International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 12 | Dec 2020 www.irjet.net p-ISSN: 2395-0072

SURVEY ON KEY FRAME EXTRACTION AND OBJECT DETECTION

Varsha Ghait[1], Sonal Karekar[2], Neha Lagad[3], Komal Mohare[4], Meera Thorat[5]

[1,2,3,4]Dept. of Information Technology, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
[5]Professor, Dept. of Information Technology Engineering, Pimpri Chinchwad College of Engineering,Pune, Maharashtra, India

---

Abstract - *Key frame extraction technique aims at retrieving relevant image from a large/small video. Generally it consists of different frames of a video, from which based on few required factors specific frames will be identified for further work. In this paper we are doing the survey on different techniques for key frame extraction and person identification. The identification as well as classification of the diversity of materials is a key visual competence that computer vision systems focus on in recent years. Understanding the identification of materials in distinct images includes a deep process which needs neural networks. After doing the survey on variety of diverse methods of Key frame extraction, we have presented comparative analysis of the methods depending on their important features and result.*

Motivation behind the paper is to develop a system for missing child identification. In India, large numbers of children are reported missing every day. Many of them remain untraced because of many obstacles, so developing such a system is the future scope of our paper.

***Key Words*: Key Frame, Video Summary, Object Identification**

1. INTRODUCTION

Key Frame extraction plays an important role in many video processing applications. In these methods, key frame extraction is usually modelled as a typical clustering process that divides one video shot into several clusters and then one or several frames are extracted based on low or high level features. Due to digital platform video data has been increased rapidly. Therefore, doing analysis on such large quantity videos is very difficult and time consuming. Key frame extraction is a powerful tool that analyse a video contents by selecting a set of key frames based on primary features. Video combines all the form of multimedia information. Key Frame extraction helps to reduce the useless information of video. Key frames are essential to analysis on large amount of video frame sequences. Object detection also plays an important role in finding certain entity in set of frames quickly. It deals with detecting instances of objects based on defined parameters. In this paper different key frame extraction techniques and object identification techniques are described.

2. SURVEY ON KEY FRAME EXTRACTION METHODS & OBJECT DETECTION

In this section, we will discuss different key frame extraction methods and object detection techniques

Table -1: Comparative Study

Sr. No.	Paper Title	Methods Used	Key Points of the paper
1	A survey on key frame extraction methods of a MPEG[1]	In this paper, author proposed a model of perceived motion energy to model patterns in video and a scheme to extract key frames based on this model. Motion is used as an important feature in this. The motion is considered as the most important quality for the presentation of actions in videos and used for determining key frames. The model is a combination of both motion characteristi	• Key frame extraction • Video summarizatio n Feature extraction

		cs and motion intensity. There are two motion vectors which is used for motion compensation in MPEG stream in each macro block of 'B' frame is called as motion vector field (MVF).	
2	Comparative study of motion detection methods for video surveillance systems[2]	In this paper, the CDnet video dataset* were used as a standard that includes many challenging problems, which varied between basic simple scenes to complex scenes affected by bad weather and dynamic backgrounds. Testing of twelve change detection methods was done, ranging from simple temporal differencing to more advanced methods, and the results were precisely evaluated by using several performance metrics	• Background subtraction • Moving object detection • Non-human Object Detection
3	Background subtraction using running gaussian average: a color channel comparison[3]	In the paper, this method is adapted to the(R,G,B) colour space, the results were compared with the single component intensity colour space	
4	Video Based Action Detection and Recognition Human using Optical Flow and SVM Classifier[4]	In this paper, the SVM classifier is used on KTH dataset and results are tested 600 videos performed by 25 persons. This paper states the action descriptors use for modelling behaviour for normal human activity detection from sequence of video. The whole process is mainly divided into tracking, human detection ,activity detection, evaluation of accuracy and recognition	• Detection of Activities • Action Recognition • SVM classifier
5	Faster R-CNN: Towards Real-Time Object Detection with Region Proposal	In this paper, Faster R-CNN: Towards Real-Time Object Detection with Region	Faster RCNN : • RPN • Detection Network • Roi pooling • Anchor Boxes

