

A New Content Based Image Retrieval Model Based on Wavelet Transform

Davar Giveki¹, Ali Soltanshahi², Fatemeh Shiri³, Hadis Tarrah⁴

¹Iranian Research Institute for Information Technology (IranDoc), Department of Information Engineering, Tehran, Iran

²Department of Computer Science, University of Tehran, Tehran, Iran

³Department of Electrical Engineering, Tarbiat Modares University, Tehran, Iran

⁴Department of Electrical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran
Email: Giveki@students.irandoc.ac.ir, Ali.soltanshahi@gmail.com, fatima.shiri@gmail.com, hs.tarrah.88@gmail.com

Received January 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Searching interested images based on visual properties of images is a challenging problem and it has received considerable attention from researchers in the fields like image processing, computer vision and multimedia systems in the last 20 years. While the importance and the effect of the image features like color, texture and shape have been taken into account in many papers, there have not been many studies on the importance of the color spaces on the performance of Content Based Image Retrieval (CBIR) systems. In this paper we first experimentally study the effect of choosing color space on the performance of content based image retrieval using Wavelet decomposition of each color channel. To this end, the retrieval results of different color spaces like RGB, YUV, HSV, YCbCr and Lab are analyzed. Then as a result a new Content Based Retrieval model using Wavelet Transform in Lab color space and Color Moments is proposed. In order to increase the efficiency of the proposed model some division schemes are taken into account which improves the performance of the proposed model. The proposed model tackles one of the important restrictions in content based image retrieval, namely, the challenge between the accuracy of retrieval and its time complexity. The experimental results on two databases [19] [24] demonstrate the superiority of the proposed model compared to existing models.

Keywords

CBIR, Wavelet Transform, Color Moments, Image Division, RGB Color Space, HSV Color Space, YUV Color Space, YCbCr Color Space, Lab Color Space

1. Introduction

In the recent years, due to the progress of digital technology in producing digital cameras and availability of high-speed Internet connections, the amount of digital images has been grown rapidly. Those factors have provided a fast and simple way to generate and propagate visual content worldwide. That means that a huge amount of visual information becomes available every day to a growing number of users. Much of that visual information is available on the Web, which has become the largest and most heterogeneous image database so far. So there is an urgent demand for image retrieval systems [1] [2], which could be satisfied by content-based image retrieval (CBIR) systems. In CBIR systems, extracting image features like color, shape and texture is a very important step. This step is done by image descriptors. There are many papers reviewing and analyzing image descriptors [3]-[5]. Studying these papers give readers insight to choose a set of proper descriptors based on the task at hand. However, a very important point is choosing the color space in which the image features are supposed to be extracted.

Color is one of the widely used appearance features and utilizing more color spaces does not imply benefit of performance enhancement. That is because the poor performance color spaces influence on the high ones. So it is significant to evaluate the performance of different color spaces for CBIR. There are many color spaces in the field image processing. Each color space has its own advantages and disadvantages. In this paper we try to address the above point by experimentally studying the results of image retrieval using wavelet transform.

Wavelet transform has been shown to be a very important technique in image processing and computer vision. Using multi resolution strategy an image can be decomposed and represented on different resolutions and scales with different amount of details [6] [7]. Wavelet transform have been successfully applied to image denoising [8], image compression [9] and texture analysis [10]. In [11] authors propose a new CBIR system using color and texture features. In this paper texture features are extracted using 2D-DWT and Euclidean distance measure was used as a similarity measure between a query image and images in the database. In [12] wavelet basis was used to characterize each query image and also to maximize the retrieval performance in a training data set. Also, a regression function was used to estimate the best wavelet filter for each query image. Finally, a simple image characterization based on the standardized moments of the wavelet coefficient distributions was introduced. The authors claim that this feature extraction model is really fast. Experimental results show significant retrieval performance on medical image data set, a texture data set, a face recognition data set, and an object picture data set. In [13] discrete wavelet transform and edge histogram were used to describe the content of the image. First wavelet coefficients were extracted in horizontal, vertical and diagonal directions and then Edge Histogram Descriptor (EHF) was applied on selected wavelet coefficients to gather the information of dominant edge orientations. The combination of DWT and EHD techniques increases the performance of image retrieval system for shape and texture based search. In [14] authors propose a new model in which wavelet transform and color coherence vector are combined to construct image feature descriptors. First, Wavelet coefficients of image are computed using Daubechies wavelets. Then a quantization step is applied and finally coherence vector of the wavelet coefficients are computed. The retrieval results obtained by the model on a database contained 500 images showed the efficiency of the proposed algorithm. As it can be seen wavelet transform has been successfully employed in designing CBIR systems. Therefore, in this paper wavelet transform is used to capture image features.

