Content-Based Natural language processing using Semantic Assisted Visual Hashing

Tanmayee Tushar Parbat[1], Honey Jain[1], Rohan Benhal[2], Sajeeda Shikalgar[3]

[1]B.E IT, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India
[2]BBA IT, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India
[3]Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

ABSTRACT

This is a new technology to support scalable content-based image retrieval (CBIR]), hashing has been recently been focused and future directions of research domain. In this paper, we propose a unique unsupervised visual hashing approach called semantic-assisted visual hashing (SAVH). Distinguished from semi-supervised and supervised visual hashing, its core idea emphatically extracts the rich semantics latently embedded in auxiliary texts of images to boost the effectiveness of visual hashing without any explicit semantic labels. To expand the reach, a unsupervised framework is advanced to learn hash codes by simultaneously preserving visual similarities of images, integrating the semantic assistance from texts on modeling high relationships of inter images and defining the correlations between images and shared contents.

Keywords : CBIR, Semantic Assistance, Visual Hashing, Text Auxiliaries, Unsupervised Learning.

I. INTRODUCTION

Advances in social media and mobile computing technology, the past decade has witnessed a tremendous growth in the availability of Web images. Consequently, there has been an increasing interest in the information retrieval and multimedia computing communities to study smart image retrieval techniques.

In particular, techniques for content-based image retrieval CBIR[1], where only visual image is used as query, are gaining in importance due to a wide range of applications such as video copyright management and visual recommendation systems To provide high quality content-based search services over huge volume of image collections, both efficiency and effectiveness are important issues. Advanced indexing structure is essential to scale the big data space and facilitate accurate search.

Its performance degrades greatly when the dimension of features to be goes high. Furthermore, both inverted file and tree structure consume large amount

of memory when storing corresponding data structures. This issue becomes even more serious when image collection scale is large.

According to the way about how to generate hash function, existing SFVH [10][15], can be further categorized into two major families: data-independent, and data-dependent hashing. SPH learns the hash functions by preserving the similarities of images in the mapped hash codes, while STH extends it for out of sample queries via linear support vector machine training in Hamming space.

An image is an artifact that represents visual perception, for example a two-dimensional picture, that has a similar appearance to some subject usually a physical or a person, thus providing a depiction of it. An image is a picture that has been created or copied and stored in electronic form. An image can be described in terms of vector graphics or raster graphics.

II. COMMON IMAGE FILE FORMATS ONLINE INCLUDE

JPEG is an image file produced according to a standard from the Joint Photographic Experts Group, an ISO/IEC group of experts that develops and maintains standards for a suite of compression algorithms for image files. JPEGs usually have a .jpg file extension.

As one of the technologies to support fast and accurate image search, visual hashing has received huge attention and became a very active research domain in last decade [8], [9].BMP is native file format of the Windows platform is like the parent format to the above three. BMP formats do not allow for image compression.BMP images are crisp and precise, but being pixel dependent they don't scale well.

Fig.1 2D-Image

Image Acquisition

Image acquisition in image processing can be broadly defined as the process of retrieving an image from some source, usually a hardware based source, so it can be passed through whatever processes need to occur.
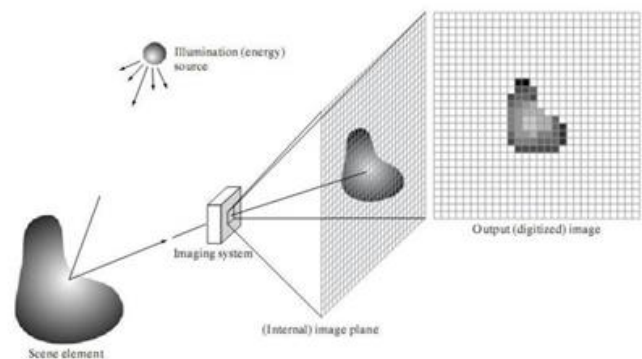

Fig.2. Image Acquisition Process

Performing image acquisition in image processing is always the first step in workflow sequence because, without an image, no processing is possible. The input images are taken from file. These images are different format like jpg, tiff, gif mostly we are using jpg format because it will accept black image and color image.

