ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICRITCSA

M S Ramaiah Institute of Technology, Bangalore

Vol. 5, Special Issue 2, October 2016

Event Image Retrieval using Decision Trees

S. Regina Lourdhu Suganthi[1], Dr. Hanumanthappa M[2]

Research Scholar, Department, Department of Computer Science and Applications, Bengaluru, India [1]

Professor, Department, Department of Computer Science and Applications, Bengaluru, India [2]

Abstract: Efficient Image data storage and tagged Image Archive are quintessential in organizations due to the growing volume of images that are captured during various events that are conducted in organizations. Acquired Images could be indexed instantly using time stamp, image content and image context by means of object identification and association, face detection and matching, object feature extraction and identification. The indexing done using the foresaid criteria transforms each image into a feature vector with domain dependent attributes and is stored in a relational database with the image URL to felicitate query based image retrieval. The domain specific user query is designed to enable the query engine to extract each key attribute as per the weight associated in the proposed model. The efficiency and relevance of an image retrieval system rely on the visualization and the presentation of the search results. Thus the query matching with the feature vector is implemented by building a hierarchical decision tree.

Keywords: CBIR, SIFT, QBIR, HOG, ROI, SSIM.

I. INTRODUCTION

Medical images, Satellite images, Web images, Enterprise images and Personal images are increasing exponentially. These images need to be organized and maintained as an archive with relevant descriptions in order to be used in future. Unlike the text, the images are hard to be interpreted and even more difficult to make machine interpret the images. Image understanding is a discipline that attempts to analyze the images and interpret an image by characterizing and identifying the objects through extracting features that are invariant to geometric and photometric changes[1][2]. Image content and image context must be used to annotate the images. This involves supervised and unsupervised machine learning processes.

Thus, the process integrates machine learning, data mining, information retrieval and computer vision. The complexity of building the image archive and the image retrieval system can be reduced, if the scope of the image data set and the objective of the retrieval is well known. Thus Content Based Image Retrieval System (CBIR) that facilitates Query based Image search based on known objects and scene associations. Image Retrieval system is largely domain specific.

As the focus here is Enterprise image collection, it is accessible by the users within the organizational Intranet. The users are generally classified as browsers, surfers and searcher based on the degree of clarity in intent of search goal[3]. A browser will have no clarity in their search objective and thus the query will be incoherent, a surfer will have moderate clarity and the search will be refined through various stages and finally, a searcher will have absolute clarity in building the query. The retrieval system must influence and accommodate all the three types of users. The query modality designed for this domain is based on fixed attributes that are used to index the images. The attributes used are image time stamp, person of interest, object that describe a scene, strategic objects, indoor/outdoor classification and event context. A general query processing can be categorized based on the system perspective as text, content and interaction. The query processing used for the image retrieval is interactive-composite.

The paper is organized as follows: Section 2 presents the literature survey of the related topics. Section 3 presents the proposed indexing and query processing system, Section 4 discusses the experimental results of an organizational data set and Section 5 concludes with scope of the work.

II. RELATED WORK

The multiplicity of information and complexity of image content is proven to be challenging for the computer to understand and interpret an image. To retrieve an image from an image collection, image content should be automatically understood and interpreted by the machine. Image mining systems [12] can be classified as function-driven systems and information driven systems. Function driven mining systems are based on specific application module functionality whereas information driven mining systems focus on pixel level, object level, semantic level and pattern level features[4] [5]. Thus, Image understanding process is an advanced image processing technique that focus on identification of image content through supervised and un-supervised machine learning techniques.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICRITCSA

M S Ramaiah Institute of Technology, Bangalore

Vol. 5, Special Issue 2, October 2016

Enterprise collection generally include images taken during various events organized during the course of an year. The important feature of an image content could be person/people of interest or product of interest. Face detection and identification is essential in identifying the person of interest and tagging the images.

One of the earlier approach of face recognition based on eigen faces visualized the problem as a template matching problem. The first and foremost real-time face detection system is due to Viola-Jones[11]. The system take the input as faces and non-faces and train the classifier by extracting features to identify faces.

