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. 3, Issue 5, May 2015

Facial Image Annotation Search using CBIR and Supervised Label Refinement

Tayenjam Aerena, Shobha T.

M.Tech Student, Dept of CSE, The Oxford College of Engineering, Bangalore, India

Associate Professor, Dept of CSE, The Oxford College of Engineering, Bangalore, India

ABSTRACT: Image annotation plays an important role in image retrieval and management. Auto face annotation aims to automatically assign a face with the denomination of the corresponding person. However, the face recognition presents conundrum in the field of image analysis and computer vision, and it has received a great deal of attention over the last few years because of its applications in different domains. Mining web facial images on the Internet has emerged as a promising paradigm towards auto face annotation. Therefore, a search assistance tool has been proposed, which helps the face image search based on annotation more efficient. In a face image database which has several face images of many personalities and the labels are not placed properly, which means some images are without label and some are without proper label, in this situation annotation based search will fail. To overcome this problem, an effective supervised label refinement (SLR) approach is proposed which helps in refining the labels of web facial images with human manual refinement effort. To further speed up the proposed scheme, clustering-based algorithm is used. Also the proposed system is enhanced for the two name problem, sometimes a person may have two names and hence confusing the system; this issue is solved in this system. The proposed technique is applied to the automated image-annotation task in our experiments, and hence a demonstration would be made to show that our technique is empirically efficacious and promising for mining web facial images. This technique may be applied in real world applications like social media portals (e.g., Facebook) to automatically annotate users' uploaded photos to facilitate the search and management of online photo albums.

KEYWORDS: Texture, Feature Extraction, GLCM, Co-Occurrence Matrices, Image Annotation

I. INTRODUCTION

The aim of image annotation is to assign keywords to images automatically. Therefore, during the image retrieval process, the client may query images based on keywords and thus automatically detecting faces of persons from an image and further assigning keywords based on the corresponding name of the person [1]. To list and retrieve individual photos predicated on a translation of "who" is in the photos, annotation (or tagging) of faces is necessary. Notwithstanding, manual face annotation is a period devouring and conflictingly inconsistent assignment that regularly forces foremost limitations on precise searching through individual photos containing their fascinating persons As an option, automatic face annotation solutions have been formulated.

Retrieval-based face annotation has emerged to be a promising standard in mining gigantic web facial images for automated face annotation approach. An annotation paradigm customarily encounters two key challenges. The principal test is the means by which to proficiently recover a short rundown of most related facial pictures from facial picture databases, and the second is how to viably perform annotation by abusing these related facial pictures and their impotent labels, which are often noisy and incomplete [2].

The fundamental supposition of a retrieval based annotation approach towards automated photo tagging is that similar/identical images would share the common/similar tags. In light of this supposition, one can assail automated photo tagging, a long-standing testing in multimedia and computer vision, by mining an immensely gigantic collection of web/social images [3]. Face annotations can be formulated as a data classification problem from a machine learning and data mining perspective. Face annotation is proximately related to face detection and recognition

In recent years, the indispensability for efficient content-based image retrieval has incremented tremendously in many application areas, mainly in web image relegation and retrieval. Image retrieval from a large image database is currently becoming a paramount and a challenging research topic mainly on its efficacy and speed. Although CBIR has

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. 3, Issue 5, May 2015

been a very active research area since the 1990s, there are still a vast number of challenging issues due to the complexity of image data. These issues are cognate to long-standing challenges among several interdisciplinary research areas, such as computer vision, image processing, image database, machine learning, etc. In a typical CBIR, image retrieval is rudimentary predicated on the visual content such as color, shape, texture, etc.. However, the texture has been considered as one of the most popular and precise features in image retrieval [12]. Though grayscale textures have come up to provide sufficient information to solve many tasks, the color information also needs to be utilized. Therefore, in recent researches color information has come into the picture.

