

# Image Segmentation and Detection using Watershed Transform and Region Based Image Retrieval

Niket Amoda<sup>1</sup>, Ramesh K Kulkarni<sup>2</sup>

<sup>1</sup>Department of Electronics and Telecommunication Engineering,  
Vivekanand Institute of Technology, University of Mumbai,  
M.G. Road Fort, Mumbai, India

**Abstract:** Early image retrieval techniques were based on textual annotation of images. Annotating images manually is a cumbersome and expensive task for large image databases, and is often subjective, context-sensitive and incomplete. Content based image retrieval, uses the visual contents of an image such as color, shape, texture, and spatial layout to represent and index the image. The Region Based Image Retrieval (RBIR) system uses the Discrete Wavelet Transform (DWT) and a Watershed Segmentation algorithm to segment an image into regions. Though k-means clustering algorithm is very fast and simple to implement, but it provides only coarse image segmentation. Better retrieval results can be expected by employing a more sophisticated segmentation technique. For this purpose, a novel Texture Gradient based Watershed Segmentation technique is developed. The Watershed Transform is a well established tool for the segmentation of images. However, it is often not effective for textured image regions that are perceptually homogeneous. In order to properly segment such regions the concept of the Texture Gradient is introduced and is implemented using a Non Decimated Wavelet Packet Transform. A marker location algorithm is subsequently used to locate significant homogeneous textured or non textured regions. A marker driven Watershed Transform is then used to properly segment the identified regions. The experimental results demonstrate the superiority of this technique over k-means clustering.

**Keywords:** Content based image retrieval, K-Means Algorithm, Discrete Wavelet Transform, Region Based Image Retrieval.

## 1. INTRODUCTION

Early image retrieval techniques were based on textual annotation of images. Content based image retrieval, uses the visual contents of an image such as color, shape, texture, and spatial layout to represent and index the image. In typical content based image retrieval systems, the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database.

To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities / distances between the feature vectors of the query

example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search the image database.

Recent retrieval systems have incorporated users' relevance feedback to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results.

A visual content descriptor can be either global or local. A global descriptor uses the visual features of the whole image, whereas a local descriptor uses the visual features of regions or objects to describe the image content. To obtain the local visual descriptors, an image is often divided into parts first.

Some of the widely used techniques for extracting color, texture, shape and spatial relationship features from images are now described briefly.

There are two generally used methods for image segmentation: discontinuity based method and similarity based methods.

### 1.1 Discontinuity Detection Methods

The interface between two different contiguous homogeneous regions is usually marked by a discontinuity in grey-level, colour or texture. Grey-level discontinuities can be located using simple edge detectors that look for local gradient maxima. The gradient of an image function  $f(x, y)$  at location  $(x, y)$  is given by:

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (1)$$

### 1.2 Similarity Based Methods

Similarity based methods of segmentation are based on similarity comparison between pixel and regions. There are four commonly used similarity based segmentation techniques: Thresholding, Region Growing, Region Splitting and Merging and Clustering.

## 2. REGION BASED IMAGE RETRIEVAL

The Region Based Image Retrieval (RBIR) system uses the Discrete Wavelet Transform (DWT) and a Watershed

Transform in place of k-means clustering algorithm to segment an image into regions. Each region is represented by means of a set of features and the similarity between regions is measured using a specific metric function on such features. The implementation of RBIR can be divided into two parts: Image Pre-processing and Image Retrieval

### 2.1 Image Pre-processing

The detailed algorithm for the pre-processing stage is given below:

