Rotation Invariant Fuzzy Shape Contexts Based on Eigenshapes and Fourier Transforms for Efficient Radiological Image Retrieval

Alaidine Ben Ayed[1], Mustapha Kardouchi[1], and Sid-Ahmed Selouani[2]

[1] Université de Moncton, Campus de Moncton, 18 avenue Antonine-Maillet, Moncton, NB, Canada E1A 3E9
alaidine.ben.ayed@umoncton.ca
[2] Université de Moncton, Campus de Shippagan, 218 boulevard J-D Gauthier, Shippagan, NB, Canada E8S 1P6

Abstract. This paper proposes a new descriptor for radiological image retrieval. The proposed approach is based on fuzzy shape contexts, Fourier transforms and Eigenshapes. First, fuzzy shape context histograms are computed. Then, a 2D FFT is performed on each 2D histogram to achieve rotation invariance. Finally, histograms are projected onto a lower dimensionality feature space whose basis is formed by a set of vectors called Eigenshapes. They highlight the most important variations between shapes. The proposed approach is translation, scale and rotation invariant. Classes of the medical IRMA database are used for experiments. Comparison with the known approach rotation invariant shape contexts based on feature-space Fourier transformation proves that the proposed method is faster, more efficient, and robust to local deformations.

Keywords: Image retrieval, Fuzzy Shape Contexts, Fourier transform, Eigenshapes, Radiological images.

1 Introduction

One of the most vivid fields of computer vision research is medical image processing [1] [2]. Medical image tools are used by physicians for diagnosis. So many works proposed new techniques of medical image processing [3] [4] [5] [6].

Medical image retrieval is a branch of medical image processing. The concept of content based image retrieval is used in many applications such as breast cancer diagnosis systems [7] [8] [9]. Each image in the database needs to be described by features providing its signature. Features extraction is based on visual characteristics. The best features when dealing with simple radiological image retrieval is shape information. In fact, using gray level based approaches is not sufficient. They are in most of cases coupled with edge detection techniques [10]. This work deals with shape descriptors. It improves the rotation invariant shape contexts based on feature-space Fourier transformation [11]. First, fuzzy shape context histograms are computed. Then, a 2D FFT is performed on each

A. Elmoataz et al. (Eds.): ICISP 2012, LNCS 7340, pp. 432–441, 2012.

2D histogram. Next, data is projected onto a more representative feature space highlighting the most important variations between shapes. Eigenshapes form the basis of the new space. This will be more detailed in the next section.

This paper is organized as follows: Section 2 presents the proposed approach. Section 3 presents experimental results. The conclusion comes in section 4.

2 Rotation Invariant Fuzzy Shape Contexts Based on Eigenshapes and Fourier Transforms

2.1 Previous Work

S. Belongie et al. initially proposed the Shape context feature descriptor used for shape matching and object recognition [12] [13] [14]. The proposed approach inspired many authors to propose variants of this descriptor [15] [16]. S. Yang and Y. Wang proposed the rotation invariant shape contexs based on feature-space Fourier transformation [11]. The main idea is to extract the shape of the object and pick up n points. They do not need to be key points such as corners. Shape context at a given point p_i is an histogram providing the distribution of vectors originating from p_i over relative positions by considering p_i as the center of a log polar coordinate system (Figure 1). This distribution provides a reach description of the shape localized at that point.

Shape context at a given point p_i is defined as follows:

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in bin(k)\} \tag{1}$$

Indeed, coordinates and tangents at each point are used to compute a set $\{(r_{ij}, \alpha_{ij})|i, j = 1, 2 \ldots, n\}$ of magnitudes and angles. For a point p_i , magnitudes are obtained by first computing distances l_{ij} between p_i and the remaining points:

$$l_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \tag{2}$$

Then, a log scale is performed on all distances. Note that a length normalization is needed. Thus, every magnitude is divided by the mean distance r_0. Finally r_{ij} is determined as:

$$r_{ij} = \frac{\log(l_{ij})}{r_0} \tag{3}$$

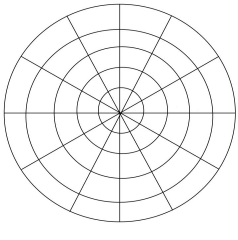


Fig. 1. Log polar grid with 60 bins used to compute shape contexts

Angles α_{ij} are defined as follows:

$$\alpha_{ij} = \arctan \frac{y_j - y_i}{y_j + y_i} \tag{4}$$

The obtained set $\{(r_{ij}, \alpha_{ij}) | i, j = 1, 2 \ldots, n\}$ is used to compute the 2D histogram defining the shape context. Application of a 2D FFT on this 2D histogram provides rotation-invariance.

