Local Radon Descriptors for Image Search

Morteza Babaie¹, H.R. Tizhoosh² Amin Khatami³ and M.E. Shiri¹

Abstract-Radon transform and its inverse operation are important techniques in medical imaging tasks. Recently, there has been renewed interest in Radon transform for applications such as content-based medical image retrieval. However, all studies so far have used Radon transform as a global or quasiglobal image descriptor by extracting projections of the whole image or large sub-images. This paper attempts to show that the dense sampling to generate the histogram of local Radon projections has a much higher discrimination capability than the global one. In this paper, we introduce Local Radon Descriptor (LRD) and apply it to the IRMA dataset, which contains 14.410 xray images as well as to the INRIA Holidays dataset with 1,990 images. Our results show significant improvement in retrieval performance by using LRD versus its global version. We also demonstrate that LRD can deliver results comparable to wellestablished descriptors like LBP and HOG.

Keywords—Image retrieval; image descriptor; Radon projections; local projections; IRMA dataset;

I. MOTIVATION

Content-Based Image Retrieval (CBIR) refers to a class of algorithms that use the image content (pixels or more commonly their features) to search for similar images in an archive when a query (input) image is given by the user. Numerous CBIR methods have been introduced that are based on feature descriptors [1]–[3]. Generally, we can distinguish between feature descriptors that are either based on "dense sampling" (features calculated from many small image windows with overlap), or feature descriptors that are based on "keypoint detection" (features calculated from small windows around selected points). One may add a third type based on global image features (FFT spectral features, Radon projections etc.) [4], [5]. The former two types of descriptors, however, have been more widely used and reported in literature.

Local binary patterns (LBP) and the histogram of gradients (HOG) and their extensions are representative for successful methods, which extract local information from images by means of dense sampling. Feature extraction methods such as SIFT and SURF have been commonly used as well. They are based on keypoint detection scheme which may fail to detect reliable keypoints in medical images [6], [7]. Dense sampling methods, in contrast, deliver reportedly higher performance [6].

The LBP generally captures the difference between each pixel and its neighbouring pixels (i.e., 8 pixels) as a binary vector which is then converted to decimal integers that are counted to construct a histogram [8]. There are numerous

extensions of the LBP by incorporating a different number of neighbors [9] and different number of histogram bins [10]. Uniform LBP reduces the number of bins from 256 to 59 by a selective grouping of patterns. LBP has been successfully applied in various applications such as face detection [11] and object detection [12]. It also has been used for medical image retrieval in several reports [13]–[15].

The HOG descriptor, on the other hand, captures the gradient information in densely sampled local windows by counting the number of discrete directions in n bins. It was initially introduced for pedestrian detection [16]. However, its power of discrimination made it applicable to a variety of fields in computer vision such as face recognition and image retrieval [17], [18].

II. RADON TRANSFROM

Radon transform integrates projection values from various directions in order to reconstruct an object (or the human body) from its parallel projections [19]. The inverse Radon transform along with filtered backprojection has been used to create body images from projections captured at different directions [20]. Given the spatial intensities f(x,y), the Radon transform can be given as

$$\mathbf{R}(\rho,\theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\delta(\rho - x\cos\theta - y\sin\theta)dxdy, \quad (1)$$

where $\delta(\cdot)$ is the Dirac delta function. In Figure 1, the Radon transform is illustrated for a small window. The Radon transform of a sample image (its sinogram containing projections in 180 directions) is depicted as well.

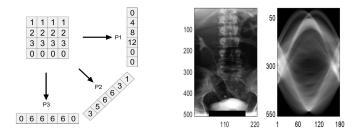


Fig. 1. Left: Three Radon projections of a 4×4 sample window at $0^\circ,45^\circ$ and 90° with zero padding. Right: a sample x-ray image and its sinogram (180 Radon projections).

Department of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran
² KIMIA Lab, University of Waterloo, ON, CANADA

³ Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Australia

Several works have recently developed approaches to medical image retrieval using "Radon barcodes" in which they binarize a certain number of equidistant projections [5], [21], [22]. We find the idea of using *projections* quite intriguing, specially for the medical field. However, the literature only contains projections-based descriptors that use "global" information (Radon projections of the entire image). Of course, the performance of such methods, restricted to global information, is expectedly rather limited.

There have been several works that use the term "local Radon transform" in connection with grayscale images (and not binary ones). However, they generally mean "localized" application of Radon transform on a certain part of the image or a sub-image [23], [24].

