International Journal of Innovative Research in Engineering Science and Technology

APRIL 2018 ISSN 2320 –981X

A Study on Video Instance Retrieval Techniques

[1]Ms.Priyanka.K.N, [2]Mrs.P.Amudha

1.Department of Computer Science and Engineering,
Avinashilingam Institute for Home Science and Higher Education for Women,
Coimbatore, India
priyankananjappan21@gmail.com

2. CORRESPONDING AUTHOR:
Mrs.P.Amudha
Associate Professor, Avinashilingam Institute for Home Science and Higher Education for Women,
Coimbatore, India
amudharul@gmail.com

Abstract: - Image processing has an increasing growth in various fields. The objective of image processing is to visually enrich or statistically evaluate some aspect of an image not readily ostensible in its original form. In prevailing systems, there are many techniques which are available for extracting information from images. Object detection is the process to find the instances of the real world like faces, animals, buildings, trees, etc. It is used in image retrieval, security and surveillance. There are many object detection techniques are available to identify the objects. This paper presents a study on the two techniques, Window technique and Query adaptive multiple instance learning (QMIL) technique. These techniques extract features from the given input to identify any instance. This technique gives more efficient outcome.

Keywords: Image retrieval, Object detection, Window technique, QMIL technique.

1. INTRODUCTION

Detecting and retrieving an object from a video is very difficult task. The object to be detected is the Object of Interest (OOI). For example, in camera surveillance system detecting or retrieving a particular person in a recorded video, it is very difficult task to detect particular person from a video. Several algorithms are combined in order to overcome this difficulty in retrieving particular objects from a video.

Several mechanisms and algorithms are combined together such as Scale Invariant Feature Transform (SIFT), Boundary Preserving Dense Local Regions (BPLR), Local Binary Pattern (LBP), Query adaptive multiple instance learning (QMIL).

QMIL technique is used for image retrieval in which window technique is used for obtaining accurate result

Local Bit Pattern has a special feature that classifies the texture. It improves the efficiency of the detection. Scale invariant feature transform detects and describes image with its local features. It is applicable for object detection, 3D modeling, video tracking, an identification of gestures. BPLR is used in the segmentation of the frame. It is the combination of 0's and 1's that extracts the color from the OOI and matches with the objects in the video instance.

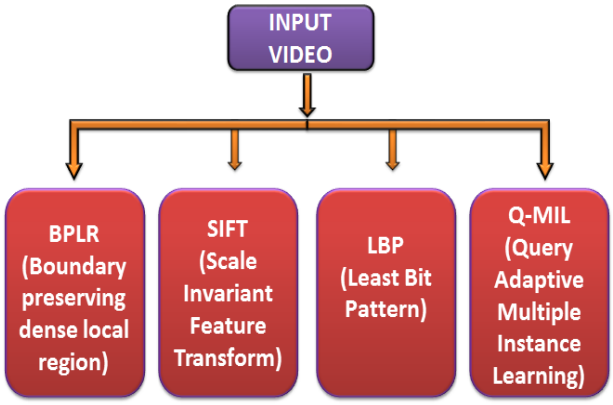


Fig-1: Various algorithms used for instance retrieval.

The four steps for image retrieval are: segmentation, bit pattern, query image and retrieval. Bit pattern matches each frame with the object of interest to give the result. Query image gives accuracy in the output, if there is a change in font or other feature. Finally it retrieves the image with MIL algorithm.

2. Boundary Preserving Dense Local Region (BPLR)

BPLR (Boundary Preserving Dense Local Region) is a technique which is used to extract the color from the video frames. This algorithm helps to explore the instance's features like color, shape, size, texture etc. For example in the figure 2, an instance from a video has detected and the features of the given input object like color, shape, size, texture, etc. will be detected in order to get a clarity image of the input object.

In [1], BPLR technique describes particular region and detects the dense local images. Video frames searching process is done through embedded marketing instance. Search process will be done in offline by the video deliverers. BPLR techniques are used to identify the embedded objects throughout the video frames. It is used to minimize the search time and space for improving computational efficiency.

In [5], in order to produce repeatable shape preserving regions they introduced a dense local detector. Without prior class knowledge its generic bottom up extraction makes it applicable. It improves the distinctiveness and also segments that in to indirectly generate regions.

