MCM-CBIR: Multi Clustering Method for Content Based Image Retrieval

Hadjer LACHEHEB
LRIA Laboratory, Computer Science Department, USTHB
Algiers, Algeria
hlacheheb@usthb.dz

Saliha AOUAT
LRIA Laboratory, Computer Science Department, USTHB
Algiers, Algeria
saouat@usthb.dz

Izem HAMOUCHENE
LRIA Laboratory, Computer Science Department, USTHB
Algiers, Algeria
ihamouchene@usthb.dz

ABSTRACT

Image retrieval systems are designed to provide the ability of searching and retrieving images in huge image databases. A content based image retrieval system (CBIR) is used to offer such tasks based on the content of the image. In this paper we propose a new method of CBIR system based on a learning technique. Our method uses k-means clustering to reduce data and to improve the system performance. The specificity of our approach is the use of each feature vector separately in the clustering process in order to obtain different clustering on the same database, differently to other approaches that combine features vectors to cluster the database. For this reason we call it multi-clustering approach. The advantage of this approach consists in keeping the performance of the features and getting several views of the database due to the separation of features. The experimental results show the efficiency of our approach.

Keywords

CBIR, visual features, color, texture, shape, k-means clustering.

1. INTRODUCTION

The increasing number of digital images makes the information management hardest. CBIR is the process of retrieving images from a huge database on the basis of visual features. These features are automatically extracted, such as color, texture and shape. The greater parts of CBIR systems are based on a typical architecture shown (see Fig.1). This architecture includes two phases. The first phase is the offline phase or Data insertion [Tor]. This phase consists of the extraction of features vectors or descriptors of each image in the database. The second phase is the online or Query processing [Tor]. In this phase we extract the query feature vectors and compare it to all feature vectors of the database. We compare the query image to the other images in the database with computing a similarity measure between their feature vectors. Afterwards, the obtained images will be ordered following a decreasing order of similarity. The visual contents or Features are the representation of any distinguishable characteristic of an image [Sub02].

These features require three levels: low, middle and high. Low level features are the visual content that can be extracted from the information obtained in the pixel level such as color, texture, and shape.

One of the attractive parts in an image is the color. The color space is the identification of the color; it is generally represented using three elements. For instance, RGB (red, green, blue) color space, HSV (hue, saturation, value), CIE L*a*b*[Col04].

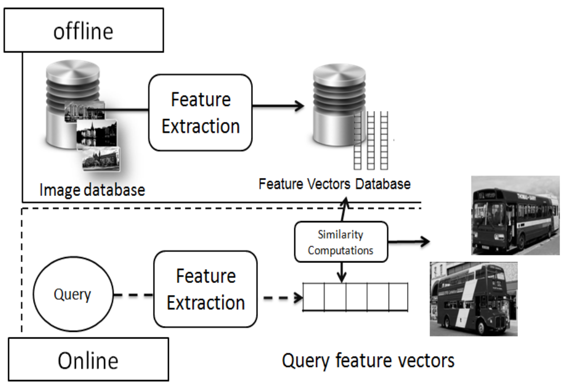

Figure 1. Typical architecture of CBIR systems.

The texture is a powerful feature that is neglected generally by a lot of CBIR systems. But, until now a clear definition about texture does not exist. However, we can define it as the repeated model that has the proprieties of homogeneity, coarseness, contrast, directionality, line-likeness, regularity, roughness [Zha12]. The other important descriptor or feature is a shape.

Shape is the characteristic surface represented by a contour or an outline. Shape descriptors need a good segmentation into regions or objects. Zhang and Lu [Zha04] classify shapes feature methods for boundary based and region based. All these descriptors need similarity metrics for the purpose of comparing the query image features vectors and feature vectors of the images in the database. A similarity measure determines the distance between the feature vectors (low level features) representing the images. Some of famous measurements are Euclidean distance, Minkowski distance, Manhattan distance, and histogram intersection [Swa91]. Many researchers use other methodologies or tools of artificial intelligence like machine learning. Arthur Samuel [Sam59] defines the machine learning as a field that gives computers the ability of learning without being programmed. Two major categories are introduced in machine learning which are supervised and unsupervised learning. In supervised learning a supervisor is going to teach the computer on the other side in the unsupervised learning the computer learns by itself. Clustering is a famous technique of unsupervised learning. One of the commonly used clustering techniques is K-means clustering. K-means is used to find K different clusters in a database of N objects, where similarities between objects in the same cluster are minimized, and between objects of the other clusters are maximized [Tan05], [Jai11].

