ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

Vol. 3, Issue 8, August 2015

# A Literature Survey on Query Adaptive Image Search using Hash Code

Vidya B. Phad, S. V. Patil

Student, Dept. of E&TC, J. T. Mahajan College of Engineering, Faizpur. NMU University, Maharashtra, India

Associate Professor, Dept. of E&TC, J. T. Mahajan College of Engineering, Faizpur, NMU University, Maharashtra, India

**ABSTRACT:** Images have a greatest importance in computer vision. Query image search based on visual similarities has been lively topic of research in recent years. Increasing use of images on the internet, there is a robust need to develop techniques for efficient and scalable image search. Usually hashing methods are used to embed high dimensional image features into Hamming space, where search can be performed in real-time based on Hamming distance of compact hash codes. There are large numbers of images sharing same hamming distances to a query image, so that fine-grained ranking is very important. This paper proposed query adaptive image retrieval system to retrieve image with equal hamming distance to the query. This is achieved by firstly extracting the features of an individual image and then features are embedded into hash codes and store in a database. Query adaptive weights are then calculated by evaluating the proximity between a query and the semantic concept classes. In this system images are represented using the popular bag-of-visual words (BOW) framework, local invariant image descriptors are extracted and quantized as per a set of visual words. By using Flicker image dataset for experiments, it shows clear improvement from our proposed approach.

**KEYWORDS**: Query-adaptive image search, scalability, binary hash codes, weighted Hamming distance.

## I. INTRODUCTION

With the wide range of images on the Internet, there is a strong need to develop techniques for efficient and scalable image search. While traditional image search engines heavily rely on textual words related to the images, scalable content-based search is receiving increasing attention. Apart from providing better image search experience for ordinary Web users, large-scale similar image search has also been demonstrated to be very helpful for solving a number of very hard problems in computer vision and multimedia such as image classification [1].

Generally a large-scale image search system consists of two Parts an effective image feature representation and an efficient search mechanism. We know that the quality of search results relies heavily on the representation power of image features. After that, an efficient search mechanism, is critical since existing image features are mostly of high dimensions and current image databases are huge, on top of which exhaustively comparing a query with every database sample is computationally prohibitive.

In this paper we represent images using the popular bag-of- visual-words (BoW) framework [6], where local invariant image descriptors (e.g., SIFT [2]) are extracted and quantized based on a set of visual words. The BoW visual features are then embedded into compact hash codes for efficient search. For this, we consider one of the most well-known hashing methods is Locality Sensitive Hashing (LSH). Hashing is preferable over tree-based indexing structures (e.g., kd-tree [7]) because it generally requires greatly reduced memory and also works better for high-dimensional samples. With the hash codes, image similarity can be efficiently measured (using logical XOR operations) in Hamming space by Hamming distance, an integer value obtained by counting the number of bits at which the binary values are different. In large scale application hamming space should be set as small number.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

# International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 8, August 2015**

The main contribution of this paper is to computes query-adaptive weights for each bit of the hash codes. To this end, we collect a set of semantic concept classes that cover many semantic aspects of image content (e.g., scenes and objects). Bitwise weights for each of the semantic classes are learned offline using a novel formulation that maximizes intra-class sample similarities and stores inter-class relationships. We show that the optimal weights can be computed by iteratively solving quadratic programming problems. These pre-computed class-specific bitwise weights are then utilized for online computation of the query-adaptive weights, through rapidly evaluating the similarities of a query image to the image samples of the semantic classes. After that, weighted Hamming distance is applied to evaluate similarities between the query and images in a target database. This weighted distance is known as query-adaptive Hamming distance. We observe that during online search it is unnecessary to compute the weighted Hamming distance based on real-valued vectors (weights imposed on the hash codes), which would bury one of the most important advantages of hashing.


**Fig. 1**. Search result lists in a Flicker image dataset, using a “sunset scene” query image (left). Top and bottom rows respectively show the most similar images based on traditional Hamming distance and our proposed query-adaptive weighted Hamming distance. It can be clearly seen that our method produces more relevant result by ranking images at a finer resolution. Note that the proposed method does not permute images with exactly the same code to the query (three images in total for this query), i.e., Hamming distance = 0. This figure is best viewed in color.

