Asian Journal of Information Technology 15 (16): 2756-2762, 2016
ISSN: 1682-3915


# Object-Based Image Retrieval Using Enhanced SURF

[1]Tawfiq A. Al-asadi and [2]Ahmed J. Obaid
[1]Department of Computer Science, College of Information Technology,
Babylon University, Babel, Iraq
[2]Department of Computer Science, College of Education, Kufa University, Najaf, Iraq

**Abstract:** Image retrieval is a key challenges in many image database application and still an active field in computer vision application. There are many proposed image retrieval systems that retrieve images based on image contents such as colors, texture, shapes and feature descriptor. The main task for image retrieval system is to create a system that capable to retrieve images that are semantically related to user's query from an image database. When user interest to retrieve images that contain a particular objects instead of retrieve similar images which might not related images to his interesting this is called object based on image retrieval. The goal of Objects-based is to retrieve images based on objects that appear in those query images from large database. In this study we use enhanced Speeded UP Robust Features (SURF) algorithm as main step to extract features from interested query objects and then checked and matched result to retrieve related images from image dataset. Speeded UP Robust Features (SURF) is a scale and rotation invariant detector and descriptor feature algorithm and was applied successfully in many Image retrieval systems due its robust against different image transformation. Finally the result written in a report and images saved along with user's query time stamp.

**Key words:** Object Based Image Retrieval (OBIR), feature matching, object recognition, SURF, image database

## INTRODUCTION

In recent years, the fast growth of image collections and databases has established a need for user-friendly image retrieval system that can retrieve images based on user's query. Image Retrieval system can be defined as the system that find all images in a given image database that depict scenes of some predefined user specification called queries (Bay *et al*., 2008). Image retrieval algorithms can be divided into two categories: text based and content based approaches. Content Based Image Retrieval (CBIR) is an important field in image retrieval and is a complement approach to text based content. Content Based Image Retrieval (CBIR) is an image retrieval system that is retrieve image based on its content by use low level visual content of an image like texture, shape, color and spatial information to represent and index the image and then measure similarity among images to return relevant result based on the difference between low level features. Many research work has been done in this field and many techniques were adopted to enhance the accuracy of retrieval. Object Based Image Retrieval (OBIR) is a part of image retrieval system and can be defined as the system that retrieve images from image database based on the appearance of objects in those images. Objects can be market logs, cars, stop signs, flags and others where user interest to find. Objects that used for image retrieval must satisfy conditions such as: Object can distinguish from background and has distinctive features like color or texture (Hoiem *et al*., 2005). Vision features can be classified into two classes: first, low level features include color, texture and inflexion. Second, middle level features that include shape description and object featur.

In this study we used enhanced Speeded UP Robust Features (SURF) (Asadi and Obaid, 2016) one of the robust local feature detector and descriptor algorithm, first presented by Bay *et al*. (2008). Speeded UP Robust Features (SURF) is local feature descriptor of the interested points that detected in an integral image. These descriptors can be used to detect the matching size between two images or for detection particular object while other objects exist in an image. For particular object where user interest to find in an image, feature extracted that correspond to it should be similar even though are extracted in different illumination, scale and noise in order to perform accuracy in object detection and recognition. Speeded UP Robust Features (SURF) is well known algorithm that can be used in computer vision application that is invariant to scale, rotation and illumination change and used to perform tasks of image retrieval, object

**Corresponding Author:** Tawfiq A. Al-asadi, Department of Computer Science, College of Information Technology, Babylon University, Babel, Iraq

detection, object recognition, face recognition etc. Quality of object recognition and retrieval images that contain that object is important to real time requirement. The process of Speeded UP Robust Features (SURF) that used for detection interested objects and retrieve related images divided into three main steps: detection, description and matching step. In first one where interest points are selected in distinctive location, in second step, interest points should has unique identifiers and finally in matching step where descriptor (feature) vectors are compared in order to find user's interest objects in those images. In this study we introduce an enhanced technique for detection and retrieved related images not for single object only but for many objects simultaneously and this can be reduce time and also manual object selection criteria.

