A STUDY ON ANATOMY OF ONTOLOGY BASED IMAGE RETRIEVAL

Mrs.P.Sumathi, Assistant Professor,School of IT & Science,
Dr.G.R.D.College of Science, Coimbatore.
Dr.Manickachezian, Associate Professor, Dr.Mahalingam Centre for Research and Development(CS), NGM College,Pollachi,

Abstract: **The objective of this work is to generate a database containing a large number of images and with high precision is still a difficult task. Image search engine apparently provide an effortless route but currently are limited by poor precision of images and limited on total number of images provided. Web mining deals with the retrieval of useful patterns based on the user's search requirements. Image retrieval is a promising approach because of its automatic indexing and retrieval based on their visual appearance. The similarity of images depends on the feature representation and dissimilarity function. The study approach is not used only to enhance the retrieval accuracy of web images but also able to annotated the unlabeled images.**

Key words: **Web Image Retrieval, Ontology, Hash, Crawler, Wordnet, Analysis.**

I. INTRODUCTION

As the demand for image retrieval and browsing online is growing dramatically, there are hundreds of millions of images available on the current World Wide Web. For multimedia documents, the typical keyword-based retrieval methods assume that the user has an exact goal in mind in searching a set of images whereas users normally do not know what they want, or the user faces a repository of images whose domain is less known and content is semantically complicated. In these cases it is difficult to decide what keywords to use for the query. To catch user's ontological intention, a new approach has been proposed as Intelligent Web Image Fetch using hash code generation algorithm, which simultaneously considers the ontological requirements in usability, intelligence and effectiveness [1]. Based on the proposed visual and textual-based image retrieval models, the response to the image query is very much easy and effective. Through empirical evaluations, our proposed model has delivered accurate results for semantic web image retrieval.

II. IMAGE RETRIEVAL

The process of retrieving and displaying relevant images is based on user's queries from the web or image database. Generally image retrieval methods are classified into two types.

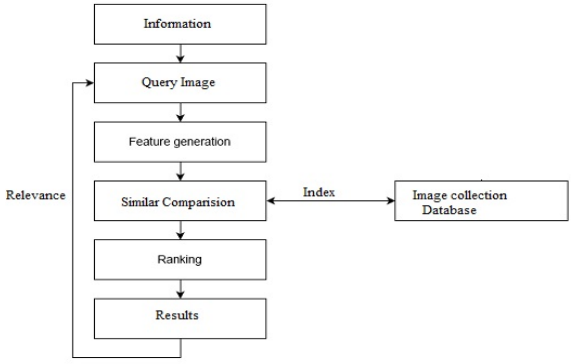
- Text based image retrieval (TBIR)
- Content based image retrieval (CBIR)

1) Text based image retrieval (TBIR)
The image is annotated by using text descriptions like creator, place, date, time, objects. The image retrieval is done by using one or more textual descriptors. The high-level retrieval involves retrieval of an image based on the name of objects, emotions and actions that can be associated with the image [2].

2) Content based image retrieval (CBIR)
Extraction of images based on image content. CBIR uses low level features such as color, texture, shape and object location that involves high level reasoning.

III. ONTOLOGY BASED IMAGE RETRIEVAL

The ontology based image retrieval uses knowledge representations which combine the features of text based image retrieval and content-based image retrieval. Ontology provides the way to organize the web information in structured way. The web contents can be understood by the computer as well as the user. There are hundreds of millions of images available on the current World Wide Web. Content-based retrieval is an efficient technique that considers low-level features of image data, the role of using some textual keywords to describe the content of an image to support the retrieval system [3]. The main purpose on ontology is to represent the image in semantic manner.