Figure 1 shows search results of a query from a Flicker image dataset. We see that the proposed approach gives clearly better result (bottom row) by ordering images with Hamming distance 1 to the query. observe that in most cases there is typically a very small number, if not zero, of images with Hamming distance 0 to search queries , as there is only one hash code satisfying this condition ( $c_a^i$ , I > 0).

## II. LITERATURE REVIEW

There are very good literature surveys on general image retrieval task. Smeulders et al. [8] and Datta et al. [9] for work from the past decade. Many people extracts simple features such as color and texture in systems developed in the early years , while more effective features such as GIST [10] and SIFT [2] have been popular recent years . In this work, we choose the popular bag-of-visual-words (BoW) [6] representation based on the local invariant SIFT features. This feature representation is efficient for image search as in various applications. Since the work in this paper is more related to efficient search, mainly reviews existing works on efficient search mechanisms, which are roughly divided into three categories: inverted file, tree-based indexing, and hashing.

### 2.1 Inverted file

Inverted index is very popular for document retrieval in the informational retrieval community. Inverted index was a initially proposed mechanism. In inverted index, a list of references to each document (image) for each text (visual) word is created so that relevant documents (images) can be quickly located given a query with several words. A main difference of document retrieval from visual search, however, is that the textual queries usually contain very few words. In Google web search, on average there are merely 4 words per query. While in the Bow representation, a single

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

**International Journal of Innovative Research in Computer and Communication Engineering**

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 8, August 2015**

image may contain hundreds of visual words, resulting in a large number of candidate images (from the inverted lists) that need further verification—a process that is usually based on similarities of the original Bow features. This largely limits the application of inverted files for large scale image search. While increasing visual vocabulary size in Bow can reduce the number of candidates, it will also significantly increase memory usage. For example, indexing 1 million Bow features of 10,000 dimensions will need 1GB memory with a compressed version of the inverted file. for the binary representation in hashing methods, the memory consumption is much lower (e.g., 48MB for 1 million 48-bit hash codes).

**2.2 Treebased indexing**

Indexing with tree-like structures [7], has been frequently applied to fast visual search. Nister and Stewenius [11] used a visual vocabulary tree to achieve real-time object retrieval in 40,000 images. Muja and Lowe [12] adopted multiple randomized *k*d-trees [7] for SIFT feature matching in image applications. One disadvantage of the classical tree-based methods is that they normally do not work well with high-dimensional feature. For example, let the dimensionality be $d$ and the number of samples be $n$, by rule $n \gg 2^d$ in order to have *k*d-tree working more efficiently than exhaustive search. There are also several works focusing on improving tree-based approaches for large-scale search , where promising image search performance has been reported. Compared with these methods, hashing has a major advantage in speed since it allows constant-time search.

**2.3 Hashing**

In review of the limitations of both inverted file and tree-based indexing, embedding high-dimensional image features into hash codes has become very popular recently. Hashing satisfies both query time and memory requirements as the binary hash codes are compact in memory and efficient in search via hash table lookup or bitwise operations. Hashing methods normally use a group of projections to divide an input space into multiple buckets such that similar images are likely to be mapped into the same bucket.