2.2 Fuzzy Shape Contexts

The main idea behind fuzzy shape contexts concept consists in considering that the belonging of a contour pixel to a given bin is not absolute. It also belongs to the surrounding bins with smaller weights. This makes the descriptor more robust to local deformations. Figure 2 shows the case of a local shape deformation supposing that a log-polar grid of four bins is used to compute shape context histograms. A pixel belongs to a given bin with weight $w_1 = 0.7$. It also belongs to the previous and next bins with weights $w_2 = w_3 = 0.15$.

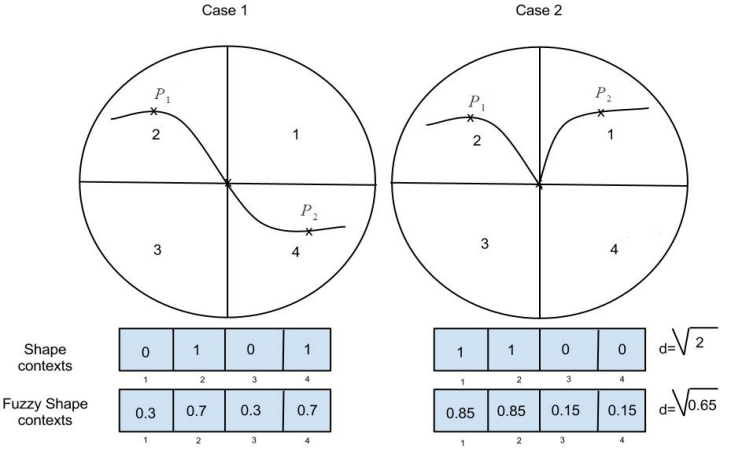


Fig. 2. A comparison between Shape contexts and Fuzzy Shape Contexts (illustration with one level-four bins)

The Euclidean distance is used to measure similarity between histograms. It is equal to $\sqrt{2}$ when shape contexts are used. However, it is equal to $\sqrt{0.65}$ when dealing with fuzzy shape contexts which are proven more robust to local deformations. Note that the difference $\delta = \sqrt{2} - \sqrt{0.65}$ is note huge. This is due to the fact that we are dealing with a local deformation.
In the rest of this work, a set of 12 equally log bins and 5 equally log radius bins is used to compute fuzzy shape contexts (Figure 3). Weights of belonging to a given bin are set empirically. There are three cases:

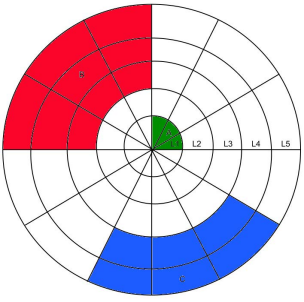


Fig. 3. Weight assignation

- a given pixel belongs to a bin of level L1 with weight $w_1 = 0.8$ and belongs to the next and the precedent bins with weight $w_1^{'} = 0.1$ for each (Eg. Bin A).
- a given pixel belongs to a given bin of level L2, L3 or L4 with weight $w_2 = 0.6$ and belongs to all the surrounding bins with weight $w_2^{'} = 0.05$ for each (Eg. Bin B).
- a given pixel belongs to a given bin of level L5 with weight $w_3 = 0.75$ and belongs to all the surrounding bins with weight $w_3^{'} = 0.05$ for each (Eg. Bin C).

2.3 Eigenshapes

For a given image, a reference point corresponding to the closest pixel to the top left image corner is fixed. The next step is to pick up other $n-1$ equidistant points. Every point is described via its fuzzy shape context which is a 2D histogram. Each histogram is then reshaped onto 1D vector which is added as a new line to the matrix representing the signature of that image. The signature is so an $n \times l$ matrix where n denotes the number of picked points and l denotes the number of bins. The next two sub-sections describe the training and recognition procedures.

