

# A Method for Image Retrieval using Combination of Color and Frequency Layers

Mohammad Reza Azodinia  
Department of Computer Graphics and Image  
Processing  
Faculty of Informatics, University of Debrecen  
Debrecen, Hungary

András Hajdu  
Department of Computer Graphics and Image  
Processing  
Faculty of Informatics, University of Debrecen  
Debrecen, Hungary

## ABSTRACT

In this paper a fast and effective noise-resistant method for image retrieval has been proposed. In this method, first, the image is decomposed into different frequency layers using complex wavelet transform so as to make it possible to extract the texture features of the image. Thereafter, in the HSV color space, each layer is quantized into 166 different colors and the color histogram is calculated for each layer. Furthermore, a number of statistical features are extracted from each sub-image using complex wavelet transform, which are used along with other features for image retrieval. In order to verify the effectiveness of the proposed method, it has been evaluated using a dataset containing 3000 images and compared to a competent method in this field. The results prove the superiority of the proposed method.

## General Terms

Image retrieval system, image processing, color histograms, texture and statistical features

## Keywords

Color feature, complex wavelet transform, Content-based image retrieval, feature extraction, histogram, image processing and texture feature.

## 1. INTRODUCTION

The global spread and success of the internet along with advances in data storage and image acquisition technologies in recent years have supported the rise of the amount of image data which is accessible to all, which has derived many researchers to make significant efforts to find efficient solutions for managing these collections of images. Apparently, one of the most basic actions which is performed on an image is the retrieval of that image. Image retrieval can be found in many applications and due to this wide range of applications, it has been the center of attention in many researches.

The content-based image retrieval (CBIR) technology has been widely used in various fields including digital libraries, biodiversity information systems (BISs), or even medical applications [1-9]. The digital museum of butterflies which is intended to create a digital collection for Taiwanese butterflies, has used a CBIR component which is based on the texture, color and patterns of images taken from these butterflies [2]. In biodiversity studies the researchers gather various data about different species, which include spatial data along with images taken from the species. In order to achieve a high-level view about the diversity of live beings and the trend of changes in their population, image-based queries can be of great use. Some studies have been conducted in this regard, such as [2] and [9]. Moreover, while many image processing techniques have been utilized in medical applications especially to make proper diagnoses of

different diseases such as breast cancer, the introduction of image retrieval techniques to these image processing systems can improve their overall efficiency [1].

The main goal of researchers who study on CBIR is to create and develop methods which provide the ability to store and retrieve images based on their contents. At the preliminary stages of creating image storage and retrieval systems, textual descriptors were used to this aim, which were created by a person who had to review the whole image dataset and assign one or more keywords that were representative of each image to it. In the retrieval process, these keywords were used to find images corresponding to queries. However, this approach suffers from two main demerits [9]. First and foremost, assigning some keywords to an image is highly subjective and can be carried out quite differently by different people. Furthermore, with the increase in the number of elements which exist in the database this process becomes harder and harder, until it becomes impractical.

To address these shortcomings, researchers considered using structural characteristics of images to retrieve them, which resulted in what we know as CBIR today. According to this approach, some features of images, such as color, texture and shape are extracted and used to index them. Swain and Ballard introduced the color histogram method in 1991 [10]. This method counts the number of pixels which have colors in each of a fixed list of colors that are made through quantization of the colors. An important advantage of histograms is that they are invariant to rotation and translation around an axis perpendicular to the image, and varies only slowly with the angle of view, therefore a three dimensional item can be retrieved easily despite a small change in the angle of view. Nevertheless, the disadvantage of this technique is that images with completely different shapes may have similar color histograms, which may occur more frequently as the number of images in the dataset increases.

In order to overcome the problems which exist in color histogram, the color coherence vector (CCV) method was proposed [11]. This method adds some extra information, which correspond to the spatial positions of the pixels, to the typical histogram. According to this approach, the pixels which belong to each bin are classified based on their local features either as coherent or incoherent and the number of pixels which belong to each category are stored. Coherent pixels are part of a big of connected component whereas incoherent pixels are part of a small connected component. Despite the superiority of the CCV method over the color histogram, it still has a major disadvantage that is the extraction of feature vectors takes too long.