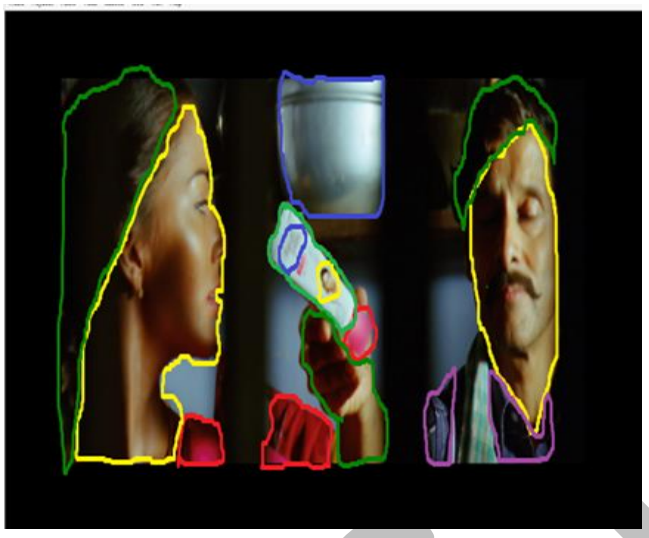


Fig-2: BPLR extracts colors from the video sequence.

In this paper, BPLR technique is used to extract from different size of patches. The combination of SIFT, PHOG (shape), LBP (texture), CIE Lab (color information) provides a robust join features.

3. Scale Invariant Feature Transform (SIFT)

SIFT splits the video frames into matrices. It rotates the frame from 0 to 360 degree. SIFT can extract the key points of the frame. This improves the efficiency and robustness.

Fig-3: OOI (Object of Interest)

In this paper, SIFT [Scale Invariant Feature Transform] describes about visual appearance information for each BPLR. Feature is extracted from different size of patches.

Fig-4: SIFT detects the object at any rotation.

In [1], SIFT is not abundant for an object with shape appearance and video resolution. The task of object recognition get succeeds, yet they are not robust to view point changes. To search partial duplicate image for matching image retrieval spatial coding with SIFT.

In [20] SIFT is mostly used in image processing and it implement the images in different point of view in different computers. It is used because of lighting and to represent images in many angles and also to develop object classifiers in advance. Flip-SIFT is used for object recognition and copy detection. The performance of FSIFT and matching the images are faster. SIFT comes from FSIFT. Compare to FSIFT, SIFT is faster. Accuracy also increased and cost is reduced up to 50%. FSIFT improves continuously. SIFT has existing features of FSIFT and additional features too. Two methods are used.1. Object detection and recognition 2.Detection in video copying. FSIFT gives same output as SIFT but with no flip. Use of FSIFT can improve detection effectiveness.

In [9], the reason for using SIFT gives region description vectors, which are invariant to offline transformation of image. Cluster region emphasis orientation of gradients rather than the position of the particular intensity within the region.

In [3], SIFT converts image data into scale invariant. In this approach features which cover the image over the scales and locations. Features are important for object recognition and identifying the objects in the cluttered background. SIFT features which extract the reference image for matching the image and recognition. These images will be stored in a database. Then the Images are compared with the feature and identify the candidate matching feature. SIFT is used for enabling accurate match and choose it from database. Separation of images are done by using high dimensional vector and then represented by image gradients with the local region. SIFT can extract the image features sensitively which will results to robustness in extracting small objects. Local features are used for matching the accurate object and also it perform for noise and blur. Invariant local features produce best outcome by using various features, which match the images successfully and improves the robustness.

In [10], they used SIFT technique for detecting a particular region of the picture. It identifies the edges and boundaries of the instance to process the image. SIFT is designed to shift a few pixels in the local region. Among the region it penalizes the intensity variations and it is less than the cross-correlation.

In [2], they recognized the object using local SIFT. The previous approaches are improved by using this SIFT features. The last stage of this affine model permits more accurate results. This recognition can be further improved by performance using new features of SIFT that incorporates texture, color, grouping with edges, sizes and offsets. By detecting different features maximum robustness is achieved.

4. Local Binary Pattern (LBP)

In [7], it evaluates the texture image and orders the LBP histogram. Ordered histogram is used to evaluate the difference between the images. Classify the content and then segment an image into regions. In a Local region, rotation invariant is extracted by using LBP.

