution images are delivered and displayed very rapidly [6]. An initial image is displayed in a fraction of a second, and a 640×480 full-resolution image (one ray per pixel) takes five seconds to generate on 24 processors on an SGI Origin-2000[1].

This integration of parallel rendering, progressive refinement and pipelined design works very effectively for monoscopic images. Closed-loop times for interaction of less than one second make it feasible to perform direct manipulation of the visualisation parameters (viewpoint, lighting, classification, cutting planes and sub-volumes). Extending the approach to work with stereoscopic displays appears, at first sight, to be straightforward. However, in practice this has not proved to be the case.

4 Extending DVR for stereoscopic display

The most obvious change required for stereoscopic display is the generation of two views, one for each eye. Our Sharp auto-stereo display (ASD) is driven

---

[1] This time was actually measured for a *pair* of stereo images; the time for a single monoscopic image is actually half of this.

from an SGI Crimson/VGXT via a videosplitter option. Each view is composed in a separate window on the SGI screen and these are fed as standard NTSC signals into the display hardware. Within the display, each view is presented on a dedicated LCD panel, and the two images are combined using a patented system which ensures that each eye sees only the correct image. The display's design ensures that a full 640 × 480 image is presented continuously to each eye, without flicker and with minimal cross-talk. With some types of auto-stereo display, multiple views are obtained by trading horizontal resolution against the number of views. Not only does this lead to lower effective horizontal resolution, but it introduces a correspondingly lower depth resolution. With the Sharp ASD, because each image is shown at full resolution, fine detail is visible and excellent depth resolution is achieved.

Such high quality is important for applications such as medical imaging, but equally means that the renderer must generate images of comparable fidelity; any unwanted artifacts will be uncompromisingly revealed by the display. Although many systems have been used for stereoscopic display, image quality has often been compromised by lower resolution, or by cross-talk between the left- and right-eye views. The human visual/perceptual system is remarkably adept at attempting to compensate for such deficiencies, but prolonged use of inadequate displays leads to visual fatigue. In the literature on stereoscopic displays for visualisation these issues have received scant attention.

In our system the left and right views are generated using the standard technique of parallel, perspective projections [8]. The offsets used for the eye positions and 'look points' are calculated to match the view plane distance of the ASD. This yields a correct perception of depth for the presented images. A novel feature of our system is that the user interface and renderer are designed to support an arbitrary number of stereo visualisations of a given data set. Each stereo pair can be classified and coloured to show different features within the volume data. These multiple visualisations can then be composed into a single 3D normalised projection coordinate space which is projected stereoscopically to yield a correct 3D view. Figure 5 shows an example.

It is important to realise that this is *not* the same as generating several stereoscopic pairs of images and then compositing these in 2D – this latter approach leads to anomalous views, resulting in distortions during viewing (see Section 8). Because we are trying to make important decisions about relative positions of features, accurate display is essential. We do not know of any other stereoscopic DVR systems which can present multiple views in this way. Being able to view several complementary, stereoscopic visualisations simultaneously in this way has proved to be a very effective technique.

5 Experiments to assess stereoscopic DVR

In an earlier paper [10] we described a set of experiments to quantify the value of stereoscopic display for a synthetically generated set of test volumes. In this section we briefly summarise the most important results, and then in the next section we consider how these have affected our use of stereoscopic display for the radiotherapy application.

5.1 Depth tests

A first experiment was designed to test whether users could differentiate very small depth differences in a series of images. The test images were generated from a *voxelised* dataset containing three small spheres enclosed in a larger transparent sphere. Three different types of image were generated by varying the classification of the data volume. They were selected to be representative of the principal effects obtainable with DVR – combinations of surface shading and transparency. Figures 4 (a) – (c) show examples of different classifications. In each case, the small spheres were made totally opaque, as we wished to assess the effect of visualising these inside another volume.

In the first classification ("Inner-only", Figure 4 (a)), the large outer sphere was made totally transparent, so that only the small spheres were visible. This provided a base case for reference, although it made judging depth very difficult. This was because there were no depth cues other than those resulting from very small variations in size due to perspective. In the next version ("Transparent", Figure 4 (b)), the outer sphere was made partially transparent, but surface effects were suppressed. The effect of this is to produce an attenuation of brightness of the inner spheres. In the third classification ("Shell", Figure 4 (c)), the surface of the outer sphere is enhanced using gradient magnitude. This produces an image in which the outer surface has a shell-like appearance.

