A Hybrid Productive-Discriminative Based Hashing Technique

P. Sandhya¹, Dr. P. Kavitha², Dr. A. Pratapa Reddy³

¹PG Scholar, Dept of CSE, Ganapathy Engineering College, Warangal, TS, India, Email: sandhyap406@gmail.com.
²Associate Professor, Dept of CSE, Ganapathy Engineering College, Warangal, TS, India, Email: kavithapachika@gmail.com.
³Professor, Dept of CSE, Ganapathy Engineering College, Warangal, TS, India, Email: ptathapreddy54@gmail.com.

Abstract: Hashing methods have proven to be useful for a variety of tasks and have attracted extensive attention in recent years. Various hashing approaches have been proposed to capture similarities between textual, visual, and cross-media information. However, most of the existing works use a bag-of-words method to represent textual information. Since words with different forms may have similar meaning, semantic level text similarities cannot be well processed in these methods. To address these challenges, in this paper, we propose a novel method called semantic cross-media hashing (SCMH), which uses continuous word representations to capture the textual similarity at the semantic level and use a deep belief network (DBN) to construct the correlation between different modalities. To demonstrate the effectiveness of the proposed method, we evaluate the proposed method on three commonly used cross-media data sets used in this work. Experimental results show that the proposed method achieves significantly better performance than state-of-the-art approaches. Moreover, the efficiency of the proposed method is comparable to or better than that of some other hashing method.

Keywords: Canonical Correlation Analysis (CCA), SCMH, DBN, LSSH.

I. INTRODUCTION

With the rapid expansion of the World Wide Web, digital information has become much easier to access, modify, and duplicate. Hence, hashing based similarity calculation or approximate nearest neighbour searching methods have been proposed and received considerable attention in recent years. Various applications such as information retrieval, near duplicate detection, and data mining are performed by hashing based methods. Due to the rapid expansion of mobile networks and social media sites, information input through multiple channels has also attracted increasing attention. Images and videos are associated with tags and captions. According to research published on eMarketer, about 75 percent of the content posted by Facebook users contains photos. The relevant data from different modalities usually have semantic correlations. Therefore, it is desirable to support the retrieval of information through different modalities. For example, images can be used to find semantically relevant textual information. On the other side, images without (or with little) textual descriptions are highly needed to be retrieved with textual query. Along with the increasing requirements, in recent years, cross-media search tasks have received considerable attention [1], [2], [3], [4], [5],[6],[7]. Since each modality having different representation methods and correlational structures, a variety of methods studied the problem from the aspect of learning correlations between different modalities. Existing methods proposed to use Canonical Correlation Analysis (CCA) [8], manifolds learning [9], dual-wing harmoniums [10], deep auto encoder [11], and deep Boltzmann machine [12] to approach the task. Due to the efficiency of hashing-based methods, there also exists a rich line of work focusing the problem of mapping multi-modal high-dimensional data to low-dimensional hash codes, such as Latent semantic sparse hashing (LSSH) [13], discriminative coupled dictionary hashing (DCDH) [14], Cross-view Hashing (CVH) [15], and so on. Most of the existing works use a bag-of-words to model textual information. The semantic level similarities between words or documents are rarely considered. Let us consider the following examples:

S1. The company announces new operating system.
S2. The company releases new operating system.
S3. The company delays new operating system.

From these examples, we can observe that although only one word differs between the three sentences, sentence S3 should not be considered as the near duplicate sentence of S1 and sentence S2. The meaning expressed by S3 is much different with S1 and S2’s. Since existing methods are usually based on lexical level similarities, this kind of issue cannot be well addressed by these methods. In short text segments (e.g., microblogs, captions, and tags), the similarities between words are especially important for retrieval. For example: journey versus travel, coast versus shore. According to human-assigned similarity judgments [16], more than 90 percent of subjects thought that these pairs of words had similar meanings. Fig. 1 illustrates a set of images retrieved from Flickr using different queries. From these examples, we can see that images may express similar concepts, even though there is little overlap in terms of annotated tags. Since users rarely annotate a single image using multiple words with similar meaning, semantic level textual similarities should be incorporated into the cross media retrieval. Motivated by the success of continuous space word representations (also called word embeddings) in a variety of tasks.

Copyright © 2017 IJIT. All rights reserved.
P. Sandhya, Dr P. Kavitha, Dr A. Pratapa Reddy

In this work we propose to incorporate word embeddings to meet these challenges. Words in a text are embedded in a continuous space, which can be viewed as a Bag-of-Embedded-Words (BoEW). Since the number of words in a text is dynamic, in [17], we proposed a method to aggregate it into a fixed length Fisher Vector (FV), using a Fisher kernel framework [18]. However, the proposed method only focuses on textual information. Another challenge in this task is how to determine the correlation between multi-modal representations. Since we propose the use of a Fisher kernel framework to represent the textual information, we also use it to aggregate the SIFT descriptors [19] of images. Through the Fisher kernel framework, both textual and visual information is mapped to points in the gradient space of a Riemannian manifold. However, the relationships that exist between FVs of different modalities are usually highly non-linear. Hence, to construct the correlation between textual and visual modalities, we introduce a DBN based method to model the mapping function, which is used to convert abstract representations of different modalities from one to another. The main contributions of this work are summarized as follows.

