Indian Journal of Science and Technology, *Vol 9(22), DOI: 10.17485/ijst/2016/v9i22/95297, June 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Classification of Event Image Set Using Mining Techniques

S. Regina Lourdhu Suganthi and M. Hanumanthappa

Department of Computer Science and Applications, Bangalore University, Bengaluru - 560056, Karnataka, India;
reginalsuganthi@gmail.com, hanu6572@bub.ernet.in

Abstract

Objective: To develop a system that will optimize storage, automatically classify and tag images using image content and image context by using classification techniques and construct image archive to enable enhanced reference and inference with a simple query. **Methods:** The approach uses both image processing and data mining techniques. Multi modal processing of the Human brain has the ability to distinguish the components in an image and understand the content instantly. To automate this processing ability by the Computer, Computer Vision attempts various techniques. Features that describe the objects of interest are extracted by converting the image into Hue, Saturation, Value (HSV) space. Further segmentation, profiling and thresholding techniques are applied. Finally K Nearest Neighbours (KNN) classifier has been used for classifying the images. Hierarchical decision tree has been built by uniquely identifying classes that could be used for tagging. **Findings:** Two levels of classification have been addressed in this paper. Statistical features are extracted from the H-plane of the HSV color space and with an appropriate threshold obtained from the histogram, images have been classified into indoor and outdoor images at the first level. Further the indoor images are classified using segmentation and profiling, as classes of images with presentation and without presentation. **Applications/Improvements:** Images can be segmented based on more image content that signify the event by watermarking technique and by extracting more features. These segments could further be used to build the decision trees that will identify distinct classes. The depth of the decision tree will enhance indexing, tagging and improve the efficiency of query based image retrieval.

Keywords: Content Based Image Retrieval (CBIR), HSV, KNN, Region Based Image Retrieval (RBIR), Region of Interest (ROI), Threshold

1. Introduction

Academic Institutions and Business organizations conduct events to enhance skill and production. The low-cost digital devices used for image acquisition activity has increased the volume of image data. This poses challenge in terms of storage and access. A system for classification and tagging of event images is a great challenge due to the varying nature of the image content and image context from one event to another. Objects of interest, background details, illumination, position of the imaging devices contribute significantly for segmentation[1] and classification. Once images are segmented and classified in terms of image content and image context, they could be tagged and placed in an image archive to be retrieved even after a longer period of time through a simple query. This will contribute to the documentation of annual activities, newsletters, and compilation activity of the organization. Various techniques namely, Segmentation, object identification, Face detection and matching, identification of indoor/outdoor images, Thresholding, Hierarchical decision tree based classification are used for Tagging the images. Here only two-level hierarchy has been addressed for classification.

** Author for correspondence*

Image content can be described as a composition of visual and semantic content. Visual content itself may be general or specific to a domain. Content Based Image Retrieval (CBIR)[2] and Region Based Image Retrieval (RBIR) improves the performance of retrieval by extracting the features of an image after segmenting the image into regions. Local features such as texture, energy play a significant role in identifying the similarity between images. Structural Similarity Measure (SSIM) is a quantitative measure to automatically predict the image quality with reference to HVS. SSIM compares local patterns of pixel intensities that have been normalized for luminance and contrast. It is used to measure perceived changes in structural information variation.

Segmentation is a very challenging area pertaining to human perception. A digital image can be partitioned into segments that share similar features such as color, texture[3], intensity within a segment. Identification of relevant features and appropriate thresholding will ease the segmentation[4]. Hidden Markov models (HMM) are well known for their effectiveness in image segmentation[5].

Image content in the context of events primarily comprise of people and objects that signify and symbolize various activities. Researchers have done tremendous work in face detection, object tracking and motion structure, whereas a general scene understanding is still under progress[6]. In the past decade many face detection techniques have been proposed. The first and foremost real-time face detection system is due to Viola-Jones The system take the input as faces and non-faces and train the classifier by extracting features to identify faces. Haar features representing different characteristics of a face are applied to the image to extract face features. Face comprise of high-dimensional data, referred to as features[7,8]. Principle Component Analysis (PCA) discover the dimensionality and yield a compact representation of these features. PCA describe the face as a set of functions known as Eigen faces. An enhancement of PCA is Multilinear PCA (MPCA)[9].

To recognize objects from an image, highly distinctive features that are invariant to image scale and rotation are extracted. These features include lines, curves, regions, corners, edge groups and key points. Once the features are extracted the recognition proceeds by comparing these features to a trained database of features using appropriate classifiers[10,11].

The paper is organized as follows: Section 2 presents the proposed framework and the detailed process flow of the framework. Section 3 discusses the experimental results of an event data set and Section 4 concludes with scope of the work.