For each classification, images were generated in which the central, small sphere was stationary, but the two outer, small spheres were moved either forwards or backwards by a small amount. If the left sphere was moved forwards, the right one was moved backwards by a corresponding amount, and vice-versa. The amount of movement was deliberately made very small, so that differences between images were very subtle. For each combination of classification and sphere positions, both monoscopic and stereoscopic views were generated.

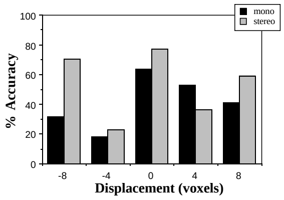


Fig. 1. Inner-only

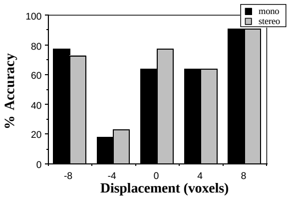


Fig. 2. Transparent

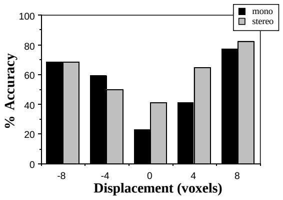


Fig. 3. Shell

The resulting thirty images were shown in a random order to 22 subjects, and they were asked to choose one of three possible answers: left sphere is closer, right sphere is closer, or all three spheres have equal depth. Subjects were not told whether a given image was monoscopic or stereoscopic.

The results, shown in Figures 1 to 3, demonstrate a clear advantage in favour of stereoscopic display. The best results were obtained in the third category, in which the outer sphere's surface was accentuated. This is perhaps not surprising, as stereo helps to accentuate depth perception when clearly delineated features are present. Most surprising was the fact that stereo was most helpful in disambiguating situations where other effects were very small. For example, in the case where all three spheres were at the same depth, the results were better with stereo.

5.2 Aliasing problems

Two types of aliasing affect stereoscopic images. These are non-correspondence aliasing and depth aliasing [13]. Non-correspondence aliasing results when features which should be visible to both eyes appear in only one image of the stereo pair. In DVR this is caused by a ray completely missing the feature in one of the views. Depth aliasing results from the discrete number of apparent depths which can be displayed for a given horizontal pixel pitch.

In common with other systems, in our earlier work we relied on progressive refinement to deliver interactive image generation times for monoscopic images. Unfortunately, under-sampling of the data to obtain initial low-resolution images does not work in stereo. This is because ray casting is a point sampling technique. If only a small number of rays are cast, then discrepancies arise between the left and right images. Features may, and often do, appear in one image and not the other, and this can cause considerable discomfort for the viewer. Aliasing is a serious problem for good quality stereoscopic display, but is often ignored. During brief periods of use the problem may go unnoticed and is often not even considered. But aliasing makes correct depth perception difficult (conflicting information is presented to the two eyes), and prolonged use results in visual fatigue.

We have found little published data dealing with this topic for volume rendering, although Adelson and Hansen [1] referred briefly to the problem. Their solution was to re-project a feature intersected by a ray in one view into the other view. They showed how this could yield accelerated rendering by reducing the work needed to obtain two views. However, it is clear that the method described in their paper will result in depth aliasing.

Sampling theory tells us that to avoid aliasing we must sample data at a rate above the Nyquist frequency. In DVR, this frequency must be chosen so that the spacing between rays, and the distance between re-sampling points along each ray, is less than the distance between data points (voxels). One possible way to address this problem is to pre-filter the volume to produce low resolution datasets, and to use these to generate initial images during progressive refinement.

To test this assertion, we computed filtered, lower-resolution versions of the full-resolution volume. We then generated images corresponding to three different levels of progressive refinement. Three images were generated with the full resolution volumes, and a further three using the lower resolution volumes which best matched the image resolutions. Subjects were shown the six different images and asked to assign two scores to each image, one for their perception of depth, and the other for viewing 'comfort'. Scores for both measures were assigned on a scale from 0 (very bad) to 4 (very good).

image resolution	volume resolution		% benefit
	full	matching	
low	1.43 ± 0.87	2.33 ± 0.96	63%
medium	2.31 ± 0.90	3.26 ± 0.78	41%
high	2.62 ± 1.24	3.52 ± 0.62	34%

Table 1. Scores for perceived depth

image resolution	volume resolution	
	full	matching
low	2.04 ± 1.17	2.09 ± 1.06
medium	2.43 ± 0.86	3.18 ± 0.65
high	3.61 ± 0.60	3.57 ± 0.78

Table 2. Scores for viewing comfort

Tables 1 and 2 show the scores (along with the standard deviations) allocated for comfort and depth respectively to each of the six images. The results confirm that depth perception improves with the sub-sampled volumes which correctly match the image resolution. The improvements are proportionately greater at lower resolutions. However, the scores assigned for 'comfort' showed only a minor improvement at the low and medium resolutions. From subjects' comments, it was clear that there was some confusion between comfort and image resolution. They strongly preferred looking at higher resolution images, but there was no clear favouring of the correctly sub-sampled volumes.

