

Segmentation of the Visible Human for high-quality volume-based visualization

Thomas Schiemann*, Ulf Tiede and Karl Heinz Höhne

Institute of Mathematics and Computer Science in Medicine (IMDM), University Hospital Hamburg-Eppendorf, Germany

Abstract

This article describes a combination of interactive classification and super-sampling visualization algorithms that greatly enhances the realism of 3-D reconstructions of the Visible Human data sets. Objects are classified on the basis of ellipsoidal regions in RGB space. The ellipsoids are used for super-sampling in the visualization process.

Keywords: multi-parametric classification, segmentation, visible human, volume visualization

Received November 21, 1996; revised February 3, 1997; accepted February 10, 1997

1. INTRODUCTION

Segmentation is the core problem for many applications in medical imaging. Although it has been worked on by numerous authors for many years, general solutions are not at hand. If we confine ourselves to specific classes of images and to special objectives, however, successful applications can be developed.

In the case of the Visible Human data sets (Spitzer *et al.*, 1996) we have three parametric photographic images, which we want to display in three dimensions using techniques of volume visualization. In the special case of volume visualization from cross-sectional images, it is well known that the best visualization results are obtained if the segmentation procedure offers results that ensure smooth coherence of the surface voxels at positions of high-intensity gradients. For scalar images (like CT or MRI) this condition is fulfilled if threshold-based methods are used (Tiede *et al.*, 1990), because in this case the visualization procedure (ray-casting) can locate surface positions and orientations with sub-voxel resolution resulting in images that show even the smallest structures in a natural way (Ney *et al.*, 1991; Pommert *et al.*, 1992). Other segmentation techniques, which are tailored for other purposes, neglect this coherence and hence often lead to visualization artefacts when small details are to be rendered.

For the colour images of the Visible Human Project this paper therefore presents an approach which generalizes the

threshold ranges in scalar space to volumes in RGB space, which can again build a basis for super-sampled ray-casting. This paper demonstrates that photographic images such as those of the Visible Human can be segmented in a way that delivers realistic images of unprecedented quality.

2. METHOD

2.1. Image preprocessing

For our experiments, we worked with the Visible Human male. The data are transverse cross-sectional photographic images of a frozen male cadaver with a resolution of 0.33 mm and slice distance of 1 mm. The images are supplemented by two sequences of high-resolution CT images from the fresh and frozen cadaver with resolutions between ~0.5–1.0 mm and slice distances of 1 mm for the frozen and between 1–5 mm for the fresh cadaver. The fresh cadaver has also been scanned by MRI, resulting in separated stacks of (mainly coronal) images, which have been obtained with three acquisition protocols (T1, T2 and PD). The MR images have a lower resolution of 1–2 mm and slice distances of up to 5 mm.

Before acquisition of the photographic images, the cadaver was cut up into different portions. This procedure leads to some shift artefacts in the resulting volume, which are visible as horizontal lines on 3-D renderings.

Next to algorithmic problems, which will be addressed later, one of the main difficulties in working with the Visible Human data is their huge size, especially that of the anatomical data. Handling even small parts of the body in full

*Corresponding author
(e-mail: schiemann@uke.uni-hamburg.de)

resolution requires very large disk storage and computer main memory (e.g. 440 Mbyte for the raw data of the head alone or 1100 Mbyte for the upper abdomen). Thus we have reduced the resolution of the anatomical images to 1 mm^3 by averaging 3×3 pixels on every slice. While this procedure has the advantage that the resulting data (49 Mbyte for the head and 121 Mbyte for the abdomen) can be processed by high-end workstations, it also introduces new segmentation problems, because partial-volume effects are created or much enhanced by the averaging procedure.

A 3-D scatter-plot of the colour distribution of the anatomical images shows that the components are highly correlated. We therefore considered the possibility of reducing the RGB data into two-channel data by principal-component analysis. But we did not ultimately choose this procedure, since the original colours of the anatomical images are still needed for later interactive segmentation and thus cannot be replaced by the artificial two-channel data.

