Use of Shape Models to Search Digitized Spine X-rays

L. Rodney Long, G. R. Thoma
National Library of Medicine, Bethesda, MD 20894
long@nlm.nih.gov, thoma@nlm.nih.gov

Abstract

We are building a biomedical information resource consisting of digitized x-ray images and associated textual data from national health surveys. This resource, the Web-based Medical Information Retrieval System, or WebMIRS, is currently in beta test. In a future WebMIRS system, we plan to have not only text and raw image data, but quantitative anatomical feature information derived from the images and capability to retrieve images based on image characteristics, either alone or in conjunction with text descriptions associated with the images. Our archive consists of data collected in the second and third National Health and Nutrition Examination Surveys (NHANES), conducted by the National Center for Health Statistics. For the NHANES II survey, the records contain information for approximately 20,000 participants. Each record contains about two thousand data points, including demographic information, answers to health questionnaires, anthropometric information, and the results of a physician's examination. In addition, approximately 10,000 cervical spine and 7,000 lumbar spine x-rays were collected. WebMIRS makes the text and images retrievable. Only raw images are returned; no quantitative or descriptive information about the images is stored in the database. We are conducting research into the problem of automatically or semi-automatically segmenting spine vertebrae in these images and determining vertebral boundaries with enough accuracy to be useful in classifying the vertebrae into categories of interest to researchers in osteoarthritis.

1. Introduction

We have described the architecture and characteristics of the WebMIRS system which provides access to the NHANES II images, in previous papers [1-2].The images in the NHANES II survey were collected primarily for the study of osteoarthritis and degenerative disc disease. Biomedical features of interest to researchers in these areas have been identified by two workshops conducted at NIH, and consist of anterior osteophytes, disc space narrowing, and subluxation. A database containing quantitative and descriptive image information that allows intelligent search for images that show varying degrees of these features is expected to be valuable for the osteoarthritis and related research communities. In addition, published research work [3-4] in the field of vertebral morphometry suggests that a database with quantitative measures of vertebral dimensions may be useful for purposes such as studies in occurrence of spinal fracture and estimation of normative values for vertebral size.

In order to derive such information directly from the images, detailed analysis and measurement of each image is required. Due to the prohibitive cost of manually carrying out such a process for large image archives such as the 17,000 images in our collection, we are

conducting research into methods for automating or semi-automating the process. In this paper, we describe work toward this goal, which consists of applying active shape modeling (ASM) techniques to the problem of locating anatomical structures (primarily vertebrae) within the images.

2. Approach

Active shape modeling (ASM). The purpose of ASM is to identify the best match to an object model from possible objects contained within a given image. The object model has two components: a *shape* model and *grayscale* model. The shape model is created from statistics (mean and covariance) of (x,y) coordinate values of points that lie on the boundaries of samples of the object. In our case, we sampled boundary points on particular vertebra (C2 and C3) boundaries in cervical spine images. The mean value of the sample shapes becomes the *mean shape* for the shape model. The covariance of the sample shapes is used to constrain the shapes to which ASM may converge. The grayscale model is created by sampling the *grayscale profile* (or grayscale difference profile) along normals to each of the points that define a sample boundary. The statistics of these samples are used to estimate probability density functions for the grayscale profiles at each boundary point in the shape model. Hence, we have a way of estimating the *expected* grayscale profile at each boundary point. ASM is an iterative algorithm; it is initialized by setting the current shape estimate to the mean shape, and choosing a priori values for scale, rotation, and position of the mean shape within the image to be searched. Then, at each step the current shape is allowed to deform by using grayscale model statistics to move each boundary point to a new position having a grayscale profile closer to the expected grayscale profile for that boundary point. The shape model covariance is used to constrain the new shape to values close to the original sample shapes. Convergence is defined by the step-to-step changes in shape, scale, rotation, and position falling below a tolerance value.

A general reference to ASM as we have used it is given in [5]. Smyth [6] has shown success in applying ASM techniques to the location of lumbar spine vertebrae acquired by dual x-ray absorptiometry (DXA). We are taking a similar approach to investigate the use of ASM for location of vertebrae within digitized x-ray film. We are extending work that we reported in [7].

The sample images for our work are chosen from the NHANES II x-ray image collection referenced above. From these samples we have collected landmark data on the boundaries of cervical spine C2 and C3 vertebrae; the mean and covariance of these landmarks comprise the basic statistical data used for the ASM searches.