Training. A set $S = \{S_1, S_2, \ldots, S_m\}$ of m images is used for training. Each image is represented by an $n \times l$ matrix. Each matrix is converted onto a column vector ζ_i. ζ_i is a $z \times 1$ vector where $z = n \times l$. Then, the average shape vector τ is computed as follows:

$$\tau = \frac{1}{m}\sum_{i=1}^{m}\zeta_i \tag{5}$$

Next, each ζ_i is normalized by subtracting the mean shape:

$$\Theta_i = \zeta_i - \tau \tag{6}$$

Then, the covariance matrix C is computed as follows :

$$C = \frac{1}{m}\sum_{n=1}^{m}\Theta_n\Theta_n^t = AA^t \tag{7}$$

Where $A = [\Theta_1, \Theta_2, \ldots, \Theta_m]$. Note that C in (7) is a $z \times z$ matrix and A is a $z \times m$ matrix. Eigenshapes are the eigenvectors U_i of AA^t.

Note that the matrix AA^t is very large so it is not practical for computations because of its dimension. Also, note that AA^t and A^tA have the same eigenvalues and their eigenvectors are related as follows: $U_i = AV_i$. Next, eigenvectors of A^tA are computed. Finally, m eigenvectors of AA^t are obtained following the relation: $U_i = AV_i$. Only k eigenvectors corresponding to the largest eigenvalues are kept. They form the basis of the new eigenshape space:

$$\Xi_k = [U_1, U_2, ..., U_k] \tag{8}$$

Each normalized shape in the training database is so projected in the new space. It is represented as a linear combination of k eigenshapes:

$$\Theta_i^{proj} = \sum_{j=1}^{k} W_j U_j \tag{9}$$

where $Wj = U_j^t\Theta_i$. Next, every normalized training shape Θ_i is represented by a vector ω^i providing its coordinates in the new eigenshape space where:

$$\omega^i = \begin{pmatrix} W_1^i \\ W_2^i \\ \vdots \\ W_k^i \end{pmatrix} \tag{10}$$

Retrieval. Now, given a query image, the goal is to retrieve the most similar image to it in the database. First, it is reshaped onto a column vector ψ . Then, it is normalized: $\theta = \psi - \tau$. The next step is to project it on the eigenshape space.

$$\theta^{proj} = \sum_{i=1}^{k} W_i U_i \tag{11}$$

where $W_i = U_i^t\theta$. Finally, θ is represented as:

$$\Omega = \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_k \end{pmatrix} \tag{12}$$

The last step is to compute $d = min_l||\Omega - \omega^l||$. The corresponding image to vector ω^l is considered as the most similar one to the query image.

3 Experiments

3.1 Image Collection

The radiological IRMA database is used for experiments. It includes images of several body parts. Figure 4 shows some IRMA database samples.

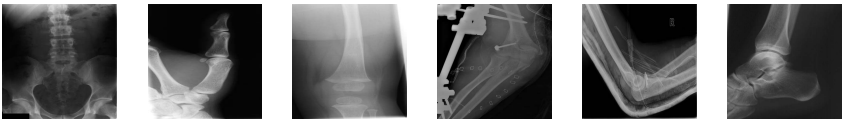
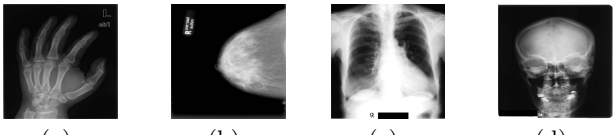


Fig. 4. IRMA samples

A set of 1000 images belonging to four classes (Hands, Breasts, Chests and Heads) is used. The number of images per class is the same. Figure 5 shows sample images of these classes.



(a) (b) (c) (d)

Fig. 5. Four classes used for performance measurement

Images in Figure 5 are randomly picked. They are used in the next sub-section as targets to evaluate the performance of the proposed approach. The Euclidean distance is used to measure the similarity between images.

3.2 Experimental Results

To evaluate the proposed approach, recall and precision measurements are used. Recall is defined as the ratio between the number of correctly retrieved images and the total number of images retrieved while precision is defined as the ratio between the number of correctly retrieved images by search and the total number of images used for test. For each measure of recall precision, the 10, 20, 40, 60, 80, 100, 150, 200 and 250 most similar images are taken in consideration. Figures 6, 7 and 8 shows recall versus precision for three tested approaches:

- FFT-RISC: Rotation-invariant shape contexts based on FFT [11].
- RISC-FFT-EIG: Rotation invariant shape contexts based on Fourier transforms and eigenshapes: histograms obtained by FFT-RISC are projected onto a new eigenshape space.

- Fuzzy RISC-FFT-EIG: Fuzzy Rotation invariant shape contexts based on Fourier transforms and Eigenshapes: Histograms obtained by FFT-RISC when using fuzzy bins are projected onto a new eigenshape space.

