International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Review Paper on Retrieval Magnets for Facial Duplication by Search Based Face Annotation

Deepika B. Patil[1], Ayesha Butalia [2]

[1]P.G. Student, Department of Computer Engineering, GMRCEM, Wagholi, Pune, India,

[2] Professor, Department of Computer Engineering, GMRCEM, Wagholi, Pune, India,

Abstract: ***Search Based Face Annotation (SBFA) is an effective technique to annotate the weakly labeled facial images that are freely available on World Wide Web. The main objective of Search based face annotation is to assign correct name labels to given query facial image. One challenging problem for search-based face annotation scheme is how to effectively perform annotation by exploiting the list of most similar facial images and their weak labels that are often noisy and incomplete. A large portion of photos shared by users on the Internet are human facial images. However some of these facial images are tagged with names properly, but many of them are not tagged properly. Also the duplicate facial images cannot be annotated. This problem has motivated to study a new technique called as "auto face annotation" which aims to annotate facial images automatically. In this paper, an effective unsupervised label refinement (URL) approach is proposed for refining the labels of web facial images using machine learning technique. The learning problem can be formulated as a convex optimization and develop effective optimization algorithms to solve the large-scale learning task efficiently. This paper also addresses the issues of duplicate human names and explores supervised/semi-supervised learning techniques to enhance the label quality. The results obtained from the proposed ULR algorithms will significantly boost the performance of the promising SBFA scheme.***

Keywords: Annotation, label refinement, search-base, web facial images, weak label**.**

1. Introduction

Due to the recognition of assorted digital cameras and therefore the rapid growth of social media tools for internet-based photo sharing, recent years have witnessed associate explosion of the quantity of digital photos captured and hold on by consumers. An oversized portion of photos shared by users on the Internet square measure human facial pictures. However some of these facial images are tagged with names properly, but many of them are not tagged properly. This problem has motivated to study a new technique called as "auto face annotation" or 'machine face annotation', which aims to annotate facial images mechanically. Auto face annotation are often useful to several real world applications. As an example, with machine face annotation techniques, on-line photo-sharing sites (e.g., Facebook) will automatically annotate users' uploaded photos to facilitate online exposure search and management. [1]

Besides, face annotation can also be applied in news video domain to observe important persons appeared within the videos to facilitate news video retrieval and report tasks. Classical face annotation approaches are often treated as an extended face recognition drawback where different classification model square measure trained from a group of well labeled facial pictures by using the supervised or semi-supervised machine learning techniques. However, the "model-based face annotation" techniques square measure restricted in many aspects. First, it's sometimes time-consuming and costly to gather an oversized quantity of human-labeled coaching facial pictures. Second, it's sometimes difficult to generalize the models once new coaching knowledge or new persons square measure side, during which associate intensive grooming process is sometimes needed. Last however not least, the annotation/recognition performance usually scales poorly when the quantity of persons/classes is extremely massive. Recently, some rising studies have tried to explore a promising search-based annotation paradigm for facial image annotation by mining the planet Wide Web (WWW), wherever an enormous range of frail labeled facial images square measure freely offered. Rather than coaching specific classification models by the regular model-based face annotation approaches, the search-based face annotation (SBFA) paradigm aims to tackle the automatic face annotation taskby exploiting Content-Based Image Retrieval (CBIR) techniques in mining huge frail labeled facial pictures on the net. [2, 3]

2. Related Work

2.1. Social Media Modeling and Computing

S.C.H. Hoi et at. presented contributions from an international selection of preeminent experts in the field. The emergence of user-centric multimedia applications on social networks has been instrumental in the creation of a new form of "social" media – created using highly accessible and scalable publishing technologies for sharing via the internet. This timely text/reference presents the latest advances in various aspects of social media modeling and social media computing research. Gathering together superb research from a range of established international conferences and workshops, the editors coherently organize and present each of the topics in relation to the basic principles and practices of social media modeling and computing. [1]

2.2 Name-It: Naming and Detecting Faces in News Videos

Shin'ichi Satoh et al. developed a Name-It, a system that knowledge source, the system is given news videos which include image sequences and transcripts obtained from audio tracks or closed caption texts. The system can then either infer the name of a given face and output the name candidates, or can locate the faces in news videos by a name. To accomplish this task, the system extracts faces from image sequences and names from transcripts, both of which might correspond to key persons in news topics. The proposed system takes full advantage of advanced image and natural language processing. The image processing contributes to the extraction of face sequences which provide rich information for face name association. The processing also helps to select the best frontal view of a face in a face sequence to enhance the face identification which is required for the processing. [2]

