Visual and Text Hashing Method with Semantic Association with Multi Intention Feature Mining Algorithms

G. Maheshwari
Assistant Professor, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India.

Prajitha.P
M.Phil Scholar, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India.

Abstract – Content retrieval from text and image search is one of the most fundamental problems in the search engines from huge sized database and machine learning research communities. The design and evaluation of an image retrieval system rely on properly defined visual features with suitable similarity matching metrics as well as correct normalization functions. In proposed system a novel unsupervised visual hashing scheme, termed as semantic-assisted visual and Text hashing (SAVTH), to effectively perform visual hashing learning with semantic assistance. The main concept of the proposed work is to extract images from the website automatically from the noisy associated texts, and stores them in the repository for future image and text based searches. Current research objective is to efficiently retrieve content based images and texts from large databases using text and image inputs.

Index Terms – Content Based Image Retrieval, Visual Hashing, Feature Mining, Text Hashing.

1. INTRODUCTION

Query by Image Content is a method which uses visual data's to search images from large scale image databases according to users' interests, has been an active and fast advancing research area since several years ago. At some stage in the past decade, remarkable progress has been made in both theoretical research and system development. But, there remain many challenging research problems that continue to attract researchers from multiple disciplines. Query by Image Content [1] uses the contents of images to represent and access the images. A typical Query by Image Content system is divided into off-line feature extraction and online image retrieval. As a result of recent advancements in digital storage technology, it is now possible to create large and extensive databases of digital images and textual contents. These collections may contain millions of images and terabytes of textual data. For users to make the most of these databases effective, efficient methods of searching must be devised. Prior to automated indexing methods, image databases were indexed according to keywords that were both decided upon and entered by a human categorizer. Unfortunately, this practice comes with two very severe shortcomings. First, as a database becomes increasingly large the manpower required to index each image becomes less practical. Secondly, two different people, or even the same person on two different days, may index similar images inconsistently. The result of these inefficiencies is a less than optimal search result for the end user of the system.

Having a computer do the indexing [2] based on a QBIC scheme attempts to address the shortcomings of human-based indexing. Since a computer can process images at a much higher rate, while never tiring, the manpower issue is solved. Additionally, as long as the algorithms [3] [4] used in the indexing procedure are kept constant, all images will be indexed consistently, solving the inherent problems resulting from fallible human-based indexing. It should also be noted that current QBIC methodologies are not without their limitations. The main objective of the current research is to improve the indexing and retrieval performance of the system from large database of image, and to compare the low level visual features between a queried image and database of images, aim to learn and fuse the extracted features from the image and textual information's using Multi-Intention Feature Mining (MIFM). To achieve all these proposed system developed a Semantic Assisted Visual Text Hashing technique.

2. PROBLEM STATEMENT

Similarity search and visual indexing [5] is one of the most fundamental problems in information retrieval, database and machine learning research communities. It is defined as the task of finding close samples for a given query to retrieve images and texts. It is of great importance to many applications, such as content-based multimedia retrieval, classification and annotation of search results in the search engines. Many hashing algorithms have been developed in recent years and the hashing methods can mainly be divided into two categories: unsupervised method. Such hashing-based methods [6] for fast similarity search can be considered as a means for embedding high dimensional feature vectors to a low dimensional

Hamming space, while retaining as much as possible the semantic similarity structure of data. Although these hashing methods have shown success in large-scale image search, there is a problem that is seldom exploited. The existing hashing approaches are defined only for the standard Euclidean distance. Several drawbacks are observed in the earlier [7]. They are as follows: The distance metric learning and hashing can be regarded as two isolated steps in this approach and the objective optimized in distance metric learning may not be appropriate for hashing For many visual hashing algorithms that rely on similarity rather than distance, such as SH, it needs a further step to convert distance to similarity and the involved radius parameter is usually sensitive. Many indexing and visual hashing algorithms are only considers the image search than text.

3. PROPOSED SYSTEM

3.1 Introduction

Content retrieval from text and image search is one of the most fundamental problems in the search engines from huge sized database and machine learning research communities. The design and evaluation of an image retrieval system rely on properly defined visual features with suitable similarity matching metrics as well as correct normalization functions.

