

Semantic Based Image Annotation Using Descriptive Features and Retagging approach

P.Nagarani¹, R.VenkataRamanaChary² and Dr.D.Rajya Lakshmi³

¹Padmasri Dr.B.V Raju Institute of Technology, Hyderabad, India
nagarani.444@gmail.com

²Padmasri Dr.B.V Raju Institute of Technology, Hyderabad, India
rvrchary@gmail.com

³GITAM Institute of Technology, Visakhapatnam, India
rdavuluri@yahoo.com

Abstract:

The Semantic based annotation of an image is very important and a difficult task in content-based image retrieval (CBIR). The low-level features of the images are described using color and the texture features and the proposed model is used for semantic annotation of images. Also the textual annotations or the tags with multimedia content are the most effective approaches to organize and to support search over digital images and multimedia databases. The quality of the tags was refined using Image retagging method. The process is given as a multi-path graph based problem, which in parallel identifies the visual content of the images, semantic correlation of the tags as well as the primary information provided by users.

The image annotation preferred because as the countless images exist in our lives it is not possible to annotate them all by hand. And so annotation by computer is a potential and promising solution to this problem precisely.

The ability to annotate images semantically based on the objects that they contain is essential in image retrieval as it provides the mechanism to take advantage of existing text retrieval systems.

Keywords:

Image Annotation, Image Retagging, Descriptive Features, multi-path graph based labeling method.

1. Introduction:

Capturing of the digital images using multimedia devices and the digital cameras is becoming increasingly more popular. Annotating the digital pictures is important to support the browsing, i.e. search pattern and retrieval of the images. Image annotation techniques depend entirely on visual properties of the image. The image annotation systems of this type work well in detecting base object classes: car, horse, motorcycle, airplane, etc.

However, certain characteristics of the image are hard to capture using strictly the visual properties, such as events, people, location etc.

Content-based image retrieval (CBIR) has been studied and explored for almost decades. CBIR retrieves images in terms of their visual features, while users often prefer intuitive text-based image searching. As there is a rapid increase of digital images on the Internet, the earlier manual annotation cannot meet the actual needs. The study of automatic semantic image annotation in image retrieval approach has become a key concept.

Image semantic annotation was intended to add keywords automatically for images to represent the semantic content for the image; the visual features can be changed into the image annotation word information, involves not only the efficient search of keywords, but also overcomes the disadvantages of manual annotation in terms of time consumption. Here the Low-level semantic features which depend on the visual content are to be described. The High-level features do not pertain to the visual content.

Numbers of methods have been proposed for content based image retrieval (CBIR). CBIR systems depend or rely on low-level features which can be automatically extracted and used to index images. The CBIR systems fail to meet user expectations because those systems are unable to index images according to the high level features as perceived by the user.

The fast growth in the development of digital image processing, database, and internet technologies, image classification and analysis are challenging problems in content-based image retrieval system (CBIR).

Most studies in this field have focused on how to group images into semantically meaningful categories or index image in a database based on low-level visual features of the images. These systems follow the paradigm of representing image via a set of features such as color, texture, and shape. However, it is difficult to describe high-level semantics using low-level features.

Semantic gap is the gap between human perception and low-level features. The gap is balanced using high-level semantics to index and retrieve images. Semantics have been used to describe low-level image features such as color, texture and shapes of objects.

The “multi-path graph” label propagation in the proposed framework focus is to characterize the particular relationship among the images with respect to a certain tag.

In the presented paper, a new interactive and useful framework was developed following two concepts. The first one is to have a more interaction to the user providing more functionalities in accordance to the most used and popular tags or the descriptive words. The second one, which is the actual intent of the work, is to try to reduce the semantic gap providing to the user some easy mechanisms to express his/her semantics.

As such the gap will slightly reduced because the high level features will be better defined and identified for each multimedia content which can also be specified by low level features automatically extracted.

2. Related work:

The earlier research on improving unreliable descriptive keywords of images to date has focused on annotation refinement, i.e., identifying and eliminating the imprecise annotation keywords produced by the automatic image annotation algorithms.

