A Semiautomatic Three-Dimensional Segmentation Method for Disarticulation of Bone Structures on Spiral Computed Tomography Images

J. Van Cleynenbreugel, D. Kratka, L. Berben, M.-H. Smet, G. Marchal, and P. Suetens

The use of binary thresholding for segmenting bone structures on spiral computed tomography images is negatively influenced by partial volume effects (PVEs) induced by the image acquisition. PVE leads to mixed voxels, making the binary decision "bone" or "non-bone" a difficult one to take. As a result, two distinct bone structures that are close to each other will often appear to be connected by this method. A typical example consists of "acetabulum/femoral head" pairs in the pelvic region. To separate them, a clinical user must interactively draw a disarticulation line. This procedure is time consuming (often interaction in 50 slices is needed) and leads to unsmooth visualization of the disarticulated areas (by three-dimensional [3D] rendering techniques). We developed a semiautomatic cutting algorithm that leads to smooth disarticulated surfaces and considerably decreases the amount of user interaction. A sheet detection operator is applied to automatically separate bone structures. Detected sheets are used as disarticulation lines. Postprocessing ensures that sheets not relevant for the application do not influence the resulting image. Our approach is encapsulated in an interactive segmentation environment based on thresholding and 3D connected-component labeling. Results are shown for pelvic region, wrist, and foot bone disarticulations.
Copyright © 1995 by W.B. Saunders Company

KEY WORDS: 3D image segmentation, 3D image user interaction, computed tomography bone segmentation.

WHEN SEGMENTING BONE on a three-dimensional (3D) computed tomography (CT) image, one often needs to distinguish between different bone structures. Labeling distinct structures is necessary for proper 3D visualization. Indeed, without this separation in distinct structures, important information could not be visible on a 3D rendition. A typical application is the visualization of acetabulum/femoral head pairs. Here fractures or the specific shape of the connecting area of these two bone structures must often be seen clearly.[1] The mandibulla/maxilla pairs, ankle-bone pairs, and wrist-bone pairs are other examples requiring separate visualization. Often, one cannot properly disarticulate such structures by simple thresholding. Indeed, because of their physical proximity with respect to the scan resolution, these structures appear to be connected on thresholded images. Therefore, after thresholding a large amount of user interaction is still necessary for the final separation. Nevertheless, this manual delineation approach[2,3] is the most common way of separation. This is a time consuming and tedious task because interaction (line drawing) may be needed in tens of slices. The result also depends on the disarticulation skill of the radiologist. Resulting rendered surfaces appear to be unsmooth in the areas in which the separation line was drawn. One alternative to manual delineation is the use of line-detection algorithms (snakes, minimal cost path[4]). Such algorithms still require user interaction in 2D to limit the search area and to correct displacements.

The semiautomatic cutting algorithm discussed in this report significantly reduces the need for user interaction. Furthermore, smoother disarticulated surfaces are obtained. In our approach, we do not focus on one individual disarticulation line. For each 2D slice, we calculate a set of possible disarticulation lines and apply user-guided connected component labeling (CCL) and dilation to determine relevant structures. Limited separation by drawing may still be required. Our algorithm is embedded in an existing bone segmentation

environment and has been tested on several clinical cases with encouraging success.

SEMIAUTOMATIC DISARTICULATION OF DIFFERENT BONE STRUCTURES

Problem Description

We denote by $I(x,y,z)$ an input 3D image volume consisting of a finite set of voxels (x,y,z), where x,y are the 2D coordinates in a CT slice, and z is the coordinate perpendicular to the slices. Segmenting bone in CT images means separating the voxels of I into bone region B and nonbone region NB. Furthermore, if several bone structures are to be distinguished from one another, different bone regions B_i are created according to the formulae

$$I = B \cup NB, \qquad (1)$$

$$B = \bigcup_{i=1}^{n} B_i, \quad \text{and} \quad \bigcap_{i=1}^{n} B_i = \varnothing. \qquad (2)$$