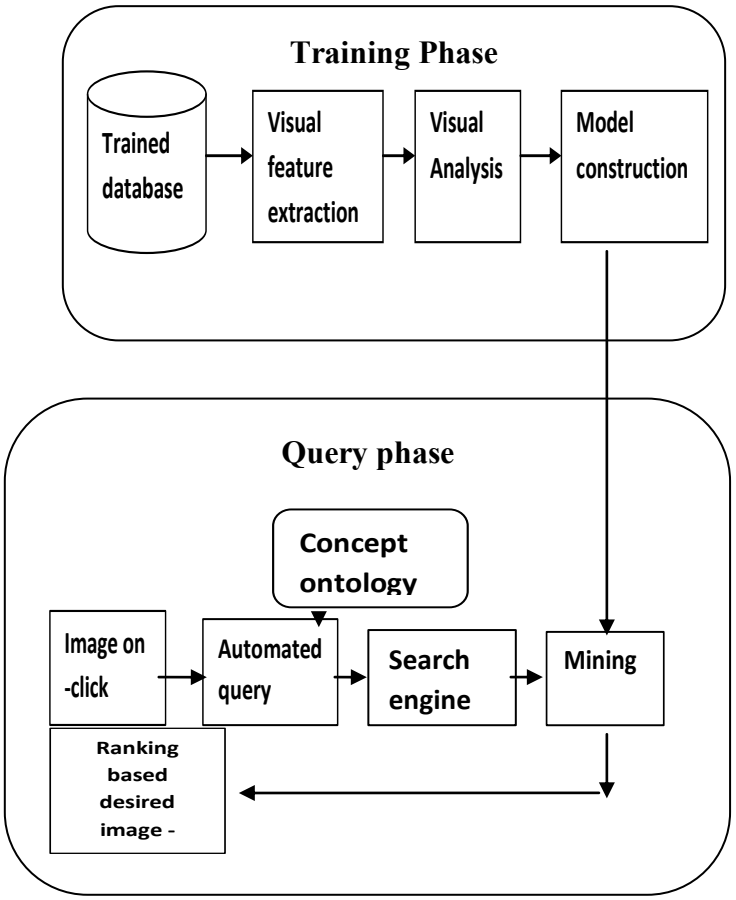


ISSN: 2278 – 7798

In the training phase the images are trained and stored in the trained database and the model is being constructed. In the query phase the input query is given in the form of query, for which the automated query generator with the help of ontology generates the query. Then the image is searched from the constructed model and the most relevant images to the input query are ranked and retrieved from the database. Hence the image is represented in machine understandable manner, this leads to retrieval task easier.

IV. MINING MODULE

In this module, a few terms or phrases from the textual descriptions of the images obtained by search are mined. Inverted indexing technique is used which contains an inverted indexing file with entries as keywords and documents as values. The documents are those which contain that specific keyword. Cluster documents by ranking salient phrases and documents containing a certain salient phrase form a cluster. Then the system projects the predicted keywords onto the ontology by computing the concept similarities [4]. Afterward, the candidate concepts can be categorized into two types, namely high-level and low-level concepts. For the user, the specific concept can be determined to serve semantic image retrieval.



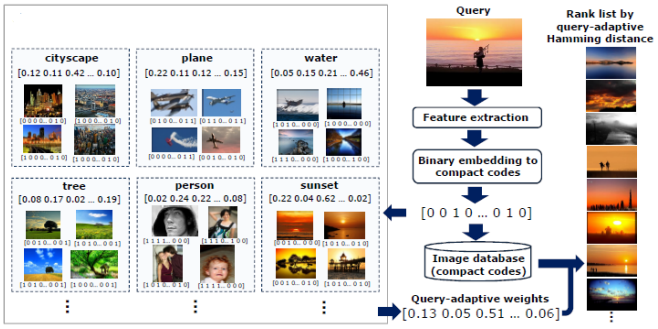
V. IMAGE RETRIEVAL PROTOTYPE

Web image retrieval prototype system consists of different components that are user interface, data collection, indexing component and image retrieval component. User interface is mainly an image browser which provides a query interface and user feedback interface. Data collection component have different modules crawler, feature extractor and image indexer. Image retrieval component have log miner, model updater and query updater. The data organization of the system mainly consists of image database that also contains metadata of images low-level and high-level features of user space model [4].

