On archiving and retrieval of sequential images from tomographic databases in PACS

Chi-Ren Shyu[a] , T. Tony Cai[b], and Lynn S. Broderick[c]

[a] School of Electrical and Computer Engineering, Purdue University

[b] Department of Statistics, Purdue University

[c] Department of Radiology, University of Wisconsin Hospital

ABSTRACT

In the picture archiving and communication systems (PACS) used in modern hospitals, the current practice is to retrieve images based on keyword search which returns a complete set of images from the same scan. Both diagnostically useful and negligible images in the image databases are retrieved and browsed by the physicians. In addition to the text-based search query method, queries based on image contents and image examples have been developed and integrated into existing PACS systems. Most of the content-based image retrieval (CBIR) systems for medical image databases are designed to retrieve images individually. But in a database of tomographic images, it is often diagnostically more useful to simultaneously retrieve multiple images that are closely related for various reasons, such as physiological contiguousness, etc. For example, high resolution computed tomography (HRCT) images are taken in a series of cross-sectional slices of human body. Typically, several slices are relevant for making a diagnosis, requiring a PACS system that can retrieve a contiguous sequence of slices. In this paper, we present an extension to our physician-in-the-loop CBIR system that allows our algorithms to automatically determine the number of adjoining images to retain after certain key images are identified by the physician. Only the key images, so identified by the physician, and the other adjoining images that *cohere* with the key images are kept on-line for fast retrieval; the rest of the images can be discarded if so desired. This results in large reduction in the amount of storage needed for fast retrieval.

Keywords: Medical image retrieval, HRCT image database, wavelets, transition detection.

1. INTRODUCTION

A high-resolution computerized tomographic (HRCT) scan of a patient with lung disease consists of between 40 to 50 cross-sectional images. Each of these images is usually a 512x512 matrix of pixels and each pixel is represented by a 2-byte word. Therefore, each image typically requires 0.5 MByte of storage, necessitating a total of 25 MBytes for a single patient scan. A typical facility in a research hospital may examine around 40 patients a day. That calls for 1 Gigabyte of storage per day. While the financial cost of this much storage may not be daunting any longer – since the prices of disks have dropped precipitously during the last few years – the huge size of the storage needed for, say, a year's worth (or, in some cases, several years' worth) of patient records can considerably slow down any attempts at automated retrieval.

In this paper we will present a scheme that reduces the needed storage considerably without reducing the retrieval effectiveness of a CBIR system such as ASSERT.[10] An additional advantage of the method presented here is that, for a given query image, in addition to retrieving the most similar physician-marked key images from the other patients, ASSERT can now also retrieve those adjacent images that *cohere* with the key images. With the method presented here, when a new patient scan becomes available for archiving, the physician chooses a small number, usually one or two, of key images as representing the entire scan. The wavelet based algorithm that we present here then selects from all of the patient images that *cohere* with the key images. Only the key images and the other images that cohere with the key images are retained for archival purposes. In contrast to some promising medical image database retrieval systems,[6–8] our approach no longer limits the CBIR system to extract image features from a single image and retrieve only individual images without including the adjacent images.

Further author information: E-mail addresses:
[a]chiren@ecn.purdue.edu, [b]tcai@stat.purdue.edu, and [c]lsbroderick@facstaff.wisc.edu

Part of the IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases VII
San Jose, California • January 1999
SPIE Vol. 3656 • 0277-786X/98/$10.00