The rest of the paper is organized as follows. In Section 2 wavelet transform is briefly discussed. Section 3 shortly points to different color spaces used in this paper. Section 4 briefly describes the proposed model for CBIR. In Section 5 the experimental results are discussed. Finally, the paper is concluded in Section 6.

2. Wavelet Transform

Wavelet transform is a signal processing technique that decomposes a signal or image into different frequency subbands at number of levels and multiple resolutions. In every level of decomposition, the high-frequency subband captures the details of the signal or images—for example, the edge information in an image. The low-frequency subband is a subsampled version of the original image, with similar statistical and spatial properties as the original signal. As a result, the low-frequency subband can be further decomposed into higher levels of resolution, and it helps in representing spatial objects in different coarser levels of accuracy in multi resolution subbands [6]. So, the wavelet transform decomposes a signal $f(x)$ with a family of functions obtained through dilations and translations of a kernel function $\psi(x)$, called the mother wavelet which is localized in both spatial

and frequency domains. This family of functions is denoted by:

$$\psi_{m,n}(x) = 2^{-\frac{m}{2}} \psi(2^m x - n) \quad (1)$$

where $m, n \in \mathbb{Z}^+$ indicate dilations and translations, respectively. To construct the mother wavelet $\psi(x)$ a scaling function $\phi(x)$ is needed:

$$\phi(x) = \sqrt{2} \sum_k h(k) \phi(2x - k) \quad (2)$$

Now, the wavelet kernel $\psi(x)$ can be defined as follows:

$$\psi(x) = \sqrt{2} \sum_k g(k) \phi(2x - k) \quad (3)$$

where

$$g(k) = (-1)^k h(N-1-k) \quad (4)$$

In the above equation N is the support of the filter. Now, the L -level decomposition of the signal $f(x)$ can be written as follows:

$$\begin{aligned} f(x) &= \sum_n c_{0,n} \phi_{0,n}(x) \\ f(x) &= \sum_n c_{L,n} \phi_{L,n}(x) + \sum_{l=1}^{L+1} \sum_n d_{l,n} \phi_{l,n}(k) \end{aligned} \quad (5)$$

The coefficients $c_{0,n}$ are given and the coefficients $c_{L,n}$ and $d_{l,n}$, both at scale l , are obtained by the coefficients $c_{l-1,n}$ at scale $l-1$ through:

$$\begin{aligned} c_{l,n} &= \sum_k c_{l-1,n} h(k-2n) \\ d_{l,n} &= \sum_k c_{l-1,n} g(k-2n) \end{aligned} \quad (6)$$

where $1 \leq l \leq L+1$. A recursive wavelet decomposition can be obtained through $h(k)$ and $g(k)$ in Equation (6). The same process can be viewed as the convolution of signal $c_{l-1,n}$ with the impulse responses $\bar{h}(n) = h(-n)$ and $\bar{g}(n) = g(-n)$ of the low- and high-pass filters H and G , respectively (also known as quadrature filters), and then by down sampling the filtered signals by a factor of 2. The 2-D wavelet and scaling functions can be expressed as the tensor products of their 1-D complements:

$$\begin{aligned} \phi_{LL}(x, y) &= \phi(x)\phi(y) & \psi_{LH}(x, y) &= \phi(x)\psi(y) \\ \psi_{HL}(x, y) &= \psi(x)\phi(y) & \psi_{HH}(x, y) &= \psi(x)\psi(y) \end{aligned} \quad (7)$$

3. Technical Color Spaces

A color space is an abstract mathematical model describing the way colors can be represented as tuples of numbers, typically as three or four values or color components (e.g. RGB and CMYK are color spaces). In the rest of this paper we used RGB, HSV, YUV, YCbCr and Lab color spaces in our experiments. For getting more information regarding different color spaces reader can refer to the references [15]-[17].