In a study conducted by Huang et al. [12] color correlogram has been used to characterize each image. The rationale behind using color correlogram has been to consider the spatial relations of pixels and overcome the shortcomings of

the color histogram. This approach is based on the assumption that the spatial correlation of each pair of colors in different images are different. This approach has proved effective, however it has the same prohibitive problem as what the CCV suffers from, i.e. suffers from a high computational complexity.

## 2. THE PROPOSED METHOD

The best metric to evaluate an image retrieval system is evaluation of the system made by its users. Consequently, using the same methods as what humans use to find similar images should be effective in order to increase the acceptability of such systems [13]. There are different spatial frequency channels in human vision each of which function as a band-pass filter. Therefore, this vision system can be considered as a filter bank which is comprised of filters each of which receive and process a specific part of images. High details parts of an image and its edges correspond to high frequencies, and parts of the image with smooth variations, which usually include the regions inside the objects or backgrounds, correspond to low frequencies. Therefore, parts of an image which have different frequencies may belong to the same object or to different objects. That being said, human visual system works like this: first, a frequency analysis is carried out on the image that is formed on the retina, which decomposes the image into different frequency layers. Thereafter, the equivalent parts of images are compared with each other which forms the basis for the general evaluation of the similarity of two images.

The proposed method is comprised of two main parts. The first part determines how the filter bank should be implemented and applied to decompose the input image into layers, and the second part proposes a method to extract features from each layer. The block diagram which illustrates how these layers are extracted has been depicted in the Figure 1.

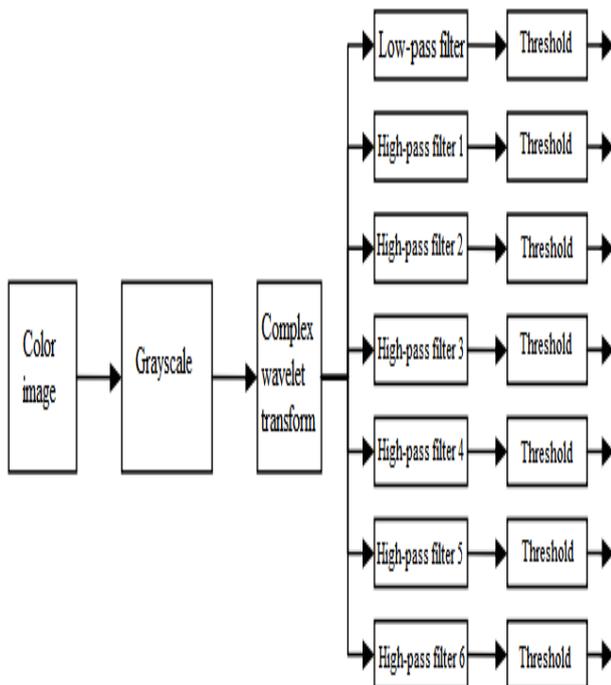


Fig 1: The block diagram of the layer extraction system.

With the extraction of these layers, each layer represents a different range of spatial frequency. The first layer determines the regions with low-frequency which correspond to backgrounds or the interior regions of objects, while the other 6 layers represent the regions which include more details in. If the variation range of the value of the pixels after the application of wavelet transform filters are set to be in the interval  $[0, 1]$ , all the threshold can be set to 0.5. For the first layer, the pixels whose value is lower than 0.5 and for the rest of the layers the pixels whose value is higher than 0.5 are taken into account and the other pixels are neglected. Moreover, in order to decrease the length of the feature vector, which results in a lower computation complexity, we can use only 3 high-pass filters out of the 6 high-pass filters mentioned before. In this case, only the first, third, and fifth filters are uses.

Thereafter, color features of each layer are extracted using ordinary color histogram method. At this stage the RGB color space is converted to HSV color space, because it complies better with human perception of colors. Then, this space is quantized into 166 distinct values. This quantization is based on [14] according to which 18 distinct values for H, 3 distinct values for both S and V and 4 distinct values for gray level are used. The reason why we have quantized the color space as described is that it is proved that this approach is quite similar to the human visual system perception of colors. Therefore, 7 histograms are created for the 7 layers as mentioned previously. Furthermore, the reason why complex wavelet transform is used is that they are very effective in extraction of texture features from images, which can be witness in Gabor filters [15], and at a higher speed compared to Gabor filters.

In the proposed method, the complex wavelet transform is applied to the image up to 4 levels, and the values of mean, standard deviation and kurtosis are calculated for the sub-images resulting from each of the 6 high-pass filters. Therefore, a vector composed of 72 entries is achieved for each image which is added to the color feature of layers which mentioned previously.