II. Related work

In [1], the authors said auto face annotation plays an important role in many real-world multimedia information and knowledge management systems. Recently there is a surge of research interests in mining weakly labeled facial images on the Internet to tackle this long-standing research challenge in computer vision and image understanding. A novel unified learning framework for face annotation by mining weakly labeled web facial images through interdisciplinary efforts of combining sparse feature representation, content-based image retrieval, transductive learning and inductive learning techniques is learned. A new search- based face annotation paradigm using transductive learning is introduced, and then propose an effective inductive learning scheme for training classification-based annotators from weakly labeled facial images, and finally unify both transductive and inductive learning approaches to maximize the learning efficacy. Its main limitation is model training using training datasets is required and is time consuming.

Multi-Instance Multi-Label learning (MIML) deals with data objects that are represented by a bag of instances and associated with a set of class labels simultaneously [3]. Previous studies typically assume that for every training example, all positive labels are tagged whereas the untagged labels are all negative [7]. In many real applications such as image annotation, however, the learning problem often suffers from weak label; that is, users usually tag only a part of positive labels, and the untagged labels are not necessarily negative. The MIMLwel approach, which works by assuming that highly relevant labels share some common instances, is proposed, and the underlying class means of bags for each label are with a large margin. It is time consuming and requires learning of large datasets. The learning problem often suffers from weak label; that is, users usually tag only a part of positive labels, and the untagged labels are not necessarily negative.

A search-based face annotation framework mainly concentrates on handling the basic issue of improving the label quality and proposed a ULR algorithm [9]. To further enhance the scalability, a clustering-based approximation solution was also proposed, which effectively accelerated the optimization task without presenting much execution corruption. The search based face annotation (SBFA) standard plans to handle the automated face annotation errand by abusing content-based image retrieval (CBIR) techniques in mining huge feebly marked facial pictures on the web. The indexing of facial feature is done with help of LSH (Locality Sensitivity Hashing) technique. After the indexing is completed, the next is to learn and refine weakly labeled data of all images. Face annotation task is performed by majority voting on the similar faces with the refined label. ULR is not accurate and reliable as it depends on majority voting.

III. Problem Statement

The face annotation search using CBIR technique uses the content of the image, specifically the texture of the image, rather than the tags or descriptions associated with the image. The tags or descriptions of an image may contain faulty data and hence making it unreliable and inaccurate. The face annotation using clustering technique is data-driven and model free. The main objective is to assign correct name labels to a given query facial image efficiently in less time. The clustering technique is capable of handling large-scale data. The supervised learning technique also enhances the label refinement technique manually and duplicate names problem is also addressed which is not present in the existing systems.

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. 3, Issue 5, May 2015

IV. PROPOSED FRAMEWORK

A. CBIR Technique

"Content-based" implies that the pursuit examines the contents of the image instead of the metadata such as keywords, tags, or descriptions connected with the image. The expression "content" in this setting might allude to colors, shapes, textures, or any other data that can be obtained from the image itself. Texture is a key segment of human visual observation. Like color, this makes it a fundamental feature to assume when querying image databases. Everybody can perceive texture yet, it is harder to characterize. Dissimilar to color, texture happens over a locale rather than at a point. It is regularly characterized simply by grey levels and thus is orthogonal to color. Texture has qualities such as periodicity and scale; it can be depicted as far as course, coarseness, contrast and so on [1]. And hence, texture is considered as an especially fascinating feature of images and resulting in a vast number of techniques for texture features extraction.

Explanations behind its advancement are that in numerous extensive image databases, conventional routines for image indexing have turned out to be lacking, difficult, and greatly tedious. These obsolete methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with an ordered description, have gotten old. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image.

B. Gray-Level Co-Occurrence Matrix (GLCM)

Texture is a property that signifies the surface and structure of an image or it can be characterized as a regular reiteration of an element or pattern on a surface. Textures may be defined as a complex visual patterns composing of entities or regions with sub-patterns listing characteristics like brightness, color, shape, size, etc. An analysis on texture describes the spatial variation of image pattern based on some mathematical procedures and models to extract information from it. Gray-Level Co-occurrence Matrix (GLCM) is one of the technique utilize for texture feature extraction, proposed by Haralick et al[12].

GLCM evaluates image properties related to second-order statistics which considers the relationship among pixels or groups of pixels (usually two).A two-dimensional GLCM matrix is widely utilized in texture analysis. This is due to the fact that a simple one-dimensional histogram may not be valuable in portraying texture features as it is a spatial property.