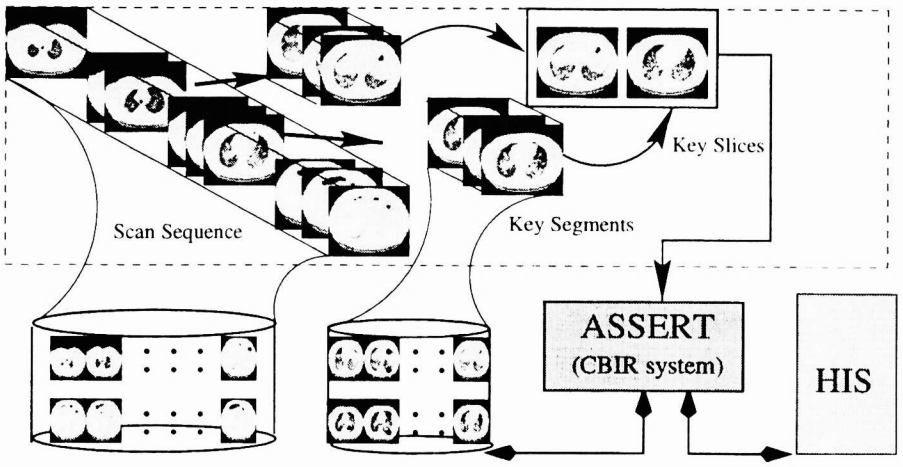

Figure 1. *The hierarchy of sequence-segment-slice structure and its integration of a content-based image retrieval system and a hospital information system.*

In this paper, in what follows we will first introduce the notion of a *key segment*, which is a sequence of contiguous images from a patient scan that *cohere* with the key images. The coherence will be established using two criteria, one sensitive to differences between the adjoining images in a scan and the other to the overall dissimilarity between a given image and a key image. In Section 3, experimental results are presented and we conclude in Section 4 with discussions and feature work.

2. KSDS: KEY SEGMENT DETECTION SCHEMA

The scanning process produces a sequence of contiguous cross-sectional medical images. Each image is called a *slice* in the sequence. Because an individual slice cannot capture the information from the third dimension, we would like to extract the *key segments* from the scanning process. A key segment contains a *key slice*, manually selected by a physician, and a certain number of contiguous slices adjacent to the key slice. Figure 1 shows the sequence-segment-slice structure. The *scan sequence* contains slices from the beginning to the end of the scan. As mentioned before, the number of slices in a scan is forty to fifty on average. A physician will typically identify one or two of these images as key images, these images are the most significant exemplars of the pathology. Since an automatic detection of key images would be an impossibly difficult task at this time, we let the physician make that determination. Using the algorithm described shortly, our system then automatically selects those images from the scan that best cohere with the key images. A key image together with the cohering images from the scan form a key segment. A key segment can be characterized by its *segment boundaries*, these being the farthest cohering images from a given key image in a scan. Note that only the key slices are given to the system for image database archiving and retrieval. During a query, when a key slice is retrieved as one of the best matching images, the physician is given the option of viewing the corresponding key segment.

KSDS is best explained with the help of the flow chart in Figure 3. The figure shows two phases: *magnitude extraction phase* and *location extraction phase*. A key step in KSDS is the detection of *key segment boundaries*. As explained in the next section, these boundaries are determined by finding two scanning transitions: *abrupt* and *gradual*. Patient scan images that are within the boundaries are said to cohere with the key images.

2.1. Important Information Extracted from Wavelet Coefficients

Wavelet bases offer a degree of localization both in space and in frequency and provide efficient representation for a wide range of images. The wavelet transform of an image provides a tool for time-frequency description of the image.

A one-dimensional orthonormal wavelet basis is generated from dyadic dilation and integer translation of two basic functions, a "father" wavelet ϕ and a "mother" wavelet ψ. The functions ϕ and ψ can be chosen to be compactly supported.

Define

$$\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k), \quad \psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k).$$

The collection $\{\phi_{J,k},\ k \in Z;\ \psi_{j,k},\ j \leq J, k \in Z\}$ is then an orthonormal basis of $L^2(R)$.

An orthonormal wavelet basis of $L^2(R^2)$ can be constructed from a one-dimensional orthonormal wavelet basis through tensor product. Define

$$\Phi_{j,m,n}(x,y) = \phi_{j,m}(x)\phi_{j,n}(y), \qquad \Psi^h_{j,m,n}(x,y) = \phi_{j,m}(x)\psi_{j,n}(y),$$
$$\Psi^d_{j,m,n}(x,y) = \psi_{j,m}(x)\psi_{j,n}(y), \qquad \Psi^v_{j,m,n}(x,y) = \psi_{j,m}(x)\phi_{j,n}(y).$$

Then the collection $\{\Phi_{J,m,n},\ m,n \in Z;\ \Psi^\sigma_{j,m,n},\ \sigma \in \{h,d,v\}, j \leq J, m,n \in Z\}$ constitutes an orthonormal basis of $L^2(R^2)$. A wavelet basis for $L^2([0,1]^2)$ is obtained in the same fashion with some minor modifications (see Daubechies[3] and Cohen, et al.[1]).

