Relational Semantic Indexing Based Retrieval of Image Using Annotation Techniques

S. Sutha¹ and Dr. C. Kavitha²

1 Assistant Professor, Department of Computer Applications, K.S.R.College of Engineering, Namakkal, Tamilnadu-India.

2 Assistant Professor, Department of Computer Science, Thiruvalluvar Government Arts College, Rasipuram-India

smsutha@yahoo.co.in, kavithachellappan@yahoo.co.in

ABSTRACT

With the rapid growth in the image count and advancement in digital imagery, there is an increasing need for effective indexing and sharing of images based on the user needs. Inductive algorithm for image annotation (IAIA) integrated the features of label correlation and visual similarity mining. Though, IAIA minimized the prediction error but the quality of image search was not improved for different combination of keywords. To improve the quality of search for different combination of keywords in the image annotation framework, Relational Semantic Indexing (RSI) technique is proposed in this paper. The segmentation process is carried out initially using gray intensity co-occurrence matrix in image annotation framework using the threshold mean method. With the segmented images in image annotation framework, the matrix structure based feature extraction uses the minimum distance classification to reduce the time taken for classification. The classified image on the annotation framework uses the RSI technique to reduce the classification time for different combination of images. Next, a relational database is created to index different combination of semantic keywords and to improve the quality of search. The entity relationship for RSI technique is also developed to easily identify the relationships with different combination of semantic keywords and the application of linear quad tree in RSI technique reduces the search time considerably. Image Annotation framework based on the RSI technique performs the experiment on the factors such as search time, quality of search rate and precision rate.

Keywords: Image Annotation Framework, Segmentation, Classification,

Keywords, Feature extraction, Relational Semantic Indexing, Threshold mean.

01. INTRODUCTION

With the availability of increasing number of images from different multimedia devices, effective mechanisms for classifying, penetrating and skimming these images are of high significant by common users. Usually, such images should undergo with indexing with the help of semantic descriptions so that the conventional methods of retrieval of images might be applied for searching of image in a precise manner. However, as it is highly impossible to annotate those images manually, automatic image retrieval in image annotation framework might be a promising solution.

In the preceding years, certain amount of research works has been concentrated on image annotation. Early work IAIA in [2], combined both the features of label correlation and visual similarity mining. With the advantage of integration, though IAIA reduced the error during prediction but the quality of image search was not improved for different combination of keywords.

2.0 RELATED WORK:

Recent years have substantiated the growth of social media and the prosperity of huge number of photo-sharing websites, including Flickr and Picasa and so on. These websites, Flickr, Picasa helps the users by rating and providing tag to the shared images. A semi automatic framework for image annotation using Locality Sensitive Hashing (LSH) [2] improved the searching combination of keywords. With the application of LSH, though search effectiveness was improved, search time increased with different combination of image annotation.

In [15], contents of the image and its corresponding textual information in the social media were approximated using semantic representations for two modalities. The framework was designed in such a way that each image was augment with relevant semantic features by applying graphs between various images. Though the rate of image retrieval was higher, more concentration was not made to the emerging social media.

Recent researchers have shown that the designing of manual tags are often insignificant and not reliable. Moreover, as many users select the most general and ambiguous tags for minimizing their involvement while selected more appropriate words, tags are considered to be noisy. To provide solution to this problem, tag completion in [12] filled the missing tags in an automatic manner and also corrected noisy tags for the images provided as input resulting in significant improvement but tag completion based on compressed sensing and matrix completion remain unaddressed.

Most methods involved in image annotation only used single photo at a time and the photos were then labeled in an individual manner. In [13], focus was made on collection of photos and used the contextual information naturally by improving both the precision and recall value. But little measures were taken to minimize the higher level of confusion between certain events. In [16], a novel image retrieval mechanism

was presented for medical images where the images from the retrieval database were efficiently annotated with the help of the pre-defined labels that resulted in the improved of retrieval factor. The higher retrieval factor was only possible with pre-defined labels but dynamism was not achieved.

