Translation and Rotation Invariant Content-based Image Retrieval System

Hossein SadeghianNejad, Jamshid Shanbehzadeh, Abdolhossein Sarafzadeh,

***Abstract*—This paper presents a novel content-based image retrieval algorithm to improve the retrieval performance in situations with local translation and rotation of image objects. This improvement is the result of employing Frobenius norm of wavelet coefficients. This norm ignores translation and rotation without negative effects on image object information. This algorithm consists of extracting wavelet coefficients of images and finding their Frobenius norm as the image features. The Frobenius norm of the query image is then compared with the norms of the images in the database. This paper uses the Corel 1000[1, 2] data set images for testing the algorithm. The evaluation criteria are recall rate, precision and F-measure.**

Keywords: content-based image retrieval, wavelet transform, feature extraction, image processing

I. Introduction

Content-based image retrieval (CBIR) refers to searching algorithms that extract similar images to a query image from a dataset of images according to their content. Due to the many image and video applications of CBIR and the development of high quality cameras with low prices, CBIR schemes have generated widespread interest [3, 4]. The idea of CBIR arose in the 1970s when images were annotated by text and so image search algorithms employed text instead of images in the retrieval process. However, annotating images suffers from a dependency on users providing efficient descriptions. These difficulties resulted in the development of algorithms that describe images based on their contents. These algorithms summarize images by the use of their features such as texture [5], color [6, 7] and shape [8, 9, and 10]. The complexities of image contents make choosing features to describe images complex too [11, 12]. Feature extraction plays an important role in CBIR [13, 14].

This paper focuses on situations where similar objects in a sub-class of image database have been translated or rotated. Our approach in defining features helps to overcome the problems arising from local translation or rotation, and also improves retrieval performance. The application of wavelet transforms (WT) and employing the F-norm of three categories of their coefficients for feature vectors results in robustness against translation and rotation.

Manuscript received on 9th December, 2011 and accepted on 23th December 2011

H. SadeghianNejad is with Tarbiat Moallem University of Tehran, Tehran, IR. Iran (E-mail: h.sadeghiyan@tmu.ac.ir).

Dr. J. Shanbehzadeh is an associate professor with Tarbiat Moallem University of Tehran, Tehran, IR. Iran (E-mail: jamshid@tmu.ac.ir).

Dr. A. Sarafzadeh is an Associate Professor and Head of Department of Computing at Unitec, New Zealand. (E-mail: hsarrafzadeh @unitec.ac.nz).

The rest of this paper is organized as follows. Section 2 talks about the algorithm. This section discusses WTs and how to extract robust image features based on WT coefficients (WTC) even if these images have been translated or rotated. Section 3 discusses the algorithm evaluation measures; recall rate (RR), precision rate (PR) and F-measure (FM). The final section presents the results of a simulation that showed algorithm robustness against translation and rotation with a suitable degree of FM.

II. Image Retrieval algorithm

Figure 1 shows a block diagram of a CBIR system. First, all images in the dataset undergo feature extraction and the feature vector (FV) of each image is generated. The features of each query image and its corresponding FV is generated as well. The FV of each query image is compared with those of the images in the dataset according to a similarity measure, and the most similar images are selected as the output results. The feature extraction consists of producing WTCs, and the FV consists of the WTC norms. The similarity measure is based on the closeness of the FV components. Next, WT and FV generation and the similarity measures are described.

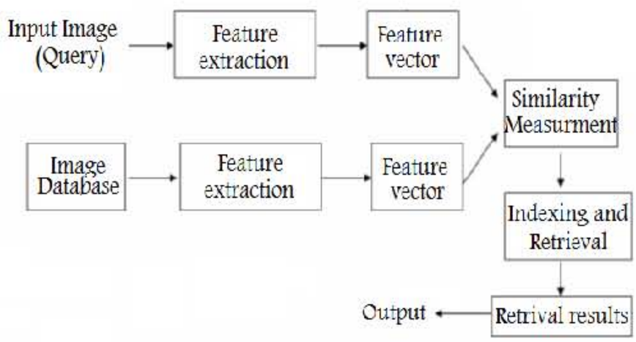


Fig 1. Block diagram of CBIR system

III. Wavelet Transform and Feature Vector Generation

WT is a powerful tool in digital image analysis. Fig. 2 shows the process of calculating the WT of an image. At first the image rows undergo low pass and high pass filtering; then the output results undergo high and low pass filtering from the column side. The final output consists of four images made up of the low pass version of the image from the row and column sides and three other images that contain the low pass version from the row (column) and high pass version from the column (row) side and the high

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

passversion from the row and column. The results are called LL, LH, HL and HH respectively.