In this paper, we introduce a new dense-sampling descriptor based on Radon transform named *Local Radon Descriptor* (LRD). We describe the details of LRD in the next section. We use two image datasets to validate the proposed approach against LBP and HOG although our primary agenda is to improve methods which have used "global" Radon projections to assemble compact descriptors. LRD is supposed to break through the limitations of globally captured projections by examining gradient change of locally extracted projections.

III. LOCAL RADON DESCRIPTOR (LRD)

As mentioned before, several approaches have been proposed to use Radon projections and/or their binary versions for retrieval of medical images. All these methods so far only employ "global" projections; they apply Radon transform on the entire image. However, we know that a major reason for the success of LBP and HOG is their capability of capturing "local" information in form of a histogram. Combining this knowledge with what Radon projections can quantify, it is obvious that one has to design a new class of projection-based descriptors that can exploit local gradient information.

To construct the LRD, we go through the following steps:

Step 1. For every selected image block (local window), n equidistant Radon projections are calculated. We specify each element in each projection by $\mathbf{R}(\rho_i, \theta_i)$, where

$$\theta_j = \left\{0, \frac{\pi}{n}, \frac{2\pi}{n}, \dots, \frac{n\pi}{n} - 1\right\},\tag{2}$$

and ρ_i refers to the *i*-th line of the projection under the *j*-th direction.

Step 2. For all projections, we build the first order derivative vector \mathbf{d}_{θ_i} to quantify projection change:

$$\mathbf{d}_{\theta_{j}}(i) = \frac{\partial}{\partial \rho} \mathbf{R}(\rho, \theta_{j}) = \mathbf{R}(\rho_{i+1}, \theta_{j}) - \mathbf{R}(\rho_{i}, \theta_{j}). \quad (3)$$

Step 3. Looking at each projection $\mathbf{p}_i = \mathbf{R}(\rho_{k=1:L}, \theta_j)$ of length L, we build n/2 pairs of projections $(\mathbf{p}_{\theta_i}, \mathbf{p}_{\theta_j})$ whereas $\theta_j = \theta_i + \alpha$.

Step 4. Selecting two derivative vectors \mathbf{d}_{θ} and $\mathbf{d}_{\theta+\alpha}$, a ratio vector \mathbf{r} is created by dividing the elements of both gradient vectors resulting in a vector in $(-\infty, +\infty)$:

$$\mathbf{r}(i) = \frac{\mathbf{d}_{\theta}(i)}{\mathbf{d}_{\theta+\alpha}(i)}.$$
 (4)

To quantize each element of \mathbf{r} and count its occurrence, we apply the *arctan* operator to map the \mathbf{r} elements from $(-\infty, +\infty)$ to $[-\pi, +\pi]$. Then, we can discretize the outcome of the $\tan^{-1}(\mathbf{r})$ in b different bins resulting in a vector \mathbf{r}_q .

Step 5. We also build a summation vector s by adding the absolute values of the derivatives:

$$\mathbf{s}(i) = |\mathbf{d}_{\theta}(i) + \mathbf{d}_{\theta+\alpha}(i)| \tag{5}$$

Step 6. We capture the information of two Radon signals in the k-th local histogram $\mathbf{h}_k(i)$ by counting the quantized \mathbf{r}_q elements and weighting them by corresponding s elements. After generating all histograms, the main descriptor is created by concatenating individual histograms \mathbf{h}_k . The length of the descriptor is proportional to the number of blocks per column n_{column} , the number of blocks per row n_{rows} , and the number of bins b per histogram in the blocks: length= $n_{\text{column}} \times n_{\text{rows}} \times b$.

Figure 2 shows the steps for creating a histogram \mathbf{h}_k for a sub-image.

IV. EXPERIMENTS

We conducted two different experiments to evaluate the LRD versus the global Radon descriptors. The first experiment was conducted on the IRMA dataset, and the second one uses the INRIA Holidays dataset. In the following, we briefly describe these datasets and then the results of LRD for retrieving images from these datasets will be reported. The Matlab code for LRD is available on our website¹.

A. Datasets

IRMA dataset - The Image Retrieval in Medical Applications (IRMA) dataset, created by the Aachen University of Technology, has been used to evaluate retrieval methods for medical CBIR tasks [25]. It contains 12,677 images in the training set and 1,733 images in the testing set. The IRMA dataset is extremely challenging mainly due to its highly imbalanced distribution of 193 different classes. As well, digitized landmarks and annotations make the retrieval even more difficult. The IRMA x-ray images come with a unified code to calculate the retrieval error. As shown in Figure 3, all images in the dataset have been assigned a 13-digit *IRMA* code. The error per query image is a number between 0 and 1. As a result, the total IRMA error for all test images is a number between 0 and 1,733. We used the Python code provided by ImageCLEFmed09 to compute the errors². The dataset can be downloaded from IRMA website³.

