

Improved Entropy and Key Point Feature Extractor for Effective Multimedia Information Retrieval

Santhosh J¹, Shafna M²

Assistant Professor, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu¹

M. Phil Scholar, Department of Computer Science, Sree Narayana Guru College, Coimbatore, Tamil Nadu²

Abstract: Information retrieval became very imperative task in current scenario, everyone have huge volume of data in the distributed environment, and retrieving those contents need extra care and attention. In this paper, a new semi-supervised lexicon based optimized information retrieval methods are proposed. Unlike the existing classification method, the proposed system decreases the need of training process and improves the information retrieval efficiency. In this proposal, the semi-supervised bag of words representation is used for information retrieval. The improved entropy and lexicon method effectively performs information retrieval process on different type of text, image and video datasets. This drastically reduces the training time and information retrieval time in the huge dataset environment. The implementation and experiments used multi-class image, video datasets and text datasets with lexicon entropy. the system significantly improve average precision, accuracy and storage efficiency by deploying high configured information retrieval method.

Keywords: Information retrieval, CBIR, classification, Web extraction, Image dataset, Dictionary Learning, Entropy Optimization, Image Retrieval, Time-series Retrieval.

I. INTRODUCTION

Information retrieval and classification are abundant in day-to-day life. Information retrieval and classification has a wide area of application, from surveillance and traffic monitoring to medical imaging. Humans are surprisingly good at information retrieval and image retrieval. Developing a machine vision system in a real-time scenario to search images in a complicated environment is a difficult task in terms of computation. Attempts have been made by researchers to improve the capability and robustness of machine vision system. The recognition and categorization of images is one of the most challenging problems in computer vision. This is due to high variability in appearance between images of the same class and the additional variability between instances of the same image with respect to viewing conditions. An object classification system must be able to detect the presence or absence of an image.

The process of retrieving information from the web is called information retrieval (IR). The Information retrieval is the activity of obtaining pertinent information to an information need from a collection of information resources. Searches can be based on metadata or full-text or other content-based indexing. IR is used to digital libraries, information filtering, and multimedia search and most visible in web search engines.

models, algebraic models and probabilistic models. Set theoretic models have four main types which are Boolean, case-based reasoning (CBR), fuzzy set and extended Boolean models. Algebraic models consist of vector space model, generalized vector space model, latent semantic indexing and neural network model. Probabilistic model covers probabilistic, inference network, and brief network model.

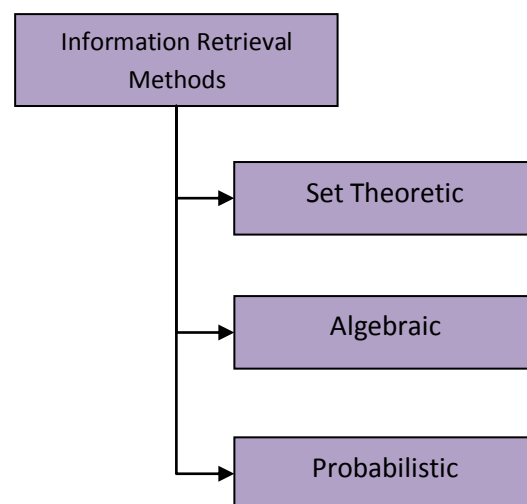


Fig 1.0 Information retrieval methods

Information Retrieval Models:

The classical information retrieval models are shown in Figure 1.3. It can be mainly divided into set theoretic

In the set theoretical model, documents and queries are represented as a set of index terms. Similarly, in algebraic



model documents and queries are represented as vectors in a t -dimensional space. In probabilistic model documents and queries representations are based on probability theory.

OBJECTIVES OF THE THESIS:

The research work aims to enhance semantic search performance in an information retrieval system. The goal is

- To develop generic semantic search technique for requester without domain knowledge and improve search performances.
- To develop specific semantic search techniques for requester with domain knowledge to find information precisely and quickly.
- The search technique decreases searching time. To develop QoS based ranking methodology to enhance semantic search performance of both generic and specific searches.