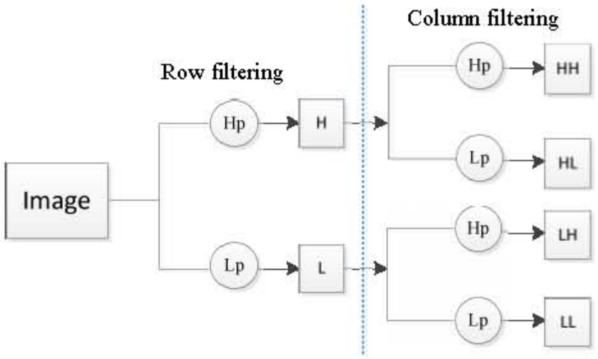


Fig.2: WT of an image

Figure 3 shows the feature selection scheme. As shown in Fig. 2, each coefficient of the FV consists of the F-norm of each output image by the WT. Equation 1 shows the F-norm of a vector. As shown in Fig. 2, we calculate the norm of the LL and LH rows and the HL columns of the wavelet coefficients. Eq. 2 shows the FV obtained by the norms. This method of FV generation ensures robustness against low levels of local translation and rotation.

$$if A \in R^n \Rightarrow norm(A) = (\sum_{i=1}^{n} (a_i)^2)^{\frac{1}{2}} \qquad (1)$$

$$V_F = \{norm(a_1), \dots, norm(a_n), , norm(h_1), \dots, norm(h_n) \; , norm(v_1), \dots, norm(v_n)\} \qquad (2)$$

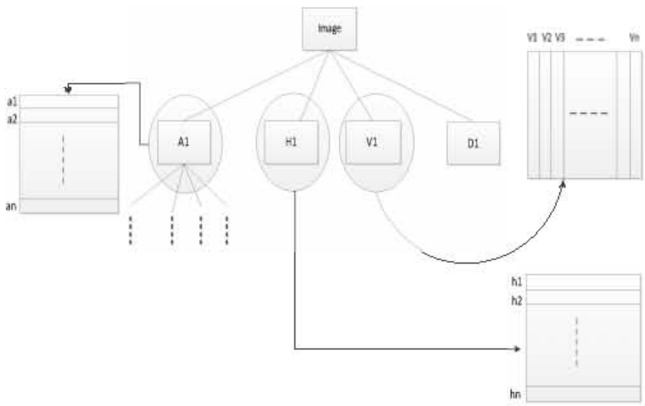


Fig3. FV generation

IV. Similarity Measure

The similarity between images in the current research is based on the ratio of two factors. The first is the minimum of two FV components and the second is their maximum value. The reason is that if two components are similar, this ratio will be high, and if they are not similar the ratio will be low [20]. Equation 3 shows the similarity measurement: in this equation $v_d^i and v_q^i$ are the ith component of FV of the query image and the image from the database. The average of all S_i is the final similarity factor, as shown in Equation 4. The most similar images to the query one are considered the desired output.

$$S_i = \begin{cases} \frac{\min(v_d^i, v_q^i)}{max(v_d^i, v_q^i)} v_d^i & or \quad v_q^i \neq 0 \\ 1 \qquad v_d^i & , \quad v_q^i = 0 \end{cases} \qquad (3)$$

$$S = \frac{\sum_{i=1}^{n} S_i}{n} \qquad (4)$$

V. SystemEvaluation Criteria.

Two criteria are normally employed to evaluate a CBIR system. These criteria are precision and recall rate, and are defined by Equations 5 and 6 respectively [21, 22]. In these equations, T is the total set of retrieved images and P is the total set of relevant images based on a desired similarity. Normally, P is a known set for each image. In the case of the Corel dataset, as all of the images have labels, the intersection of T and P can be found without human intervention. The absolute value notation shows the total number of elements of a set.