¹ http://kimia.uwaterloo.ca/

²http://www.imageclef.org/2008/medical

³https://ganymed.imib.rwth-aachen.de/irma

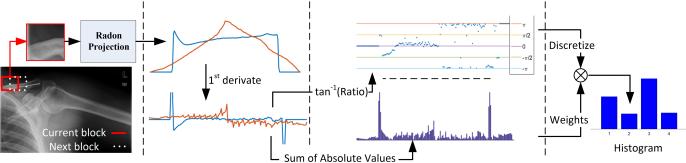


Fig. 2. General steps to extract LRD from gradients of pairs of Radon projections. The arctan of the ratio and the sum of absolute values of paired projections are used to assemble a histogram for each sub-image.

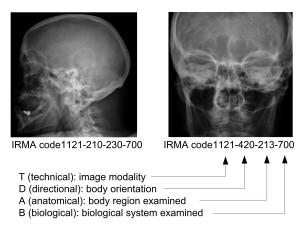


Fig. 3. Two sample x-ray images and their corresponding IRMA codes.

INRIA Holidays dataset – Jegou et al. [26] have created a searching dataset containing 1,991 high resolution images in 500 different categories such as landscapes, buildings and historical sites from different point of views and with various sizes (Figure 4). The first image in each category is supposed to be used as a query image. The performance of retrieving method is evaluated via the percentage of true retrieval against all 500 query images. The dataset can be downloaded from INRIA website⁴.

B. Parameter Setting

There are several parameters which influence the performance of LRD and hence must be investigated. Perhaps two important factors to be adjusted are the number of blocks per row, $n_{\rm rows}$, and per column, $n_{\rm columns}$. Note that these two parameters are needed to investigate in LBP and HoG as well. Additionally, the overlap percentage O_p between adjacent blocks is certainly also a factor that may affect the retrieval performance. And finally, we have to also specify how many projection pairs should be selected, and which projections should be paired to achieve the best results.

Through zero-padding, all images are resized to 256×256 (hence, $n_{\rm columns} = n_{\rm rows}$). We tested 2 different ways to adjust pair selection among n=18 projections:

1) Orthogonal selection: With n=18, we have $\theta=\{0^\circ,10^\circ,20^\circ,\ldots,170^\circ\}$. Orthogonal angles are then selected as a pair such that we have a set P of 9 paired projections:

$$P = \{(0^{\circ}, 90^{\circ}), (10^{\circ}, 100^{\circ}), \dots, (80^{\circ}, 170^{\circ})\}.$$
 (6)

2) Using characteristic projections: Based on [27], $\theta = 0^{\circ}$ and 90° are the most characteristic projection directions to retrieve images from the IRMA dataset (x-ray images generally do not exhibit rotations, and these directions provide very distinct projections). In order to exploit this observation, we paired each projection under $\pi/2$ by zero degree and all other projections by 90 degree such that we get a set $P = P_{0^{\circ}} \cup P_{90^{\circ}}$ with

$$P_{0^{\circ}} = \{(0^{\circ}, 10^{\circ}), (0^{\circ}, 20^{\circ}), \dots, (0^{\circ}, 80^{\circ})\},$$
 (7)

and

$$P_{90^{\circ}} = \{ (90^{\circ}, 100^{\circ}), (90^{\circ}, 110^{\circ}), \dots, (90^{\circ}, 170^{\circ}) \}.$$
(8)

The number of blocks and bins were tuned by exhaustive search. The best number of blocks for INRIA Holidays dataset was set to $n_{\rm columns}=n_{\rm rows}=3$ (9 sub-images) and the best number of bins was determined to be b=22. For IRMA dataset, these number were set to $n_{\rm columns}=n_{\rm rows}=5$ (25 sub-images), and b=12, respectively. The orthogonal paring did not exhibit good results such that we conducted all experiments with the characteristic projection paring P_{0° and P_{90° .

Fig. 5 shows sample histograms for LRD in comparison with LBP and HOG. Table 1 lists the histogram length (= $n_{\text{columns}} \times n_{\text{rows}} \times b$) for all descriptors for the two datasets.

C. Results

For sake of comparison, we used well-established descriptors like HOG and LBP to retrieve images as well. We did

⁴http://lear.inrialpes.fr/ jegou/data.php



Fig. 4. Sample images from the Holidays dataset.

