Evaluating Partial Shape Queries for Pathology-based Retrieval of Vertebra

Jason Shou*, Sameer Antani, L. Rodney Long, George R. Thoma
Lister Hill National Center for Biomedical Communications
U. S. National Library of Medicine, National Institutes of Health, DHHS
Bethesda, MD 20894, USA

ABSTRACT

There is a growing research interest in reliable content-based image retrieval (CBIR) methods for biomedical images. Feature representation algorithms used in indexing medical images have to address conflicting goals of reducing feature dimensionality while retaining important and often subtle biomedical features. In a collection of digitized X-rays of the spine, for example, a typical user is interested in only a small region of the vertebral body pertinent to the pathology: for this experiment, the Anterior Osteophyte (AO). To address this, we have proposed the use of partial shape matching (PSM) methods for spinal X-ray image retrieval. This paper describes an evaluation of a Procrustes distance-based PSM method for pathology-based retrieval. It was conducted on a subset of data selected from a collection of 17,000 digitized X-rays of the spine from the Second National Health and Nutritional Examination Survey (NHANES II) maintained at the U.S. National Library of Medicine. Two classifications for AO were used to evaluate the PSM algorithm, viz., Macnab classification, which is a published method, and a grading system developed by us. The evaluation based on *ground truth* established by the classifications shows that the PSM algorithm is a promising approach for content-based image retrieval of biomedical images.

Keywords: Partial Shape Matching, Procrustes Distance, NHANES II, Anterior Osteophytes, Performance Evaluation, Medical Image Databases, Content-Based Image Retrieval.

1. INTRODUCTION

There has been increasing research interest in developing content-based image retrieval (CBIR) methods for biomedical images [1]. However, due to the lack of effective automated methods, these images typically are manually annotated and retrieved using a text keyword-based search. Even such retrieval is impossible for collections of images that have not been annotated or indexed. CBIR methods developed specifically for biomedical images could offer direct searching of the rich visual content of these images.

At the Lister Hill National Center for Biomedical Communications, an R&D division of the U.S. National Library of Medicine, we maintain a collection of 17,000 digitized spine X-rays from the Second National Health and Nutritional Examination Survey (NHANES II). These X-rays, in which anterior osteophytes are present, serve as a rich data source for research into pathology-based retrieval of biomedical images. An osteophyte, a bony protuberance on normal bone surface, is a characteristic feature of degenerative joint disease of the spine. After careful study of the NHANES II image collection medical experts have previously determined that the anterior osteophyte (AO) is a pathological feature that can be reliably and consistently detected in the data set, along the anterior superior or inferior edges of the vertebral bodies. An example of superior and inferior AO is illustrated in Figure 1(a) and shown in 1(b). Points 2-3-8-7 describe a superior anterior osteophyte on a sagittal spinal X-ray and points 7-9-6-5 describe an inferior anterior osteophyte. In such cases, only a small region of the vertebral body is pertinent to the pathology.

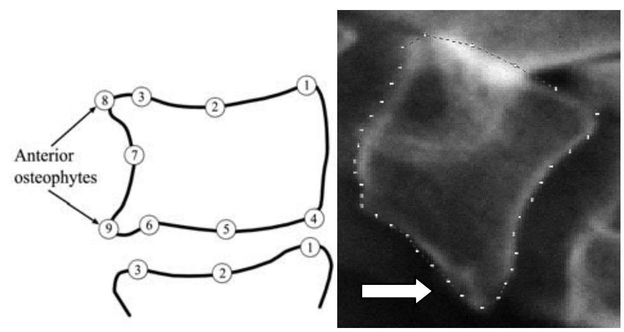


Figure 1. (a) Anterior Osteophytes illustrated on vertebral body outline. Points marked indicate regions of interest. (b) Cropped Spinal X-ray image showing inferior AO on vertebra with 36 boundary points superimposed.