The CT and MRI data are roughly registered with the anatomical data. We improved this registration using an interactive landmark-based tool (Schiemann *et al.*, 1994). The main mapping parameters which had to be refined were translation and scaling within the slices. The other parameters (scaling and translation perpendicular to the slices and rotation) were already correct. The registration error is reduced to approximately 1 voxel for the frozen CT data set, but for the images obtained from the fresh cadaver, locally larger registration errors occur due to morphological changes caused by freezing. The registration procedure allows us to create combined renderings of different modalities in a single view with object parameters split into the different modalities.

2.2. Volume-based visualization

For the task of 3-D visualization of volume data, there are in general three different approaches (Kaufman, 1991):

- Rendering of surfaces, which are represented as polygons. In a preprocessing step such as the marching cubes algorithm (Lorenson and Cline, 1987), polygons are extracted from the image volume. These surfaces can then be visualized with standard graphics packages.
- Volume rendering produces 3-D images by summing up all voxel values along the path of each ray (Drebin *et al.*, 1988; Levoy, 1988). Different weighting factors, which usually depend on intensity or gradient magnitude, can be assigned to each voxel in order to enhance the appearance of certain objects.
- Volume-based rendering of surfaces uses object membership attributes for each voxel during ray-casting in order to decide which voxels are to be rendered and how the rendering should be performed (e.g. opaque

or transparent, shading parameters, etc.) (Tiede *et al.*, 1990). The membership attributes are obtained from segmentation, which has to be done in advance.

While polygonal surface rendering has the advantage of being fast, if the number of polygons is kept low, considerable detail is usually lost during the preprocessing steps necessary for generating the surfaces. Accurate mapping of surface texture and realistic cutting are difficult to implement. On the other hand, pure volume rendering is usually very slow on standard hardware and is generally not capable of creating realistic images. We have thus concentrated on volume-based surface rendering, which can compute images with a high degree of detail and realism. Although computing times for these volume-based methods are still longer than for triangular rendering, this disadvantage will be overcome by future hardware.

In our previous work, we developed a framework for volume-based anatomy atlases (Schubert *et al.*, 1993; Höhne *et al.*, 1995). One of the key procedures of this system is a high-quality volume rendering module, which uses volume-based ray-casting methods. There are two conditions that generally have to be fulfilled for the computation of high-quality images:

- (i) The surface shading appears most realistic when the surface normals are determined by voxel intensity gradients (Höhne and Bernstein, 1986; Tiede *et al.*, 1990). Thus the segmentation procedure, which precedes visualization, has to define the object borders 'reasonably well' at the location of the intensity gradients.
- (ii) The clusters defining an object in RGB space must be expressed in a simple analytical form in order to allow a fast refined classification of super-sampled data at ray-casting time.

For scalar images (CT and MRI) these prerequisites are fulfilled by determination of an intensity range by a pair of lower and upper threshold values. First the threshold range yields object boundaries at the location of high intensity gradients, then the threshold can be interpolated in sub-voxel resolution. The strong influence of the threshold specification on the rendering results is shown in Figure 1—without threshold specification surface locations can be computed in voxel resolution only, which leads to very blocky images. Using the threshold for super-sampling produces smooth surfaces, because the surface location is computed with very high accuracy.

For volume-based rendering of the Visible Human colour images, the key problem for computation of high-quality images is a generalization of scalar thresholds to a description

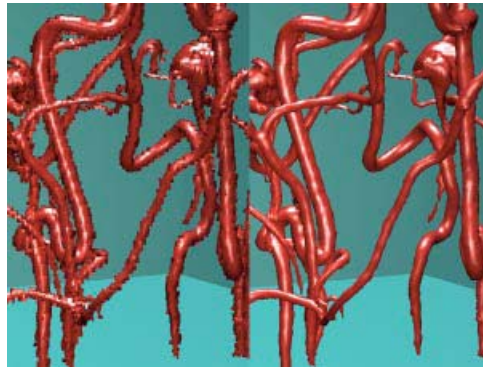


Figure 1. Influence of the threshold specification on the rendering result. Both images show a volume-based rendering of the same voxel set. Left, disregarding the threshold leads to blocky surfaces, because even super-sampling only hits the voxel boundaries. Right, regarding the threshold during super-sampling produces a smooth and realistic surface appearance.