This paper is organized as follows:

Section 2 is an overview about related works dealing with CBIR. The third section contains the details about our proposed method. In section 4 we show the experimental results and comments about their efficiency after applying our proposed technique.

2. RELATED WORKS

Many general-purpose image retrieval systems have been developed. A lot of famous systems like QBIC [Fal94], Photobook [Pen96], Blobworld [Car02], Virage [Gup97], VisualSEEK and WebSEEK [Smi96]. An important part of new approaches start to use key point features one of the well known proposition on that are SIFT descriptors proposed by Lowe in 2004 [Low04]. Another used approach is visual words where an image is represented by a histogram of visual words [Fei05],[Siv03]. Besides, using clustering techniques is an efficient and an important option added to CBIR systems since they allow reducing the time of retrieval and increasing the performance of research. K-means was early proposed over 50 years ago it is still one of the most widely used algorithms for clustering. The main reasons for the success of these techniques are the ease of implementation, the simplicity, the efficiency and the empirical success. The most efficient k-means algorithm is ElKAN's algorithm [Jai11]. For instance, SemQuery system [She02] organizes images into different groups of clusters based on their heterogeneous features. Vailaya et al. [Vai01] create a hierarchical structure about vacation images. At the top level, images are classified as indoor or outdoor. Outdoor images are then classified as a city or landscape that are further divided into the sunset, forest, and mountain closes. The SIMPLIcity system [Wan01] category images into a graph, textured photograph, or non-textured photograph. In 2012, Swapna Borde and Udhav Bohosle proposed novel techniques for image retrieval using clustering features extracted from images which are RMC, CMC, RMDC and CMDC, RMWC and CMWC. Other techniques use the advantages of transforms (wavelet and DCT) [Rai12] . Other useful MPEG-7 for searching in multimedia systems are in [Sal02,Bas10, Say05]. For instance, VITALAS (**V**ideo & image **I**ndexing and re**T**riev**A**l in the **LA**rge **S**cale) is an Industrial project started from 2007 the aim of this project is to produce a prototype for industry to retrieve and index multimedia information. This project has given an interesting result. The goal of this project focus on: First, Cross-media indexing and retrieving try to use automatic annotation and getting semantic level. Second, using techniques for large scale search. Finally, trying to improve visualization and context adaptation[Vit07].

3. THE PROPOSED METHOD

Our approach uses k-means clustering on each feature of the images separately. For each feature vector we get a different clustering. For example, if we extract color feature vector and texture feature vector and shape feature vector we get color clustering, texture clustering and shape clustering. These differences clustering allow us to have different views and levels of the retrieved images for a query which is not possible with other approaches. We call this method Multi-Clustering Content Based

Image Retrieval. The proposed approach takes four steps:
The two first steps are in the offline phase of our CBIR system and the two second steps are in the online phase.

1. Feature extraction.
2. K-means multi-clustering.
3. Query image searching.
4. Organizing results.

3.1. Feature Extraction

The first step is to extract features from the images of the database. Vi is a feature vector. We choose color and texture features for our system. Images used are in RGB color space. In addition, we propose to represent color feature with three feature vector: red histogram, green histogram and blue histogram. Also, we use the gray level co-occurrence matrix GLCM to represent texture feature.

3.1.1 Histogram

A color histogram [Swa91] is an important feature we can extract gray level histogram as we can use color histogram. For our system, we extract histograms with 256 bins of each R, G, B colors. So, as a result we get three feature vectors for color descriptors.

3.1.2 Co-occurence Matrix (GLCM).

The Grey Level Co-occurrence Matrix, or called the Grey Tone Spatial Dependency Matrix) is a table representing a number of different combinations of pixel brightness values (grey levels) that occur in an image [Har73].

At the end of this step four feature vectors are extracted

- V1: histogram of red color.
- V2:histogram of green color.
- V3:histogram of blue color.
- V4: co-occurrence matrix.