## III. PROPOSED SYSTEM

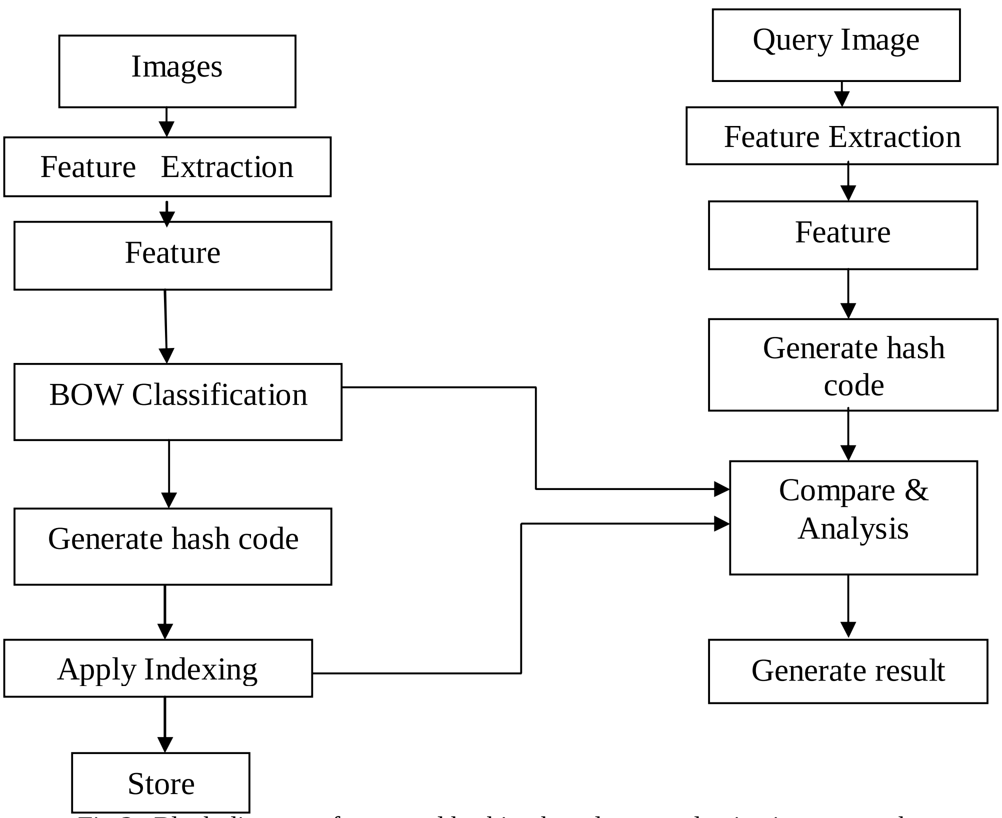


Fig.2 Block diagram of proposed hashing based query adaptive image search system.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 8, August 2015**

In this system there are two images taken as a input to the system. First is query image which is to be search and another is image from a dataset. Here we represent images using the popular bag-of-visual-words (Bow) model [6], where local invariant image descriptors (e.g., SIFT [2]) are extracted and quantized based on a set of visual words. The Bow feature vectors are then embedded into compact hash codes for efficient search and store in a feature database. Query image feature vectors and image dataset feature vectors store in database are compare. This process usually called similarity matching and calculating the hamming distance in a hamming space. Similarity of two images depends on the distance in feature space between the feature points defined by the vectors. We called this distance as a query adaptive hamming distance. Shorter the distance between two points, the images are similar. Finally, it retrieves image which is similar.

**Bow model**

The bag-of-words model (BoW model)[6] is used for image classification, by treating image features as words. Bow create a visual words, or bag of features, by extracting feature descriptor from representative image of each class. Here we use a famous descriptor is scale invariant feature transform (SIFT). SIFT converts each patch to 128-dimensional vector. Each image is a collection of vectors of the same dimension (128 for SIFT. The BoW model is then convert vector represented patches to "codeword", which also produces a "codebook". A codeword can be considered as a representative of no. of similar patches. By performing k-means clustering [13] over all the vectors, we get Codeword defined as the clusters. The number of the clusters is the codebook size .Finally, each patch in an image is mapped to a certain codeword through the clustering process and the image can be represented by the histogram of the codeword.

**Hashing techniques**

There are two hashing techniques adopted semi-supervised hashing [4] and semantic hashing with deep belief networks [6] Review from paper most of the existing hashing techniques are unsupervised. Among them, one of the most known hashing methods is Locality Sensitive Hashing (LSH) [4]. Recently, Kulis and Grauman [14] extended LSH to work in arbitrary kernel space, and Chum et al. [15] proposed min-Hashing to extend LSH for sets of features. Weiss et l. [16] proposed a spectral hashing (SH) method that hashes the input space depend on data distribution. SH also ensures that the projections are orthogonal and sample number is balanced over different buckets.

Although SH can achieve similar or even better performance than LSH with a fewer number of bits, it is important to view that these unsupervised hashing techniques are not robust enough for similar image search. The fact is that similarity in image search is not simply equivalent to the proximity of low-level visual features, but is more related to high-level image semantics (e.g., objects and scenes). Several supervised methods have been proposed to learn hash functions by minimizing reconstruction error between original feature distances and Hamming distances of hash codes. By replacing the original feature distances with semantic similarities, this method can be applied for supervised learning of hash functions.