Web pages containing images and store them into a local database. These pages have to extract both the low level visual features and the high level semantic features of the images. Whenever crawler finds the Web page containing some images, the feature extractor is applied immediately on the images and the page to obtain the features and store the images and their features in the image database. An image interface is also provided for users to specify whether a returned image is relevant or not to the user's intents.

VI. HASHING TECHNIQUE IN IMAGE RETRIEVAL

A new hashing method has been proposed to solve the problem of finding nearest neighbors in gigantic datasets. Such techniques are needed in many important applications such as content-based retrieval and matching of images matching of visual features in high dimensional spaces and other applications involving millions or billions of samples. In several solutions, the best way was to find the optimal projections for generating the binary hash bits [5]. In others we exploit the strategies like semi supervised learning graph based manifold representation, query-dependent adaptation, or joint speed accuracy optimization to significantly improve the hashing performance.



Semi-Supervised Hashing

Semi-supervised hashing method that minimizes empirical error on the labeled data while maximizing variance and independence of hash bits over the labeled and unlabeled data.

Sequential Hashing

Data dependent projection learning method similar to the concept of boosting such that each hashing function is designed to correct the errors made by the previous one sequentially[6].

Kernel Hashing

Efficient codes for large scale data of general formats with any kernel function, including kernels on vectors, graphs, sequences, sets.

VII.HASH ALGORITHM

A hash function in the unrestricted sense is a function which has usually meant a map h: D → {0, 1} n that is compressing, meaning |D | > 2n.
A collision for h: D → {0, 1} n is a pair x1, x2 ∈ D of points such that h(x1) = h(x2) but x1 ≠ x2
We consider a family H: {0, 1} k× D → {0, 1} n of functions, meaning for each K we have a map
h = HK:
D → {0, 1}n defined by
h(x) = H (K , x)
X(=X0,X1,X2,……..X15) denotes a collection of images.

X = Collision
A = Probability Adv (H)
D = Family of function
1.H(K) = K mod M
2.Key V = Itemkey(X)
3.Int i = hash(V,M)
4.i ← shaped (M)
5.Int skip = hash(V,K)
//generate Adversary collision
6. V ← shift (K)
7.A1(X) ← D
8. H = (A * H + V) / M
9.H(K(D)) = A(X)
10.While (st[i] ! = Null)
11.i = (i+skip) / M
12.st[i] = x
13. return st[i]
14.end

The algorithm works in two phase like training phase and testing phase.

In the training phase some set of the images are trained by extracting the features like color, texture, and shape features to build the main object classifier and background object models.

In the testing phase the entire work has been separated into several modules like searching, mining and ranking.

VIII. HASH - IMAGE MINING

Nearest neighbor (NN) search has been widely used in machine learning and related application areas, such as information retrieval, data mining, and computer vision. Recently, with the explosive growth of data on the Internet, there has been increasing interest in NN search in massive (large-scale) data sets. Traditional brute force NN search requires scanning all the points in a data set whose time complexity is linear to the sample size. Hence, it is computationally prohibitive to adopt brute force NN search for massive data sets which might contain millions or even billions of points. Another challenge faced by NN search in massive data sets is the excessive storage cost which is typically unacceptable if traditional data formats are used [7]. By using hashing codes, we can achieve constant or sub-linear search time complexity. Moreover, the storage needed to store the binary codes will be dramatically reduced.

XI. IMAGE RANKING

A simple ranking and annotation rejection approach to select and rank the mined cluster names and use the top ones as the output annotations. Using the textual features the goal is to re-rank the retrieve images [8]. Each feature is treated as “True” or “False”. The features define a binary feature vector for each image and ranking is based on posterior probability of the image. To re-rank images for one particular class we employ the ground truth data for that class we can re-rank the images. The results show that the proposed method has high effectiveness for image retrieval.