To clarify more this first step we illustrate an example (see Fig.2). To simplify this example we suppose that we have three feature vectors for each image in the database

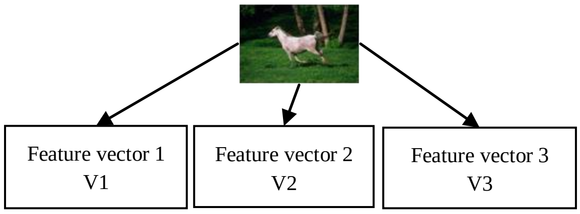

Figure 2. Illustration of the feature extraction for one image.

3.2. K-Means Multi-Clustering

After the feature extraction, k-means clustering algorithm is executed. K-means clustering is a leaning machine algorithm. This algorithm needs 80% of images for the learning phase and 20% of images for the evaluation phase.

The evaluation images are the images used to query the system. Learning k-means clustering is to find K different clusters in a database of N objects, where similarity intra cluster distance is minimized, and the distance inter clusters are maximized [Tan05], [Jai11].

The learning machine is repeated until the centers of the clusters are stable. Algorithm 1 shows the operation of k-means clustering [Tan05].

```
Algorithm k_means_clustering
Begin
  Select K points as the initial
  centroid.
  Repeat
   Form K clusters by assigning all
   points to the
   closest centroid.
   Recompute the centroid of each
   cluster.
  Until the centroids don't change.
End.
```

Algorithm 1: k-means clustering [Tan05]

At the end of this step we obtain k clusters with K stable centers and each N element assigned to its corresponding cluster. This learning is executed for each previous vectors V_i, i=1. . 4. For our example we have just V_1, V_2 and V_3 (see Fig.3).

 ISBN 978-80-86943-75-6

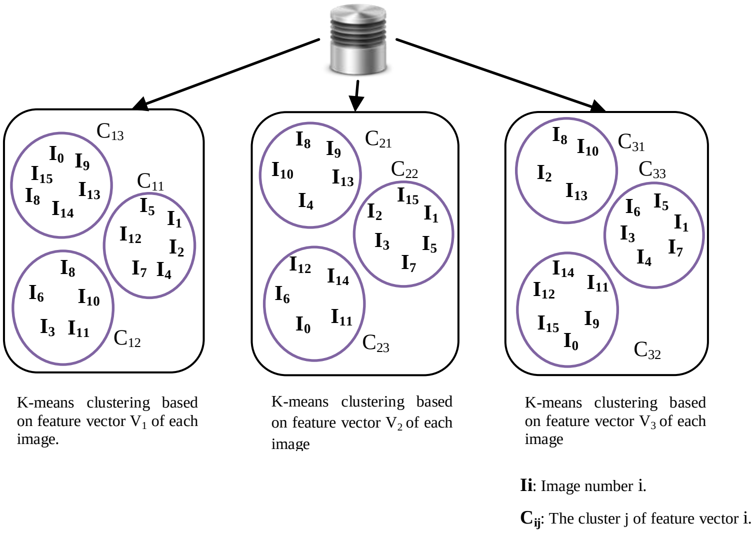

Figure 3. Example of a multi-clustering data base (offline phase)

3.3. Query Image Searching

The two previous steps are in the offline phase. Next two steps are in the online phase. First, we extract the feature vectors of the query image using the same process (see Fig.2). As a result we get four feature vectors (VQ_1 , VQ_2 ,VQ_3, VQ4).

For each feature vector, a searching process is launched. Each query feature vector VQ_i is searched in its corresponding clustering. For instance, VQ_1 is searched in V_1 clustering. As a result we get the relevant cluster of each feature vector VQ_i (see Fig.4).