	Networks[5]	Proposal Networks and Region Proposal Network (RPN) and its working with explanation of work of each layer	

Some Results from reference paper:

Method	Important Feature	Result
Perceived motion energy model	Motion	This method is faster because motion data can directly be used in the analysis of motion.
Frame Difference Method	Motion	When this method is combined with the Running Gaussian Average Method, it helps to fill up the small gaps that may have been contained within the moving object.
Classification Methods	Feature Vector	SVM classifier does the labelling and categorise with different features.
Object Detection Method	Region Proposal Network(to generate region proposals)	Test-time of image is much faster of Faster RCNN because the selective search algorithm is replaced by a separate network for region proposals.

Fig -1: Important features from reference paper

2.1 METHODS IN DETAIL

Perceived Motion Energy Model: In this model, Tianming Liu has proposed a model of perceived motion energy to model patterns in video and a scheme to extract key frames based on this model. Motion is used as an important feature. The motion is considered as the most important quality for the presentation of actions in videos and hence should be used for determining key frames. The model is a combination of both motion characteristics and motion intensity.

Video key frame selection: Key frame selection process is considered as the basic unit for video structural analysis. It eliminates the redundant frames for better representation of the video. It has a great significance and it provides user with better representation of the entire video clip. Although in best of our knowledge, there is no universal evaluation metrics which can evaluate the methods of key frame selection. The methods should be highly efficiency, low computational complexity and must be minimal and good representative of the entire video as much as possible. In our proposed work, key frame evaluation metrics employed are Fidelity (FID) and Video Sampling Error (VSE).

Running Gaussian average method of background subtraction: Running Gaussian Average method was followed by an automatic thresholding operation, in an effort to predict region changes and remove small changes in light. The RGA method outperforms the MoG method, with higher Recall, low FNR values and higher F-measure in almost all categories. This is owed to the large number of parameters required to set for the MoG algorithm, which differ with the challenging conditions presented by a video (day/night, indoor/outdoor, complex/simple background, with/without noise). Thus, the RGA method seemed to be sufficient. In this method, the background is modelled by fitting a Gaussian distribution (μ, σ) over a histogram for each pixel4; this gives the probability density function (pdf) of the background.

SVM CLASSIFIER: In this paper [4], the SVM classifier is used on KTH dataset and results are tested 600 videos performed by 25 persons. This paper states the action descriptors use for modelling behaviour for normal human activity detection from sequence of video. The whole process is mainly divided into tracking, human detection, activity detection, evaluation of accuracy and recognition. Methodology discussed in paper [4], SVM classifier takes the feature vector as a input.. The SVM classifier does the labelling and categorise with different classifiers.

3. CONCLUSIONS

In this paper we have discussed various methods for key frame extraction and object identification. All the techniques are having their own features according to the application. From this survey the best tools/technologies/algorithm will intend in project and hence by adding technical

Advancement social help can be provided through the system for automated search and also enabling instant as well as efficient ways to society.

REFERENCES

[1]. Shaving Pandey; Prashant Dwivedy; Sunil Meena; AnjaliPotnis "A survey on key frame extraction methods on a MPEG video" ,ieee2017

[2].Kamal Sehairi, Fatima Chouireb, Jean Meunier, &quot: "Comparative study of motion detection methods for video surveillance systems", JEI, 2017

[3]. Cerman Martin, "Background subtraction using running gaussian average: a color channel comparison", ieee, 2018

[4].Jagadeesh B, Chandrashekar M Patil," Video Based Action Detection and Recognition Human using Optical Flow and SVM Classifier", ieee, May 20-21, 2016

[5]. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun: "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks"