TAG BASED IMAGE RETRIEVAL (TBIR) USING AUTOMATIC IMAGE ANNOTATION

Karthikram[1], Mailaivasan[2], Parthiban[3], Ganesh[4]

[1]Assistant Professor, [2, 3, 4]Student, IT Department, Sri Manakula Vinayagar Engineering College, Pondicherry, India

Abstract

In recent days, several social networking sites are more popular with digitized images. It comprises the major portion of the databases which makes the search engines to face difficulty in searching. We present a proficient image retrieval technique, which achieves eminent retrieval efficiency. Most of the images are annotated manually, thus the visual content and tags may be mismatched. This leads to poor performance in Tag Based Image Retrieval (TBIR). Automatic Image Annotation (AIA) analyzes the missing and noisy tags and over-refines it to increase the performance of TBIR. AIA can be achieved using the Tag Completion algorithm. The images retrieved from the TBIR are ranked based on the relevancy of the tags and visual content of the images. The relevancy can be evaluated using Content Based Image Retrieval (CBIR) technique. Based on the ranks, the images are indexed in the Tag matrix. Thus the images that match the search query can be retrieved in an optimal way.

Keywords: *Image Retrieval, Automatic Image Annotation, Tag Based Image Retrieval (TBIR), Tag Completion Algorithm, Content Based Image Retrieval (CBIR), Tag Matrix*

----------***----------

1. INTRODUCTION

The internet has witnessed a great success of social media websites. It increases the number of digital images in the websites. Nowadays people are more interested in searching the relevant images directly through search engines. The search engines use the image processing techniques for finding the images from the World Wide Web.

Image processing is any form of signal processing for which the input is an image, such as a photograph, the output may be either an image or a set of characteristics or parameters related to the image. The purpose of image processing is visualization, image sharpening and restoration, image retrieval, measurement of pattern and image recognition. Image processing is classified into analog and digital image processing. Analog image processing is conducted on two dimensional signals by means of analog input and output. For this type, the analyst must apply a combination of personal knowledge and collateral data to image processing.

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data.

Feature is an interesting part of an image, and features are used as a starting point for many computer vision algorithm. Feature detection is a low-level image processing operation. Feature extraction is a special form of dimensionality reduction and transforming the input data into the set of features.

Machine learning, a branch of artificial intelligence, concerns the construction and study of systems that can learn from data. Feature learning or representation learning is a set of techniques in machine learning that learn a transformation of "raw" inputs to a representation that can be effectively exploited in a supervised learning task such as classification.

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. Manual image annotation is time consuming, laborious and expensive; to address this, there has been a large amount of research done on automatic image annotation. Additionally, the increase in social web applications and the semantic web have inspired the development of several web-based image annotation tools.

Automatic image annotation (also known as automatic image tagging or linguistic indexing) is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image.

The image retrieval systems we use in our system are Tag Based Image Retrieval (TBIR) and Content Based Image Retrieval System (CBIR).

The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. This is a system that can filter images based on their content would provide better indexing and return more accurate results. Tag Based Image Retrieval retrieves images based on the tags/keywords present in the digital media that matches with the search query of the user.

The rest of the paper is organized as follows. Section 2 describes the various works that has been done previously by various scholars. Section 3 contains the architecture and working of the existing system. Section 4 describes briefly about the proposed framework that has to be implemented for higher retrieval efficiency. Section 5 concludes with future works to be done with the proposed system.

2. LITERATURE REVIEW

The surveys for image retrieval systems are done and some of the examples are given as follows. The techniques are grouped into the following: Annotation, Searching and Retrieval techniques. Under the first group of algorithms, firstly we are going to see the Latent Dirichlet Allocation [2] (LDA) that check the images for keywords and perceptual features, which uses Map Reduce technique by Liu et al. The works of [3] Xirong Li and Snoek shows us that the neighbor voting algorithm that efficiently learns tag relevance from the votes accumulated by visual neighbors. It gives the Tag ranking and image ranking for ordering the images. Another method of annotation is TagProp (i.e.) that is trained using a nearest neighbor model discriminatively [4]. The neighbor weights can be calculated by the distance or neighbors rank. Over the various image annotation objective- guided performance measures such as macro averaging which are sensitive to infrequent keywords & hamming measures is easily affected by skewed distribution. Qi Mao proposed a unified Multi-label learning framework [11]. That is a multi-layer hierarchical structure of learning hypotheses.