**Literature review:** In this section we try to identify some of the current works in the research of object based image retrieval system. Most common systems that use segmentation of an image into homogenous regions and then compared object regions with query images as in (Sharawy *et al*., 2009; Jing *et al*., 2003; Wang, 2000). Bag of visual word has been applied by Yan-Tao Zheng for object based retrieval by using new approach called bag of visual synsets and is defined as a probabilistic concept, this concept groups similar words that has same probability which belongs to the same class and disparate from query image. Kavitha and Sudhamani (2013) used a technique that combine local and global features, local feature used for edge and corner detection of an object then use global feature (HSV) color feature for extracted an object from images. In David Garc´ýa-P´erez researchers, they used another approach for object based retrieval where objects are represented by using two dimensional structure active net that capable to adjust relevant image regions based on edge and chromatic information and used for matching process. A new approach has been introduced by Broek *et al*. (2005) by use human based 11 color categories and color correlogram for segment images and then extracted features of regions. SURF algorithm used in many computer vision and real time application for face recognition, object detection and especially in image retrieval due its consider fast matching features algorithm. Speeded UP Robust Features (SURF) has been combined with Support Vector Machine (SVM) for image retrieval as in (Sharma and Saddiqui, 2014). In the proposed method we used enhanced SURF algorithm (Asadi and Obaid, 2016) for extracting features from images in image collection and object images, model compared and find the percentage of matching for each object in all image collection then retrieved related images that include interested object, procedure will continue for next interested object. The contribution in this study is provide a user-friendly framework where user select his interesting objects to find in image collection or image database, model detect all objects in all images and retrieve images correspond to every query object separately, proposed model also checked the new entered query object if it found in object collection database then confirm user about it, our proposed model saved the result in predefined path for preview result at any time. Our proposed model can support different type of reports based on user's purpose and requirement, also provide controlled GUI for entering parameters for some metrics that used for matching and retrieved steps as shown in result and discussion section in this study.

**Overview of SURF algorithm:** Speeded UP robust Features "SURF" algorithm is a robust local feature detector and extractor as well as Scale and rotation invariant algorithm, rotation invariant is an important property of image features. There are two steps for extracting feature descriptor in Speeded UP robust Features "SURF" algorithm, to obtain SURF descriptor first detecting points and then extracting descriptor at SURF points. SURF used fast Hessian matrix for detection points. Hessian matrix determination can be used for deciding whether the points can be selected as interest points or not. Hessian matrix at point X and scale ó for an input integral image (I) and Lxx(x, ó) convolution of Gaussian second order derivative of an image at point with coordinates (x, y) is defined by the Eq.1 as follow:

$$H(x,\sigma)=\begin{bmatrix} Lxx(x,\sigma) & Lxy(x,\sigma) \\ Lyx(x,\sigma) & Lyy(x,\sigma) \end{bmatrix} \quad (1)$$

Gaussian second order derivative needs to be discretized before implement convolution with an image. Bay *et al*. (2008) proposed box filter to approximation discretized Gaussian partial derivative. $D_{xx}$, $D_{yy}$ and $D_{xy}$ represent the convolution of box filter with the image, the approximated second order Gaussian derivative calculation is made fast by using integral image and approximated Hessian matrix determination is defined by Eq. 2 as follow:

$$\det\left(H_{apprax}\right)=D_{xx}D_{yy}-\left(0.9D_{xy}\right)^2 \quad (2)$$

The approximated determinant of Hessian matrix is calculated at each scale and non-maximum suppression in 3×3×3 neighborhood is applied to find the maxima. SURF point orientation gained by using Haar wavelet response, in neighborhood of SURF point Haar wavelet response is