In proposed system a novel unsupervised visual hashing scheme, termed as semantic-assisted visual and Text hashing (SAVTH), to effectively perform visual hashing learning with semantic assistance. The main concept of the proposed work is to extract images from the website automatically from the noisy associated texts, and stores them in the repository for future image and text based searches. This helps to enhance the discriminative capability of hash codes, and thus facilitate the performance improvement of visual hashing. SAVTH works as follows:

Initially hash code learning is formulated in a unified un-supervised framework, where relaxed hash codes are learned by simultaneously preserving visual and textual similarity of images and considering the assistance of texts and its description. More specifically, the framework integrates two important assistance of auxiliary texts to effectively mitigate the inherent limitations of visual features by deploying the text mining techniques. The first assistance models high-order semantic relations of images by constructing topic hyper-graph, while the second one correlates images and latent shared topics detected via collective matrix factorization. Then, an optimization method based on MIFM is proposed to iteratively calculate the optimal solution. We specially pre-serve bits-uncorrelated constraint during iterative process to facilitate learning and simultaneously reduce information redundancy between hash bits. Finally, hash functions are constructed based on linear regression to enable out-of-sample query extension. Linear projection can support efficient hash code generation in online retrieval.

In the proposed system an effective visual indexing and information retrieval methods were studied the process of retrieving desired images from a large collection on the basis of features that can be automatically extracted from the images themselves. These systems called as Content Based Image Retrieval or Query By Image Content have received intensive attention in the literature of image information retrieval and consequently a broad range of techniques has been proposed. CBIR or QBIC performs three main tasks such as Feature extraction, Feature selection and Retrieval .The extraction task transforms rich content of images into various content features. Feature extraction is the process of generating features to be used in the selection and classification tasks. Feature selection reduces the number of features provided to the classification task. Those features which are likely to assist in discrimination are selected and used in the classification task. Features which are not selected are discarded of these three activities, feature extraction is most critical because the particular features made available for discrimination directly influence the efficacy of the classification task. The final process of the mining task is a set of features, commonly called a feature vector, which constitutes a representation of the image. The key feature of the proposed system is being summarized as follows:

Instead of considering only visual feature or equally treating images and texts, SAVTH specially exploits the auxiliary texts to assist visual hashing. Two important assistances from auxiliary texts: modeling semantic correlations of images with topic hyper-graph, correlating images and latent shared topics via collective matrix factorization, are proposed to effectively incorporate semantics into the hash codes. SAVTH is designed in a unified unsupervised mining framework, which comprehensively considers visual and text similarity preservation of images and semantic-assistance between the images and texts. An effective solution based on MIFM is proposed to calculate the optimal features and its hash codes.

Current research objective is to efficiently retrieve content based images and texts from large databases using text and image inputs. The main contribution of the present research is as follows:

1. Multi features from the images with visual indexing along with Low level features and user provided tags has been extracted and optimized and fused by using Multi-Intention Feature Mining (MIFM)

2. These features can be learned in the form of visual graph by constructing the SAVTH.

3. To bridge the gap between semantic and content, reinforcement style integrating with feature weighting learning and n-gram method is proposed. This method estimates the

most important information of features as a weight from the obtained features for the image and documents.

4. Then these features can be indexed by using SAVTH This indexing constructs chain for relevance estimation by considering the semantic dependence of user queries. This can be done prior to spectral hashing.

5. Finally, to improve the retrieval performance of images in a large database, using visual weighted compressed sensing method is proposed in which the some information of images are abstracted and reconstructed while retrieving the image. Thus it reduces the retrieval time.

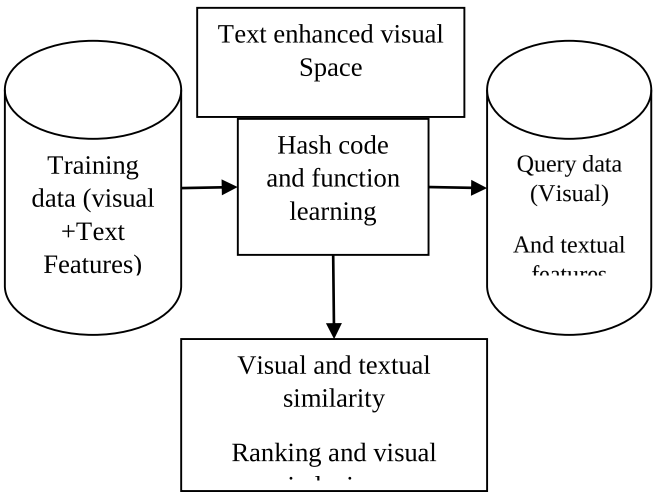

Figure 1.0SAVTH framework

Figure1.0 describes the basic framework of the SAVTH -based image and text retrieval system. The system mainly includes two core components: offline learning and hashing and online feature matching process.

