Segmentation for Image Indexing and Retrieval on Discrete Cosines Domain

Suhendro Y Irianto
Department of Informatics, Faculty of Computer Science
Darmajaya Informatics and Business Institute,
Jl. Z.A. Pagar Alam 93A, Bandar Lampung, Lampung, Indonesia
e-mail: syiranto1962@gmail.com, pasca@darmajaya.ac.id,

Abstrak

Dalam artikel ini digunakan teknik region growing untuk proses segmentasi pada citra DC. Dengan menggunakan teknik region growing pada citra DC akan mengurangi jumlah region yang akan digunakan dalam proses pengindeksan citra. Teknik ini akan mengurangi waktu pemrosesan dalam membuat kunci indeks. Teknik recursive region growing bukan merupakan metode atau teknik yang baru, akan tetapi apliksi pada citra DC untuk membuat kunci indeks cukup baru dan masih jarang digunakan oleh peneliti-peneliti lain. Hasil penelitian menunjukan bahwa citra DC yang telah disegmetasi menunjukan efektivitas lebih besar dibandingkan dengan citra DCT yang tidak dilakukan proses segmentasi dimana precisionya masing-masing sebesar 0.75 dan 0.59. Di samping itu, karena dalam penelitian menggunakan citra DC untuk pencarian citra (CBIR) maka hanya 1/64 saja dari 8x8 yang digunakan untuk indexing. Hal yang paling penting dalam penelitian ini adalah hanya digunakan satu keoefisen dari 64 koefisien DCT yang ada yaitu koefisien DC, sehingga banyak menghemat storage.

Kata kunci: *DC coefficients, region growing, segmentation, DC domain*

Abstract

This paper used region growing segmentation technique to segment the Discrete Cosines (DC) image. The classic problem of content Based image retrieval (CBIR) is the lack of accuracy in matching between image query and image in the database. By using region growing technique on DC image,it reduced the number of image regions indexed. The proposed of recursive region growing is not new technique but its application on DC images to build indexing keys is quite new and not yet presented by many authors. The experimental results show that the proposed methods on segmented images present good precision which are higher than 0.60 on all classes. So, it could be concluded that region growing segmented based CBIR more efficient compared to DC images in term of their precision 0.59 and 0.75, respectively. Moreover, DC based CBIR can save time and simplify algorithm compared to DCT images. The most significant finding from this work is instead of using 64 DCT coefficients this research only used 1/64 coefficients which is DC coefficient.

Keywords: *DC coefficients, region growing, segmentation, DC domain*

1. Introduction

In the field of digital imaging, image segmentation plays a vital role as a preliminary step for high level image processing. To understand an image, one needs to isolate the objects in it and find relation among them. The process of image partition referred as image segmentation [1]. In other words, segmentation is used to pull out the significant objects from the image. Deng [2] proposed a JSEG algorithm to segment the image based on multi scale 'Jimages'. The images which correspond to the measurements of local homogeneities at different scales are called as 'J-images'. The system has the ability to segment color textured images without supervision. First the colour inside the image is quantized to several classes. The pixels are then replaced by their corresponding colour class label which forms the class map of the image. A region growing method is then used to segment the image based on multiscale 'J-images'.

Histogram thresholding is one of the common techniques for monochrome image segmentation [2], [3]. This technique considers that an image consist of different regions corresponding to the grey level ranges. The histogram of an image can be separated using peaks (modes) corresponding to the different regions. A threshold value corresponding to the valley between two adjacent peaks can be used to separate these object. But one of the

weaknesses of this method is that, it ignores the spatial relationship information of the pixels. Adolfo [4] proposed a neural network based adaptive thresholding segmentation algorithm for monochrome image.

The main advantage of this method is that, it does not require a priori knowledge about number of objects in the image. To humans, an image is not just a random collection of pixels; it is a meaningful arrangement of regions and objects. There also exits a variety of images such as natural scenes, and paintings. Despite the large variations of these images, humans have no problem to interpret them. Image segmentation is the first step in image analysis and pattern recognition. It is a critical and essential component of image analysis system, is one of the most difficult tasks in image processing, and determines the quality of the final result of analysis. Image segmentation is the process of dividing an image into different regions such that each region is homogeneous.