As an extension work, the estimation of the semantic correlation among the annotated keywords and then those weakly-correlated ones are removed. However, this particular method can only achieve limited success since it totally ignores the visual content of the images.

To address the problem, another work proposed a content-based approach to re-rank the automatically annotated keywords of an image and only reserve the top ones as the refined results. Keeping aside the efforts, the existing methods primarily focus on selecting a coherent subset of keywords from the automatically annotated keywords.

The proposed work is first or initially related to the automatic annotation of an image also known to be the tagging, a process to identify the tags associated with an image. Many methods and algorithms were proposed before for accomplishing this work or the task.

The main idea and the scope that includes in the proposed framework is the mapping between the image contents and the semantic tags based on the collection of the labeled set of images and using a different model to predict or identify the unlabelled tags.

But on the contrary, the main concept of the work is not to predict tags for the unlabeled images, but rather to refine the imprecise tags provided by the users. But we do not have the imprecise tags always in the set of the images. Instead, the only available supervision to learn the improved tag assignment in our scenario is the imprecise, incomplete or subjective tags provided by the users.

Image retagging is the more challenging task for accomplishing this task effectively.

On the other hand, the tags associated with social images are often imprecise and incomplete, and thus, their descriptive capability to the visual content cannot be guaranteed. Therefore, the reranking-and-removing strategy in annotation refinement methods is not applicable in the image retagging scenario.

To address this difficulty, another work proposed to refine the tags based on the visual and semantic consistency residing in the social images, which assigns similar tags to visually similar images.

However, it still uses the holistic image similarity to model the image relationship and thus only results in limited performance improvements

Recently, several methods and approaches have been proposed to construct discriminative visual sub-vocabularies, which explicitly incorporate the category-specific information.

3. Proposed Framework:

The presented framework concept is based on the descriptive features such as the content description and the structural descriptions and the low-level features. Prior to the image retrieval of the system involving the annotation and retagging approach the images are to be stored in the database.

The following framework involves:

- Image Annotation using the Low-level semantic Features
- Image Retagging Approach

3.1 Image Annotation:

The aim of annotation methods is to attach textual labels to un-annotated images or the unlabelled images, as the descriptions of the content or objects in the images. The final goal of image annotation is mostly to perform image retrieval by providing users with a text based interface for search.

The image annotation preferred because as the countless images exist in our lives it is not possible to annotate them all by hand. And so annotation by computer is a potential and promising solution to this problem precisely.

The ability to annotate images semantically based on the objects that they contain is essential in image retrieval as it provides the mechanism to take advantage of existing text retrieval systems.

3.2 Image Retagging Approach:

The high-level meanings can be associated to images or image regions through image tagging, also known as captioning or annotations. Tagging improves the content of images and helps image retrieval search engines to better retrieve desired images in response to text queries.

Images can be tagged with a variety of descriptions, keywords and structured metadata. While image tags are a set of keywords, metadata is a structured way of expressing the image descriptions.

3.3 Block diagram of the Proposed Framework:

The tag-specific image representation plays an important role in the proposed framework. Since the initial tags of the social image collection contains a lot of noises, learning such similarity of tag-specific images is a challenging task. Specifically, the Descriptive Features are used to represent the visual content of the images and then learn a tag-specific visual vocabulary for each tag. The purpose of the visual sub-vocabulary related to a target tag is to select the most informative visual words to represent the corresponding tag.

Therefore, two conditions are required for selecting the words which are visible:

1. The words which are visible in the tag-specific visual Sub-vocabulary should appear many times in images which are labeled with the given tag than in images without specifying the tags

2. Those were able to specify or locate on the particular location/object although location/object is surrounded by non-visible background.

The figure-1 Block Diagram represents the overview of the process involving in the proposed framework.

A huge collection of the captured images or the digital images were collected and stored in the database and were considered as the Input Query. The images were provided initially with the user-specified tags. The Images were then categorized into different groups based on the context, region etc. The retrieval of the images is done based on the descriptive features known to be the semantic features. Then the Image Retagging approach is done to annotate the images with the precise tags. Finally the retagged Image results were found with more relevant information.