Offline Learning and hashing. This phase aims to learn hash codes of database images and simultaneously generate hash function for query image. It consists of four main steps. First, visual and text features of images are extracted to transform image pixels to mathematical vector representations. Then, a text-empowered visual graph is constructed with the assistance of topic hyper-graph, and latent semantic topics are detected under guidance of text information's. then, hash codes of database images are mined in the SAVTH framework which preserves relationships and similarities of images and that between images and semantic topics. At last, visual and text hash functions are generated with respect to the hash codes within the MIFM model.

Online feature matching. Visual feature of query image and texts are extracted in the user search module. Then, it is located into binary codes with visual hash functions. Finally, the matching between query image and set of images from database are calculated and database images are returned in order of ranking ascending.

3.2 User Query Image :

The user query image and text are extracted for the content retrieval. The given query is mined and evaluated for content retrieval. For a given query image, let use D S, where S is the complete image set. To improve the discriminative power of the least-ranked visual features, two kinds of query samples are considered D+ and D-. D+ holds the most relevant samples for the corresponding keyword, referred to as positive group; I− represents negative group in which the samples are irrelevant to the keyword of concern but may look similar to the positive query samples; D = D+ D−. the given query samples are matched with the visual and text hash codes and finds the neighbor images.

3.3 Feature Extraction:

The proposed system performs effective feature selection technique to find appropriate images and documents. The feature is described as a function of one or more measurements, each of which specifies some quantifiable property of an object and is computed such that it quantifies some significant characteristics of the object. The various features described and employed are defined as follows:

General features: Application independent features such as color, texture and shape. According to the abstraction level, they can be further divided into:

Pixel-level features: Features calculated at each pixel, eg. color, location.

Text features: Features calculated over the results of the associated text and documents.

Global features: Features calculated over the entire image or just regular sub-area of an image

Category-specific features: Application dependent features such as human faces, fingerprints and conceptual features. These features are often producing least-ranked features for a specific category. On the other hand, all features can be coarsely classified into least-ranked features and high-level features. Least-ranked features can be extracted directed from the original images, whereas high-level feature extraction must be based on low level features. The primary goal of current research is to develop an approach for automatically combining least-ranked visual features.

3.4 Proposed Multi Intention Feature Mining:

The goal of Multi Intention Feature Mining (MIFM) method is to define a suitable multi-feature model for the visual representation of a specific Image dataset. The core of this method is a learning process toward an optimal combination model by assigning each involved low level feature space F_j a proper weight α_j. This can be achieved by optimizing an objective function or a set of objective functions for variable α. Since several representative samples are used for a good visual representation of a keyword, the interest of each single query sample may conflict with others. Thus, an objective function for each query sample in R is constructed and a multi-objective optimization strategy is used to find the solution that can achieve a common optimum for all these functions. Based on the distance matrix DM given in a set of objective functions can be constructed for the optimization of a multi feature model. Each objective function is formed as weighted linear combinations of feature-specific distances. Considering DM total number of m objective functions can be constructed.

3.5 Visual Hashing:

Visual Hashing (VH) is proposed to design compact binary codes for approximating nearest neighbor search based on Visual graph mining. Besides the common property of maintaining sample similarity in the reduced hamming space, VH requires the codes to be balanced and uncorrelated. The VH codes satisfy the following process:

First perform weighted MIFM algorithm to compute image features and performs the visual hash extraction process with the Semantic based similarities and second perform ranked feature mining considering the semantic based similarity to update the feature weights and then update the node similarities based on the new content similarity.

