IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

A Survey on Visual Hashing and Indexing Techniques for Content-Based Image Retrieval

Mrs. G. Maheshwari[1], Prajitha.P[2]

Assistant Professor, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India[1]

M. Phil Scholar, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu, India[2]

Abstract: Interest in the potential data and multimedia content retrieval has increased enormously over the last few years. Users in many professional fields are exploiting the opportunities offered by the ability to access and manipulate remotely-stored data's in all kinds of new and exciting ways. Visual data hashing and indexing is an important technique for content based image and information retrieval, which is a part of information retrieval system. Content-based image retrieval (CBIR) systems are very effective in many types of applications. Despite extensive research efforts for decades, how to discover and incorporate semantic information of images still poses a formidable challenge to real-world CBIR systems. In this paper, we surveys the tools and techniques are used to perform content based image and information retrieval.

Keywords: Content-based Image Retrieval, Information Retrieval, Data Mining, Semantics Modeling, Hashing, Indexing.

I. INTRODUCTION

Content-based image retrieval (CBIR) is a method that utilizes visual contents in the process of search images from large scale image databases according to users' interests It has been an active and fast advancing research area since several years [1]. Over a few years, remarkable progress has been made by CBIR in both theoretical research and system development. However, there remain many challenging research problems that continue to attract researchers from multiple applications. Content-based retrieval uses the contents of images to represent and access the images. A typical content-based retrieval system is divided into two categories such as offline feature extraction and online image retrieval [2]. The CBIR system framework is illustrated in fig 1.0.

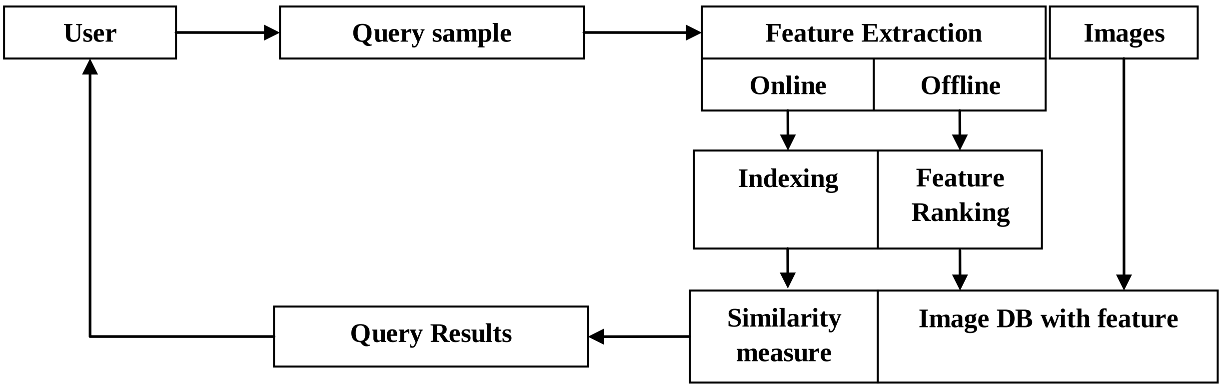

Fig 1.0 Content based image retrieval framework

In offline stage, the system automatically extracts visual attributes such as color, tags of the image, texture and spatial information of every image in the database based on its pixel values and stores them in a different database within the system called feature database. The feature data ,which is also known as image signature for each of the visual attributes of each image is very much smaller in size compared to the image data, thus the feature database contains an abstraction of the images in the image database. One advantage of a signature over the original pixel values is the significant compression of image representation. However, a more important reason for using the signature is to gain an improved correlation between image representation and visual semantics. In on-line image retrieval, the user can submit a query example to the retrieval system in search of desired images. The distances (i.e., similarities) between the feature vectors of the query example and those of the media in the feature database are then computed and ranked. Information Retrieval is made effective when there is a proper indexing process. Finally, the system ranks the search

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

results and then returns the results that are most similar to the user query. If the user is not satisfied with the search results, they can provide relevance feedback (RF) [3] to the retrieval system, which contains a mechanism to learn the user's information needs. CBIR system runs on a totally different standard, which retrieves stored images from a set of images by comparing features automatically extracted from the images themselves. The most important features used in the CBIR system are mathematical measures of color, texture or shape; hence virtually all current CBIR systems, whether commercial or experimental, operate at level by level.