II. PROBLEM DEFINITION

Information retrieval that handles the bindings of semantic search is facing problem due to low precision, recall and long time. Recent researches have been focused on this issue to improve web search performance. In order to combat these issues, an efficient semantic search system with analytical proof is essential to compare the proposed work to be more effective than the existing approach.

III. LITERATURE REVIEW

Multimedia search and technologies refer to the semantic search and technologies designed in the purpose of retrieving audios, videos (Tian et al 2012), and images (Wang et al 2012 & Tian et al 2010). Multimedia searching has become an important research field, particularly for understanding E-commerce and entertainment Web search. However, there is a major need for new multimedia Web search techniques and systems. People continually raise the bar of expectation for sophisticated Web search engines.

Due to the rapidly increasing usage of digital multimedia data, a single piece of information can often be richly conveyed using multiple correlated media, thus enabling users to concurrently receive information from multiple sources (Tjondronegoro & Spink 2008). Text is currently one of the most intuitive methods for searching images since most of the available search engines still rely on text, and users usually think of topics, keyword, or other high-level concepts in the form of text to write queries. Current Web search engines, such as Google Image (<http://images.google.com>), MSN image (<http://search.msn.com/images>), and Yahoo! Image (<http://images.search.yahoo.com>), have been fixed at the point of text analysis relevant to the multimedia material. However, the major limitation of this approach is the tedious and ineffective nature of manually annotating

every temporal segments of a video or every region of interests in an image.

Moreover, the process of automatic annotation by mapping low-level features into high-level semantic concepts is generally difficult as it needs machine learning and interpretations. For example, systems can use modelled domain knowledge to make sense and interpret the semantic meaning of video data by observing visual, audio and text features (Duan et al 2003). Yoshitaka and Ichikawa (1999) surveyed query methods that utilize low-level features from image such as shape, spatial relation, color and texture; and video, including object motion, spatial-temporal relations. They have compared these features-based queries with semantic-based and knowledge assisted retrievals including „query-by-subject (i.e. keyword) and „query-by-subject/object (i.e. derivation knowledge).

The emerging applications and services from these types of retrievals include: finding images that are visually similar to a chosen picture or sketch; summarizing videos with thumbnails of key frames; finding video clips of a particular event, story, or somebody; and making a two-minute skim of an hour-long program (Chang et al 1998). Kherfi et al (2004) provided a comprehensive discussion on the issues, techniques and systems of image retrieval from the Web and reviewed some prototypes. Lu et al (2005) presented Content-based audio retrieval (CBAR) leads to a more accurate classification than what can be achieved by content based image retrieval systems. For example, audio stream can be classified into music, speech, and noise, as well as some semantic details like cheering, applause, and laughter. This raises a challenge for bridging the gaps between low-level features and high-level semantic. Labra et al (2010) applied semantic web technologies and collaborative tagging to multimedia web information systems.

The proposed system uses a search engine that combines both kinds of meta-data to locate the desired content of text, images and video. It will also provide browsing capabilities through the ontology concepts and the developed tags. Richang et al (2013) presented a novel concept named Mediapedia which aims to construct multimedia encyclopedia by mining web knowledge. The Mediapedia distinguishes itself from traditional encyclopedia in its multimedia presentations and demonstration, complete automatic production, dynamic update and the flexible framework where each module is extensible to potential applications.

This approach employed the affinity propagation algorithm in producing the exemplars from image pool, while using latent semantic analysis to associate exemplars to Wiki pages and utilizing document lattice model to perform Wiki pages summarization and assembled them for multimedia encyclopedia. This study can be deemed as an attempt at constructing Mediapedia by leveraging on web information the implementation results not as good as expected. This may be aroused by the assumption that the distribution of images from Flickr can automatically make



a tradeoff between several queries. Conversely, this is not the truth for all the concepts. Enhancement can be created by taking into account the tags in producing the exemplar, and leveraging the images embedded in Wikipedia to facilitate better association and so on. An alternative approach to tackle the problem is to start this Mediapedia from Wikipedia, by identifying different senses of the concept by Wikipedia first and then associating them with images /audios.

