ijITECH International Journal of Innovative Technologies

WWW.IJITECH.ORG

ISSN 2321-8665
Vol.03,Issue.02,
June-2015,
Pages:0110-0113

Weakly Image Label Extraction using Data Mining Techniques

G. MANJULA[1], A. E. KOKILA[2]

[1]Assistant Professor, Dept of CSE, Priyadarshini College Of Engineering and Technology, Nellore, AP, India

[2]PG Scholar, Dept of CSE, Priyadarshini College Of Engineering and Technology, Nellore, AP, India.

Abstract: We mine weakly labeled facial images which are freely accessible on the World Wide Web to investigate a framework of search-based face annotation. There is one major problem for SBFA scheme is how to efficiently implement annotation by exploiting the list of most similar facial images and their weak labels which are often noisy and incomplete. To solve this problem, we suggest an operational unsupervised label refinement (ULR) method for refining the labels of web facial images via machine learning techniques. We convey the learning difficult as a convex optimization and develop efficient optimization algorithms to resolve the large-scale learning task efficiently. We further more propose a clustering-based approximation algorithm that can increase the scalability considerably and speeds up the scheme. We have accompanied a wide set of experimental studies on a large-scale web facial image test bed, in which positive results showed that the proposed ULR algorithms can significantly increase the performance of the SBFA scheme.

Keywords: Face Annotation, Content-Based Image Retrieval, Facial Image Test Bed, Label Enhancement, Web Facial Images, Weak Label.

I. INTRODUCTION

Various digital cameras and the rapid growth of social media tools are more popular for internet-based photo sharing, recent years have witnessed an explosion of the number of digital photos captured and stored by consumers. A large portion of photos shared by users on the Internet are Human facial images. Some of these facial images are tagged with names, but many of them are not tagged properly. This has motivated the study of auto face annotation, an important technique that aims to annotate facial images automatically. Auto face annotation can be beneficial to many real-world applications. For example, with auto face annotation techniques, online photo-sharing sites (e.g., Facebook) can automatically annotate users' uploaded photos to facilitate online photo search and management. Besides, face annotation can also be applied in news video domain to detect important persons appeared in the videos to facilitate news video retrieval and summarization tasks. Classical face annotation approaches are often treated as an extended face recognition problem, where different classification models are trained from a collection of well labeled facial images by employing the supervised or semi-supervised machine learning techniques. However, the "model-based face annotation "techniques are limited in several aspects. First, it is usually time-consuming and expensive to collect a large amount of human-labeled training facial images. Second, it is usually difficult to generalize the models when new training data or new persons are added, in which an intensive retraining process is usually required. Last but not least, the annotation/recognition performance often scales poorly when the number of persons/classes is very large.

Recently, some emerging studies have attempted to explore a promising search-based annotation paradigm for facial image annotation by mining the World Wide Web(WWW), where a massive number of weakly labeled facial images are freely available. Instead of training explicit classification models by the regular model-based face annotation approaches, the search-based face annotation(SBFA) paradigm aims to tackle the automated face annotation task by exploiting content-based image retrieval(CBIR) techniques in mining massive weakly labeled facial images on the web. The SBFA frame work is data-driven and model-free, which to some extent is inspired by the search-based image annotation techniques for generic image annotations. The main objective of SBFA is to assign correct name labels to a given query facial image. In particular, given a novel facial image for annotation, we first retrieve a short list of top K most similar facial images from a weakly labeled facial image database, and then annotate the facial image by performing voting on the labels associated with the top K similar facial images. One challenge faced by such SBFA paradigm is how to effectively exploit the short list of candidate facial images and their weak labels for the face name annotation task. To tackle the above problem, we investigate and develop a search-based face annotation scheme as shown in Fig.1. In particular, we propose a novel unsupervised label refinement (URL) scheme by exploring machine learning techniques to enhance the labels purely from the weakly labeled data without human manual efforts.

We also propose a clustering-based approximation (CBA) algorithm to improve the efficiency and scalability. As a summary, the main contributions of this paper include the following:

- We investigate and implement a promising search based face annotation scheme by mining large amount of weakly labeled facial images freely available on the WWW.
- We propose a novel ULR scheme for enhancing label quality via a graph-based and low-rank learning approach.
- We propose an efficient clustering-based approximation algorithm for large-scale label refinement problem.
- We conducted an extensive set of experiments, in which encouraging results were obtained.