The first formula expresses the thresholding, and the second, the disarticulation operation. Without loss of generality, we assume $n = 2$. If two bone structures have no neighbouring voxels, thresholding followed by CCL solves the disarticulation problem immediately. This condition is seldom valid because in many situations the real distance between different bone structures is comparable to the scan resolution. The borders between B and NB seldom align to the voxel boundaries, resulting in border voxels containing a mixed amount of bone and soft tissue.[5] This so-called partial volume effect (PVE) causes narrow sheets of soft tissues surrounded by bone to appear as bone in the thresholded image. Thresholding I leads to false connections in such cases. Generally, PVE is a function of the scan resolution, but because the scan resolution in the z direction is several times lower than in the x-y slice, PVE also depends on the orientation of a sheet. We have discussed these issues in more detail in another report.[6]

Figure 1 schematically illustrates a typical example. A narrow sheet of soft tissue ST is misclassified as bone, when a threshold th_i is applied. The middle picture represents a possible situation in the y-z plane. The upper picture illustrates the CT number variations along a line in a 2D slice corresponding to this situation. The soft-tissue sheet is not aligned with boundary voxels, but skewed. Because of PVE, the 2D-slice pixels in the area between the two bone structures contain CT numbers larger than the threshold, th_i,. However, increasing th_i often results in unwanted undersegmentation at other locations.

The upper part of Fig 1 illustrates that although all CT numbers are larger than the threshold th_i, there still exists a "valley" (or sheet in 3D) of lower grey values, corresponding to soft tissue (ST). Our algorithm exploits this fact. To disarticulate two different bone structures, we use a sheet operator that assigns the highest response to such gray value "valleys." This is illustrated in the lower part of Fig 1. However, sheet operators also respond to edges. Because the grey-value variations inside bone structures are lower than between different bone structures, we can first eliminate most of the nonrelevant sheet response by sheet
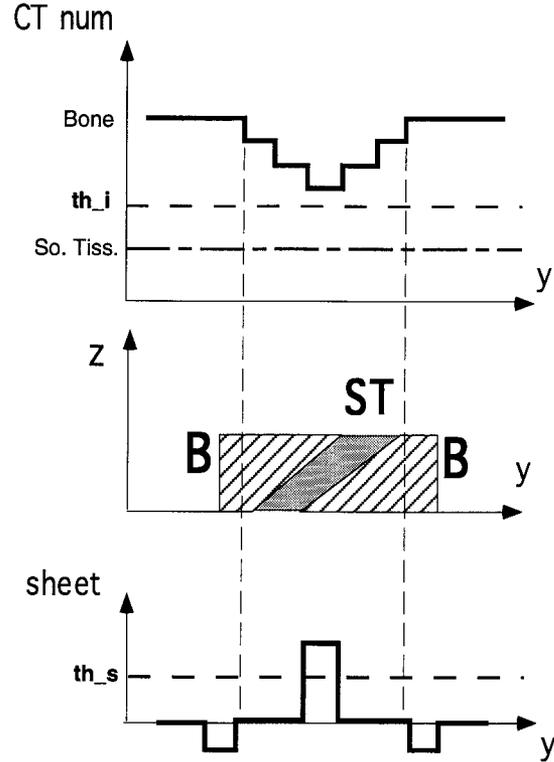

Fig 1. A small sheet of ST surrounded by bone (BO) (middle) and its response on CT (upper) illustrating the influence of PVE is shown. When thresholding using th_i, a false connection does occur. However sheet detection (lower) could be used to benefit from the inherent information available in the soft tissue "valley." In this example, the sheet response is obtained by convolution of the upper profile with the mask [1 -2 1].

thresholding (th_s). User guided CCL can further be used to find the regions of connected bone pixels, surrounded by thresholded sheets. The border between such regions will in most of the cases correspond to the presumed disarticulation line. Now we describe our method hinging on Fig 2 to illustrate the mathematical techniques. The bone structure B_1 to be disarticulated (Fig 2A) is the scaphoid bone.

