RESEARCH ARTICLE OPEN ACCESS

Face Recognition by Using Multi-Probe LSH Indexing

Ms. Tanuja Kiran Rajput [1], Prof. Dr. B. D. Phulpagar [2]
Department of Computer Science and Engineering
PESMCOE
Pune-India

ABSTRACT

WEB data including images are growing rapidly. For example, the photo sharing website Flickr, Facebook etc has over 5 billion images. There is an emerging need to retrieve relevant content from such massive databases. So our task is to recognize a face with a proper name by using a high dimensional Multi-probe locality sensitive hashing.

Locality-sensitive hashing (LSH) is a basic primitive in several large-scale data processing applications, including nearest neighbour search, de-duplication, clustering, etc. Locality sensitive hashing (LSH) has been used extensively as a basis for many data retrieval applications. LSH is based on hashing the data points to a number of buckets such that similar points are more likely to map to the same buckets. Recently, locality sensitive hashing (LSH) and its variations have been proposed as indexing techniques for approximate similarity search. A significant drawback of these approaches is the requirement for a large number of hash tables in order to achieve good search quality. This paper proposes an indexing scheme called Multi-Probe LSH that overcomes this drawback. Multi-probe LSH is built on the well known LSH technique, but it intelligently probes multiple buckets that are likely to contain query results in a hash table.

Our algorithm allows one to reduce the number of hash tables, and is hence memory efficient, while achieving high accuracy.

Keywords :- Face annotation, indexing, label refinement, Multi-Probe LSH.

I. INTRODUCTION

A large collection of photo images which are usually untagged are present on Internet. This has become a great challenge for end users to browse and search. This is time consuming and costly for large photo collections. Instead of tagging images manually, we are tagging it automatically with SBFA.

This plays an important role in many real world multimedia information and knowledge management systems. Recently there is surge of research interest in mining weakly labeled web facial images for search based face annotation. The main challenge in this paper is how to remove the duplicate names from given image. A duplicate name is practical issue in real world applications. We assume first that each name belongs to a unique single person. We can learn the similarity between two different names according to the web pages to determine how the different names belong to the same person.

This paper proposes a new indexing scheme called Multi-Probe LSH that overcomes this drawback. Multi-probe LSH is built on the well known LSH technique, but it intelligently probes multiple buckets that are likely to contain query results in a hash table. Our method is inspired by and improves upon recent theoretical work on Basic LSH designed to reduce the space requirement. Our evaluation shows that the multi-probe LSH method substantially improves upon previously proposed methods in both space and time efficiency. To achieve the same search quality, Multi-Probe LSH require less time than Basic LSH method while reducing the number of hash tables by an order of magnitude. Multi-Probe LSH uses less query time and fewer number of hash tables.

II. RELATED WORK

We studied several groups of research work.

For facial feature representation, C. Siagian et al. extract the GIST texture features [1] to represent the extracted face. As a result, each face can be represented by a d-dimensional feature vector.

Z. Li et al. [2] proposes a saliency and Gist features for target detection in satellite images. This study shows that the proposed target search method can reliably and effectively detect highly variable target objects in a large dataset of images.

W. Dong works to index the extracted features of the faces by applying some efficient high-dimensional indexing technique to facilitate the task of similar face retrieval in the subsequent

step. In our approach, we adopt the locality sensitive hashing (LSH) [3] a very popular and effective High-dimensional indexing technique.

X. Gu and Q. Lv works to index the extracted features of the faces by applying some efficient high-dimensional indexing technique to facilitate the task of similar kind of face retrieval in the subsequent step. This method can accurately predict the average search quality and latency on a small sample dataset. They experimented with three different datasets including audio, images and 3D shapes to evaluate their methods. An improved method of locality sensitive hashing for indexing large scale and high-dimensional features. In this approach, they adopt the locality sensitive hashing (LSH) [4], [5] a very popular and effective High-dimensional indexing technique.