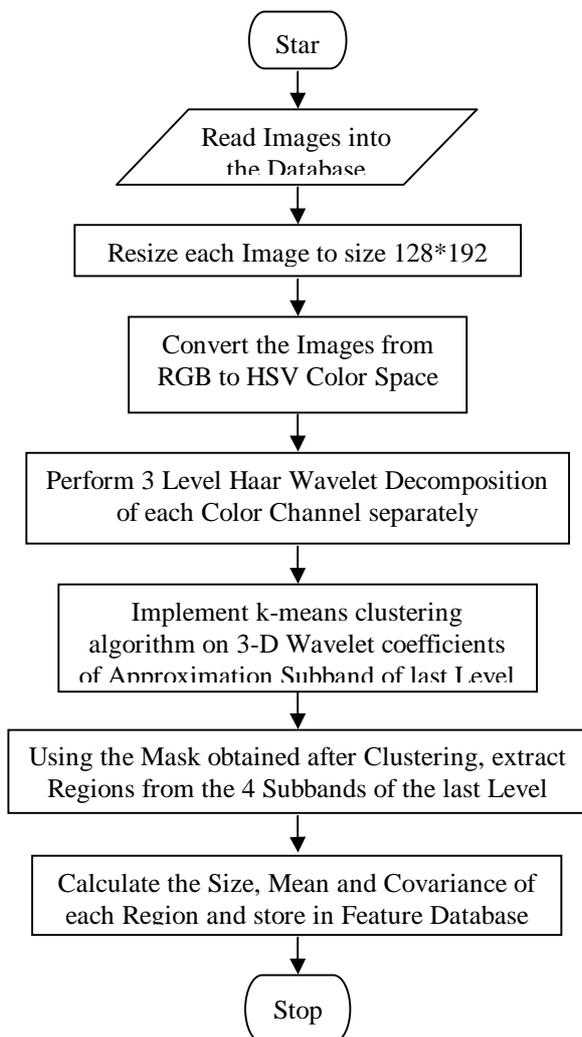


Figure: 1 Flow Diagram

The algorithm requires that all the images in the database and the query image be of the same size. A size of 128\*192 was chosen for all of the images. If the image had a different size, it was first resized to 128\*192 and then the pre-processing operations were carried out on the image.

In the RBIR application, each image is divided into the corresponding color channels (i.e. H, S and V) and the DWT was applied separately to each color channel. The  $j$ -th Wavelet coefficient of subband B ( $B \in \{LL, LH, HL, HH\}$ ), where L stands for “low” and H for “high”)

### 2.2 Watershed Transform

The Watershed Transform is a unique technique for segmenting digital images that uses a type of region growing method based on an image gradient. The concept of Watershed Transform is based on visualizing an image in three dimensions: two spatial coordinates versus gray levels. In such a “topographic” interpretation, we consider three types of points:

- A. Points belonging to a regional minimum.
- B. Points at which a drop of water, if placed at the location of any of those points, would fall with certainty to a single minimum.
- C. Points at which water would be equally likely to fall to more than one such minimum.

For a particular regional minimum, the set of points satisfying condition (B) is called the *catchment basin* or *watershed* of that minimum. The points satisfying condition (C) form crest lines on the topographic surface and are termed *divide lines* or *watershed lines*. The principal objective of segmentation algorithms based on these concepts is to find the watershed lines.

#### 2.2.1 Advantages of Watershed Transform

The Watershed Transform effectively combines elements from both the discontinuity and similarity based methods. Since its original development with grey-scale images, the Watershed Transform has been extended to a computationally efficient form (using FIFO queues) and applied to colour images. The main advantages of the Watershed method over other previously developed segmentation methods are

- A. The resulting boundaries form closed and connected regions. Traditional edge based techniques most often form disconnected boundaries that need post-processing to produce closed regions.
- B. The boundaries of the resulting regions always correspond to contours which appear in the image as obvious contours of objects. This is in contrast to split and merge methods where the first splitting is often a simple regular sectioning of the image leading sometimes to unstable results.
- C. The union of all the regions forms the entire image region.

#### 2.2.2 Disadvantage of Watershed Transform

The main disadvantage of the Watershed Transform is that for most natural images it produces excessive *over-segmentation* as shown in Figure 2



Figure: 2 Result of simple Watershed Transform of Lena test image

### 2.3 Image Retrieval

The detailed algorithm for the Image Retrieval phase is given below.