X.CONCLUSION

Searching an image in the web is an important technique which will be useful to us in various applications. The entire process is formulated in a divide and conquer framework where a query image is provided along with the uncaptioned image to improve both the effectiveness and efficiency [9]. A novel attempt of image retrieval not only evaluates how much data can help us in image understanding, but also illustrates that search, as an important technique itself which can help us to retrieve correct relevant images of our study. This process will be useful when a particular image of interest has to retrieve from the huge database [10]. In future we would like to implement the mining which uses the textual descriptions of the image in order to cluster documents using the salient terms or phrases in the image.

REFERENCES

[1]K. Barnard, P. Duygulu, N.de Freitas, D. Blei and M. Jordan, “Matching Words and Pictures”, J. Machine Learning Research.
[2]F. Schroff, A, Criminisi and A.Zisserman, “Harvesting Image Databases from the Web”, http://www.robots.ox.ac.uk/~vgg/data/mkdb.
[3] Chakrabarti S. “Mining the web: Analysis of Hypertext and Semi Structured Data”, morgan Kaufmann, San Francisco, CA.
[4]R. He , N. Xiong, L.T. Yang, and J. H. Park, “Using Multi-Modal Semantic Association Rules to Fuse Keywords and Visual Features Automatically for Web Image Retrieval”, Information. Fusion, 2010.
[5]R. He, N. Xiong, T. H. Kim and Y. Zhu,” Mining Cross-Modal Association Rules for Web Image Retrieval”, International Symposium on Computer Science and its Applications, 2011.
[6]V. Mezaris, I. Kompatsiaris, and M. G. Strintzis, “An Ontology Approach to Object-Based Image Retrieval”, in Proc. IEEE ICIP, 2010.
[7]M.L. Kherfi, D. Ziou, A. Bernardi, “Image Retrieval from the World Wide Web: Issues, Techniques, and Systems”, ACM Computing Surveys Vol.36, No. 1, 2008.
[8]Web Ontology Language (OWL), http://www.w3.org/2004/OWL/.
[9]W.H. Lin, R. Jin and A. Hauptmann, “Web Image Retrieval Re-Ranking with Relevance Model”, Proc. IA-DIS Int’l Conf.,

ISSN: 2278 – 7798

[10]CHEN, H. M. AND COOPER, M. D. "Using Clustering Techniques to Detect Usage Patterns in a Web Based Information System," Journal of the American Society for Information Science and Technology
[11]DAVISON, B.D. 2004. Course Website of CSE345/445 http://www.cse.lehigh.edu/~brian/course/searchengines.

Mrs.P.Sumathi M.Sc.,M.Phil., working as Assistant Professor with eight years of experience in teaching, She has published five papers in national level conference and two papers in journals. She has participated in seminars/ conference/workshops/refresher courses at national Level. Now she is pursuing Ph.D Computer Science in Dr. Mahalingam center for research and development at NGM College Pollachi.

** **Dr. R.Manicka chezian** received his M.Sc., degree in Applied Science from P.S.G College of Technology, Coimbatore, India in 1987. He completed his M.S. degree in Software Systems from Birla Institute Technology and Science, Pilani, Rajasthan, India and Ph.D degree in Computer Science from School Computer Science and Engineering, Bharathiar University, Coimbatore, India. He served as a Faculty Maths and Computer Applications at P.S.G College of Technology, Coimbatore from 1987 to 1989.Presently, he has been working as an Associate Professor of Computer Science in N G M College (Autonomous), Pollachi under Bharathiar University, Coimbatore, India since 1989. He has published seventy five papers in international/national journal and conferences: He is a recipient of many awards like Desha Mithra Award and Best Paper Award. His research focuses on Network Databases, Data Mining, Distributed Computing, Data Compression, Mobile Computing, Real Time Systems and Bio-Informatics.