Many content-based image retrieval (CBIR) systems have been developed since the early nineties. A recent article published by Smeulders [5], reviewed more than 200 references in this ever changing field. Readers are referred to that article and some additional references [4] for more information. Most of the CBIR projects aimed at general-purpose image indexing and retrieval systems focus on searching images visually similar to the query image or a query sketch. They do not have the capability of assigning comprehensive textual description automatically to images, because of the great difficulty in recognizing a large number of objects.

Researches done by [5]-[7] in the last few years used all coefficients DCT for image indexing and retrieval. This method caused low speed retrieving process due to more coefficients processed compare to the purposed method which is only 1/64 (a DC coefficients) of DCT coefficients. Therefore the purposed method performs 64 times faster in indexing and retrieving compared to DCT based image retrieval. Besides that the proposed method also provides a number of potential advantages and features, which can be briefly summarized as: (i) low complexity and low computing cost; (ii) high processing speed, (iii) easy to implement inside the JPEG compressed domain, and thus providing the additional advantage that compressed images can be directly retrieved without full decompression.

Some researchers have attempted to use machine-learning techniques for image indexing and retrieval [2], [5]. The system internally generated much segmentation or groupings of each image's regions based on different features combination, then learned which combinations best represented the semantic categories given as examples by the user. The system requires the supervised training of various parts of the image. Most of CBIR uses all regions in the image to match between image query and image in the database [7],[8]. Therefore their works are inefficient and time consuming due to the use more DCT coefficients used. In order to work out of these problems, this paper used Region Growing Segmentation to search and retrieve images from database. The remainder of the paper is organized as follows. Section 2 describes the research methods used in this work. Section 3 discusses the results and discussion. Section 4 describes the conclusions and some remark for the future works.

2. Research Methods
2.1. Database and JPEG

In this work more than 5,000 images used as ground truth (image database) which are collected from Internet and other sources. This work also used JPEG images, as it has many advantages compared to other format. In JPEG images, to compute the DCT images of a list of length n=8 and the 2D DCT of an 8 x 8 array. Rather than taking the transformation of the image as a whole, the DCT is applied separately to 8 x 8 blocks of the image, this called as a DCT block. In calculating a DCT block, actually the work does not actually have to divide the image into blocks. Since the 2D DCT is separable, it can be partitioned each row into lists of length 8, apply the DCT to them, rejoin the resulting lists, and then transpose the whole image and repeat the process.

DCT-based image compression relies on two techniques to reduce the data required. First is quantization, and the second is entropy coding of the quantized coefficients. Quantization is the process of reducing the number of possible values of a quantity, thereby reducing the number of bits needed to represent it. Entropy coding is a technique for representing the quantized data as compactly as possible. A function then developed to quantize images and to calculate the level of compression provided by different degrees of

quantization.   JPEG uses a combination of spatial-domain and frequency-domain coding. The image is divided into 8 x 8 blocks, each of which is transformed into the frequency domain using the discrete cosine transform (DCT). Each block of the image is thus represented by 64 frequency components. The signal tends to concentrate in the lower spatial frequencies, enabling high-frequency components, many of which are usually zero, to be discarded without substantially affecting the appearance of the image. The main source of loss of information in JPEG is a quantization of the DCT coefficients. A table of quantization coefficients is used, one per coefficient, usually related to human perception of different frequencies. The quantized coefficients are ordered in a zig-zag sequence, starting at the upper left (the DC component), since most of the energy lies in the first few coefficients. The final step is entropy coding of the coefficients, using either Huffman coding or arithmetic coding.

2.2 DC Images Technique

This work used only DC coefficients instead of using all DCT coefficients to construct image indexing and retrieving.   Discrete Cosines coefficient is one of 64 DCT coefficients as an image can be   constructed as 8 x 8 arrays or block. To build indexing keys, as illustration given the   N-blocks of an image to construct indexing keys by using only DC coefficients of every image in the database using the following equation:

$$H = \sum_{i=1}^{64} hi \qquad (1)$$

where $hi = \sum_{i=1}^{N} \dfrac{DC_i}{N}$,   is the ith DC coefficient of every block and N is the number of block in the image.   In building indexing keys, this approach only considers one DC coefficient in every block, which means it only needs 1/64 of the times needed by the full extraction and indexing process. Due to the limitations of hardware speed, the work only used a database of 5,000 images.