2.3 Naming People in News Videos with Label Propagation

P.T. Pham et al. improved video content annotation and search tasks by Labeling persons appearing in video frames with names detected in a corresponding video transcript. They implement a face naming method that learns from labeled and unlabeled examples using iterative label propagation in a graph of connected faces or name-face pairs. By incorporating the unlabeled data points during the learning process, this method can work with few labeled data points. Moreover, they presents variations of this model that better cope with a large number of data by reducing the time and space complexity. On BBC News videos, the label propagation algorithm yields better results than a Support Vector Machine classifier and a nearest neighbor classifier trained on the same labeled data. Furthermore, we show that when anchor detection precedes the label propagation process, it helps boosting the face naming performance. Reusing labeled examples from different broadcasts, they manage to name 70% of the faces with a precision of 85%. [3]

2.4 Automated Annotation of Human Faces in Family Albums

L. Zhang et al. presented a learning framework to automate the face annotation in family photograph albums. Firstly, methodologies of content based image retrieval and face recognition are seamlessly integrated to achieve automated annotation. Secondly, face annotation is formulated in a Bayesian framework, in which the face similarity measure is defined as maximum a posteriori(MAP) estimation. Thirdly, to deal with the missing features, marginal probability is used so that samples which have missing features are compared with those having the full feature set to ensure a non-biased decision. The experimental evaluation has been conducted within a family album of few thousands of photographs and the results show that the proposed approach is effective and efficient in automated face annotation in family albums. [4]

2.5 Approaches, Challenges and Future Direction of Image Retrieval

Hui Wang et al. discussed the evolution of the retrieval approaches focusing on development, challenges and future direction of the image retrieval. It highlights both the already addressed and outstanding issues. The explosive growth of image data leads to the need of research and development of Image Retrieval. However, Image retrieval researches are moving from keyword, to low level features and to semantic features. Drive towards semantic features is due to the problem of the keywords which can be very subjective and time consuming while low level features cannot always describe high level concepts in the users' mind. Hence, introducing an interpretation inconsistency between image descriptors and high level semantics that known as the semantic gap. This paper also discusses the semantic gap issues, user query mechanisms as well as common ways used to bridge the gap in image retrieval. [5]

2.6 Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments

Gary B. Huang et al. created most face databases under controlled conditions to facilitate the study of specific parameters on the face recognition problem. These parameters include such variables as position, pose, lighting, background, camera quality, and gender. While there are many applications for face recognition technology in which one can control the parameters of image acquisition, there are also many applications in which the practitioner has little or no control over such parameters. This database, Labeled Faces in the Wild, is provided as an aid in studying the latter, unconstrained, recognition problem. The database contains labeled face photographs spanning the range of conditions typically encountered in everyday life. The database exhibits "natural" variability in factors such as pose, lighting, race, accessories, occlusions, and background. In addition to describing the details of the database, they provide specific experimental paradigms for which the database is suitable. This is done in an effort to make research performed with the database as consistent and comparable as possible. They provide baseline results, including results of a state of the art face recognition system combined with a face alignment system. [6]

2.7 Fisher Vector Faces in the Wild

Simonyan et al. makes two contributions: first, and somewhat surprisingly, we show that Fisher vectors on densely sampled SIFT features, i.e. an off-the-shelf object recognition representation, are capable of achieving state-of-the-art face verification performance on the challenging "Labeled Faces in the Wild" benchmark; second, since Fisher vectors are very high dimensional, we show that a compact descriptor can be learnt from them using discriminative metric learning. This compact descriptor has a better recognition accuracy and is very well suited to large scale identification tasks. [7]

2.8 Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection

Peter N. Belhumeur et al. developed a face recognition algorithm which is insensitive to large variation in lighting direction and facial expression. Taking a pattern classification approach, we consider each pixel in an image as a coordinate in a high-dimensional space. We take advantage of the observation that the images of a particular face, under varying illumination but fixed pose, lie in a 3D linear subspace of the high dimensional image space if the face is a Lambertian surface without shadowing. However, since faces are not truly Lambertian surfaces and do indeed produce self-shadowing, images will deviate from this linear subspace. Rather than explicitly modeling this deviation, we linearly project the image into a subspace in a manner which discounts those regions of _the face with large deviation. Our projection method is based on Fisher's Linear Discriminant and produces well separated classes in a low-dimensional subspace, even under severe variation in lighting and facial expressions. [8]