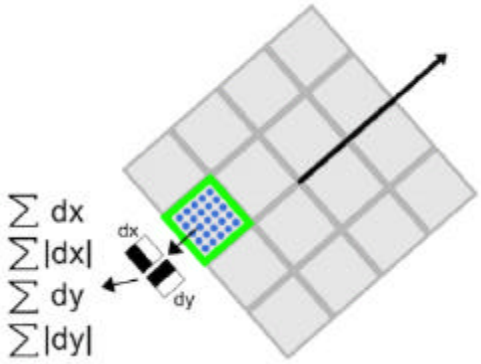


Fig. 1: Feature descriptor at interest point

calculated in both x and y direction in circular neighborhood of 6s (where s is scale factor at which interest point was identified). Dominant orientation is estimated by calculating sum of all response within a sliding orientation window and longest vector is chosen as the dominant orientation. Then descriptor vector is computed by choosing square regions (size 20s), this region chosen around the interest point is split into smaller 4×4 sub-regions and Haar wavelet in x and y direction are calculated within each sub-regions at 5×5 regularly placed sample points as shown in Fig. 1.

Wavelet response are summed up in each region, by use this values a four dimensional feature vector V = ($\Sigma d_x$, $\Sigma d_y$, $\Sigma d_x$, $\Sigma d_y$) constructed per 4×4 regions to form descriptor vector that used for matching step. Descriptor length used for matching step can be 64 or 128 length, in our proposed method we used length of 128 to produce perfect matching. Finally matching has been done between descriptor vectors of each object in object database and collection of images (image database), matching score calculated based on distance among vectors by using distance metrics such as Euclidian distance, in our proposed model we used enhanced SURF (Asadi and Obaid, 2016) to determine the number of matching features between object and tested image then given a decision about it if exist or not, descriptor of points which is uniquely identifies points for matching even if tested image get distortion due illumination change, noise, rotation change, scale change and view point change. Our model used another algorithm for eliminated outliers features, details of our proposed model is discuss in next section.

## MATERIALS AND METHODS

**Proposed methodology:** The proposed method used enhanced SURF algorithm as shown in Fig. 2 as follow: We use an image dataset that selected from (CIT, Babylon University) archiving images for web purposes which

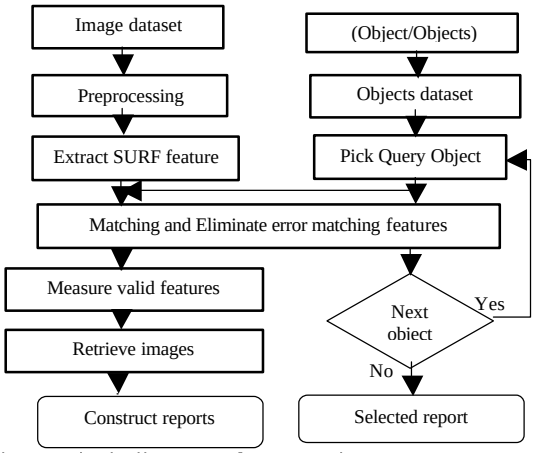


Fig. 2: Block diagram of proposed system

contain 18 folders that include images used in 18 Websites, objects images are selected from (CIT, CS Dept., Babylon University), following the information about our selected dataset.

- Total folders dataset: 18 image folders
- Total images in one folder: 50-100 images
- Total selected objects images :18 objects
- Total size in local disk: 1.3 GB
- Total images in image dataset: 1500 images

In this model user have selected his interested (object/objects) to be found in image database (image dataset) by retrieve all images contain it, our interested objects has been stored in a database called objects database. There is an option to be check all objects in database or just particular objects which may selected by user. When user add new object image using this system, object will be checked if it is found in object database or not, then it will be added to database in case is not found. Before this system has work starting there is a simple preprocessing step to filtering invalid images in image dataset, such as icons, social, gif etc. images. Generally invalid images has small size for example might not exceed 25 kb will not consider. After system we be starting, first object picked and then SURF interested points selected from it, in our proposed model strongest feature has selected based on predefined threshold, Feature descriptor also called object representative because its can distinguish and recognize objects in the image, and this process same as in image dataset, were constructed. Matching done between first picked object and all images in dataset, in matching process features include inlier and outliers features from the total number of feature descriptor that used for matching process, to eliminate the outlier features, RANdom SAmple Consensus (RANSAC) algorithm has applied to map points between points pairs