Figure 6 shows the recall precision curve for the Hand sample image (a) showing that Fuzzy RISC-FFT-EIG and RISC-FFT-EIG outperform significantly the FFT-RISC approach. Even when considering the best 250 retrieved images, precision rate remains superior to 90 %.

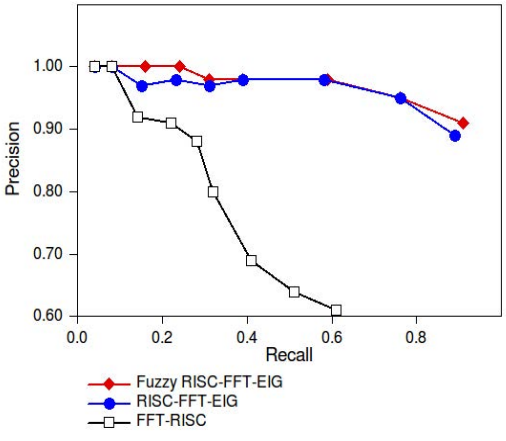


Fig. 6. Recall Vs. Precision for the Hand sample (a)

Recall and precision curve for the Breast sample image (b) is illustrated by Figure 7. The precision rate is equal to 100 % for the five first measurements for all of the three approaches. Fuzzy RISC-FFT-EIG and RISC-FFT-EIG provide better recognition rates than FFT-RISC when recall is higher than 0.4

Figure 8 shows recall precision curve for the Chest image sample (d). For the first measure, the precision rate is equal to 100 % for all of the tested approaches. Then, it is higher when using FFT-RISC. However the Fuzzy RISC-FFT-EIG and RISC-FFT-EIG outperform when recall is higher than 0.5.

To further prove the performance of the proposed approach, the average of the precision rate per class is computed considering the best 200 images retrieved. Results are illustrated by Table 1 showing that Fuzzy RISC-FFT-EIG and RISC-FFT-EIG outperform the FFT-RISC approach. This is due to elimination of noisy data.

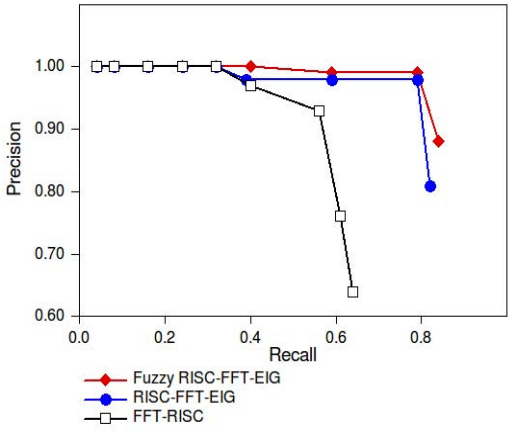


Fig. 7. Recall Vs. Precision for the Breast sample (b)

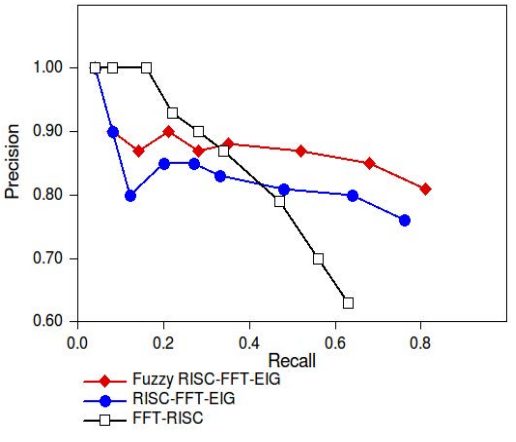


Fig. 8. Recall Vs. Precision for the Chest sample (c)

Table 1. Average of the precision rate per class considering the best 200 images retrieved

Image	FFT-RISC	RISC-FFT-EIG	Fuzzy-RISC-FFT-EIG
Hands	83.09	98.08	98.64
Breasts	85.65	93.47	94.1
Chests	98.49	97.01	97.92
Heads	95.67	94.06	94.02
Average	**90.72**	**95.65**	**96.16**

3.3 Discussion

Experimental results show that better results are obtained when histograms are projected in a new eigenshape space. The average of precision rate per class considering the best 200 images retrieved with RISC-FFT-EIG reaches 95.65 % while it is equal to 90.72 % with FFT-RISC. Other point to note is that using fuzzy shape contexts ameliorates results. In this case, the recognition rate is 96.16 %. Indeed, fuzzy shape contexts are more robust to local deformations. Note that there is no significant gap between results obtained by RISC-FFT-EIG and Fuzzy RISC-FFT-EIG approaches. In fact, local deformations do not affect significantly the performance of retrieval.