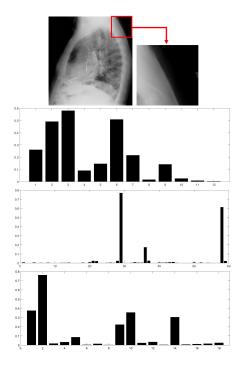


Fig. 5. Top: A sample x-ray and one of its sub-images. Bottom (left to right): LRD, LBP and HOG histograms of the sub-image.

Table 1. Sub-image and bin settings for LBP, HOG and LRD $(n_{\rm columns} \times n_{\rm rows} \times b).$

	IRMA dataset	Holidays dataset
LBP	$12 \times 12 \times 59$	$5 \times 5 \times 33$
HOG	$14 \times 14 \times 18$	$3 \times 3 \times 24$
LRD	$5 \times 5 \times 12$	$3 \times 3 \times 22$

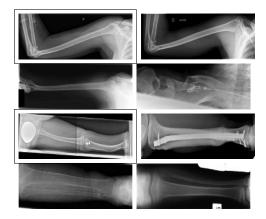


Fig. 6. Sample retrieval results for LRD (left image is the query image). The top three results are shown. Top: The first two results are correct (IRMA errors: 0, 0, and 0.58), bottom: the first results is wrong (IRMA errors: 0.24, 0.45 and 0.45).

the same parameter tuning steps for number of blocks and the number of bins for HOG and the cell size in LBP. The results of applying all three methods on the IRMA dataset is compared with the best global Radon (GR) results as reported in literature [27]. The maximum error is 1733 (if the retrieval for all test images delivers wrong macthes). Hence, beside IRMA error, we can also report the accuracy (= $1 - \frac{\text{error}}{1733}$).

Fig. 6 shows sample retrieval results for LRD. We display the top three matches for query images whereas for error calculation we only consider the first match.

Table 2, shows the comparison between these four retrieving methods (LRD, global Radon projections, HOG and LBP) as well as their descriptor lengths. The retrieval times via k-

NN (with k=1) are provided as well. LRD performance is slightly below the LBP and close to HOG. However, we can observe significant improvement in local implementation of Radon projections versus global methods. The IRMA error is considerably reduced (accuracy increase of more than 10%) and the LRD descriptor length is 4 times shorter than its global counterpart.

Although LBP and HOG are more accurate for the IRMA images, their descriptor lengths are 28 and 12 times, respectively, longer than LRD's, a fact that is also reflected in computation times. Another interesting point is, for the INRIA Holidays dataset, that LRD results are the same as HOG's. Also for this dataset, LRD delivers a shorter descriptor.

Table 2. Retrieval results (GR stands for global Radon). LRD considerably improves the performance of projection-based descriptors as proposed via global methods. As well, LRD does in fact deliver comparable results with state-of-the-art descriptors like LBP and HOG.

Method -	IRMA Dataset			Holidays Dataset		
	Error (Accuracy)	Length	Time	True	Length	Time
LBP	244.81 (85.87%)	8496	0.50s	48.00%	825	0.03s
HOG	266.36 (84.46%)	3528	0.23s	40.00%	216	0.0s
LRD	287.77 (83.39%)	300	0.02s	40.02%	198	0.0s
GR	472.31 (72.74%)	1200	0.11s	06.00%	2400	0.1s

V. SUMMARY

Radon projections can be used to construct descriptors for image retrieval. However, global projections (projecting the entire image) may not deliver expressive features suitable for image retrieval. In this paper, we introduced Local Radon Descriptor (LRD) as a new descriptor for image retrieval. It is designed to capture Radon projections in local neighbourhoods by looking at the ratio and sum of derivatives of paired projections. LRD exhibits significant advantage over reported works on projection-based approaches to image retrieval that are globally implemented. As well, LRD is a very compact descriptor. Experimental results on two datasets demonstrated the potentials of LRD for both medical and non-medical images.