of two images (Asadi and Obaid, 2016). RANSAC is an iterative model to valuation parameters of mathematical model from set of observed data that include outliers, and can be defined as an outlier detection method (Strutz, 2016). RANSAC depend on distance threshold, pair $P_{i\,(a)}$, $P_{i\,(b)}$ a pair of points in image a and image b simultaneously and is an inlier just when the distance between $P_{i\,(b)}$ and projection of $P_{i\,(a)}$ based on transformation matrix falls within predefined threshold. Distance used in RANSAC algorithm is define by Eq. 3 as follow:

$$d = \sum_{i}^{N} = 1 \text{Min}\left(D(pi(b), \varphi(pi(a):M)), T\right) \quad (3)$$

Where:

$P_{i\,(a)}$ = A point in image a

$P_{i\,(b)}$ = A point in image b

$\varphi(p_i(a): M)$ = Projection of point on image a based on transformation matri M

D = Distance measure

T = Predefined threshold

N = total number of interested points (feature point)

Measuring valid features include the formula for counting the number of inlier feature descriptor to the total number of features that used for matching process, when the percentage of inlier feature exceed outlier we can declare the object is found in images, otherwise when outliers features exceed number of inlier features then object is not found, so we have total number of feature descriptor F where is $F_{in}$ total number of inlier feature descriptor and $F_{out}$ is the total number of outlier feature descriptor, P is a measure for deciding whether object is found or not also we call it (metric for valid feature) and is defined in Eq. 4 and 5 as follow:

$$P_{in} = \frac{F_{in}}{F} \times 100\% \quad (4)$$

$$P_{our} = \frac{F_{our}}{F} \times 100\% \quad (5)$$

Predefined Threshold ($T_R$) for valid feature, then valid feature must satisfy conditions: $P_{in} \geq T_R$ and $P_{out} \geq T_R$.

Finally, retrieved images will be shown and result has been stored for each query object along with time stamp information. There are many report can be constructed based on user requirement, in our system framework we build three type of reports, first report contain the information about all images names correspond to all

Fig. 3: Object-based image retrieval main GUI

query objects in object dataset. Second, is a summary report that contain information about directories names opposite to all query objects, third one, include the information for only selected query object from object dataset, selected query object done by user when need to retrieve images only for selected and interested single query object. The details of these reports and final result are discussed in details in next section (Result and discussion section).

**Proposed object based retrieval algorithm**

**Object based retrieval algorithm:**

Step1: Preprocessing image dataset
Outer loop = length (Source directories).
Inner loop = length (images in each directory).
Read (image) from current directory.
If Size (image) = 50 KB and Ext (gif, tif, icon)
Then Eliminate.

Step2: Detect strongest points
For I = length of (Objects in Dataset/selected)
$P_a$ = Read first (Object/Selected Object)
For J = length of (Source directories)
For L = length (images in Current directory)
$P_b$ = Read (image from current directory)
Detect feature point in ($P_a$, $P_b$)
Selected Strongest features from ($P_a$, $P_b$)

Step3: Extract feature descriptor
Extract feature for ($P_a$, $P_b$)

Step4: Matching and eliminate outliers
Match = Match feature for ($P_a$, $P_b$)
IF Match has sufficient features then
Apply (RANSAC) algorithm
Else go to step 1

Step5: Detect object in images
IF $P_{in} \geq T_R$ and $P_{in} \geq T_R$ Then
Store image in result database

Step6: Retrieved images and report constructed
Show images in result database
Write result to Reports

End

## RESULTS AND DISCUSSION

Object based image retrieval simulation has been done by using Matlab, main GUI consist of many option for entering and selecting interesting objects, our selected image dataset consist of source directory include 18 image directories, Fig. 3 show the main interface for our system.

