# Spine x-ray image retrieval using partial vertebral boundaries

Zhiyun Xue, L. Rodney Long, Sameer Antani, George R. Thoma
*National Library of Medicine, National Institutes of Health*
*{xuez,rlong,santani,gthoma}@mail.nih.gov*

## Abstract

*The anterior osteophyte (AO) is a bony spur on the vertebra and is symptomatic of osteo-arthritis of the spine. We present advances in our research into matching vertebral boundaries based on pathological (severity) and visual similarity. Proposed image retrieval methods are based on partial shape matching (PSM) that use landmarks along sagittal vertebral outlines that are consistent with those used by medical experts. Besides the two PSM methods that are improved algorithms of our previously developed methods, a new PSM method that is based on a simple but effective localized shape feature is proposed. The methods are evaluated and tested on a dataset of 856 segmented vertebrae and their performance is compared using precision-recall and average precision graphs. The best approaches are combined and integrated into our Web-based Spine Pathology & Image Retrieval System (SPIRS).*

## 1. Introduction

The National Health and Nutrition Examination Survey (NHANES) is a program conducted regularly by the National Center for Health Statistics (NCHS) with the aim of determining the prevalence of selected diseases and the associated risk factors. The National Library of Medicine (NLM) maintains the data collected by the second survey (NHANES II). This database contains 17,000 digitized cervical and lumbar spine x-ray images (sagittal view) as well as related text metadata: demographic information, anthropometric data, and health and medical history [1]. This large data collection is a valuable resource for the study and education of bone morphometry and musculoskeletal diseases. It is also a valuable resource for research and development in content-based image retrieval (CBIR), a technique that addresses the problem of indexing and retrieval of visual data through visual features instead of text.

For spine x-rays, the pathologies are often along the vertebral boundaries and searching this database on geometric characteristics of the vertebral bodies has been one of our research topics. Osteophytes have been found to be one of several key types of spine diseases that are prevalent in this collection. An osteophyte refers to a pointy growth on the vertebra and is a sign of bone degeneration in the spine. In this paper, we focus on the anterior osteophytes (AO), of which the bony protuberance occurs along the superior or inferior parts of the anterior vertebral boundary. Because of the localized property of AO pathology, shape-matching methods that use the whole vertebral shape of vertebra are often not effective and cannot differentiate the subtle differences between the critical "corner". We have been investigating partial shape matching (PSM) that focuses on specific segments of the vertebra boundary instead of the whole shape. The PSM has been implemented in our Web-based Spine Pathology & Image Retrieval System (SPIRS) [2] (http://archive.nlm.nih.gov/spirs), a system that provides both text-based and content-based capabilities to search images. For visual content retrieval, it allows users to select any of the segmented vertebral shape, identify the local boundary segment of interest, assign greater weights on the points at the boundary segment of interest and retrieve similar vertebrae based on this partial shape query using weighted Euclidean metric. The PSM algorithm implemented in SPIRS can be applied to any segment along the boundary and has a better performance than whole shape matching approaches for classifying the severity of osteophyte [3]. However, using local information under a more restricted setting could result in better matches.

In this paper, we present three PSM methods. Two of which are advances to our previously proposed methods. The other is a new PSM method that is based on a simple angle feature set that can effectively represent the local characteristics, such as the bending direction and turning angle, of the tip of the corner segments associated with AO pathology. The three PSM algorithms are quantitatively evaluated using a dataset of 896 vertebrae graded by a medical expert with one of three severity levels (slight, moderate, and severe) of AO. Given a query vertebra, we seek to retrieve vertebrae that have a similar shape on the specific corner that indicate the existence of AO and consequently have similar severity grade. Two of the three PSM algorithms whose retrieval performances are among the best in all the severity degree categories are further combined and tested on the ground truth dataset. These PSM methods are also integrated into SPIRS for visual examination of their retrieval performance.

In the following section, we describe properties of the segmented vertebral boundary and the 36 point landmark model that is used to represent the vertebral contour. We describe the PSM methods in Section 3. The evaluation results and discussion are given in Section 4. Section 5 discusses the integration of the new methods into SPIRS and a subjective evaluation of these methods.

## 2. Vertebra Representation

Segmentation of vertebrae in the spine x-ray images is an important step toward pathology-based vertebral retrieval. We have investigated multiple automatic and semi-automatic algorithms [4], and compared them to those manually marked by an expert. After the segmentation of the vertebral boundary, 9 landmark points on the boundary that mimic the 9-point model used by radiologists for marking relevant pathology are automatically extracted. Figure 1a illustrates the existence of AO using the 9-point model, in which points 2-3-8-7 and points 7-9-6-5 indicate the presence of superior AO and inferior AO respectively. As illustrated in Figure 1(a), points 1 and 4 specify the upper and lower posterior corner of the vertebra respectively; points 3 and 6 specify the upper and lower anterior corner of the vertebra respectively; points 2 and 5 indicate the median point along the upper and lower vertebra edge respectively; point 7 is the median point along the anterior vertical boundary of the vertebra; and points 8 and 9 mark the presence of the upper and lower anterior osteophytes. For normal vertebra, points 8 and 9 will overlap with points 3 and 6 and the corner angles on the vertebral body are approximately right angles. Since the 9-point model is often too sparse for representing vertebra for useful shape retrieval, we created a denser, 36-point model (Figure 1(b)) by 36 salient points including this 9 radiologist landmark points. More details on the landmark point localization can be found in [5]. In the 36-point representation, the points are numbered from 1 to 36 starting from the superior posterior corner of the vertebra and proceeding counterclockwise along the boundary. There are roughly three new boundary points between each pair of the 9 landmark points in the 36-point representation.

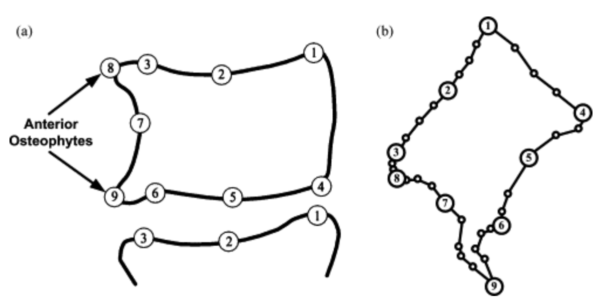



Figure 1. Vertebral shape description: (a) 9-point model (b) 36-point model

## 3. Partial Shape Matching Methods

Compared to other shape matching problems, the special challenges in this task are: 1) high shape similarity exhibited across vertebrae; and 2) subtle peculiarity that indicates AO pathology. In our previous research work, we explored several whole shape matching approaches, such as Fourier descriptors, polygonal approximation, geometric global shape properties, and invariant moments for spine shape retrieval. However, these whole shape matching methods are not effective in retrieving similar AO pathological spine shapes, as the retrieval results are adversely affected by the similarities of non-AO parts of the shape. To deal with the difficulty in effectively representing the subtle differences and capturing discriminative characteristics among the high shape similarity, we have been working on partial shape matching and PSM has been demonstrated to be a more effective method for vertebral pathology-based shape matching than whole shape matching methods. Our previous work on PSM [6] allows querying on specific intervals (such as one of the four corners) along the vertebral boundary shape and searches for the best matching intervals on other whole shapes. We describe this method for the current context in Section 3.2. The other PSM method we developed previously embeds the shape into a shape space to achieve fast indexing [3]. It can be applied to any user-specified segment along the boundary. In this paper, to further improve the retrieval accuracy, we restrict both the query and the similarity matching to a specific interval of interest, by taking advantage of the prior knowledge that AOs are only expressed along the two anterior "corners" of the vertebral outline (in the sagittal view) and by utilizing the vertebral landmark model described in Section 2. In addition to modifying previous works to limit the matching to a certain boundary interval, we also propose a new set of simple but effective features for AO similarity matching.

The first step is to specify the segment of interest. It should not be too long, or we will include some points that are not of interest for AO pathology; and should not be too short, or we may exclude some critical points that carry AO information. After inspection of a number of AOs in various individuals, we derived the heuristic method, for the 36-point model, of selecting the interval between point 17 and point 28 as the segment of interest for inferior AO characterization. Part of our rationale for this choice is the observation that points 17 and 28 approximate the locations of the median points along the anterior vertical boundary and the lower horizontal boundary of the vertebra (which are point 7 and 5 in the 9-point model), respectively. Similarly, we selected the interval between point 5 and point 16 as the segment of interest for superior AO. After the specification of the segment of interest, the following three PSM methods are developed and tested.