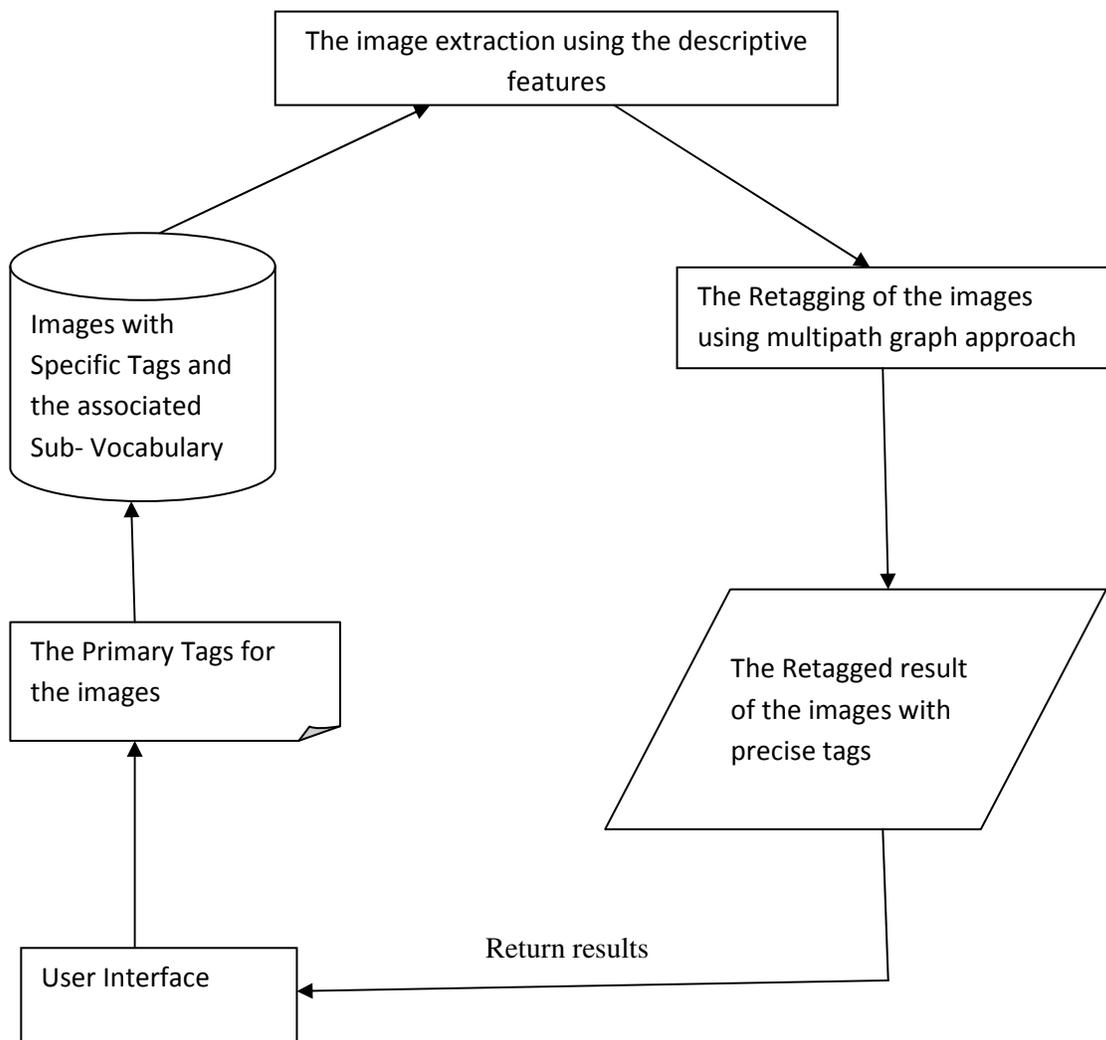


Fig 1. Block Diagram

3.4 Multipath Graph:

A graph-based method was proposed and used to improve the quality of tags by providing the visual appearance of the images over multiple tag-specific similarity graphs, semantic consistency of the tags as well as initial information provided by users. The multiple tag-specific similarity graphs in a collaborative way is performed. To construct tag-specific similarity multipath graph, we propose a robust method to learn tag-specific visual sub-vocabulary from a collection of social images with noisy tags is observed.