A typical system allows users to formulate queries by submitting an example of the type of image being sought, though some offer alternatives such as selection from a palette or sketch input. The system then identifies those stored images whose feature values match those of the query most closely. As a result of recent advancements in digital storage technology, it is now possible to create large and extensive databases of digital imagery. These collections may contain millions of images and terabytes of data. For users to make the most of these databases effective, efficient methods of searching must be developed. 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 based on a CBIR 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 used in the indexing procedure are kept constant, all images will be indexed consistently, solving the inherent problems resulting from fallible human-based indexing. In this paper, the data mining techniques that are used in the CBIR systems are reviewed. The techniques in the literature are classified into relevance feedback system, indexing methods and visual hashing techniques. Under these categories the list of techniques are reviewed and finally the problem outline are defined with future enhancement ideas.

II. LITERATURE REVIEW

Content based image retrieval is researched by several authors. And many techniques are deployed with several performance considerations. The literature part is defined by three categories which are CBIR with RF (Relevance Feedback methods), indexing methods for CBIR and visual hashing techniques for CBIR.

A. Relevance Feedback Methods: A relevance feedback (RF) approach allows a user to interact with the retrieval algorithm by providing the information of which images user thinks are relevant to the query. Authors in [4] analyzed and used relevance feedback technique for image retrieval. Various parameter estimation techniques have been proposed for relevance feedback. Additionally, in the paper, the methods that perform optimization on multilevel image content model have been formulated. However, these methods only perform relevance feedback on low-level image features and fail to address the images' semantic content. In content based image retrieval, there are two types of input has collected one is text and another from the content features. There is a problem in the text based image retrieval, that some images may not have appropriate tags to describe the images. So the results may not accurately generated and this task is challenging in CBIR system, to overcome this problem, a solution called RF is proposed in [5] [6], this uses to reduce the possible errors and improves the accuracy. Later with RF, Bayesian classifier is used in [7][8] which deals with positive and negative feedbacks from the user to mine both related and unrelated results. Authors in [9] proposed vector model, which is the most popular models in the information retrieval. Numerous effective retrieval techniques have been developed for this method, and this is considered as the most effective one in the relevance feedback category. Most of the previous relevance feedback research can be classified into two approaches as: reweighting and query point movement which was studied in [10].

B. Indexing Methods: Authors in [11] proposed a generative model, called the citationtopic (CT) model, for modeling linked documents that explicitly considers the relations among documents. This perspective actually reflects the process of writing a scientific article: the authors probably first learn knowledge from the literature and then combine their own creative ideas with the learned knowledge to form the content of the paper. Furthermore, to capture the indirect relations among documents, the proposed model contains a generative process to select related documents where the related documents are not necessarily directly linked to the given document. The CT model was applied to the document clustering task and the experimental comparisons against several state-of-the-art approaches demonstrate very promising performances. With the advent of digital databases and communication networks, huge repositories of textual data have become available to a large public. Today, it is one of the great challenges in the information sciences to develop intelligent interfaces for human-machine interaction which supports computer users in their quest for relevant information. Although the use of elaborate ergonomic elements like computer graphics and visualization has proven to be extremely fruitful to facilitate and enhance information access, progress on the more fundamental question

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

of machine intelligence is ultimately necessary to ensure substantial progress on this issue. In order for computers to interact more naturally with humans, one has to deal with the potential ambivalence, impreciseness, or even vagueness of user requests and has to recognize the difference between what a users might say or do and what she or he actually meant or intended. One typical scenario of human machine interaction in information retrieval is by natural language queries: the user formulates a request, e.g., by providing a number of keywords or some free-form text and expects the system to return the relevant data in some amenable representation, e.g., in form of a ranked list of relevant documents. Many retrieval methods are based on simple word matching strategies to determine the rank of relevance of a document with respect to a query. Yet, it is well known that literal term matching has severe drawbacks, mainly due to the ambivalence of words and their unavoidable lack of precision as well as due to personal style and individual differences in word usage.
Latent Semantic Analysis (LSA) is an approach to automatic indexing and information retrieval that attempts to overcome these problems by mapping documents as well as terms to a representation in the so{called latent semantic space. This kind of automatic indexing usually takes the multi high dimensional vector space representation of documents based on term frequencies as a starting point and applies a dimension reducing linear projection. The specific form of this mapping is determined by a given document collection and is based on a Singular Value Decomposition (SVD) of the corresponding term/document matrix. The general claim is that similarities between documents or between documents and queries can be more reliably estimated in the reduced latent space representation than in the original representation. The rationale is that documents which share frequently co-occurring terms will have a similar representation in the latent space, even if they have no terms in common. LSA thus performs some sort of noise reduction and has the potential benefit to detect synonyms as well as words that refer to the same topic. In many applications this has proven to result in more robust word processing. Although LSA has been applied with remarkable success in different domains including automatic indexing - Latent Semantic Indexing (LSI), it has a number of deficits, mainly due to its unsatisfactory statistical foundation. The primary goal is to present a novel approach to LSA and factor analysis called Probabilistic Latent Semantic Analysis (PLSA) that has a solid statistical foundation, since it is based on the likelihood principle and defines a proper generative model of the data [12]. This implies in particular that standard techniques from statistics can be applied for questions like model fitting, model combination and complexity control. In addition, the factor representation obtained by PLSA allows to deal with polysemous words and to explicitly distinguish between different meanings and different types of word usage. Image auto-annotation, i.e., the association of words to whole images, has attracted considerable attention. In particular, unsupervised, probabilistic latent variable models of text and image features have shown encouraging results, but their performance with respect to other approaches remains unknown.

Authors in [13] applied and compared two simple latent space models commonly used in text analysis, namely Latent Semantic Analysis (LSA) and Probabilistic LSA (PLSA). Annotation strategies for each model are discussed. Remarkably, it was found that on an 8000-image dataset, a classic LSA model defined on keywords and a very basic image representation performed as well as much more complex, state-of the-art methods. Furthermore, non-probabilistic methods like LSA and direct image matching techniques outperformed PLSA on the same dataset. The potential value of large image collections can be fully realized only when effective methods for access and search exist. Image users often prefer to formulate intuitive text- based queries to retrieve relevant images, which requires the annotation of each image in the collection. Automatic image annotation has thus emerged as one of the key research areas in multimedia information retrieval. It reduces the issues like manual caption generation, costly etc. Based on latent space models in text analysis, generative probabilistic models for auto-annotation have been proposed in [14], this includes variations of PLSA and permission to make digital or hard copies of all or part of. The work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Later, Latent Dirichlet Allocation (LDA) models use a latent variable representation for unsupervised learning of co-occurrences between image features and words in an annotated image collection, and later employ the learned models to predict words for unlabeled images. The latent space representation can capture high-level relations within and across the textual and visual modalities. Specific assumptions introduce variations in the ways in which co-occurrence information is captured. However, with a few exceptions, most previous works assume that words and visual features should have the same importance in defining the latent space. There are limitations with this view. First, the semantic level of words is much higher than the one of visual features extracted even by state-of- the-art methods. Second, in practice, visual feature co occurrences across images often do not imply a semantic relation between them. This results in a severe degree of visual ambiguity that in general cannot be well handled by existing joint models. For auto-annotation, in the paper [14] is ultimately interested in defining a latent space that is consistent in semantic terms, while able to capture multimodal co-occurrences. A novel approach to achieve the above goal has been presented, based on a linked pair of PLSA models. The definition of the latent space is constraint by focusing on textual features first and then learning visual variations conditioned on the space learned from text.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

As like the above, the authors in [15] proposed a model which consistently outperforms previous latent space models, while retaining the elegant formulation of annotation as probabilistic inference. Hofmann (2001) presented a novel statistical method for factor analysis of binary and count data which is closely related to a technique known as Latent Semantic Analysis. In contrast to the latter method which stems from linear algebra and performs a SVD of co-occurrence tables, the proposed technique uses a generative latent class model to perform probabilistic mixture decomposition. This results in a more principled approach with a solid foundation in statistical inference. More precisely, it was proposed to make use of a temperature controlled version of the Expectation maximization algorithm for model fitting, which has shown excellent performance in practice. Probabilistic Latent Semantic Analysis has many applications, most prominently in information retrieval, natural language processing, machine learning from text and in related areas. Authors in [16] presented perplexity results for different types of text and linguistic data collections and discusses an application in automated document indexing. The experiments indicate substantial and consistent improvements of the probabilistic method over standard Latent Semantic Analysis work includes search by text, search by image feature similarity, search by segment features, search for specific types of images using more compressive methods and search by image sketch. A few systems combine text and image data. Search using a simple conjunction of keywords and image features is provided in Blob world. Here the image segment color is translated in a pre-processing step into one of a handful of color categories. Thus, image feature search becomes a text search and standard database systems can be used for the query. This is efficient, but potential for more sophisticated use of image features is limited. Web seer uses similar ideas for query of images on the web, but also indexes the results of a few automatically estimated image features. These include whether the image is a photograph or a sketch and notably the output of a face finder. Going further, some text and histogram data were integrated in the indexing. Others have also experimented with using image features as part of a query refinement process. The model which builds on is developed. Others also argue for statistical models of data for image retrieval. Finally, the work in [17] mentions the area of using associated text for image understanding. For the image retrieval component of this work, it was insisted that browsing was well supported. This is in contrast with many existing systems where the main access to the images is through query. This puts the burden on the user to pose the correct question and the system to provide the correct prompts. It is simply easier for the user to find an image of interest if some structure can be imposed on the collection and exposed to the user. Other work emphasizing this philosophy includes the application of multidimensional scaling using the Earth Mover's distance to image displays. The interest in browsing leads us to consider a hierarchical model which imposes a course to fine, or general to specific, structure on the image collection. Such structure is part of semantics and therefore a hierarchical system is proposed that is better poised to capture semantics than a flat one

C. Visual Hashing methods:

Single Feature Visual Hashing (SFVH): In the CBIR systems, there are several hashing methods are used in the literature. This reviews the single feature visual hashing techniques. The SFVH is categorized into two major types which are data-independent and data dependent hashing. Under Data dependent hashing, Locality sensitive hashing (LSH) [18] is used this most typical one. This used the random vectors like standard Gaussian distribution. This maps the points with high similarity and high probability. The data dependent hashing schemes are developed to study the hash functions according to the uniqueness of underlying data distribution by using machine learning methods. Spectral hashing (SPH) is proposed in [19], which is an optimal way to accelerate similarity search. It preserves the image similarities in the hash code format. As like SPH, the anchor graph hashing (AGH) [20], and self-taught hashing (STH) [21] are deployed. The similarities between these hashing techniques are the unsupervised hashing nature. The linearSVM [22] methods are developed with the hamming space training, which extends the STH technique. This improves the performance by eliminating the previous hashing technique by adopting sample queries in the training model. The affinity graph hashing method (AGH) approximates the affinity graph with low rank matrix, and learns the hash functions by binarizing the hashing functions. Iterative quantization (ITQ) [24] is proposed to reduce the quantization loss by rotating the learned hash codes. The types of SFVH and its techniques are described with its advantage in table 1.0

Table 1.0 Single Feature Visual Hashing techniques

Category	**SFVH**					
Techniques	**LSH**	**SPH**	**AGH**	**STH**	**Linear SVM**	**ITQ**
Abbreviations	Locality Sensitive Hashing	Spectral hashing	affinity graph hashing	self-taught hashing	linear support vector machine	Iterative quantization
Advantages	Useful for high dimensional data	Easy semantic retrieval	unsupervised hashing nature	Self learning process	Mapping hash codes based on the similarity	reduce the quantization loss

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

Multiple Features Visual Hashing (MFVH): The interpretation and performance on visual contents in MFVH is significantly better. The learning performance is high in this type of hashing technique. There are numerous schemes are conducted hashing with multiple feature fusion considerations. Authors in [25] proposed a Sequential update for multi-view spectral hashing (SU-MVSH) to sequentially learn hash functions by solving the successive maximization of local variances. In this method, the multiple features are integrated with the help of divergence minimization from the view specific distance matrices. Authors in [26] present multi-view anchor graph hashing (MVAGH) by extending AGH to handle multiple image representations. This uses the fusion similarity matrix with the binary codes. The MFH a multiple feature hashing defines the learning process simultaneously by preserving the local structural information by consideration the every feature in the feature set. The types of SFVH and its techniques are described with its advantage in table 2.0

Table 2.0 Multiple Feature Visual Hashing techniques

Category	**MFVH**				
Techniques	**SU-MVSH**	**MVAGH**	**LCMH**	**MVLH**	**MVAH**
Abbreviations	Sequential update for multi-view spectral hashing	Multi-view anchor graph hashing	Multiple Feature Hashing	Multi-view latent hashing	Multi-view alignment hashing
Advantages	Sequentially learns hash functions	Supports multiple image representation	Preservers the local structural information's	Performs cross-modal similarity search	Non-negative factorization

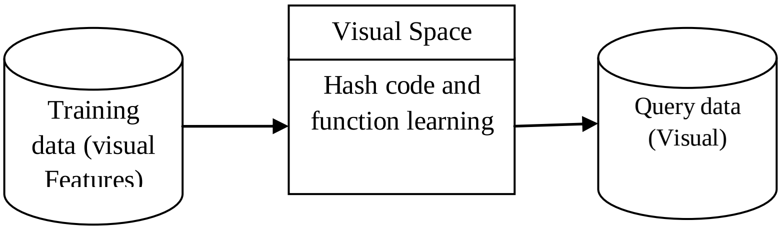

Fig 2.0 the structure of SFVH and MFVH schemes

The basic structure of SFVH and MFVH schemes are shown in fig 2.0. Authors in [27] expressed the hashing learning on multiple visual features within multi graph framework, where multiple visual features are integrated with appropriate ranks. Multimodal features in binary representation are incorporated in Multi-view latent hashing (MVLH) [28]. The multi view features are ranked and the weighted features are gathered. The weights for multiple feature fusion are learned according to the reconstruction error with each view. Multi-view alignment hashing (MVAH) [29] learns hash codes with regularized kernel non-negative matrix factorization. It considers both the hidden semantics and joint probability distribution of multiple visual features.

Multi-modal Hashing (MMH): Multi modal hashing technique supports visual and text based queries with semantic enhancements, but the MMH is not completely providing support for the CBIR systems. The learning features of the MMH are supports to both visual and text features. The basic structure of **MMH** scheme is shown in fig 3.0.

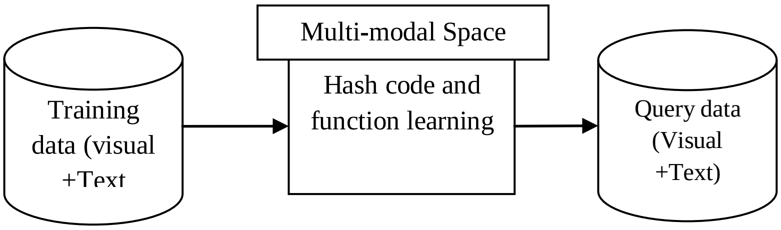

Fig 3.0 the structure of MMH scheme

Unsupervised Cross-Modal Hashing: Unsupervised Cross-Modal Hashing technique supports either visual or text based queries with limited semantic enhancements, but the UCMH is partly providing support for the CBIR systems. The learning features of the UCMH are supports to both visual and text features. The UCMH helps to locate the heterogeneous modalities into the hamming codes.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

Table 3.0 Unsupervised Cross-Modal Hashing techniques

Category	**UCMH**				
Techniques	**CVH**	**IMH**	**LCMH**	**LSSH**	**CMFH**
Abbreviations	cross-view hashing	inter-media hashing	Linear cross modal hashing	latent semantic sparse hashing	Collective matrix factorization hashing (
Advantages	Minimizing Hamming distances of similar samples and maximizing that of dissimilar samples	Preserve intra-similarity of each individual modality	scalable multi - media search across different modalities	Calculates the weights	Provides multiple modalities for single sample

Types of MMH and its techniques are described with its advantage in table 3.0 and the basic structure of **MMH** scheme is shown in fig 4.0.

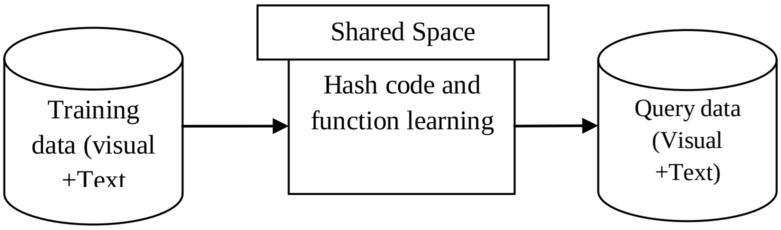

Fig 4.0 the structure of MMH scheme

Semantic-Assisted Visual Hashing (SAVH): Authors in [30] proposed a new and effective hashing framework named as SAVH. The idea is leveraging the associated texts of images to assist the visual hashing using unsupervised learning. This integrates the additional discriminative data into the collected visual codes. The main advantage of this technique is, it gives an effective offline learning system, which reduces the time complexity and perfectly suitable for the real application scenarios. However, the system doesn't contain the spatial information's and only suitable for the image retrieval. The basic structure of **SAVH** scheme is shown in fig 5.0.

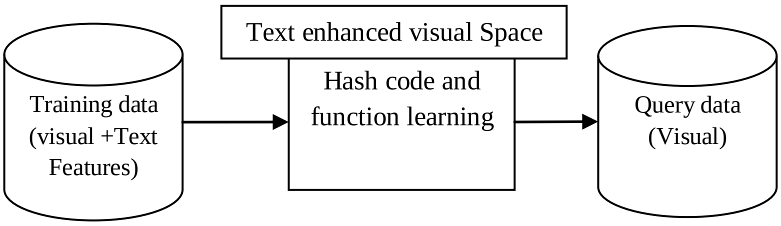

Fig 5.0 the structure of SAVH scheme

Table 4.0 Visual hashing techniques comparison table

Features ↓	**Methods→**	**SFVH**	**MFVH**	**MMH**	**UCMH**	**SAVH**
Query		Visual	Visual	Visual+Text	Visual or Text	Visual
Learning Feature		Visual	Visual	Visual+Text	Visual+Text	Visual+Text
Learning Space		Visual	Visual	Multimodal	Shared	Text enhanced
Semantic Enhancement		No	No	Yes	Limited	Yes
CBIR		Yes	Yes	No	Partly	Yes

The table 4.0 shows the overall comparison of the visual hashing techniques. The major drawback of SFVH and MFVH is that they only take the features from visual features into the consideration. Due to the semantic gap, image relations characterized by low-level visual feature cannot effectively describe rich image semantics, consequently making the hash codes less semantically meaningful.

DOI10.17148/IJARCCE.2017.6760

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

III. CONCLUSION

Similarity search 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. It is of great importance to many multimedia applications, such as content-based multimedia retrieval, classification and annotation. 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 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. This paper provides the overview of the available techniques and methods for effective image mining under CBOR concept. The further work can be developed with consideration of the above technique features.