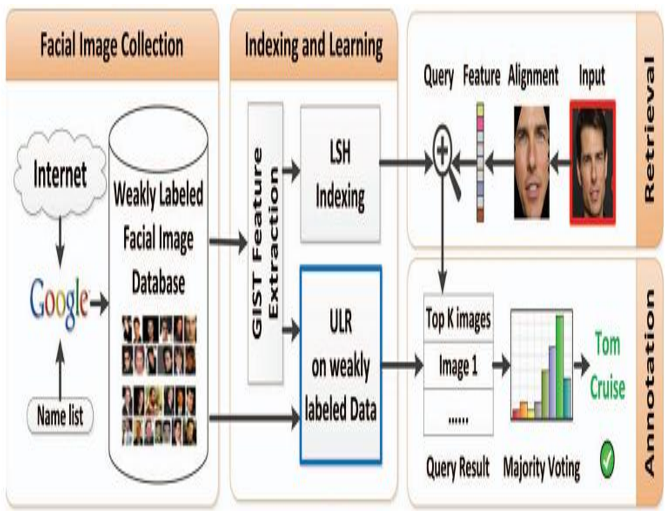


Fig.1. The system flow of the proposed search-based face annotation scheme.

II. FRAME WORK FOR SEARCH BASED FACE ANNOTATION

The proposed framework of search-based face annotation, which consists of the following steps:

- facial image data collection;
- face detection and facial feature extraction;
- high-dimensional facial feature indexing;
- learning to refine weakly labeled data;
- similar face retrieval; and
- Face annotation by majority voting on the similar faces with the refined labels.

The first four steps are usually conducted before the test phase of a face annotation task, while the last two steps are conducted during the test phase of a face annotation task, which usually should be done very efficiently. We briefly describe each step below.

- The first step is the data collection of facial images as we gathered a collection of facial images from the WWW by an existing web search engine according to a name list that contains the names of persons to be collected. As the output of this crawling process, we shall obtain a collection of facial images, and each of them is associated with some human names. Given the nature of web images, these facial images are often noisy, which do not always correspond to the right human name. Thus, we call such kind of web facial images with noisy names as weakly labeled facial image data.
- The second step is to preprocess web facial images to extract face-related information, including face detection and alignment, facial region extraction, and facial feature representation. For face detection and alignment, we adopt the unsupervised face alignment technique proposed in. 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.
- The third step is 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), a very popular and effective high-dimensional indexing technique.

Besides the indexing step, another key step of the framework is to engage an unsupervised learning scheme to enhance the label quality of the weakly labeled facial images. This process is very important to the entire search based annotation framework since the label quality plays a critical factor in the final annotation performance. All the above are the processes before annotating a query facial image. Next, we describe the process of face annotation during the test phase. In particular, given a query facial image for annotation, we first conduct a similar face retrieval process to search for a subset of most similar faces from the previously indexed facial database. With the set of top K similar face examples retrieved from the database, the next step is to annotate the facial image with a label or a subset of labels by employing a majority voting approach that combines the set of labels associated with these top K similar face examples. We focus our attention on one key step of the above framework, i.e., the unsupervised learning process to refine labels of the weakly labeled facial images.

III. EXPERIMENTS ON PROPOSED SYSTEM

A. Experiment Test Bed

In our experiments, we collected a human name list consisting of popular actor and actress names from the IMDB website: http://www.imdb.com. In precise, we collected these names with the billboard: “Most Popular People Born In yyyy” of IMDb, where yyyy is the born year. For example, the webpage2presents all the actor and actresses who were born in 1975 in the popularity order. Our name list covers the actors and actresses who were born between 1950 and 1990. To enlarge the retrieval database, we extended the name number from 400 to1,000. We submitted each name from the list as a query to search for the related web images by Google image search engine. The top 200 retrieved web images are crawled automatically. After that we used the Open CV toolbox to detect the faces and adopt the DLK algorithm to align facial images into the same well-defined position. The no-face-detected web images were ignored. As a result, we collected over 100,000 facial images in our database. We refer to this database as the “retrieval database,” which will be used for facial image retrieval during the auto face annotation process.

To evaluate varied number of persons in database, we divided our database into two scales: one contains 400 persons and about 40,000 and the other contains 1,000 persons and about 100,000 images. We denote them by "DB0400" and "DB1000," respectively. For the "test data set," specifically, we randomly chose 80 names from our name list. We submitted each selected name as a query to Google and crawled about 100 images from the top 200th to 400thsearch results. Note that we did not consider the top 200 retrieved images since they had already appeared in the retrieval data set. This aims to examine the generalization performance of our technique for unseen facial images. Since these facial images are often noisy, to obtain ground truth labels for the test data set, we request our staff to manually examine the facial images and remove the irrelevant facial images for each name. As a result, the test database consists of about 1,000 facial images with over 10 faces per person on average.