Methods for querying vertebrae through visual means would be immensely useful to radiologists, bone morphometrists, medical students, and researchers of musculoskeletal diseases. Pathology in a vertebra can be captured in the outline of its body seen in a 2D projection in the sagittal spinal X-ray. In query-by-sketch and query-by-example paradigms of shape-based CBIR, traditional *whole* shape matching methods employed for retrieval of

---

* Jason Shou is currently with Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick, NJ 08901, USA

these shapes from a database are unable to retrieve vertebrae that precisely match the pathology classification of the query vertebral shape. This may be due to users' tendency to focus on the particular segment of interest on the shape boundary while disregarding shape characteristics in other boundary regions. The *whole* shape similarity methods match the entire boundary, often resulting in less than acceptable results [2]. To address this drawback, we have proposed the use of partial shape matching (PSM) methods in retrieving images of interest by allowing the user to sketch or identify the specific region of interest on the vertebral boundary.

This paper describes an evaluation of a particular partial shape matching method developed as part of the broader research into CBIR of medical images that has been reported earlier [1]. This shape matching method uses shapes described by a fixed number of boundary points and is based on the Procrustes distance [3]. The remainder of the paper is organized as follows. Section 2 presents a brief overview of PSM methods and PSM application to vertebral shapes. Section 3 describes the osteophyte classification schema, *ground truth* (reference set) and test set development. Section 4 details the performance evaluation strategy. Section 5 presents results and analysis. We conclude with Section 6 wherein we also briefly discuss our continuing and future work.

2. BACKGROUND ON PSM

There is little in the literature about incomplete or partial shape matching or on recognizing occluded objects. There are two broad classes of PSM techniques. One uses a fixed point description [4] and the other uses a variable number of boundary points for the shape data.

Variable Point PSM

A recent contribution by Petrakis et al [5] presents an approach for variable point open shape matching using Dynamic Programming (DP). Inflection points on the boundary serve as inputs to the shape representation method. The extracted shape features are length, area and rotation angle. The approach is to match merged sequences of consecutive small segments in a shape with larger segments of another shape, with an associated merging cost. DP selects the most promising merges and searches for the least cost match path from a table. Another approach by Mori et al [6] also employs DP as a variable point PSM method. They propose wedge wave feature extraction. The features are based on wedge wave convexity which consists of four factors: location on the contour, direction in which the contour is facing, scale of convexity and class of convexity (sharp or dull). Partial matching is enabled by extending DP matching to allow one-to-zero correspondence as well as one-to-one correspondence. The costs associated with one-to-zero correspondence share the same approach as in [5].

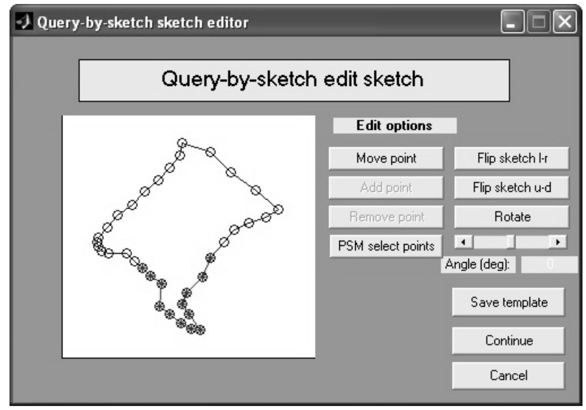


Figure 2. Selecting region of interest on the query shape.

Fixed Point PSM

This method is based on the Procrustes distance [3]. The Procrustes matching process does a point-to-point match for which it requires the same number of points on the user selected query shape and the target shape in the database. This is achieved by using a fixed number of points to describe the vertebral shapes. The method allows the user to select a region of interest along the vertebral boundary, highlighted in Figure 2. The selected query partial shape is then compared with every other region in the target shape. The matched shapes are listed in decreasing order of similarity with the region of interest highlighted, as shown in Figure 3.

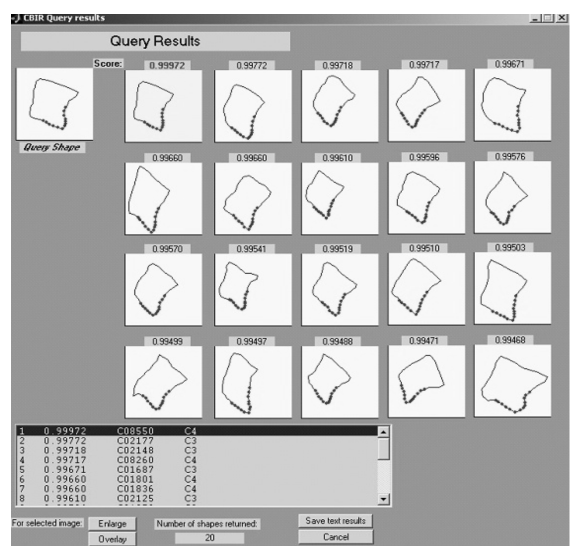


Figure 3. Result of PSM on inferior AO osteophyte .

Technique Overview

The Procrustes distance method performs a linear transformation on one shape to find the best match between two shapes. This is represented by the Equation 1, where x,y and x',y' are n boundary point coordinates of shapes X and X'.

$$P = \sum_{i=1}^{n} \left\| \begin{bmatrix} S \cdot \cos\theta & -\sin\theta & T_X \\ \sin\theta & S.\cos\theta & T_Y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}_X - \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix}_{X'} \right\|^2 \quad (1)$$

Equation 1. Procrustes Distance

The matching process translates shape X by T_X,T_Y such that the center of gravity of the two shapes coincide. Next, the shape X is scaled by S and rotated by θ for the minimum sum of squared distances between the boundary points of the two shapes. The subtraction represents the Euclidean distance measure between two 2D points. In the context of shape space theory [3], this method finds the closest chord distance between two shapes.

3. DATA SET DEVELOPMENT

The data set for this experiment was generated from a total of 206 spinal X-ray images (106 cervical and 100 lumbar films) from subjects who were 60 years of age or older, selected from the NHANES II collection. The age criterion was used because of higher incidence of degenerative joint disease in this population. Two classification schemes for anterior osteophytes were established by a medical expert to evaluate the accuracy of the PSM algorithm. One is Macnab's classification, established by Macnab and his coworkers in 1956 on radiological and pathological bases [7, 8]. Two types of osteophytes are adapted from Macnab's classification: claw and traction. Claw spur rises from the vertebral rim and curves toward the adjacent disk. It is often triangular in shape and curved at the tips. Traction spur protrudes horizontally, is moderately thick, does not curve at the tips, and never extends across the intervertebral disk space.

The second classification is a grading system which was defined by the medical expert consistent with reasonable criteria for assigning severity levels to AO. Three grades of AO are *slight*, *moderate*, and *severe*. Slight grade includes normal, where the vertebral body is squarish. It may have a slight protuberance, where the tip of the osteophyte is round and no narrowing is observed at the base of the protuberance. Moderate grade is characterized by evident protuberance from the ideal horizontal or vertical edge of the vertebra. The bounding edges of the AO form an angle of at least 45 degrees and the osteophyte has a relatively wider base than severe grade. Severe grade is characterized by presence of hook, the angle is less than 45 degrees and has a narrow base, or protrudes far (about 1/3 of the length of the horizontal border) from the *normal* (ideal 90 degree) vertebral corner, as shown in Table 1 below. Vertebrae from the selected 206 spinal X-ray images were segmented using an Active Contour Segmentation algorithm. This algorithm, which implements orthogonal curves as gridlines prevents ambiguity in selection of boundary points due to intersecting grid lines at vertebra corners [9]. The process yields 896 segmented vertebrae, composed of 407 cervical (C3-C7) and 489 lumbar (L1-L5) vertebrae. Each vertebra is represented by 36 points.

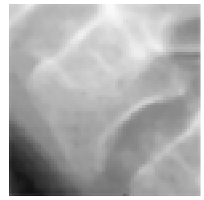
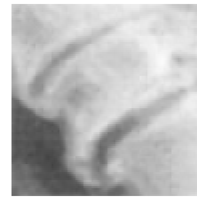
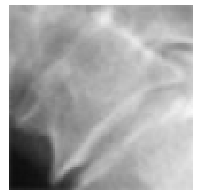
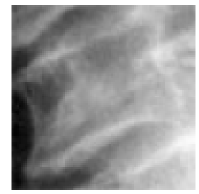
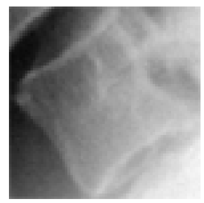
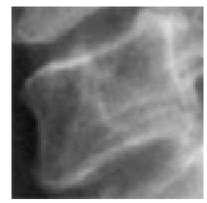
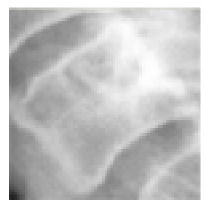
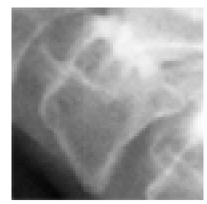
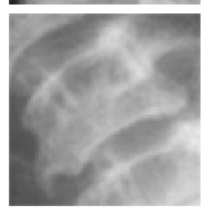
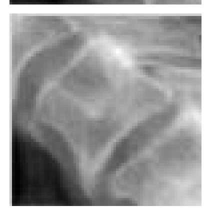
Pathology	**Example Images**	
Claw		
Traction		
Slight		
Moderate		
Severe		

Table 1. Macnab's classification and osteophyte severity grading scheme developed by us.

Ground Truth

Ground Truth or reference data is recorded into a MS Access database table based on the medical expert's observation of the selected spinal X-ray images, shown in Fig. 4. While necessary for our evaluation, caution must be taken in considering this set as a *gold standard*. Bias and inconsistencies of the medical expert may confound this reference set to some degree, especially in cases that are borderline between the definitions. As with any medical diagnosis, the classification must be regarded as an opinion. Ideally a ground truth set should be developed through some form of consensus on multiple expert opinions and such a development is a goal for us.

filename	vertebra	location	claw	traction	grade
C01235	C3	inferior	☑	☐	moderate
C01235	C3	superior	☐	☐	slight
C01235	C4	inferior	☑	☐	severe
C01235	C4	superior	☐	☑	slight
C01235	C5	inferior	☐	☐	slight

Record: 1 of 1901

Figure 4. *Ground Truth* table.

Queries

The medical expert also developed a set of queries for each classification type described above. The query shapes were taken from the segmented data set. A set of 28 queries was developed using combinations of the classifications and locations of the osteophytes, listed in Table 2, below. With 2 instances of each classification, one cervical vertebra shape and the other a lumbar, the queries are developed for Superior and Inferior Claw type with Slight, Moderate and Severe pathologies; and similarly for Superior and Inferior Traction type. In addition, queries were developed for Superior and Inferior normal vertebrae.

	Slight	**Moderate**	**Severe**
Superior Claw	C01630_C3 L01659_L5	C01770_C4 L01267_L2	C01281_C5 L02201_L1
Inferior Claw	C01286_C4 L01804_L5	C01286_C3 L01367_L4	C01235_C4 L01818_L4
Superior Traction	C08142_C7 L01275_L5	C01351_C6 L01277_L3	L02204_L4 C02171_C5
Inferior Traction	C01292_C5 L01293_L4	C08550_C4 L01277_L2	C01706_C5 L01281_L1
Superior Normal	C08555_C3 L01235_L2		
Inferior Normal	C08550_C6 L01267_L1		

Table 2. Vertebral shapes selected for queries.

It should be noted that even though classification is available for inferior and superior locations of AO, the current form of our PSM algorithm does not, by design, distinguish one from the other. As such, an *anterior superior* query could, for example, result in matches from the posterior or inferior region of vertebra.