Fig. 4: Main GUI that show all interesting objects


Fig. 5: Object and origin images (BMW poster)

Fig. 6: Detected Object in origin image

Main GUI include many options for matching process when user would like to check single object correspond to single or all images in dataset, also option found for selected directory, another option available when user would like to match all objects in object dataset, view options for our system in main GUI is 18 objects, in case there are many objects more than 18 objects then could be view in montage view or by use sub-plot option. Figure 4 show the option when user show all objects in object dataset.

As shown in Fig. 4 user have option to select single object to be checked in all directories and show relevant images based on it, GUI has option to select source directories when user would like to retrieve images from some of sub-directories in source directory, there are three options to select sub-directories, PACK1, PACK2 and PACK3 by user. Object can be recognize and detected in single image as shown in Fig. 5 and 6 as follow:

Both image has been selected from Google for application purpose, the result of detection object shown in the following figure (Fig. 6).

Fig. 7: Result of object based image retrieval

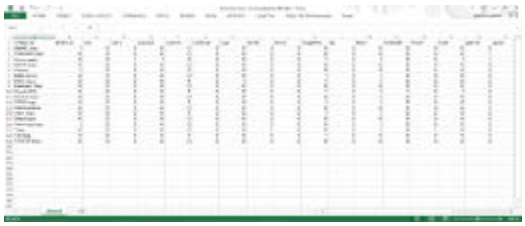

Fig. 8: Result of first directory in an excel file

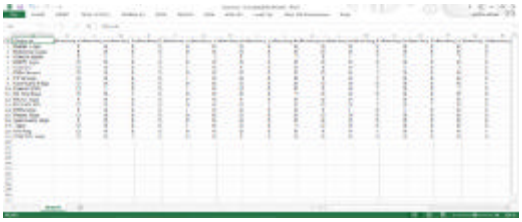

Fig. 9: Summary result for image dataset

When user would like to retrieve all images that contain BMW logo as shown in Fig. 7 then user will get beside retrieve images two types of report, one contain summary result and other for matching result. Summary report contain the image directories name and correspond object name and the entries values of this report contain only two values (1) if that object found in image directory at all and (0) if it's not found. Details reports, consist of 18 reports as our selected source directory which consist of 18 image directory, each report has a particular directory name and contain all image name in that directory as well as object names, and entries of report is two values (1) if object is found in correspond image and (0) for others. Option for retrieving images in main GUI is also supported to view retrieved images for each query object, user can select query object and then related images can be view up to 18. If the number of images >18 then sub-plot option and montage view can be supported to view all related result. Figure 7 show the retrieved images for BMW logo query object.

Report can be constructed based on user interest, Fig. 8 show the result of object-based image retrieval in an excel mode.

Result of query objects for all directories can be constructed in a single sheet in a summary report that give only two values, 1 if the corresponding object is found in any images in that directory otherwise 0, Fig. 9 show the summary result for our image dataset.

**Performance evaluation:** The most common measures used in retrieval system for performance evaluation are precision and recall. Recall can be defined is the ability measure of system to present all relevant images, in other words how many relevant images we retrieved out of known total images, Recall (R) is the ration of number of retrieved relevant images to the total number of relevant images in dataset, to understand this measure and how do we use in our model, we suppose that A is the total number of relevant images we have retrieved and C represent the number of relevant images in our image dataset, formula of Recall can be written as follow:

$$\text{Recall}(R) = \frac{\text{No. of relevant images retrieved}}{\text{No. of relevant images dataset}} = \frac{A}{C} \quad (6)$$

While Precision can be defined by the ability to present only relevant images. Precision (P (is the ration of number of relevant images retrieved to the total number of images retrieved (relevant and non-relevant), in other words to give some explanation about this measure, we suppose A is the number if relevant images retrieved and B is the total number of irrelevant images retrieved, then A+B give to us the total number of relevant and non-relevant images, formula of Precision can be written as follow:

$$\text{Precision}(P) = \frac{\text{No. of relevant images retrieved}}{\text{Total No. of retrieved images}} = \frac{A}{A+\sigma} \quad (7)$$