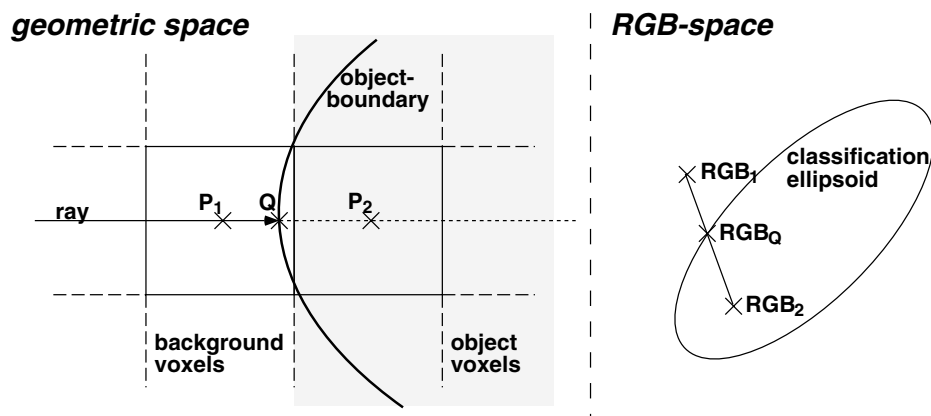


Figure 2. Using the ellipsoidal description of an object for sub-voxel computation of the surface location: voxel P_2 with colour RGB_2 carries the object label, voxel P_1 with colour RGB_1 is outside the object to be visualized. The object surface is assumed to be at the point Q , which divides the line between P_1 and P_2 in the same ratio as RGB_Q divides RGB_1 and RGB_2 .

in RGB space. Pure enumeration of the RGB-tuples is not practical. The cluster should have a simple formal description for efficient computation during ray-tracing. Attempts with a box-shaped classification in RGB space gave unsatisfactory results. Since objects in RGB space look rather ellipsoidal (see Figure 3), we have decided to describe the classified regions by ellipsoids in RGB space. For generalization of the scalar procedure for sub-voxel surface location it is thus sufficient to find the point where a straight line between two RGB-tuples intersects the ellipsoid (Figure 2). The required distance computations are done in Euclidean space with equal weights for all RGB components.

In addition to the threshold generalization we have described, the visualization system has also been adapted in the following way to handle and take advantage of the Visible Human images:

- Instead of assigning an artificial colour, texture mapping is performed by using the RGB-tuple at the surface location for the Phong illumination model.
- The surface normal is calculated as the normalized sum of the intensity gradients of each RGB component.
- Rendering parameters can be taken from different data sets (anatomical, CT or MRI) depending on the object (e.g. the bone surface location and normal can be obtained from CT but the colour comes from the RGB data). This option requires well-registered data volumes, since misregistration leads to considerable artefacts.

2.3. Segmentation on the basis of RGB ellipsoids

Since the segmentation has to be done in great detail, automatic segmentation is not possible and we aim instead at a semi-automatic procedure. The segmentation is adapted from