B. Calculation on Varied Top K Retrieved Images and Top T Annotated Names

This experiment aims to examine the relationship between the annotation performance of varied values of K and T, respectively, for top K retrieved images and top T annotated names. To ease our discussion, we only show the results of the ULR algorithm. The face annotation performance of varied K and T values are illustrated in following Fig.2.

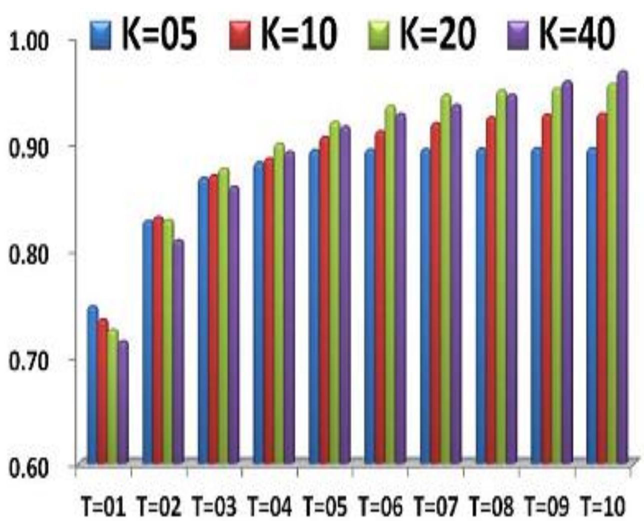
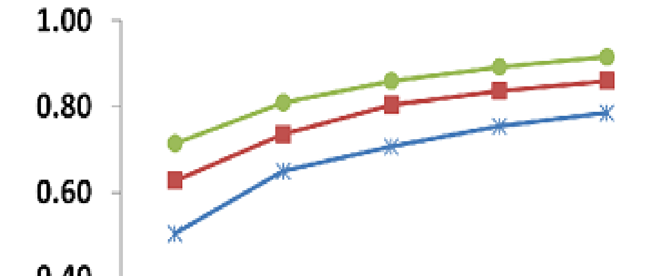


Fig.2. Performance of varied K and T values.

Some observations can be drawn from the experimental results. First of all, when fixing K, we found that increasing T value usually leads to better hit rate results. This is not surprising since generating more annotation results certainly gets a better chance to hit the relevant name. Second, when fixing T, we found that the impact of the K value to the annotation performance fairly depends on the specific value. In specific, when T is small (e.g., T = 1), increasing the K value leads to the decline of the annotation performance; but when T is large (e.g., T>5), increasing the K value often boosts the performance of top T annotation results. Such results can be explained as follows: When T is very small, for example, T=1, we prefer a small K value such that only the most relevant images will be retrieved, which, thus, could lead to more precise results attop-1 annotated results. However, when T is very large, we prefer a relatively large K value since it can potentially retrieve more relevant images and thus can improve the hit rate at top T annotated results.

C. Calculation on Varied Numbers of Images Per Person In Database

This experiment aims to further examine the relationship between the annotation performance and the number of facial images per person in building the facial image database. Unlike the previous experiment with top 100retrieval facial images per person in the database, we created three variables of varied-size databases, which consist of top 50, 75, and 100 retrieval facial images per person, respectively. We denote these three databases asP050, P075, and P100, respectively.



	T=01	T=02	T=03	T=04	T=05
P050	0.505	0.651	0.708	0.755	0.786
P075	0.628	0.736	0.804	0.836	0.859
P100	0.715	0.809	0.859	0.892	0.916

Fig.3. the test results of average annotation performance.

The above Fig.3 shows the test results of average annotation performance. It is clear that the larger the number of facial images per person collected in our database, the better the average annotation performance can be achieved. This observation is trivial since more potential images are included into the retrieval database, which is beneficial to the annotation task. We also noticed that enlarging the number of facial images per person in general leads to the increases of computational costs, including time and space costs for indexing and retrieval as well as the ULR learning costs.