III. VISUAL FEATURES EXTRACTION

Visual Feature extraction starts from an initial set of measured data and builds derived values determine to be informative and non-redundant. Feature extraction is commonly related to dimensionality reduction and also, visual features usually have high dimensions. Visual Hashing basic idea is to map the raw high-dimensional visual features into binary codes, that visual similarities of images can be efficiently measured by simple but efficient bit-wise operations.
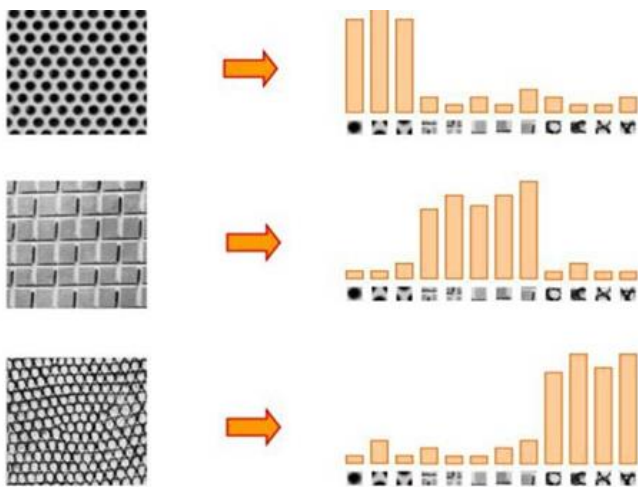

Fig.3. Feature Extraction Process

Classification

The classification performance is largely dependent on the descriptiveness and discriminativeness of feature descriptors. The hyper-graph is constructed based on the extracted visual features. Effectively preserving visual similarities of images in binary hash codes is essential to visual hashing. The text enhanced visual graph is constructed. The visual hash code learning is used to measure semantic similarity in Hamming space keep consistent with shared topic distributions.
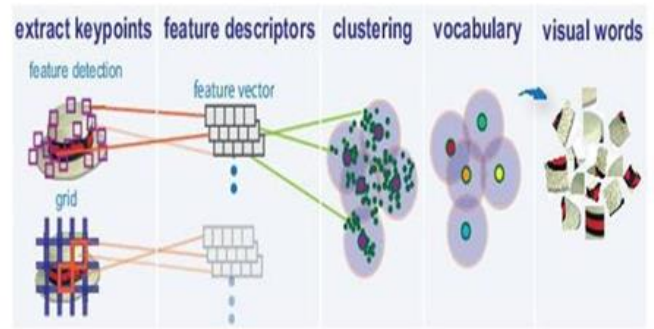

Fig.4. Classification Process

IV. RESULT EVALUATION

SAVH is specially designed for content based image retrieval CBIR [1] without any labeled images. SAVH has an significant advantage that its offline learning can effectively important semantics involved in text, while online hashing requires only visual image as input. Finally, it is to extract the semantic text and visual image from the semantic assisted visual hashing. Results highlight various advantages of SAVH and defines that SAVH significantly outperforms various methods.
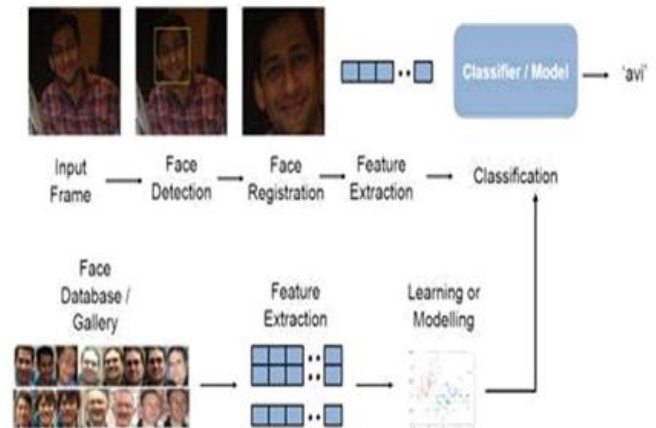

Fig. 5. Result Evaluation Process

V. EXISTING SYSTEM

According to strategies of leverage different kinds of visual features, existing unsupervised visual hashing schemes can be classified into several independent families: single feature visual hashing (SFVH [10][15])

and multiple feature visual hashing (MFVH[3]). Due to the lack of discriminative capability on representing high ranged semantics, the learned hash codes are not capable to characterize semantic correlations of images and the ones between images and semantics involved in the image database. It should be denoted that, one of the visual modality in MFVH [3] can be substituted with text modality. In such case, MFVH [3] becomes multi-modal hashing (MMH[3][4]).