## 3. EXPERIMENTAL RESULTS

In order to evaluate the proposed method, it has been implemented using Matlab running on a PC with an Intel CORE 2 Due 2.20 GHz processor and 4GB of RAM, on which the extraction of each feature from an image takes about 2 seconds. A collection including 3000 images borrowed from Hamshahri dataset was used in our experiments [16]. The main reason for using this dataset was the fact that this dataset is extracted from an online news website so the images have come with very precise details about the category or categories the image belong to and each image is supported with a list of keywords which describe its contents. This gave us the ability to precisely verify the effectiveness of the proposed method. Although the Hamshahri dataset includes 36 main categories with many more sub-categories, we used 5 main categories in our evaluations including sports, buildings, animals, weather forecast, and accidents.



Fig 2: A number of pictures selected from the dataset used to evaluate the proposed method.

A number of pictures which belong to this dataset have been illustrated in the Figure 2.

Two common metrics, namely precision, and recall have been used to evaluate the proposed method. These metrics have been used in various applications and their definitions have been adapted to make them suit the application properly. We have used the definitions given in [17]:

$$P_j = \frac{\text{Number of retrieved and relevant elements in the first } j \text{ positions}}{j} \quad (1)$$

$$R_j = \frac{\text{Number of retrieved and relevant elements in the first } j \text{ positions}}{\text{Total number of relevant elements in the collection}} \quad (2)$$

We have used the tags provided for each image, as mentioned before, in order to verify whether the image is relevant to the query or not.

According to the aforementioned equations,  $P_j$  shows an almost increasing behavior when  $j$  increases while the opposite is true for  $R_j$ . As we mentioned before, color correlogram is one of the best methods in CBIR, so we have compared our proposed method with this method. The results have been shown in the Table 1.

Table 1. The results of experiments conducted using the proposed method (Pr) and the color correlogram (CC) method.

Category	P <sub>5</sub>	P <sub>10</sub>	P <sub>15</sub>	P <sub>20</sub>	R <sub>5</sub>	R <sub>10</sub>	R <sub>15</sub>	R <sub>20</sub>	Method
Sports	1	1	1	0.94	0.22	0.47	0.79	0.8	Pr
	1	0.94	0.81	0.84	0.17	0.39	0.71	0.78	CC
Buildings	1	1	1	1	0.14	0.33	0.59	0.63	Pr

	0.89	0.74	0.86	0.86	0.12	0.28	0.45	0.57	CC
Animals	1	0.99	0.99	0.98	0.1	0.21	0.31	0.46	Pr
	1	0.96	0.83	0.76	0.09	0.21	0.25	0.9	CC
Weather Forecast	1	0.99	0.99	0.87	0.14	0.22	0.32	0.37	Pr
	0.96	0.92	0.85	0.86	0.15	0.19	0.28	0.31	CC
Accidents	0.98	0.98	0.98	0.98	0.04	0.09	0.21	0.29	Pr
	0.81	0.82	0.89	0.85	0.03	0.06	0.11	0.14	CC

As it is implied from the Table 1, the proposed method shows a consistent superiority over the color correlogram method which is a competent method in the field of CBIR.

Some of the images retrieved using the proposed method in response to the query image shown as the top-leftmost image in the Figure 3 are depicted in this figure.



Fig 3: Some of the images retrieved in response to the query image shown in the top-leftmost position.

#### 4. CONCLUSION

In this paper we proposed a method for the retrieval of images which is developed based on human visual system. The proposed method first decomposes each image into different frequency bands, each of which represents different details of the image. In order to compare to images, the features of the corresponding sub-bands are used. To this aim, complex wavelet transform has been used which like Gabor filters covers almost the entire frequency space, and have a high computational speed. In addition, the HSV color space is quantized so that matches the human visual system to high extent.

Our proposed method was applied on a dataset including about 3000 images and the results compared to those of color correlogram method, which proved the superiority of the proposed method.

In the future, the authors will try to combine the proposed method which is presented in this paper with image retrieval methods that take image shapes into account in computing the similarity of images. One of the most important issues to which we note mostly is the time complexity of the algorithm. Furthermore, parallelization methods and distributed architectures which can reduce the computation time are of highest priority in our future researches. The authors intend to concentrate on Hadoop [18] framework and to find solutions based on MapReduce programming model [19].