The following figure illustrates as:

Each database depending upon different category is provided with a graph and altogether combined or collaborated using the multi-path similarity graph method.

The multi-graph label propagation in the proposed framework focus is to characterize the particular relationship among the images with respect to a certain tag.

The Image Retagging approach involves the multi-path graph which can be observed as follows:

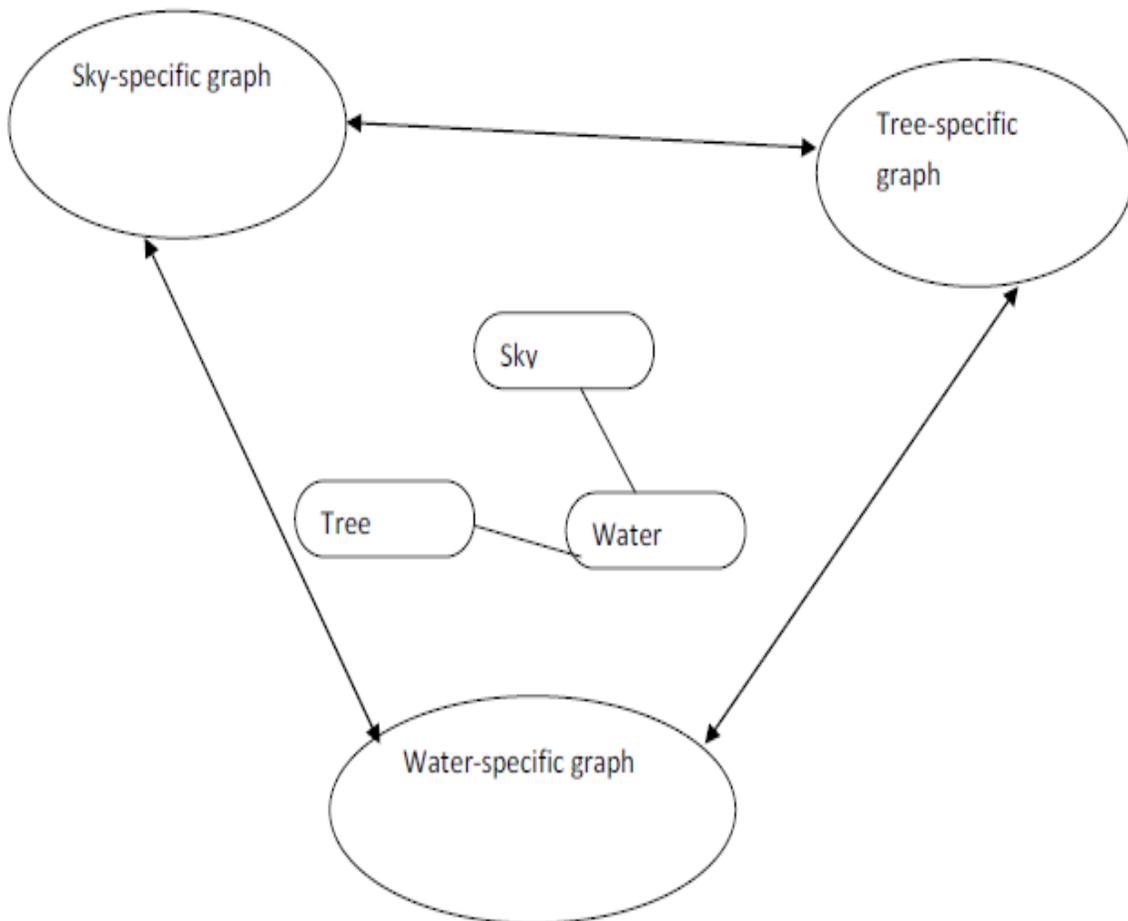


Fig.2.Retagging Approach involving multi-path graph

4. Methodology:

The proposed method is observed by the following real-world social images.

