Advances in Computational Sciences and Technology
ISSN 0973-6107 Volume 10, Number 5 (2017) pp. 1211-1219

http://www.ripublication.com

Data base Search on Web Facial Images using Unsupervised Label Refinement

Pankaj Agarkar
JJTU Research Scholar , Zunzuno Rajsthan
Department of Computer Engineering , Pune
Email: pmagarkar@gmail.com

Dr. S.D. Joshi
JJTU Research Guide , Zunzuno Rajsthan
Department of Computer Engineering , Pune
Email : sdj@live.com

Abstract

Auto face annotation is important in many real world knowledge management system and multimedia information. Face annotation is a field of face detection and recognition. Mining weakly labeled web facial images on the internet has emerged as a promising paradigm towards auto face annotation. A know framework search based face annotation (SBFA) is there to mine weakly labeled facial images that is available on World Wide Web (WWW). A one challenging problem with search based face annotation scheme is how effectively perform annotation by listing similar facial images and their weak labels which are noisy and incomplete. To treacle this problem proposed approach is unsupervised label refinement (URL) for refining the labels of web facial images. To speed up the proposed scheme a clustering-based approximation algorithm is used which improve the scalability. The proposed UPL algorithm can significantly boost the performance of the promising SBFA scheme.

Keywords: Face annotation, web facial images, weak label, label refinement

I. INTRODUCTION

Digital photo albums are growing explosively in both number and size due to rapid popularization of digital cameras and mobile phone cameras. A large portion of photos shared by users on Internet are human facial images. User share and access large volume of information on social networking sites like Facebook, Flicker. Face annotation for effective management of personal photos in online social networks (OSNs) is currently of considerable practical interest. These large collections require the annotation of some semantic information to facilitate browsing, manipulation and sharing of photos. The existing OSNs only support manual face annotation, a task that can be considered time-consuming and labor-intensive. A new framework collaborative face recognition (FR) framework [1], to improving the accuracy of face annotation is adapted. The accuracy of face annotation is improved by effectively making use of multiple FR engines available in an OSN. The large number of human facial images shared over the different social real world applications some of these images are tagged properly with names, but many of them are not tagged properly. This is motivated the study of *auto face annotation.* Auto face annotation is an important technique which automatically gives name of relevant person.

Auto face annotation technique is 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.

The paper is attempting to explore a promising search-based annotation paradigm for facial image annotation by mining the World Wide Web (WWW), where massive number weakly labeled facial images are freely available. The search-based face annotation (SBFA) paradigm aims to tackle the automated face annotation task by exploiting content-based image retrieval (CBIR) [2] techniques in mining massive weakly labeled facial images on the web. The main objective of SBFA [3], [4], [5] is to assign correct name labels to a given query facial image. In this novel facial image for annotation, we first retrieve a short list of top *K* most similar facial images from a weakly labeled facial images database, and then annotate the facial image by performing voting on the labels associated with the top *K* similar facial images.

One of the challenges faced by SBFA paradigm is how to effectively utilize the short list of candidate facial images and their weak labels for the face name annotation task. To deal with this problem we inspect and develop a search-based face annotation scheme. We propose a work of fiction unsupervised label refinement (URL) scheme to enhance the labels purely from the weakly labeled data without human manual efforts. A cluster-based approximation (CBA) algorithm is adapted to improve the efficiency and scalability.

II. RELATED WORK

The first group of related work is on the topics of face recognition and verification, which are classical research problems in computer vision and pattern recognition. Biometric-based technologies include identification based on Physiological characteristics and behavioral traits. Face recognition appears to offer several advantages over other Biometric method; facial images can be easily obtained with a couple of inexpensive fixed cameras. Good face recognition algorithms and appropriate preprocessing of the images can compensate for noise and slight variations in orientation, scale and illumination.

Face recognition is used for two primary tasks:

A. Verification (one-to-one matching):

When presented with a face image of an unknown individual along with a claim of identity, ascertaining whether the individual is who he/she claims to be.

B. Identification (one-to-many matching):

Given an image of an unknown individual, determining that person's identity by comparing (possibly after encoding) that image with a database of (possibly encoded) images of known individuals.