Summary:

From this survey identify some common drawbacks of traditional information retrieval models and semantic search technologies shown as below;

Identify user's perceptions: A user perception plays an important role in semantic search engines. The existing semantic search engines are not able to satisfy user's requirements due to the difference between each person's subjective perceptions regarding the objective world.

Ignorance of evolving knowledge structure: Knowledge structures in the semantic search engines cannot be frequently updated to suit the change of user's requirements and the change of external environment.

Low precision and high recall: Some search engine cannot show their significant performance in improving precision and low recall.

Lack of semantics and lack of evaluations: Since many semantic search methodologies are only in the phase of conceptual model development, many of them have not been tested via the reasonable number of experiments.

Ambiguity/Polysemy: A polysemous word has more than one meaning. When searching for documents with a word such as "play", related to a theatre piece, a search engine can return unrelated results such as a set of games for children.

Lack of synonym relations: Words are synonymous if they have the same meaning. Words "irritated" and "annoyed" are very closely related; however, when searching for one of these words, found items will hardly contain the other word.

Lack of consensus: The lack of consensus in the use of tags, especially as granularity is concerned, makes a traditional tagging system quite inefficient. To describe a particular item, different users may consider terms at different levels of generality/ specificity. For example, a user can tag a photograph as "bird", and another user can tag the same photo as "eagle".

IV. PROPOSED SYSTEM

In this paper, we propose a new semi-supervised bag of words representation for fast textual and multimedia information retrieval. The improved entropy and lexicon method effectively performs information retrieval process on different type of text, image and video datasets. This drastically reduces the training time and information retrieval time in the huge dataset environment. For effective feature extraction KIFT algorithm is proposed

(Key Invariant Feature transform), which reduces the feature extraction and matching time in the multimedia information retrieval.

The proposed system has the following steps.

1) Feature extraction: Feature extraction, in which multiple features, such as KIFT descriptors are extracted from each image. That way, the feature space is formed where each image is represented as a set of features.

2) Lexicon learning: Lexicon learning, in which the extracted features are used to learn a Lexicon of representative features. These features are called as words or codeword's.

3) Feature quantization and encoding: feature quantization and encoding, in which each feature is represented using a codeword from the learned lexicon and a histogram is extracted for each image. That way, the histogram space is formed where each image is represented by a constant dimensionality histogram vector. The proposed information retrieval architecture uses the above three steps along with the descriptor extraction for fast image and content retrieval.

KIFT Algorithm:

Input: Video (V), frame (F)

Output: Key point descriptor (K), matched frames and combined video

Step 1: get the frames and extract features of the image.

I=Image.getBytes(Gf.FI)

Notations:

Gf=Getframes

FI=Frame Index

I-index value

Step 2: Then extracts KIFT features of the key frames.

■ Scale-space extrema detection

Definition:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Where,

■ Keypoint detection

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

Keypoints are detected using scale-space extrema in difference-of-Gaussian function D

D definition:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

Step 3: store the frame in the database.

$$= L(x, y, k\sigma) - L(x, y, \sigma)$$

Save(imageid i,image im)

Save descriptors (imageid i,image im,local_descriptor d)

Step 4: perform the step 2 for the test image

Step 5: Match the image with already stored image by KIFT method.

The above algorithm shows the basic steps of KIFT, which performs with the test samples. The implementation and



experiments used multi-class image, video datasets and text datasets with lexicon entropy. The proposed system significantly improve average precision, accuracy and storage efficiency by deploying high configured information retrieval method.

V. EXPERIMENTAL RESULTS

We evaluate the proposed method using extensive experiments on two multi-class image datasets, the 23-Scene dataset and the CBIR dataset. Then, we also provide an experimental evaluation of the method using one text dataset, and one multi-class image dataset, The 23-Scene dataset contains 13 different scene categories: Animals, flowers, celebrity, outdoor, store, industrial, Building, vehicles, landscape images etc., the total number of images is 100 and each category has 10 to 20 images. From each image we densely extract KIFT descriptors.

Performance evaluation of proposed semi-supervised KIFT feature extraction for information retrieval is compared with the existing feature extraction algorithms.

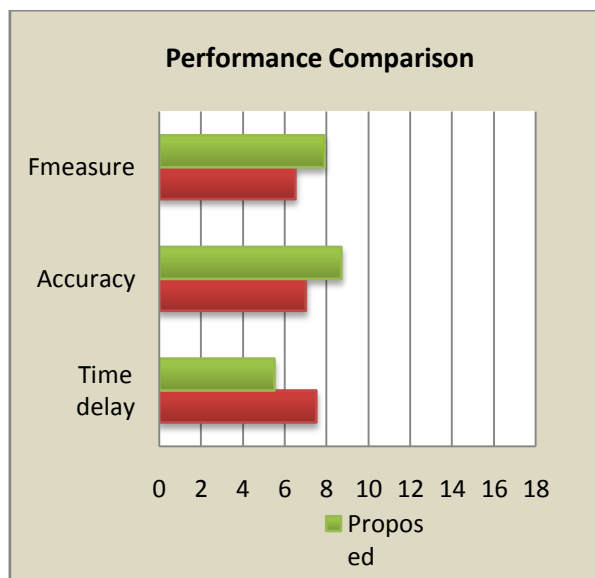


Fig 2.0 Efficiency analysis of proposed system

From the fig 2.0 it shows the performance measure based on the efficiency and the proposed approach KIFT yields maximum efficiency by considering several parameters such as time, accuracy, F-measures.

CONCLUSION

In this paper, we proposed semi-supervised dictionary learning method with improved entropy and lexicon methods, which improves the information retrieval efficiency and decreases the time at the time of extraction. We experimented the outcome of the proposed method to improve the retrieval performance using Different text and image datasets, a collection of time-series datasets, a text dataset and a video dataset. First, for

a given lexicon size, it can improve the information retrieval efficiency.

REFERENCES

- [1] A. Andoni and P. Indyk, "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," in 47th Annual IEEE Symposium on Foundations of Computer Science, 2006, pp. 459–468.
- [2] M. M. Baig, H. Gholamhosseini, and M. J. Connolly, "A comprehensive survey of wearable and wireless ECG monitoring systems for older adults," *Medical & Biological Engineering & Computing*, vol. 51, no. 5, pp. 485–495, 2013.
- [3] W. Bian and D. Tao. (2009) The COREL database for content based image retrieval. [Online]. Available: <https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval>
- [4] Y.-L. Boureau, F. Bach, Y. LeCun, and J. Ponce, "Learning mid-level features for recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2559–2566.
- [5] F. K.-P. Chan, A. W.-C. Fu, and C. Yu, "Haar wavelets for efficient similarity search of time-series: with and without time warping," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 3, pp. 686–705, 2003.
- [6] D. Chatzakou, N. Passalis, and A. Vakali, "Multispot: Spotting sentiments with semantic aware multilevel cascaded analysis," in *Big Data Analytics and Knowledge Discovery*, 2015, pp. 337–350.
- [7] 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.
- [8] K. Eguchi and V. Lavrenko, "Sentiment retrieval using generative models," in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2006, pp. 345–354.
- [9] D. Gorisse, M. Cord, and F. Precioso, "Locality-sensitive hashing for chi2 distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 402–409, 2012.
- [10] X. He, D. Cai, and J. Han, "Learning a maximum margin subspace for image retrieval," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 2, pp. 189–201, 2008.