6 Implications for stereo DVR

These results have two important consequences for interactive DVR. First, stereoscopic viewing yields a demonstrable improvement in depth perception for volume rendered images containing nested transparent objects. Because DVR images can take some time to generate, this benefit may be quite important in our medical imaging application.

Second, although the effects of aliasing can be mitigated by using multi-resolution volumes, depth aliasing cannot be removed because it results from pixel replication. Users are not happy looking at low resolution images in stereo because of correspondence problems. Given that DVR uses ray casting, which is a point sampling technique, there does not appear to be any easy solution to this problem. Using filters, or tri-linear interpolation, helps, but does not remove the problem. The practical consequence of this is that during progressive refinement, monoscopic images should be used. This is easily arranged by displaying a single image for both eyes, even though both images are computed. Once full resolution is reached, the correct second view can be switched in, giving the full benefits of extra depth information. Although not ideal, this is a practical compromise, which gives the benefits of rapid closed-loop interaction, and stereoscopic viewing of complete images.

7 Visualisation of radiotherapy plans

Figures 5 and 6 show examples of the volume renderer applied to a particular set of patient data. These images need to be viewed in colour, and may be accessed at `http://www.cs.man.ac.uk/aig/ASD/`.

In Figure 5 we show four different visualisations composited into a single view. For clarity, we show this monoscopically, although our system generates a correct stereo rendering of this picture. The upper-left picture shows the dose field, coloured red, resulting from four treatment beams. Organs are coloured green, yielding a yellow colour where the two intersect. A bright yellow indicates a high dose. The bottom left image shows the corresponding hot (red) and cold (blue) spots. In the bottom-right image, we have added the organs (green). Finally, in the top-right image, we have added the pelvic girdle, obtained from the original CT scans. The option to produce a number of visualisations showing different classifications of the data, correctly composited for stereo viewing, greatly aids our ability to relate the different features to each other.

Figure 6 contains an example of a stereo rendering. The two side-by-side images are arranged for cross-eyed viewing (with the right-eye view on the left). If you are able to fuse these images it should be evident that the stereo view gives a much clearer interpretation of the depth than looking at either image individually. We should emphasise that on the Sharp auto-stereo display, these images can be viewed very comfortably. In this picture, we see several nested features. The tumour is coloured green. This is surrounded by the 'planning tumour volume' (PTV), coloured light grey. The PTV is the region which the clinicians have decided should receive a high dose, to allow for the possibility of movement of the organs during a course of treatment. This is therefore the target region. In addition, we see the bladder (the somewhat rectangularly shaped blob towards the bottom of the image) and a section of the rectum (the tube-shaped feature at the top). The white lines represent the centres of the radiation beams. As previously, hot and cold spots are shown.

8 Use of head tracking to control viewpoint

An alternative, complementary technique for depth perception is the use of rotation. This can take a number of forms: direct user control of motion, automatic 'rocking' of the image between two angular positions, and – where suitable hardware is available – head tracking.

The Sharp ASD has a head tracking facility. In normal use, this is coupled to a unique image steering mechanism which ensures that the left- and right-eye images remain correctly aligned with the user's eyes. An important feature of this is that a full resolution image is continuously available to each eye, unlike other ASD systems in which horizontal resolution is reduced to obtain stereoscopic views and/or look-around capability. The Sharp system also provides the option to feed the user's head position into the host computer. This means that it can be employed to change the viewing parameters. With a little calibration, it is possible to provide a ‘look-around’ capability, in which the 3D objects appear stationary in front of the viewer as the head is moved from side to side. Without this, a lateral movement of the head produces a strange apparent distortion of the displayed view – this is the effect referred to previously in Section 4.

To assess this technique, we have generated sequences of full resolution, stereoscopic images (such as those shown in Figure 5) to cover a range of viewing positions in small angular increments. In look-around mode, the view displayed is chosen to correspond to the user’s head position, and this is updated dynamically. To provide a comparison, a rocking mode has also been implemented, in which the view oscillates at a user-controllable rate. With both methods of control, the display can be switched between stereoscopic and monoscopic viewing.