The second group of techniques covers the searching techniques like the system proposed by Najlae Idrissi in which the user can reach desired texture by navigating into a hierarchy of sub collections previously held. Co-occurrence matrices [9] are used to extract the intrinsic properties of the texture features. A framework [10] which has the components such as Keyword based search, User specific search , Collaborative filtering and ranking model given by Rani Borkar.

The final group among the three techniques is retrieval system which utilizes many algorithms to retrieve back the images from the net. In this group, the first technique we are going to see is Gray Level Co-occurrence Matrix [5] (GLCM) which uses shape features, dynamic dominant color and texture of an image by Babu Rao. Ying Liu and Xiaofang Zhou described the automatic texture segmentation algorithm [6] which is based on the features derived from wavelet domain that efficiently captures the textured region in arbitrary images. Xiang-Yang Wang used color, texture and shape information together for higher retrieval efficiency. Fast Color Quantization [7] algorithm obtains dominant colors and its percentages when clusters merging. Then the steerable filter decomposition is used to extract the spatial Textual features and pseudo-Zernike moments are used for shape descriptor. Mahmoud R. Hejari and Yo-Sung Ho applied nonlinear modified discrete radon transform [8]. Rotation-invariant texture feature, directionality and regularity are used for similarity assessment. Ning et al. proposed a framework based on domain dependent ontology [12] to perform semantic retrieval in image annotation. Ying Liu proposed a block-wise automatic texture segmentation algorithm [13] based on texture feature in wavelet domain.

Finally, our work is nearly related to Tag Completion [1], the goal is to automatically fill in the missing tags as well as correct noisy tags for given images and they represent an image-tag relation by tag-matrix, with both observed tags and visual similarity of image. Tag completion automatically complete Tag matrix with real number to indicate the probability of assigning tags to the images. Lei Wu et al. proposed a Tag Completion Algorithm where the input is observed Tag Matrix, parameters and convergence threshold and the output will be the complete Tag Matrix T.

3. TAG COMPLETION

The former system used the following architecture [1] for retrieving the images from the database. The architecture explains the working of the search engine. As per the architecture, the images and the manually added tags are present in the database of the website. The database consists of a set of images with previously assigned tags in it. When the user enters his search query in the search engine, the system finds the image related to query by looking into the tags/keywords present in the images. The Tag Completion Algorithm generates a tag matrix which has the relevancy of the images and previously added tags. The algorithm then fills the Tag matrix automatically by updating the relevance score of tags to all the images. Tag based image search or other search techniques can be then used to find the images from the tag matrix to display the results.

The problem in this system is that the automatic image annotation assigns tags to the images or corrects the tags after finding the images based on the tags already present in the image itself. Thus it increases the difficulty in finding the relevant images. When searching an image, at that time only the tags are added to the images retrieved from the search. To overcome this we proposed a new framework to assign tags before searching an image from a database.

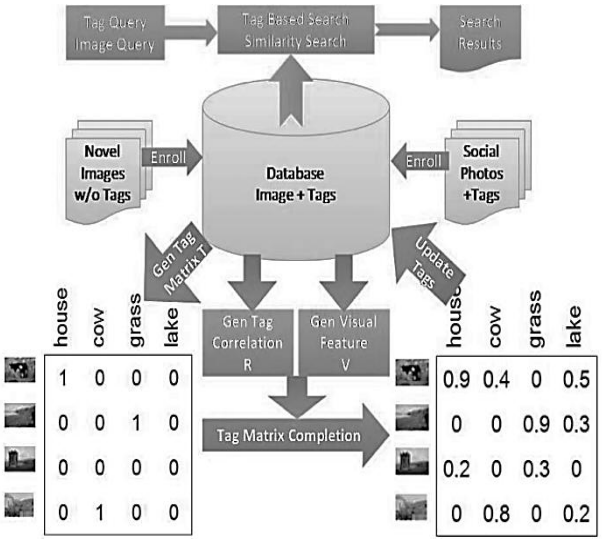

Fig -1: Framework for Tag Matrix Completion with the relevancy of the tags and its content [1].