$$\text{Precision} = \frac{|P \cap T|}{|T|} \qquad (5)$$

$$\text{Recall} = \frac{|P \cap T|}{|P|} \qquad (6)$$

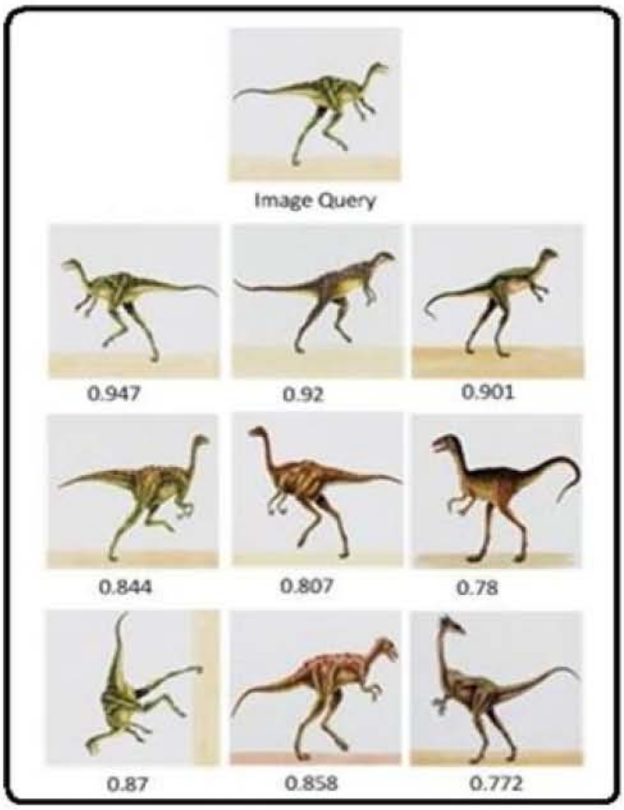
A suitable retrieval system should have a high recall rate and precision rate at the same time. As presented by Equation 7, M-Measure can be used to show system suitability [23]. This factor will have a high value if both the recall rate and the precision rate have high values at the same time; otherwise it will have a low value. Consequently, F-measure is a more suitable value for describing system performance.

$$F - \text{measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (7)$$

VI. Experimental Results

For this research we performed simulations on the Corel database, which contains different categories of images. Importantly, many of these images contain similar objects that have been translated and/or rotated. All of the images in the dataset have labels from which similar images can be identified. Figure 4 shows the retrieved images that are similar to the two sample images; the average similarity of the retrieved images is written under each image. It can be seen that the variation in the similarity of images under different rotations and translationswas very low. This result shows the robustness of the proposed algorithm against translation and rotation.

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)