2. Event Image Classification and Tagging

Image processing techniques are largely application specific. Here, the objective of creation of an archive of image data set taken during an event is for future reference with the perspective of organizational documentation. The key focus here is thus effective storage and efficient query based retrieval.

The images acquired during the events are first pre-processed to eliminate redundant and corrupted images. The pre-processed images are then subjected to feature extraction and classification. Performance of tagging and retrieval distinctly rely on identifying classes that describe unique classes. The uniqueness of the class is described by its features. Thus, for classifying the images into its constituent classes, a random training image set has been chosen. Each image is transformed to HSV space and Statistical features are extracted from this training image set. These features are used by the classifier for training and the rest of the test images are then classified into its appropriate classes. Features that describe the objects of interest and person of interest could then be used for further classification. In this work only two primary levels of classifications are considered. At the first level, the classification is based on indoor and outdoor images. This consideration is due to the fact that, the presentations in events happen generally in indoor auditorium. At the second level the indoor images could further be classified as images with presentation or images without presentation. By associating images with presentation to stage event combined with the person of interest would describe the scene appropriately for tagging. The process flow of the work has been depicted in Figure 1.

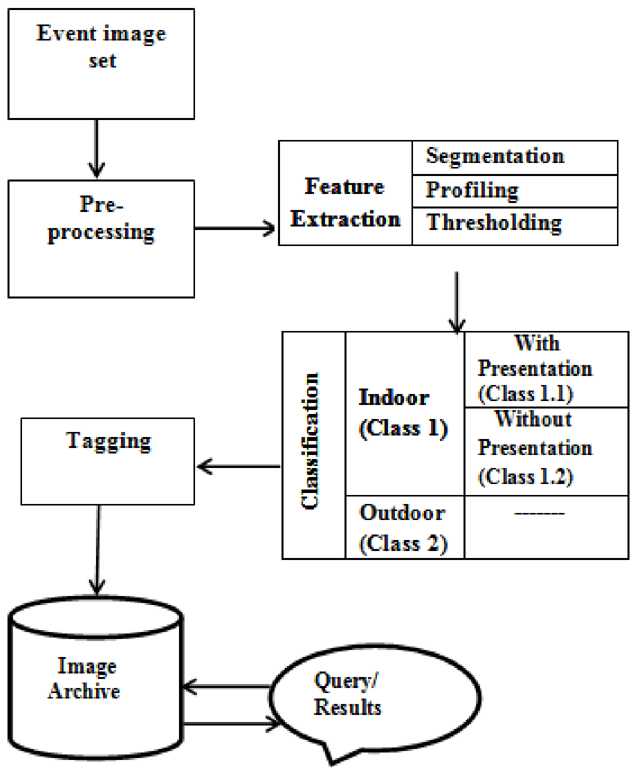

Figure 1. Process Flow Diagram.

2.1 Pre-Processing

Pre-processing in terms of digital images include image enhancement to have a subjective visual appearance, on the other hand pre-processing in the context of data mining include elimination of redundancy. Understanding of the above mentioned facts significantly achieve effective storage and quality archive. The time-stamp, segmentation and structural similarity of the images have been used to eliminate redundancy[12].

2.2 Feature Extraction

To classify and tag the images using the image content, features that describe the objects of interest are extracted. To discriminate the images as indoor and outdoor images at the first level, each image is transformed into HSV space and the statistical texture features from the image histogram of the H plane of the HSV color space namely, the Average gray level (m), Average contrast (σ), Measure of smoothness (R), Measure of skewness (μ_3), Measure of uniformity (U), Entropy (e) are computed. These features combined with an additional features namely the Sum of average gray level, average contrast, image size/M*N, where MxN is the image dimension, image size are used. In the second level of classification, the image is segmented and horizontal profiling technique is used in the segmented region of interest.

2.3 Classification

The classification is carried at two-levels. At the first level, the images are grouped into two classes namely, indoor (Class 1) and the other as outdoor (Class 2) image set. The six statistical features from the H plane and the other two additional features extracted are used to train the classifier. KNN classifier is used to classify the images by identifying appropriate but random training image set. The results of the classifier is refined using a threshold function as given in the Algorithm 1.