Algorithm 1: Tag Completion Algorithm [1] (TMC)

1. INPUT:
   - Observed tag matrix: $\check{T} \epsilon IR^{nxm}$.
   - Parameters: γ, η, λ, and μ
   - Convergence threshold: ε
2. OUTPUT: the complete tag matrix T.
3. Complete the tag correlation matrix $R=\check{T}^T\check{T}$.
4. Initialize $w_1 = 1_d$, $T_1 = \check{T}$, and t = 0.
5. **repeat**
6. Set t = t + 1 and stepsize $\eta_t = 1/t$.
7. Compute $\check{T}_{t+1}$ and $\hat{w}_{t+1}$ according to (8).
8. Update the solutions T_{t+1} and w_{t+1} according to (9) and (10).
9. **until** convergence: $\| £(T_t,w_t) – £(T_{t+1},w_{t+1}) \| <= \varepsilon \| £(T_t,w_t) \|$

At each iteration t, given the current solution T_t and w_t, we first compute the sub gradients of the objective function £(T,w).

$$G=T_tT_t^T – V\ diag(w_t)V^T\ ,\ H=T_t^TT_t – R.$$

$$\blacktriangledown_T £(T_t,w_t) =2GT_t + 2\lambda T_tH + 2\eta(T_t-\check{T}) + \mu\Delta,$$

$$\blacktriangledown_w £(T_t,w_t) = -2diag(V^TGV) + \gamma\delta,$$

Where $\Delta \epsilon IR^{nxm}$ and $\delta \epsilon IR^d$ are defined as

$$\Delta_{i,j} = sgn(T_{i,j}),\ \delta_i = sgn(w_i)$$

$$Tt+1 = Tt - \eta_t \blacktriangledown_T £(T_t,w_t),$$

$$W_{t+1} = \prod_{\Omega}(w_t - \eta_t \blacktriangledown_w £(T_t,w_t)),$$

4. PROPOSED FRAMEWORK

With the ever-growing number of images on the Internet (such as in the online photo sharing Website, the online photo forum, and so on), retrieving relevant images from a large collection of database images has become an important research topic. Over the past decades, many image retrieval systems have been developed, such as Text-Based Image Retrieval (TBIR) and Content-Based Image Retrieval (CBIR).

By exploiting such rich semantic textual descriptions of Web images, the TBIR has been widely used in popular image search engines (e.g., Google, Bing, Flickr and Yahoo). Specifically, a user is required to input a keyword as a textual query to the retrieval system. Then the system returns the ranked relevant images whose surrounding texts contain the query keyword, and the ranking score is obtained according to some similarity measurements between the query keyword and the textual features of relevant images. However, the retrieval performance can be very poor, particularly when the textual features of the Web images are sparse and noisy in a high-dimensional space. To overcome the defects in the former system, we propose a new framework for the image retrieval system.

The techniques to be used in the proposed system are in the following order for an optimal retrieval of a search query:

1. Automatic Image Annotation.
2. Web Crawler
3. Tag Based Image Retrieval (TBIR).
4. Content Based Image Retrieval (CBIR).
5. Tag Matrix.

4.1 Automatic Image Annotation

Automatic image annotation (AIA) is the process by which a computer system automatically assigns metadata in the form of adding captions or keywords or tags to a digital image. Content based ranking of images is harder than for textual documents because they do not have words in that part. Image search techniques are working mainly based on using annotations and semantic tags that are present in the images. However, tags are entered by the users manually which consumes a large amount of time for the tons and tons of images present in the database. Thus, AIA has been a most challenging task in the past decades.