And the work in [6] by D. Wang et al. applied an unsupervised label learning which is different from the above previous works in two main aspects. First of all, their work aims to solve the general content-based face annotation problem using the search-based paradigm, where facial images are directly used as query images and the task is to return the corresponding names of the query images.

III. SYSTEM ARCHITECTURE

A. Architectural Diagram:

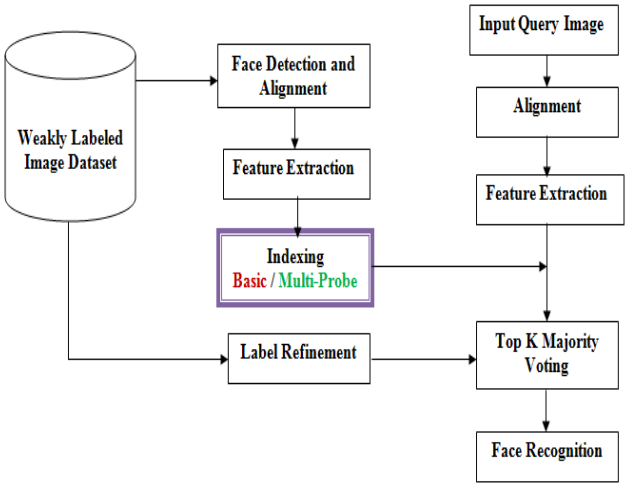

Fig. 1: Architectural Diagram for Face Recognition

B. Processing Steps for the System:

1. Input Data Collection:

We collected a human name list consisting of popular actor and actress names from the http://wlfdb.stevenhoi.com/ website.

As the output of this crawling process, we shall obtain a collection of facial images, each of them is associated with some human names automatically.

Our main input which is given by user is a single Image as a query.

2. Facial Feature Extraction:

There are many techniques for facial feature extraction. Such as Gabor, LBP, EBGM, BIC, GIST etc. For facial feature representation, we extract the GIST texture features to represent the extracted faces. As a result, each face can be represented by a d-dimensional feature vector.

3. Facial Feature Indexing:

Also there are many techniques for indexing. One is Basic LSH and Multi-Probe. Among these two, Multi-Probe LSH is better than Basic LSH, so we adopt the multi-probe locality sensitive hashing (LSH) a very popular and effective high-dimensional indexing technique for similar face retrieval.

4. Refine Weakly Labeled Data:

We apply the ULR Unsupervised Label Refinement method to refine the raw weak labels.

5. Top K Image Majority Voting:

All these steps are applied on single query image. We search for the query facial image to retrieve the top K similar images and use their associated names for voting toward auto annotation.

6. Output: Face Annotation:

After applying all the algorithms which are listed above, the output will be a Query image is properly tagged by proper name.

IV. EXPERIMENTAL RESULTS

The performance of a similarity search system can be measured in three aspects: search quality and search speed. Search quality is measured by recall.

$$\textbf{Recall} = |\, A(q) \cap I(q) \,| \,/\, |\, I(q) \,|$$

Where, q: query object,
I (q): set of ideal answers,
A (q): actual answers.

In the ideal case, the recall score is 1.0, which means all the k nearest neighbours are returned.

Search speed is measured by query time, which is the time spent to answer a query.

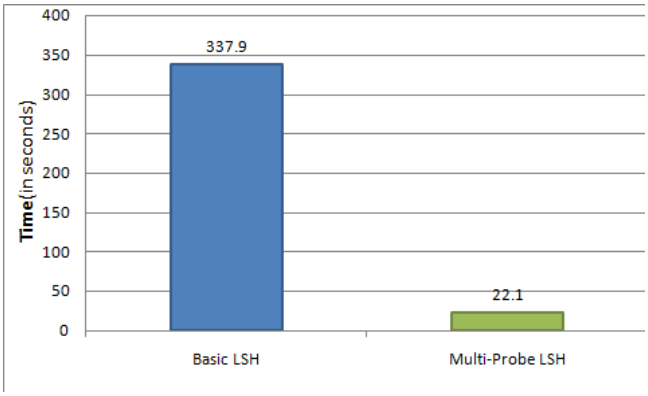

Fig. 2: Average Running Time between Basic LSH and Multi-Probe LSH.

