e-ISSN: 2349-9745
p-ISSN: 2393-8161
Scientific Journal Impact Factor (SJIF): 1.711
International Journal of Modern Trends in Engineering and Research
www.ijmter.com

FACE EXPRESSION IDENTIFICATION USING IMAGE FEATURE CLUSTRING AND QUERY SCHEME
ANNOTATING FACIAL IMAGES AUTOMATICALLY

S.Revathi [1], B.Arulkumar[2],B.Preetha[3]
[1]Computer Science and Engineering, Vivekanandha College Of Engineering For Women
[2]Computer Science and Engineering, Vivekanandha College Of Engineering For Women
[3]Computer Science and Engineering, Vivekanandha College Of Engineering For Women

Abstract - Web mining techniques are used to analyze the web page contents and usage details. Human facial images are shared in the internet and tagged with additional information. Auto face annotation techniques are used to annotate facial images automatically. Annotations are used in online photo search and management. Classification techniques are used to assign the facial annotation. Supervised or semi-supervised machine learning techniques are used to train the classification models. Facial images with labels are used in the training process. Noisy and incomplete labels are referred as weak labels. Search-based face annotation (SBFA) is assigned by mining weakly labeled facial images available on the World Wide Web (WWW). Unsupervised label refinement (ULR) approach is used for refining the labels of web facial images with machine learning techniques. ULR scheme is used to enhance the label quality using graph-based and low-rank learning approach. The training phase is designed with facial image collection, facial feature extraction, feature indexing and label refinement learning steps. Similar face retrieval and voting based face annotation tasks are carried out under the testing phase. Clustering-Based Approximation (CBA) algorithm is applied to improve the scalability. Bisecting K-means clustering based algorithm (BCBA) and divisive clustering based algorithm (DCBA) are used to group up the facial images. Multi step Gradient Algorithm is used for label refinement process. The web face annotation scheme is enhanced to improve the label quality with low refinement overhead. Noise reduction is method is integrated with the label refinement process. Duplicate name removal process is integrated with the system. The indexing scheme is enhanced with weight values for the labels. Social contextual information is used to manage the query facial image relevancy issues.

Keywords- Auto Face Annotation ,Query Facial Image , K-Means Clustering, Noise Reduction, Weakly Labeled Facial Images

I. INTRODUCTION

Generally speaking, image content may include both visual and semantic content. Visual content can be very general or domain specific. General visual content include color, texture, shape, spatial relationship, etc. Domain specific visual content, like human faces, is application dependent and may involve domain knowledge. Semantic content is obtained either by textual annotation or by complex inference procedures based on visual content. This concentrates on general visual content descriptions.

Advances in image acquisition and storage technology have lead to tremendous growth in significantly large and detailed image databases. The WWW is regarded as the largest global image repository. An extremely large number of image data such as satellite images, medical images, and digital photographs are generated every day. These images, if analyzed, can reveal useful information to the human users. Unfortunately, there is a lack of effective tools for searching and finding useful patterns from these images. Image mining systems that can automatically extract semantically meaningful information from image data are increasingly in demand. The fundamental challenge of image mining is to determine how the low level, pixel representation contained in a raw image or image sequence can be efficiently and effectively processed to identify high level spatial objects and its relationships.

A visual content descriptor can be either global or local. A global descriptor uses the visual features of the whole image, whereas a local descriptor uses the visual features of regions or objects to describe the image content. To obtain the local visual descriptors, an image is often divided into parts first. The simplest way of dividing an image is to use a partition, which cuts the image into tiles of equal size and shape. A simple partition does not generate perceptually meaningful regions but is a way of representing the global features of the image at a finer resolution. A better method is to divide the image into homogenous regions according to certain criterion. A more complex way of dividing an image is to undertake a complete region segmentation to obtain semantically meaningful regions.

II. LITERATURE SURVEY

2.1. Naming Persons in News Videos with Label Propagation

Massive digitisation efforts combined with the advent of media sharing websites and the explosion of user generated content have resulted in an overwhelming amount of video material available online. To date, access to this wealth of information is hampered by the lack of good indexing mechanisms. At best, archivists provide a few, often subjective, keywords or tags – by far insufficient for a directed large-scale search. Likewise, jumping to the fragment of interest within a video is normally not possible, as the manually added tags are not time-coded, forcing users to fall back to random sampling of the video timeline or to feed forward search through the video in a linear fashion.Methods that learn from multimodal data are becoming increasingly popular. By exploiting information from multiple modalities, the effort of constructing a large, carefully collected and annotated training corpus can often be avoided or reduced. This is achieved using a graph based algorithm that learns the name face alignments jointly from labeled and unlabeled examples. Variant versions of the label propagation model are proposed here that deal with time and space complexity of processing large datasets. In addition, we show the portability of labeled seeds across different broadcasts. We evaluate and compare with state-of-the-art labeling approaches on a larger dataset, i.e., on 9 broadcasts of BBC news.