Using LBP, Global matching schema is used for feature extraction and implements the scheme. It is used to identify the shortest distance between them. LBP features can be used to evaluate the matching distance.

According to [7], LBP is a grey–scale texture operator. LBP based method is classified as feature extraction, histogram creation and classification. In feature extraction non-linear filter is used to form the pattern and measure the pixel. Histogram creation is a time consuming process. LBP histogram is used to calculate rotation invariance by clustering each row into one. To speed up the matching schema, it was proposed by feature set reduction method.

In [4], the conventional operator has long histograms which are sensitive to image rotation, the small spatial area of support, a loss of local textural information and noise sensitivity. The common LBP coding improved and classification results are achieved. LBP has computationally a simple, training-free and data independent method.

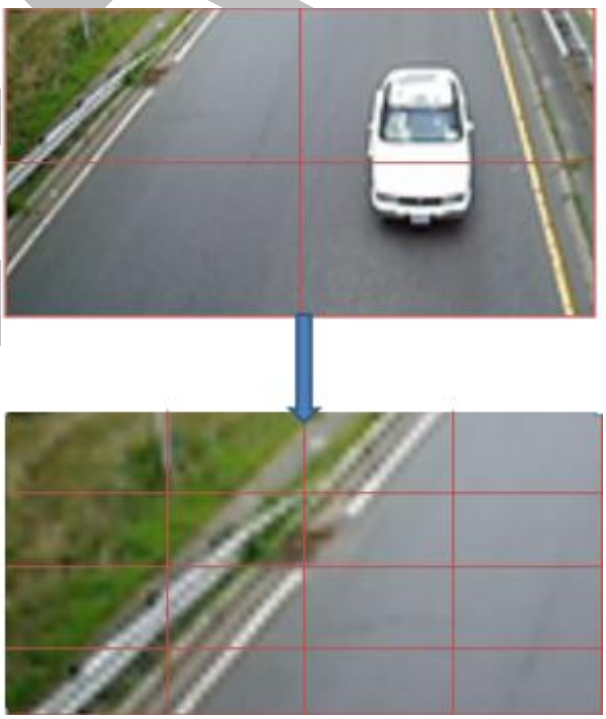


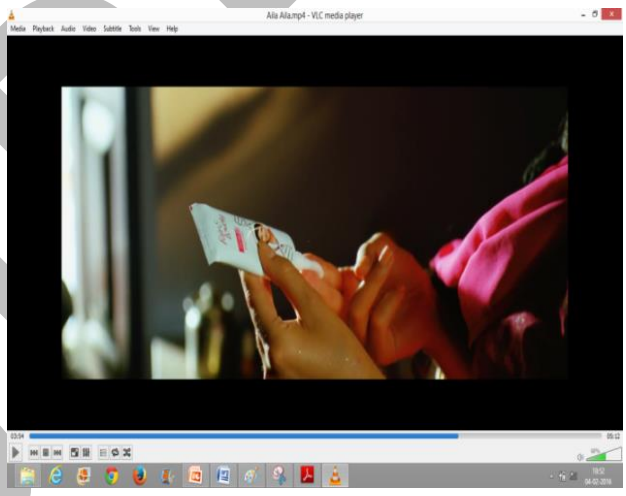
Fig-5: Divides the frame into segments.

In [8], Based on facial image analysis, it has been one of the most popular and successful application in recent years. LBP methodology can be well illustrated in facial image analysis and most of its recent variations are proposed in this area. The information pattern can be encoded by using their discrimination capability. LBP technique is used in facial image analysis through two methods. First it discriminate different classes, Second through the fast processing extracted the raw images. Finally in order to avoid the expensive classifier make use of low dimensionality.

5. Query adaptive multiple instance learning (QMIL).

In [6], MIL algorithm is used to solve object based image retrieval problem. It uses supervised learning but in certain image collection they also get information regarding unlabelled data. Hence MISSL (Multiple Instance Semi Supervised Learning) is widely used. A framework presented for MISSL by using multiple instance and semi-supervised property. Using MISSL, GMIL (Graph based Multiple Instance Learning algorithm) has been developed.

It performs on three sets of data -Labeled data, unlabelled data, Semi-labeled data and it works simultaneously to plot the information in the graph. For solving object based image retrieval problem GMIL is used and superiority is given by their results. GMIL algorithm is reduced to novel standard MIL algorithm (GMIL-M) in the same framework.