D. Calculation of Clustering-Based Estimate

In this test, we aim to calculate the acceleration performance of the two suggested clustering-based approximation systems (BCBA and DCBA) on the large databaseDB1000. A noble approximation is expected to achieve a high reduction in running time with a small loss in annotation performance. Therefore, this experiment evaluates both running time and annotation performance. The running time of CBA scheme mainly consists of three parts: 1) the time of constructing the similarity matrixC;2) the time of clustering; and 3) the total time of running ULR algorithm in each subset. First of all, the proposed CBA scheme could suggestively decrease the running time for the label refinement task. Second, increasing the value of cluster number q_c generally leads to less running time, however, the reduction becomes marginal where q_c is larger than some threshold. Third, the running time of the division clustering algorithm is a bit smaller than the one of is setting the K-mean algorithm. The reasons leading to this phenomenon are twofold: one is there is no need for multi loops in each bisection step of DCBA, another is the similarity matrix is directly used for MST building without extra computation. Two observations can be drawn from the results. First, although the approximation algorithms (BCBA, DCBA) slightly degrade the final annotation performance, their performances are still much better than the other compared algorithms for small T value. Considering the reduction in running time, the proposed clustering-based approximation scheme is a good approximation for the ULR algorithm, which could significantly improve the scalability of search-based face annotation framework. Second, the performance difference between BCBA and DCBA are statistically marginal, but the average performance of BCBA is a bit better than DCBA.

E. Label Enhancement on Artificial Data Set

In this test, we target to calculate the label refinement performance of various algorithms. We built an artificial data set that consists of nine classes in 2D space with 20 samples for each class. To introduce noise into the label matrix, we arbitrarily mislabeled half of the whole dataset. Given the data set and the noisy label matrix, we calculated the enhanced label matrixes using the four algorithms. Several observations can be drawn from the above results: first, the MKL and CL algorithms work well for the classes with a smaller amount of noise, but they fail for the classes where more samples are mislabeled and widely distributed. Second, by taking the graph information, both LPSN and ULR could handle all the classes better. Clearly, by finding the extreme value in each label vector, we can recover the ideal label matrix from the refined label matrix FULR. Third, for the proposed ULR algorithm, we also consider the distortion with the original label matrix and the scarcity of each label vector. As a result, ULR can attain more stable and sparse refined label matrix that is more appropriate for our face annotation problem.

IV. BOUNDARIES

Regardless of the hopeful results, our work is limited in various aspects. First, we adopt each name relates to a unique single person. Duplicate name can be a practical issue in real-life situations. One future enhancement is to extend our method to address this practical problem. For example, we can learn the similarity between two different names according to the web pages so as to find how likely the two dissimilar names belong to the same person. Second, we assume the top retrieved web facial images are related to a query human name. This is clearly true for celebrities. However, when the query facial image is not a well-known person, there may not exist many relevant facial images on the WWW, which consequently could affect the performance of the suggested annotation solution. This is a common limitation of all existing data-driven annotation methods. This might be partly solved by exploiting social contextual information.

V. CONCLUSION

We examined a promising search-based face annotation framework, in which we concentrated on solving the critical problem of improving the label quality and suggested a ULR algorithm. We also suggested clustering-based approximation solution, which successfully accelerated the optimization task without introducing much performance degradation. After a wide set of experiments, we found that the proposed system achieved hopeful results under a range of settings. These results also indicated that the proposed ULR method significantly exceeded the other regular approaches in literature. Future work will discourse the issues of duplicate human names and explore supervised/semi-supervised learning methods to further increase the label quality with reasonable human manual enhancement efforts.

VI. REFERENCES

[1] J.Y. Choi, W.D. Neve, K.N. Plataniotis, and Y.M. Ro, “Collaborative Face Recognition for Improved Face Annotation in Personal Photo Collections Shared on Online Social Networks,”IEEE Trans. Multimedia,vol. 13, no. 1, pp. 14-28, Feb. 2011.

[2] Z. Wu, Q. Ke, J. Sun, and H.-Y. Shum, “Scalable Face Image Retrieval with Identity-Based Quantization and Multi-Reference Re-Ranking,” Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR),pp. 3469-3476, 2010.

[3] J. Cui, F. Wen, R. Xiao, Y. Tian, and X. Tang, “EasyAlbum: An Interactive Photo Annotation System Based on Face Clustering and Re-Ranking,”Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI),pp. 367-376, 2007.

[4] P.T. Pham, T. Tuytelaars, and M.-F. Moens, “Naming People in News Videos with Label Propagation,”IEEE Multimedia, vol. 18, no. 3, pp. 44-55, Mar. 2011.

[5] Z. Cao, Q. Yin, X. Tang, and J. Sun, “Face Recognition with Learning-Based Descriptor,” IEEE Conf. Computer Vision and Pattern Recognition (CVPR),pp. 2707-2714, 2010.

[6] J. Tang, R. Hong, S. Yan, T.-S. Chua, G.-J. Qi, and R. Jain, “ImageAnnotation by KNN-Sparse Graph-Based Label Propagation overNoisily Tagged Web Images,”ACM Trans. Intelligent Systems andTechnology,vol. 2, pp. 14:1-14:15, 2011.