In the ideal case, the recall score is 1.0, which means all the k nearest neighbours are returned.

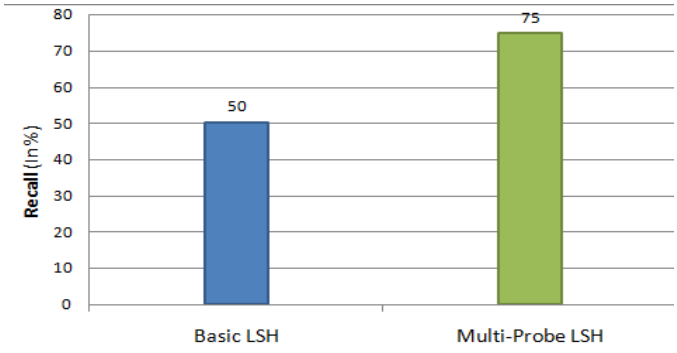

Fig. 3: Recall Graph between Basic LSH and Multi-Probe LSH in Percentage.

V. CONCLUSIONS

The Basic LSH is having some drawbacks, like many number of hash tables are formed hence memory space requirement is high. And the time required for indexing is less than Basic LSH. The results in Figure 1 shows that the Multi-Probe LSH method is substantially more time efficient than the Basic LSH approach. Also the Recall is 25% more i.e. Multi-Probe LSH is more accurate in terms of Recall than Basic LSH.

ACKNOWLEDGMENT

The Every orientation work has an imprint of many people and it becomes the duty of author to express deep gratitude for the same.

I take this opportunity to express my deep sense of gratitude towards my esteemed guide Prof. (Dr). B. D. Phulpagar for giving me this splendid opportunity to select and present this seminar.

I thank Prof. (Ms). S. A. Itkar, Head, Department of Computer Engineering, for opening the doors of department towards realization of seminar report, all the staff members, for their indispensable support, priceless suggestions and for most valuable time lent as and when required.

I am grateful for the cooperation and constant encouragement from my honourable Prof. (Ms). D. V. Gore. Her regular suggestion made my work easy and proficient. With all respect and gratitude, I would like to thank all the people, who have helped me directly or indirectly.

REFERENCES

[1] C. Siagian and L. Itti, "Rapid Biologically Inspired Scene Classification Using Features Shared with Visual Attention," IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 29, no. 2, pp. 300-312, Feb. 2007.

[2] Z. Li and L. Itti, "Saliency and Gist Features for Target Detection in Satellite Images," Proc IEEE Transactions on Image Processing, Vol. 20, No. 7, July 2011.

[3] W. Dong, Z. Wang, W. Josephson, M. Charikar, and K. Li, "Modeling LSH for Performance Tuning," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM), pp. 669-678, 2008.

[4] X. Gu, Y. Zhang, L. Zhang, D. Zhang, J. Li, "An Improved Method of Locality Sensitive Hashing for Indexing Large-Scale and High-Dimensional Features," Proc. ACM Conf. Signal Processing, Vol. 93 Issue 25, pp 2244-2255, July, 2012.

[5] Q. Lv, W. Josephson, Z. Wang, M. Charikar, K. Li, "Multi-Probe LSH: Efficient Indexing for High Dimensional Similarity Search," ACM 978-1-59593-649-3 July 2009.

[6] D. Wang, Steven C.H. Hoi, Ying He, and Jianke Zhu, "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation, "IEEE Transactions on

Knowledge and Data Engineering, Vol. 26, No. 1, January 2014.

[7] https://en.wikipedia.org/wiki/Locality-sensitive_hashing

[8] Rafael C. Gonzalez, Richard E. Woods, "Digital Inage Processing", 3rd Edition, Pearson, 2011.