The biggest challenges in multi-view imaging include a proper definition with clear representations that provides the user with the intrinsic geometry for the visual information. These geometric images can be approximated using sparse image representation with a precise selection of a good dictionary. A method was presented in [3] with sparse representation of stereo images for studying joint representation of stereo images. Though multi view feature matching was improved it can be applied only on omnidirectional images.

A novel method was presented for graph indexing [17] for better image retrieval using hypergraph which improved the scalability of images being retrieved. A structural learning algorithm was instantiated in [18] for implementing large-scale image classification and annotation using structural learning algorithm. But the results, accuracy obtained for large scale classification was not precise.

With the increasing use of Internet, the rate at which the archival multimedia data is also increasing with the broad application of digital video hardware. In [4], a corner based approach was presented which detected text and caption from videos in an efficient manner using the decision tree for learning the classification criteria. Though robustness was provided, focus was not made on word segmentation and text recognition.

In [5], a hybrid method was presented to robustly identify and localize the texts present in natural scenery images. Moreover, with the application of conditional random field (CRF), the non-text elements were filtered out in an efficient manner by considering both unary and binary component properties. But the time taken to filter with larger amount of images increased correspondingly.

The rest of this paper is organized as follows. In Section 2, we introduce image annotation framework based on relational semantic index for automatic image retrieval. The linear quad tree construction is presented for image annotation framework. Section 3 provides an experimental evaluation with the detailed discussion involved in Section 4. Section 5 briefly reviews some of the related work in image annotation by different researchers. Finally Section 6 concludes with the concluding remarks.

3.0 DESCRIPTION OF PROPOSED ALGORITHM:

3.1 Relational Semantic Indexing

In this section, we first briefly explain about the design consideration of image annotation framework based on relational semantic indexing. The main goal of Relational Semantic Indexing technique in the image annotation framework is to reduce the search time for different combinations of semantic keyword. The RSI technique also works to improve the quality of search by developing an entity relationship model. To develop the relationship between different semantic keywords, RSI technique is developed on image annotation framework.

In this paper, an image annotation framework for automatic image retrieval is developed for training a large number of images using multiple word phrases jointly. The following highlights some of the main aspects of our proposed work

- 1. Our minimum distance classification that includes different class centers effectively classifies the grey-level vector points with minimal runtime. It is worth noting that the grey-level vector points retrieve all combination of objects by reducing the search time.
- 2. Our linear quad tree construction can significantly reduce the search time by partitioning the two dimensional space into four quadrants.
- 3. The application of relational semantic indexing dramatically encapsulates the function for varying combination of keywords.

3.1 Threshold Mean Segmentation Method

Initially, the threshold mean segmentation method is used for segmenting different size images. The threshold means method is implemented on the proposed image annotation framework for effective discrimination of foreground from the background images. A threshold value with 'T' is selected and followed by this the gray level image is converted into binary image. The converted binary image contains all the essential information about the location and outline of the objects for segmenting the images with minimal time.

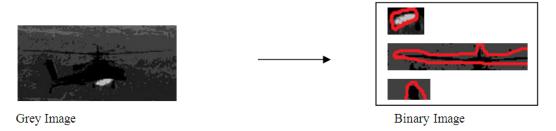


Fig- 1 Segmentation using Threshold Mean

Fig 1 clearly segments different parts of the aircraft grey images into the binary images. With the threshold mean value, the objects position and shape is also identified. The threshold mean segmentation method identifies the intensity characteristics of the size of the objects and number of objects present in the grey image. The threshold mean value of the pixels is used for segmenting the objects for RSI based image annotation framework.

3.2 Feature Extraction process

With the image segmented using threshold mean segmentation method, features are extracted that forms the representative of various classes of objects for RSI based image annotation framework. Gray Intensity Co-occurrence matrix is used for extracting the spatial relationship of annotated image pixels. The spatial relationship in the image annotation framework contains the (x, y) elements and these elements is

used for evaluating the co-occurrence matrix. Gray Intensity Co-occurrence matrix identifies the pixel with value 'x' that occurred in specific spatial relationship to a pixel with value 'y'.

