E-ISSN: 2321–9637
Volume 1, Issue 4, November 2013
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

REVIEW ON OBJECT DETECTION IN VIDEO USING SEQUENCE ALIGNMENT AND JOINT COLOR & TEXTURE HISTOGRAM

PALLAVI M.SUNE[1], PROF. A.P. THAKARE [2]
Department of Computer science & Engineering. /Amravati University, India
Department of Electronics & Telecommunication / Amravati University, India
Pallavi.sune@gmail.com, apthakare40@rediffmail.com

ABSTARCT
Video indexing requires the efficient segmentation of video into scenes. In this paper we review and discuss about the video scene segmentation, object tracking, mean shift. The video is first segmented into shots and a set of key-frames is extracted for each shot. Videos are composed of many shots that are caused by different camera operations, e.g., on/off operations and switching between cameras then Object tracking is one of the key technologies in intelligent video surveillance and how to describe the moving target is a key issue. A novel object tracking algorithm is presented by using the joint color texture histogram to represent a target and then applying it to the mean shift framework. In TV-series and movies also indicate that the scene detection method accurately detects most of the scene boundaries while preserving a good tradeoff between recall and precision. The target representation model effectively extracts the edges and corners, which are important and robust features of the object while suppressing the smooth background features.
Keywords: scene detection, object tracking, mean shift.

1. INTRODUCTION

In recent years the extended use of videos in several applications such as internet-TV and video on demand, as well as the thousand TV-series and movies produced every year has led to a significant increase in the availability and the amount of video information. Recently, multimedia information has been made overwhelmingly accessible with the rapid advances in communication and multimedia computing technologies. The requirements for efficiently accessing mass amounts of multimedia data are becoming more and more important. Video scene change detection is a fundamental operation used in many multimedia applications such as digital libraries and video on demand (VOD), and it must be performed prior to all other processes. Video data can be divided into different shots. A shot is a video sequence that consists of continuous video frames for one action. Scene change detection is an operation that divides video data into physical shots. Video indexing, retrieval and analysis seem quite difficult due to this huge amount of data constantly produced. Video scene segmentation provides the most efficient solution so far. [1] However, to proceed with scene segmentation, low level segmentation of the video must be first applied. Real-time object tracking is critical tasks in computer vision applications are particularly important and visual object tracking has drawn increasing interest in recent years. There exist many applications, e.g. video surveillance in airports, schools, banks, hospitals and traffic, e-health cares, and robotics. Object surveillance may provide crucial information about the behavior, interaction, and relationship between objects of interest. Since automated video surveillance systems comprise modules from low-level to highlevel processing, e.g., object detection, tracking, event analysis and classification, robustness and efficiency in each module Among the various tracking algorithms, mean shift tracking algorithms have recently become popular due to their simplicity and efficiency.

The smallest physical segment of a video is the shot and is defined as an unbroken sequence of frames recorded from the same camera. The visual content of each shot of the video can be represented by one or multiple frames, called key-frames.[2] The number of key-frames cannot be predetermined because due to content variation it may be different for each shot. For example for a static shot where there is little object

E-ISSN: 2321–9637
Volume 1, Issue 4, November 2013
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

motion, one key-frame may represent the shot quite adequately, whereas when there is high camera and object motion, more key-frames are needed for a good representation. Several approaches have been proposed for key-frame extraction. In the authors detect multiple frames using unsupervised clustering based on the visual variations in shots. A variant of this algorithm is presented in where the final number of key-frames depends on a threshold parameter which defines two frames to be similar.

Currently, a widely used form of target representation is the color histogram, which could be viewed as the discrete probability density function (PDF) of the target region. [3] Color histogram is an estimating mode of point sample distribution and is very robust in representing the object appearance. However, using only color histograms in mean shift tracking has some problems. [4] First, the spatial information of the target is lost. Second, when the target has similar appearance to the background, color histogram will become invalid to distinguish them. For a better target representation, the gradient or edge features have been used in combination with color histogram. The texture patterns, which reflect the spatial structure of the object, are effective features to represent and recognize targets. Since the texture features introduce new information that the color histogram does not convey, using the joint color-texture histogram for target representation is more reliable than using only color histogram in tracking complex scenes.