To apply these measures, we select the first result from our table to calculated and evaluate our proposed Object-Based image retrieval system, first example of Babel logo, the total number of images that contain this logo is 322 and the number of relevant images retrieved is 278, number of irrelevant images retrieved is 66 then precision in this case is (0.80) and so on for other values, Recall is slightly difference from Precision, C is number of relevant images in image dataset, then in our example C is 322 then Recall (R) is (0.86), the formula for calculating the precision and recall as follow:

$$P = \frac{278}{278+66} = 0.80 \quad (8)$$

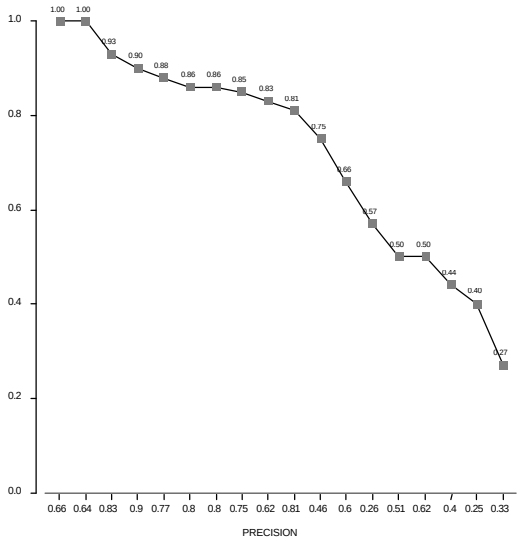


Fig. 10: Precision-Recall graph

Table 1: Evaluation result of object-based retrieval system

| Query object | Object-name | Images | Recall | Precision |
|---|---|---|---|---|
| 1 | Babel Logo | 322 | 0.86 | 0.80 |
| 2 | National Logo | 10 | 0.50 | 0.62 |
| 3 | Green apple | 8 | 0.75 | 0.46 |
| 4 | BMW logo | 18 | 0.66 | 0.60 |
| 5 | Canary | 2 | 1.00 | 0.66 |
| 6 | Eiffel tower | 7 | 0.57 | 0.26 |
| 7 | FIFA logo | 23 | 0.86 | 0.80 |
| 8 | Germany Flag | 9 | 0.44 | 0.40 |
| 9 | DAESH ISIS | 11 | 1.00 | 0.64 |
| 10 | Kit Kat logo | 14 | 0.85 | 0.75 |
| 11 | NIDO logo | 6 | 0.83 | 0.62 |
| 12 | MONALIZA | 10 | 0.90 | 0.90 |
| 13 | NBA logo | 55 | 0.81 | 0.81 |
| 14 | Pepsi logo | 5 | 0.40 | 0.25 |
| 15 | Samsung logo | 16 | 0.93 | 0.83 |
| 16 | Tiger | 22 | 0.27 | 0.33 |
| 17 | UK flag | 20 | 0.50 | 0.51 |
| 18 | UNICEF logo | 42 | 0.88 | 0.77 |
| Total | 600 | | | |

$$P = \frac{278}{322} = 0.80 \quad (9)$$

Then from result of Table 1 for each query object we got pair of values for Precision and recall we can draw Precision- Recall graph as shown in figure (Fig. 10).

In Table 1 show the evaluation result for our proposed method include the predefined 18 query object in object dataset.

As explained in Table 1, there are 18 objects used for evaluation in our selected dataset, images are counted which contain the predefined logos and images, the second column entries in Table 1 represent the count number of images that contain this logo or image in our selected dataset.