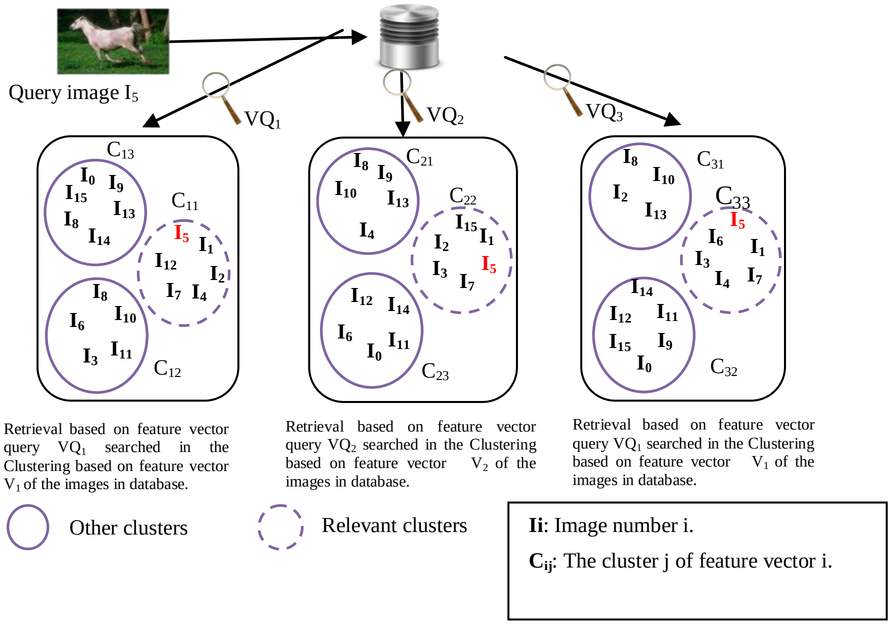

Figure 4. Query processing (online phase)

 ISBN 978-80-86943-75-6

3.4. Organizing Results

The last step is to organize the results from the most to the less relevant. We propose two ways in this step. First, organization by levels. Second, organization by the clusters union. Let us carry on the previous example, we take $\mathbf{I}_5$ as image query. As a result we have three relevant clusters (C_{11}, C_{22}, C_{33}) (see Fig.4).

3.1.3 First Proposed Organization (By Levels).

In this organization we propose to separate the results on three levels.In the first level we display the intersection of all relevant clusters. In our case $C_{11} \cap C_{22} \cap C_{33} = \{I_5, I_1, I_7\}$. For the second level.The intersection two by two of relevant clusters that do not exist in the previous level. For our example: $((C_{11} \cap C_{22}) \cup (C_{22} \cap C_{33}) \cup (C_{11} \cap C_{33})) - (C_{11} \cap C_{22} \cap C_{33}) = \{I_2, I_4, I_3\}$. Third level. The images that exist in just one cluster and do not exist in the previous two levels. For our example $C_{11} \cup C_{22} \cup C_{33} - ((C_{11} \cap C_{22}) \cup (C_{22} \cap C_{33}) \cup (C_{11} \cap C_{33})) - (C_{11} \cap C_{22} \cap C_{33}) - C_{11} \cap C_{22} \cap C_{33} = \{I_{12}, I_{15}, I_6\}$.

After, for each level we compute the similarity measure (Euclidian distance) and order the retrieved images in a descending order.

3.1.4 Second Proposed Organization (By a Clusters Union).

The second proposed organization is to group all images in just one level and compute the similarity measure and order the retrieved images in a descending order. In our example, $C_{11} \cup C_{22} \cup C_{33} = \{I_1, I_2, I_4, I_7, I_{12}, I_{15}, I_3, I_6\}$.

4. RESULTS AND DISCUSSION

This section is a presentation of our experimental results. In addition, we are going to compare our results to other clustering CBIR system methods like WaveQ [Geb07] and discuss these results. To experiment our approach, we use the Corel database [Wan01]. The database contains 1000 images dived on 10 classes as follows: African people and villages, Beach, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains and Glaciers, Food. We choose to extract color and texture features. To process our method we first divide our dataset into two parts, where 80% of the images are used during the training data set (offline phase) and 20% of the images are used for the query phase (online phase). During the offline phase, we pre-process the images, extract their feature vectors and construct a new database containing features vectors of each image. After, we apply k-means clustering for each feature vector. In the online phase, we calculate the distance between the query image and the images in the database. Finally, we sort these distances and return all the found relevant images after using our two organization methods presented in section 3. Moreover, a presentation of the results without using texture features (only color features) and then we add other results using both features (color and texture). We evaluate our system with computing recall and precision[Per01].

$$\text{Recall} = \frac{\text{number of releveant images retrived}}{\text{number of relevant images in collection}} \quad (1)$$

$$\text{Precision} = \frac{\text{number of relevant images retrival}}{\text{total number of images retrieved}} \quad (2)$$

4.1.Without Texture Feature

We execute our system on several images without using texture features. So we use just three feature vectors (red, green, blue). The results are impressive and in most case relevant 20 images are displayed on the top of the retrieved images.

