

An Integrated Methods for Refined Facial-Image Retrieval using Sift and Click-Through Data

C. Balasubramanian, J. Raja Sekar



Abstract: In the modern era, the emergence of social networking sites paved the way to the people to upload lot of images online. Social sites like Instagram and Flickr allow users to add semantic information to the images in the form of tags. Often these tags are the firsthand semantic data for retrieving the images from the Internet. When a user searches for images in the web, the images with tags relevant to user query are retrieved. Most of the time, these semantic data are not relevant to the content of the image and hence the user gets irrelevant images in contrast to their intended search. This is more common in facial image search. In this paper, we propose an integrated approach to refine the relevance of retrieved facial images. Using Scale Invariant Feature Extraction (SIFT) the facial features of semantically most relevant image and click-through data of all the retrieved images are used to rank and present a meaningful search result. Along with facial features and click through information, the co-occurrence of related tags is also considered. Also we propose the construction of inverted index structure to improve the search performance.

Keywords: facial images; inverted index; SIFT, click-through data, co-occurrence.

I. INTRODUCTION

We are living in an era of images. Every person is having a digital camera in the form of a mobile phone aka smart phone. Taking pictures is increasing and becoming inevitable. The images taken by the people are uploaded in social networks abundantly. These images often carry lot of information in the form of tags and labels. Most of the commercial search engines obtain user query as input and provide the search image results mainly based on the relevance to the query with tags and labels. This suffers either from irrelevance of the query or inappropriateness of the information associated with the images.

The work can be defined as:

- We proposed a refined retrieval and re ranking of facial images based on users' search query. We have constructed an inverted index structure to speed up the image retrieval. We have methodically blended SIFT, semantic and views features of images to improve result efficiency.
- The semantic features of images are extracted and a

co-occurrence matrix of tags is constructed. This matrix is used to re rank the images

- The SIFT features from each retrieved image is extracted and it is compared with the selected reference query image
- For effective re ranking, we combine SIFT and semantic features with views information to calculate relevance score for all images

This paper is organized in the following way. Section 2 consists of discussions about the related works in image retrieval and re ranking. Section 3 consists of proposed method and its features. Section 4 discusses about the training part of the system while Section 5 discusses about the testing part. In Section 6, the results or implementation and performance evaluation is presented.

II. RELATED WORK

Social media websites such as Flickr and Instagram enables its users to add semantic information such as tags to the images they upload

A. Vote based process

Images are retrieved based on tags and their re ranking [1] [2] and [3] is done by relevance score by considering the votes. Keeping this as a basis, Lee and N eve [6] proposed a new method of calculating relevance score by neighbor voting.

B. Ranking of relevant images

Gao et al. [7] proposed a hyper graph learning method, which analyses the relevance score of all the query results.[8] [11] proposed Support Vector Machine(SVM) classifier for the purpose. [8] proposed ranking plan based on both visual information and semantic information while re ranking the images.

III. PROPOSED SYSTEM

The system is implemented in two phases; training and testing. In training phase, SIFT features, semantic features and view features of images are extracted. Semantic feature is nothing but the co-occurrence of query and tags of an image. The training phase is done in two parts. In the first part image re ranking is carried out using the co-occurrence terms for the particular query. In the second part, the re ranking of image is carried out by calculating a score for measuring the relevance between the query and the image.

Manuscript published on 30 December 2019.

* Correspondence Author (s)

C. Balasubramanian*, Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India. Email: bala.vnr@gmail.com

Dr. J. Raja Sekar, Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

The system design can be conceptually depicted in Figure.1

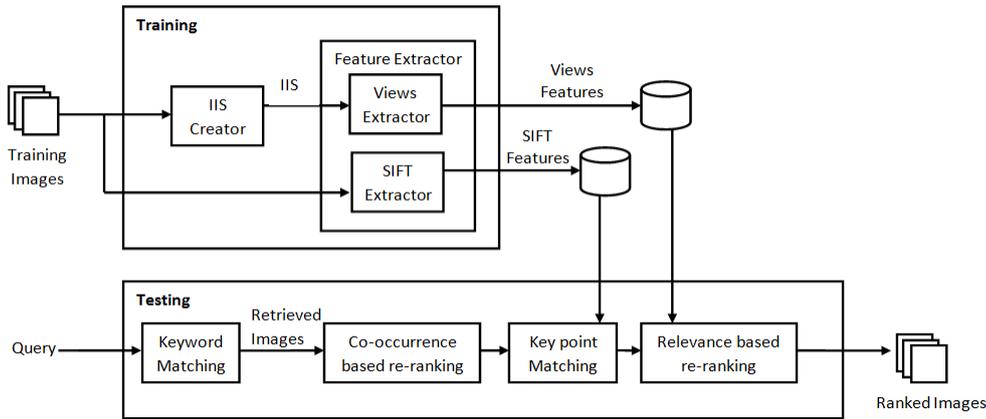


Fig 1. Proposed System Design

IV. TRAINING PHASE

A. Inverted Index Structure Construction

In order to retrieve the images faster an inverted index structure (IIS) is constructed. The image repository may contain enormous amount of pictures and images uploaded by many of the users. The basic element of the IIS is tags. The image repository may contain unique images as well as unique tags. Each image may be tagged with many tags. At the same time various images may be tagged with a common tag. Let the tag set $\Gamma = \{\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_G\}$ consists of unique tags in the image repository. The IIS of the images can be defined as $I = \{I_1, I_2, I_3, \dots, I_G\}$ where I_i is a collection of images having the tag Γ_i .