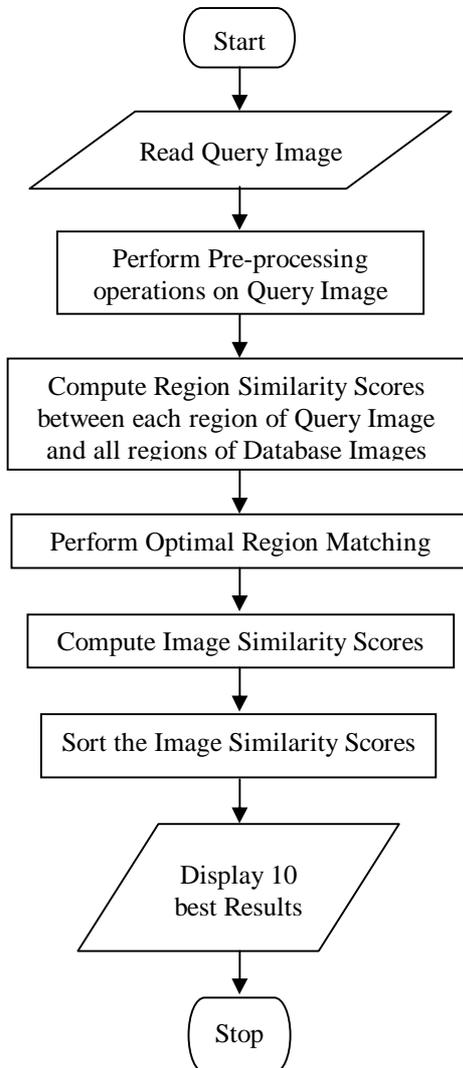


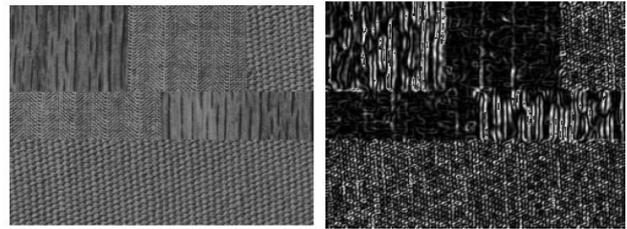
Figure: 3 Flow Diagram

After reading the query image, the same pre-processing operations of image resizing, RGB to HSV conversion, DWT decomposition, Watershed clustering and feature extraction must be performed on the query image. At the end of the pre-processing operation, the sizes, mean vectors and covariance matrices of the regions of the query image would be obtained.

### 3. IMPEMENTATION

The watershed transformation is performed on a gradient image  $g$  extracted from the original image. However the problem with the conventional intensity gradient is it is not able to detect the interfaces between homogeneously textured image regions. This is because the gradient image highlights the variations within the textures rather than showing the change between textured regions, as shown in Figure 4 below. A Texture Gradient is therefore

required to detect the texture boundaries.



4(a) Original Texture Image 4(b) Intensity Gradient

Fig: 4 Gradient of Texture Image

A marker based method is used to control the over-segmentation. The various steps involved in marker controlled watershed segmentation are shown below.

**Step 1:** The first step is to read the combined Texture and Intensity Gradient.

**Step 2:** Next step is to compute the Foreground Markers. These are connected blobs of pixels within each of the objects in the image. A variety of procedures could be applied to find the Foreground Markers. In the present work, morphological techniques called "opening-by-reconstruction" and "closing-by-reconstruction" are used to "clean" up the image. These operations will create flat maxima inside each object. Opening-by-reconstruction is erosion followed by a morphological reconstruction whereas closing-by-reconstruction is dilation followed by morphological reconstruction. These operations will remove small blemishes without affecting the overall shapes of the objects. Good Foreground Markers can be obtained by computing the regional maxima of the resulting Gradient Image.

**Step 3:** Some of the mostly-covered and shadowed objects will not be marked in the previous step, which means that these objects will not be segmented properly in the end result. Also, the Foreground Markers in some objects go right up to the objects' edge. Hence it is necessary to clean the edges of the marker blobs and then shrink them a bit. This can be done by a closing followed by erosion. This procedure tends to leave some stray isolated pixels that must be removed.

**Step 4:** Next the background locations need to be marked. In the cleaned-up image, the dark pixels belong to the background, so thresholding is a suitable operation to start with.

**Step 5:** The background pixels will be in black, but ideally the background markers shouldn't be too close to the edges of the objects that are being segmented. So the next step is to "thin" the background by computing the "skeleton by influence zones", or SKIZ, of the foreground. This can be done by computing the watershed transform of the distance transform of thresholded image, and then looking for the watershed ridge lines of the result.

**Step 6:** The next step is to modify the Gradient Image so that it has regional minima only in certain desired locations i.e. at the Foreground and Background Marker pixels.

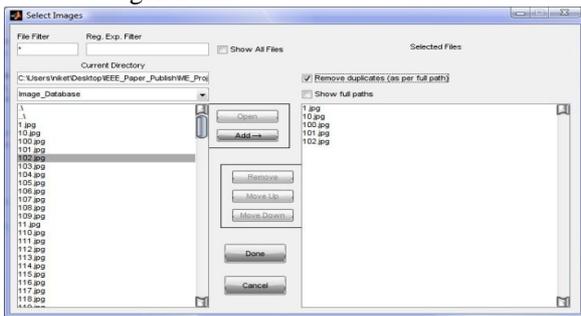
**Step 7:** The final step is to give this Modified Gradient Image as input to the Watershed Transform Algorithm.