1. We propose to incorporate continuous word representations to handle semantic textual similarities and adopted for cross-media retrieval.
2. Inspired by the advantages of DBN in handling highly non-linear relationships and noisy data, we introduce a novel DBN based method to construct the correlation between different modalities.
3. A variety of experiments on three cross-media commonly used benchmarks demonstrate the effectiveness of the proposed method. The experimental results show that the proposed method can significantly outperform the state-of-the-art methods.

II. EXISTING SYSTEM

Along with the increasing requirements, in recent years, cross-media search tasks have received considerable attention. Since, each modality having different representation methods and correlational structures, a variety of methods studied the problem from the aspect of learning correlations between different modalities. Existing methods proposed to use Canonical Correlation Analysis (CCA), manifolds learning, dual-wing harmoniums, deep autoencoder, and deep Boltzmann machine to approach the task. Due to the efficiency of hashing-based methods, there also exists a rich line of work focusing the problem of mapping multi-modal high-dimensional data to low-dimensional hash codes, such as Latent semantic sparse hashing (LSSH), discriminative coupled dictionary hashing (DCDH), Cross-view Hashing (CVH), and so on.

Disadvantages of Existing System:
1. Most of the existing works use a bag-of-words to model textual information. The semantic level similarities between words or documents are rarely considered.
2. Existing works focused only on textual information.
3. Another challenge in this task is how to determine the correlation between multi-modal representations.

III. PROPOSED SYSTEM

We propose a novel hashing method, called semantic cross media hashing (SCMH), to perform the near-duplicate detection and cross media retrieval task. We propose to use a set of word embeddings to represent textual information. Fisher kernel framework is incorporated to represent both textual and visual information with fixed length vectors. For mapping the Fisher vectors of different modalities, a deep belief network is proposed to perform the task. We evaluate the proposed method SCMH on three commonly used data sets. SCMH achieves better results than state-of-the-art methods with different the lengths of hash codes.

Advantages of Proposed System:
1. We introduce a novel DBN based method to construct the correlation between different modalities.
2. The proposed method can significantly outperform the state-of-the-art methods.

IV. IMPLEMENTATION

A. Admin
In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as upload images, view uploaded images, view the searching history, view all image ranking and view all users, search images and logout.

B. Search History
This is controlled by admin; the admin can view the search history details. If he clicks on search history button, it will show the list of searched user details with their tags such as user name, user searched for image name, time and date.

C. Rank of images
In user’s module, the admin can view the list of ranking images. If admin click on list of ranking images, then the server will give response with their tags image and rank of image.

D. Upload Images
In this module, the admin can upload n number of images with hash code and the Hash code will be generated based on

Fig1. System Architecture
the image name and contents. If the image is same and then the hash code will be similar for all corresponding images. Admin want to upload new image then he has enter some fields like image name, image color, image description, image type, image usage, browse the image file and upload. After uploading successfully he will get a response from the server. Initially new uploaded image rank is zero. After viewing that image rank will re-rank.

E. User

In this module, there are n numbers of users are present. User should register before doing some operations. And register user details are stored in user module. After registration successful he has to login by using authorized user name and password. Login successful he will do some operations like view my details, search images based on Hashing code of the group of images, request secrete key and logout. The user click on my details link then the server will give response to the user with all details such as user name, phone no, address, e mail ID and location. An end user can search the images based on Images based on keyword or hash code and gets the details like image name, image color, image usage and image type. And server will give response to the user, then that image rank will be increased.

V. CONCLUSION

In this work, we propose a novel hashing method, SCMH, to perform the near-duplicate detection and cross media retrieval task. We propose to use a set of word embeddings to represent textual information. Fisher kernel framework is incorporated to represent both textual and visual information with fixed length vectors. For mapping the Fisher vectors of different modalities, a deep belief network is proposed to perform the task. We evaluate the proposed method SCMH on three commonly used data sets. SCMH achieves better results than state-of-the-art methods with different the lengths of hash codes. In NUS-WIDE data set, the relative improvements of SCMH over LSSH, which achieves the best results in these datasets, are 10.0 and 18.5 percent on the Text! Image and Image! Text tasks respectively. Experimental results demonstrate the effectiveness of the proposed method on the cross-media retrieval task.

VI. REFERENCES