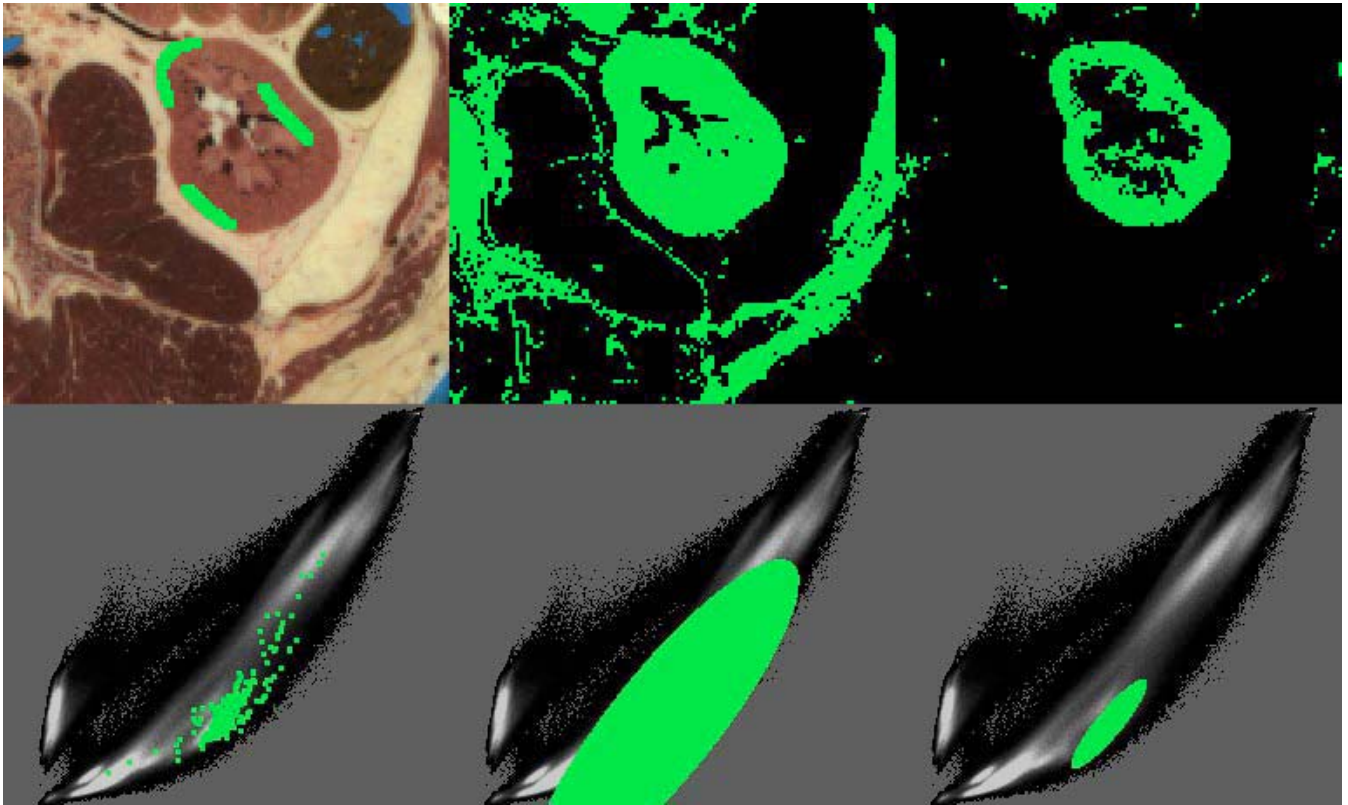


Figure 3. Interactive classification in RGB space for the left kidney. The intermediate results are shown in the image space (upper row) and on 2-D projections along the blue axis of the 3-D scatter-plot (lower row). The user first outlines some typical areas of the kidney (left column). A preliminary result is obtained when all marked triples are used (middle). By restriction to the ‘substantial’ triples, the final sharp classification is computed (right).

an interactive method, which has proven to be successful for many purposes with scalar images like CT or MRI (Höhne and Hanson, 1992; Schiemann *et al.*, 1992). The procedure is based on thresholding followed by binary mathematical morphology and connected component labelling. While the latter two steps remain unchanged for colour images, the intensity range describing an object is replaced by an ellipsoid in RGB space. Determination of the ellipsoid is done by the following steps (Figure 3):

- (i) The user outlines a (usually small) typical region of the object to be segmented. This results in a set of colour triples $(\text{RGB})_i$.
- (ii) In most cases, the sample set $(\text{RGB})_i$ will contain ‘incorrect’ triples, which belong to other structures, but have also been selected during outlining. In order to get rid of these triples, the median RGB triple μ_{RGB} and the median μ_d of all Euclidean distances $d[\mu_{\text{RGB}} - (\text{RGB})_i]$ are computed. For determination of the ellipsoid, only

those triples closer to μ_{RGB} than an interactively specified multiple f of the median μ_d are used.