We test for (K= 5) or five clusters. From 69 tested images average recall is above 0.77 and average precision is over 0.33. In addition, we get 94% of querying images having a precision above 0.50. Also, we get 52% of images having more than 0.90 recall value. We notice that for five classes we get better results.

4.2.With Texture Feature

In this part we add the co-occurrence matrix to the three other features. The results are good and promising. We test for (K=5) and we get the following results the average of recall is above 0.918. Comparing to the previous tests without texture, texture features increase the recall. This means that the number of retrieved relevant images increases.

4.3. Comparison With Other Systems

We will compare our system with WaveQ system. WaveQ uses a clusering method for this database. An execution is shown (see Fig.5) where we can notice the efficiency of our system comparing to WaveQ. WaveQ gives as results the image query at first and then a dish in the second position. Also, the other images are semantically the same buses. Our system displays the image query as first image and the four other images are buses. In addition, our results are visually the same (red and white buses color) and semantically the same (bus in the city).

 ISBN 978-80-86943-75-6

Figure 5. An example comparing WaveQ system and our system. Set of Images number 1 are results of WaveQ system. Sets numbers two are the five first retried images of our system without texture. Sets numbers three are the five first images results of our system with texture.

We can see another test of our system. First, A multi level organization is displayed (see Fig.6). In addition, we extract images intersection of the relevant clusters images (see Fig.7).

Figure 6. An example of levels organization results in our system. First five relevant images are displayed for each level

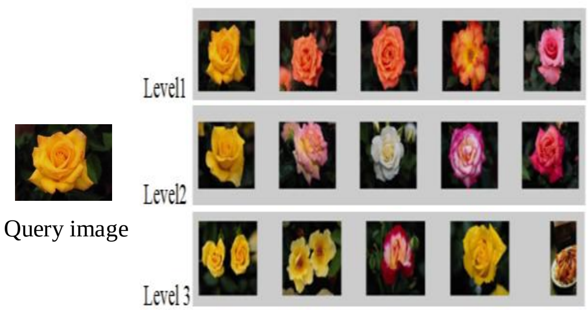

Figure 7. An example of second proposed organization method of our system. Twenty first relevant image are displayed

5. CONCLUSION AND FUTURE WORK

In this paper a new approach for content based image retrieval is proposed. This method uses K-means clustering separately for each feature vector of an image. This is done in order to keep the efficiency of each feature without combining them to get one feature vector. Also, the proposed approach avoids the computing of similarity measures for the entire database, we just calculate those of the relevant clusters. The results show the effectiveness of our approach and give a good Recall and a promising precision values in database of 1000 images. In addition, our system gives very satisfactory retrieved images (20 first images look alike the query). Comparing to WaveQ our system gives visually better results.

As future work, we will test this method with different color space, color features and texture features. In addition, using others similarity measures.

References

[Bas10] Bastan, M. , Cam, H., Gudukbay, U. and Ulusoy, O. BilVideo-7: An MPEG-7-Compatible Video Indexing and Retrieval System," *IEEE Multimedia*, vol. 17, no. 3, pp. 62–73, 2010.

[Car02] Carson, C., Belongie, S., Greenspan, H. and Malik, J. Blobworld: image segmentation using expectation-maximization and its application to image querying. Pattern Analysis and Machine Intelligence, IEEE Transactions on.Vol. 24, pp.1026-1038, 2002.

[Col04] The Color Guide and Glossary, Communication, measurement and control for Digital Imaging and Graphic Arts X-rite, 2004.

[Fal94] Faloutsos, C., Barber, R., Flickner, M., Hafner, J., Niblack, W., Petkovic, D. and Equitz, W. Efficient and effective Querying by Image Content. Journal of Intelligent Information Systems.Vol. 3, pp.231–262 , 1994.

[Fei05] Fei-Fei, L. and Perona, P. A Bayesian hierarchical model for learning natural scene categories. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 2, pp. 524 – 531, 2005.

[Geb07] Gebara, D. and Alhajj, R. WaveQ: Combining Wavelet Analysis and Clustering for Effective Image Retrieval. Advanced Information Networking and Applications Workshops, 2007, AINAW '07. 21st International Conference on. pp.289–294, 2007.

[Gup97] Gupta, A. and Jain, R. Visual information retrieval. Commun. ACM. 40, 70–79, 1997.

