ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

ARP-GIST Descriptor for Search Based Face Annotation

Priyanka P. Taware, Dr. P. R. Futane

M.E. Student, Department of Computer Engineering, SCOE, Vadgaon(Bk), Pune, Savitribai Phule Pune University, Pune, India

Professor, Department of Computer Engineering, SCOE, Vadgaon(Bk), Pune, Savitribai Phule Pune University, Pune, India

ABSTRACT: A large portion of photos shared by users on Internet are human facial images. Some of these images are tagged with names, but many of them are not tagged properly. To overcome this problem, Auto face annotation technique is used. It aims to annotate facial image automatically. It uses a structure of Search Based Face Annotation (SBFA) by mining weakly named facial pictures that are totally accessible on the World Wide Web (WWW). Search based face annotation technique is used by mining weakly labeled facial images. In this technique, it is making use of ARP GIST descriptor that uses efficiency of angular partitioning for recording the finer specifics of picture. The designed algorithm keeps up tough spatial layout as well as provides versatility in every block, getting balance in spatial constraints as well as flexibility. To evaluate performance, it uses a weakly labeled facial dataset., in which the result showed that the ARP-GIST descriptor gives the finer details of images hence the accuracy of relevant images are more than previous GIST descriptor.

KEYWORDS: Search-based face annotation, label refinement, machine learning, content based image retrieval, feature descriptor.

I. INTRODUCTION

Due to the prevalence of different computerized cameras and the fast development of online networking devices for web based photograph sharing, late years have seen a blast of the quantity of computerized photographs caught and put away by shoppers. An extensive bit of photographs shared by clients on the Web are human facial pictures. Some of these facial pictures are labeled with names; however a significant number of them are most certainly not labeled legitimately. This has confidently investigated of auto face annotation, an imperative procedure that means to expound facial pictures consequently. Traditional face annotation methodologies are frequently regarded as an improved face acknowledgment issue, where unique order models are prepared from an accumulation of well labeled facial pictures by utilizing the managed or semi-managed machine learning procedures. The "model-based face annotation"[1] methods are constrained in a few view points. Traditional face annotation methodologies are frequently regarded as a broaden face acknowledgment issue, where distinctive grouping models are prepared from a gathering of well labeled facial pictures by utilizing the directed or semi-directed machine learning methods. In any case, the model-based face annotation procedures are restricted in a few aspects. First, it is as a rule tedious and costly to gather a lot of human-marked preparing facial pictures. Second, it is for the most part hard to sum up the models when any new preparing information on the other hand new persons are included, in which a serious retraining procedure is typically required. To wrap things up, the annotation/acknowledgment execution regularly scales inadequately at the point when the quantity of persons/classes is important. The digital components user for capturing images as well as assets for sharing these captured images are getting better day by day such as social media tools/sites are also improving quickly. High quality of image is shared on these resources, but lots of time images shared by any user are without label that's why it will become confusing in identifying names of the people in the picture if any specific person knows the person. Face annotation is nothing but labeling facial image. In these days various methods are implemented are implemented for face annotation. For automated face image annotation with no human intervention, auto face annotation is used [10].

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

Face annotation can be also be used for video clips, like annotation of facial pictures from news video is performed after that it displayed on television so that user can identify person in TV [1].

The design platform annotation has more limitations such as, it is more time intensive as well as very costly to collect huge amount of persons branded training facial image. It is more hard to make generalizations the design at the time of addition of new people , in which need of re-training procedure is must as well as last the annotation efficiency is gotten to be lacking when the types of persons is more. To solve this "Auto face explanation" is imperative methodology which quickly offers name to related individual images. This strategy is more helpful to various genuine application for (e.g. Facebook) which annotates pictures updated by the user for taking care of online record and also questions the pictures. Search for platform annotation is utilized for face picture annotation by investigation the World Wide Web (WWW), where a considerable measure weakly labeled facial Images is easily available. The automated face annotation procedure is granted as confounded aim of research based face annotation by utilizing content-based image recuperation connected with investigation assortment of pitifully marked facial pictures on the web [3, 4].

The main aim of search-based face tagging is for allocating proper name to provided facial picture [1]. Search for search based Face tagging structure is information motivated and additionally free from design, which is motivated via look dependent picture annotation techniques for picture annotation. The essential aim of Search-Based Face Annotation(SBFA) is to assign proper associative name brand to image taken as feedback to input query. In this methodology top k most same face images are recovered from information source, as well as individually procedure of explanation is directed on the purposes behind larger part voting arrangement.