- (iii) The ellipsoid’s centre is taken from the median of the restricted sample set, the axes directions are taken as the principal axes computed by covariance analysis and the extent of the ellipsoid is determined such that all restricted triples are inside the ellipsoid.

During a practical session, the first goal is to define an ellipsoid that covers the object to be segmented. Due to the huge colour feature space, several outlining steps will be necessary and ‘incorrect’ triples will be included in the ellipsoid (Figure 3, left). In this stage of the procedure, the multiple f is set to an upper limit, so that it has no further influence (Figure 3, middle). Then, the second goal is to refine the ellipsoid by reducing f (Figure 3, right). In our experiments, we ascertained that values between 2 and 4 are a good choice for f .

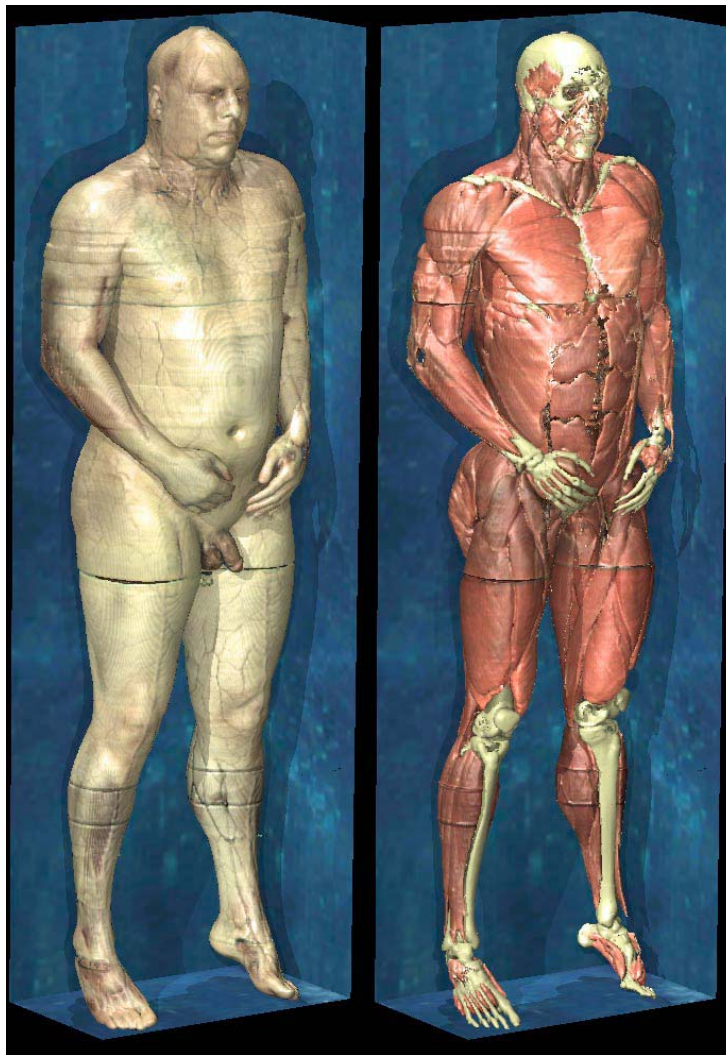


Figure 4. Views of the Visible Human male: skin surface and bone and muscles.

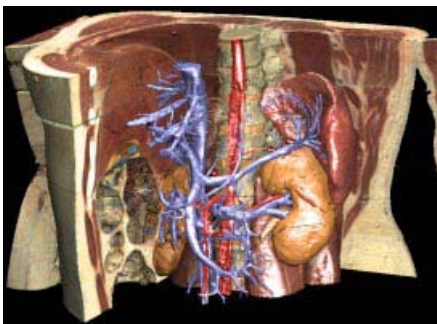


Figure 5. Simulation of gastroscopy showing the typical surface structures of the stomach.

The problem of wrongly marked pixels in the first step is reduced significantly if the full resolution (0.33 mm) images are used. This is due to the fact that the full resolution images show much sharper object boundaries, which can thus be more readily identified in the marking step. In addition, partial-volume effects on the full resolution images are not very strong. Thus, the cluster of marked RGB-tuples is much sharper than that generated from images with reduced resolution.