After the optimal region assignment has been performed, the next step is to compute the Image Similarity Score. This score is obtained by simply adding the region similarity scores of the matched regions. The final step is to sort the image similarity scores so obtained and then display the images having the least distance from the query image.

**3.1 Adding Images to the Database:**

The steps involved in adding images to the database are:

- 1) Run the ‘Feature Database Generation Population’ program. A Matlab GUI appears as shown in Figure 5
- 2) Select the images from the folder Image\_Database and click on ‘Add’ to generate a feature database.
- 3) The images selected will appear in the listbox.
- 4) Click on the ‘Done’ after selecting the images to be added to the database.
- 5) Feature Database of the image will be generated.

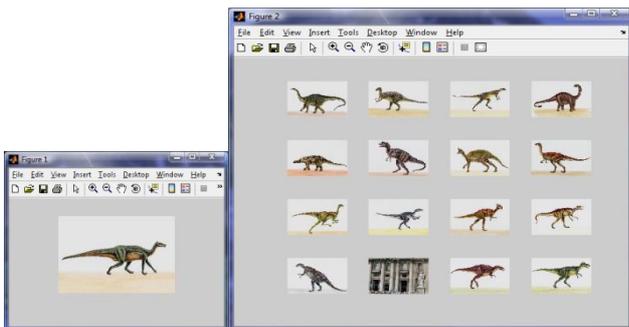


**Figure: 5** Feature Database population

**3.2 Image Retrieval :**

The steps involved in searching for images are:

- 1) Start the ‘Region Based Image Retrieval’ program. A GUI appears as shown in Figure 6 (a).
- 2) Select a query image from the folder Image\_Query.
- 3) The query image will be displayed Figure 6 (b) along with the 16 matches of the most similar images available in the database.



**Figure: 6(a)**

**Figure: 6(b)**

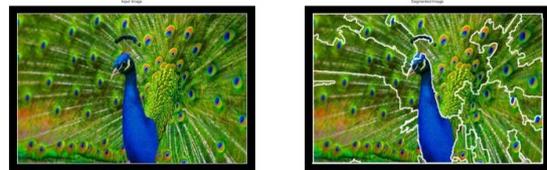
**Figure: 6** Region Based Image Retrieval

**4. RESULT**

The Figure7 below shows the results of segmenting the ‘Lena’ test image and the ‘Peacock’ test image using the texture gradient based watershed segmentation technique. The results show that this technique produces effective intensity and texture based segmentation of natural images.



**Figure: 7(a)** Lena Test Image



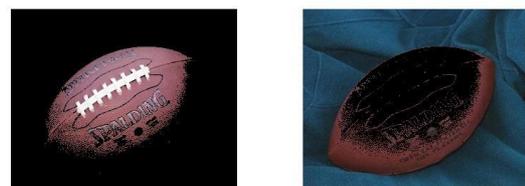
**Figure: 7(b)** Peacock Test Image

**Figure: 7** Result of segmenting the Lena and Peacock Test Images

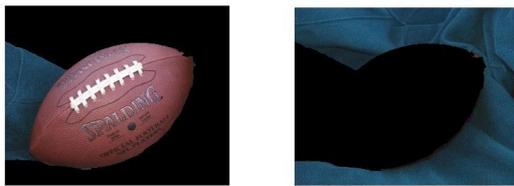
**4.1 Comparison of Watershed segmentation with K-means clustering using a ‘football’ test image**



**Figure: 8(a)** Football Test Image



**Figure: 8(b)** Regions obtained using k-means clustering with k =2



**Figure: 8 (c)** Regions obtained using Watershed Segmentation

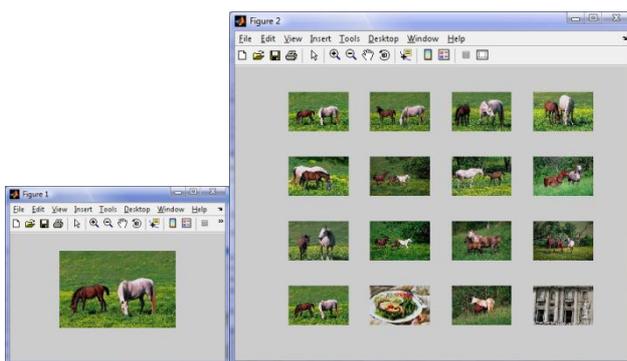
**Figure: 8** Results of segmentation “Football” Test Image

The experimental setup used was a PC with a 3.0 GHz Pentium 4 Processor and 512 MB of DDR 400 RAM running MATLAB 7 on Windows XP Professional.