2.3 Region Growing Segmentation Technique

In this work, Region growing Image segmentation used as first key process in numerous applications of computer vision. It partitions the image into different meaningful regions with homogeneous characteristics using discontinuities or similarities of image components, the subsequent processes depend on its performance. In most cases, the segmentation of colour image demonstrates to be more useful than the segmentation of monochrome image, because colour image expresses much more image features than monochrome image. In fact, each pixel is characterized by a great number of combinations of R, G, B chromatic components. However, more complicated segmentation techniques are required to deal with rich chromatic information in the segmentation of colour images. A variety of segmentation techniques have been proposed in the literature. However, most techniques are kind of dimensional extension directly inherited from the segmentation of monochrome image [9]. The spatial compactness and colour homogeneity are two desirable properties in unsupervised segmentation, which lead to image-domain and feature-space based segmentation techniques.

The segmentation of images has always been a key problem in computer vision. Up to the early nineties bottom-up techniques like edge detection and split-and-merge algorithms were the primary focus of research. However, by that time people realized that perfect segmentation would not be possible without incorporation of higher level knowledge. Thus the focus shifted towards model based techniques like snakes   and methods based on geometric models [10].

Region growing algorithm starts from an initial, incomplete segmentation and try to aggregate the yet unlabelled pixels to one of the given regions. The initial regions are usually called seed regions or seeds. The decision whether a pixel should join a region or not is based on some fitness function which reflects the similarity between the region and the candidate pixel. As proposed in [11], the order in which the pixel is processed is determined by a global priority queue which sorts all candidate pixels by their fitness values. This approach elegantly mixes local (fitness) and global (pixel order) information. There is an abundance of literature on

image segmentation, and a number of review articles highlighting them. Methods also have been defined for post processing the low-level segmentation to further regularize the segmentation output, such as Markov Random Fields [12].

Automatic image segmentation is one of the primary problems of early computer vision, has been intensively studied in the past [11]. The existing automatic image segmentation techniques can be classified into four approaches, namely: thresholding techniques, boundary-based methods, region- based methods, and hybrid techniques. Region-based techniques rely on the assumption that adjacent pixels in the same region have similar visual features such as grey level, colours value, or texture. A well-known technique of this approach is split and merges [13]. Obviously, the performance of this approach largely depends on the selected homogeneity criterion. Instead of tuning homogeneity parameters, the seeded region growing (SRG) technique is controlled by a number of initial seeds [14]. Given the seeds, SRG tries to find an accurate segmentation of images into regions with the property that each connected component of a region meets exactly one of the seeds. Moreover, high-level knowledge of the image components can be exploited through the choice of seeds. This property is very attractive for semantic object extraction toward content-based image database applications. However, SRG suffers from another problem: how to select the initial seeds automatically for providing more accurate segmentation of images. The algorithm of region growing segmentation technique can be describes as follows:

Input : **image I**
create an (empty) set S of segments
    stage 0: i:=0;
  for all DC coefficients P in I
      create a new segment R_p of level 0
        (consisting only of P)
   put R_p in S
  repeat
   stage i:
   for all segments R_i of level i in S
    repeat
           find a segment $\acute{R}^j$ of level j ≤ i in S,
             R_i and $\acute{R}^j$ are neighboured and
             R_i Ú $\acute{R}^j$ is homogeneous enough
      remove R_i and $\acute{R}^j$ from S
      redefine R_i:= R_i Ú $\acute{R}^j$ of level i+1
    until no such $\acute{R}^j$ can be found
    add R_i to S
  i:=i+1
  until stage i-1 has created no new segment

In Markov Random Fields [12] algorithm for image segmentation has been drawn considerable attention due to its ability to integrate texture, colour, and edge information in an optimal manner to devise a robust labeling of the image into homogeneous regions [15]. These methods still depend on the assumption that the pixels belonging to the object of interest share a common set of low-level image attributes, thereby allowing the object to be extracted as a single entity. If an object is composed of multiple regions of differing texture or colour then the object is divided into regions corresponding to each of these, and these sub regions must then be re-assembled through some contextual-based post processing to segment the complete object from the image. By employing an additional constraint upon the segmentation that encourages it to find a human, it would be possible to only extract the regions corresponding to the human in the image. This additional constraint can be provided through information regarding the desired shape of the final retained region.