Computing times for the ellipsoid itself are negligible, but for a fast binary decision of whether a triple is inside the ellipsoid or not, we use a binary volume of size 256^3 , in which all triples inside the ellipsoid are set. Updating of this volume can take several seconds depending on the size of the ellipsoid.

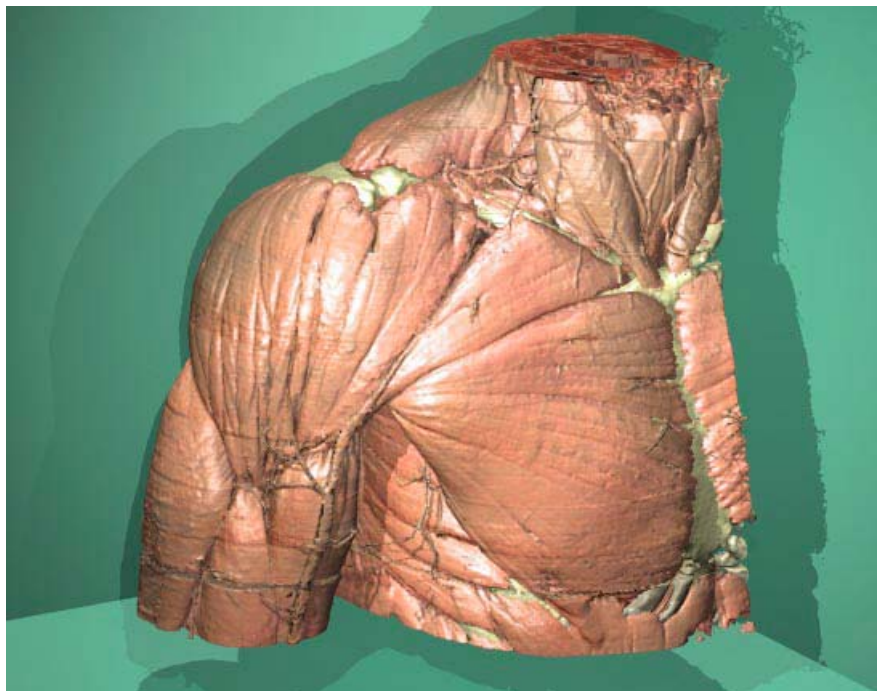


Figure 6. Abdominal anatomy as obtained with the interactive segmentation procedure. Some objects (vasculature, bone) are coloured artificially for better distinction.

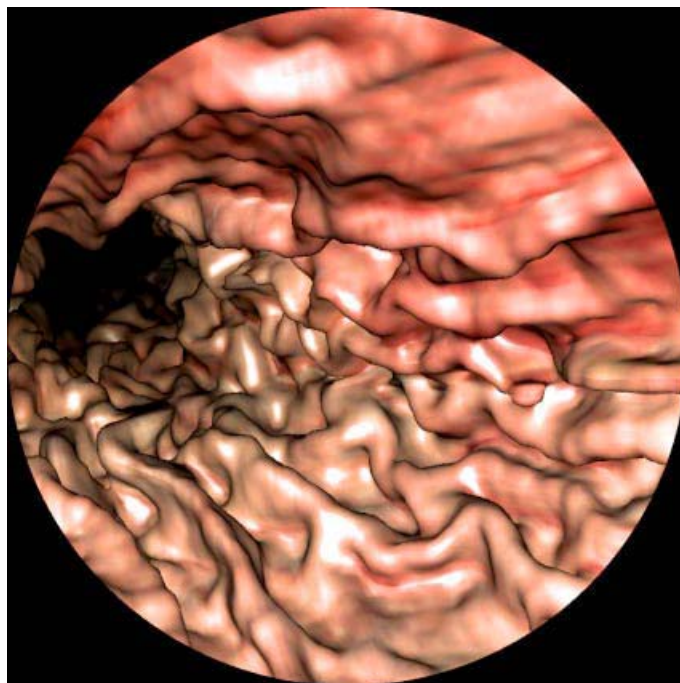


Figure 7. The Visible Human's right shoulder. Photo-realistic quality is achieved exhibiting even small vessels.