Algorithm: SAVTH Algorithm

Input: Input graph G

1. Initializes the process
2. Iterate every step

{

Compute semantic similarity for all image pairs

Compute semantic similarity for all text pairs

Compute semantic similarity for all group pairs;

Compute semantic similarity for all tag pairs;

} until converge or stop criteria satisfied

3. Perform feature learning to update W Wm* 1

4. Update image similarities

Output: S, similarity scores of query image and database images.

Algorithm: MIFM Integration of Semantic and Concept Similarity

Input: G, Graph

1. Construct the visual hash Graph
2. Find the similarity of each object using SAVTH algorithm
3. Initialize similarity score
4. Iterate{
5. Calculate the semantic similarity for image pairs via enhanced SAVTH algorithm
6. Perform feature learning to update W Wm* 1
7. Search for new top k similar image objects based on the new similarity weighting
8. Update the new image similarities
9. Compute semantic based similarity for all group and tag pairs via weighted semantic feature mining algorithm
10. } until converge or stop criteria satisfied

Output: S, similarity scores of each dataset

An image is represented as a point in a D-dimension feature space with either a single type of feature or a combination of multiple types of features. Image similarity can be estimated from image content features from visual hash indexes. To build a bridge between the content and semantics similarity of text and images, the integration mechanism with the MIFM is used.

4. IMPLEMENTATION AND RESULTS

Dataset: The implementation has performed with 10 image dataset from the different image repository. The dataset is namely Birds (100 training images), roses (81 training images), Flower (106 training images), Face (133 training images), Fruit (89 training images), House (100 training images), Lake (100 training images), Mountain (110 training images), Plane and Sunset (130 training images). This finally has chosen 300 images randomly from each category for training purpose with its textual descriptions. The images are real world, and with high intra-class variability. Using the different set of databases, it is possible to evaluate retrieval results. The datasets are crawled from the Google website.

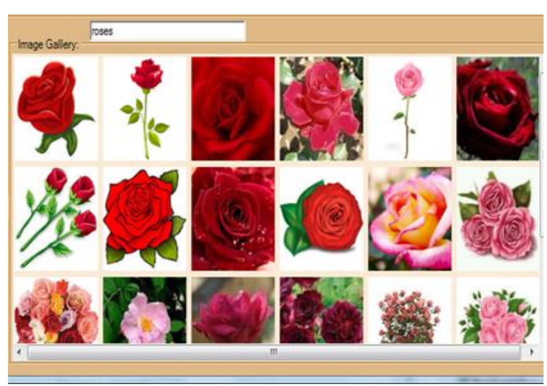

Figure 2.0 sample dataset under the flower category

The figure 2.0 shows the sample dataset crawled from the url https://www.google.com/search?q=flowers.

Performance analysis:

The results presented in this section were obtained from experiments using image data set of approx 800 color images crawled from the Google images. These images are primarily divided into 10 data sets like birds, Sunset, Plane, Mountain, Rose, House, Fruit, Flower, Face, Car. For finding the best combination of image's features of the current image data set. This turns off the Region feature and 8 combinations of Global features is performed.

Performance Evaluation Metrics of Image Retrieval

Precision- P is defined as the ratio of the number of retrieved relevant images to the total number of retrieved images.

$$P = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}$$

Recall - R is defined as the ratio of the number of retrieved relevant images to the total number of relevant images in the whole database.

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in the database}}$$

Table 1.0 precision and recall for 2 types of dataset for the proposed system

	Dataset1		Dataset 2	
No.of Images	Precision	Recall	Precision	Recall
10	13.31	4.35	14.23	4.40
20	13.24	7.92	14.22	8.60
30	13.47	11.68	14.75	12.82
40	13.54	14.97	15.00	16.88
50	13.83	18.57	14.97	20.40
60	14.15	22.32	14.91	23.95
70	14.28	25.80	14.86	27.30
80	14.43	29.57	14.74	30.39
90	14.32	32.60	14.89	34.08
100	14.34	36.10	14.90	37.55

The 1.0 shows the various types of image dataset and its precision and recall. This is increased when the dataset is increases. This has been conducted for 5 types of dataset such as flower, animal, sunset, nature and sea. The overall accuracy, precision and recall is compared with the existing system and given in table 2.0.