Histogram of Oriented Gradient (HOG) feature could be used to identify objects of definite shape[9]. Locally normalized Histogram of Oriented Gradient (HOG) descriptors provide edge orientations for pixels within a region after dividing the image into small connected regions. Each region is classified into its angular bin based on the gradient orientation. Then the adjacent regions are grouped into blocks. The normalized histograms of these blocks represent the HOG descriptors. These are widely used to detect objects in an image.

Automatic image tagging has been addressed through various techniques namely decision trees and rough sets, rules generated through support vector machine and decision trees[6][10]. The size of the decision trees are independent of the database size and thus scale well for large databases.

III. PROPOSED WORK

Image mining systems are generally categorized as:

- Personal collection
- Domain-specific collection
- Enterprise collection
- Archive collection
- Web image collection

Each of these collections have distinct features that demands different mining design goals[7][8]. The focus of the proposed work is to design an image mining system for an archive of an enterprise collection.

Images taken during various events conducted in an Enterprise must be automatically annotated using image content and event context to enable simple query based retrieval for future reference. Future reference of these image collection is generally for the printing of magazines, newsletters and need based surfing.

The proposed process architecture for automatic annotation of Enterprise event image collection and query based retrieval is shown in fig 1.

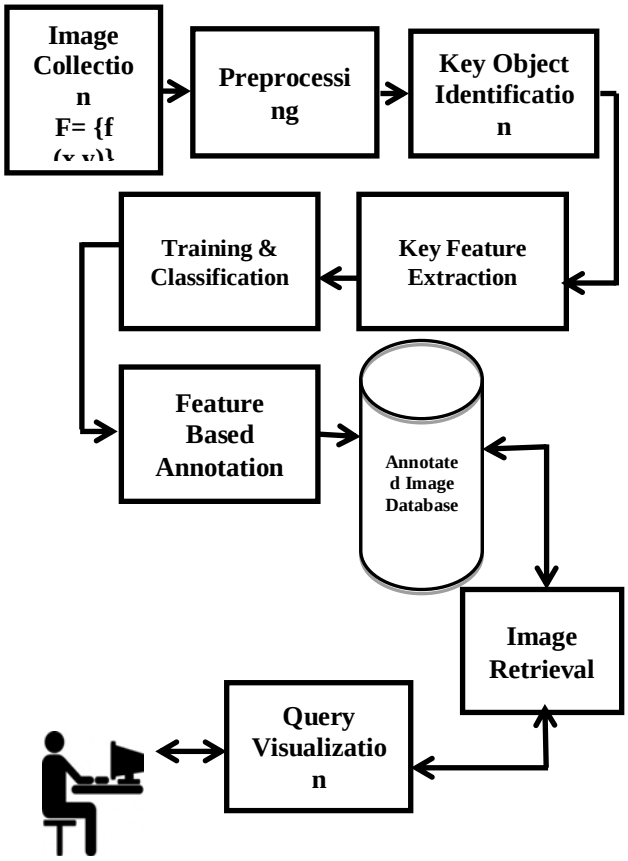

Fig. 1: Process Architecture

A. Image Collection:
The image collection is transferred to a central server of the enterprise instantly after an event. The images are renamed with the time stamp and the event name in a fixed format and stored in the server.

B. Preprocessing:
Each event image collection is generally composed of images taken during fixed time duration and thus includes multiple images of same shots. To eliminate redundant images and noisy images SSIM (Structural SIMilarity) index global measure and noise reduction filters are used. This undoubtedly reduces the storage space.

C. Key Object Identification:
This process solely depend on the domain knowledge of the experts to identify key objects that will form the basis for scene classification. These key objects are segmented using ROI-Region of Interest and stored in a database as reference image set. The object features are extracted and used as training data to identify the respective object classification and automatic annotation of images in that image set.