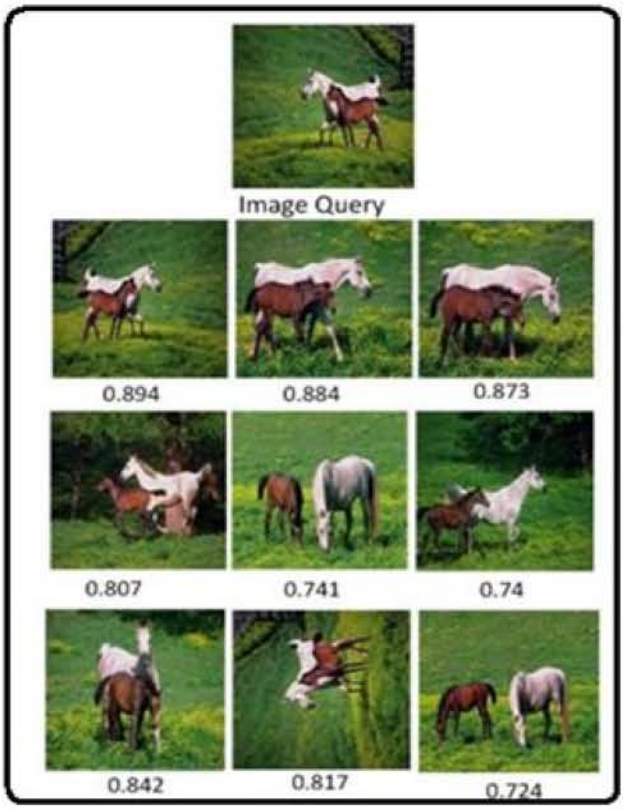
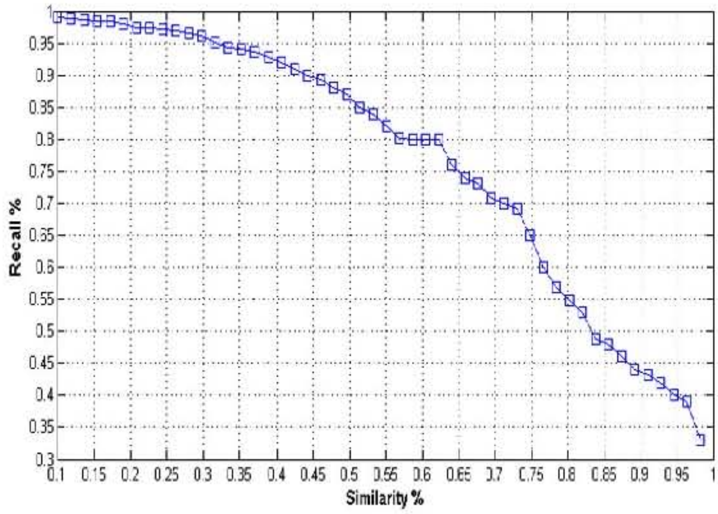
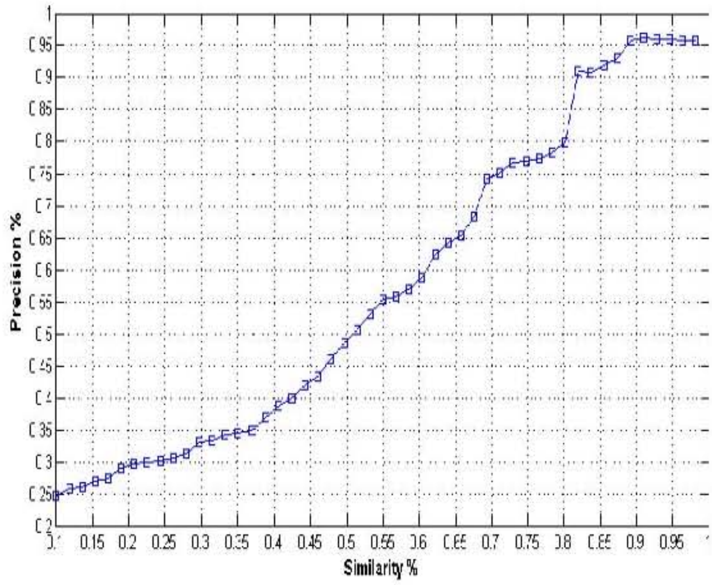


Fig 4. Similar images to query and the precision

While the new algorithm performs well under translation and rotation conditions, its overall performance is comparable to similar schemes. Figure 5 shows the average RR and PRaccording to the similarity measure. It is evident that for high similarity the RR is low but the PR is high, and vice versa. The best value for both RR and PR is the point where the F-measure is atits maximum value. Figure 6 shows that this point is where RR and PR are near 0.7 and 0.76 respectively. This gives an F-measure equal to 0.73.



a.Recall rate based on similarity



b. Precision rate based on similarity

Fig 5. Recall and precision rates based on similarity

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

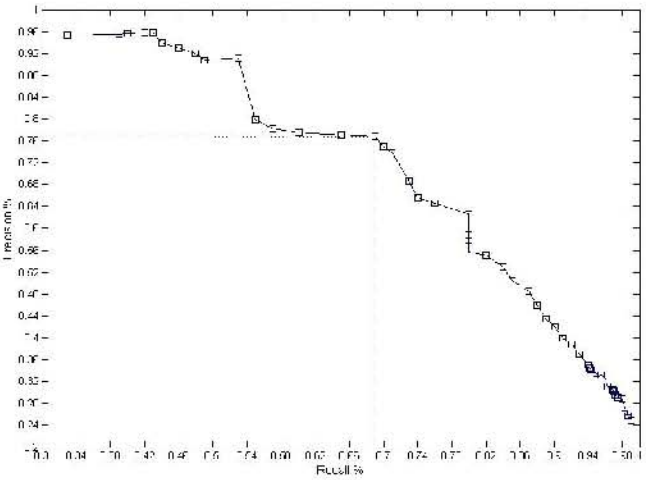


Figure 6. The best point of F-measure based on recall rate and precision rate

REFERENCES

[1] Jia Li, James Z. Wang, ``Automatic linguistic indexing of pictures by a statistical modeling approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pp. 1075-1088, 2003.