Preliminary results from this once again provide convincing evidence of the value of stereoscopic display. Although the rocking motion gives a clear sense of depth, continuous movement makes it impossible to study the data carefully. For this, rocking must be temporarily suspended. In look-around mode, it is possible to freeze the view simply by keeping the head reasonably still. Changing view is intuitively easy and feels quite natural. Distortions which would otherwise result from head motion are eliminated. If stereoscopic viewing is enabled, then depth perception persists when motion is stopped. However, if monoscopic viewing is used then the image immediately becomes ‘flat’ and depth interpretation is difficult.

The combination of head tracking and stereoscopic viewing is very attractive. However, it is not cost-free. As noted previously, monoscopic presentation should be used during progressive refinement. Because a full resolution image takes several seconds to generate, the look-around mode is only feasible with pre-computed images. Also, there is an associated cost for the necessary hardware in the display itself. We plan to conduct a more complete set of tests with a representatively large sample of users to better quantify the benefits of this technique.

9 Concluding remarks

We have demonstrated that there are important benefits in using stereoscopic presentation for DVR. Our experiments with depth perception show a clear advantage for stereo, even in cases where depth relationships are very subtle, and where transparency makes judgement difficult. However, several problems arose, of which the most important are non-correspondence and depth aliasing. These arise particularly with progressive image refinement, during which we revert (temporarily) to monoscopic viewing.

The problems of aliasing should not be underestimated with stereoscopic displays. They are frequently ignored because the human visual/perceptual system attempts to make sense of anomalous images, but they are present nonetheless, and for good quality images anti-aliasing must be used. Without this precaution, visual fatigue – evident in our multi-resolution aliasing experiments – will result, making stereo displays difficult to view for extended periods.

There are, of course, other approaches to volume rendering. Candidates of particular relevance for the kinds of visualisations with which we are concerned here are splatting [14], and 3D texture mapping (α-blending) [3]. The latter, in particular, has the potential to generate images quite rapidly using modern display hardware. It would be instructive to investigate the extent to which these methods suffer from aliasing problems when used with a high-quality stereoscopic display.

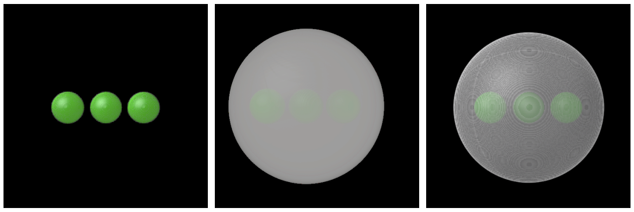
Acknowledgements

This research is funded by the UK's Engineering and Physical Sciences Research Council under grant GR/L02685. The auto-stereoscopic display was developed, and lent to our group, by the Imaging Technology Group of Sharp Laboratories of Europe, Oxford, UK. We are very grateful to Dr David Ezra and his staff, and to Sharp Laboratories, for their support.

References

1. S.J. Adelson and C.D. Hansen. Fast steroscopic images with ray-traced volume rendering. In Arie Kaufman and Wolfganag Krueger, editors, *1994 Symposium on Volume Visualization*, pages 3–9, Washington D.C., October 1994. ACM Press.
2. A.R. Hounsell et al. Dose calculations in multileaf collimator fields. *British Journal of Radiology*, 65:321–326, 1992.
3. Brian Cabral, Nancy Cam, and Jim Foran. Accelerated volume rendering and tomographic reconstruction using texture mapping hardware. In Arie Kaufman and Wolfganag Krueger, editors, *1994 Symposium on Volume Visualization*, pages 91–98, Washington D.C., October 1994. ACM Press.
4. David Ezra, Graham J. Woodgate, Basil A. Omar, Nicolas S. Holliman, Jonathan Harrold, and Larry S. Shapiro. New autostereoscopic display system. In *Proc. SPIE 1995 Conference*, volume 2409, pages 31–40, 1995.
5. D. Gordon, M.A. Peterson, and R.A. Reynolds. Fast polygon scan conversion with medical applications. *IEEE Computer Graphics and Applications*, 14(6):20–27, November 1994.
6. D.J. Hancock and R.J. Hubbold. Distributed parallel volume rendering on shared-memory systems. In Bob Hertzberger and Peter Sloot, editors, *High-performance computing and networking*, number 1225 in Lecture Notes in Computer Science, pages 157–164. Springer-Verlag, April 1997. (Proc. HPCN Europe 1997. ISBN 3-540-62898-3.).
7. D.J. Hancock and R.J. Hubbold. Efficient image synthesis on distributed architectures. In Rae Earnshaw and John Vince, editors, *The Internet in 3D, Information, Images and Interaction*, pages 347–364. Academic Press, 1997. ISBN 0-12-227736-8.
8. Larry F. Hodges and David F. McAllister. Computing stereoscopic views. In David F. McAllister, editor, *Stereo Computer Graphics and Other True 3D Technologies*, chapter 5, pages 71–89. Princeton University Press, 1993.
9. A.R. Hounsell and J.M. Wilkinson. Dose calculations in multileaf collimator fields. In *Proc. XIth International Conference on Use of Computers in Radiation Therapy*, pages 234–235. Manchester, UK, 1994.