B. Feature Extraction

In this work, three features namely, SIFT, semantic and views are considered for representing the image data set.

a. Extracting the SIFT features

Scale Invariant Feature Transform algorithm is one of the feature detection algorithms generally used to extract and describe features in images. This technique is helpful for finding reliable matching among various images [9] [10]. The extracted features are not varying with respect to any image transforms such as scaling, rotation etc. SIFT features can be described as [15],

- DOG – Difference of Gaussian: This method helps to identify key points in the image which are invariant to transforms

$$H(a, b, \sigma) = DG(a, b, \sigma) \times \text{Img}(a, b) \quad (1)$$

Where $\text{Img}(a, b)$ is the input image

Scale space extrema detection is used in DOG. $DG(a, b, \sigma)$ is calculated from the variations of different scales which is differentiated by a continual factor c :

$$DG(a, b, \sigma) = (DG(a, b, c\sigma) - DG(a, b, \sigma)) \times \text{Img}(a, b) \quad (2)$$

- Keypoint Localization: Adjacent pixels are compared and the pixels with lower contrast are rejected

b. Extracting the views

Each image is associated with click data. That is the number of times the image is being clicked by the user. It is assumed that while searching the image the users normally click on the image which they intend to search. Hence this will increase the relevance factor for the image. This concludes that more the clicks, more popular the image is. Let C_i is the number of clicks on image i , the normalized clicks is calculated by

$$C_i = ((c_i - c_{min}) / (c_{max} - c_{min}) (R_{max} - R_{min})) + R_{min} \quad (3)$$

Where C_{max} denotes maximum clicks in the dataset, C_{min} denotes minimum clicks in the dataset, R_{max} and R_{min} denotes maximum and minimum range in the dataset

c. Extracting the semantic features

Extracting the semantic features involves in calculating co-occurrence words. Co-occurrence is nothing but the order of appearance of words. The co-occurrence of $CO(q) = \{co_1, co_2, co_3, \dots, co_L\}$ where L is the total number of words co-occurring for the input query q . For a query q , set of images will be retrieved with tag as q . Retrieved images will have other tags also. For each tag t , the total number of images that are tagged with the tag t and query q provides the correlative frequency. The weight for co-occurrence word is calculated by

$$CO_i = \exp\{(-(\max(\log(N(q)), \log(N(co_i))) - \log(N(q, co_i))) / (\log(n) - \min(\log(N(q)), \log(N(co_i))))\} \quad (4)$$

where n is the n represents the number of images and $N(q, co_i)$ denotes the correlative frequency between q and i th co-occurrence word.

V. TESTING PHASE

A. Key word Matching

The first step in testing is retrieving the image set as a result of user query. Given the input query q, the image set S consisting of images tagged with q. This image set may consist of irrelevant images which are tagged with q

B. Re-ranking based on Co-occurrence

The semantic confirmation matrix SC is constructed and using the matrix, the semantic relevance is calculated for the image tagged with query q and the query q itself. The mean co-occurrence similarity between ith image and q is calculated as

$$sci = (\sum tag(tm) * COm) / \sum tag(tm) \tag{5}$$

$$where\ tag(tm) = \begin{cases} 1 & \text{if image is tagged with tm,} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

C. Key point matching

For first query image, the SIFT and facial features are extracted. This is then compared with all the retrieved images from the image dataset [14].

D. Relevance based re-ranking

The set R consists of relevance score r for each image which is calculated by

$$r_i = sci + C_i \tag{7}$$

Then R is sorted based on highest relevance score.

VI. RESULTS AND DISCUSSIONS

A. Dataset

Around 10,000 facial images along with their associated information are collected from the website www.flickr.com [8]

B. Results

The dataset consisted of different facial images. Let us assume that suppose the query given is ‘captain’, the images that are having tag as ‘captain’ is retrieved. The retrieved result is shown in the figure 2(a). The Ci for each image is calculated and the images are sorted based on Ci ie., from the highest value to the lowest. The sorted result is shown in the figure.2(b).

The image with highest Ci is considered as the query image and whose SIFT features are extracted. Similarly the SIFT features of all the retrieved images are extracted and compared with the features of the query image. After selecting and re-ranking images, we get more relevant images on top.



Fig 2(a). Retrieved images



Fig 2(b). Images sorted based on click data

The obtained images are shown in the figure 3(a). For these images, the relevance score is calculated using the views features. The image set is then sorted based on the relevance score from highest to the lowest. The obtained results are shown in the figure 3(b).