#### 5. ACKNOWLEDGMENT

This work was supported in part by the OTKA grant NK101680, and by the projects TAMOP-4.2.2.C-11/1/KONV2012-0001, TAMOP-4.2.2.A-11/1/KONV-2012-0045, and TAMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program' supported by the European Union and

the State of Hungary, co-financed by the European Social Fund.

## 6. REFERENCES

- [1] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler. A Review of Content-Based Image Retrieval Systems in Medical Applications – Clinical Benefits and Future Directions. *International Journal of Medical Informatics*, 73(1):1–23, Feb2004.
- [2] J. S. Hong, H.Y. Chen, and J. Hsiang. A Digital Museum of Taiwanese Butterflies. In *Proceedings of the Fifth ACM Conference on Digital Libraries*, pages 260–261, San Antonio, Texas, United States, 2000. ACM Press.
- [3] L. D. Bergman, V. Castelli, and C. Li. Progressive Content-Based Retrieval from Satellite Image Archives. *D-Lib Magazine*, 3(10), October1997.
- [4] B. Zhu, M. Ramsey, and H. Chen. Creating a Large-Scale Content-Based Air photo Image Digital Library. *IEEE Transactions on Image Processing*, 9(1):163–167, January2000.
- [5] J. C. French, A. C. Chapin, and W. N. Martin. An Application of Multiple Viewpoints to Content-based Image Retrieval. In *Proceedings of the 3rd ACM/IEEE-CS joint conference on digital libraries*, pages 128–130, Washington, DC, USA, 2003.
- [6] J. Z. Wang and Y. Du. Scalable integrated region-based image retrieval using irm and statistical clustering. In *Proceedings of the 1st ACM/IEEE-CS joint conference on digital libraries*, pages 268–277, 2001.
- [7] Y. Wang, F. Makedon, J. Ford, L. Shen, and D. Goldin. Generating fuzzy semantic metadata describing spatial relations from images using the r-histogram. In *Proceedings of the 4th ACM/IEEE-CS joint conference on digital libraries*, pages 202–211, 2004.
- [8] R. S. Torres, C. B. Medeiros, M. A. Goncalves, and E. A. Fox. A Digital Library Framework for Biodiversity Information Systems. *International Journal on Digital Libraries*, 6(1):3–17, February 2006.
- [9] Y.Rui and T.S.Hung, Image retrieval: current technique promising directions and open issues, *Journal of Visual Communication and Image Representation*, vol.10, pp.39-62, 1999.
- [10] M.Swain and D. Ballard, “Color indexing,” *Int. J. Comput. Vis.*, vol. 7, pp. 11-32, 1991.
- [11] G. Pass and R. Zabih, “Histogram refinement for content – based image retrieval,” in *Proc. IEEE Workshop on Applications of Computer Vision*, 1996, pp. 96-102.
- [12] J. Huang et al., “Spatial color indexing and applications,” *Int. J. Comput. Vis.*, pp.245-268, 1999.
- [13] A.Mojsilovic et al., “Matching and retrieval based on the vocabulary and grammar of color patterns,” *IEEE Trans. Image Processing*, vol. 9, pp. 38-54, Jan. 2000.
- [14] J.R. Smith and S. F. Chang, “VisualSeek: A fully automated content-based image query system”, *Proc. Int. Conf. on Image Proceeding*, 1996.
- [15] B. S. Manjunath and W. Y. Ma, “texture features for browsing and retrieval of image data”, *IEEE Trans. Patt. Anal. Mach. Int. Special Issue on Digital Libraries*, Vol. 18, No. 8, pp. 837-842, August 1996.
- [16] Abolfazl AleAhmad , Hadi Amiri , Ehsan Darrudi , Masoud Rahgozar , Farhad Oroumchian, Hamshahri: A standard Persian text collection, *Journal of Knowledge-Based Systems*, Vol. 22 No.5, p.382-387, Elsevier, July 2009.
- [17] H. Muller, W. Muller, D.M. Squire, S. M. Mailent and T. Pun, Performance evaluation in contentbased image retrieval: overview and proposals, *Pattern Recognition Letters*, vol.22, pp.593-601, 2001.
- [18] <http://hadoop.apache.org/>
- [19] D. Jeffrey, S. Ghemawat. MapReduce: Simplified data processing on large clusters. *Symposium on Operating System Design and Implementation*, 2004.