The proposed method of this research is applying region growing segmentation on DC images. The region growing segmentation is not new method to segment an image [2],[9], but applying it to DC image is a quite new and have not done before by other authors. This proposed method has three significant merits namely reduce storage usage, more accurate and effective in matching of image region.
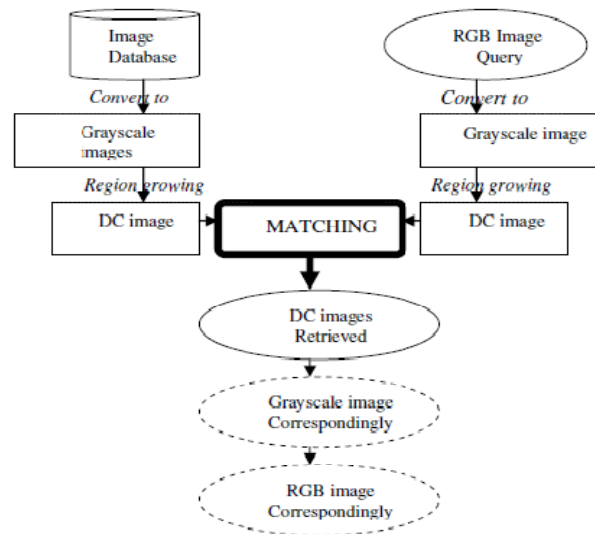
Figure 1. The query process of region growing based image retrieval system

The query process has been established as follows: Firstly, user queries an RGB image in the system, RGB image then is converted to grayscale image. Secondly, by utilizing region growing algorithm this image will be segmented into meaningful regions. Finally based on these regions, the minimum distance between them will be calculated and compared to the image regions in the database. Figure.1 illustrates query process diagram of the image retrieval system proposed. Once a query is specified, it scores each segmented image based on how closely and satisfies to the images in the database. The score *i* for each atomic query (segmented image) is calculated by using the following equation.

$$d(Hq - Hk) = \frac{1}{64} \sum_{i=0}^{63} |hqi - hki|$$

(2)

where H_q and H_k are query indexing key and image indexing keys, respectively. The distance is equal to 0, if the image is identical in all the regions. We then rank the images according to overall score and return to the twenty best matches.

3.   Results and Discussion

In this experiment, 5,000 of JPEG images used and consist of 10 classes which consist of *bear, bike, building, car, cat, flower, model/celebrity, mountain, sky*, and *texture*. The work evaluates only the top twenty images ranked in terms of the similarity measures by using precision and recall parameters. Precision is the ratio of the number of relevant images retrieved to the total number of retrieved (both irrelevant and relevant) images retrieved. Whilst, Recall is the ratio of the number of relevant images retrieved to the total number of relevant images in the database.

$$Precision = \frac{\sum relevant\ images\ retrieved}{\sum number\ of\ images\ retrieved}$$

$$Recall = \frac{\sum relevant\ images\ retrieved}{\sum relevant\ images\ in\ the\ class}$$
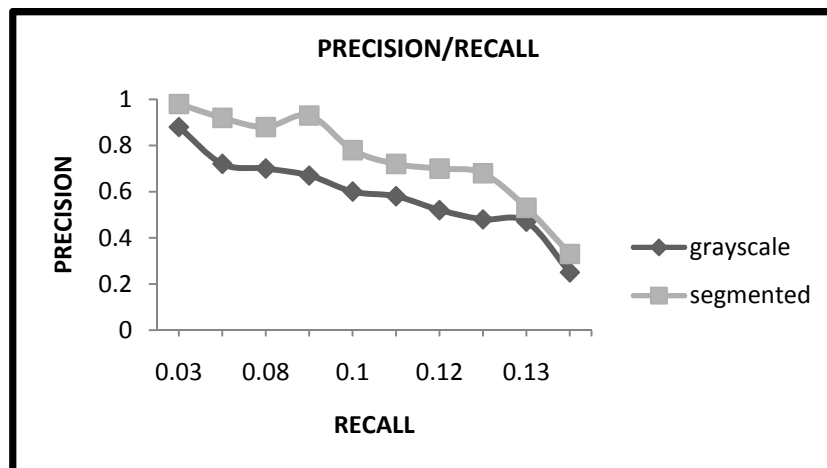
Figure 2. The effectiveness of image retrieval by grayscale and segmented images based methods produced by the work