4. EVALUATION STRATEGY

In this section we describe the strategy for evaluating PSM performance in querying various types of AO. The approach is to study the top 10 results returned for each of the query types described in Table 2. The results are then saved as a comma separated text file. The metadata for each result include rank number, a file identifier, vertebra number, and normalized similarity distance. This text file is then converted into a MS Access table, called RESULTS. In addition a visual representation of the partial match region, as shown in Figure 3, is used to manually determine the correctness of anterior inferior or superior region matches. The MS Access RESULTS table is then manually annotated to reflect the results. An enhancement to the PSM matching technique that will make it sensitive to AO position is under development. A database join operation is then performed on RESULTS table with the GROUND TRUTH table on osteophyte pathology, position, and severity. Resulting data on number of matched results for each classification type; (viz., normal, claw, traction,) and severity; (viz., slight, moderate, and severe) are recorded. These allow a detailed statistical analysis on the algorithm performance and correlation between pathologies.

5. DISCUSSION

We present a summary of results and analysis of algorithm performance through charts and result tables. Four perspectives were obtained on the aggregate results. These are indicated in Figures 5, 6, and 7 and in Table 3. The charts indicate the proportion of retrieved vertebra that matches particular criteria of the query shape. These criteria are:

a. Severity of pathology depicted in the query shape, shown in Figure 5;
b. Macnab's classification indicated on the query shape, shown in Figure 6; and
c. Severity combined with Macnab's classification indicated on the query shape, shown in Figure 7.

Finally, correlation between Macnab class and pathology severity of retrieved vertebrae with the severity and location of pathology indicated on the query is studied in Table 3. It is important to note that the PSM algorithm uses only shape information to retrieve "matching" vertebrae. It is the intention of this evaluation to determine the extent to which shape information alone (along with the selected PSM algorithm) is able to correctly retrieve intended vertebra.

Figure 5 shows the proportion of 3 grades of severity exhibited by the retrieved vertebrae for each queried grade. For queries on slight grade, 88.3% of vertebrae were correctly retrieved. It should be noted that slight grade includes normal vertebrae since distinction between normal and slight is highly subjective. Results were relatively mixed for moderate and severe grades (45% correct, for each). From the data it appears that the PSM algorithm favors slight grade; 45% of slight grade vertebra are retrieved in queries for moderate grade and 26.3% in those for severe queries. This confusion is further analyzed in the chart in Figure 7.

Figure 6 shows the proportion of retrieved vertebrae whose Macnab class matched with that indicated in the query shape. Again, we see a 87.5% match for normal vertebra, 43.3% for claw and 49.2% for traction. It is also interesting to note that claw and traction have nearly identical confusion with normal vertebrae (near 30%) and with each other (near 17.5%). This is despite lack of algorithmic design to identify the class types based on features described in Section 3. These results are viewed as encouraging because raw partial shape information is able to separate the classes reasonably well, and the results show a near uniform value for inter-class confusion. Figure 6 also indicates a low confusion (less than 6%) for those vertebrae that have been marked as dual-category by the medical expert. We believe that incorporating the shape characteristics that define claw and traction pathologies and severity into the algorithm would greatly improve the results.

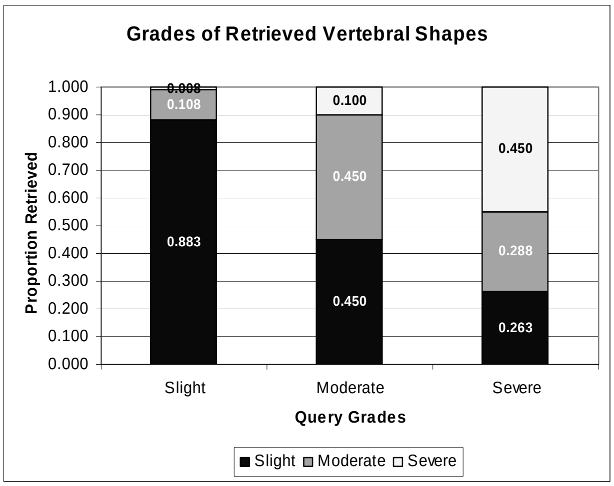


Figure 5. Proportion of 3 grades of severity exhibited by the retrieved vertebrae for each queried grade