4. Proposed Model

In this section a brief description of our proposed model is presented. At first, each image in database is converted to the aforementioned color spaces. Then, for each color channel features of the images in database are extracted using wavelet transform. In the next step a feature vector for each image is composed of wavelet coefficients extracted from each image channel. So, for image I in the data base feature vector f_I is defined as follows:

$$f_I = (w_{c_{first}}, w_{c_{second}}, w_{c_{third}}) \quad (8)$$

where $w_{c_{first}}$ are first and second order moments of the wavelet coefficients of the first channel of the image I

in color space C . After that all the feature vectors will be kept in order to be used in next steps of retrieval process. Another important building block of image retrieval systems is similarity measure. In this paper L_2 norm is used as a similarity measure. Once the query image is presented, its feature vectors are extracted and used for computing the similarity between this image and database images. Suppose that I is a database image with feature vector $f_I = (w_{C_{first}}, w_{C_{second}}, w_{C_{third}})$ and Q is a query image with feature vector $f_Q = (w'_{C_{first}}, w'_{C_{second}}, w'_{C_{third}})$. Now, the similarity between these two images is computed using the following equation:

$$Sim(Q, I) = \sqrt{(w'_{C_{first}} - w_{C_{first}})^2} + \sqrt{(w'_{C_{second}} - w_{C_{second}})^2} + \sqrt{(w'_{C_{third}} - w_{C_{third}})^2} \quad (9)$$

Figure 1 shows the steps for image retrieval using proposed model. So, in order to do a comparative study of the effect of choosing color space on the performance of the CBIR systems in each color space the image features are extracted in the same way using the proposed model. Also conversion from RGB color space and other color spaces has been done using the known models in the literature [18].

5. Experimental Results

The database used in our experiments is Wang database [21]. It consists of 1000 images in 10 categories and 100 images in each category (**Table 1**). In order to compare the effect of the color spaces on the performance of proposed CBIR model, statistical measures namely Precision and Recall were used. These measures can be computed as follows:

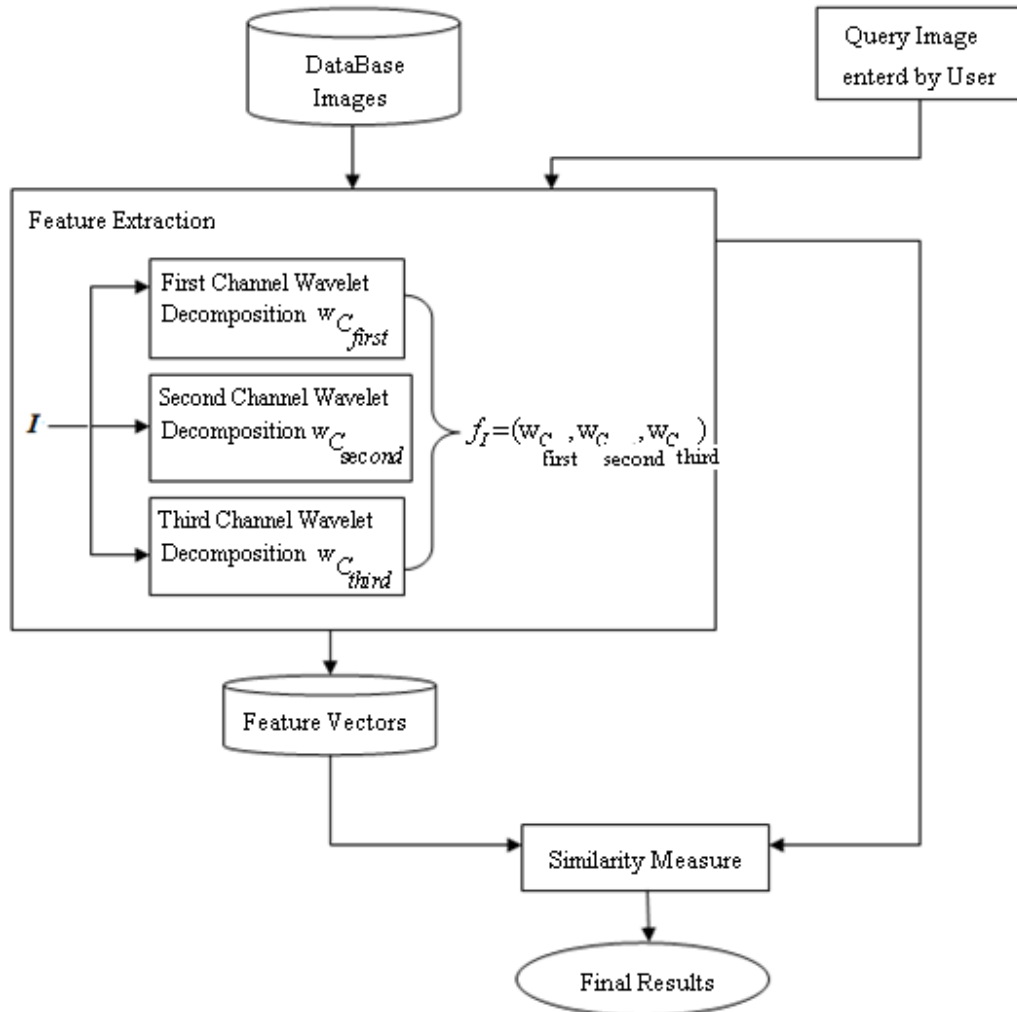


Figure 1. Algorithm of the proposed method.

$$precision = \frac{|TP|}{|TP + FP|}, \quad recall = \frac{|TP|}{|TP + FN|} \quad (10)$$

As it can be seen from **Table 2**, the most promising results were achieved using Lab color space.

In order to compare the effect of the color spaces on the performance of proposed CBIR model, statistical measures namely Precision and Recall were used. These measures can be computed as follows:

In the next stage, in order to get better results, we decided to capture information from different parts of the images. In this way extracted features can be supposed as local features since they have been extracted from different parts of image not from the entire image at once. These local features give us more discriminative power so that we can get better results. In **Figure 2(a)**, image is 9 parts with the same size while in **Figure 2(b)**, the image is divided into 9 different parts and the size of the central part is 9/25 of the area of the image. In **Figure 2(a)**, the image is divided into 5 parts. Each part is 1/4 area of the image. In **Figure 2(d)**, The image is divided into 5 parts. Each part is 1/2 area of the image except the central part which is 4/9 area of the image. While extracting image features one can assign more weights to the central parts because in most of the images the core objects are likely to be appeared in the central part.

Table 1. Ten classes of the image dataset.

Category	Semantic name
1	African people and village
2	Beach
3	Building
4	Buses
5	Dinosaurs
6	Elephants
7	Flowers
8	Horses
9	Mountains and glaciers
10	Food

Table 2. The average precision and recall of the methods.

Semantic name	RGB-Wavelet		HSV-Wavelet		YUV-Wavelet		YCbCr-Wavelet		Lab-Wavelet	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
African people and village	51.34	45.28	53.26	46.35	38.29	33.79	56.68	54.17	58.73	55.36
Beach	42.52	37.46	44.59	37.28	35.27	30.24	47.67	43.24	48.94	46.27
Building	46.41	39.62	41.36	37.64	38.21	34.18	49.57	42.67	53.74	45.37
Buses	86.54	79.37	88.45	85.37	80.69	75.36	92.68	88.34	95.81	91.38
Dinosaurs	92.68	88.14	94.54	91.23	89.84	84.23	97.21	92.13	98.36	95.72
Elephants	61.24	59.26	65.19	60.27	46.27	40.27	61.75	52.84	64.17	58.12
Flowers	72.34	68.53	77.15	73.27	69.44	63.12	81.37	80.36	85.64	83.57
Horses	67.68	55.79	69.14	65.31	66.37	58.14	77.92	72.82	80.31	76.19
Mountains and glaciers	37.27	31.68	41.58	36.74	27.26	20.45	49.17	46.32	54.27	50.67
Food	59.32	44.85	62.14	57.27	51.27	47.38	61.73	57.82	63.14	62.57
Average	61.787	54.998	63.74	55.309	54.321	47.716	67.575	63.071	70.311	66.522