The excellent precision of 0.98 has been demonstrated by applying region growing   technique for bear class, and worst precision of 0.25 for texture class. Interesting result,   grayscale method shows that the best precision of 0.88 also for bear, and the lowest precision of 0.32 for cat class as shown in table 1.   The experimental result shows that the proposed method on segmented images presents good precision which are higher than 0.50 on all classes except for texture class. Figure 2. shows that all precision and recall values of segmented images greater than grayscale image (no segmentation applied). This results demonstrate that region growing segmentation offers more effective method for image indexing and retrieving compared to un-segmented image or grayscale image with full DCT coefficients as some previous researchers done [1], [3], [6], [14]. Since the proposed method used only object without its background as well as used 1/64 of image size, so that it will save time and huge of storage in image processing as general, especially for Content Based Image Retrieval. For further precision and recall detail , it can be seen in table 1.

Table. 1. Precision and recall on grayscale and segmented images

Class	Grayscale images		Segmented images	
	Precision	Recall	Precision	Recall
Bear	0.88	0.03	0.98	0.01
Bike	0.70	0.08	0.92	0.02
Build	0.60	0.10	0.88	0.03
Cars	0.58	0.11	0.93	0.04
Cat	0.25	0.19	0.78	0.06
Flower	0.47	0.13	0.72	0.07
Model	0.67	0.08	0.70	0.08
Mount	0.48	0.13	0.68	0.09
Sky	0.72	0.07	0.53	0.12
Text	0.52	0.12	0.33	0.17
Average	**0.60**		**0,75**	

Table 1 also illustrates the average (mean) of precision 0.60 and 0.75 for grayscale and segmented images, respectively. The experiment results show that applying of region growing technique   (segmented images) gives better average precision of ten classes in the database. The most significant of this work, it used only DC coefficients instead of using all DCT

coefficients or pixels which saved 1/64 storage (image size) to construct image indexing and retrieval and to segment images which some authors used previously [6], [11], [16]. DC coefficient is one of 64 DCT coefficients as an image can be constructed as 8 x 8 array or block. To build indexing keys, as illustration given the N-blocks of an image, it can be used to construct indexing keys by using only DC coefficients of every image in database using this equation:

$$H = \sum_{i=1}^{64} hi \tag{3}$$

where $hi = \sum_{i=1}^{N} \dfrac{DCi}{N}$, is the ith DC coefficient of every block and N is the number of block in the images. In building indexing keys, this approach only considers one DC coefficient in every block, which means it only needs $1/64^{th}$ of the time taken by the full extraction, indexing process. Due to the limitations of hardware speed, the work only used a database of 5,000 images. Compare to the previous work, this research has two merits simpler algorithm and faster processing. Whilst, the effectiveness image retrieval of this work and most other image retrieval researches is not comparable due to the result of image retrieval has very much influenced by characteristics of image database used in the research.

4. Conclusion and Future Works

New approach has been proposed for an image retrieval system based on region growing segmentation on DCT compress domain. It is presented as a different way to develop image indexing by using of DCT descriptors. The method has been carried out for compressed images database to verify its performance in JPEG standard stream line. The proposed method of region growing segmentation on DC images offers huge storage and time saving for Image indexing and retrieving.

From this work, it could be concluded that segmentation, while imperfect, is an essential step and very useful in building indexing keys. In summary, this indexing key method is a promising method for image retrieval on segmented image on compress domain. This new approach could be used for image indexing by other segmentation methods. For the near future, it will be used another segmentation approaches such as Support Vector Machine, Fuzzy logic, and Split Merge to improve speed of image indexing and Retrieval.

Query

Grayscale

Segmented

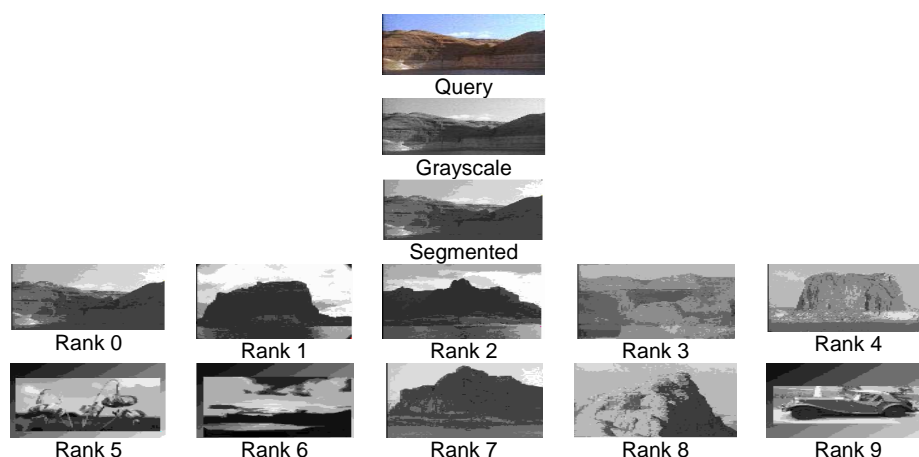|				
Rank 0	Rank 1	Rank 2	Rank 3	Rank 4
Rank 5	Rank 6	Rank 7	Rank 8	Rank 9