Cooccurence matrix
$$\left(C_{\Delta x}C_{\Delta y}(i,j)\right) = \sum_{g=1}^{n} \sum_{h=1}^{m} \begin{cases} 1 & \text{If } (g,h) = i \text{ and } I(g + \Delta x, h + \Delta y) = j \\ 0 & \text{otherwise} \end{cases}$$
 (1)

With annotated image I taken for consideration, 'i' and 'j' are the image intensity values of the annotation image with 'g' and 'h' represent the spatial positions of the annotated image I. Δx , Δy in Gray Intensity Co-occurrence matrix depends on the path and the space of the object in the spatial relationship. The extracted features are provided as the input to the classification process in the image annotation framework. The classification process is briefly explained in the subsequent section.

3.3 Minimum Distance Classification

With the features being extracted, the RSI technique uses the minimum distance classification that contains different types of class centers such as, $C_i = 1,2 ... n$. In general to classify the objects from the grey-level vector 'v' points, the minimum distance MinDistance, is described as,

$$MinDistance = \sqrt{(Cooccurence\ matrix - v)^t(Cooccurence\ matrix - v)^{1/2}}$$
(2)

RSI Process:

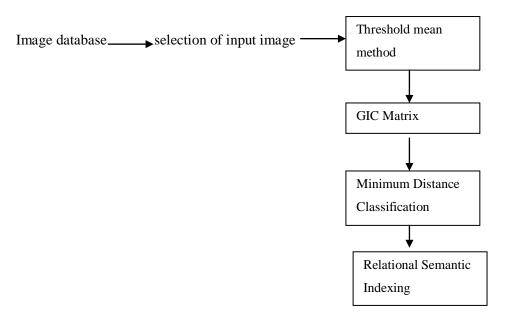


Fig 2 Flow diagram of RSI based Image Annotation

As illustrated in Fig 2, with the application of RSI based image annotation framework, the search quality can be improved for different combination of semantic keywords. The keyword retrieves all combination of objects from the classified class and reduces the search time using the relational semantic indexing. The classified data objects initially perform the segmentation and feature extraction process. Followed by this, the classified object class is now indexed into the relational table for easy retrieval of quality information according to the requests placed by the users'. An elaborate discussion of relational semantic indexing is provided in the forthcoming subsection.

3.4 Relational Semantic Indexing

The application of Relational Semantic Indexing enables to register secondary class method at the database server for building an index structure. The semantic keyword relational indexing encapsulates the function for retrieving different combination of keywords. The relational semantic functions are listed in the table given below,

Function	Task
Rel_Index_Create ()	Create relational semantic index structure
Rel_Index_Open()	Open index structure for operations
Rel_Index_Fetch()	Fetch relational information based on user keywords
Rel Index Close ()	The relational index closes the function after image retrieval

Table 1 A model of relational semantic functions

A semantic keyword relational indexing method encapsulates the functions for creating, opening, fetching and closing. The row based processing is carried out in the relational indexing method to add different class of objects into the indexing table. Moreover, the indexing table is effectual in fetching different combination of user keywords in RSI technique using the different functions as provided in table 1. RSI technique map the index structure to fetch the different combination of keywords using built in relational (i.e.,) pixel values. The relational values are designed in such a way that it operates on top of the relational query language.

3.4.1 Entity Relationship model

The entity relationship model in RSI technique takes the functions provided in table1 to plot the relationship diagram.