- 1) The images that can be seen (or) visually similar images often reflect similar theme and thus are typically annotated with similar tags. The observation, referred to as visual consistency, has been widely explored in visual category learning but the correlations among tags are generally not utilized.
- 2) The tags associated with social images do not appear in isolation. Instead, they appear correlatively and naturally interact with each other at the semantic level. For example, the presence of tag “dog” often occurs together with the presence of tag “animal” while rarely co-occurs with “vegetable”. It is referred to as semantic consistency assumption. It also has been applied in some multi-label or contextual learning algorithms
- 3) With the general knowledge that human-beings share most of common concepts in the semantic space, the user-provided tags of an image, despite imperfect, still reasonably reveal the primary semantic theme of the image content.

The tags can be regarded as prior information and use them to guide our retagging process. Actually, image retagging can be regarded as a multi-label learning problem from imperfectly-labeled image set, since each social image is typically associated with multiple (noisy) tags.

The steps involved in the proposed framework were:

Step 1: A large collection of digital images were taken as an Input

Step 2: The Images were categorized into different groups and are given user-specified tags

Step 3: The images were then annotated and classified depending upon the Semantic description
(Low-level features)

Step 4: Also the Image Retagging approach is used to identify the correct set of data related to the images

Step 5: Finally the retagged Image results were found with more relevant information.

The figure-3 gives the precise idea of the proposed framework in a graphical manner involving the steps that take place in the process of the framework.