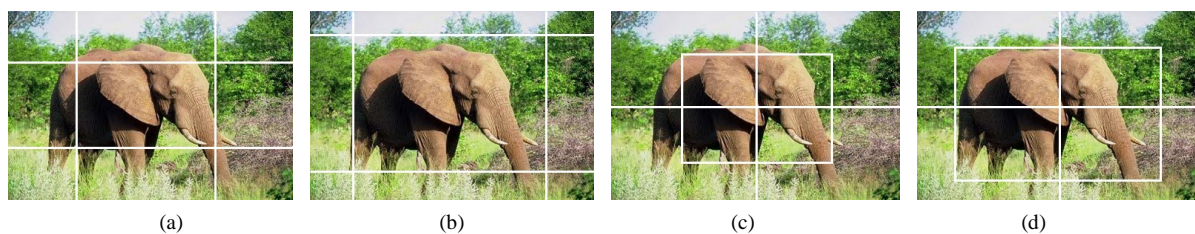


Figure 2. Image Division with different schemes.

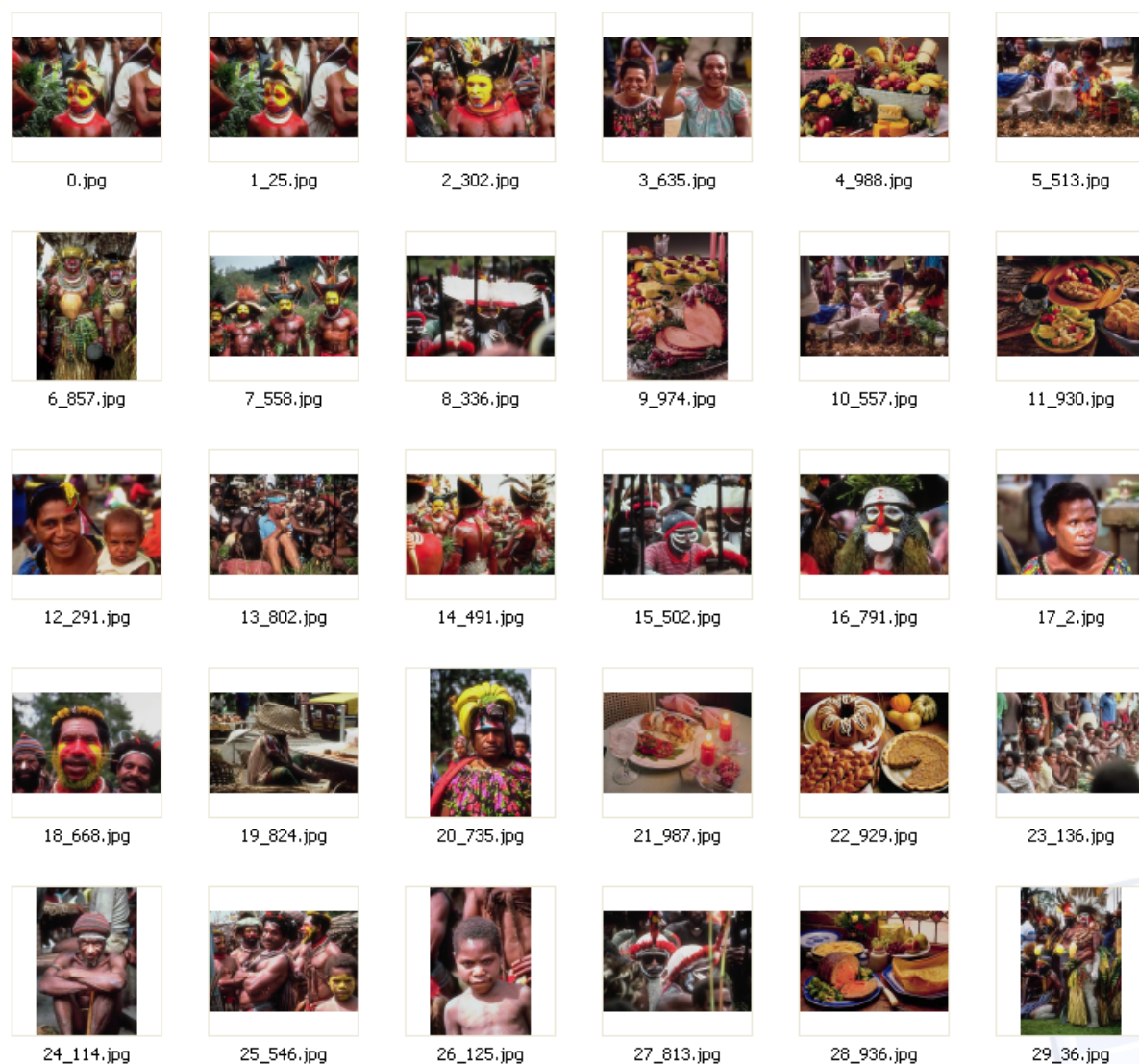


Figure 3. The retrieval results of an image from the Wang database [21].

Here we used color moments as another image features. These moments can be computed as follows:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij}, \quad \sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2 \right)^{\frac{1}{2}}, \quad s_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right)^{\frac{1}{3}} \quad (11)$$

where f_{ij} is the value of the i^{th} color component of the image pixel " j ", and N is the total number of pixels in the image. Therefore, using wavelet transform in the Lab color space and computing color moments the feature vector

Table 3. The average precision and recall of the proposed model using wavelet transform in lab color space and color moments.

Semantic Name	Precision	Recall
African people and village	65.51	58.34
Beach	59.37	54.75
Building	60.93	52.91
Buses	97.58	92.73
Dinosaurs	99.21	96.67
Elephants	67.39	63.46
Flowers	93.42	87.94
Horses	92.35	85.42
Mountains and glaciers	62.83	56.28
Food	66.74	63.84
Average	76.533	71.234

of an image is constructed. Using this proposed model we could achieve a high performance as it can be seen from **Table 3**.

Compared to other models reported in [20]-[23] our proposed model provides promising results. We also successfully implemented our proposed model on Li database [24] and we could achieve high rate of accuracy. **Figure 3** shows the retrieval results of the proposed model. It should be noted that final results of the first 29 similar images are presented.

6. Conclusion and Future Work