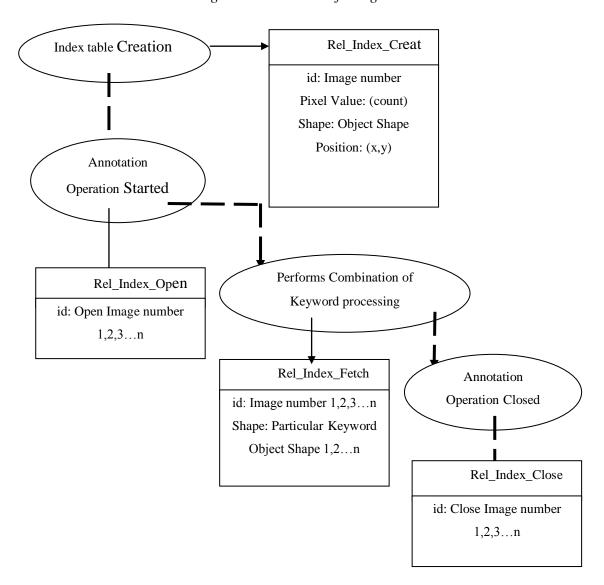


Fig 3 Entity Relationship model in RSI technique

Each image annotation framework uses the entity relational model followed in Fig 3 to fetch the semantic key word based images to the users using the RSI technique. The most frequently used function in the relational indexing in the Rel_Index_Fetch. With the use of the functional, Rel_Index_Fetch, the indexing significantly simplifies the image fetching process in the RSI based image annotation framework for different combination of semantic keywords. In addition, the relational semantic indexing uses the linear quad tree for partitioning the two dimensional space into four quadrant in order to reduce the search time. The algorithmic steps involved in the design of linear quad tree is described below

3.4.1.1// Linear Quad tree Construction

Begin

Step 1: Resolution of image size is initialized

Step 2: $C_{x,y}^*$ denote linear code of pixels

Step 3: $x \leftarrow 0$, $y \leftarrow 0$, $\Delta x \leftarrow -1$, $\Delta y \leftarrow -1$

Step 4: Partition the coordinated into four partitions in square form

Step 5: If (x, y) = indexed table value

Step 6: Fetch the objects from relational semantic indexing table

Step 7: End If

Step 8: Repeat step 5 to 7 on all partitioned coordinates

Step 9: Semantic keyword related objects fetched from the stored database images

End

Let us consider a linear quad tree image with of n*n sized pixels for performing relational semantic indexing. The linear code pixels are initially evaluated for easier search result with minimal time consumption. Linear quad tree in RSI technique perform the image annotation process for different resolution images from the database.

3.4.2 Semantic Keyword Addition in RSI

The semantic keyword addition on image annotation framework takes 'n' combination of keywords to the fetch the user result. The probability of event on the keywords 1,2,... are selected as,

Keyword Combination =
$$\sum_{i=1}^{n} (K|(x,y))$$
 (3)

The above equation proceeds with the assumption that all keywords 'K' are assigned autonomously and RSI technique fetches the user result with higher precision rate. Relational indexing handles different combination of keywords with (x, y) pixel points for effectual quality of search. The combinatorial interpretation in RSI technique chooses 'n' semantic keywords and result with potentially high search result on larger database.

4.0 EXPERIMENTAL EVALUATION

Image Annotation Framework based on the Relational Semantic Indexing (RSI) technique performs the experimental work using MATLAB coding. RSI indexing is experimented using the ImageNet database according to the WordNet hierarchy.RSI is compared against the Inductive Algorithm for Image Annotation (IAIA) and Semi automatic framework with Locality Sensitive Hashing (LSH).

Each meaningful concept in WordNet, possibly described by multiple words phrases named as synset. There are more than 100,000 synsets in WordNet in which

the ImageNet provide on average 1000 images to illustrate each synset. Images of each concept are quality-controlled and human annotated for experimental work. In its completion, ImageNet offer tens of millions of cleanly sorted images for experiment the factors such as search time, recall ratio, precision rate, classification rate on image annotation and quality of search rate.

The Search time using RSI is the time taken to retrieve the image on image annotation framework with respect to clock cycles per image as given below in (4)

$$ST = CCT * No. of images * avg CPI$$
 (4)

The search time (ST) is measured as the product of clock cycle time (CCT) with the number of images given for retrieval $(No.of\ images)$ and the average clock cycles per image $(avg\ CPI)$. It is expressed in terms of milliseconds.

Let us consider an ImageNet database with 1000 images. The Recall ratio (as shown in eqn (5)) based on Relational Semantic Indexing (RSI) measures the number of relevant images retrieved to the total number of relevant images in the ImageNet database. It is expressed in terms of percentage (%).