## IV. QUERY-ADAPTIVE SEARCH

Hashing based, scalable image search can be performed in Hamming space using Hamming distance. Hamming distance can be defined as distance between two hash codes where total number of bits at which the binary values are different. In this paper, we propose to learn query adaptive weights [5] for each bit of the hash codes, so that images with the same Hamming distance to the query can be ordered in a finer resolution. Specific indices/locations of the bits with different values are not considered. For example, given three hash codes P = 1100, Q = 1111, and R = 0000, the Hamming distance of P and Q is equal to that of P and R, regardless of the fact that R differs from P in the first two bits while Q differs in the last two bits. Due to this nature of the Hamming distance, practically there can be hundreds or even thousands of samples sharing the same distance to a query. Going back to the example, suppose we knew that the first two bits are more important (discriminative) for P, then Q should be ranked higher than R if P was the query.

## V. CONCLUSION

In this paper, we have reviewed a novel framework for query adaptive image search. By collecting a large set of predefined semantic concept classes, our approach is able to predict query-adaptive bitwise weights of hash codes in real-time, with which search results can be rapidly ranked by weighted Hamming distance at finer-grained hash code

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

level. In this system BOW model is used for image classification, SIFT extracts the features and then embedded in to compact hash code for efficient search. We computes query adaptive image weights for each bit of hash code. Each hash code has unique similarity to query. This system proposes a means to rank images at a finer resolution. This system is useful in hospitals, multimedia, web search etc.

## REFERENCES

[1] Yu-Gang Jiang, Jun Wang, Xiangyang Xue, and Shih-Fu Chang, "Query-Adaptive Image Search with Hash Codes", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 2, FEBRUARY 2013.
[2] D. Lowe, "Distinctive image features from scale-invariant keypoints,"International Journal on Computer Vision, vol. 60, no. 2, pp. 91–110 ,2004.
[3] J. Wang, S. Kumar, and S.-F. Chang, "Semi-supervised hashing for scalable image retrieval," in IEEE Conf. on Computer Vision and Pattern Recognition, 2010.
[4] B. Kulis and K. Grauman, "Kernelized locality-sensitive hashing for scalable image search," in IEEE International Conference on Computer Vision, 2009.
[5] Y.-G. Jiang, J. Wang, and S.-F. Chang, "Lost in binarization: Query adaptive ranking for similar image search with compact codes," in ACM International Conference on Multimedia Retrieval, 2011.
[6] Csurka, G., C. R. Dance, L. Fan, J. Willamowski, and C. Bray. Visual Categorization with Bags of Keypoints. Workshop on Statistical Learning in Computer Vision. ECCV 1 (1–22), 1–2.
[7] J. L. Bentley, "Multidimensional binary search trees used for associative searching," Communications of the ACM, vol. 18, no. 9, pp. 509–517, 1975.
[8] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349–1380, 2000.
[9] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: ideas, influences, and trends of the new age," ACM Computing Surveys, vol. 40, no. 2, 2008.
[10] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," International Journal on Computer Vision, vol. 42, pp. 145–175, 2001.
[11] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in IEEE Conf. on Computer Vision and Pattern Recognition, 2006.
[12] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in International Conference on Computer Vision Theory and Applications, 2009, pp. 331–340.
[13] Deza, Elena; Deza, Michel Marie (2009). Encyclopedia of Distances. Springer. p. 94."Cluster analysis". March 2, 2011.
[14 ]Lowe, David G. (1999). "Object recognition from local scale-invariant features"CCV.1999.790410.

## BIOGRAPHY

**Vidya B. Phad** is a Student in the Electronics and Telecommunication Department, J. T. Mahajan College of engineering Faizpur, India. North Maharashtra University. She received Bachelor degree in Electronics and Telecommunication in 2010 from, MET'S IOE Nasik, India.

**Mr. S. V. Patil** is Associate Professor in Electronics and Telecommunication Department, J. T. Mahajan College of engineering, Faizpur, India.