In this paper it is utilizing Angular Radial Partitioning GIST(ARP-GIST) descriptor, built on the first GIST descriptor; the proposed ARP GIST descriptor utilizes the convenience of angular partitioning to record the same details of facial pictures. With the DFT change as well as magnitude of its coefficients, the ARP GIST empowers positional invariance of facial structures inside a rectangular block. The proposed procedure keeps up rough spatial layout, as well as offers adaptability in every block, getting a balance in spatial freedom and constraints.

Related work which is required for this research is given in section 2, the implementation details in section 3 where system architecture and modular description is given. Section 4 discusses about the algorithm. In next section 5, discussion about the Experimental setup is given, where 6 discuss the Results and at last provide a conclusion and future work in section 7.

II. RELATED WORK

In [2], author described that for given a facial image, it first retrieves the top n similar objects using CBIR technique and then use their labels for face annotation. For this, it uses a Weak Label Regularized Local Co-ordinate Coding(WLRLCC) technique, which uses the principle of local co-ordinate coding by learning sparse features, and employs the idea of graph-based create label regularization to improve the weak label of the similar facial images. To minimize the running time without introducing much degradation of annotation performance it uses an offline approximation scheme AWLRLCC. A WLRLCC algorithm aims to achieve a new discriminative feature representation and refined label information in a unified learning scheme. In [3], the author proposed the method to improve the face retrieval it uses a technique to utilize automatically detected human attributes that contain semantic cues of the face photos. Due to the popularity of digital devices and the increase of social network sharing services, there are largely growing user photos available. Among all these photos many of them are human faces. Goal of this paper is to point the problem of large scale Content Based Face Image retrieval. It tries to find similar face images from a large image database. It uses a technique to utilize automatically detected human attributes that contain semantic cues of the face photos. For this, it uses two methods, Attribute enhanced sparse coding and Attribute embedded inverted indexing to improve the face retrieval. In [6], the author focused on face recognition using weak supervision in the form of captions. There is a large increase of image and video data available on Internet. One of the challenges is the need for tools that automatically analyze the visual content and enrich it with semantically meaningful annotations. This paper presents methods for face recognition using a collection of images with captions. First it retrieves all faces of a particular person in a data set, after that it establishes the correct connection between the names in the captions and the faces in the images. For both tasks it compares generative and discriminative probabilistic models. For the retrieval task, it also uses the benefit of query expansion. In [7], for mining social images that are often connected with multimodal contents, it used a machine learning technique. In this, it used a novel Unified Distance Metric Learning scheme. Along with the popularity of digital media and high quality mobile cameras as well as the advances of Internet technologies, users can easily upload their images and photos over WWW. Because of this popularity of large number

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

of social images associated with labels are available on social network. Mining social images on the web has become an important research topic. In this paper, it uses a machine learning framework for mining social images that are often associated with multimodal contents. In this, it uses novel Unified Distance Metric Learning scheme. It investigates an emerging retrieval-based annotation paradigm for automated photo tagging by mining massive social images freely available on the web. The basic notion of a retrieval based annotation approach towards automated photo tagging is that similar images would share the common labels.

The goal of this research is to design a search technique for face annotation by using Unsupervised Label Refinement (ULR) approach and Angular Radial Partitioning descriptor.