AIA methods require a set of training images, from which annotations for the images are determined. The process involved in the AIA is as follows. The training dataset has been previously loaded to the system in order to process the given input image. This technique uses the local and global features for estimating the presence of the training dataset in the given image. It splits the given image into various combination of images based on scaling by Red, Blue, Green and various colour histograms. Then the images are compared

with the dataset, the unwanted features present in the images are left unnoticed and if there is the presence of similar features are added tags and retrieved back with tags annotated automatically.

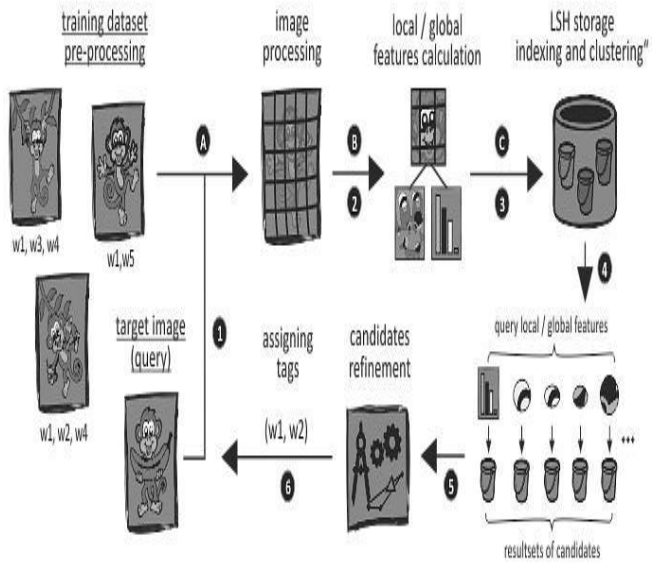

Fig -2: Automatic Image Annotation technique with a sample dataset of images.

4.2 Web Crawler

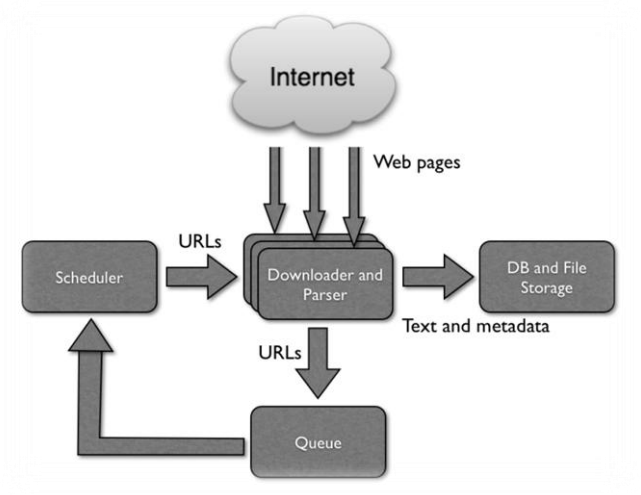

Fig -3: Architecture of a Web crawler

WebCrawler is a software system used for locating publicly available webpages. The crawler follows one link which is found from another link and sends back the data about the webpages to the webserver. It grabs the web addresses from the previously grabbed webpages. Web crawler is the central part of the search engines. The working of the WebCrawler is explained with the help of WebCrawler architecture.

The crawler downloads the data from the webpage and then grabs the link if any and adds it to the queue. The pages are scheduled for downloading data. After the data is downloaded, it is indexed with its link in the database as text/metadata. If the search word is found in the index then the result is retrieved to the user as text or digital media present in that page.

4.3 Tag Based Image Retrieval (TBIR)

Tag Based Image Retrieval (TBIR) is the application of computer vision techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases.

The name Tag Based means, it looking for the tag that the user entered as a search query in the browser of any system in the world. It looks the similar tag that has been attached with the image and retrieves the image to the user. It didn't check the content of the image; it only checks the tag in the image. TBIR is the most efficient technique in image retrieval but it is dependent the tags. The tags are added manually by the users during the time of uploading. To overcome that, automatic annotation technique is applied to add tags automatically to the images during the upload of an image in any database. The tags are the text or annotation present in the image.