Face recognition techniques

Face recognition techniques can be broadly divided into three categories: methods that operate on intensity images, those that deal with video sequences, and those that require other sensory data such as infra-red imagery.

a) Face Recognition from Intensity Images

Face recognition methods for intensity images fall into two main categories: Feature-based and Holistic. Feature-based approaches first process the input image to identify and extract (and measure) distinctive facial features such as the eyes, mouth, nose, etc., and then compute the geometric relationships among those facial points, thus reducing the input facial image to a vector of geometric features and Holistic approaches attempt to identify faces using global representations, i.e., descriptions based on the entire image rather than on local features of the face.

b) Face Recognition from Video Sequences

A video-based face recognition system typically consists of three modules: one for detecting the face; a second one for tracking it; and a third one for recognizing it. Most of these systems choose a few good frames and then apply one of the recognition techniques for intensity images to those frames in order to identify the individual.

The second group is about the studies of generic image annotation. The classical image annotation approaches usually apply some existing object recognition techniques to train classification models from human-labeled training images or attempt to infer the correlation/probabilities between images and annotated keywords. Given limited training data, semi-supervised learning methods have also been used for image annotation [9], [10], and [11]. For example, Wang et al. [9] proposed to refine the model-based annotation results with a label similarity graph by following random walk principle. Similarly, Pham et al. [10] proposed to annotate unlabeled facial images in video frames with an iterative label propagation scheme. Although semi-supervised learning approaches could leverage both labeled and unlabeled data, it remains fairly time-consuming and expensive to collect enough well-labeled training data to achieve good performance in large-scale scenarios. Recently, the search-based image annotation paradigm has attracted more and more attention [3].

The third group is about face annotation on personal/ family/social photos. It focused on the annotation task on personal photos, which often contain rich contextual clues, such as personal/family names, social context, geotags, timestamps and so on. These techniques usually achieve fairly accurate annotation results, in which some techniques have been successfully deployed in commercial applications, for example, Apple iPhoto, Google Picasa, Microsoft easy Album and Face book face auto tagging solution. Jae young choi et al. [1] they proposed a novel collaborative framework of face recognition for improving the accuracy of face annotation. Multiple FR engines available in online social networks (OSN"s) are used for effective FR. This paper includes two main tasks, first is the selection of expert FR engines to recognize query face images. And second is the merging of multiple FR results, generated from different FR engines, into single FR results. These works implement the viola-Jones face detection algorithm for detecting facial images in personal photos.

The fourth group is about the studies of face annotation in mining weakly labeled facial images on the web. Some studies consider a human name as the input query, and mainly aim to refine the text-based search results by exploiting visual consistency of facial images. For example, Ozkan and Duygulu [12] proposed a graph-based model for finding the densest sub-graph as the most related result. Following the graph-based approach, Le and Satoh [13] proposed a new local density score to represent the importance of each returned images, and Guillemin et al. [14] introduced a modification to incorporate the constraint that a face is only depicted once in an image. On the other hand, the generative approach like the Gaussian mixture model was also been adopted to the name-based search scheme and achieved comparable results. Recently, a discriminate approach was proposed in [15] to improve over the generative approach and avoid the explicit computation in graph-based approach. By using ideas from query expansion, the performance of name-based scheme can be further improved with introducing the images of the “friends” of the query name. Unlike these studies of filtering the text-based retrieval results, some studies have attempted to directly annotate each facial image with the names extracted from its caption information. For example, Berg et al. [16] proposed a possibility model combined with a clustering algorithm to estimate the relationship between the facial images and the names in their captions.

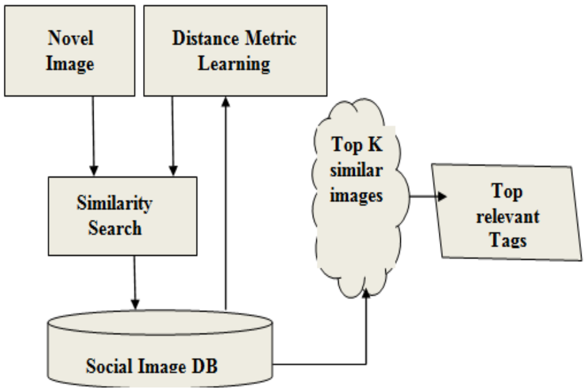

Figure 1. Retrieval-Based Annotation Approach

Fig.1 shows that retrieval based approach are applied with distance metric learning also various different techniques are implemented with these retrieval based or search based face annotation.

The fifth group is about purifying web facial images, which aims to leverage noisy web facial images for face recognition applications. Usually these works are proposed as a simple preprocessing step in the whole system without adopting sophisticated techniques.