## 3.1 Method 1: PSM-PD

This method is based on the Procrustes distance (PD) [7]. Following steps are performed to find the best match between two shapes: 1) calculate the centroid of each shape; 2) scale each shape to have equal size; 3) translate the two shapes so that their centroids coincide; and 4) align the two shapes by rotation, are performed. The Procrustes distance is computed as the minimum sum of squared distances between the boundary points of the two shapes after the alignment and is illustrated by Equation 1, where $(\mathrm{x}, \mathrm{y})$ and $(\mathrm{x}', \mathrm{y}')$ are n boundary point coordinates of shapes of A and B, respectively, and shape A is translated by $(\mathrm{T_x}, \mathrm{T_y})$, scaled by S, and rotated by θ. The Procrutes matching process does a point-to-point match and requires shapes with one-to-one point correspondence. This fixed point matching suits our method in which the segment of the interest is fixed.

$$P = \sum_{i=1}^{n} \left\| \begin{bmatrix} S\cdot\cos\alpha & -\sin\alpha & T_x \\ \sin\alpha & S\cdot\cos\alpha & T_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}_A - \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix}_B \right\|^2 \quad (1)$$

## 3.2 Method 2: PSM-SF

In [6], we proposed a corner-guided PSM method that uses a modified dynamic programming (DP) technique to retrieve spine vertebrae. It uses a multiple open triangle shape representation as illustrated in Figure 2, which does not require equally distributed shape data points. Each open triangle can be measured by the average length of the two line segments and the angle between them. Let $\mathrm{l_q}$ and $\mathrm{l_t}$ ($\theta_\mathrm{q}$ and $\theta_\mathrm{t}$) denote the lengths (angles) of two open triangles, respectively, the length similarity is calculated as:

$$S_l(l_q, l_t) = \frac{4c_q c_t + (c_q{}^2-1)(c_t{}^2-1)}{(c_q{}^2+1)(c_t{}^2+1)} \quad (2)$$

where $c_q = l_q/l_q{}^0$ and $c_t = l_t/l_t{}^0$, $l_q{}^0$ and $l_t{}^0$ are the mean of the length of all the line segments on the two open triangles, respectively. The angle similarity is defined as:

$$S_\theta(\theta_q, \theta_t) = \cos(\theta_q, \theta_t) \quad (3)$$

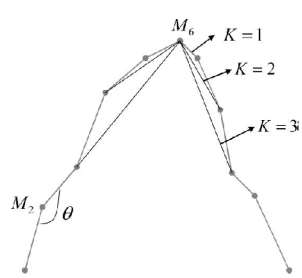



Figure 2. Multiple open triangles of a partial shape with 11 data points

For one partial shape, there might be several open triangles associated with each point, as shown in Figure 2. There is only one open triangle associated with $M_2$. There are possibly five open triangles associated with $M_6$. For each point, the overall length (angle) similarity is the average of all the individual length (angle) similarities of the open triangles associated with that point.

The shape matching was achieved using DP. To overcome computational overhead, it limits the possible search regions to four corners based on the rectangular appearance of vertebral shape and the 9-point landmark model. When matching it starts from a corner which is a point in the middle of the whole matching segment rather than from the first point of the matching segment. The capability of point merging was also implemented during the DP process. Therefore, the total cost used by DP consists of the cost contributed by the corresponding length and angle features as well as any merging cost. Please refer to [6] for additional information on this method. In this paper we limit this PSM method to the corresponding specified segment-of-interest of query and target vertebra.

## 3.3 Method 3: PSM-AD

Besides above two PSM methods, we also propose a new set of partial shape features which we call *angle difference* (AD) features. Different from the PSM-PD method that treats the entire segment of interest equally, it focuses on the corner head area on the segment of interest which is the most sensitive area with respect to AO shape similarity matching. The PSM-AD method uses a new set of angle based features that is different from the PSM-SF method which uses the overall length of the sides and the overall angle between them for the open triangles associated with each point as features. This new set of features are easy to compute and can implicitly take into account critical shape characteristics, such as the bending direction of the corner tip, the narrowing of the base, and the vertebral protuberance along the vertical or horizontal boundary near the corner head. In addition, the similarity matching is based on simple Euclidean distance for comparing the query feature vector with the feature vectors of the vertebrae in the database instead of using DP. The resulting retrieval process is much faster than that of the PSM-SF method even with this linear comparison approach (though efficient indexing and optimization can further increase the retrieval speed), which is a very attractive property for a large database.