Anyhow, it requires text modality at both stages of offline learning and online hashing. Due to its constraint, the scheme cannot meet the requirement of CBIR [1] in practical retrieval applications. Where only visual image is uploaded as query. However, the main design aim of various UCMH approaches is to enable multimedia retrieval process across heterogeneous modalities. It assumes that each and every type of the involved modality contributes equally to cross-modal retrieval. According to the way about how to generate hash function, existing SFVH can be further categorized into two major families: data-independent [16] hashing, and data-dependent hashing [10], [15].

Thus the assumption makes the method less discriminative and effective. In addition, discriminative information latently involved in original visual feature may be lost in hashing process due to mandatory heterogeneous modality correlation. The most significant limitation of SFVH [10][15] and MFVH[3] is that they only take the features from visual modality into account. Due to the semantic gap, relations between the images are characterized by low- level visual feature cannot effectively describe rich image semantics, therefore making the hash codes less semantically meaningful.

Disadvantages
✓ Semantic Information might be lost. Considers only text input query.

✓ Both the inverted file and tree structure consume large amount of memory, when storing corresponding data structures.
✓ The power of the search strategies is very difficult to optimize the retrieval quality.

VI. PROPOSED METHODOLOGY

Unsupervised visual hashing scheme is proposed and also termed as semantic-assisted visual hashing (SAVH), to effectively perform visual hashing learning with semantic assistances. The key idea of this technique is to extract semantics automatically from the noisy associated texts to enhance the discriminative capability of hash codes, and thus facilitate the performance improvement of visual hashing. The core idea of SAVH is to effectively leverage the semantics embedded in the associated informative texts about images to assist visual hashing. With the assistance of texts, the generated visual hash codes and functions can awake high-level semantics, and thus they will be much discriminativeness. SAVH technology works as follows: First, hash code learning is formulated in a unified unsupervised structure, where hash codes are learned by simultaneously preserving visual similarity of images and considering the assistance of texts associated with images. More exactly, our framework integrates two important assistance of auxiliary texts to effectively mitigate the latent limitations of visual features. 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.

Initially, visual and text features of images are extracted to transform image pixels to mathematical vector representations. After that, a text-enhanced visual graph is constructed with the assistance of topic hyper graph, and latent semantic topics are detected under guidance of text information n. Then, hash

codes of database images are learned in a framework which preserves correlations between images and semantic topics. In the proposed work, retrieval performance increases steadily with hash code length. It is to extract the image as well as content related to the hash code. Finally it lists high ranking data among retrieved data. Also it gets inter-related images as well as semantics embedded with those images. SAVH can yield a better performance than many compared approaches with longer hash codes.

Advantages:
- ✓ Fast query response.
- ✓ More effective.
- ✓ Low storage consumption.
- ✓ An important advantage that its offline learning can effectively leverage semantics involved in text.
- ✓ It retrieves highly ranked data.

Algorithm 1: Hashing Algorithm

Input: Database images: fIngNn=1, query image q.
Output: Hash codes of database images: Y, hash functions: F. Image retrieval results for image query q.
Step 1 : Offline Learning
- ✓ Extract features of database images, obtaining Y(1); Y(2);
- ✓ Compute visual graph Laplacian matrix LG;
- ✓ Compute topic hypergraph Laplacian matrix LTHG ;
- ✓ Learn relaxed hash codes;
- ✓ Construct hash functions G ;
- ✓ Insert database images into binary hash codes with G;

Step 2:Online Hashing
- ✓ Extract visual features of query image;
- ✓ Project query visual features into the hash codes;
- ✓ Calculate the Hamming distances between hash codes of query image and that of database images;