In Figure 7, we analyze the correlation between the proportion of Macnab's classification types in retrieved vertebrae for queries that exhibited a combination of features from the Macnab's classification and the grading system. It is apparent from the results that the algorithm suffers greatly by the lack of training on particular shape characteristics of claw and traction. The method does fairly well in matching cases with pathology that is indicated by gross shape patterns. The algorithm tends to confuse moderate grades in claw or traction with normal vertebrae or those with slight grade since these could have a similar overall shape. Local angles on the osteophyte tip and their direction tendency can distinguish vertebra pathology and its severity. These can be especially unclear to the untrained algorithm in case of vertebra with moderate grade pathology. It should be noted that the figure ignores cases which have been marked as exhibiting both claw and traction pathologies causing the bars to not add up to 1.0

Table 3 shows the correlation of osteophyte severity and location with osteophyte severity and Macnab's classification types of retrieved vertebrae. The results reinforce the analysis that while retrieved results are promising, further improvement in the shape matching algorithm to allow for these features could greatly improve results.

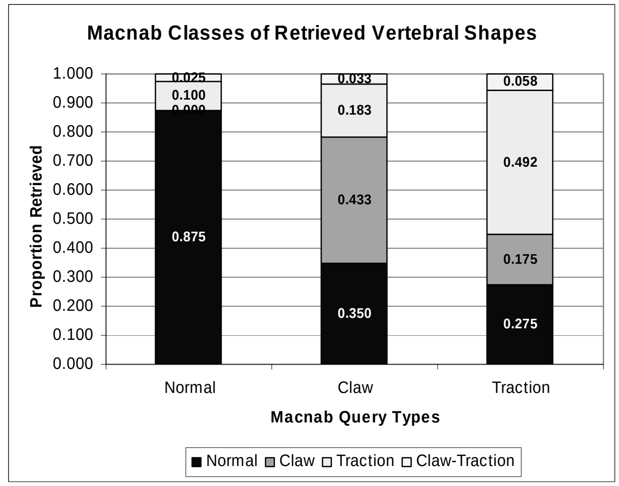


Figure 6. Proportion of retrieved vertebrae with their Macnab class correctly matching that indicated in the query shape.

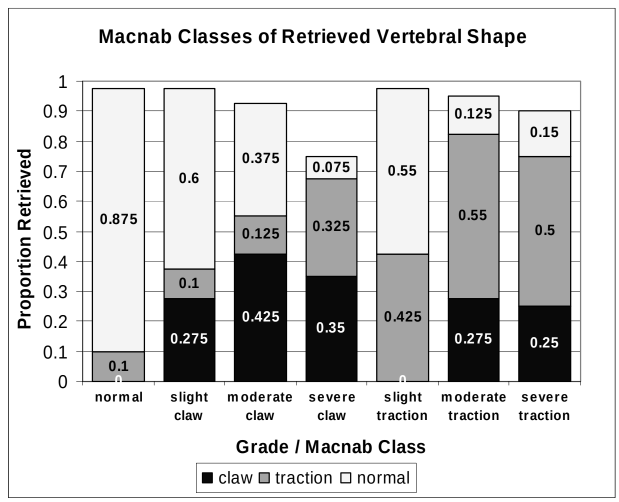


Figure 7. Correlation between the proportion of Macnab's classification types in retrieved vertebrae for queries that exhibited a combination of features from Macnab's classification and the grading system.

6. CONCLUSIONS AND FUTURE WORK

We have presented an evaluation of a partial shape matching algorithm for CBIR of vertebra shapes. The algorithm was evaluated on its ability to determine the pathology severity and differentiate among Macnab's classification types using partial shape of a vertebra that characterizes the pathological features. The intent is to study the extent a fixed point partial shape matching algorithm is capable in capturing query shape characteristics. The evaluation has yielded promising results from PSM and pointed to the need for further work on the topic. It has provided important insights into the behavior of the algorithm in classifying the vertebra on its severity and pathology. These indicate the need to incorporate the heuristics used to identify moderate and severe pathology and the Macnab's classification types into the PSM algorithm.