The new set of features is extracted as follows. First, given the specified interval of interest, the algorithm automatically detects the "tip" point of the corner (the farthest protruding point on the specified segment), by finding the point whose distance to the center of mass of the vertebral polygon is the greatest among all the points in the segment of interest. This "tip" point ($\mathrm{P_0}$) together with the two end points of the segment of interest ($\mathrm{P_1}$ and $\mathrm{P_2}$) are three critical points and the angle difference features are calculated based on these three critical points as illustrated in Figure 3. Please note the numbers of points between these three critical points are usually different among the vertebrae. Denote the first three points along the segment between $P_0$ and $P_1$ (or $P_2$) that are the closest to $P_0$ to be $P_{11}$, $P_{12}$, and $P_{13}$ (or $P_{21}$, $P_{22}$, and $P_{23}$), respectively. For each pair of points ($P_{1i}$ and $P_{2i}$, $i = 1,2,3$ ), three angles are computed as follows:

$$\alpha_{0i} = \textit{ the angle between } \overline{P_0P_{1i}} \textit{ and } \overline{P_0P_{2i}}$$
$$\alpha_{1i} = s_{1i} * \textit{the angle between } \overline{P_0P_{1i}} \textit{ and } \overline{P_0P_1}$$
$$\alpha_{2i} = s_{2i} * \textit{the angle between } \overline{P_0P_{2i}} \textit{ and } \overline{P_0P_2}$$

Where $s_{1i}$ ($s_{2i}$) is +1 if $\overline{P_0P_{1i}}$ ($\overline{P_0P_{2i}}$ ) is on the left of $\overline{P_0P_1}$ ($\overline{P_0P_2}$) and is -1 if $\overline{P_0P_{1i}}$ ($\overline{P_0P_{2i}}$ ) is on the right of $\overline{P_0P_1}$ ($\overline{P_0P_2}$). The feature set (a vector of length 9) created by this scheme is straightforward. However, they are effective for representing the AO pathology and differentiating the subtle differences among them through the various angle measurements and their signs, as demonstrated by the experimental evaluations described in the next two sections.

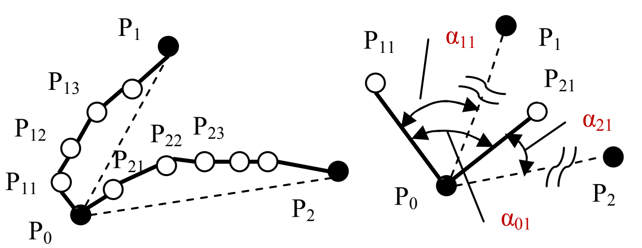



Figure 3. Partial shape features

# 4. Evaluation

First, we quantitatively evaluate the performance of the proposed three PSM methods using a "ground truth" dataset labeled by one medical expert. Based on the evaluation results, a combined method (PSM-COM) of the PSM-PD and PSM-AD methods was developed and tested. Since objective evaluation judges the retrieval performance from the aspect of classification accuracy not visual similarity, a subjective evaluation was also carried out after the integration of these PSM methods into SPIRS.