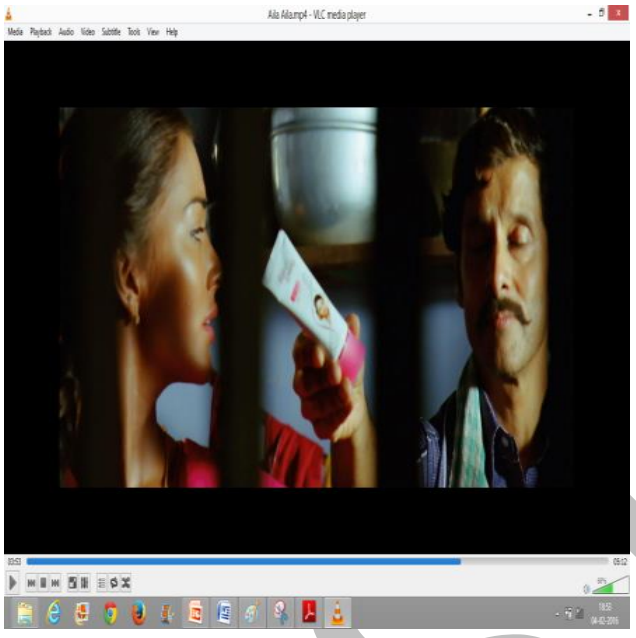


Fig-6: Detection of the objects by features: color, size, and texture.

The content based image retrieval takes single image object as input and deliver similar image from the image repository. The user gives only the object of image as input and rest of the images are irrelevant. The main use of MIL is to classify the bags based on the similarity. We consider semi-supervised property and multiple instance property to get a direct solution for MISSL problem.

In [11], MIL [Multiple Instance Learning] contains an infinite number of instances in a single bag. It takes an instance in the form of audio or image and then it is divided into pieces, which are saved as an instance. In a bag it may have at least one positive instance. For example, if a pedestrian is taken as an image and it can be divided into multiple patches by using sliding window technique for creating positive bags. Patches can be selected as a random for subsampling the bag. Query bag model correctly fit the data. Instances are sampled based on training. It is taken as a sample and produces an optimal solution.

CONCLUSION

The objective of this paper is to give a review about following various techniques BPLR, LBP, SIFT and QMIL. BPLR is used for similar type of video but it can be enhanced by using clustering method. It supports different types of videos to process. MIL can be used for giving many inputs and process that gives more accurate result. MIL applications are used in different domains like computer vision, audition, text, and informatics. The study indicates that Query based multiple instance learning is best for retrieving an instance from a video.

REFERENCES

[1] Ting-Chu Lin_, Jau-Hong Kaoy, Chin-Te Liuy, Chia-Yin Tsaiz and Yu-Chiang Frank Wang. "Video Instance Search For Embedded Marketing"

[2] David G. Lowe. "Object Recognition from Local Scale-Invariant Features"

[3] David G. Lowe. "Distinctive Image Features from Scale-Invariant Key points"

[4] Li Liu, Paul Fieguth, Gangyao Kuang. "Generalized Local Binary Patterns for Texture Classification".

[5] Jaechul Kim and Kristen Grauman University of Texas at Austin. "Boundary Preserving Dense Local Regions".

[6] Changhu Wang, Lei Zhang, Hong-Jiang Zhang. "Graph-based Multiple-Instance Learning for Object-based Image Retrieval".

[7] Zhenhua Guo, Lei Zhang, David Zhang. "Rotation invariant texture classification using LBP variance (LBPV) with global matching". Pattern Recognition, vol. 43, pp. 706–719, 2010

[8] Di Huang, Caifeng Shan, Mohsen Ardebilian, Yunhong Wang, and Liming Chen. "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey"

[9] Josef Sivic and Andrew Zisserman." Video Google: A Text Retrieval Approach to Object Matching in Videos". Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set 0-7695-1950-4/03 $17.00 © 2003 IEEE

[10] Josef Sivic and Andrew Zisserman. "Efficient visual search of videos cast as text retrieval". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE.

[11] Boris Babenko, Piotr Dollar, Serge Belongie. "Multiple Instance Learning with Query Bags"

LIST OF FIGURES