3. Results

To make an initial assessment of ASM performance on these images, we have concentrated on one boundary area (C2/C3) and on one complete vertebra (C3) in the cervical spine. The model we have constructed represents the inferior (bottom) edge of C2, which is the topmost vertebra structure consistently capable of being visually detected in the images; and the complete boundary of C3. To model these shapes, we used 81 landmark points, which is roughly consistent with the number of landmarks used by Smyth [6] for DXA images of the lumbar spine. Our model used 40 1462x1755 cervical spine images as input for a total of 3240 manually-collected points. Not only does this data collection involve a nontrivial amount of labor, but unless care is taken, the collected points may not exhibit any mathematical

consistency as edge points. For example, points which are selected as edge points strictly on the basis of their visual appearance may lie several pixels from the edge point as defined by maximum gradient, or by maximum derivative along a given line segment (such as a line segment normal to a boundary). On the other hand, trying to collect points in a completely automated fashion with a conventional edge detector results in many spurious points that have to be manually removed. For these reasons, we created a software tool (LMARK) that allows manual collection of the landmark points, but provides computer assistance.   LMARK operation is illustrated in Figures 1-2, which show the two main windows of the tool. Figure 1 is the image display window that shows the region of interest (ROI) around a particular vertebra as well as the landmark points as they are collected. The image in the ROI may have its contrast enhanced by histogram equalization. In Figure 1 one of the points is "current" and has an intersecting line segment that is approximately normal to the vertebra edge. The window in Figure 2 displays the smoothed derivatives of the grayscale values along this segment. An additional window (not shown) displays a graph of the smoothed grayscale values along this same segment. The user specifies on the command line (not shown) which point along the profile should be take as the edge, or landmark, point. Typically, the point with maximum absolute derivative is selected, which would be point 26 in Figure 2. When this point is selected, it is marked in red on the graphs in the grayscale profile and the derivative windows, and the point also appears marked in the image window. If the image window shows the selected point is in fact not on the edge (as sometimes happens due to noise or a digitized scratch on the image) the user may adjust the choice of the selected point.

Figure 3 shows the results of consistency testing for boundary detection of C2/C3 in these 40 images. That is, we built a model using the entire 40 image set, then used the model to detect C2/C3 in each of the images that contributed to the model. Presumably, if the image data is of sufficient quality, and the boundary detection method is sound, we should get reasonably good boundary locations. The Figure 3 plot shows the results obtained, in the form of average point-to-point errors as measured between boundary points for the converged boundary solution for each image, and the "known truth" boundary points that were manually collected. For 16 of the 40 images, the errors were less than 15 pixels, which corresponds to an error of less than 1/10 inch in the original image. Figure 4 illustrates one of the cases for which convergence was accurate.   The converged solution overlays the truth boundary so well that the two are indistinguishable in the figure.   For the remaining images most of the erroneous solutions included a boundary segment close to the truth boundary, connected to one or more segments not on the truth boundary, but rather on other visually-detectable edges, such as tissue/background edges or even edges produced by grayscale intensity variation within a vertebra. Figure 5 illustrates such a case, where part of the converged solution lies on a tissue/background boundary, part on curves formed by contrast differences that are actually within the interior regions of vertebrae, and part lies on true vertebral boundaries.

To get an indicator of how we might expect the ASM algorithm to perform on at least a subset of the cervical spine images, we used only the 16 images that performed best in the above test to build a C2/C3 model. The results of a consistency test that searched each of these 16 images for the model created using the same 16 images is shown as the bottom, solid-line curve in Figure 6. With the exception of one image, the ASM solutions were all better than the ASM solutions for these same images when the 40-image model was used. The solution for the exceptional image exhibited strong sensitivity to initial positioning of the model on the image, and a tendency to attract to the C3/C4 area, rather than C2/C3. For this image, the error value shown in Figure 6 is the mean error resulting from 10 different random initial positionings of

the model. With this one exception, the consistency errors remained below the 1/10 inch level for mean point-to-point error.

We then used this 16-image model to perform cross-validation testing by iteratively (1) deleting one image from the model; (2) building a model with the remaining 15 images; and (3) searching for the model shape within the image that was omitted. This is expected to give an indicator of how the ASM algorithm could be expected to perform on images that it has not been trained on, but which are similar to the 16-image test set. These cross-validation results are shown as the top (dashed) line in Figure 6. For each image, the error in the ASM cross-validation test exceeds the error in the consistency test, as would be expected, since the consistency model was built with all 16 images: each image being searched itself contributed to the statistics of the model.