Algorithm 1:
Step 1: Set Class 1 = Indoor, Class 2 = Outdoor
Step 2: Read Image Set
Step 3: For each image in the Image set repeat steps 3 to 5
Step 4: Convert the image to HSV Color space
Step 5: Extract only the H-Plane of the image
Step 6: Compute m, σ, R, μ_3, U and e
Step 7: Train the KNN Classifier with a random sub set of the image set features
Step 8: Test the Image set and classify using KNN
Step 9: For each image in the image set repeat step 10 to Step 11
Step 10: Compute S = (m + σ + imagesize/M*N)
Step 11: if (Class =1 and S>300) then
    Class = 2
    else if (Class=2 and S>250)
    Class = 2
    else
    Class = 1
Step 12: End Algorithm

Events in Academic Institutions generally consist of Lecture sessions, Inaugural session and Presentations. Thus the indoor images that are classified as Class 1 are now processed in the second level to be grouped as images with presentation (Class 1.1) and images without presentation (Class 1.2). Mid rectangular image region from each image of Class1 is now segmented and with an appropriate threshold that has been extracted by understanding the histogram of the H-plane in the HSV color space, a binary image is constructed. Complement of this region with horizontal profiling identifies the presence of a power point presentation in the image giving rise to Class 1.1 and Class 1.2. The procedure to categorize the images into these classes is given by the Algorithm 2.

Algorithm 2:
Step 1: Set Class 1.1 = With Presentation, Class 1.2 = Without Presentation
Step 2: Read Class1 Image Set
Step 3: For each image I in the Image set repeat step 4 to 5
Step 4: Convert the image to HSV Color space
Step 5: Extract only the H-Plane of the image
Step 6: Let T1 and T2 are the threshold from the Histogram of the H-Plane
Step 7: Obtain the binary image B
If (I(i,j)>T1 and I(i,j)<T2)
B(i1,j1)=1
else
B(i1,j1)=0
Step 8: Complement B
Step 9: Segment the image
Step 10: Perform horizontal profiling for the segmented region
Step 11: Count the number of horizontal rows with number of zeros
Step 12: If (count > T)
Class1.1 = I
else
Class1.2 = I
Step 13: End Algorithm

2.4 Image Tagging

To tag the images based on image content and event context, in-depth classification that include Chief guest, Celebrity, Guest of honour, Head of the organization and other important personalities and objects such as logo, podium, inaugural lamp that signify different activities in an event must be considered. In this paper we have only taken two levels. Further levels of classification will be addressed sooner with face identification and matching, object detection and recognition by training the classifiers with the features from the region of interest. Hence, at this level, tagging is only based on the above discussed algorithms 1&2, resulting in three classes C1.1,C1.2, C2.

3. Experimental Results

Event E: Two-day National Seminar held at an academic institution. Image set: 743 images. The image time-stamp has been used to classify the image set into Day 1 & Day 2 image sets. Taking advantage of the fact that generally in events multiple shots of the same scene are captured in continuous intervals, similarity check is done between two consecutive image pairs. Each image has been segmented into 3x3 equal segments with padding, each segment pair has been then compared with its SSIM index and one of the two similar image is retained and the other is eliminated. This resulted in 17% of redundant images thus reducing the image set to 702 images. Day 1 Image Set size = 245 Images and Day 2 Image Set size = 457 Images.

3.1 Feature Extraction and Classification

Features extracted from a random set of indoor images and outdoor images are shown in Table 1 with its Class

Table 1. Extracted Features: μ3, U & E are normalized by $(L-1)^2$, where L=256

File name (.JPG)	m (0-255)	σ (0-100)	R (0-1)	μ_3	U	E	Image size (Bytes)	S	Class
IMG_2623	110.06	47.48	0.03351	3.80868	0.02291	5.9636	1051031	333	2
IMG_2624	129.33	64.09	0.05941	0.39581	0.00476	7.8121	975541	291	2
IMG_2625	168.62	72.87	0.07550	-3.29297	0.00872	7.5491	851649	321	2
IMG_2608	38.62	62.27	0.05629	7.39228	0.04285	5.7982	566081	109	1
IMG_2609	51.83	67.11	0.06478	6.61299	0.01938	6.4006	656764	164	1
IMG_2610	48.86	51.10	0.03860	3.92496	0.01129	6.9750	881909	151	1
IMG_2612	63.15	56.44	0.04670	3.13836	0.00803	7.3638	981840	162	1
IMG_2716	121.57	63.90	0.05909	0.95968	0.00563	7.6457	975061	270	2
IMG_2717	110.10	57.04	0.04766	1.91185	0.00576	7.6447	1108796	251	2
IMG_2719	89.48	74.40	0.07845	5.47658	0.00838	7.4261	721057	284	2
IMG_2631	88.39	53.70	0.04246	-0.66765	0.01243	6.7376	760776	245	1
IMG_2635	35.98	39.10	0.02298	1.85481	0.01598	6.5389	819161	130	1
IMG_2636	58.38	53.47	0.04212	3.15794	0.00902	7.1506	917806	149	1
IMG_2638	55.75	51.49	0.03918	2.98762	0.00898	7.1920	934808	169	1
IMG_2639	75.17	50.51	0.03775	1.25421	0.00622	7.4935	1133926	182	1

specified. Few outdoor images (Class 2) are shown in Figure 2(a), (b) and few indoor images (Class 1) are shown in Figure 2(c),(d). Class 1 images are further classified using Algorithm 2 into images with presentation (Class 1.1) and images without Presentation (Class 1.2).