An orthonormal wavelet basis has an associated exact orthogonal discrete wavelet transform (DWT) that is norm-preserving and transforms an image into the wavelet coefficient domain in $O(n)$ steps. The interested reader is referred to Daubechies[2] and Strang[11] for further details about the wavelets and the discrete wavelet transform. Wavelet bases are well localized and the wavelet transform can compact the energy of an image into a small number of large wavelet coefficients and thus achieves data compression (see DeVore, et al.[5] and Meyer[9]).

The detail information in an image I can be extracted by applying a two-dimensional discrete wavelet transform with the scaling function Φ and a wavelet function Ψ. The wavelet transform yields coefficients that characterize the rate of gray level changes in an image at different resolutions and in three principal directions: diagonal, horizontal and vertical with different weightings * for each pixel. Using the index j to denote a resolution level, the relationship between the image and its wavelet coefficients is thus given by

$$I(x,y) \approx \sum_{m,n} s_{J,m,n}\Phi_{J,m,n}(x,y) + \sum_{\sigma\in\{v,d,h\}}\sum_{j=1}^{J}\sum_{m,n} d^\sigma_{j,m,n}\Psi^\sigma_{j,m,n}(x,y) \tag{1}$$

In this equation, $s_{J,m,n}$ are the coefficients representing the smooth part of the image at resolution level J and $d^\sigma_{j,m,n}$ are the coefficients representing details along three directions, vertical, horizontal, and diagonal at resolution level j. These coefficients can be naturally organized as a matrix. In this paper, we will use $M^h_{j,k}$, $M^d_{j,k}$, and $M^v_{j,k}$ to denote the three detail coefficient matrices at level j for slice k.

There are two important relationships between a coefficient matrix and its corresponding image: 1. the location of a coefficient in a matrix is related to the location of the corresponding pixel-block in an image. 2. the deeper the resolution level, the bigger the pixel-block. For example, shown in Figure 2, a coefficient in the first-level coefficient matrix "covers" an $8 - by - 8 - pixel$ block in the corresponding image if the applied wavelet has filter length 8. At level 2, a coefficient in the matrix "covers" a $16 - by - 16 - pixel$ block. Furthermore, it is well-known to use the leading coefficients to represent images in image compression. It has been proved that small coefficients are negligible in representing images.[4] Therefore, the space translation of the leading coefficients that show the significant magnitudes on gray-scale changes will capture the changes of detail information from two consecutive slices.

The wavelet-based information extraction process is depicted in Figure 3. In this figure, slice k is fed into KSDS. Its *J-level* two-dimensional discrete wavelet transform generates J coefficient matrices for each detail part. The scheme then sorts the coefficients for all coefficient matrices and picks up the N_c leading coefficients. All coefficients in the set of N_c leading coefficients are called *significant coefficients*. The leading N_c coefficients are then used to threshold the coefficient matrices. All coefficients less than the N_c^{th} leading coefficients are set to 0. The model then records the locations, (x,y) lists, for those non-zero coefficients in each decomposition level as shown in Figure 3. Since the positions of coefficients are *highly related* to the related positions of the pixels on the images, without applying inverse discrete wavelet transform, the system therefore will compute the distances related to the locations as well as the magnitude differences of leading coefficients between two slices.

Now we are ready to utilize both magnitude and location information to detect the segment transition.

*The QMF (quadrature-mirror-filter) corresponding to the wavelet determines different weights for pixels in computing the magnitude of gray-scale changes in detail parts of the image.