The square matrix of a GLCM may reveal certain properties about the spatial distribution of the gray-levels in the texture image. It was defined by Haralick et al. in 1973. It indicates how frequently a pixel value known as the reference pixel with the intensity value i occur in a specific relationship to a pixel value known as the neighbor pixel with the intensity value j. In this way, each element (i,j) of the matrix is the number of occurrences of the pair of pixel with value i and a pixel with value j which are at a distance d relative to each other.

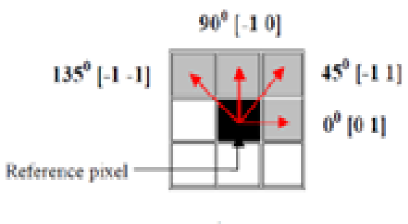

Fig.1. Co-occurrence matrix with neighboring pixels

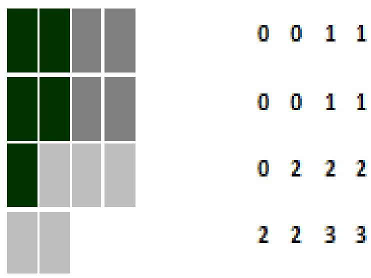

Fig.2. Determination of gray levels of an image

The spatial relationship between two neighboring pixels can be determined from multiple points of view with different offsets and angles, the default one being between a pixel and its immediate neighbor to its right. There are 4 possible spatial relationships (0°; 45°; 90° and 135°). Therefore, considering one neighboring pixel (d=1) along four

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. 3, Issue 5, May 2015

possible directions as {[0 1] for 0°; [-1 1] for 45°; [-1 0] for 90° and [-1 -1] for 135°]}. Every component of the GLCM is the quantity of times that two pixels with gray tone i and j are neighborhood in distance d and direction θ.

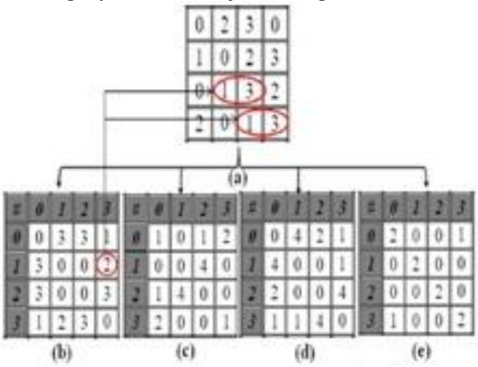

Fig: 3 GLCM construction based on(a)test image along four directions(b)0°(c) 45°(d) 90° (e) 135°

GLCM, Mathematically

For a given image I of size K×K, the elements of a G×G gray-level co-occurrence matrix M_{CO} for a displacement vector **d**(= dx, dy) is defined as

$$M_{CO} = \sum_{x=1}^{K}\sum_{y=1}^{K}\begin{cases} 1, \text{if } I(x,y) = i \text{ and } I(x+d_x, y+d_y) = j \\ 0, \text{otherwise} \end{cases}$$

Consider co-occurrence matrix M_{CO} as $P_{i,j}$..An Entry $C_{i,j}$ in $P_{i,j}$ is a count of no. of times that f(x,y) = i and f(x+1,y+1)=j, where d=1. After obtaining the matrix $P_{i,j}$, various feature can be computed from it.

Haralick defined the GLCM as a matrix of frequencies at which two pixels, separated by a certain vector, occur in the image. Varying the vector used allows the capturing of different texture characteristics. Table 1 shows the feature calculated from the matrix $P_{i,j}$.

Feature	Formula
Energy	$\sum_i \sum_j P^2(i,j)$
Entropy	$\sum_i \sum_j P(i,j) log P(i,j)$
Contrast	$\sum_i \sum_j (i-j)^2 P(i,j)$
Homogeneity	$\sum_i \sum_j \frac{P(i,j)}{1+\lvert i-j \rvert}$

Fig.4. Features calculated from the normalized co-occurrence matrix $P_{i,j}$

C. Semi-automatic image annotation

The semi-automatic image annotation uses the manual user mediation to correct errors of the automatic methods. The automatic methods here may refer to an application or processes. The correction done by the user is based on similarities to a search with relevance feedback. The framework is composed of four parts (see Fig. 3.2): an application, the human collaboration module, a block to update the human facial database and a technique for automatic image annotation.