The Method

Thresholding. For the volume, I, a user interactively chooses a threshold, th_i, so that the total number of false openings and of false closings in the segmented bone is minimized (Fig 2A). This leads to the image volume, $T(x,y,z)$:
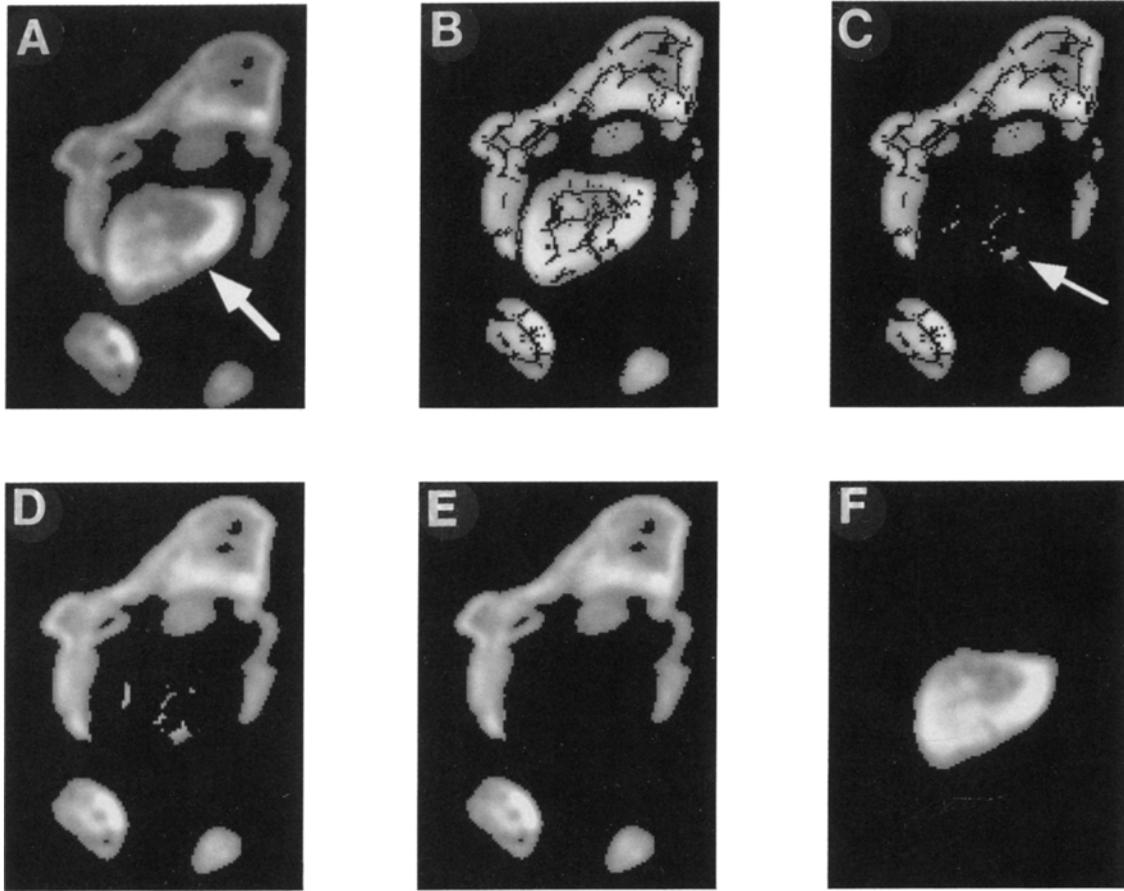
if $(I(x, y, x) > = th\_i)$ then

$$T(x, y, z) = I(x, y, z) \quad \text{and} \quad [x, y, z] \in B$$

or else

$$T(x, y, z) = 0 \quad \text{and} \quad [x, y, z] \in NB$$

Gaussian smoothing and sheet detection. From I, a sheet image, S, is created. It is well known that sheet-like

Fig 2. The semiautomatic disarticulation method applied to a scaphoid bone from a wrist joint is shown. Each subfigure shows the corresponding part from the same 2D slice through different volumes: (A) result of thresholding (remarking the connection of the scaphoid bone), (B) combination of sheet detection and thresholding, (C) deselection of the larger part of the scaphoid, (D) dilation of the complement, (E) removal of flying voxels, (F) taking the complement of E leads to the bone structure.