REFERENCES

- [1] L. Nanni, A. Lumini, and S. Brahnam, "Survey on {LBP} based texture descriptors for image classification," *Expert Systems with Applications*,
- [2] G.-H. Liu, L. Zhang, Y.-K. Hou, Z.-Y. Li, and J.-Y. Yang, "Image retrieval based on multi-texton histogram," *Pattern Recognition*, vol. 43, no. 7, pp. 2380 – 2389, 2010.
- [3] A. Kumar, J. Kim, W. Cai, M. Fulham, and D. Feng, "Content-based medical image retrieval: A survey of applications to multidimensional and multimodality data," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1025–1039, 2013.
- [4] S. Krig, Computer vision metrics. Springer, 2014.
- [5] H. R. Tizhoosh, "Barcode annotations for medical image retrieval: A preliminary investigation," in 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 818–822.
- [6] M. Kashif, T. Deserno, D. Haak, and S. Jonas, "Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment," *Computers in Biology and Medicine*, vol. 68, pp. 67–75, 2016.
- [7] D. Sargent, C. Chen, C.-M. Tsai, Y.-F. Wang, and D. Koppel, "Feature Detector and Descriptor for Medical Images," *Proceedings of the SPIE International Symposium on Medical Imaging* 2009, vol. 7258, pp. 1–8, 2009.
- [8] M. Pietikainen and T. Ojala, "Nonparametric texture analysis with complementary spatial operators," Series in Machine Perception and Artificial Intelligence, Volume 40: Texture Analysis in Machine Vision, pp. 3–18, 2000.
- [9] X. Shi, A. Zhu, J. Burt, W. Choi, R. Wang, T. Pei, B. Li, and C. Qin, "An experiment using a circular neighborhood to calculate slope gradient from a DEM," *Photogrammetric Engineering & Remote Sensing*, vol. 73, no. 2, pp. 143–154, 2007.
- [10] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [11] T. Ahonen and M. Hadid, A.and Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.

- vol. 39, no. 3, pp. 3634 3641, 2012.
- [12] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE transactions on* pattern analysis and machine intelligence, vol. 28, no. 4, pp. 657–662, 2006.
- [13] C. Yao and S. Chen, "Retrieval of translated, rotated and scaled color textures," *Pattern Recognition*, vol. 36, no. 4, pp. 913–929, 2002.
- [14] S. Murala, R. P. Maheshwari, . R. Balasubramanian, S. Murala, and R. Balasubramanian, "Directional local extrema patterns: a new descriptor for content based image retrieval," *Int J Multimed Info Retr*, vol. 1, no. March, pp. 191–203, 2012.
- [15] P. Srivastava and A. Khare, "Integration of wavelet transform, Local Binary Patterns and moments for content-based image retrieval," *Journal* of Visual Communication and Image Representation, vol. 42, pp. 78– 103, 2017.
- [16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. I, pp. 886– 893 2005
- [17] R. Hu and J. Collornosse, "A performance evaluation of gradient field HOG descriptor for sketch based image retrieval," *Computer Vision and Image Understanding*, vol. 117, no. 7, pp. 790–806, 2013.
- [18] O. Déniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," *Pattern Recognition Letters*, vol. 32, no. 12, pp. 1598–1603, 2011.
- [19] J. Radon, "On determination of functions by their integral values along certain multiplicities," Ber. der Sachische Akademie der Wissenschaften Leipzig, (Germany), vol. 69, pp. 262–277, 1917.
- [20] G. Beylkin, "Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform," *Journal of Mathematical Physics*, vol. 26, no. 1985, p. 99, 1985.
- [21] H. R. Tizhoosh, S. Zhu, H. Lo, V. Chaudhari, and T. Mehdi, *MinMax Radon Barcodes for Medical Image Retrieval*. Springer International Publishing, 2016, pp. 617–627.
- [22] S. Zhu and H. Tizhoosh, "Radon features and barcodes for medical image retrieval via svm," in 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 5065–5071.
- [23] M. Krause, J.-M. Hausherr, and W. Krenkel, "(micro)-crack detection using local radon transform," *Materials Science and Engineering: A*, vol. 527, no. 26, pp. 7126–7131, 2010.
- [24] R. Pourreza, T. Banaee, H. Pourreza, and R. D. Kakhki, "A radon transform based approach for extraction of blood vessels in conjunctival images," in *Mexican International Conference on Artificial Intelligence*. Springer, 2008, pp. 948–956.
- [25] T. Tommasi, B. Caputo, P. Welter, M. Güld, and T. Deserno, "Overview of the CLEF 2009 medical image annotation track," CEUR Workshop Proceedings, vol. 1175, 2009.
- [26] H. Jegou, M. Douze, and C. Schmid, "Hamming Embedding and Weak Geometry Consistency for Large Scale Image Search â Extended version," *Eccv*, no. October, pp. 304–317, 2008.
- [27] M. Babaie, H. R. Tizhoosh, and M. E. Shiri, "Retrieving Similar Cases from Big Image Data Using Radon Barcodes with Single Projections," *International Conference on Pattern Recognition Applications and Methods*, 2017.