D. Key Feature extraction:
Various feature elements in an image signify different scene descriptions. Once the key objects are identified the

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICRITCSA

M S Ramaiah Institute of Technology, Bangalore

Vol. 5, Special Issue 2, October 2016

relevant features related to the key objects are trained and detected automatically in the image collection. Face portion of the person of interest with various postures of uniform size M x N is cropped from the chosen face images. This set is used as a training set for feature extraction. To detect objects whose aspect ratio do not vary much, Cascade Object Detector is used to train the classifier by sliding a window over the image in stages. The ROIs that are of different in size and orientation are provided to the classifier, to identify key objects that will form the basis for scene classification. These key objects are segmented using ROI-Region of Interest and stored in a database as reference image set. SIFT features are extracted from these reference images and stored.

E. Training and Classification:

The feature detected from the faces in each image is stored as a vector comprising of the top-left co-ordinates of the face rectangle and the length and breadth of the rectangle. This data is used to crop the face present in the image. The co-ordinates of the 4-dimensional vector is moderately adjusted to obtain the face image of size M x N. This is mapped to a M*N x 1 vector (cv) and the mean of the trained face image set is subtracted. The respective principle component s of this detected face image is obtained and then z=norm (cv-s) is computed. The minimum of z with an appropriate threshold identifies the presence of the trained face in the test image. This information aids in classifying the image into a class that is notified as the presence of the Person of Interest to index. On the same lines the object of interest is identified by using the KNN Classifier and this information is used to classify the image into an associated class that is notified as the presence of the Object of Interest. Domain specific objects and the associated classes could further be used for more accurate classification.

F. Feature Based Annotation:

To ease the retrieval process, the images are annotated and the annotations are stored as a relational table with the image URL. The proposed annotation strategy with the resulting table design is shown in fig 2.

Event Id	Event Name	No of Days

Event Id	Time Stamp	URL

URL	POI	Session	OOI	Location

Event Id	Time Stamp

Fig 2: Annotation Strategy

G. Image Retrieval:

The First level of search is done by constructing a decision tree using the global characteristics namely Event Title and the Time Stamp. The classes for classification are C1, C2, C3, where the class C1 (Event Title is known and the time stamp is known) reduces the search to a very minimal image data set, the class C2 (Event Title is not known and the Time Stamp is known or Event Title is known and the Time Stamp is not known) pools image groups into a single but larger set and then refines the search, on the other hand the class C3 (Both Event Title and the Time Stamp are not known) leave the search to the entire collection. The decision tree is shown in fig 3.

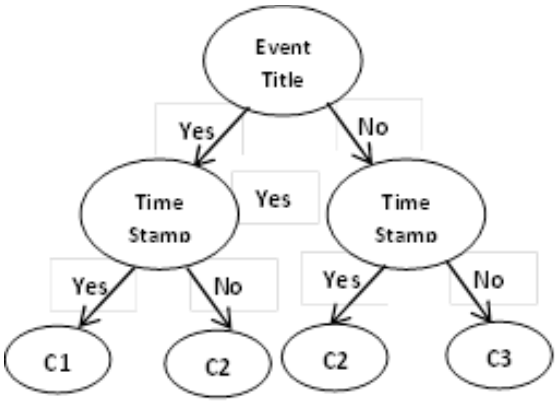

Fig 3 : First Level Decision Tree

At the second level of search the local image characteristics are used as searching strategy with association rule mining and assigned dynamic weight namely Event context, Person of Interest (POI), Object of Interest (OOI) and the Location of the event. The weighted sum is computed from the user query after matching with the image annotation and the search results are presented as the query result in the order of the weight.

IV. EXPERIMENTAL RESULTS

An image set of five events organized in an Institution have been considered. The image set consists of images that were taken during seminar, workshop and cultural events. These images comprise of segments that include people of interest (namely Chief guest, Head of the Institution), Object of interest, lecture presentations and images with various objects that describe the scene that are related to the event context. Each image has been reduced to $1/5^{th}$ of its original size to ease the processing. Redundant images and noisy images are eliminated from the image set during pre-processing using structural similarity index. The resulting images in the image set are converted to gray scale images for processing.