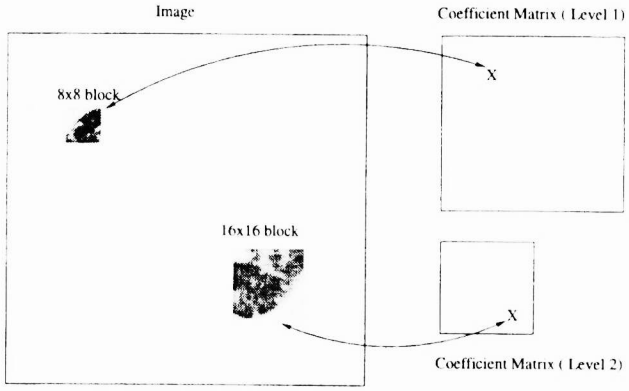

Figure 2. *The relationship between the coefficient matrices and the corresponding image when the applied wavelet has filter length 8.*

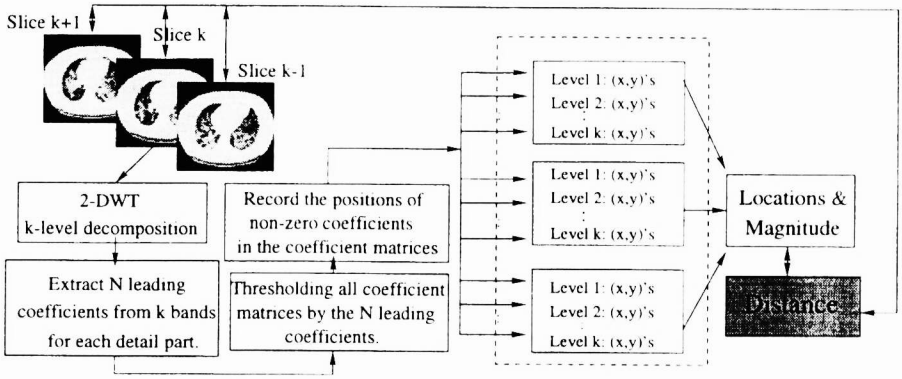

Figure 3. *The flow chart for distance measurements in KSDS.*

2.2. Segment Transition Detection

We refer to the process of human-assisted detection of transitions from one key segment to other insignificant segments as *segment transition detection*. The detection process can be categorized into two approaches: *abrupt transition* and *gradual transition*. The former can be detected by comparing the distance between two consecutive slices while the latter can be detected by accumulating the distances from one key slice to the slice in the segment boundary.

The concept of the distance measurement is illustrated in Figure 4. This figure shows an example of two-level decomposition for the horizontal details from slices k and $k-1$. The system starts from slice k and extracts the information on N_c leading coefficients. If there are $N_{j,k}$ leading coefficients in $M^h_{j,k}$, $3N_{j,k}$ leading coefficients are extracted from the corresponding coefficient matrix, $M^h_{j,k-1}$. For each leading coefficient in $M^h_{j,k}$, (x_i, y_i), the system picks up the nearest neighboring coefficient of (x_i, y_i) in $M^h_{j,k-1}$ among the $3N_{j,k}$ leading coefficients. The selected coefficient in $M^h_{j,k-1}$ has index $q(i)$ with the corresponding index i in slice k.

The distance between two consecutive slices is determined by the differences in location, (x_i, y_i) and $(x'_{q(i)}, y'_{q(i)})$ as well as the differences in magnitude, $A(i)$ and $A(q(i))$. We define the distance in the horizontal direction by the following equation:

$$D^h_k = \sum_{l=1}^{J} \sum_{i=1}^{N_{l,k}} 2^{l/2} \left(w_{location} \left(|x_i - x_{q(i)}|^p + |y_i - y_{q(i)}|^p \right) + w_{magnitude} \, |A(i) - A(p(i))|^p \right)^{\frac{1}{p}} \tag{2}$$

where $N_{l,k}$ is the number of significant coefficients at level l, $2^{l/2}$ is a normalization factor, and $w_{location}$, $w_{magnitude}$ are weights for location and magnitude distances, respectively. The overall distance is the summation of three distances contributed by three detail parts. In our experiment, we use both L_1 and L_2 distances.