structures in 3D density volumes strongly respond to second directional derivatives. To suppress noise and to detect smoother sheets, we convolve I with a 3D Gaussian mask, g, first.

$$G(x, y, z) = g(x, y, z, \sigma_x, \sigma_y = \sigma_x, \sigma_z) **I(x, y, z) \qquad (3)$$

where σ_i is the variance of the Gaussian function in direction i, and ** is the symbol for convolution. Next, a discrete mask $(1 \; -2 \; 1)$ is convolved with G in the 13 discrete 3D directions, D_{13}, for taking second directional derivatives. The result of applying this operator to a one-dimensional case has already been shown in Fig 1. The maximum of the 13 convolution responses, and the direction $v_d \in D_{13}$ at which this maximum occurs, is retained for each voxel v. To suppress double response to the same sheet, nonmaxima suppression (NMS)[7,8] is used. This means that the maximum found for a given voxel v will be kept only if it is also larger than the maxima for its neighboring voxels in the direction v_d. So the volume S is defined by,

$$S(x, y, z) = NMS\left[\max_{d \in D_{13}} \{[1 \quad -2 \quad 1]_d **G\}\right] \qquad (4)$$

Combination of sheet volume, S, and thresholded volume, T. Now we create an image volume $C(x,y,z)$, which combines T and S. One parameter, the sheet threshold, th_s, must be chosen interactively, so that the presumed disarticulation line is not interrupted. $C(x,y,z)$ is formed in the following way:

if $S(x, y, z) < th\_s$, then $C(x, y, z) = T(x, y, z)$;

if $S(x, y, z) \geq th\_s$, then $C(x, y, z) = 0$.

The value of th_s must be adequate for a given application. Too high a sheet threshold results in an incomplete disarticulation line and does not eliminate the drawing completely. Too low a sheet threshold results in an overfragmented image. In this case, several fragments must be removed separately by CCL. Figure 2B shows a slice of image C corresponding to Fig 2A. Note how the disarticulation line is made up from sheets.

Semiautomatic disarticulation of two different bone structures. Every voxel satisfying $C(x,y,z) = 0$ is a possible member of a disarticulation line. We apply user-guided

CCL in C to obtain a 3D region W, containing connected nonzero voxels. Region W must be part of only one bone structure, $W \subset B_1$, and some of its border voxels must be neighboring the presumed disarticulation line. By knowing W, we can completely determine B_1 and its complement, respectively, in the entire bone region B in a final labeling step (see next step).

We have developed two methods: 2D CCL and adapted 3D CCL. It will be clear from the following discussion why a true 3D CCL is not feasible in our method. It is surely not an issue of user interface, but related to the fact that there is no 100% guarantee that all disarticulation voxels have a zero response in C.

In 2D CCL, the region W is formed as a union of 2D regions W_z. To create W_z in a slice z, the user has to select (with a mouse click) a seed pixel $s_z \in B_1$. s_z becomes the initial member of W_z. Region W_z is grown by adding the four orthogonal neighbors n of a pixel $p \in W_z$, that belong to the same z slice and satisfy $C(n) \neq 0$. A result is shown in Fig 2C: almost the entire scaphoid visible on this slice is part of W_z and made invisible.

In adapted 3D CCL, the region W is also formed as a union of 2D regions W_z; however, the seed pixels for a slice $z + 1$ can be derived automatically from the W_z found in the slice z. Adapted 3D CCL is mainly meant to propagate information found in one slice towards its neighboring slices. Suppose voxel $p[x,y,z] \in W_z$ and there exists a 2D environment U_p of p such that $U_p = \{q \,|\, \text{if } C(q) \neq 0 \text{ then } q \in W_z\}$. Then the voxel $p'[x,y,z + 1]$ can be then taken as the seed of $W_{(z+1)}$. The parameter describing the digital environment U_p is user defined and is meant to exclude the voxels neighboring the disarticulation line from the 3D operation. In this way, we take into account that one bone structure in the z slice (eg, acetabulum) can be connected to another structure (eg, femoral head) in the slice at $z + 1$. Because of the lower resolution in the z direction, the connection area can be several pixels wide. Therefore, the voxels lying close to the disarticulation line must be excluded to generate seed pixels in the neighboring slices.