10. R.J. Hubbold, D.J. Hancock, and C.J. Moore. Autostereocopic display for radiotherapy planning. In Scott S. Fisher, John O. Merritt, and Mark T. Bolas, editors, *Stereoscopic Displays and Virtual Reality Systems*, volume 3012, pages 16–27. SPIE – The International Society for Optical Engineering, February 1997.
11. Arie E. Kaufman. Introduction to volume visualization. In Arie E. Kaufman, editor, *Volume Visualization*, chapter 1, pages 1–18. IEEE Computer Society Press, 1991.
12. M. Levoy. A hybrid ray tracer for rendering polygon and volume data. *IEEE Computer Graphics and Applications*, 10(2):33–40, March 1990.
13. William F. Reinhart. Gray-scale requirements for anti-aliasing of stereoscopic graphic imagery. In *Proc. SPIE 1992*, volume 1669, pages 90–100, 1992.
14. L. Westover. Footprint evaluation for volume rendering. *ACM Computer Graphics*, 24(4):367–376, August 1990.



(a) (b) (c)

Fig. 4. (a) Inner-only, (b) transparent, (c) surface.

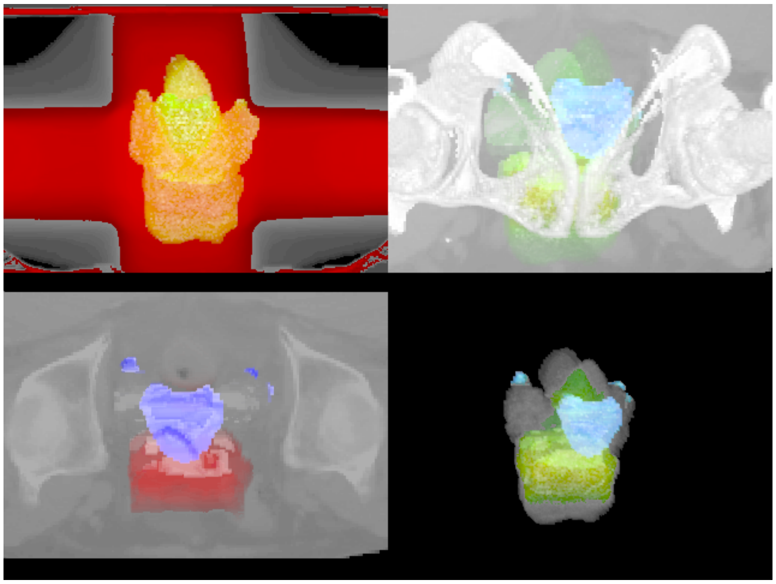


Fig. 5. Four views composited into a single image. The compositing is performed in 3D, although this is not apparent in this monoscopic view.

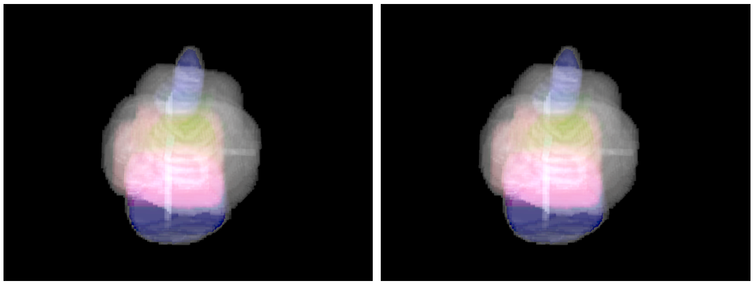


Fig. 6. Stereoscopic pair showing nested features, arranged for cross-eyed viewing.