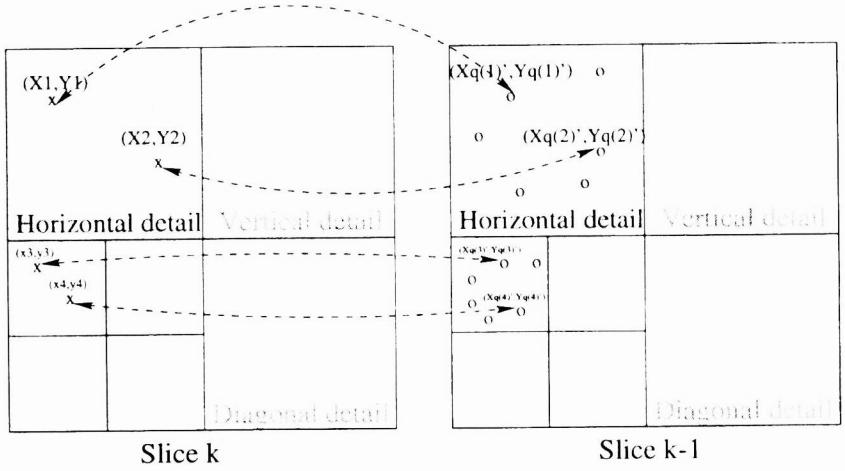

Figure 4. *The location-related distance measurement.*

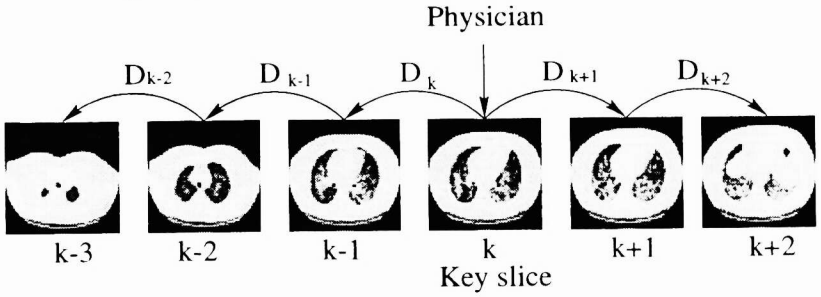

Figure 5. *Distances for transition detection.*

Therefore, the distances contributed by three detail parts can be obtained as follows:

$$D_K = w_h D_j^h + w_d D_j^d + w_v D_j^v \tag{3}$$

where h, d, and v stand for horizontal, diagonal, and vertical parts, respectively. w_h, w_d, and w_v are elements of the weighting vector.

The detection schema is established using two criteria, one sensitive to the differences between the adjoining images in a scan and the other sensitive to the overall dissimilarity between a given image and a key image. The former is called *abrupt transition detection* and the latter is *gradual transition detection*. Assuming the key slice has index K, the distance between slices K and $K-1$ is denoted by D_K as shown in Figure 5.

The *base distance* is defined as follows:

$$D_{base} = \frac{D_K + D_{K+1}}{2} \tag{4}$$

The distance computation will be operated in forward and backward directions until either of the following stop criteria is met.

$$D_j \geq (1 + r_1) D_{base} \tag{5}$$

$$\frac{1}{|m|} \sum_{j=0}^{m} D_{K+j} \geq (1 + r_2) D_{base} \tag{6}$$

In our experiment, we pick 0.3 for both r_1 and r_2 from the training results of current image database.

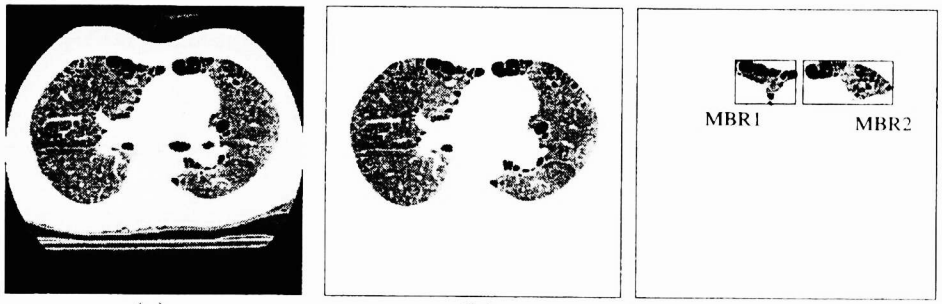

Figure 6. *Three sets of images: (a) the whole image (b) lung extracted image (c) PBR image.*