2.9 Names and Faces in the News

Tamara L. Berg et al. shows quite good face clustering is possible for a dataset of inaccurately and ambiguously labelled face images. Their dataset is 44,773 face images, obtained by applying a face finder to approximately half a million captioned news images. This dataset is more realistic than usual face recognition datasets, because it contains faces captured "in the wild" in a variety of configurations with respect to the camera, taking a variety of expressions, and under illumination of widely varying color. Each face image is associated with a set of names, automatically extracted from the associated caption. Many, but not all such sets contain the correct name. They cluster face images in appropriate discriminant coordinates. They use a clustering procedure to break ambiguities in labeling and identify incorrectly labelled faces. A merging procedure then identifies variants of names that refer to the same individual. The resulting representation can be used to label faces in news images or to organize news pictures by individuals present. They report results for (a) the original dataset (b) the datasets resulting from our clustering, merging and cleaning process, without using cluster information (c) the datasets resulting from our clustering, merging and cleaning process, including their cluster structure. [9]

2.10 Naming Every Individual in News Video Monologues

Jun Yang et al. approaches problem of Naming Every Individual in News Video Monologues with a statistical learning method. Naming every individual person appearing in broadcast news videos with names detected from the video transcript leads to better access of the news video content. Two categories of information extracted from multiple video modalities have been explored, namely features, which help distinguish the true name of every person, as well as constraints, which reveal the relationships among the names of different persons. The person naming problem is formulated into a learning framework which predicts the most likely name for each person based on the features, and refines the predictions using the constraints. Experiments conducted on ABC World New Tonight and CNN Headline News videos demonstrate that this approach outperforms a non-learning alternative by a large amount. They have described a machine learning model for naming every individual person in broadcast news videos. Trained based on some manually named people, this model predicts the most likely name of each individual who is giving a monologue speech based on features extracted from multiple modalities including speech, overlaid text, transcript, etc. Equivalence constraints relating the names of different persons are used to refine the predictions. Experiments have shown that this approach significantly outperforms a non-learning alternative in terms of the accuracy of the names assigned, particularly when the constraints are applied. [10]

3. Proposed Work

The SBFA framework is data-driven and model-free, that 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 question facial image. Specially, given a unique facial image for annotation, there is a tendency to initial retrieve a brief list of prime K most similar facial pictures from a frail labeled facial image information, then annotate the facial image by performing pick on the labels related to the highest K similar facial pictures. One challenge moon-faced by such SBFA paradigm is however to effectively exploit the list of candidate facial pictures and their weak labels for the face name annotation task. To tackle the on top of drawback, a tendency to investigate and develop a search-based face annotation theme will be proposed. The proposed system of search-based face annotation scheme is as shown in figure 1. Specially, a novel unsupervised label refinement (URL) scheme is proposed by exploring machine learning techniques to enhance the labels strictly from the frail labeled information without human manual efforts. There is a tendency to additionally propose a clustering-based approximation (CBA) algorithmic program to boost the efficiency and scalability. [11, 12, 13]