An important issue that this study reinforces is the need for a standardized ground truth data set that can be used to evaluate the methods. The ground truth set should be developed from multiple expert inputs that are suitably combined. Additionally, the evaluation results must be validated by a number of medical experts to minimize the effect of inter- and intra-observer variability. Our work in this area and further development of shape algorithms is continuing as a part of a larger framework of developing CBIR for medical images.

Acknowledgements

The PSM code was initially developed by Xiaoqian Xu of Brigham Young University based on an earlier prototype of whole shape matching developed by Daniel Krainak of Northwestern University. The segmentation program was initially developed by Xiaoning Qian of Yale University. All work was done as a part of their student internships at the U.S. National Library of Medicine.

REFERENCES

1. Long LR, Antani S, Lee DJ, Krainak DM, Thoma GR. **Biomedical Information from a National Collection of Spine X-rays: Film to Content-based Retrieval.** Proc. of the IS&T/SPIE Med. Imaging 2003 – PACS and Integrated Med. Info. Systems: Design and Evaluation, San Diego, CA. 2003 Feb., SPIE Vol. 5033:70-84.
2. Antani S, Long LR, Thoma GR, Lee DJ. **Anatomical shape representation in spine x-ray images**. Proc. 3rd IASTED International Conf. on Visualization, Imaging and Image Proc. (VIIP 2003), Sept. 2003; Benalmadena, Spain; Vol. 1:510-15.
3. Dryden IL, Mardia KV. **Statistical Shape Analysis.** John Wiley & Sons, Inc. 1998.
4. Antani S, Xu X, Long LR, Thoma GR. **Partial Shape Matching for CBIR of Spine X-ray Images.** Proc. IS&T/SPIE Electronic Imaging - Storage and Retrieval Methods and App. for Multimedia 2004. San Jose, CA; Jan. 2004; SPIE Vol. 5307: 1-8.
5. Petrakis EGM, Diplaros A, Milios E. **Matching and retrieval of distorted and occluded shapes using dynamic programming.** IEEE Trans. on Pattern Analysis and Machine Intelligence. 2002 Nov., 24(11):1501-16.
6. Mori K, Ohira M, Obata M, Wada K, Toraichi K. **A Partial Shape Matching Using Wedge Wave Feature Extraction.** Proc. of 1997 IEEE Pacific Rim Conf. on Communications, Computers and Signal Proc. -- '10 Years PACRIM 1987-1997 - Networking the Pacific Rim', 1997 August, Vol. 2: 835-838.
7. Heggeness MH, Doherty BJ. **Morphologic study of lumbar vertebral osteophytes.** South Med J. 1998 Feb;91(2):187-9.
8. Pate D, Goobar J, Resnick D, Haghighi P, Sartoris DJ, Pathria MN. **Traction Osteophytes of the Lumbar Spine: Radiographic-Pathologic Correlation.** Radiology. 1988 Mar;166(3):843-6.
9. Tagare H. **Deformable 2-D template matching using orthogonal curves**. IEEE Trans. on Medical Imaging, 16(1), Feb. 1997, 108-117.

Query Vertebra	Retrieved Vertebrae							
Severity + Location	Grades of Severity				Macnab's Classifcation Types			
	Slight	Moderate	Severe		Normal	Claw	Traction	Claw-Traction
Slight + Inferior	0.800	0.183	0.017		0.583	0.133	0.233	0.050
Slight + Superior	0.967	0.033	0.000		0.767	0.050	0.183	0.000
Moderate + Inferior	0.300	0.550	0.150		0.125	0.500	0.250	0.125
Moderate + Superior	0.600	0.350	0.050		0.375	0.200	0.425	0.000
Severe + Inferior	0.150	0.350	0.500		0.025	0.475	0.400	0.100
Severe + Superior	0.375	0.225	0.400		0.200	0.375	0.425	0.000

Table 3. Correlation of osteophyte severity and location with osteophyte severity and Macnab's classification types of retrieved vertebrae.