The precision rate based on Relational Semantic Indexing (RSI) is the ratio of retrieved images that are relevant (as shown in eqn (6)). Let RIR represents the relevant images retrieved, RINR denotes the relevant images not retrieved and IRR denotes the irrelevant records retrieved with 800 images retrieved, the recall ratio and precision rate is given as below in (5) and (6)

$$RR = \frac{RIR}{(RIR + [1000 - RIR])} * 100$$
(5)

$$PR = \frac{RIR}{(RIR + [1000 + (800 - RIR)])} * 100$$
(6)

The classification rate on image annotation is the classification of various annotated images using minimum distance classification as given in (2). The quality of search rate is the rate at which the quality search is provided using RSI on image annotation framework. It is expressed in terms of percentage (%).

5.0 RESULT ANALYSIS OF RSI

The Image Annotation framework based on the Relational Semantic Index (RSI) technique is compared against the existing inductive algorithm for image annotation (IAIA) [1] and Locality Sensitive Hashing (LSH) [2]. The simulation results using MATLAB are compared and analyzed through table and graph form given below.

No. of test images	Search Time (ms)		
(MB)	RSI	IAIA	LSH
100	23	28	31
200	26	32	35
300	32	34	38
400	28	33	34
500	34	38	41
600	37	40	43
700	40	45	48
800	39	42	42

Table 2 Tabulation for search time

Table 2 provides a few statistics on search time with respect to 800 test images for experimental purpose. Consequently, the results obtained using proposed RSI is compared with the existing IAIA [1] and LSH [2].

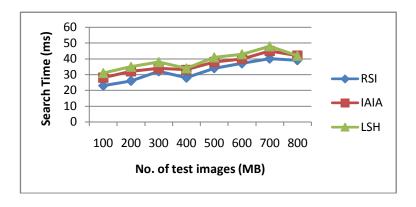


Fig 4 No. of test images versus search time

Fig 4 shows a graph representing the search time using our proposed method RSI technique and comparison is made with two other methods, IAIA [1] and LSH [2]. The graph shows the search time with respect to 800 test images taken for experimental purpose. It is observed that the search time using RSI is better as an entity relationship is developed to identify the relationships with different combination of semantic keywords by minimizing the search time. It is also observed that the performance of search time are affected using the existing IAIA [1] and LSH [2]. This is because the Locality Sensitive Hashing (LSH) though improves the searching combination of keywords but the time of search is increased for different combination of image annotation. This improves the search time using RSI by 6-23% when compared to IAIA. In addition, using RSI, the keyword retrieves all combination of objects from classified class and reduces the search time using the relational semantic indexing by 16-34% when compared to LSH [2].

No. of test images (MB)	Recall (%)		
	RSI	IAIA	LSH
100	76	71	67
200	72	70	66
300	70	68	65
400	74	69	67
500	65	62	60
600	60	57	55
700	58	56	50
800	55	54	48

Table 3 Tabulation for recall

Table 3 shows that while both the existing image annotation methods are stable with different test images, our approach, RSI is able to obtain significantly more accurate results in terms of recall value.

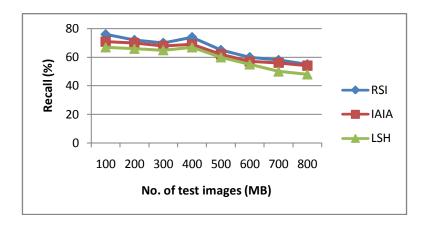


Fig 5 No. of test images versus recall

Fig 5 represents the recall ratio using RSI by varying the number of test images from 100 to 800 for experimental purpose. The simulation shows that the recall rate is higher using the proposed RSI than when compared to the existing inductive algorithm for image annotation (IAIA) [1] and Locality Sensitive Hashing (LSH) [2]. This is because of the application of relational database for indexing different combination of semantic keywords which in turn increases the recall ratio in an optimal manner based on the number of test images by $2-8\,\%$ when compared to IAIA [1]. In a similar manner, with the application of linear quad tree, the two dimensional space is efficiently partitioned into four quadrant form that increases the recall ratio by $7-13\,\%$ when compared to LSH.