2.2. Distance Metric Learning from Uncertain Side Information for Automated Photo Tagging

Although content-based image retrieval has been extensively studied, searching image and photo by textual queries remains one of the most common and imperative functions for most intelligent multimedia systems. For many real-world multimedia systems, raw images and photos are often not associated with text labels or human tags. Recently, due to the popularity of social networks and social web, massive tagged images have been available on the web, which is referred to as “social images/photos”. Existing DML methods work only with explicit side information, which is given either in the forms of class labels or pairwise constraints. Besides, existing DML methods also assume that the given side information is clean and perfect. For example, in our application, the tags and contents generated by users for images are often erroneous and more importantly cannot be used directly as the explicit side information. This motivates us to study a new approach of distance metric learning from uncertain/implicit side information.

To this end, in this paper, we present a novel Probabilistic Distance Metric Learning (PDML) framework, which aims to learn distance metrics from noisy and uncertain side information for automated photo tagging tasks. The proposed framework consists of two steps: (1) an unsupervised learning approach for discovering probabilistic side information from hidden erroneous and implicit side information contained in rich user-generated content of social image data; and (2) a PDML approach for learning an optimal distance metric from probabilistic side information.

2.3. Mining Social Images with Distance Metric Learning for Automated Image Tagging

Along with the popularity of digital cameras and high quality mobile devices as well as the advances of internet technologies, users can easily upload their images and photos over the World Wide Web (WWW).The traditional approach typically has two steps: (1) representing images by extracting visual features and (2) pre-

training recognition models by building classification models from a collection of manually-labeled training data. The idea of the retrieval-based paradigm is to first retrieve a set of *k* most similar images for a test photo from the social image repository and then to assign the test photo with a set of *t* most relevant tags associated with the set of *k* retrieved social images.

The crux of the retrieval-based photo tagging paradigm is to effectively identify and retrieve a set of top *k* similar photos from social image database, which mainly relies on two key components: (1) a feature representation scheme to extract salient visual features and (2) a distance measure scheme to compute distances for extracted features.To this end, this paper presents a novel unified distance metric learning (UDML) framework, which aims to learn effective metrics from implicit side information of social images towards the application of automated photo tagging. Besides, this framework also unifies both inductive and transductive metric learning approaches together in a systematic approach.

2.4. Tag Localization with Spatial Correlations and Joint Group Sparsity

Multimedia understanding and retrieval is a long standing research problem in the field of computer vision and multimedia. Nowadays, confronted with the huge number of social images on the Web, traditional image-level tagging methods tend to become less effective because global image matching approach can hardly handle the diversity and arbitrariness of Web image content. Several related efforts have been made on this research topic. Multiple instance learning techniques and graph models have been exploited and shown some effectiveness in region-level annotation. The common tags of images containing the target region and sparsely selected regions will be re-assigned to the target region according to the reconstruction coefficients. It is worth noting that basic regions in the dictionary are implicitly assumed to be independent with each other. Contextual relationships among these semantic regions/ objects, e.g., co-occurrence and spatial correlations, are ignored. Besides, when reconstructing regions within an image they individually encode each region and again ignore the intrinsic correlations among encoding regions. Hence, to overcome these drawbacks we propose a joint region reconstruction model which extends group sparse coding with collaborative encoding ability and integrates spatial correlations among basic regions into the training dictionary.

2.5. Image Annotation by kNN-Sparse Graph-based Label Propagation over Noisily-Tagged Web Images

Recent years have witnessed the proliferation of social media and the success of many photo-sharing websites, such as Flickr and Picasa. These websites allow users to upload personal images and assign tags to describe the image contents. With the rich tags as metadata, users can more conveniently organize and access these shared images.Actually a graph is a gathering of pairwise relations, while the relation among visual images is essentially an estimation based on human cognition system. It has been found in neural science that the human vision system seeks a sparse representation for the incoming image using a few words in a feature vocabulary. This motivates us to construct the so-called sparse graph through the sparse reconstructions of the samples. The semi-supervised label inference for semantic concepts is then conducted on this sparse graph.

The kNN-sparse graph-based semi-supervised learning method has the following advantages: 1) it can remove most of the semantically-unrelated links to avoid the propagation of incorrect information, since each sample only has links to a small number of most probably semantically-related samples; 2) it is robust to noisy elements in the visual features; 3) it is naturally effective for discrimination since the sparse graph characterizing the local structure can convey important information for classification; and 4) it is practical for large-scale applications since the sparse representation can reduce the storage requirement while the approximate kNN-sparse graph construction is much more efficient than normal sparse graph construction.

III.PROBLEM DESCRIPTION

Due to the popularity of various digital cameras and the rapid growth of social media tools 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 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 framework 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.

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. 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. We note that a short version of this work had appeared in SIGIR2011.

- This journal article has been significantly extended by including a substantial amount of new content.