- ✓ Rank Hamming distances and return high ranked retrieval results.
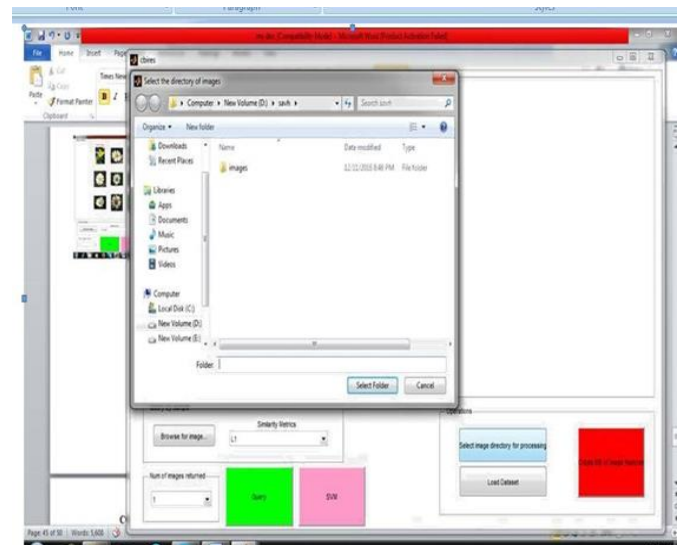
VII. EXPERIMENTAL RESULTS
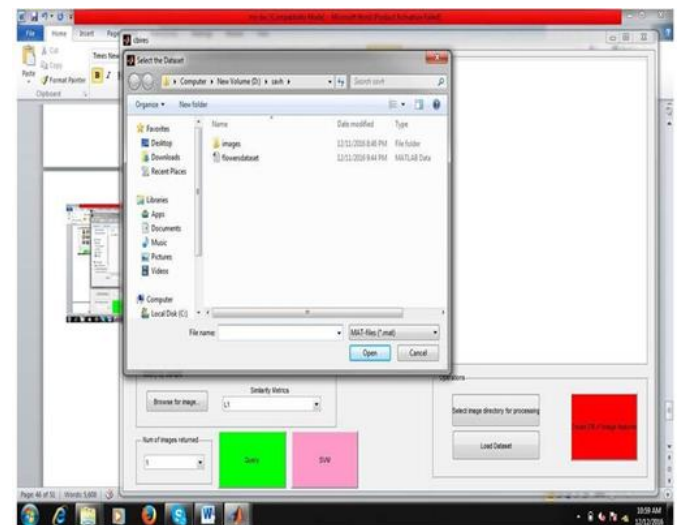

Fig. 6: Selecting Image Directory
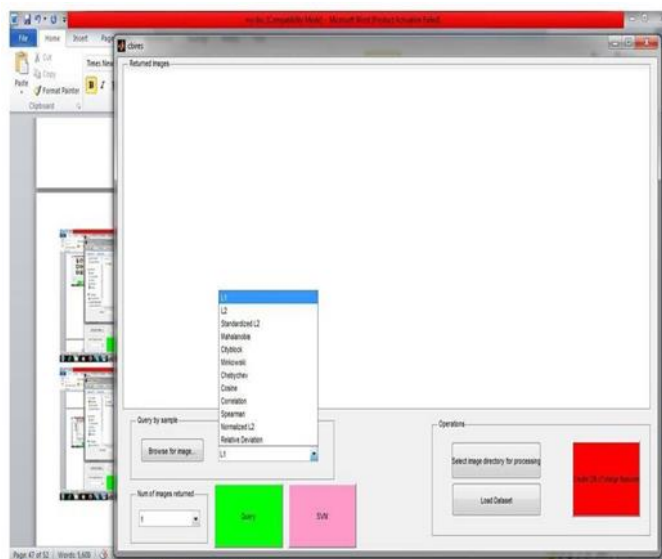

Fig.7 loading the Dataset

Fig.8 Selecting Metrics

VIII. CONCLUSION AND FUTURE SCOPE

Most existing single feature and multiple feature hashing approaches for CBIR[1] approaches build their schemes with only visual features. They also ignore the valuable semantics involved in the related texts. Although Unsupervised cross-modal hashing approaches can latent text for retrieval task across heterogeneous modalities, they equally treat visual and text, and still fail to fully take advantages of text. Different from them, this method proposes an effective hashing framework, SAVH. Our idea is leveraging the associated texts of images to benefit the visual hashing using unsupervised learning. SAVH can organize extra discriminative information into the generated visual hash codes and their functions. while its online hashing requires only visual image as input query.