The ellipsoidal classification in RGB feature space is generally not sufficient for segmentation of an object without further processing steps in image space, because several structures are classified together. Connected component labelling and binary mathematical morphology are used to achieve further spatial decomposition of the voxels in the geometric space covered by the initial ellipsoid.

Instead of the proposed basic ellipsoidal specification, the segmentation step could also be performed with more sophisticated classification algorithms. However, for the method presented here, it is essential that all computations on the classification result be performed very quickly, because they are required several times for each image point during 3-D rendering. Since more elaborate classification methods usually do not result in easily formalized descriptions (such as ellipsoids), these classification procedures cannot be used for our purpose.

The proposed segmentation procedure is successful for all major constituents of the body, such as the brain or abdominal organs. For a detailed subdivision of these structures, e.g. into gyri of the brain, manual segmentation is still needed, due to a lack of image properties for this purpose.

3. RESULTS

Although the availability of the Visible Human images is rather recent, a number of systems for their exploration does already exist. Unfortunately, most of these systems present the images only as stacks of orthogonal slices, which can be browsed through with different techniques. The number of real 3-D applications (Hong *et al.*, 1996; Seymour, 1996) is still limited.

We have applied our method to several regions of the Visible Human male (Figure 4). The major constituents of the head, the upper abdomen and the right shoulder have been segmented with the interactive segmentation procedure described above. Most structures (e.g. brain, kidneys, muscles and intestine) can be segmented in a few minutes on a DEC 3000 workstation with 256 Mbyte of main memory. Extended organs like the liver with many links to other structures need additional interaction on single-slice sequences.

The main areas of application for techniques using the Visible Human data sets are teaching (especially in anatomy and radiology) and simulation. Figure 6 gives an example of a typical teaching image showing a view of the upper abdomen with kidneys, spleen, muscles, vertebral column and abdominal vasculature. The cut planes show the anatomical colours of the underlying data volume. Texture mapping of the anatomical colours is used for shading of most 3-D rendered structures, while for bone and blood vessels, artificial colours are chosen in order to enhance their appearance.

Figure 7 shows a view of the muscles of the right shoulder, which have been segmented from the anatomical data. The bone structures have been segmented from the frozen CT data. Texture mapping is not performed for the bone, but an artificial colour is again chosen for enhancement. The shoulder region is hit by several shift artefacts, which are visible as horizontal lines.

The possibility of simulation is shown with the example of virtual endoscopy (Figure 5). From a perspective inside the stomach pictures resembling those obtained during real gastroscopy are generated with realistic colours and great anatomical detail.

Further 3-D renderings of the Visible Human's head and abdomen have been published elsewhere (Schiemann *et al.*, 1996; Tiede *et al.*, 1996).

4. CONCLUSION

We have proposed a new combined approach of interactive classification and volume-based visualization of RGB volumes such as those of the Visible Human. The pictures show that the rendering quality obtained with the method described here is superior to that of other approaches published so far.

Our first steps in segmentation for high-resolution volume rendering show that the Visible Human data set allows a new quality of anatomical imaging when used in a state-of-the-art visualization system. Nevertheless, a huge amount of work still has to be done to achieve the detailed segmentation necessary for a complete 3-D anatomical atlas. But unlike clinical imaging, which requires new segmentation for every case, segmentation of the Visible Human has to be done only once. If this work is completed, the results will surely have great impact on many applications of visualization such as education and simulation.

ACKNOWLEDGEMENTS

We are grateful to all the members of our department who have supported this work and special thanks go to Rainer Schubert for anatomical and Uwe Pichlmeier for statistical support. Our thanks also go to Victor Spitzer, University of Colorado School of Medicine and Michael Ackerman, National Library of Medicine, for providing the Visible Human data set.