Figure 3. Result examples of the system for segmented images retrieved with RGB image query, RGB image converted into grayscale, the grayscale image then partitioned by region growing technique.

References

[1]   N. A. Mat-Isa, M. Y. Mashor, N. H. Othman. Seeded Region Growing Features Extraction Algorithm: Its Potential Use In Improving Screening For Cervical Cancer. *International Journal of the Computer, The Internet And Management.* 2005; 13(1): 61-70.

[2]   T.P. Minka, R.W. Picard.1997. Interactive Learning Using a Society of Models. *Pattern Recognition.* 1997; 30(3): 565-573.

[3]   Mehmet Sezgin, Bu¨lent Sankur. Survey over Image Thresholding Techniques and Quantitative Performance Evaluation. *Journal of Electronic Imaging.* 2004;13(1): 146–165.

[4]   Adolfo Martı´nez-Uso, Filiberto Pla, Pedro Garcı´a-Sevilla. Unsupervised Colour Image Segmentation by Low-Level Perceptual Grouping. *Pattern Analysis and Aplication.* Springer. 2011. 123-132. DOI 10.1007/s10044-011-0259-1.

[5]   A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain. Content-Based Image Retrieval at the End of the Early Years. *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 2000; 22(12): 1349-1380.

[6]   Jun Sun, Yan Wang1,Xiaohong Wu, Xiaodong Zhang, Hongyan Gao. A New Image Segmentation Algorithm and Its Application in Lettuce Object Segmentation. *TELKOMNIKA.* 2012; 10(3): 557-563. e-ISSN: 2087-278X

[7]   Nidhi. Singhai, Shishir K. Shandilya. A Survey On: Content Based Image Retrieval Systems. *International Journal of Computer Applications.* 2010; 4(2): 22-26.

[8]   Ritendra Datta, Dhiraj Joshi, Jia Li, James Z. Wang. Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys.* 2008; 40(2): Article 5.

[9]   J. Philbin, M. Isard, J. Sivic, A. Zisserman. *Descriptor Learning For Efficient Retrieval.* European Conference on Computer Vision. 2010; 5(2): 213-217.

[10]   A.K. Jain. Fundamentals of Digital Image Processing. Upper Saddle River. NJ: Prentice Hall. 1989.

[11]   Anny Yuniarti, Anindhita Sigit Nugroho, Bilqis Amaliah, Agus Zainal Arifin. Classification and Numbering of Dental Radiographs for an Automated Human Identification System. *TELKOMNIKA.* 2012; 10(1): 137-146.

[12]   N. A. Mat-Isa, M. Y. Mashor, N. H. Othman. Seeded Region Growing Features Extraction Algorithm: Its Potential Use In Improving Screening for Cervical Cancer. *International Journal of the Computer, The Internet and Management.* 2005: 13;(1): 61 -70.

[13]   Kanchan Deshmukh,G. N. Shinde. An Adaptive Neuro-Fuzzy System For Color Image Segmentation. *Journal Indian Institute of Science.* 2006;86(1): 493–506.

[14]   R. Adams, L. Bischof. Seeded Region Growing. *IEEE Transaction on Pattern Analysis and Machine Intelligent.* 1994;16(6): 641-647.

[15]   Z. Tu,S.-C. Zhu. Image Segmentation by Data-driven Markov Chain Monte Carlo. *IEEE Transaction on Pattern Analysis.* 2002; 24(5): 657–673.

[16]   Mustafa Ozden, Ediz Polat. A Color Image Segmentation Approach For Content-Based Image Retrieval. *Pattern Recognition.* 2007; 40(2): 1318 – 1325.