## 4.1 Ground truth dataset

All three PSM methods were evaluated using a dataset of 856 vertebrae segmented from 204 spinal x-ray images. Among the 856 vertebrae, 400 are cervical (C3-C7) vertebrae and 456 are lumbar (L1-L5) vertebrae. Each vertebra is represented using 36 boundary points. Both the inferior anterior "corner" and superior anterior "corner" of these vertebrae were labeled with AO severity levels by a medical expert. In this grading system, there are three severity degrees: slight (which includes normal), moderate, and severe, based on the extent of protuberance, the narrowing of the base, the degree of the angle and the presence of hook, key criteria identified by medical experts for rating the severity levels of the AO pathology. For inferior AO, among 856 vertebrae, 518 of them were graded as "slight", 234 of them were graded as "moderate", and 104 of them were graded as "severe"; for superior AO, 740 of them were "slight", 85, "moderate", and 32, "severe". Therefore, the superior AO data is much more unbalanced than the inferior AO data. Since this dataset was graded by only one medical expert whose bias and inconsistencies may affect the accuracy and reliability of the dataset to some degree (especially the borderline cases), caution must be taken in considering this set as a *gold standard*. However, given a reasonably large dataset as the inferior AO group, it can be considered a good test bed for evaluating our algorithms.

## 4.2 Evaluation measures

*Precision* and *recall* are two commonly-used measures for evaluation of retrieval performance. *Precision* is defined as the number of relevant images retrieved divided by the total number of retrieved images, and *recall* is defined as the number of relevant images retrieved divided by the total number of existing relevant images. We considered a retrieved image to be relevant if the corner of interest in the retrieved image had the same AO severity grade as the corresponding corner of interest in the query image. In our evaluation we used both, the precision-recall and the average-precision graphs, respectively. Since the numbers of vertebrae are not balanced across all grade categories, we generated both graphs for each grade class. For example, for the vertebrae labeled "slight", we used each one of them as the query vertebra and searched for similar vertebrae in the entire dataset. We then generated the precision-recall graph and the average-precision graph by averaging recall and precision, respectively, over all the query vertebrae that are graded "slight".

## 4.3 Results and discussion

The precision-recall and the average-precision graphs for each of the three grades generated using the inferior AO data are given in Figure 6. The results of the PSM-PD, PSM-SF, and PSM-AD method are shown in solid lines, dash-dotted line, and dashed lines, respectively. These results appear to satisfy a common criterion for many users: return relevant results in the top few of the returned images. The performance of all methods is better for "slight" grade than the "moderate" and "severe" grades. This may be due to the higher number of cases being "slight" in the dataset. While PSM-PD achieves the best performance for both "slight" grade and "moderate" grade, PSM-AD achieves the best performance for "severe" grade which is significant since this class is presumably of most interest. Compared to PSM-PD and PSM-AD, the PSM-SF does not perform well for both "moderate" and "severe" grades. Based on these observations, a weighted approach (PSM-COM) in which the similarity values are obtained by averaging the normalized similarity values calculated using

PSM-PD and PSM-AD methods. The retrieval performance of PSM-COM is shown with "++" line which is the second best in each of the severity groups. As PSM-PD treats the points along the segment of interest equally while PSM-AD emphasize the six points at the corner head, the combined method tries to balance the tradeoff between the significance of the corner tip area and that of the rest part of the segment-of-interest perceived by users when they judge the visual similarity exhibited among a variety of vertebrae partial shapes. Figure 4 shows the proportion of three grades of severity exhibited by the top 10 retrieved vertebra for each queried grade using the method PSM-COM. It demonstrates the confusion among the severity grades.

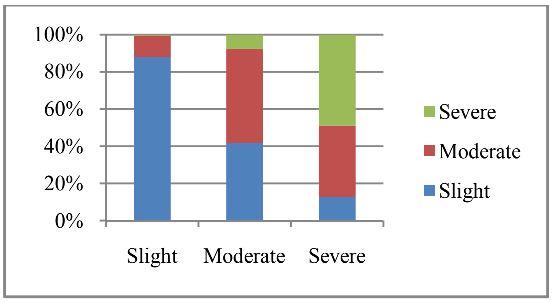



Figure 4. Proportion of three grades of severity exhibited by the top 10 retrieved vertebrae for each queried grade using the method PSM-COM

### 4.4 Subjective evaluation

In the objective evaluation described above, the similarity is defined as "being in the same severity category as the query" based on the assumption that the shapes in the same class are visually similar. To directly assess the system from the aspect of "being looking similar to the query", human perception has to be examined. To this end, a preliminary subjective test was carried out and three human subjects participated in the test. They are all engineers who have certain knowledge on content based image retrieval and two have worked on topic of shape matching. Each engineer was asked to randomly select 10 query vertebrae and evaluate each of the top 6 returned results on a subjective similarity scale from 1 (no or little similarity) to 5 (high similarity). The score for the first returned result was not counted since it is always the same as the query for this algorithm. The average score across the engineers were around 3.90 for all four methods which indicate the retrieval performance are good with respect to visual similarity for most of the examined cases.