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. 3, Issue 5, May 2015

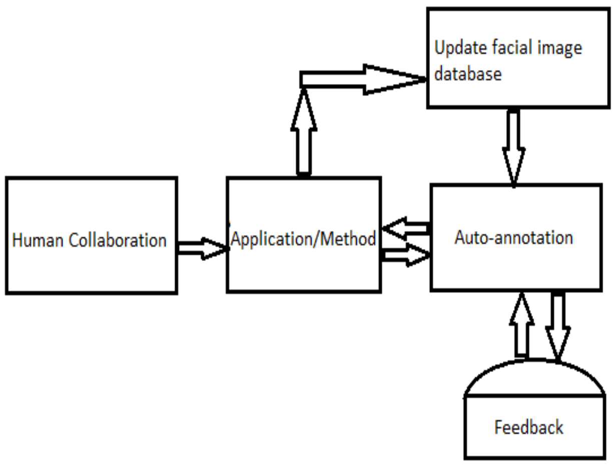

Fig.5. Block diagram of the semi-automatic annotation method

D. System Architecture

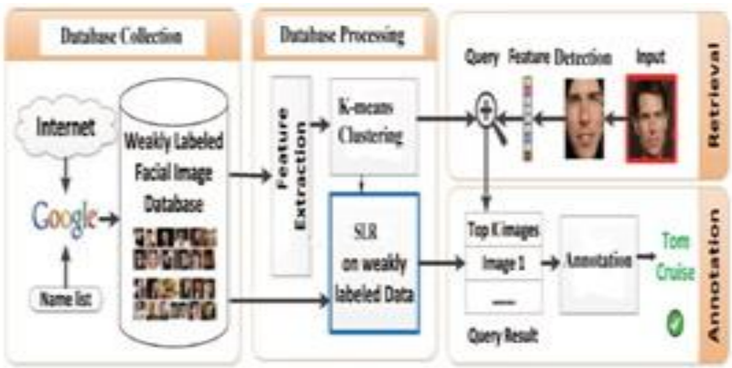

Fig. 6. System Architecture

The system consists of a weakly labeled facial image database, where the images are primarily assumed downloaded from the World Wide Web (WWW). The input query to the system issued by the user is an image of a person, whose annotation is to be done. Using face detection technique namely OpenCV, the face of the person is detected from the image. This system uses the Grey Level Co-occurrence Matrix (GLCM) algorithm to extract feature from the face image. The GLCM extracts the texture feature of the face image. Texture is a property that signifies the surface and structure of an image or it can be characterized as a regular reiteration of an element or pattern on a surface. GLCM evaluates image properties related to second-order statistics which considers the relationship among pixels or groups of pixels (usually two). A two-dimensional GLCM matrix is widely utilized in texture analysis. The square matrix of a GLCM may reveal certain properties about the spatial distribution of the gray-levels in the texture image. The square matrix of a GLCM is also known as a co-occurrence matrix. After analyzing the co-occurrence matrix, co-occurrence vector is generated. The feature extraction technique is also carried out on the weakly labeled facial images contain in the face database.