In most cases, W is built up without any further user interaction. Disarticulation lines will be found automatically from sheets when CCL is applied. However when CCL fails on slice z, the condition $W \subset B_1$ is not automatically valid and, therefore, manual separation by drawing must be used on slice z to delineate the proper part W_z of W.

Final labeling. The region, W, is a part of the bone region, B. To obtain B_1 from B and W, the next procedure is followed.

First, W is deselected. This means that its voxels are not considered any longer for further computational operations, until W is selected again. Fig 2C shows the result of this deselection for the scaphoid bone. Because of the construction of W, this deselection can be done by 3D CCL. Pointed out in the center part of Fig 2C are flying voxels, obviously belonging to B_1 but not to W. Second, the remaining bone structure is dilated using a 3- × 3-pixel structuring element, e, with values 1.[9] Figure 2D shows the result of this operation. Third, as the regions of flying voxels of B_1 are not connected to the complement of B_1, they can be deselected by 3D CCL. This is shown in Fig 2E. Finally, by taking the complement of the thus remaining bone structure of B, B_1 is obtained (Fig 2F).

To use this approach to disarticulate n bone structures in B, the method needs to be applied at most $n - 1$ times.
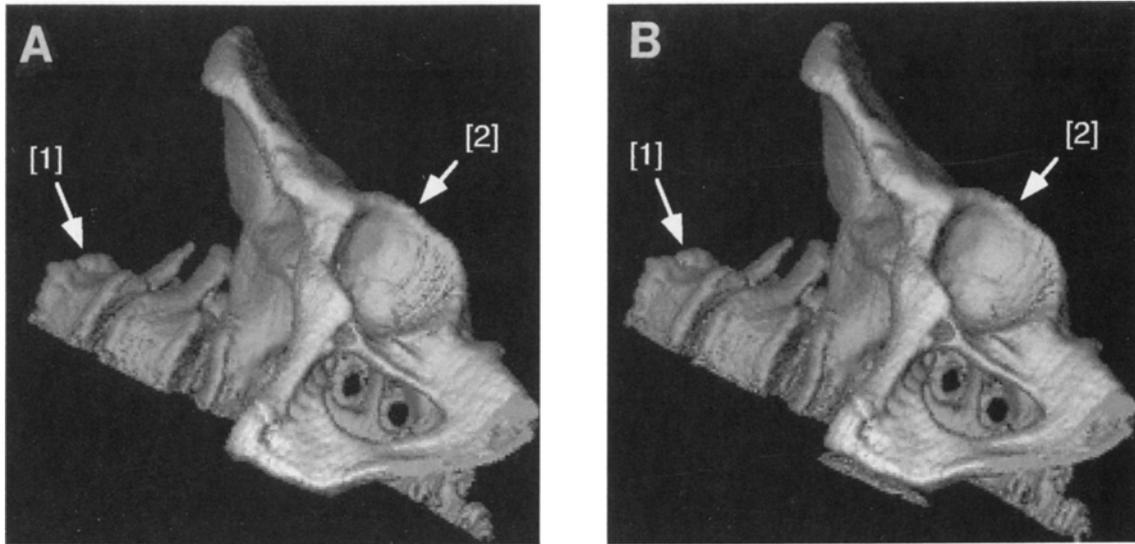
RESULTS

Here we describe a validation test of our algorithm on ten 3D spiral CT image volumes of eight different patients, (on the average, 115 slices per volume). Two patients were scanned twice using different clinically relevant acquisition parameters. In seven images, the purpose was to separate the left and/or right femoral head from its acetabulum; in two images, the radius wrist bone needed to be disarticulated from the ulna wrist bone on a right and a left hand; finally, the talus bone in a foot had to be separated from the tibia, fibula, and calcaneus bones. The amount of user interaction needed on slice level is compared for manual disarticulation and for our semiautomatic cutting algorithm. Both the manual and semiautomatic disarticulation were done by an experienced radiologist.