Analysis. We make the following observations: (1) Our work to date is at a preliminary stage. Results should be considered suggestive, but not conclusive. One reason for this is that the number of images sampled (40) is quite small relative to our total image population (10,000 cervical spine images). (2) Results obtained with the 16-image model suggest that, for at least a significant subset of our images, mean point-to-point error bounds of 0.2-0.3 inch may be achievable. Even though this level of accuracy may not be of direct use for the long-term goal of vertebra feature classification, it is sufficient to support systems that automatically identify and label individual vertebra ("C2", "C3", etc.), and may improve as we continue our work. (3) The C2/C3 model that we are using is itself only a partial model for the cervical spine. One implication of this is that, due to similarities between the vertebral bodies, the ASM algorithm cannot always be expected to attract to the correct region in the vertebral column. Occasionally, for example, cases occur where the model converges to C3/C4, rather than C2/C3. These cases are expected to diminish in number or disappear as we expand the model to represent more of the cervical spine. (4) The question of initial conditions for the model remains largely unexplored. In almost all of the cases, we used the default initial state generated by the commercial software we were using. The sensitivity of the convergence to this initial state is an area requiring further investigation: how should we initialize the model scale, rotation, and positioning to insure or enhance its chances of lying in a region of convergence for the algorithm? In previous work [8] we demonstrated some initial success in getting basic orientation information from cervical spine images that may be useful in initializing ASM. We developed an algorithm that automatically places an orthogonal coordinate system with one axis U at the base of the skull and the perpendicular axis V positioned to intersect this axis at the point O (coordinate system origin) of greatest bone density in the spine (as defined by brightest grayscale concentration. One of our areas of research is to investigate whether this U/V system location algorithm is sufficiently robust for its use in initializing ASM searches.

For this work we have used MATLAB 5.3 and the ASMToolkit for MATLAB. This toolkit has been a valuable resource in initiating work in this research area, but our work has now progressed to the point where certain apparent limitations in the software have surfaced. Examples are limitations on the number of points that our models may include, inability to precisely know or set initial conditions, and inability to know quantitative convergence measures.

4. Conclusion

Our goal of indexing digitized x-ray images of the cervical spine for content relevant to researchers in the field of spine disease and injury consists of many steps of which the first is to first obtain segmentation of vertebral structures using active shape modeling (ASM). Evaluation of the usefulness of this technology for indexing a large image collection of digitized x-rays such as ours is continuing.   A key concern in our work is to assess the degree to which computerized methods may assist in the accurate derivation of such x-ray image content and to create a concept of human-machine interaction that exploits the strengths of both the human participant and computer capability for solution of this particular problem.

5. References

1. Long LR, Pillemer SR, Lawrence RC, Goh G-H, Neve L, Thoma GR. World Wide Web platform-independent access to biomedical text/image databases. *Proceedings of SPIE Medical Imaging 1998: PACS Design and Evaluation: Engineering and Clinical Issues,* SPIE Vol. 3339, San Diego, CA, February 21-26, 1998, pp. 52-63.
2. Long LR, Pillemer SR, Lawrence RC, Goh G-H, Neve L, Thoma GR. WebMIRS: Web-based Medical Information Retrieval System. *Proceedings of SPIE Storage and Retrieval for Image and Video Databases IV*, San Jose, CA, January 28-30, 1998, SPIE vol. 3312, pp. 392-403.
3. Rosol MS, Cohen GL, Halpern EF, Chew FS, Kattapuram SV*,* Palmer WE, Dupuy DE, Rosenthal DI. Vertebral morphometry derived from digital images. *American Journal of Roentgenology*, Vol. 167, December 1996, pp. 1545-1549.
4. Hedlund LR, Gallagher JC. Vertebral morphometry in diagnosis of spinal fractures. *Bone and Mineral*, Vol. 5, 1988, pp. 59-67.
5. Cootes TF, Taylor CJ. Statistical models of appearance for computer vision. Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, September 13, 1999. Draft report, available at http://www.wiau.man.ac.uk.
6. Smyth PP, Taylor CJ, Adams JE. Vertebral shape: automatic measurements with active shape models. Radiology, May 1999, vol. 211, no. 2, pp. 571-578.
7. Long LR, Thoma GR. Segmentation and image navigation in digitized spine x-rays. Proceedings of SPIE Medical Imaging 2000, San Diego, California, February 26-29, 2000, (Forthcoming).
8. Long LR, Thoma GR. Segmentation and feature extraction of cervical spine x-ray images. Proceedings of SPIE Medical Imaging 1999: Image Processing. Vol. 3661, San Diego, CA, February 20-26, 1999, pp. 1037-1046.