[2] James Z. Wang, Jia Li, Gio Wiederhold, SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries, IEEE Trans. on Pattern Analysis and Machine Intelligence, (2001) vol 23, no.9, p 947-963.

[3] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux, "Wavelet optimization for content-based image retrieval in medical databases", Medical Image Analysis 14 (2010) 227–241

[4] Datta, R., Joshi, D., Li, J., and Wang, J. Z. 2008. Image retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv. 40, 2, Article 5 (April 2008), 60 pages DOI = 10.1145/1348246.1348248

[5] Chuen-Horng Lin, Rong-Tai Chen, Yung-Kuan Chan, "A smart content-based image retrieval system based on color and texture feature", Image and Vision Computing 27 (2009) 658–665

[6] P.S.Suhasini , Dr. K.Sri Rama Krishna, Dr. I. V. Murali Krishna, "CBIR Using Color Histogram Processing", Journal of Theoretical and Applied Information Technology, (2009) Vol6. No1. (pp 116 – 122)

[7] Chuen-Horng Lin, Der-Chen Huang, Yung-Kuan Chan, Kai-Hung Chen, Yen-Jen Chang," Fast color-spatial feature based image retrieval methods", Expert Systems with Applications 38 (2011) 11412–11420

[8] S.Arivazhagan, L.Ganesan, S.Selvanidhyananthan, " Image Retrieval using Shape Feature", international journal of imaging science and engineering (IJISE),GA,USA,ISSN:1934-9955,VOL.1,NO.3, JULY 2007

[9] George Gagaudakis, Paul L. Rosin, "Incorporating shape into histograms for CBIR", Pattern Recognition 35 (2002) 81–91

[10] Jun Wei Han, Lei Guo, "A shape-based image retrieval method using salient edges", Signal Processing: Image Communication 18 (2003) 141–156

[11] Remco C. Veltkamp, Mirela Tanase, "Content-Based Image Retrieval Systems: A Survey"

[12] Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying Ma, "A survey of content-based image retrieval with high-level semantics", Pattern Recognition 40 (2007) 262– 282

[13] Khairul Muzzammil Saipullah, Deok-Hwan Kim, "A robust texture feature extraction using the localized angular phase", Multimed Tools Appl DOI 10.1007/s11042-011-0766-5

[14] Dmitriy Bespalov, William C. Regli, Ali Shokoufandeh, "Local feature extraction and matching partial objects", Computer-Aided Design 38 (2006) 1020–1037

[15] J. Berger, r. Coifman, and M. Goldberg, "Removing noise from music using local trigonometric bases and wavelet packets", Jour Audio Eng. Soc., Dec., 1994 pp. 808-818

[16] V. Bruni, D. Vitulano, "Combined image compression and denoising using wavelets", Signal Processing: Image Communication 22 (2007) 86–101

[17] D. Lee Fugal," Conceptual Wavelets in Digital Signal Processing", Space & Signals Technologies, San Diego, 2009

[18] Andras Hajdu, Tamas Toth," Approximating non-metrical Minkowski distances in 2D", Pattern Recognition Letters 29 (2008) 813–821

[19] Jun Ye," Cosine similarity measures for intuitionistic fuzzy sets and their applications", Mathematical and Computer Modelling 53 (2011) 91–97

[20] Y. M. Latha1, B.C. Jinaga and V.S.K. Reddy, "A Precise Content-based Color Image Retrieval: Lifting Scheme", ICGST-GVIP Journal, (2008), 25-32

[21] Tania DiMascio, DanieleFrigioni, LauraTarantino, "VISTO: A new CBIR system for vector images", Information Systems 35 (2010) 709–734

[22] L. Egghe, "The measures Precision, Recall, fallout and miss as a function of the number of retrieved documents and their mutual interrelations", Information Processing and Management 44 (2008) 856–876

[23] N. Lazarevic-McManus, J.R. Renno, D. Makris, G.A. Jones, "An object-based comparative methodology for motion detection based on the F-Measure", Computer Vision and Image Understanding 111 (2008) 74–85

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)