The Vision toolbox of MATLAB software has been used to detect faces and using Region of interest, the objects of interest has been marked and features are extracted to train the classifier to detect the objects from the test images. SIFT feature has been used to detect strategic points in objects to identify indoor and outdoor images. These are used to create a relational table as mentioned in figure 2. The time stamp and the event title for each event with its

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

ICRITCSA

M S Ramaiah Institute of Technology, Bangalore

Vol. 5, Special Issue 2, October 2016

specific event id has been generated to store images in the respective group. Association rules are constructed using scene dynamics.

To identify the event sessions, associations of the following kind are used:

Inaugural Session ⟶ Burning lamp
Product Launch ⟶ Product & Guest
Lecture ⟶ Presence of PPT

Based on the above said features, the relational database has been constructed. The query engine is executed at two levels. At the first level the decision tree shown in figure 3 has been used to identify the group in order to reduce the search time. If the user input is not sufficient to identify the group, the query is refined in successive iterations to identify the group. At the second level, the requested images are listed based on the associated weights.

Query 1:

Level 1 : Event Title : Mobile Computing
Time Stamp : September 2010
Result : Identified Image Set 1

Level 2 : POI : Nil
Session : Lighting of Lamp
OOI : Burning Lamp
Result : 4 Images shown in fig 4.

Fig 4: Query Result

V. CONCLUSION

Annotation schemes and Retrieval techniques are generally built for specific problem domain. Hence, the system must define the annotation scheme based on the objects of interest and association rules based on the problem domain. The current system is limited to few levels of annotation. This could be enhanced and extended to more levels by building more association rules that combine the specific object and the image scene. The current system could also be extended to include drop lists with specific objects and face images to be used as query attributes to enhance the search and retrieval process.

REFERENCES

[1] Barbora Zahradnikova, Sona Duchovicova and Peter Schreiber, "Image Mining: Review and New Challenges", International Journal of Advanced Computer Science and Applications, Vol. 6, No. 7, 2015.
[2] Chesta Agarwal, Abhilasha Sharma, "Image Understanding Using Decision Tree Based Machine Learning", Proceedings of the 5th International Conference on IT & Multimedia, IEEE Explore, Jan 2012.
[3] Datta, R., Joshi, D., Li, J., and Wang, J. Z. " Image retrieval: Ideas, influences, and trends of the new Age", ACM Comput. Surv. 40, 2, Article 5, April 2008.
[4] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.
[5] S. Regina Lourdhu Suganthi, Hanumanthappa. M, "Classification of Event Image Set Using Mining Techniques", Indian Journal of Science and Technology, Vol 9(22), DOI: 10.17485/ijst/ 2016/ v9i22/95297, June 2016.
[6] Kun-Che Lu And Don-Lin Yang, "Image Processing and Image Mining using Decision Trees", Journal Of Information Science And Engineering 25, 989-1003 , 2009.
[7] Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, Larry Davis, "Real-time foreground–background segmentation using codebook model", Elsevier, 2005.
[8] LeLu, Gregory Hager, "Dynamic Foreground/ Background Extraction from Images and Videos using Random Patches", Advances in Neural Information Processing Systems, 2006, 929--936.
[9] Li-li Huang, Jin Tang,Si-bao Chen, Chris Ding and Bin Luo, "An efficient algorithm for feature selection and feature correlation", IScIDE, volume 7751 of Lecture Notes in Computer Science, page 639-646. Springer, 2012.
[10] Manoj P. Patil Satish R. Kolhe, "Automatic Image Annotation Using Decision Trees And Rough Sets" International Journal of Computer Science and Applications, Vol. 11, No. 2, pp. 38 – 49, 2014.
[11] Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Cumputer Vision 57(2), 2004.
[12] Yu-Jin Zhang, "Image Classification and Retrieval with Mining Technologies", Chapter VI, 2009. Handbook of Research on Text and Web Mining Technologies, 2009.