Applying manual separation, one or two disarticulation lines had to be drawn on 313 slices of the test set. On the contrary, our semiautomatic technique required drawing a disarticulation line on only 64 slices. Moreover, these lines are much shorter, about one eighth of the average length in the manual case. The efficiency of our semiautomatic algorithm is higher when spiral CT acquisition parameters (slice thickness and table feed) correspond to the measurements of finer details. The results of our experiments are summarized in Table 1.

Table 1. Summary of Experiments

Image	Patient	Acquistion (slice-thickness/ table feed)	No. of Slices Manual Disarticulation	No. of Slices Semiautomatic Disarticulation
Pelvis1	p1	2/2	32	2
Pelvis2	p2	2/2	14	1
Pelvis3	p3	5/5	32	7
Pelvis4	p4	5/5	45	33
Pelvis5	p5	2/2	40	1
Pelvis6	p5	5/6	40	14
Pelvis7	p6	2/2	7	1
Wrist8	p7	1/1	14	0
Wrist9	p7	2/2	27	0
Foot10	p8	2/2	62	5

Column 3 shows the relevant spiral CT acquisition parameters. Column 4 shows the number of slices that required drawing when manually disarticulating. Column 5 shows the number of slices that required drawing when our semiautomatic algorithm is applied.

Fig 3.   3D rendition of a pelvic region (spine, 1; acetabulum 2): (A) Result of visualizing a manual segmentation; (B) Result of visualizing the result of our semiautomatic approach.

Using manual disarticulation, user-guided 2D CCL needs to be applied on each slice after drawing the disarticulation line. On the other hand, the adapted 3D CCL option implemented as part of our semiautomatic approach does not require the pointing out of such seed pixels in every slice.

Because the cutting is needed only in a certain 3D volume, the user can restrict the sheet calculation to a 3D region of interest. This feature increases the speed of noninteractive preprocessing and postprocessing (steps 1, 2, 3, and 5) significantly. The speed of the interactive step 4 depends on the number of user interactions (Table 1).

We have also compared the smoothness of the rendered surface near the area of the disarticulation. The same surface rendering method is used to depict the results of manual and semiautomatic disarticulation. In 80% of our cases, the rendered disarticulation areas show a smoother appearance when the semiautomatic method was used; the other cases are inconclusive. Examples are shown in Figs 3, 4, and 5.

CONCLUSION

We have designed a method for semiautomatic disarticulation of joint bone structures. This algorithm is implemented as part of an
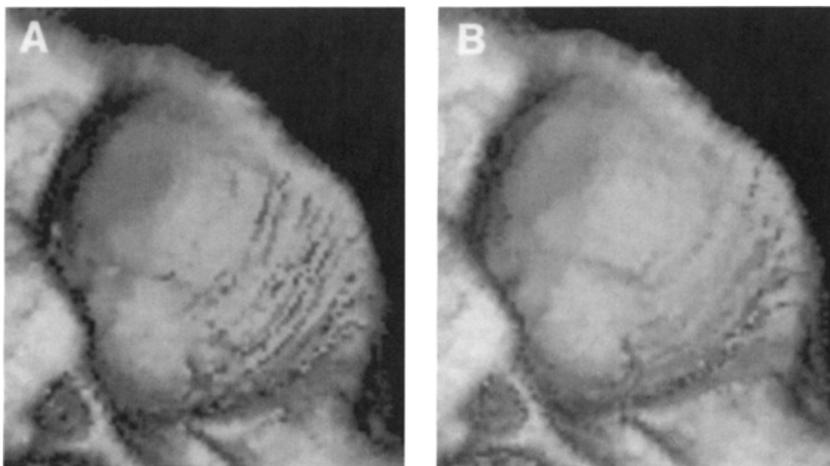
Fig 4.   Detail of the rendered acetabula of Fig 3.
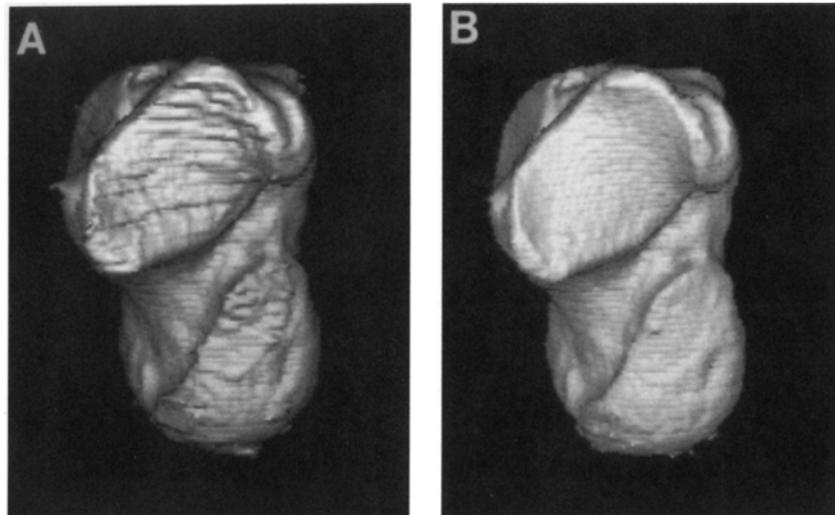
Fig 5.   3D renditions of a talus foot bone: (A) Result of visualizing a manual segmentation; (B) Result of visualizing the result of our semiautomatic approach.