III. SEARCH BASED FACE ANNOTATION

The below fig.2 illustrates the system flow of the proposed framework of search-based face annotation.

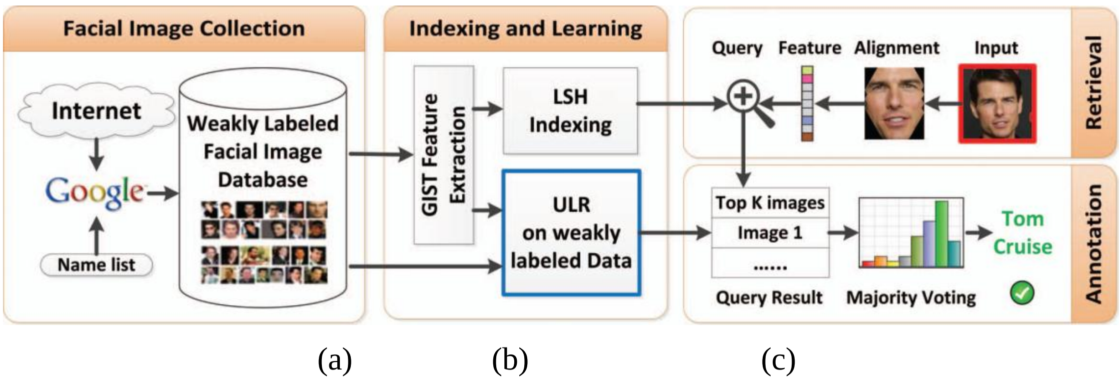

Fig.2 The system flow of the search-based face annotation scheme. (a) Collect weakly labeled facial images from WWW (b) The crawled web facial images, including face detection, face alignment, and feature extraction for the detected faces; (c) Search for the query facial image to retrieve the top K similar images and use their associated names for voting toward auto annotation.

Search based face annotation framework consist of the following steps.

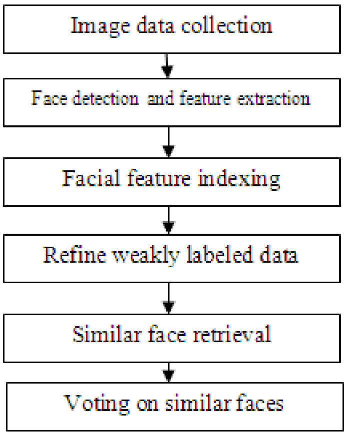

Fig. 3 Search based face annotation

Step 1: Facial image data collection;

Step 2: Face detection and facial feature extraction;

Step 3: High-dimensional facial feature indexing;

Step 4: Learning to refine weakly labeled data;

Step 5: Similar face retrieval.

Step 6: Face annotation by majority voting on the similar face with the refined labels.

Step 1 to step 4 are conducted before test phase of face annotation task. Step 5 and step 6 are conducted during the test phase of face annotation task.

The step 1 is the data collection of facial images as shown in fig.2 (a), in which facial images are collected from the WWW by an existing web search engine i.e. Google according to name list that contains the names of persons to be collected. Output of the process is a collection of facial images; each of them is associated with some human names. 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 step 2 is the preprocess web facial images to extract face-related information, which includes face detection and alignment, facial region extraction and facial feature representation. For face detection and alignment, we adopt the unsupervised face alignment technique [6]. For facial feature representation, it extracts the GIST texture features [7] to represent the extracted faces. As a result, each face can be represented by a d-dimensional feature vector.

The step 3 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. The locality sensitive hashing (LSH) [8], a very popular and effective high-dimensional indexing technique is adapted.

The step 4 is the unsupervised learning scheme to enhance the label quality of the weakly labeled facial images. This process is very important in the entire framework because the label quality is a critical factor.

The step 5 is the process of similar face retrieval to search for a subset of most similar faces (top K similar face) from the previously indexed facial database.

In the step 6 the majority voting approach is applied on set of top K similar face retrieved from database to annotate the facial image with a label.

IV. UNSUPERVISED LABEL REFINEMENT

A. *Preliminaries*

The Some denotes, X∈ 〖IR〗 ^(n*d) $X \in IR$ is the extracted facial image features, where *n* and *d* represents the number of feature dimensions, respectively. Ω= {n_1, n_2,…………. n_m } $\{n_1, n_2, \ldots\ldots n$ the list of human names for annotation, where *m* is the total number of human names. Y∈ 〖 [0, 1] 〗 ^ (n*m) he initial raw label matrix to describe the weak label information, in which the i^{th} row represents the label vector of the i^{th} facial image x_i∈ 〖IR〗 ^d In this application *Y* is the noisy and incomplete. For each weak label value Y_{ij}, $Y_{ij}\neq 0$ indicates that the i^{th} facial image x_i has label name n_j, while $Y_{ij}=0$ indicate that the relationship between i^{th} facial image x_i and j^{th} name is unknown.