No. of test images (MB)	Precision (%)		
	RSI	IAIA	LSH
100	81	73	67
200	77	72	66
300	75	70	65
400	79	71	67
500	70	64	60
600	65	59	55
700	63	58	50
800	60	56	48

Table 4 Tabulation for precision

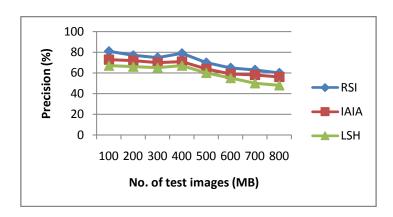


Fig 6 No. of test images versus precision

Table 3 and Fig 6 shows the precision rate for automatic image retrieval using the proposed RSI and comparison is made with two other existing methods namely, inductive algorithm for image annotation (IAIA) [1] and Locality Sensitive Hashing (LSH) [2]. From the Fig 6 it is evident that the precision rate for image retrieval is increased using the RSI though with the increase in the number of test images the precision rate decreases. But comparatively, the precision rate is higher using the RSI technique. This is because with the application of semantic keyword addition on image annotation framework, it takes 'n' combination of keywords to fetch the user result and therefore increases the precision rate by 7-10% when compared to IAIA. In addition, with the assumption that all K keywords are assigned autonomously, RSI technique fetches the user result with higher precision rate by 14-20% when compared to LSH [2].

Class Centers (n)	Classification rate on image annotation (%)			
	RSI	IAIA	LSH	
2	45	38	38	
4	52	42	39	
6	55	49	44	
8	58	54	48	
10	63	59	53	
12	67	62	57	
14	72	64	61	
16	76	68	63	

Table 5 Tabulation for classification rate on image annotation

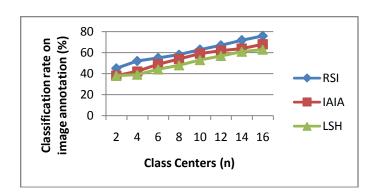


Fig 7 No. of class centers versus classification rate on image annotation

To evaluate the classification rate on image annotation, several experiments were conducted using Matlab as shown in table 5 and Fig 7. The number of class centers (i.e., n) taken into consideration was 2 to 16 class centers based on different interval of time. The results show that the classification rate on image annotation is significantly improved using RSI technique. The Fig 7 shows that with the increase in the number of class centers, the classification rate on image annotation is also increased. This is due to the fact that with the application of minimum distance based classification, the classification of various annotated objects are performed in an efficient manner by 6-16 % when compared to IAIA [1] and 14-25 % when compared to LSH [2].

6.0 CONCLUSION

This work advances automatic image retrieval using relational semantic indexing technique on image annotation framework. First, it plays a significant role in the quality of search for different combination of keywords and hence reduces the classification time for different combination of images. This is achieved by building an entity relationship model in image annotation framework by constructing linear quad tree. With the application of minimum distance classification, the tasks of

classifying various annotated objects get reduced with time using RSI technique. Second, with the introduction of relational semantic indexing, is effectual in fetching the user combination of keywords in RSI technique using the functions. Moreover, using the entity relationship model, search time is reduced with the application of partition of two dimensional spaces into four quadrant form. Experimental results demonstrate that the proposed RSI technique outperforms two existing works by improving the classification rate by 25 % and increases the precision rate by 20 %.