Image	Size	Time Taken for Segmentation (seconds)	
		<i>K-means Clustering</i>	<i>Watershed Segmentation</i>
Football	256*320* 3	1.6410 (k=2)	4.3590
Lena	512*512* 3	7.0780 (k=2)	24.3600

To test the RBIR application, a database consisting of 180 general images, was used. The 180 images could roughly be categorized into 9 groups, each group consisting of 20 similar images. In addition 9 query images, each query corresponding to one of the groups were taken.

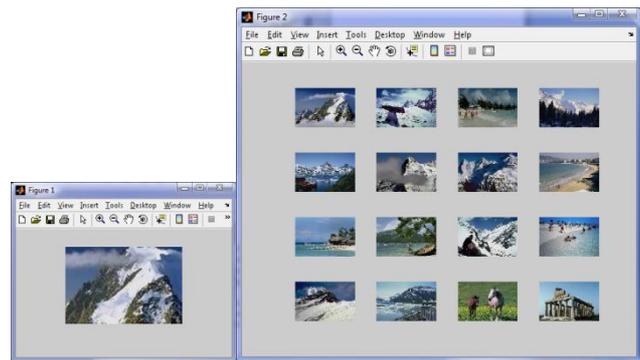
Figure 9 shows an example of the results obtained with the RBIR application. From a semantic point of view, the results obtained are particularly good i.e. all the images in this particular example are of horses.



**Figure: 9** RBIR Results

#### 4.2 Partial Match Queries

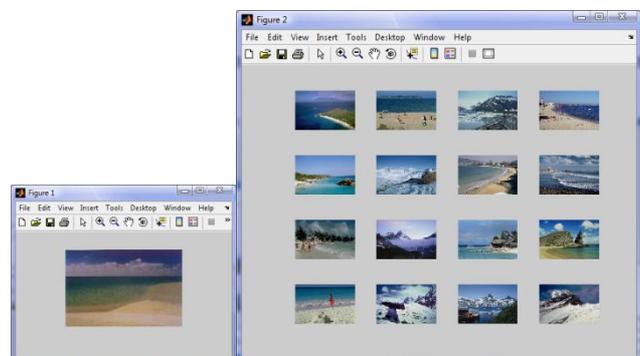
A partial match query is a query that specifies only part of the image. In Figure 10, the query image is obtained by cropping a database image. As the results show, the RBIR application gave the complete image as the very first match.



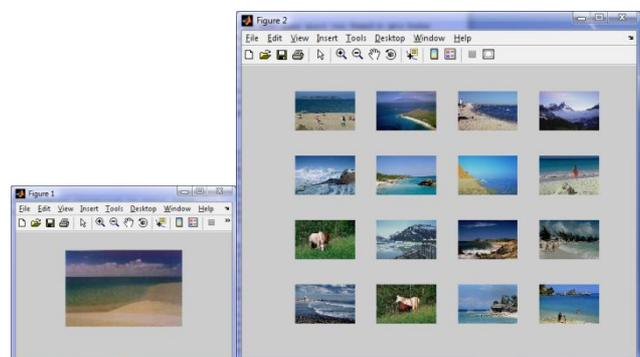
**Figure: 10** Partial Match Querying

#### 4.3 Scanned Queries

If the query image is scanned, it may suffer artifacts such as color shift, poor resolution, dithering effects, and misregistration. To consider the effect of scanned images on the retrieval effectiveness, the query image was first printed and then subsequently scanned. The scanned image appeared fuzzier, darker and slightly misregistered compared to the original. Figures 11 (a) and (b) show the results obtained with the original query image and the scanned query image respectively. It can be observed that there is a slight degradation in the quality of the results obtained.