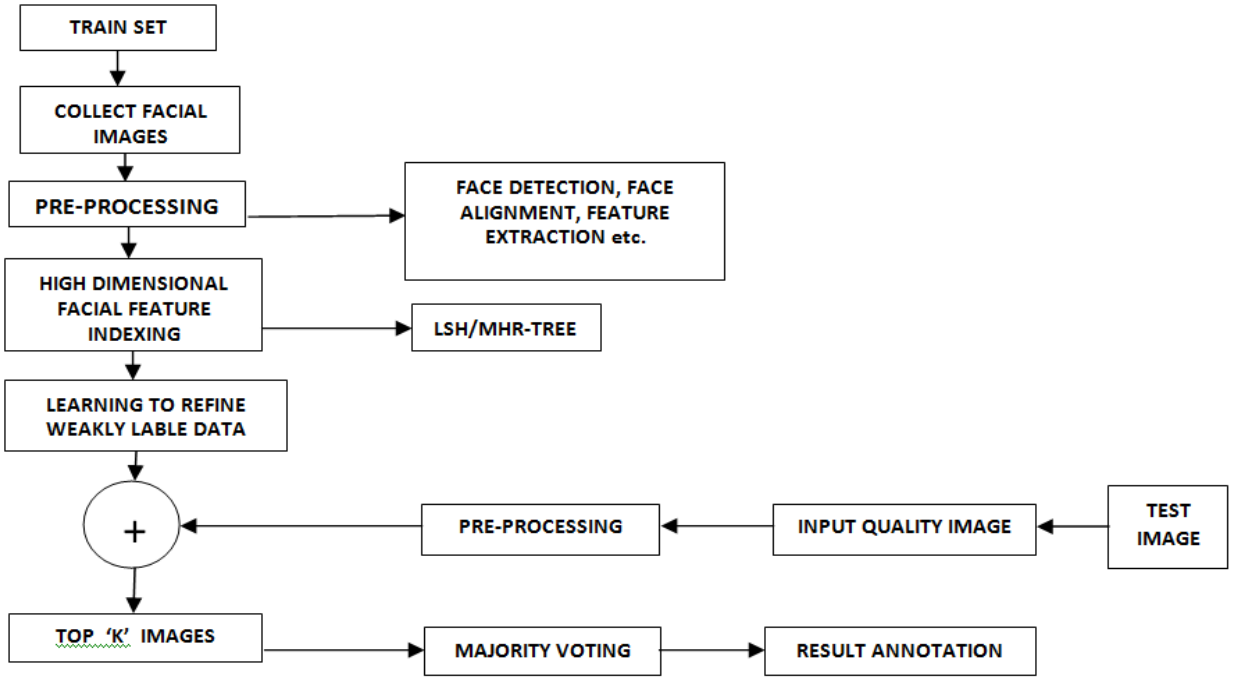

Figure 1: System Flow of Search Based Face Annotation Scheme

4. Conclusion

In this paper Search Based Face Annotation (SBFA) technique is proposed 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 main focus is on tackling the critical problem of enhancing the label quality and proposed an Unsupervised Label Refinement (ULR) algorithm. By giving a novel facial image for annotation, it 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. The proposed novel ULR scheme will enhance the label quality of weakly labeled facial images and will achieve the promising results under a variety of settings which will significantly surpassed the other regular approaches in literature. Also the issues of0020duplication of facial images can be easily removed with the help of this technique

References

[1] S.C.H. Hoi, J. Luo, S. Boll, D. Xu, and R. Jin, "Social Media Modeling and Computing", Eds. Springer, 2011.)
[2] S. Satoh, Y. Nakamura, and T. Kanade, "Name-It: Naming and Detecting Faces in News Videos," In Proceeding of IEEE Conference, vol. 6, no. 1, pp. 22-35, Jan.-Mar. 1999.
[3] P.T. Pham, T. Tuytelaars, and M.-F.Moens, "Naming People in News Videos with Label Propagation," In Proceeding of IEEE Conference, vol. 18, no. 3, pp. 44-55, Mar. 2011.
[4] L. Zhang, L. Chen, M. Li, and H. Zhang, "Automated Annotation of Human Faces in Family Albums," Proc. 11th ACM Int'l Conf. Multimedia (Multimedia), 2003.
[5] HuiHui Wang, DzulkifliMohamad, N.A. Ismail, "Approaches, Challenges and Future Direction of Image Retrieval ",Journal of Computing, Volume 2, Issue 6, June 2010.
[6] Gary B. Huang, MarwanMattar, Tamara Berg, and Erik Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments" Journal of Computing, Volume 1, Issue 4, June 2011.
[7] Karen Simonyan, Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman, "Fisher Vector Faces in the Wild", ERC Grant VisRec no. 228180 and EU Project AXES ICT- 269980.
[8] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces versus Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, July 1997.
[9] T.L. Berg, A.C. Berg, J. Edwards, M. Maire, R. White, Y.W. Teh, E.G. Learned-Miller, and D.A. Forsyth, "Names and Faces in the News," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 848-854, 2004.
[10] J. Yang and A.G. Hauptmann, "Naming Every Individual in News Video Monologues," Proc. 12th Ann. ACM Int'l Conf. Multimedia (Multimedia), pp. 580-587. 2004.
[11] 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.
[12] 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.
[13] 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.

[14] 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.

[15] S.C.H. Hoi, R. Jin, J. Zhu, and M.R. Lyu, "Semi-Supervised SVM Batch Mode Active Learning with Applications to Image Retrieval," ACM Trans. Information Systems, vol. 27, pp. 1-29, 2009.

[16] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Survey, vol. 35, pp. 399-458, 2003.

[17] G.B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face0020Recognition in Unconstrained Environments," technical report07-49, 2007.

[18] H.V. Nguyen and L. Bai, "Cosine Similarity Metric Learning for Face Verification," Proc. 10th Asian Conf. Computer Vision (ACCV '10), 2008.

[19] Y. Yang, Y. Yang, Z. Huang, H.T. Shen, and F. Nie, "Tag Localization with Spatial Correlations and Joint Group Sparsity," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 881-888, 2011