This paper gives an experimental study of the effect of choosing color space on the performance of a CBIR system while wavelet transform is employed as image feature descriptor. The final results show that Lab color space provides the most promising results. Therefore, a new Content Based Retrieval model using Wavelet Transform and Lab color space and Color Moments is proposed. The proposed model tackles one of the important restrictions in content based image retrieval, namely, the challenge between the accuracy of retrieval and its speed. The experimental results on a Wang database [19] and Li database [24] demonstrate the superiority of our proposed. In future, incorporating other image descriptors accompanied by an ensemble of classifiers aiming at achieving more promising results on larger databases will be the direction of our studies.

References

- [1] Kherfi, M.L., Ziou, D. and Bernardi, A. (2004) Image Retrieval from the World Wide Web: Issues, Techniques, and Systems. *ACM Computing Surveys*, **36**, 35-67. <http://dx.doi.org/10.1145/1013208.1013210>
- [2] Datta, R., Joshi, D., Li, J. and Wang, J.Z. (2008) Image Retrieval: Ideas, Influences, and Trends of the NEW Age. *ACM Computing Surveys*, **40**, 1-60.
- [3] Yang, M., Kpalma, K. and Ronsin, J. (2010) A Survey of Shape Feature Extraction Techniques. *Pattern Recognition*, 1-38.
- [4] Penatti Otavio, A.B., Valle, E. and Torres, R.da.S. (2012) Comparative Study of Global Color and Texture Descriptors for Web Image Retrieval. *Int. J. Via. Commun. Image R*, 359-380.
- [5] Deselaers, T., Keysers, D. and Ney, H. (2008) Features for Image Retrieval: An Experimental Comparison. *Information Retrieval*, **11**, 77-107.
- [6] Mallat, S.G. (1989) A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674-693.
- [7] Sarck, J.L., Murtagh, F.D. and Bijaoui, A. (1998) Image Processing and Data Analysis: The Multiscale Approach.