IV. PROPOSED SYSTEM

The web face annotation scheme is enhanced to improve the label quality with low refinement overhead. Noise reduction is method is integrated with the label refinement process. Duplicate name removal process is integrated with the system. The indexing scheme is enhanced with weight values for the labels. Social contextual information is used to manage the query facial image relevancy issues.

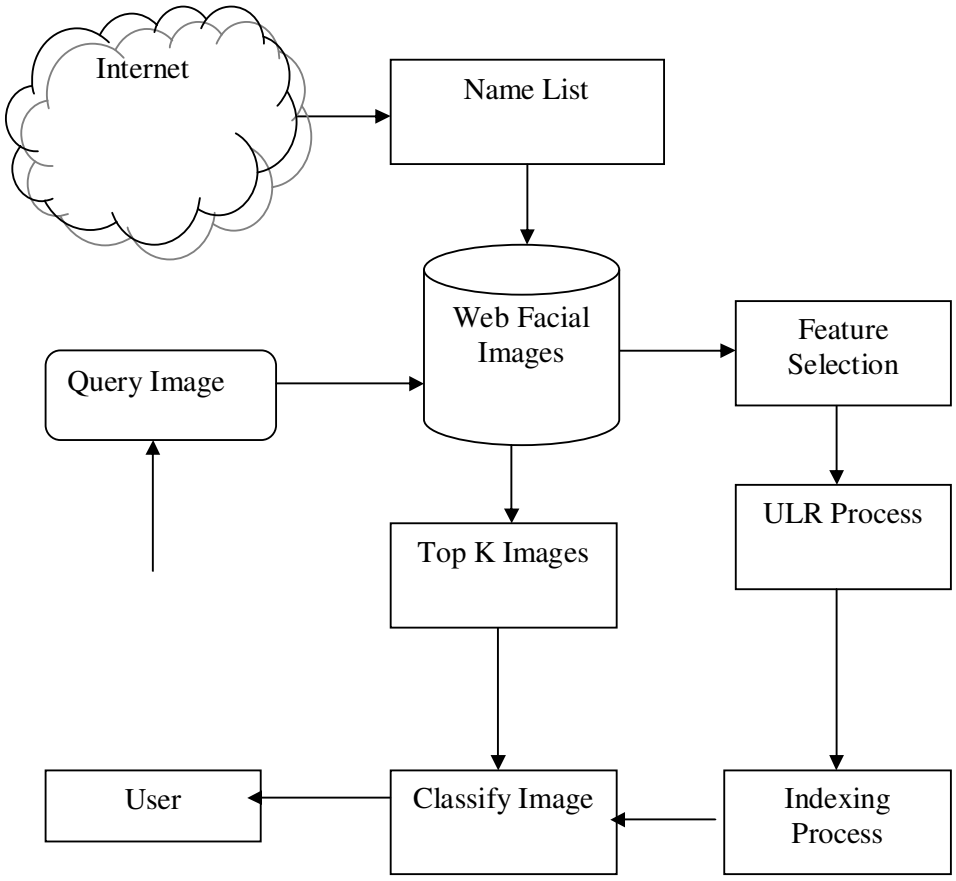

Figure 1.Architecture Diagram

V. CONCLUSION

Web facial images are assigned with annotated text values. Search Based Face Annotation (SBFA) method is used to assign annotated text for the facial images. Clustering techniques are used to group up the image features. The face annotation scheme is enhanced to improve the quality with minimum refinement overheads. Efficient large scale learning mechanism is applied in the system. Annotated label quality is high in the system. Duplicate name analysis model is performed to improve the accuracy level. Label assignment is achieved with minimum refined images.

REFERENCES

[1] D. Wang, S.C.H. Hoi, Y. He and J. Zhu, "Retrieval-Based Face Annotation by Weak Label Regularized Local Coordinate Coding," Proc. 19th ACM Int'l Conf. Multimedia (Multimedia), pp. 353- 362, 2011.
[2] Dayong 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.
[3] J. Tang, R. Hong, S. Yan 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.
[4] Konstantinos Klimis S. Ntalianis and Stefanos D. Kollias, "Mining User Queries with Markov Chains: Application to Online Image Retrieval", IEEE Transactions On Knowledge And Data Engineering, February 2013.
[5] 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.

[6] M. Guillaumin, T. Mensink and C. Schmid, "Face Recognition from Caption-Based Supervision," Int'l J. Computer Vision, vol. 96, 2011.
[7] 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, pp. 197-206, 2011.
[8] P.T. Pham, Tuytelaars and M.-F. Moens, "Naming People in News Videos with Label Propagation," IEEE Multimedia, vol. 18, no. 3, Mar. 2011.
[9] Pengfei Zhu, Wangmeng Zuo, Lei Zhang and David Zhang, "Image Set-Based Collaborative Representation for Face Recognition", IEEE Transactions On Information Forensics And Security, July 2014.
[10] 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.