3. EXPERIMENTAL RESULTS

In the field of medical image database retrieval systems, such as ASSERT, the clinical useful information consists of gray level variations in highly localized regions of the image. As shown in Figure 6(a), the artifacts outside the lung regions will affect the accuracy of key segment detection because those pixels are not diagnostically meaningful. Therefore, we discard the background pixels and keep only the lung regions shown in Figure 6(b). However, the pathology bearing regions (PBRs) in medical images tend to be not well-defined and present great difficulties in any attempts at automatic segmentation. It is for this reason that we have had to implement a physician-in-the-loop system to ask the physician to delineate the PBRs. Shown in Figure 6(c) are two PBRs marked by a physician with paraseptal emphysema.[12] In this section, we present three sets of experiments designed to test the utility of KSDS. These three sets of images are as follows:

1. Whole image: image includes all pixels in a 512x512 matrix.
2. Lung extracted image: image includes only the pixels inside the lung region(s).
3. PBR image: image includes only the pixels inside the minimum bounding rectangles (MBRs) of PBRs.

In each experiment set, two measurements are computed: 1. the accuracy of key segment detection and 2. the storage saving achieved by keeping only the slices in the key segment. Our evaluation uses a testbed containing 3,516 HRCT lung images from 78 patients' scans, producing 117 key segments for each set.

Our goal is to find accurate segment boundaries for each key segment. To measure KSDS's accuracy we compare its boundaries to those marked by a physician. The physician first selects key slices that provide pathological information for the diagnosis by looking at a display of all scans. The physician then scans backwards to mark the left boundary P_L and then right to mark the right boundary, P_R, of the key segment surrounding the key slice. To compute the accuracy we compare the system's boundaries, S_L and S_R to the physicians' using the following formula:

$$Accuracy = \left(1 - \frac{|P_L - S_L| + |P_R - S_R|}{|P_R - P_L|}\right) \times 100\% \tag{7}$$

The purpose of storing key segments is to radically reduce the storage requirements over storing all slices from the scan, while still providing physicians with the required three-dimensional view. To evaluate how well our method achieves this goal we present the storage reduction of each filter length of Symlets [†]. The percentage of savings is defined as follows:

$$Storage = \left(1 - \frac{|S_R - S_L|}{Total\ number\ of\ slices\ in\ one\ scan}\right) \times 100\% \tag{8}$$

Figure 7(a) presents the accuracy results. Our accuracy metric does not tell us whether our method underestimates or overestimates the boundaries. We analyzed the results and found that approximately 80% of the times our method overestimates the boundaries, including more slices than the physician. This result is invariant of which filter length,

[†]In our experiments, we use Symlets to have consistency with human's perception system. To capture the performance differences in coarseness of image variation, we also compare the results with different filter lengths of Symlets.

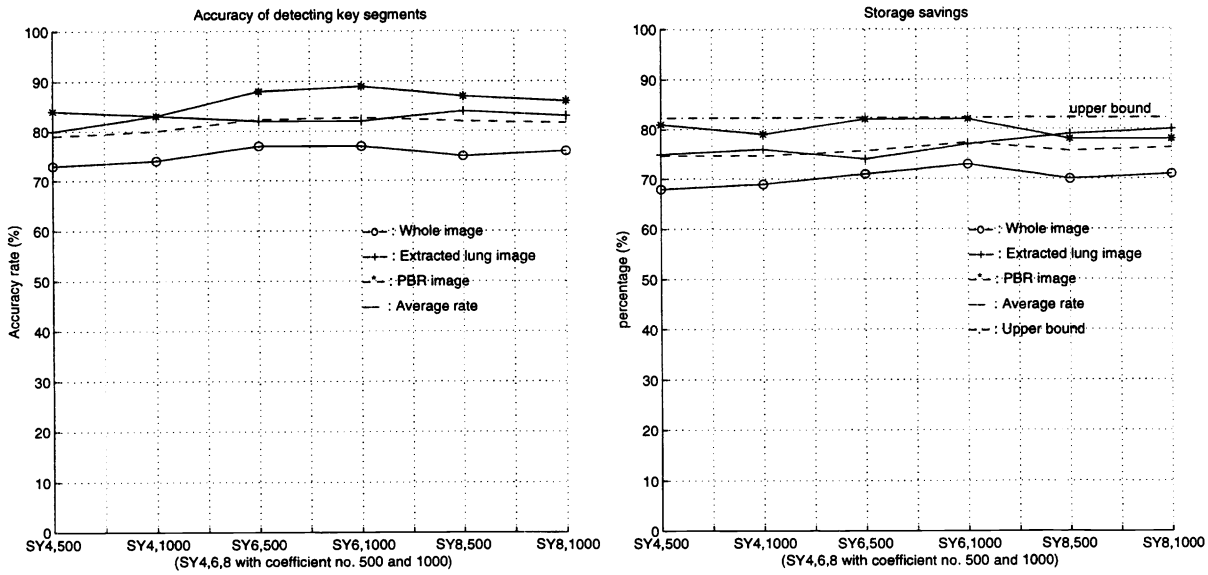

Figure 7. *(a) Accuracy rates for detecting key segments. (b) Storage saving. (In both figures, labels of x-axes stand for the filter length of Symlets 4, 6, 8 with leading coefficient number 500 and 1000.)*

how many coefficients used, and the parameter settings. From the chart, the average accuracy rates range from 77% to 82%. Most of times, the PBR image set gives better results than the other two image sets. The accuracy of whole image set is relatively lower than the other two sets. This also supports the empirical observation of our previous work[10] which presented localized approaches are superior to global approaches. Figure 7(b) shows the rate of storage saving. On average, the rate ranges from 74% to 76%. The best performance our KSDS can achieve is the *upper bound* depicted in the chart. The upper bound is equivalent to the storage saving when a physician manually selects all boundaries of key segments. In the testbed, there are 18 different lung diseases. Since the lung pathologies have so many visual patterns, the filter lengths of Symlets do not give much information for us. However, experiments with 1,000 coefficients perform slightly better than the experiments with 500 coefficients.

4. CONCLUSIONS AND FUTURE WORK

This paper has focussed on the wavelet-based transition detection methods for extracting key segments in a series of cross-sectional HRCT images. Our evaluation was performed in the context of ASSERT, our CBIR system for retrieving medical images. A discussion of retrieval and archival in ASSERT are presented in[10] . Our results obtained an average accuracy of 81%, typically off by only one or two slices from the boundaries selected by the physician. Because the key segments are used to present a three-dimensional view, the absence or addition of one or few slices has minimal impact on their ability to interpret the pathology information. In the immediate future we plan to apply learning algorithms to select the weights for location and magnitude differences in Equation 2 and the values of r-parameters for thresholding the distance measurement for transition detection in Equations 5 and 6. We are also exploring the possibility of incorporating KSDS in the attribute space of our CBIR system to improve the retrieval accuracy.

ACKNOWLEDGMENTS

This work is supported by NSF under grant number IRI9711535, NIH under grant number 1 R01 LM06543-01A1, and the Showalter Fund.