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. 3, Issue 5, May 2015

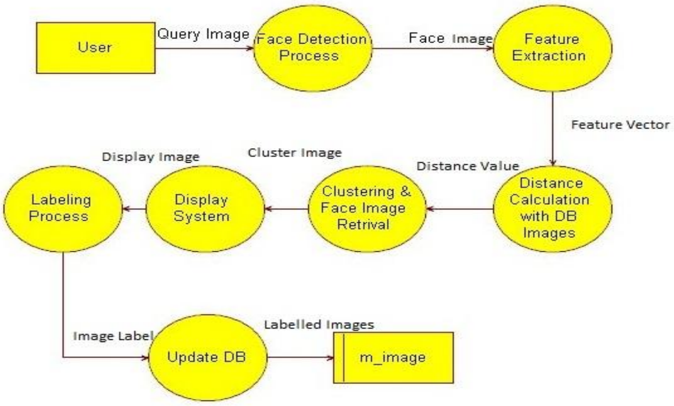

Fig: 7. Data flow diagram of Facial Image Annotation

After the feature has been extracted, distance calculation between the co-occurrence values of the query image and the co-occurrence values of the face database images is carried out. The resulting distance values signify the amount of similarity between the query image and the images in the database. The images are now grouped into a cluster using K-means clustering with the query image as the mean value. The calculation now reveals the top-k most similar images with the query image. After which the images in the cluster are flashed on the display system. The final face annotation process consists of labeling the facial images using human manual effort by exploring machine learning techniques namely Supervised Label Refinement. The SLR uses affordable human manual mediation to correct errors of the automatic methods. The correction done by the user is based on similarities to a search with relevance feedback. After the labeling process, the changes are updated in the face database. This system aims to ameliorate the existing search based face annotation task with inevitable accuracy and more expeditious rate.

VI. PSEUDO CODE

Step 1: Read the image submitted by the user.
Step 2: Detect the face from the image.
Step 3: Calculate the width and height of image
Step 4: Get the packed pixel content of image as an RGB array.
Step 5: Transform the RGB array into grayscale array.
Step 6: Initialize the co-occurrence matrix for gray level.
Step 7: For each pixel compare the gray level of another pixel.
Step 8: Conduct co-occurrence analysis and convert into value.
Step 9: Calculate distance between query image and image from database using co-occurrence value.
Step 10: Cluster the similar images using the distance values.
Step 11: Manually annotate the clustered images.

VII. EXPERIMENTAL RESULTS

The experimental results shown in the figures describe the improved web facial annotation search. Fig.8 shows the original images contained in the face database.

Fig.8. Original images in the database

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. 3, Issue 5, May 2015

Fig.9 depicts the result obtained after using OpenCV for face detection from the original images contained in the face database.

Fig.9. Images after face detection

Fig.10. shows the table consisting of the vectors obtained after using Grey Level Co-occurrence Matrix (GLCM) for feature extraction from each of the images. The GLCM algorithm uses texture feature of the image for image identification, unlike the existing system which uses the label tag and descriptions of the image to identify an image. The square matrix of a GLCM may reveal certain properties about the spatial distribution of the gray-levels in the texture image. The square matrix of a GLCM is also known as a co-occurrence matrix. After analyzing the co-occurrence matrix, co-occurrence vector is generated. The table also shows the Euclidean distance values between the images and the query image.

img_id	img_name	annotate	feature1	feature2		feature3	f1_dist	f2_dist	c_id
2	1_1.jpg	sachin tendulkar		8.0E-4-8.4921-0.2491-8.8...	143 b...	0	0	1.84	2
3	2_1.jpg	sachin tendulkar		0.0354-8.4894-0.3219-8.2...	139 b...	0	0	1.65	2
4	3_1.jpg	sachin tendulkar		0.0027-15.1327-0.2171-8....	142 b...	0	0	4.24	1
5	4_1.jpg	sachin tendulkar		0.0027-15.1327-0.2171-8....	142 b...	0	0	4.24	1
6	5_1.jpg	sourav ganguly		0.0084-13.2015-0.237-8.6...	141 b...	0	0	1.67	2
7	6_1.jpg	sourav ganguly		0.0583-8.0441-0.3231-8.3...	140 b...	0	0	2.03	2
8	7_1.jpg	sourav ganguly		0.0054-9.7311-0.2879-8.6...	142 b...	0	0	1.63	2
9	8_2.jpg	sourav ganguly		0.0157-6.297-0.3492-8.33...	140 b...	0	0	2.56	2

Fig.10. Database showing vectors obtained after feature extraction and distance calculated values

Fig.11 shows the query image search window. The user submits an image as the query, hence using the content of the image as the main search criteria and not labels/tags associated with the image.