REFERENCES

- [1] Yi Yang, Fei Wu, Feiping Nie, Heng Tao Shen, Yueting Zhuang, and Alexander G. Hauptmann., "Web and Personal Image Annotation by Mining Label Correlation With Relaxed Visual Graph Embedding., " IIEEE Transactions On Image Processing, Vol. 21, No. 3, March 2012
- [2] Jinhui Tang., Shuicheng Yan., Chunxia Zhao., Tat-Seng Chua., Ramesh Jain., "Label-specific training set construction from web resource for image annotation," Signal Processing., Elsevier Journal., 2012
- [3] Ivana sic, and Pascal Frossard," Dictionary Learning for Stereo Image Representation", IEEE Transactions On Image Processing, VOL. 20, NO. 4, APRIL 2011
- [4] Xu Zhao, Kai-Hsiang Lin, Yun Fu, Yuxiao Hu, Yuncai Liu, and Thomas S. Huang," Text From Corners: A Novel Approach to Detect Text and Caption in Videos", IEEE Transactions On Image Processing, Vol. 20, No. 3, March 2011
- [5] Yi-Feng Pan, Xinwen Hou and Cheng-Lin Liu," A Hybrid Approach to Detect and Localize Texts in Natural Scene Images", Ieee Transactions On Image Processing, Vol. 20, No. 3, March 2011
- [6] Horng-Horng Lin, Jen-Hui Chuang, and Tyng-Luh Liu," Regularized Background Adaptation: A Novel Learning Rate Control Scheme for Gaussian Mixture Modeling", IEEE Transactions On Image Processing, Vol. 20, No. 3, March 2011
- [7] Yu-Gang Jiang, Qi Dai, Jun Wang, Chong-Wah Ngo, Xiangyang Xue and Shih-Fu Chang," Fast Semantic Diffusion for Large-Scale Context-Based Image and Video Annotation", IEEE Transactions On Image Processing, Vol. 21, No. 6, June 2012
- [8] Jing Liu , Bin Wang , Hanqing Lu , Songde Ma," A graph-based image annotation framework", Pattern Recognition Letters 29 (2008) 407–415, Elsevier
- [9] Yohan Jin, Kibum Jin, Latifur Khan, B.Prabhakaran," The Randomized Approximating Graph Algorithm for Image Annotation Refinement Problem", IEEE Computer Society on Computer Vision and Pattern Recognition, Jan 2008
- [10] Min Wu, and Bede Liu," Data Hiding in Binary Image for Authentication and Annotation", IEEE Transactions On Multimedia, VOL. 6, NO. 4, AUGUST

2004

- [11] Antonio Torralba, Bryan C. Russell, and Jenny Yuen," LabelMe: Online Image Annotation and Applications", IEEE Proceedings, Vol. 98, No. 8, August 2010
- [12] Lei Wu, Rong Jin, Anil K. Jain," Tag Completion for Image Retrieval",IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. Xx, No. Xx, January 2011
- [13] Liangliang Cao, Jiebo Luo, Henry Kautz and Thomas S. Huang," Image Annotation Within the Context of Personal Photo Collections Using Hierarchical Event and Scene Models",IEEE Transactions On Multimedia, Vol. 11, No. 2, February 2009
- [14] S. Navid Hajimirza, Michael J. Proulx, and Ebroul Izquierdo, Senior, IEEE," Reading Users' Minds from Their Eyes: A Method for Implicit Image Annotation", IEEE Transactions on Multimedia, Jan 2012
- [15] Yin-Hsi Kuo, Wen-Huang Cheng, Hsuan-Tien Lin, and Winston H. Hsu," Unsupervised Semantic Feature Discovery for Image Object Retrieval and Tag Refinement", IEEE Transactions On Multimedia, Vol. 14, No. 4, August 2012
- [16] Jian Yao,_, Zhongfei (Mark) Zhanga, Sameer Antani, Rodney Longb, George Thomab," Automatic medical image annotation and retrieval", Neuro Computing, Elsevier, Mar 2008
- [17] Salim Jouili, SalvatoreTabbone," Hypergraph-based image retrieval for graph-based representation", Pattern Recognition, Elsevier April 2012
- [18] Peixiang Dong, KuizhiMei, NanningZheng, HaoLei, JianpingFan," Training inter-related classifiers for automatic image classification and annotation", Pattern Recognition, Elsevier, Oct 2013.

BIOGRAPHY:



S.Sutha is currently working as Assistant Professor in the Dept of MCA at K.S.R College of Engineering. She received her BSc and MCA Degree from Periyar University and MPhil degree from Bharathiar University. Presently she is a Research Scholar from AnnaUniversity.Her Research interest includes image processing, Retrieval and Networks. She is a Member of ISTE.



Dr.C.Kavitha is currently working as Assistant Professor in the Dept of Computer Science at Thiruvalluvar Government Arts College-Rasipuram.Her area of interest is image retrieval and processing, Data mining. She is a Member of ISTE.