Table 2.0 performance comparison table

Techniques	**Dataset counts**	**Accuracy**
SAVH	20	83.146
	30	90.667
	40	83.404
	60	85.514
SAVTH	20	85.955
	30	91.333
	40	85.319
	60	87.85

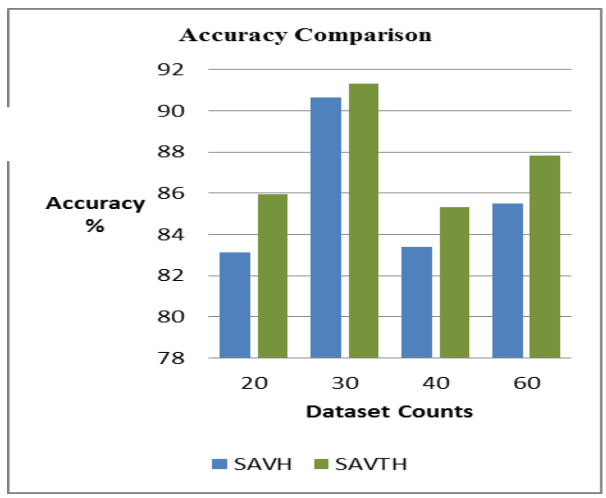

Figure 3.0 accuracy comparisons between existing and proposed system

5. CONCLUSION

In the current paper, an efficient Content based image and text retrieval system is proposed based on semantic assisted visual and text hashing with hyper graph approach. Unlike existing visual hashing technique, which uses only to build the visual graph for images, current research aims to optimize the semantic assisted visual and text hashing technique with multi intention feature mining (MIFM) based on the pair-wise similarities of image labels or tags of features from the MIFM method. MIFM method is employed to fuses the extracted features color descriptors. Hyper-graph has been learnt and the gap of semantic and concept information has been linked based on the SAVTH algorithm. The proposed algorithm computes the similarity or relevance in weighted information of features from the available features of Images. In order to mine the images in the database appropriately based on given query, indexing has been done by using ranked feature indexing (RFI) which defined keyword relevance of user query as a connectivity measure between query states modeled after the user queries has been issued. After performing RFI, An effective binary code has been generated and resulted by textual and visual hashing and thus the intrinsic structure of images is captured. Finally the retrieval performance of the images against large database is achieved and improved by using visually weighted compressed sensing approach which actually recovers the sparse data of the image and reconstructed. The visually weighted compressed sensing scheme generates the better visual improvements by the introduction of the visual weights. Extensive evaluations are conducted on publicly available datasets such as Google images and from various websites for different tasks and the comparison results have shown the dominance of the current research to many other methods.

REFERENCES

[1] Mahajan, Vishal, and Nishant Kashyap. "QUERY BY IMAGE CONTENT: PRINCIPLE, TECHNIQUES AND APPLICATION OF IMAGE RETRIEVAL." *IJCEM)* 3, no. 2 (2016).

[2] Grycuk, Rafał, Marcin Gabryel, Marcin Korytkowski, and Rafał Scherer. "Content-based image indexing by data clustering and inverse document frequency." In *International Conference: Beyond Databases, Architectures and Structures*, pp. 374-383. Springer, Cham, 2014.

[3] Rahmani, Md Khalid Imam, M. A. Ansari, and Amit Kumar Goel. "An Efficient Indexing Algorithm for CBIR." In *Computational Intelligence & Communication Technology (CICT), 2015 IEEE International Conference on*, pp. 73-77. IEEE, 2015.

[4] Kacem, Neila, Abir Gallas, and Ezzeddine Zagrouba. "General bases indexing using approximate structure techniques." In *Image Information Processing (ICIIP), 2013 IEEE Second International Conference on*, pp. 225-230. IEEE, 2013.

[5] Liu, Li, Mengyang Yu, and Ling Shao. "Unsupervised local feature hashing for image similarity search." *IEEE transactions on cybernetics* 46, no. 11 (2016): 2548-2558.

[6] Zhang, Xiaofan, Wei Liu, Murat Dundar, Sunil Badve, and Shaoting Zhang. "Towards large-scale histopathological image analysis: Hashing-based image retrieval." *IEEE Transactions on Medical Imaging* 34, no. 2 (2015): 496-506.

[7] G. Maheshwari, Prajitha.P " A Survey on Visual Hashing and Indexing Techniques for Content-Based Image Retrieval." *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 6, issue 7 (2017): 357-363.