**Figure: 11 (a)** Result with Original Query

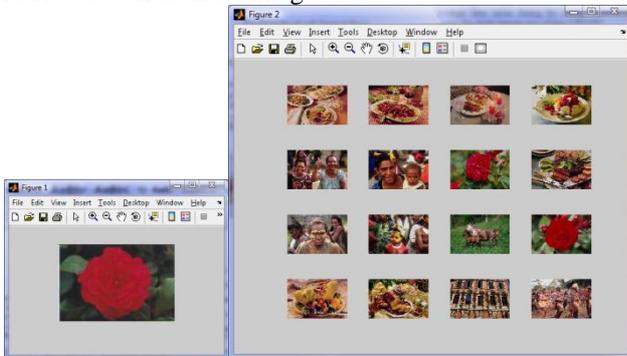


**Figure: 11 (b)** Result with Scanned Query

**Figure: 11** Scanned Queries

#### 4.4 Difficult Queries

The effectiveness of the RBIR application is confirmed when considering “difficult” queries, i.e. queries having a low number of similar images in the database. Figure 12 shows the results for a query having only two similar images in the database. The RBIR system is able to retrieve both of these images.



**Figure: 12** Difficult Queries

#### 4.4 Search Time

For an image retrieval application, the time taken for retrieval is an extremely important parameter. On the other hand, the time taken for pre-processing is not as important since the pre-processing operations have to be carried out only once.

The first entry in table below shows the average time taken to perform the pre-processing operations on the database of 180 images of size 128\*192. The second entry shows the average time taken during the image retrieval phase. Here again, the query image was of size 128\*192. The experimental setup consisted of a PC with a 3.0 GHz Pentium 4 processor and 512 MB of DDR 400 RAM running MATLAB 7 on Windows XP Professional.

Pre-processing Stage	25.2 seconds
Image Retrieval Stage	1.7 seconds

### 5. CONCLUSIONS

From the above experiment, it is clear that the watershed segmentation technique takes about three times the time taken by the k-means clustering technique. Although the HSV color space was found to give better results compared to the RGB color space in, in our experiments the RGB and HSV color spaces were found to give almost equivalent results. Eventually, it was decided to use the HSV color space because it gave better results than the RGB color space in case of “difficult queries”. The figure below shows that when using the RGB color space, only one of the two matches was retrieved. On the other hand, in the HSV color space, both the matches were retrieved.

#### REFERENCES

[1] D. Lowe, “Object recognition from local scale-invariant features,” in ICCV, 1999, pp. 1150–1157.

[2] Y.J. Zhang “A survey on evaluation methods for image segmentation”, Pattern Recognition 29 (8) (1996) 1335 - 1340

[3] A. Jain, “Data clustering: 50 years beyond k-means,” Pattern Recognition Letters, vol. 31, no. 8, pp. 651 – 666, June 2010.

[4] W. Zhao, H. Ma, Q. He, "Parallel K-Means Clustering Based on MapReduce," in: Cloud Computing, vol. 5931, pp. 674-679, 2009.

[5] W. D. Arthur, S. Vassilvitskii, “K-means++: the Advantages of careful seeding,” in Proc. 2007 Symposium on Discrete Algorithms, pp.1027-1035.

[6] Rafael C. Gonzalez, Richard E. Woods, " Digital Image Processing" , Second Edition, Prentice Hall Upper Saddle River, New Jersey 07458, TA1632.G66 2001, 698-740

[7] Fast Multiresolution Image Querying, International Conference on Computer Graphics and Interactive Techniques, 1995: Charles E. Jacobs, Adam Finkelstein, David H. Salesin

[8] Content-based Image Retrieval, A report to the JISC Technology Applications Programme, 1999: John Eakins, Margaret Graham

[9] Fundamentals of Content-based Image Retrieval, Multimedia Information Retrieval and Management - Technological Fundamentals and Applications, Springer, 2002: Dr. Fuhui Long, Dr. Hongjiang Zhang, Prof. David Dagan Feng

[10] Image Retrieval – Current techniques, Promising directions and Open issues, Journal of Visual Communication and Image Representation, 1999: Yong Rui, Thomas S. Huang, Shih-Fu Chang

[11] Wavelet Based Texture Analysis and Segmentation for Image Retrieval and Fusion, Thesis, University of Bristol, 2002: Paul R. Hill

[12] WINDSURF: A Region Based Image Retrieval System, Proceedings of the 10th International Workshop on Database & Expert Systems Applications, 2000: Ilaria Bartolini, Paolo Ciaccia, Marco Patella