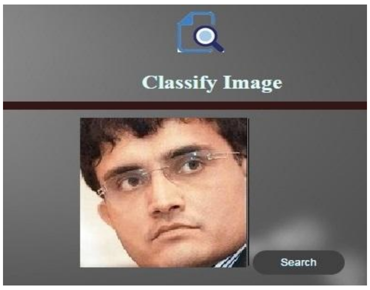

Fig.11. Query Image Search

Fig.12 shows the images retrieved as per the query fed. The images shown here are those images obtained using K-means clustering technique with the query image submitted by the user as the mean value. The result obtained after K-means clustering consists of irrelevant data and hence, Supervised Label Refinement was proposed.

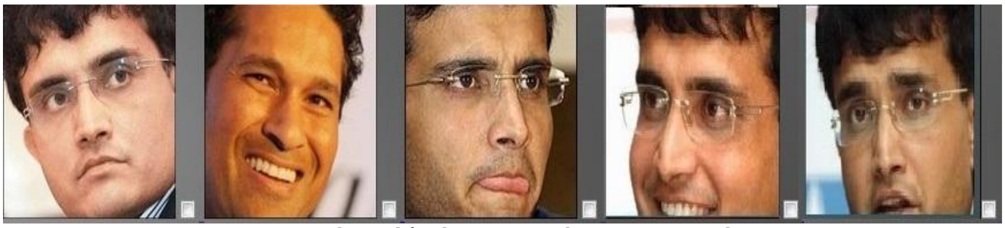

Fig.12. Images obtained for the query search using K-means clustering

Fig.13 shows the final result after SLR has been implemented. The SLR uses human manual collaboration to filter the irrelevant images obtained after K-means clustering. The correction done by the user is based on similarities to a search with relevance feedback. The SLR successfully incremented the accuracy levels of the facial image annotation search.

Fig.13. Final Results after SLR

VII. Conclusion and Future Work

A promising Web Facial Image Annotation framework has been examined which concentrates on the tackling of critical problem of enhancing label quality of facial images in a face database and therefore, proposed a Supervised Label Refinement (SLR) technique. The SLR technique increases the accuracy of the auto-face annotation technique using human manual collaboration. The proposed framework combines the benefits of content-based image retrieval (CBIR) and SLR to result in a satisfactorily reliable and accurate system. A face detection technique called OpenCV has been implemented that detects only the face of the person in the image. A feature extraction technique namely Gray Level Co-occurrence Matrix (GLCM) algorithm had also been introduced in the system. The GLCM algorithm uses texture feature of the image for image identification, unlike the existing system which uses the label tag and descriptions of the image to identify an image. To further increase the efficiency and scalability of the technique, a clustering algorithm namely K-means clustering has also been proposed, which successfully accelerated the optimization task without presenting much performance degradation. In future work, the system can be enhanced for identifying facial images in videos and annotate the image tag to it. The system can also be enhanced to be used for military purposes.

REFERENCES

[1] D. Wang, S.C.H. Hoi, and Y. He, “A Unified Learning Framework for Auto Face Annotation by Mining Web Facial Images,” Proc. 21st ACM Int’l Conf. Information and Knowledge Management (CIKM), pp. 1392-1401, 2012

[2] 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.

[3] 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.

[4] Z. Lin, G. Ding, and J. Wang, “Image Annotation Based on Recommendation Model,” Proc. 34th Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), pp. 1097-1098, 2011

[5] 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.

[6] J. Cui, F. Wen, R. Xiao, Y. Tian, and X. Tang, “Easy Album: 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.

[7] Y.-Y. Sun, Y. Zhang, and Z.-H. Zhou, “Multi-Label Learning with Weak Label,” Proc. 24th AAAI Conf. Artificial Intelligence (AAAI), 2010.

[8] J. Fan, Y. Gao, and H. Luo, “Multi-Level Annotation of Natural Scenes Using Dominant Image Components and Semantic Concepts,” Proc. 12th Ann. ACM Int’l Conf. Multimedia (Multimedia), pp. 540-547, 2004.

[9] Dayong Wang, Steven C.H. Hoi, Member, IEEE, 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.

[10] P.Mohanaiah, P. Sathyanarayana, L GuruKumar, ”Image Texture Feature Extraction Using GLCM Approach,” International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.