4 Conclusion

Shape context has been proven a very powerful shape descriptor. It is translation and scale invariant. Rotation invariance is achieved by application of 2D FFTs on the 2D histograms.

This work proves that using fuzzy bins makes the descriptor more robust to local deformations. Also, projecting data onto a lower dimensionality space highlighting the most important variations between shapes reduces time execution. In addition to that, better recognition rates are obtained. This is due to elimination of noisy data. Note that the major limitation of the proposed descriptor is the fact that it can not be used when dealing with images having many textures. The proposed approach can be improved if weights are set in respect to the linear distance between each pixel and the surrounding bins.

Acknowledgment. This work was supported by the New Brunswick Innovation Foundation (NBIF) and the Natural Sciences and Engineering Research council of Canada (NSERC). Authors would like to thank University of Ashen for providing the IRMA database.

References

1. Akgl, C.B., Rubin, D.L., Napel, S., Beaulieu, C.F., Greenspan, H., Acarl, B.: Content-based image retrieval in radiology: current status aExperimentsnd future directions. Digit. Imaging 24, 208–222 (2011)
2. Müller, H., Michoux, N., Bandon, D., Geissbuhler, A.: A review of Content-based image retrieval systems in medical applications-clinical benefits and future directions. International Journal of Medical Informatics 73(1) (2004)
3. Šajn, L., Kukar, M.: Image processing and machine learning for fully automated probabilistic evaluation of medical images. Computer Methods and Programs in Biomedicine 104(3), 75–86 (2011)
4. Krefting, D., Vossberg, M., Hoheisel, A., Tolxdorff, T.: Simplified implementation of medical image processing algorithms into a grid using a workflow management system. Future Generation Computer Systems 26(4), 681–684 (2010)

5. Mahmoudi, S.E., Akhondi-Asl, A., Rahmani, R., Faghih-Roohi, S., Taimouri, V., Sabouri, A., Soltanian-Zadeh, H.: Web-based interactive 2D/3D medical image processing and visualization software. Computer Methods and Programs in Biomedicine 98(2), 172–182 (2010)
6. Martínez, A., Jiménez, J.J.: Tracking by means of geodesic region models applied to multidimensional and complex medical images. Computer Vision and Image Understanding 115(8), 1083–1098 (2011)
7. Wei, L., Yang, Y., Nishikawa, R.M.: Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. Pattern Recognition 42(6), 1126–1132 (2009)
8. Chen, D.R., Huang, Y.L., Lin, S.H.: Computer-aided diagnosis with textural features for breast lesions in sonograms. Computerized Medical Imaging and Graphics 35(3), 220–226 (2011)
9. Kuo, W.J., Chang, R.F., Lee, C.C., Moon, W.K., Chen, D.R.: Retrieval technique for the diagnosis of solid breast tumors on sonogram. Ultrasound in Medicine and Biology 28(7), 903–909 (2002)
10. Bottigli, U., Golosio, B.: Feature extraction from mammographic images using fast marching methods. Nuclear Instruments and Methods in Physics 487(1-2), 209–215 (2002)
11. Yang, S., Wang, Y.: Rotation invariant shape contexts based on feature-space Fourier transformation. In: Fourth International Conference on Image and Graphics (2007)
12. Belongie, S., Malik, J.: Matching with Shape Contexts. In: IEEE on Content based Access of Image and Video Libraries, CBAIVL 2000 (2000)
13. Belongie, S., Malik, J., Puzicha, J.: Shape Context: A new descriptor for shape matching and object recognition (2001)
14. Belongie, S., Malik, J., Puzicha, J.: Shape Matching and Object Recognition Using Shape Contexts. IEEE Transaction on Pattern Analysis and Machine Intelligence 24(4) (2002)
15. Diplaros, A., Gevers, T., Patras, I.: Color-Shape Context for Object Recognition. In: IEEE Workshop on Color and Photometric Methods in Computer Vision, in Conjunction with the 9th Int. Conf. Computer Vision (2003)
16. Kortgen, M., Park, G.-J., Novotni, M., Klein, R.: 3D Shape Matching with 3D Shape Contexts