This method opens up several promising directions for further exploration. Notably, it is interesting for further validation of the effectiveness of SAVH when more associated modalities are involved. For instance, geographical location of image, social correlation of images, and etc. Moreover, it will be interesting to investigate the effectiveness of visual image on assisting hashing for text retrieval.

IX. REFERENCES

[1]. J.-H. Su, W.-J. Huang, P. S. Yu, and V. S. Tseng, ―Efficient relevance feedback for content-based image retrieval by mining user navigation patterns.‖ IEEE Trans. Knowl. Data Eng., vol. 23, no. 3, pp. 360–372, 2011.

[2]. P. Wu, S. C. H. Hoi, P. Zhao, C. Miao, and Z. Liu, ―Online multi-modal distance metric learning with application to image retrieval,‖ IEEE Trans. Knowl. Data Eng., vol. 28, no. 2, pp. 454–467, 2016.

[3]. J. Song, Y. Yang, Z. Huang, H. T. Shen, and J. Luo, ―Effective multiple feature hashing for large-scale near-duplicate video retrieval.‖ IEEE Trans. Multimedia, vol. 15, no. 8, pp. 1997– 2008, 2013.

[4]. J. Song, Y. Yang, Z. Huang, H. T. Shen, and R. Hong, ―Multiple feature hashing for real-time large scale near-duplicate video retrieval,‖ in Proc. ACM Int. Conf. Multimedia (MM), 2011, pp. 423– 432.

[5]. V. Jagadeesh, R. Piramuthu, A. Bhardwaj, W. Di, and N. Sundaresan, ―Large scale visual recommendations from street fashion images,‖ in Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (SIGKDD), 2014, pp. 1925–1934.

[6]. J. Sivic and A. Zisserman, ―Video Google: A text retrieval approach to object matching in videos,‖ in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR), 2003, pp. 1470–1477.

[7]. P. Ciaccia, M. Patella, and P. Zezula, ―M-tree: An efficient access method for similarity search in metric spaces,‖ in Proc. Int. Conf. Very Large Data Bases (VLDB), 1997, pp. 426– 435.

[8]. J. Wang, H. T. Shen, J. Song, and J. Ji, ―Hashing for similarity search: A survey,‖ CoRR, vol. abs/1408.2927, 2014.

[9]. L. Gao, J. Song, X. Liu, J. Shao, J. Liu, and J. Shao, ―Learning in high-dimensional multimedia data: the state of the art,‖ Multimedi Syst., pp. 1–11, 2015.

[10]. Y. Weiss, A. Torralba, and R. Fergus, ―Spectral hashing,‖ in Proc. Advances in Neural Information Processing Systems (NIPS), 2008, pp. 1753–1760.

[11]. M.D. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. afner, D. Lee, D. Steele, and P. Yanker,―Query by Image and Video Content: The QBIC System,‖Computer, vol. 28, no. 9, pp. 23-32, Sept. 1995.

[12]. R. Fagin, ―Combining Fuzzy Information from Multiple Systems,‖Proc. Symp. Principles of Database Systems (PODS), pp. 216-226, June 1996.

[13]. R. Fagin, ―Fuzzy Queries in Multimedia Database Systems,‖ Proc. Symp. Principles of Database Systems (PODS), pp. 1-10, June 1998.

[14]. J. French and X-Y. Jin, ―An Empirical Investigation of the Scalability of a Multiple Viewpoint CBIR[1] System,‖ Proc. Int'l Conf. Image and Video Retrieval (CIVR), pp. 252-260, July 2004.

[15]. D. Zhang, J. Wang, D. Cai, and J. Lu, "Self-taught hashing for fast similarity search," in Proc. Int. ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2010, pp. 18–25.