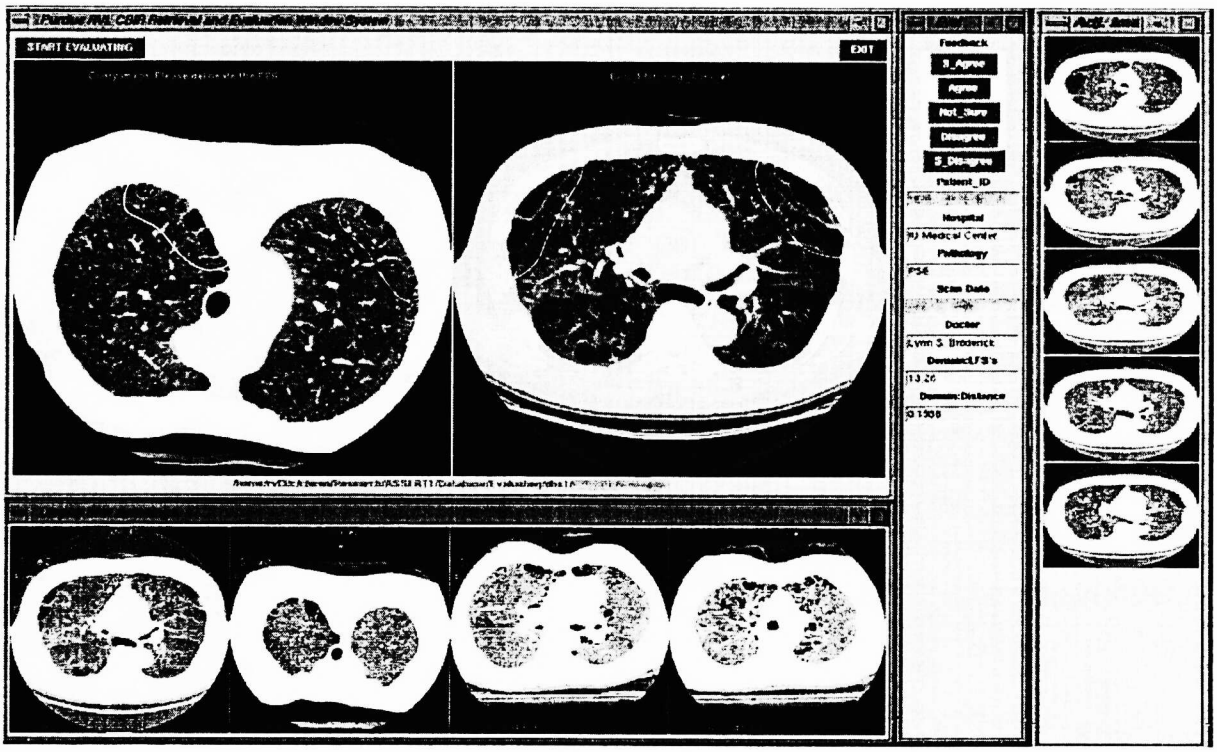

Figure 8. *The adjacent images of the best matched image are displayed in the right column of the main window. There are 5 slices in the key segment that coheres the retrieved image pathologically.*

REFERENCES

1. Cohen, A., Daubechies, I., Jawerth, B. & Vial, P., Multiresolution analysis, wavelets, and fast algorithms on an interval. *Comptes Rendus Acad. Sci. Paris* (A). 316., pp. 417-421, 1993.
2. Ingrid Daubechies *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
3. Daubechies, I. , Two recent results on wavelets: wavelet bases for the interval, and biorthogonal wavelets diagonalizing the derivative operator. In Schumaker L.L. & Webb G. (eds), *Recent Advances in Wavelet Analysis*. Academic Press, pp. 237 - 258, 1994.
4. R. A. DeVore, B. Jawerth, and B. Lucier, Image Compression Through Wavelet Transform Coding, *IEEE Trans. on Information Theory*, Vol. 38, No. 2, pp. 719-746, 1992.
5. DeVore, R. and Popov, V., Interpolation of Besov Spaces. *Trans. Amer. Math. Soc.*, 305, pp. 397-414, 1988.
6. C. C. Hsu, W. W. Chu, and R. K. Taira, A knowledge-based approach for retrieving images by content, *IEEE Trans. on Knowledge and Data Engineering*, Vol. 8, No. 4, pp. 522-532, 1996.
7. P. M. Kelly, T. M. Cannon, and D. R. Hush, Query by image example: the CANDID approach, in *SPIE Vol. 2420 Storage and Retrieval for Image and Video Databases III*, pp. 238-248, 1995.
8. F. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, Z. Protopapas, Fast Nearest Neighbor Search in Medical Image Databases, in *Proc. of the Conference on Very Large Data Bases (VLDB)*, pp. 215-226, 1996.
9. Meyer, Y. , *Wavelets and Operators*, Cambridge University Press, Cambridge, 1992.
10. C. R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. Aisen and L. Broderick, Local versus Global Features for Content-Based Image Retrieval, in *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries*, pp. 30-34, Santa Barbara, CA, June 1998.
11. Strang, G., Wavelet And Dilation Equations: A Brief Introduction. *SIAM Review*, 31(4), pp. 614 - 627, 1992.
12. W. R. Webb, N. L. Muller, and D. P. Naidich, *High-Resolution CT of The Lung*, second edition, Lippincott-Raven, Philadelphia, 1996.