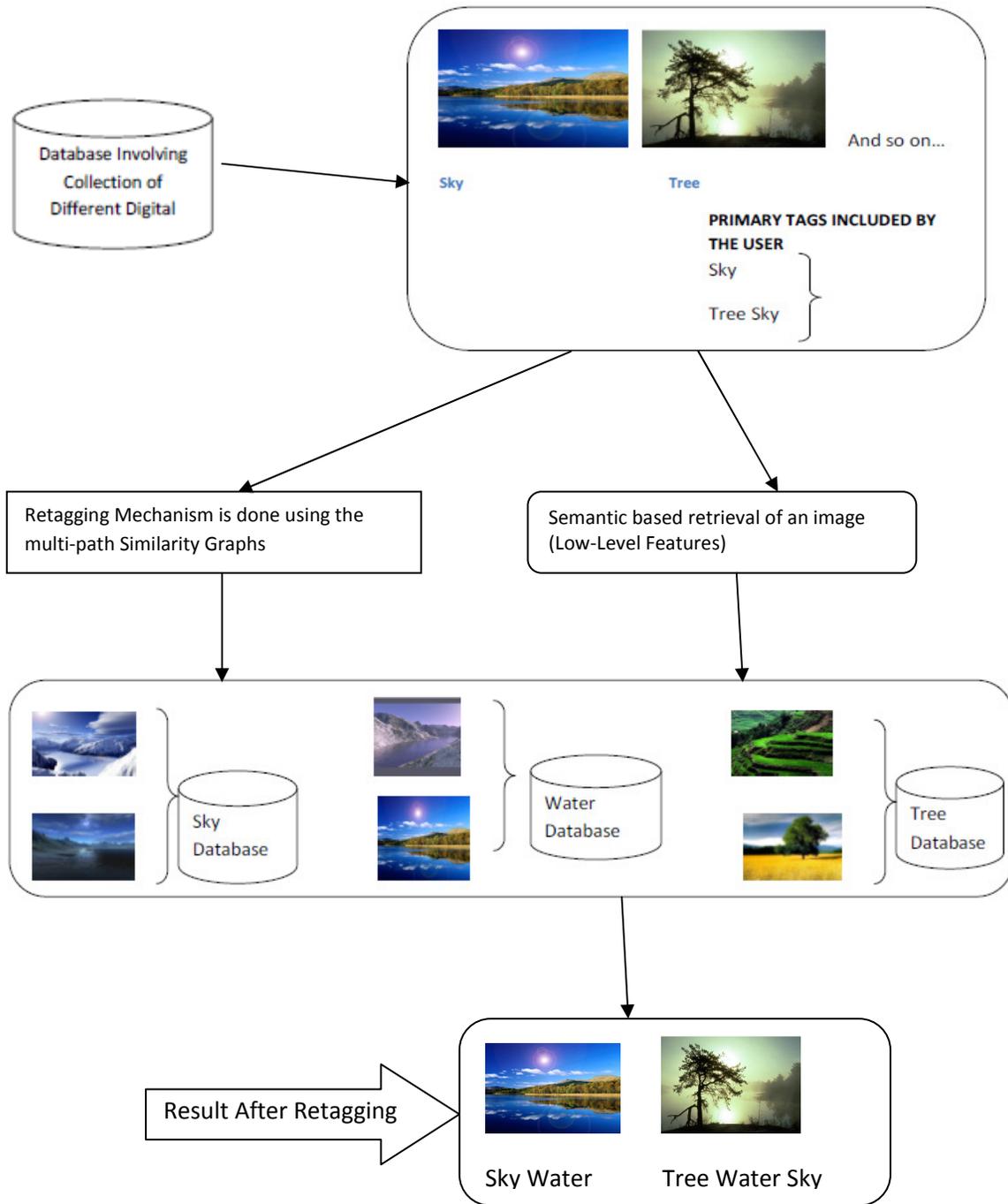


Fig. 3. Annotation of the Image involving the Retagging Approach

5. Problem Formulation:

A training sample of images were considered, **Where $S=\{X_1, X_2, X_3, \dots, X_n\}$, 'n' is the size of the image set.** All unique initial tags appearing in this collection form a tag set,

$P= \{ Y_1, Y_2, Y_3, \dots, Y_m\}$, where 'm' denotes the total number of unique tags.

The user-defined tags for the whole image collection can be presented in a **binary matrix**,

$$\mathbf{B} \in \{0, 1\}^{m \times n}$$

whose element indicates the presence of tag in image. To represent the final retagging results, another matrix is defined.

'**B**' can be expressed as:

$$\mathbf{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots, \mathbf{b}_n \},$$

$$\text{and } \mathbf{b}_k = \{ \mathbf{b}_{1,k}, \mathbf{b}_{2,k}, \mathbf{b}_{3,k}, \dots, \mathbf{b}_{M,k} \}^T,$$

Where **T** denotes matrix transpose.

Here, \mathbf{b}_n can be interpreted as a binary encoded tag-based representation of the image I_m .

The problem formulation is based on the descriptive features such as the content description and the structural descriptions and the low-level features.

The following framework involves:

- Image Annotation using the Low-level semantic Features
- Image Retagging Approach

The Image re-tagging scheme/method involves considering the Initial tags of the images which are user-defined and the images which are not provided with the tags and then labeling them without even considering the clutter background.

To represent the visual content of an image, the Proposed model is used, which was given to capture both color and texture information of small patches in natural color images, and has been successfully applied to image coding, indexing, and retrieval.

Then the following is done:

1. Select a Image I_m with probability $P(I_m)$ from S
2. Sample a Image Category C_k with probability $P(C_k | I_n)$ from the selected image I_n

3. For each of the images, Sample S_n from a Image Database conditioned on the image category C_k .
4. For each of the selected image, the annotation is done along with the retagging process and the relevant images with precise tags are observed.

The similarity of the tags-specific information considered as the Image pair $(I_m | I_n)$ specified above:

$$\text{Equality } (I_m | I_n) = \sum_{k=1}^K b_{k,m} b_{k,n} F_k$$

F_k - denotes the inverse frequency of the word w_j , calculated using the equation:

$$F_k = 1 + 1/N_k$$

N_k - denotes the number of images annotated per word w_j

The **Image Retagging** consists of the following:

$R_l(Y)$ - Loss term (To specify the User-defined tags deviation)

$R_s(Y)$ - Regularization term (To specify semantic consistency)

$R_v(Y)$ - Regularization term (To specify semantic consistency)

The Results can be formulated by using the minimization function as:

$$\text{Min } (F(Y)) = R_l(Y) + R_s(Y) + R_v(Y)$$

6. Conclusions:

The presented framework concept is primarily to attach textual labels to un-annotated images or the unlabelled images, as the descriptions of the content or objects in the images. The final goal of image annotation is mostly to perform image retrieval by providing users with a text based interface for search.