The local binary pattern (LBP) technique is very effective to describe the image texture features. LBP has advantages such as fast computation and rotation invariance, which facilitates the wide usage in the fields of texture analysis, image retrieval, face recognition, image segmentation, etc. Recently, Robust Object Tracking Using Joint Color-Texture Histogram LBP was successfully applied to the detection of moving objects via background subtraction. In LBP, each pixel is assigned a texture value, which can be naturally combined with the color value of the pixel to represent targets. We adopt the LBP scheme to represent the target texture feature and then propose a joint color-texture histogram method for a more distinctive and effective target representation. The major uniform LBP patterns are used to identify the key points in the target region and then form a mask for joint color-texture feature selection. In the authors transform this task into a graph partitioning problem. A shot similarity graph is constructed, where each node represents a shot and the edges between shots depict their similarity based on color and motion information. Then the normalized cuts method is applied to partition the graph for each shot, all key frames are merged into a larger image and the similarity between shots is computed by comparing these shot images. A similar approach is presented where a scene transition graph is constructed to represent the video and the connectivity between shots. The task of road detection is difficult for computer vision, because the road appearance is affected by a number of factors that are not easily measured and change over time and space, such as the road materials, illumination, and weather conditions. The more road scenarios where the vision system can provide reliable road information, the more frequently the vehicle can operate in the automated driving mode, the higher degree of autonomy the vehicle can achieve.

2. LITERATURE REVIEW & RELATED WORK:

According to Chunzhao Guo, Seiichi Mita, and David McAllester proposed a Robust stereo vision based drivable road detection and tracking system that was designed to navigate an intelligent vehicle through challenging traffic scenarios and increment road safety in such scenarios with advanced driver-assistance systems (ADAS).[5] An autonomous vehicle through its environments or increment road safety with advanced driver-assistance systems (ADASs), the host vehicle must perceive the structure of that environment, modeling world features that are relevant to navigation. One of the primary perception tasks is to provide a description of the drivable road so that the intelligent navigation system can plan appropriate actions. The detected free road space can provide significant context information to reduce the region of interest, reweight hypotheses, and remove false positives for other functions such as vehicle detections. The novelty lies in the formulation of road detection as a MAP problem in MRF with an unsupervised learning technique, which not only improves the accuracy by ensuring the local consistency of the road classification but enhances the robustness as well by adapting to different road environments. The limitation is there are no clear boundaries between the road and nonroad regions, to solve this problem, we could segment the image based on the color and texture information first so that the lost boundary can be captured between different segments based on the perceptual meaning of the scene. According to Vasileios T. Chasanis, Aristidis C. Likas, and Nikolaos P. Galatsanos proposed the Scene Detection in Videos Using Shot Clustering and Sequence Alignment system the video is first segmented

E-ISSN: 2321–9637
Volume 1, Issue 4, November 2013
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

into shots and set of key frames is extracted for each shot [8]. The visual content of each shot of the video can be represented by one or multiple frames called key frames and the one key frame may represent the shot quite adequately. The video is segmented into shots and the spectral clustering algorithm is employed to extract the key-frames of the corresponding shots. Next, shots are grouped with respect to their visual similarity and labeled according to the group they are assigned. Finally, a sequence alignment algorithm is implemented to identify high dissimilarities between successive windows of shot labels.

3. PROBLEM DEFINE

In this we have to explain that first object tracking is done by observing color pattern but it doesn't track the object continuously. Same case will happen with only observing texture pattern but same problem is happened here also. That's why we are considering joint color texture histogram method to Track the object with scene detection.

4. PROPOSED WORK

The working in the scene Detection in videos converted into numbers of frames using clustering algorithm. Scene with object detection in videos using Sequence alignment and joint color & texture histogram the video is form of number of shot in that shot the numbers of frame created then used Spectral Clustering Algorithm is employed to extract the key frames of the corresponding shots