4.4 Content Based Image Retrieval (CBIR)

Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases. Content-based image retrieval is opposed to concept-based approaches.

"Content-based" means that the search analyses the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colours, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because most web-based image search engines rely purely on metadata and this produces a lot of garbage in the results. Also having humans manually enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Thus a system that can filter images based on their content would provide better indexing and return more accurate results.

4.5 Tag Matrix

The Tag matrix is a multidimensional array which has the ranking for the content of the images and it's relevancy of tags present in it. It arranges the images based on the integer values which have a maximum value. It stores a large amount of

image and its value for a temporary time for that particular keyword given in the query.

5. CONCLUSIONS

In this paper, we present a new image search engine framework which consists of Automatic Image Annotation (AIA), Web Crawler, Tag Based Image Retrieval (TBIR), Content Based Image Retrieval (CBIR), and Tag matrix techniques. The AIA plays an important role in the proposed framework. Since the annotation of the tags is done using Tag Completion Algorithm and the tags are annotated automatically. Thus it minimizes the work of the user to add the tags manually by himself. TBIR thus will give an increase in efficiency in retrieving the image from the database than the former system in which the tags are annotated manually. Tag Matrix serves the purpose of indexing the images.

REFERENCES

[1]. Tag Completion for Image Retrieval, Lei Wu, Member, IEEE, Rong Jin, Anil K. Jain, Fellow, IEEE, March 2013.
[2]. Web-scale Image Annotation, Jiakai Liu, Rong Hu, Meihong Wang. Yi Wang, and Edward Y. Chang.
[3]. Learning Social Tag Relevance by Neighbor Voting, Xirong Li, Student Member, IEEE, Cees G. M. Snoek, Member, IEEE, Marcel Worring, Member, IEEE, November 2009.
[4]. TagProp: Discriminative Metric Learning in Nearest Neighbor Models for Image Auto-Annotation, Matthieu Guillaumin, Thomas Mensink, Jakob Verbeek and Cordelia Schmid.
[5]. Content Based Image Retrieval Using Dominant Color, Texture and Shape, M.BabuRao, Dr.B.PrabhakaraRao, Dr.A.Govardhan.
[6]. Automatic Texture Segmentation for Texture-based Image Retrieval, Ying Liu, Xiaofang Zhou.
[7]. An effective image retrieval scheme using color, texture and shape features, Xiang-Yang Wang, Yong- Jian Yu, Hong-Ying Yangs.
[8]. An Efficient Approach to Texture-Based Image Retrieval, Mahmoud R. Hejazi, Yo- Sung Ho, 1 Oryong-dong, Buk-gu, Gwangju.
[9]. Bridging the Semantic Gap for Texture-based Image Retrieval and Navigation, NajlaeIdrissi, Jos´eMartinez, DrissAboutajdine, Rabat – Agdal.
[10]. Learn To Personalized Image Search from Photo Sharing Websites, Rani Borkar, MehjbeenTamboli, and PoojaWalunj.
[11]. Objective-Guided Image Annotation, Qi Mao, Ivor Wai-Hung Tsang, and ShenghuaGao.
[12]. Semantic Based Image Retrieval in Remote Sensing Archive: An Ontology Approach, NingRuan, Ning Huang, Wen Hong.
[13]. Texture Segmentation Based on Features in Wavelet Domain for Image Retrieval, Ying Liu, Si Wu, Xiaofang Zhou.

BIOGRAPHIES

A. Karthikram is working as an Assistant Professor of department of Information Technology in Sri Manakula Vinayagar Engineering College.

A. Mailaivasan pursuing final year in the degree of Bachelor of Technology in the stream of Information Technology from Sri Manakula Vinayagar Engineering College of Pondicherry University [2010-2014].

V. Parthiban pursuing final year in the degree of Bachelor of Technology in the stream of Information Technology from Sri Manakula Vinayagar Engineering College of Pondicherry University [2010-2014].

D. Ganesh pursuing final year in the degree of Bachelor of Technology in the stream of Information Technology from Sri Manakula Vinayagar Engineering College of Pondicherry University [2010-2014].