## 5. System integration

We have included above PSM into SPIRS which has 4513 indexed vertebrae. In SPIRS, the query image can be selected by either random selection or by specifying patient image name. A query vertebral shape can then be selected from the loaded query x-ray image or from the retrieved vertebrae obtained by the previous query. The user can then select any part or parts of the vertebral boundary to generate partial shape query. Since the segment of interest that may exhibit significant pathology for inferior and superior AO has been pre-specified for the proposed PSM methods, the manual user selection of partial shape query is not necessary and the specified segment will be highlighted automatically in blue color when one of the proposed PSM methods is selected. The retrieval results of one example visual query using the PSM-COM method are given in Figure 5.

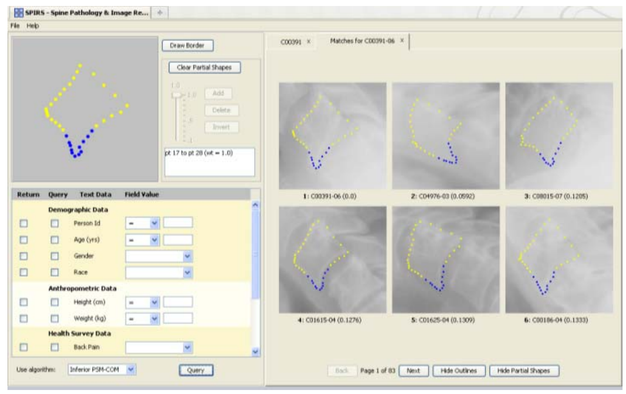


Figure 5. Retrieval results in SPIRS using PSM-COM

## 6. Conclusion

Searching vertebrae in a large collection of spine x-ray images that are relevant to pathology is potentially important for providing assistance to radiologists and bone morphometrists. As with other medical CBIR research, one prominent challenge we encountered for such searching task is the so called "semantic gap", which refers to the disparity between the human high-level perception of the image content and the low-level image representations provided by computer algorithms. We recognized that AO is one type of pathology that can be captured in the outline of a vertebra and can be reliably identified by the vertebral shape. However, the vertebrae in the dataset have high shape similarity and the dissimilarities that characterize the pathology are subtle. To narrow the discontinuity between a user's semantic understanding and the image feature representation, in this paper, we work on partial shape matching by limiting the search space to the specific area that AO exhibits so that the confusion added by other non-significant area will be alleviated. We developed three PSM methods and evaluated their performance both objectively and subjectively.

## Acknowledgement

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill

National Center for Biomedical Communications (LHNCBC).


## References

[1] Long, L. R., Antani, S., and Thoma, G.R. 2005. Image informatics at a national research center. *Computerized Medical Imaging and Graphics* 29, 2-3 (2005), 171-193. DOI= http://dx.doi.org/10.1016/j.compmedimag.2004.09.015

[2] Hsu, W., Antani, S., Long, L. R., Neve, L., and Thoma, G.R. 2009. SPIRS: a Web-based image retrieval system for large biomedical databases. *Int. J. Medical Informatics* 8, 1 (2009), S13-S24. DOI= http://dx.doi.org/10.1016/j.ijmedinf.2008.09.006

[3] Tagare, H. D., Qian, X., Fulbright, R. K., Long, R., and Antani, S. 2004. Shape based retrieval in NHANES II. In *Proceedings of the 12th Annual ACM international Conference on Multimedia* (New York, NY, USA, October 10 - 16, 2004). MULTIMEDIA '04. ACM, New York, NY, 776-779. DOI=http://doi.acm.org/10.1145/1027527.1027709