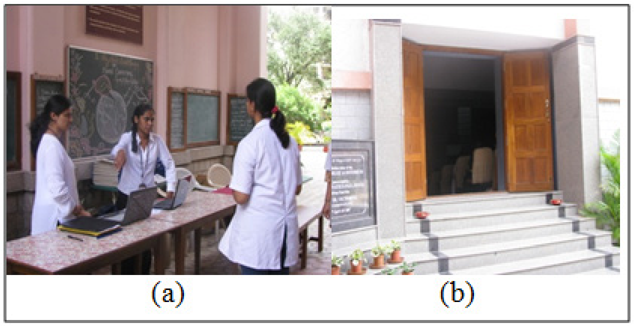

Figure 2. (a) & (b). Outdoor Images Class 2 (IMG_2624.JPG, IMG_2625.JPG).

Figure 2. (c) & (d). Indoor Images Class 1 (IMG 2631.JPG: With Presentation, IMG_2625.JPG: Without Presentation).

Actual Image Statistics of day1 and day2 are given in Table 2a and Table 2b. The validity of the test results are shown in the Table 3 where TP=True Positive, FP= False positive, FN= False Negative and TN= True Negative. Recall and Precision are the basic measures used in evaluating search strategies. Recall is the ratio of the number of relevant images classified divided by the total number of relevant images in the database, while precision is the number of relevant images classified divided by the total number of relevant and irrelevant images.

Table 2a. Day 1 Statistics

Day 1	Class 2	Class 1	
		Class1.1	Class 1.2
245	30	46	169

Table 2b. Day 2 Statistics

Day2	Class 2	Class 1	
		Class 1.1	Class 1.2
457	37	102	318

Table 3. Test Results

	Class 1	
Test Results	TP=127	FP=15
	FN=21	TN=472
	Class 1.1	Class 1.2

4. Conclusion

Voluminous collections of images are increasing challenge in organizations. Classification and Tagging of these image collections would be useful for creating an archive in various perspectives. It finds its application in even documenting written material that is present on leaves (Olai Chuvadi). The classification addressed in this paper could further be extended in depth and breadth by identifying more image content by classifying an image into significant and representational objects of the scene and including image context. This would enhance the search with a simple query.

5. References

1. Ganesan P, Rajini V, Sathish BS, Kalist V, Khamar Basha SK, Satellite Image Segmentation Based on YcbCr Color Space. Indian Journal of Science and Technology. 2015 January; 8(1): 35–41.
2. Ruba AA. Salamah. Efficient Content Based Image Retrieval. Thesis. Islamic University – Gaza, Deanery of Higher Studies, 2010, 82p.
3. Guru DS, Sharath YH, Manjunath S. Texture Features and KNN in Classification of Flower. IJCA Special Issue on Recent Trends in Image Processing and Pattern Recognition. RTIPPR, 2010; 21–29.
4. Hanumanthappa M, Regina S, Suganthi L. Identification and Elimination of Similar Images in an Event Image Set. Proceedings of International Conference on Current Trends in Advanced Computing. International Journal of Computer Applications. 2014 May; 0975–8887.
5. Zhuowen Tu, Song-Chun Zhu. Image segmentation by data-driven Markov chain Monte Carlo. IEEE Transactions on Pattern analysis and Machine Intelligence. 2002 May; 24(5):657–73
6. Antonio Torralba, Aude Oliva. Depth estimation from Image structure. IEEE Transactions on Pattern analysis and Machine Intelligence. 2002 September; 24(9): 1226–38.

7. Ole Helvig Jensen. Implementing the Viola-Jones Face Detection Algorithm. Thesis, 2008; 1–36. ISBN 87-643-0008.
8. Yi-Qing Wang. An Analysis of the Viola-Jones Face Detection Algorithm, Image Processing On Line. 2014; 128–48.
9. Mengmeng Zhang, Zeming Li, ChangNian Zhang, Huihui Bai. Adaptive Feature Extraction and Image matching Based on Haar Wavelet Transform and SIFT. International Journal of Digital Content Technology and its Applications. 2012 April; 6(7):1–8
10. Uma S, Maheswari, Ramakrishnan R. Sports Video Classification Using Multi Scale Framework and Nearest Neighbor Classifier. Indian Journal of Science and Technology. 2015 March; 8(6):1–7
11. Mohd Jabarullah B, Nelson Kennedy Babu C. BPNN-hippoamy Algorithm for Statistical Features Classification. Indian Journal of Science and Technology. 2015 July; 8(14):1–17
12. Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, Eero P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions On Image Processing. 2004 April; 13(4): 600–12.