REFERENCES

- Drebin, R. A., Carpenter, L. and Hanrahan, P. (1988) Volume rendering. *Comput. Graphics*, 22, 65-74.
- Höhne, K. H. and Bernstein, R. (1986) Shading 3D-images from CT using gray level gradients. *IEEE Trans. Med. Imag.*, 5, 45-47.

- Höhne, K. H. and Hanson, W. A. (1992) Interactive 3D-segmentation of MRI and CT volumes using morphological operations. *J. Comp. Assis. Tomogr.*, 16, 285–294.
- Höhne, K. H., Pflessner, B., Pommert, A., Riemer, M., Schiemann, T., Schubert, R. and Tiede, U. (1995) A new representation of knowledge concerning human anatomy and function. *Nature Med.*, 1, 506–511.
- Hong, L., Kaufman, A., Wei, Y.-C., Viswambharan, A., Wax, M. and Liang, Z. (1996) 3d virtual colonoscopy. In *IEEE Biomedical Visualization Symp.*, pp. 26–32.
- Kaufman, A. (ed.) (1991) *Volume Visualization*. IEEE Computer Society Press, Los Alamitos, CA.
- Levoy, M. (1988) Display of surfaces from volume data. *IEEE Comp. Graphics Appl.*, 8, 29–37.
- Lorensen, W. E. and Cline, H. E. (1987) Marching cubes: a high resolution 3D surface construction algorithm. *Comp. Graphics*, 21, 163–169.
- Ney, D., Fishman, E. K., Magid, D., Robinson, D. D. and Kawashima, A. (1991) Three-dimensional volumetric display of CT data: effect of scan parameters upon image quality. *J. Comp. Assis. Tomogr.*, 15, 875–885.
- Pommert, A., Bomans, M. and Höhne, K. H. (1992) Volume visualization in magnetic resonance angiography. *IEEE Comp. Graphics Appl.*, 12, 12–13.
- Schiemann, T., Bomans, M., Tiede, U. and Höhne, K. H. (1992) Interactive 3D-segmentation. In Robb, R. A. (ed.), *Visualization in Biomedical Computing II*, Chapel Hill, NC, *Proc. SPIE*, Vol. 1808, pp. 376–383.
- Schiemann, T., Höhne, K. H., Koch, C., Pommert, A., Riemer, M., Schubert, R. and Tiede, U. (1994) Interpretation of tomographic images using automatic atlas lookup. In Robb, R. A. (ed.), *Visualization in Biomedical Computing 1994*, Rochester, MN, *Proc. SPIE*, Vol. 2359, pp. 457–465.
- Schiemann, T., Nuthmann, J., Tiede, U. and Höhne, K. H. (1996) Generation of 3D anatomical atlases using the Visible Human. In Kilcoyne, R. F., Lear, J. L. and Rowberg, A. H. (eds), *Computer Applications to Assist Radiology, Proc. SCAR '96*, pp. 62–67. Symposia Foundation, Carlsbad, CA.
- Schubert, R., Höhne, K. H., Pommert, A., Riemer, M., Schiemann, T. and Tiede, U. (1993) Spatial knowledge representation for visualization of human anatomy and function. In Barrett, H. H. and Gmitro, A. F. (eds), *Information Processing in Medical Imaging, Proc. IPMI '93, Lecture Notes in Computer Science*, Vol. 687, pp. 168–181. Springer-Verlag, Berlin.
- Seymour, J. (1996) A new 3d database and atlas based on the visible human. In *2nd Southeastern Medical Informatics Conf.*, Gainesville, FL.
- Spitzer, V., Ackerman, M. J., Scherzinger, A. L. and Whitlock, D. (1996) The Visible Human male: a technical report. *J. Am. Med. Inf. Assis.*, 3, 118–130.
- Tiede, U., Höhne, K. H., Bomans, M., Pommert, A., Riemer, M. and Wiebecke, G. (1990) Investigation of medical 3D-rendering algorithms. *IEEE Comput. Graphics Appl.*, 10, 41–53.
- Tiede, U., Schiemann, T. and Höhne, K. H. (1996) Visualizing the Visible Human. *IEEE Comput. Graphics Appl.*, 16, 7–9.