Fig. 3. (a) Re-ranking based on keypoint matching



Fig. 3. (b) Re-ranked images based on relevance score

C. Performance

The quality of the images retrieved are evaluated and compared with other techniques such as views based re-ranking and co-occurrence based re-ranking. Our work has a significant improvement both in relevancy of the intended search and better results.



The visualization of the quality is depicted in the figure 4.

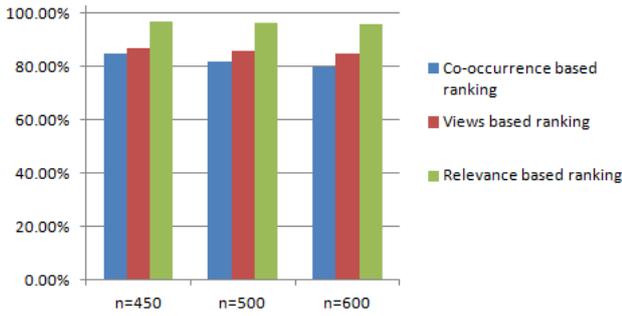


Fig.4. Relevancy of results

VII. CONCLUSION

In this work, we have proposed an integrated approach for refined facial-image retrieval. In this approach, we have considered the SIFT features, semantic features and click-through data of a image. The results demonstrate that our approach is efficient and time saving.

REFERENCES

1. Liu, X. Hua, I. Yang, M. Wang, and H. Zhang, "Tag Ranking", in Proc.Int. Conf. World Wide Web, 2009, pp. 351 - 360.
2. X. Li, C. Snoek, and M. Worring, "Learning Tag Relevance By Neighbor Voting For Social Image Retrieval", in Proc. ACM in Conf. Multimedia Inform. Retrieval, 2008, pp. 180 - 187.
3. D. Mishra, U. P. Singh, and V. Richhariya, "Tag Relevance For Social Image Retrieval In Accordance With Neighbor Voting Algorithm" , Int. J. Comput.Sci. Netw. Secur., vol.14, no. 7 pp.50 - 57, 2014.
4. K. Yang, M. Wang, X. Hua, and H. Zhang "Social Image Search With Diverse Relevance Ranking", in Proc. Int. Conf.Magn. Magn. Mater., 2010,pp. 174-184.
5. Dan Lu, Xiaoxiao Liu, and Xueming Qian, "Tag-Based Image Search by Social Re-ranking", IEEE Trans. Multimedia, vol. 18, no. 8, pp. 1628 - 1639, Aug. 2016.
6. S. Lee, W. D. Neve, and Y. M. Ro, "Visually weighted neighbor voting for image tag relevance learning", Multimedia Tools Appl., vol. 72, no. 2, pp. 1363 - 1386, 2013.
7. Y. Gao et al., "Tag-based social image search with visual-text joint hyper graph learning", in Proc. ACM Int. conf Multimedia Inform.Retrieval, 2011, pp. 1517 - 1520.
8. D. Wu, J. Wu, and M. Lu, "A two-step similarity ranking scheme for image retrieval", in Proc. IEEE 6th Int. Symp. Parallel Architectures, Algorithms Program.,Jul. 2014, pp.191 - 196.
9. David G. Lowe., "Distinctive image features from scaleinvariant keypoints" in International journal of computer vision, 60(2):91 - 110 2004.
10. T. Ojala, M. Pietikinen, and T. Menp, Multiresolution grayscale and rotation invariant texture classification with local binary patterns,IEEE Trans. Pattern Analysis and Machine Intelligence, vol.24, no.7,pp.971-987, 2002
11. L. Chen, D. Xu, and I. Tsang, "Tag-based image retrieval improved by augmented features and group - based refinement", IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1057 - 1067, Aug.2012.
12. R. Cilibrasi and P. Vitanyi, "The Google similarity distance", IEEE Trans.Knowl. Data Eng., vol. 19, no. 3, pp. 1065 - 1076, Mar. 2007.
13. Flickr. <http://www.flickr.com>
14. Hirdesh Kumar, Padmavati, "Face Recognition using SIFT by varying Distance Calculation Matching Method" in international journal of computer applications pp.0975 - 888, 2012, pp. 21-26
15. Trasha Gupta, , Lokesh Garg, "Face Recognition Using SIFT ", in International Journal of Emerging Technology and Advanced Engineering, 2014, pp. 358-363 .

AUTHORS PROFILE



Mr. C. Balasubramanian is working as assistant professor in the department of CSE at Mepco Schlenk Engineering College, Sivakasi, India. He has the teaching experience of over 14 years. He has published two papers in the international journals and presented around fifteen technical papers in various conferences.



Dr. J. Raja Sekar is working as associate professor in the department of CSE at Mepco Schlenk Engineering College, Sivakasi, India. He has the teaching experience of over 25 years. He has published fourteen papers in the international journals and presented around forty three technical papers in various conferences.