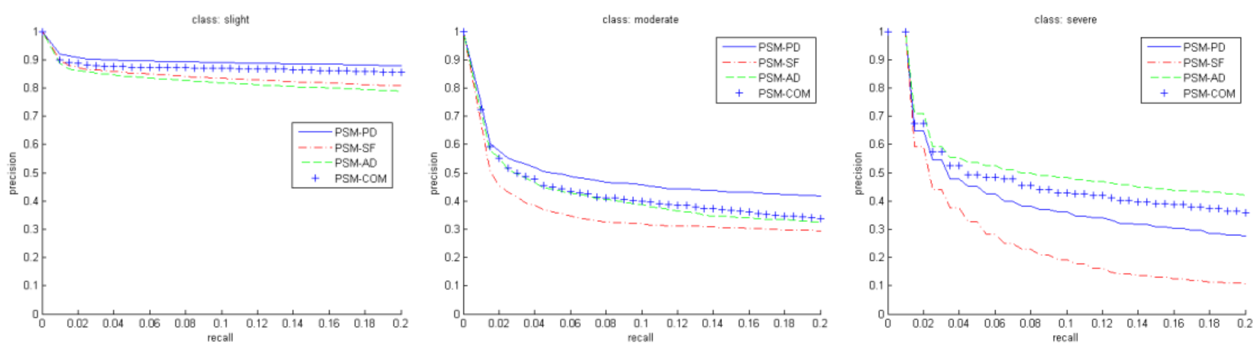
[4] Antani, S., Long, L.R., Thoma, G.R. 2002. A biomedical information system for combined content-based retrieval of spine x-ray images and associated text information. In *Proceedings of the Indian Conference on Computer Vision, Graphics, and Image Processing* (Ahmedabad, Gujarat, India, December 16-18, 2002). ICVGIP ’02. 242-247.

[5] Xu, X., Lee, D. J., Antani, S., and Long, L. R. 2003. Localizing contour points for indexing an x-ray image retrieval system. In *Proc. 16th IEEE Symp. Comput.-Based Med. Syst.* (New York, NY, USA, Jun. 26–27, 2003). CBMS’03. IEEE, New York, NY, 169–174. DOI= http://doi.ieeecomputersociety.org/10.1109/CBMS.2003.1212784

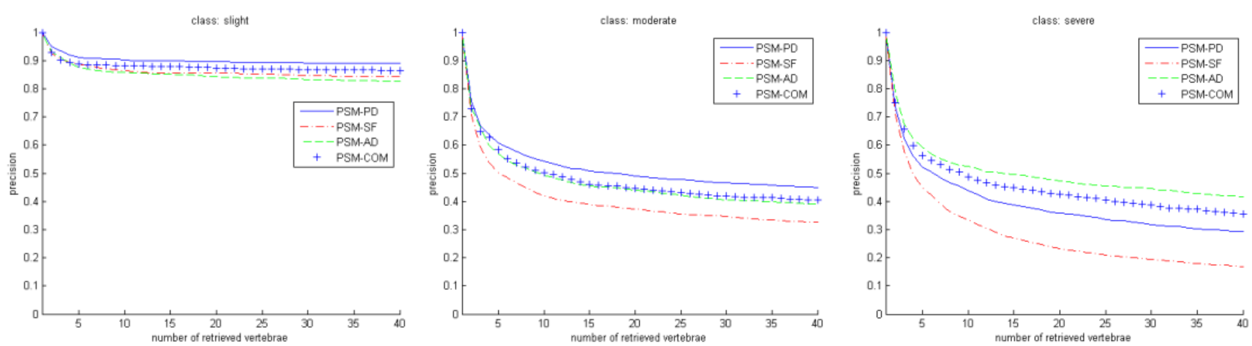
[6] Xu, X., Lee, D. J., Antani, S., and Long, L. R. 2008. A spine X-Ray image retrieval system using partial shape matching. *IEEE Transactions on Information Technology in Biomedicine* 12, 1 (2008), 100-108. DOI = http://dx.doi.org/10.1109/TITB.2007.904149

[7] Dryden I. L. and Mardia K. V. 1998. Statistical shape analysis. John Wiley & Sons, Inc.

[8] Bai, X., Yang, X., Latecki, L. J., Liu, W., and Tu, Z. 2010. Learning context-sensitive shape similarity by graph transduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 5 (2010), 861-874. DOI = http://doi.ieeecomputersociety.org/10.1109/TPAMI.2009.85


(a) Precision-recall graphs




(b) Average-precision graphs

Figure 6. Retrieval performance for inferior AO