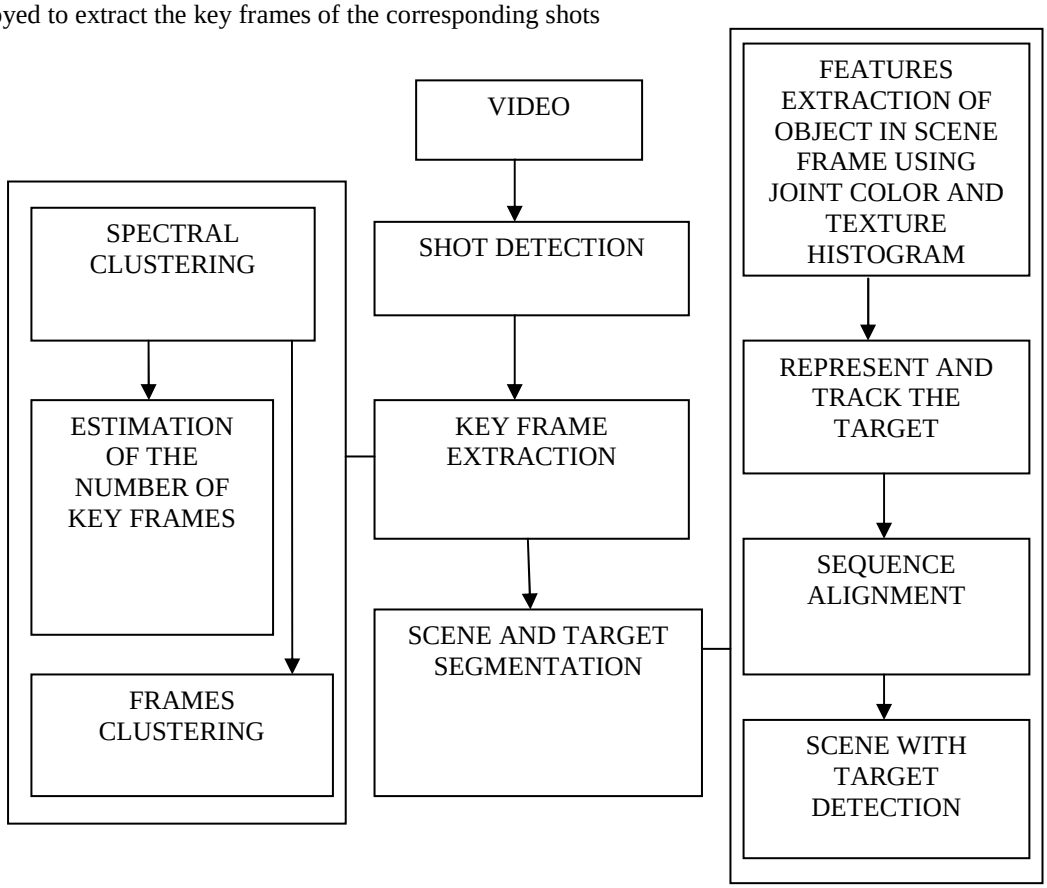

Fig.1. Data Flow diagram of the proposed system.

E-ISSN: 2321–9637
Volume 1, Issue 4, November 2013
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

Next shots are grouped with respect to their Visual similarity and labeled according to the group they are assigned then number of Scene and target segmentation apply local binary pattern (LBP) Schema to represent the target texture features and then propose a features extraction of object in scene frame using Joint color & texture histogram method for a more distinctive and effective way track the target representation. Finally, a sequence alignment algorithm is implemented to identify high dissimilarities between successive windows of shot labels and then scene with target detection in the frames. Now, all the frames with target detection are converted to form same video.

5. IMPLEMENTAION

Object tracking can be defined as the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc. in order to extract useful information. Tracking is a significant and difficult problem that arouses interest among computer vision researchers. The objective of tracking is to establish correspondence of objects and object parts between consecutive frames of video. The object tracking algorithm utilizes extracted object features together with a correspondence matching scheme to track objects from frame to frame. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video.