III. IMPLEMENTATION DETAILS

A. *Introduction:*

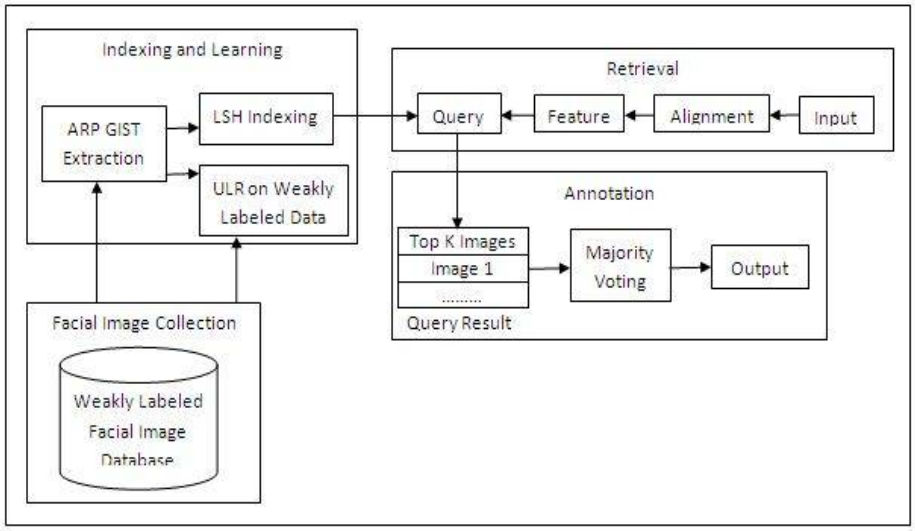

Fig 1: Overview of proposed system

It proposes an effective Unsupervised Label Refinement (ULR) approach for refining the labels of web facial images using machine learning techniques. It formulates the learning problem as a convex optimization.

This system is going to use ARP GIST descriptor [4] which utilizes the efficiency of angular partitioning to capture the finer details of images. The proposed method not only conserve rough spatial layout, but it also provides flexibility in each block.

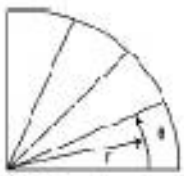

Fig 2: Angular Radial Partitioning

Angular Radial Partitioning is applied in content based image retrieval. It uses both angular and radial partitioning that is similar to the polar coordinate system. Spatial layout is an important thing of a scene image. It carries crucial information related to its category. The original GIST descriptor is calculated on an N-by-N grid. Within every block on the grid, it calculates average intensity value to represent the feature in that block. But it fails to represent spatial structures effectively within a block as the averaging operator often renders different structures identical. It gives the result in mismatch among scene categories.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

For better representation of scene, in this system it is using ARP GIST feature representation technique. In this, it further divides every block into *A* angular bins. It does not only extract the course features of scene but it also extracts the finer features of the scene image.

B. *Advantages of System:*

- The proposed ARP GIST descriptor utilizes the usefulness of angular partitioning to capture the finer details of images.
- With the DFT transform and magnitude of its coefficients, the ARP GIST allows positional invariance of face structures within a rectangular block [4].

C. *Detailed description about proposed framework:*

Proposed Framework is divided into different phases –

- Facial Image Collection
- Face detection and facial feature extraction
- Indexing and Learning
- Retrieval
- Annotation

Facial Image Collection: Here proposed system is going to create facial image dataset. This will contain facial images of different persons. This phase will collect images from Google. According to a name list names of persons are to be collected. Output of this process, it will get a collection of facial images, every facial image is associated with some human names. These facial images are often noisy, which do not always correspond to the correct human name. This type of facial images with noisy names called as weakly labeled facial image data.

Face detection and facial feature extraction: Next is to pre-process web facial images to extract face-related information. Here, it will first detect facial part within the image. Then it aligns detected facial area. After alignment it extracts ARP GIST texture features of the align image for future processing.

High dimensional facial feature indexing: Next step is to index the extracted features of the faces by applying an efficient high-dimensional indexing technique to provide the task of similar face retrieval. It uses the locality sensitive hashing (LSH).

Learning to refine weakly labeled data: In addition to an indexing step, another important step of the system is to engage an unsupervised learning scheme to improve the label quality of the weakly labeled facial images. This process is very important to the entire search based annotation framework as the label quality plays a critical role in the final annotation performance.

Similar face retrieval: It first conducts a similar face retrieval process to search for a subset of most similar faces (typically top K similar face examples) from the earlier indexed facial database.

Face annotation: The next step is to annotate the facial image with a label by employing a majority voting approach. It combines the set of labels associated with these top K similar face images.

D. *Unsupervised Label Refinement By Learning on weakly labeled data*

Let the extracted facial image features denoted by $X \in IR^{n*d}$, where n and d represent the number of facial images and the number of feature dimensions respectively.

The initial raw label matrix is represented by $Y \in [0, 1]^{n*m,}$ which is used to describe the weak label information. Here, Y is often noisy and incomplete.

If $Y_{ij} \neq 0$, it indicates that the i^{th} facial image x_i has the label name n_j. And if $Y_{ij} = 0$, indicates that the relation between i^{th} facial image x_i and j^{th} name is unknown.