[Har73] Haralick, R.M., Shanmugam and K., Dinstein, I.: Textural Features for Image Classification. Systems, Man and Cybernetics, IEEE Transactions on. SMC-3, pp.610–621, 1973.

[Jai11] Jain, M., Singh and S.K. A Survey On: Content Based Image Retrieval Systems Using Clustering Techniques for Large Data sets. International Journal of Managing Information Technology.Vol. 3, pp.23–39 , 2011.

[Low04] Lowe, D. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision.Vol. 60, pp.91–110, 2004.

[Pen96] Pentland, A., Picard, R.W. and Sclaroff, S. Photobook: Content-based manipulation of image databases. International Journal of Computer Vision.Vol. 18, pp.233–254, 1996.

[Per01] Müller, H. , Müller, W., Squire,S. Marchand-Maillet, D. M. and Pun, T. performance evaluation in content based image Retrieval: overview and proposals, Pattern Recognition Letters, vol. 22, no. 5,pp. 593–601, April 2001 .].

[Rai12] Raikwar, A.K. and Jain, S.Article: Content based Image Retrieval using Clustering. International Journal of Computer Applications. 41, pp.29–33, 2012.

[Sal02] Salembier, P. and Sikora, T. Introduction to MPEG-7: Multimedia Content Description Interface. New York, NY, USA: John Wiley & Sons, Inc., 2002.

[Sam59] Samuel, A.L. Some studies in machine learning using the game of checkers. IBM J. Res. Dev.Vol. 3, pp.210–229 , 1959.

[Say05] Saykol, E., Güdükbay, U. and Ulusoy, Ö. A histogram-based approach for object-based query-by-shape-and-color in image and video databases, *Image Vision Comput.*, vol. 23, no. 13, pp. 1170–1180, 2005.

[She02] Sheikholeslami, G., Chang, W., and Zhang, A. SemQuery: semantic clustering and querying on heterogeneous features for visual data. Knowledge and Data Engineering, IEEE Transactions on.Vol. 14, pp.988–1002, 2002.

[Siv03] Sivic, J. and Zisserman, A. Video Google: a text retrieval approach to object matching in videos. Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. Vol.2, pp. 1470 –1477 ,2003.

[Smi96] Smith, J.R. and Chang, S.-F. VisualSEEk: a fully automated content-based image query system. Proceedings of the fourth ACM international conference on Multimedia. pp. 87–98. ACM, New York, NY, USA , 1996.

[Sub02] Subramanya, S.R., Teng, J.-C. and Fu, Y. Study of Relative Effectiveness of Features in Content-Based Image Retrievals. Proceedings of the First International Symposium on Cyber Worlds (CW'02). IEEE Computer Society, Washington, DC, USA , pp.0168-175, 2002.

[Swa91]Swain, M.J. and Ballard, D.H. Color indexing. International Journal of Computer Vision.Vol. 7, 11–32 ,1991.

[Tan05] Tan, P.-N., Steinbach, M. and Kumar, V. Introduction to Data Mining, (First Edition). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA ,2005.

[Tor] Torres, R.D.S., Falcão, A.X. Content-Based Image Retrieval: Theory and Applications. Revista de Informática Teórica e Aplicada. Vol. 13, pp.161–185.

[Vit07] VITALAS (Video & image Indexing and reTrievAl in the LArge Scale) http://vitalas.ercim.eu/.

[Vai01] Vailaya, A., Figueiredo, M.A.T., Jain, A.K. and Zhang, H.-J. Image classification for content-based indexing. Image Processing, IEEE Transactions on.Vol.10, pp. 117–130, 2001.

[Wan01] Wang, J.Z., Li, J. and Wiederhold, G. SIMPLIcity: semantics-sensitive integrated matching for picture libraries. Pattern Analysis and Machine Intelligence, IEEE Transactions on.vol. 23, pp.947–963, 2001.

[Zha04] Zhang, D. and Lu, G. Review of shape representation and description techniques. Pattern Recognition.Vol. 37, pp.1–19 ,2004.

[Zha12] Zhang, D., Islam, M.M. and Lu, G. A review on automatic image annotation techniques. Pattern Recognition. Vol. 45, pp. 346–362 ,2012.

 ISBN 978-80-86943-75-6