- [8] Hill, P., Achim, A. and Bull, D. (2012) The Undecimated Dual Tree Complex Wavelet Transform and Its Application to Bivariate Image Denoising Using a Cauchy Model. *19th IEEE International Conference on Image Processing (ICIP)*, 1205-1208. <http://dx.doi.org/10.1109/icip.2012.6467082>
- [9] Kalra, M. and Ghosh, D. (2012) Image Compression Using Wavelet Based Compressed Sensing and Vector Quantization. *IEEE 11th International Conference on Signal Processing (ICSP)*, **1**, 640-645.
- [10] Kokareh, M., Biswas, P.K. and Chatterji, B.N. (2005) Texture Image Retrieval Using New Rotated Complex Wavelet Filters. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, **35**, 1168-1178.
- [11] Balamurugan, V. and Anandha Kumar, P. (2008) An Integrated Color and Texture Feature Based Framework for Content Based Image Retrieval Using 2D Wavelet Transform. *IEEE International Conference on Computing, Communication and Networking*, 1-16. <http://dx.doi.org/10.1109/iccnet.2008.4787734>
- [12] Quellec, G., Lamard, M., Cazuguel, G., Cochener, B. and Roux, C. (2012) Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval. *IEEE Transactions on Image Processing*, **21**, 1613-1623.
- [13] Agarwal, S., Verma, A.K. and Singh, P. (2013) Content Based Image Retrieval Using Discrete Wavelet Transform and Edge Histogram Descriptor. *International Conference on Information Systems and Computer Networks (ISCON)*, 19-23. <http://dx.doi.org/10.1109/iciscon.2013.6524166>
- [14] Wang, Y. and Zhang, W. (2012) Coherence Vector Based on Wavelet Coefficients for Image Retrieval. *IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, **2**, 765-768. <http://dx.doi.org/10.1109/CSAE.2012.6272878>
- [15] Hu, J.-L., Deng, J.-B. and Sui, M.-X. (2009) Color Space Conversion Model from CMYK to LAB Based on Prism. *IEEE International Conference on Granular Computing*, 235-238. <http://dx.doi.org/10.1109/grc.2009.5255123>
- [16] Pratt, W.K. (2001) Digital Image Processing. 3rd Edition, PIKS Inside. Wiley. <http://dx.doi.org/10.1002/0471221325>
- [17] Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F. and Phillips, R.L. (1993) Introduction to Computer Graphics. Addison-Wesley Longman, Amsterdam.
- [18] Ford, A. and Roberts, A. (1998) Color Space Conversions.
- [19] Wang Database. <http://wang.ist.psu.edu/docs/related.shtml>
- [20] Meng, F., Guo, B. and Fang, Y. (2010) Novel Image Retrieval Model based on Interest Points. *3rd International Congress on Image and Signal Processing CISP*, 1582-1585.
- [21] Lin, Ch.-H., Chen, R.-T. and Chan, Y.-K. (2009) A Smart Content Based Image Retrieval System Based on Color and Texture Features. *Image and Vision Computing*, **27**, 658-665.
- [22] Huang, P.W. and Dai, S.K. (2003) Image Retrieval by Texture Similarity. *Pattern Recognition*, **36**, 665-679.
- [23] Jhanwar, N., Chaudhurib, S., Seetharamanc, G. and Zavidovique, B. (2004) Content Based Image Retrieval Using Motif Co-Occurrence Matrix. *Image and Vision Computing*, **22**, 1211-1220.
- [24] Li's Database. <http://sites.stat.psu.edu/~jiali/index.download.html>