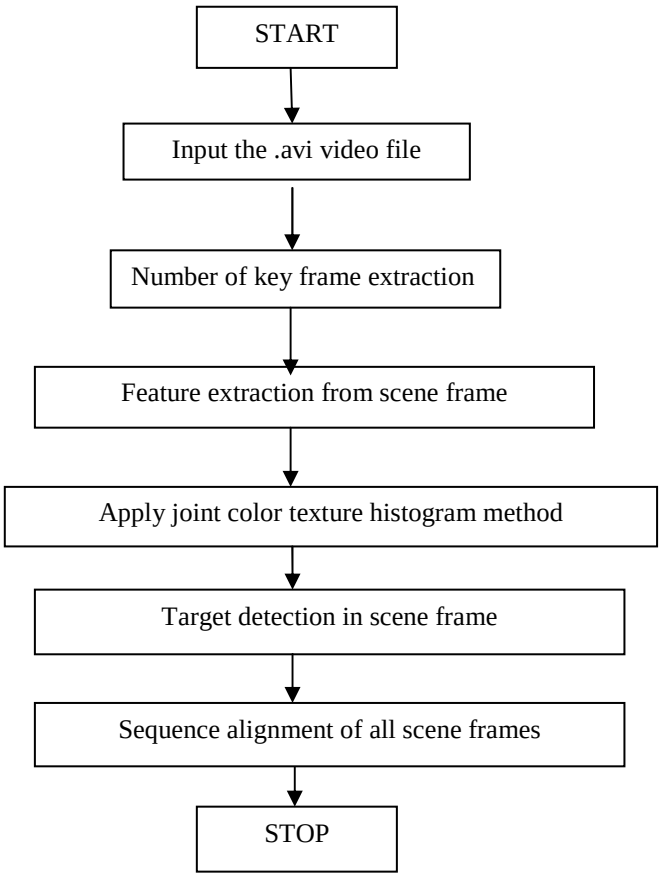

Fig.2. Flowchart diagram

E-ISSN: 2321–9637
Volume 1, Issue 4, November 2013
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

6. APPLICATION

- Traffic monitoring: In some countries highway traffic is continuously monitored using cameras. Any vehicle that breaks the traffic rules or is involved in other illegal act can be tracked down easily if the surveillance system is supported by an object tracking system.
- Robot vision: In robot navigation, the steering system needs to identify different obstacles in the path to avoid collision. If the obstacles themselves are other moving objects then it calls for a real-time object tracking system.
- Automated video surveillance: In these applications computer vision system is designed to monitor the movements in an area, identify the moving objects and report any doubtful situation. The system needs to discriminate between natural entities and humans, which require a good object tracking system.

7. CONCLUSION

From the discussion, it can be seen that object tracking has many useful applications in the robotics and computer vision fields. Several researchers have explored and implemented different approaches for tracking. The success of a particular approach depends largely on the problem domain. In other words, a method that is successful in robot navigation may not be equally successful in automated surveillance.

ACKNOWLEDGMENT

The authors would like to thank Prof. A. P. Thakare for his helpful discussions and valuable suggestions for this paper.

REFERENCES

1 Y. Zhai and M. Shah, “Video scene segmentation using Markov chain Monte Carlo,” IEEE Trans. Multimedia, vol. 8, no. 4, pp. 686–697,2006.
2 Z. Rasheed and M. Shah, “Detection and representation of scenes in videos,” IEEE Trans. Multimedia, vol. 7, no. 6, pp. 1097–1105, Dec. 2005.
3 D. Ramanan, D. Forsyth and A. Zisserman, Tracking people by learning their appearance, IEEE Trans. Patt. Anal. Mach. Intell. 29(1) (2007) 65–81.
4 Z. Kim, “Realtime object tracking based on dynamic feature grouping with background subtraction,” Proc. IEEE Comput. Vis. Pattern Recog., pp. 1–8, 2008.
5 Chunzhao Guo, Seiichi Mita, and David McAllester,”Robust Road Detection and Tracking in Challenging Scenarios B.ased on Markov Random Fields With Unsupervised Learning” IEEE Trans on Intelligent Transportation systems, vol. 13,NO. 3, Sept 2012.
6 Jiafu Jiang and Hui Xiong, “Object Tracking Based on Multi-feature Mean-shift Algorithm” National Conference on Information Technology and Computer Science (CITCS 2012)
7 Zulfiqar Hasan Khan, Irene Yu-Hua Gu, and Andrew G. Backhouse “Robust Visual Object Tracking Using Multi-Mode Anisotropic Mean Shift and Particle Filters” IEEE Trans on Circuits and systems for Video Technology, vol.21, NO. 1, Jan 2011.
8 Vasileios T. Chasanis, Aristidis C. Likas, and Nikolaos P. Galatsanos “Scene Detection in Videos Using Shot Clustering and Sequence Alignment”,IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 11, NO. 1, JANUARY 2009.
9 Jiffng ning, Lei Zhang and David Zhan “ROBUST OBJECT TRACKING USING JOINT COLOR-TEXTURE HISTOGRAM” Journal of Pattern Recognition and Artificial Intelligence Vol. 23, No. 7 (2009) 1245–1263 c_ World Scientific Publishing Company.