## CONCLUSION

Traditional Content Based Image Retrieval (CBIR) system used low level features such as color, texture and shape for retrieval similar images. CBIR is a challenges task in recent years in the manner of perform good accuracy for retrieval processes. Many proposed methods and modification has been done in this field to obtain acceptable result. In recent year interest point detection and extraction algorithm has been used for retrieval similar images such as SIFT, SURF. SURF algorithm applied in many computer vision application for face recognition, object detection and image retrieval. SURF can perform fast by compare it with other algorithm that used for matching between images like SIFT. SURF is also robust to invariance when happen in images such as scale changes, rotation and noise. SURF also combine with other techniques such as color moment and HSV histogram to develop its work for image retrieval and object detection. There are many research work focused for extracted objects from query images and then matched extracted objects to other images in image database. This techniques also faced many challenges based on method used for segmentation which may produce fractal segments and also the accuracy depend on the algorithm and type of images used for this purpose. When users would like to retrieve images that has particular objects then object based image retrieval has used for this task to retrieve images that as possible contain or include query objects in its scene. Proposed system used enhanced version of SURF algorithm for extracting feature descriptor from one or more query object at same time and matching it with all images has been found in image database. By compare proposed method with other systems where time is decreased and user able to submit more than one object at same time and also get independent result for each query. Proposed method can be used for image mining for clustering and classification images based on its content. We take in consideration to provide different types of reports can e used for mining purpose and to view result for each independent image directory and combine result based on interesting result.

## REFERENCES

Asadi, T.A.A. and A.J. Obaid, 2016. Object detection and recognition by using enhanced Speeded Up Robust Feature. Intl. J. Comput. Sci. Network Sec., 16: 66-71.

Bay, H., A. Ess, T. Tuytelaars and L. van Gool, 2008. Speeded-Up Robust Features (SURF). Comput. Vision Image Understand., 110: 346-359.

Bay, H., T. Tuytelaars and L.V. Gool, 2006. Surf: Speeded up Robust Features. In: European Conference on Computer Vision, Leonardis, A., H. Bischof and A. Pinz (Eds.). Springer, Berlin, Germany, ISBN: 978-3-540-33833-8,pp: 404-417~ pp: 404-417.

Broek, D.V.E.L., E.M.V. Rikxoort and T.E. Schouten, 2005. Human-Centered Object-Based Image Retrieval. In: International Conference on Pattern Recognition and Image Analysis, Sameer, S., S. Maneesha, A. Chid and P. Perner (Eds.). Springer, Berlin, Germany, ISBN: 978-3-540-31999-3, pp: 492-501.

Hoiem, D. R. Sukthankar, H. Schneiderman and L. Huston, 2005. Object-Based image retrieval using the statistical structure of images. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, July, 2004, USA., pp: 490-497.

Jing, F., M. Li, L. Zhang, H.J. Zhang and B. Zhang, 2003. Learning in Region-Based Image Retrieval. In: International Conference on Image and Video Retrieval, Bakker, E.M., S.L. Michael, S.H. Thomas, N. Sebe and X.S. Zhou (Eds.). Springer, Berlin, Germany, ISBN: 978-3-540-45113-6, pp: 206-215.

Kavitha, K. and M.V. Sudhamani, 2014. Object based image retrieval from database using combined features. Proceedings of the 2014 5th International Conference on Signal and Image Processing (ICSIP), January 8-10, 2014, IEEE, New York, USA., ISBN:978-0-7695-5100-5, pp: 161-165.

Sharawy, G.A., N.I. Ghali and W.A. Ghoneim, 2009. Object-based image retrieval system based on rough set theory. IJCSNS Int. J. Comput. Sci. Network Sec., 9: 160-165.

Sharma, S. and A.M. Siddiqui, 2014. Image retrieval using speeded up robust feature: An effort to improvement. Intl. J. Comput. Sci. Network Secur., 14: 102-107.

Strutz, T., 2016. Data Fitting and Uncertainty. 2nd Edn., Springer, Vieweg, Berlin, Germany.

Wang, J.Z., 2000. SIMPLIcity: A region-based retrieval system for picture libraries and biomedical image databases. Proceedings of the 8th ACM International Conference on Multimedia, October 30-November 3, 2000, ACM, New York, USA., ISBN: 1-58113-198-4, pp: 483-484.