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]. Yue, J., Li, Z., Liu, L., & Fu, Z. (2011). Content-based image retrieval using color and texture fused features. Mathematical and Computer Modelling, 54(3), 1121-1127.
[6]. Zhou XS, Huang TS., Relevance feedback in image retrieval: A comprehensive review. Multimedia Syst; 8: 536-544, 2003.
[7]. P.Chundi, U.Dayal, , M.Sayal,, M.Hsu, “A Document Clustering Method and System:” google patents,US20077181678, 2007.
[8]. D. Melas and S. Wilson, “Double Markov Random Fields and Bayesian Image Segmentation,” IEEE Trans. Signal Processing,vol. 50, no. 2, pp. 357-365, Feb. 2002.
[9]. I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathomas and P.Yianilos, “The Bayesian Image Retrieval System, PicHunter: Theory, Implementation and Psychophysical Experiments”, IEEE Trans. Image Processing, vol. 9, no. 1, pp. 20-37, 2000.
[10]. Buckley, C., Salton, G., Allan, J., & Singhal, A. (1995). Automatic query expansion using SMART: TREC 3. NIST special publication sp, 69-69.
[11]. Ishikawa Y, Subramanya R and Faloutsos C (1998). MindReader: query databases through multiple examples. Int'l Conf. on Very Large Data Bases (VLDB), NY.
[12]. Guo, Z., Zhu, S., Chi, Y., Zhang, Z., & Gong, Y. (2009, July). A latent topic model for linked documents. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (pp. 720-721). ACM.
[13]. Hofmann, Thomas. "Probabilistic latent semantic indexing." Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 1999.
[14]. Monay, Florent, and Daniel Gatica-Perez. "PLSA-based image auto-annotation: constraining the latent space." Proceedings of the 12th annual ACM international conference on Multimedia. ACM, 2004.
[15]. Monay, Florent, and Daniel Gatica-Perez. "PLSA-based image auto-annotation: constraining the latent space." Proceedings of the 12th annual ACM international conference on Multimedia. ACM, 2004.
[16]. Monay, Florent, and Daniel Gatica-Perez. "On image auto-annotation with latent space models." Proceedings of the eleventh ACM international conference on Multimedia. ACM, 2003.
[17]. Hofmann, Thomas. "Unsupervised learning by probabilistic latent semantic analysis." Machine learning 42.1 (2001): 177-196.
[18]. Barnard, Kobus, and David Forsyth. "Learning the semantics of words and pictures." Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 2. IEEE, 2001.
[19]. M. Raginsky and S. Lazebnik, “Locality-sensitive binary codes from shift-invariant kernels,” in Proc. Advances in Neural Information Processing Systems (NIPS), 2009, pp. 1509–1517.
[20]. Y. Weiss, A. Torralba, and R. Fergus, “Spectral hashing,” in Proc. Advances in Neural Information Processing Systems (NIPS), 2008, pp. 1753–1760
[21]. W. Liu, J. Wang, and S. fu Chang, “Hashing with graphs,” in Proc. Int. Conf. Machine Learning (ICML), 2011, pp. 1–8.
[22]. 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.
[23]. R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, “Liblinear: A library for large linear classification,” J. Mach. Learn.Res., vol. 9, pp. 1871–1874, 2008.
[24]. Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, “Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 12, pp. 2916–2929, 2013.
[25]. L. Gao, J. Song, F. Nie, Y. Yan, N. Sebe, and H. T. Shen, “Optimal graph learning with partial tags and multiple features for image and video annotation,” in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR), 2015, pp. 4371–4379.
[26]. S. Kim, Y. Kang, and S. Choi, “Sequential spectral learning to hash with multiple representations,” in Proc. European Conf. Computer Vision (ECCV), 2012, pp. 538–551.
[27]. S. Kim and S. Choi, “Multi-view anchor graph hashing,” in Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), 2013, pp. 3123–3127.
[28]. J. Cheng, C. Leng, P. Li, M. Wang, and H. Lu, “Semi-supervised multi-graph hashing for scalable similarity search,” Comput. Vis. Image Underst., vol. 124, no. 0, pp. 12–21, 2014.
[29]. X. Shen, F. Shen, Q.-S. Sun, and Y.-H. Yuan, “Multi-view latent hashing for efficient multimedia search,” in Proc. ACM Int. Conf. Multimedia (MM), 2015, pp. 831–834.
[30]. L. Liu, M. Yu, and L. Shao, “Multiview alignment hashing for efficient image search,” IEEE Trans Image Process., vol. 24, no. 3, pp. 956–966, 2015.
[31]. Zhu, L., Shen, J., Xie, L., & Cheng, Z. (2017). Unsupervised visual hashing with semantic assistant for content-based image retrieval. IEEE Transactions on Knowledge and Data Engineering, 29(2), 472-486.