The aim of the unsupervised label refinement problem is to learn a refined label matrix $F^* \in IR^{n*m}$. This matrix is expected to be more correct than the initial raw label matrix Y.

To achieve this task of label refinement initially it has the raw label matrix Y and the data examples X. The solution for this problem is to construct a graph-based learning solution based on a main assumption of "label smoothness". The basic idea of this concept is that any two images which share the more similar visual information, the chances are more that they share the same label.

This label smoothness function can be formulated using a loss function $E_s(F,W)$:

$$E_s(F,W) = \frac{1}{2}\sum_{i,j=1}^{n} W_{ij}||F_{i*} - F_{j*}||_F^2 = tr(F^T L F),$$

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

IV. ALGORITHMS

Bisecting K-means Clustering Based Approximation(BCBA):
Input : $C \in R^{m*n}, q_c \in \mathbb{N}, L_{loop} \in N$
Output: clustering result list L_{list}
Add M_0 to L_{list} ; I * M_0 contains all the points * 1
Repeat
Remove the largest cluster M_1 from L_{list}
For i = 0 to t do
Bisect M_1 to $M_1^{(1)}$ & $M_2^{(i)}$
Compute sum of square error ;
Select the result with the lowest SSE I value;
Add $M_1^{(1)}$ & $M_2^{(i)}$ to L_{list} ;
Until $|L_{list}| = q_c$;

V. RESULTS

The system is built using Java framework (version jdk 8) on Windows platform. The Netbeans (version 8.1) is used as a development tool. The system doesn't require any specific hardware to run, any standard machine is capable of running the application.

- *Dataset :*

This research uses images downloaded from internet as well as images which are given in IMDb dataset.

- *Results:*

Fig. 3 shows the Accuracy graph. This shows the accuracy comparison between the GIST and ARP-GIST in which it can clearly shows that the accuracy is increased while using ARP GIST. It gives the more accurate images of person. ARP-GIST descriptor gives 5% more accurate images than GIST descriptor. Fig. 4 shows the memory comparison graph which shows comparison in memory consumed by existing system[1] and proposed system. In this the memory utilization of proposed system is reduced. Proposed system required less memory as compared to existing system. Fig. 5 shows the time comparison graph which shows the time required for the system to generate the output. From figure it will clearly see that the proposed system needs less time to do the operations.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