The main idea and the scope that includes in the proposed framework is the mapping between the image contents and the semantic tags based on the collection of the labeled set of images and using a different model to predict or identify the unlabelled tags. But on the contrary, the main concept of the work is not to predict tags for the unlabeled images, but rather to refine the imprecise tags provided by the users.

The proposed framework extracts relevant images from the huge collection of database and refines the quality of the tags .Finally the precise annotated images with refined tags were observed.

References:

- [1] A semantic annotation algorithm based on image. International Conference on Computer Science and Software Engineering.
- [2] IEEE Transactions on Supervised Learning of Semantic Classes for Image Annotation and Retrieval
- [3] Automatic linguistic indexing of pictures by a Statistical modeling approach IEEE Transactions
- [4] Image annotation by combining multiple evidence & Word Net in Proc. ACM Int. Conf. Multimedia
- [5] Tagging over Time: Real-World Image Annotation by Lightweight Meta-Learning ACM International Conference Multimedia
- [6] Automated Image Annotation Using Global Features and Robust Nonparametric Density Estimation
- [7] R.Venkata Ramana Chary , Dr. D. Rajya Lakshmi, Dr. K.V.N. Sunitha Published an International Journal on “Image Retrieval and Similarity Measurement Based on Image Feature” in International Journal of Computer Science and Technology (IJCT) Volume 2 ,Issue 4 OCT-DEC 2011.PP385-389 ISSN :0976-8491(ONLINE)|ISSN:2229-4333(PRINT).
- [8] A Graph-Based Image Annotation Framework, Pattern Recognition Letters
- [9] A New Baseline for Image Annotation, Proc. European Conference Computer Vision
- [10] A Model for Learning the Semantics of Pictures, Advances in Neural Information Processing Systems
- [11] Refining Image Annotation Using Contextual Relations between Words, Proc. Sixth International Conference Image and Video Retrieval
- [12] D. Liu, X.-S. Hua , L. Yang, M.Wang, and H.-J. Zhang, “Tag ranking ,”in Proc. ACM Int. Conf. World Wide Web.
- [13] Image Retrieval: Ideas, Influences, and Trends of the New Age, ACM Computing Surveys
- [14] Resolving tag ambiguity, in Proceeding ACM International Conference Multimedia
- [15] R.Venkata Ramana Chary, Dr. K.V.N. Sunitha, Dr. D. Rajya Lakshmi Published an International journal “Quality Measures on Multispectral Images” in international journal of Image Processing and Applications(IJIPA) ,volume 2(1),2011,July- December 2011 .pp 101-105 ISSN:0975-8178
- [16] Supervised learning of semantic classes for image annotation and retrieval, IEEE Transactions on Pattern Analysis and Matching
- [17] Semi-supervised multi-label learning by constrained non-negative matrix factorization, in Proc. 21st National Conference Artificial Intelligence
- [18] G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos, “Supervised learning of semantic classes for image annotation and retrieval,”IEEE Trans. Pattern Anal. Mach.

Authors:

P. Nagarani is pursuing Masters in Software Engineering branch in the Department of Information Technology at Padmasri Dr. B.V.Raju Institute of Technology, (Affiliated to JNTU Hyderabad), Narsapur, Medak (Dist.), AP, India. Received Bachelor's Degree in Information Technology from Sri Padmavathi Mahila Visvavidyalayam (Women's University), Tirupati, Chittoor (Dist.).AP, India.



R.Venkata Ramana Chary is working as an Associate Professor in the Department of CSE at Padmasri Dr. B.V.Raju Institute of Technology, Narsapur, Medak, AP, India. Presently he is pursuing Ph.D from GITAM University, AP, India. His research interest is in Image Processing, Performance Measurements, Algorithms Analysis and Programming Techniques. He is having 15 years of teaching Experience.



Dr. D. Rajya Lakshmi is working as Professor and HOD in the Department of IT at GITAM University, Visakhapatnam, AP, India. Her research areas include Image processing and Data mining. She is having about 17 years of teaching and research experience. Her PhD was awarded from JNTUH, Hyderabad, AP in the area of Image Processing.