in-house developed 3D segmentation software[10] that runs on high-end Unix workstations (IBM RS6000/370). We have applied this technique to the segmentation of bone on spiral CT images. The tests shows that the user interaction is significantly reduced when compared with manual delineation. On the average, the disarticulation of a femoral head from an acetabulum on a 250 slices spiral CT dataset takes less than half an hour, including the time to load the data into the workstation's memory.

As the efficiency of our method increases with finer detail acquisitions (spiral or incremental), it can be extrapolated that for example on magnetic resonance imaging (MRI), this tech-

nique will probably be beneficial only when high-resolution 3D acquisitions are considered. Contrary to spiral CT, such 3D MRI acquisitions take very long acquisition times.

Moreover the rendered surface near the area of semiautomatic disarticulation clearly shows to be smoother than the corresponding area of the manual disarticulation when identical 3D surface rendering techniques are used. A medical model validation protocol defined in the context of the PHIDIAS (laser photopolymerization models based on medical imaging, a development improving the accuracy of surgery) project is applied to the resulting surfaces also.[11]

REFERENCES

1. Fishman EK, and Ney DR: Computers for clinical practice and education in radiology. RadioGraphics 13:463-475, 1993

2. Crawford CR, Starshak RJ, Waisman RC: Three-dimensional display of mathematically disarticulated objects. Computer Assisted Radiology Conference, Berlin, Germany, July 1989

3. Gerber JD, Magid D, Fishman EK: Simulated femoral repositioning with three-dimensional CT. J Comput Assist Tomogr 15:121-125, 1991

4. Ezrielev J, Dzik S: An image editor for 3d-CT reconstruction system. SPIE Medical Imaging IV, 1233:67-76, 1990

5. Mankovich NJ, White SC: The appearance of bony lesions on 3-D CT reconstructions: A case study in variable renderings. SPIE Image Capture Formatting Display 1653: 443-450, 1992

6. Kratka D, Van Cleynenbreugel J, Suetens P, et al:

Binary bone segmentation: Problems and improvements. Pattern Recognition Image Anal 5:114-120, 1995

7. Canny J: Acomputational approach to edge detection. IEEE Trans Pattern Anal Machine Intelligence 8:679-698, 1986

8. Lacroix V: A three module strategy for edge detection. IEEE Trans Pattern Anal Machine Intelligence 10:803-810, 1988

9. Serra J: Image Analysis and Mathematical Morphology. London, UK, Academic, 1982

10. Van Cleynenbreugel J, Vandermeulen D, Smet M-H, et al: An interactive environment for efficient 3D segmentation and flexible 3D visualization. Laboratory for Medical Imaging Research. Leuven, Belgium, January 1995 (internal report)

11. Smet M-H, Marchal G, Suetens P: Medical model requirements and validation of medical models. PHIDIAS Mid Term Report Annex H, PHIDIAS Consortium, Leuven, Belgium, September 1994