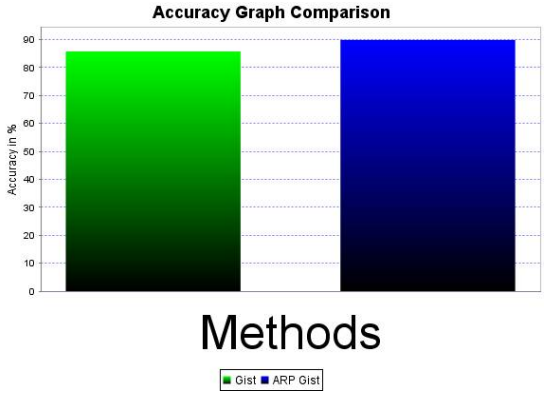

Fig.3. Accuracy Graph

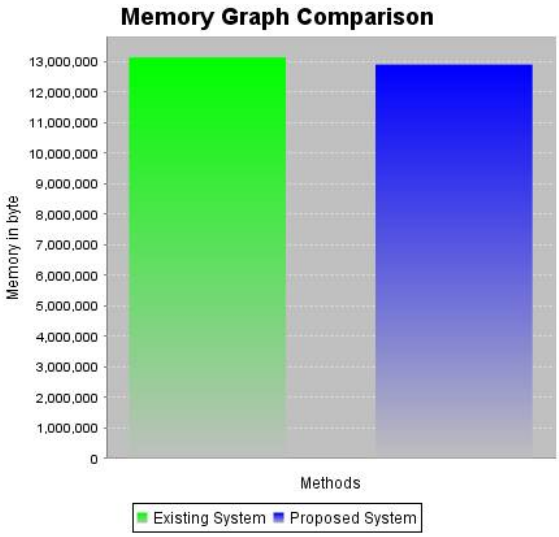

Fig. 4. Memory Graph

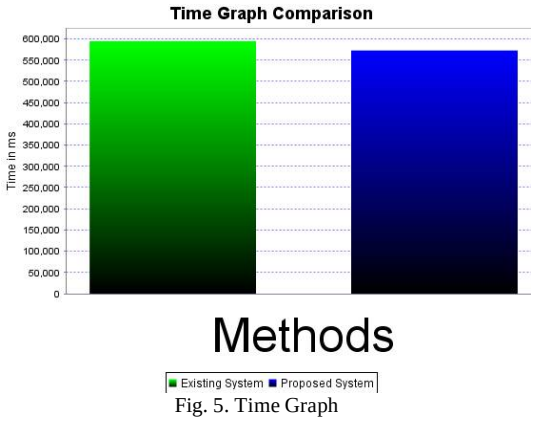

Fig. 5. Time Graph

VI. CONCLUSION AND FUTURE WORK

In this, the Search-based Face Annotation technique is implemented using ARP-GIST descriptor and ULR algorithm. It reveals that this proposed method outperforms the other existing parallel methods. The search-based face annotation system aimed on retrieving the short list of candidate facial images for facial name annotation task. It used Unsupervised Label Refinement(ULR) technique for refining labels. It also used an ARP-GIST descriptor which is built on the original GIST descriptor. The implemented ARP-GIST descriptor makes use of usefulness of angular partitioning to record the finer details of facial images. The implemented results have proved that the accuracy of retrieved images had been increased by using ARP-GIST descriptor method as compared with GIST descriptor. Also it consumed less memory by proposed system.

In future, this work can be extended on algorithm to minimize the time complexity for annotating facial images.

REFERENCES

1. D. Wang, S.C.H. Hoi,Y. He, and J. Zhu , "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation ", IEEE Transaction on Knowledge and Data Engineering, Vol.26, No.1., January 2014
2. Dayong Wang, Steven C.H. Hoi, Ying He, Jianke Zhu,Tao Mei, and Jiebo Luo, "Retrieval-Based Face Annotation by Weak Label Regularized Local Coordinate Coding ", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 36, NO. 3, MARCH 2014..

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2016

3. Bor-Chun Chen, Yan-Ying Chen, Yin-Hsi Kuo, and Winston H. Hsu, "Scalable Face Image Retrieval Using Attribute-Enhanced Sparse Codewords ", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 5, AUGUST 2013
4. Wei Liu, Serkan Kiranyaz and Moncef Gabbouj, "Robust Scene Classification by GIST with ANGULAR RADIAL PARTITIONING", ISCCSP, 2012.
5. J D. Wang, S.C.H. Hoi, and Y. He, Unified Learning Framework for Auto Face Annotation by Mining Web Facial Images,", Proc.21st ACM Intl Conf. Information and Knowledge Management(CIKM), pp. 1392-1401 , 2012.
6. Matthieu Guillaumin, Thomas Mensink, Jakob Verbeek, Cordelia Schmid, "Face Recognition from Caption-Based Supervision ", Intl J. Computer Vision, vol. 96, pp. 64-82, 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 Intl Conf.Web Search and Data Mining (WSDM 11), pp. 197-206, 2011.
8. 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.
9. Wei Dong, ZheWang,William Josephson, Moses Charikar, Kai Li, "Modeling LSH for Performance Tuning", in ACM,2008.
10. 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

BIOGRAPHY

Priyanka Pradip Taware is pursuing her Master of Engineering in Computer Networks from Sinhgad College of Engineering Pune, Maharashtra, India. She has done Bachelor of Engineering in Computer Science and Engineering from Shri Vitthal Education and Research Institute's College of Engineering(SVERI's CEP), Pandharpur (Solapur University) and Diploma in Computer Engineering from Institute of Technology and Engineering(ITE), Malegaon(Bk), Baramati, Pune, MSBTE

Dr. Futane Pravin Ramkrishn graduated from Government College of Engineering, from Amravati University in 1997 and completed M.E. ELEX in Computer Engineering from Pune University in 2002. He received his Ph.D. in PG Department of Computer from Amravati University, Amravati in 2016. He has a research project Sign Language Domain especially Indian Sign language used by deaf and mute people. He works with Sinhgad College of Engineering, Pune. His research interests are in the field of Sign Language, Database and Image Processing. He is Member of Local Inquiry Committee, University of Pune and Member of Staff Selection Committee, University of Pune.