V. CLUSTERING-BASED APPROXIMATION

The problem were considered as *n*m*, where *n* is the number of facial images in the retrieval database and *m* is the number of distinct names. For small problem, it can be efficiently solve by MGA-based algorithm. For large problem it can adapt the CDA-based algorithm. However, when *n* is extremely large, the CDA-based algorithm can be computationally intensive. One of the solutions to it is to adapt parallel computation. But speedup of the parallel computation approach is depending on hardware capability. To advance the scalability and efficiency in algorithm here is the clustering-based approximation solution.

Clustering concept could be applied in two different levels:

A. "image-level" which is used to directly separate all the *n* facial images into set of clusters.

B. "name-level" which is used to first separate the *m* names into clusters, then to further split retrieval database into different subsets according to name-label clusters.

"image-level" clusters would be more time—consuming than "name-level" because the number of facial images *n* is much larger than the number of names *m*.

VI. APPLICATIONS

Face annotation finds its application in the field of

- Achieve relatively high performance without user interaction.
- When user interaction is included, reduce to an acceptable level.
- Online photo album management and also in video domain.

VII. CONCLUSION

This paper presents an extensive survey on face annotation techniques foe web facial images. Here is the investigation on promising search based face annotation framework in which focused on tackling the critical problem of enhancing the label quality and proposed a URL algorithm. The paper proposed a clustering based approximation to improve the scalability which successfully accelerated the optimization task.

VIII. FUTURE WORK

The work is limited in the assumption that each name correspond to a unique single person. Future work will address the issues of duplicate human names. In the future direction we can learn the similarity between the two different names according to the web pages so as to determine how likely the two different names belong to the same person.

RRFERENCES

[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] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[3] X.-J. Wang, L. Zhang, F. Jing, and W.-Y. Ma, "AnnoSearch: Image Auto-Annotation by Search," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1483-1490, 2006.

[4] L. Wu, S.C.H. Hoi, R. Jin, J. Zhu, and N. Yu, "Distance Metric Learning from Uncertain Side Information for Automated Photo Tagging," ACM Trans. Intelligent Systems and Technology, vol. 2, no. 2, p. 13, 2011.

[5] P. Wu, S.C.H. Hoi, P. Zhao, and Y. He, "Mining Social Images with Distance Metric Learning for Automated Image Tagging," Proc. Fourth ACM Int'l Conf. Web Search and Data Mining (WSDM '11), pp. 197-206, 2011.

[6] J. Zhu, S.C.H. Hoi, and L.V. Gool, "Unsupervised Face Alignment by Robust Nonrigid Mapping," Proc. 12th Int'l Conf. Computer Vision (ICCV), 2009.

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

[8] 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.

[9] 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.

[10] P. Pham, M.-F. Moens, and T. Tuytelaars, "Naming Persons in News Video with Label Propagation," Proc. VCIDS, pp. 1528-1533, 2010.

[11] J. Tang, R. Hong, S. Yan, T.-S. Chua, G.-J. Qi, and R. Jain, "Image Annotation by KNN-Sparse Graph-Based Label Propagation over Noisily Tagged Web Images," ACM Trans. Intelligent Systems and Technology, vol. 2, pp. 14:1-14:15, 2011.

[12] D. Ozkan and P. Duygulu, "A Graph Based Approach for Naming Faces in News Photos," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1477-1482, 2006.

[13] D.-D. Le and S. Satoh, "Unsupervised Face Annotation by Mining the Web," Proc. IEEE Eighth Int'l Conf. Data Mining (ICDM), pp. 383-392, 2008.

[14] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, "Automatic Face Naming with Caption-Based Supervision," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2008.

[15] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, "Face Recognition from Caption-Based Supervision," Int'l J. Computer Vision, vol. 96, pp. 64-82, 2011.

[16] T.L. Berg, A.C. Berg, J. Edwards, and D. Forsyth, "Who's in the Picture," Proc. Neural Information Processing Systems Conf. (NIPS), 2005.