Figure 1. Landmark data collection. Image display window showing some collected landmarks. The line segment through the last point is approximately normal to the vertebra boundary.
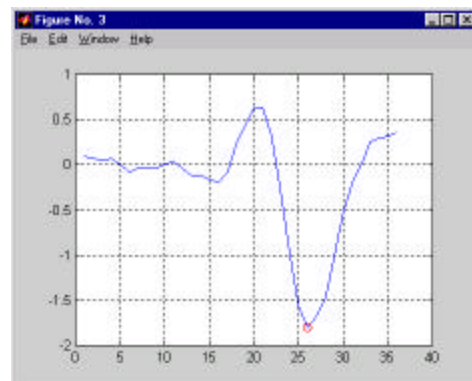
Figure   2.   Landmark   data collection. Smoothed derivatives of grayscale values along a normal line segment.   Point 26 has maximum absolute value.
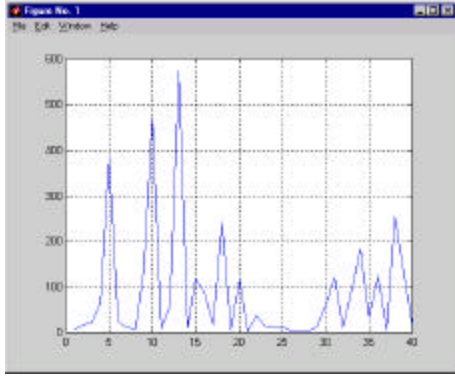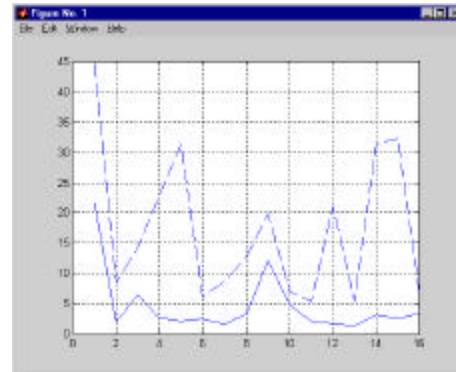
Figure 3. Errors on 40 image test. Vertical axis is mean point-to-point distance between converged solution and "truth" (manually-collected landmarks). Horizontal axis is image index number. Many of the large errors appear to correspond to (erroneous) convergence to local minima.
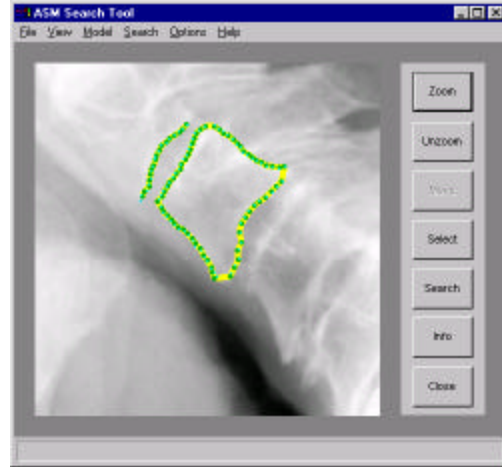
Figure 4. Example of good convergence for the C2/C3 40-image model. The boundary of the converged shape overlays the truth boundary. (Avg. point-point error = 1.67 pixels.)

Figure 5. Example of erroneous convergence for the C2/C3 40-image model. The anterior of the converged solution has fixed on a tissue/background boundary; the superior and inferior segments of the solution have fixed on edge curves that lie within the C3 and C4 vertebrae, respectively. The posterior segment of the solution partially overlaps the truth shape. (Avg. point-point error = 64.96 pixels).

Figure 6. Solid line: consistency test results for the C3/C4 model built with the best performing 16 images in Figure 4. Dashed line: cross-validation test results for this model. Vertical axis is mean point-to-point distance between converged solution and "truth" (manually-collected landmarks). Horizontal axis is image index number.