Survey on Various Approaches used in Video Shot Boundary Detection

Anantrao Metange, Dr. Shrikant Chavate, Rupali Chopade

Abstract— The shot boundary detection is the crucial task in the field of video processing for the task to achieve the effective results in the applications like, video retrieval, browsing and indexing. Looking into the research carried out in the last couple of decades, the precise detection of shot boundary is still a challenging task when the video frames are affected with illumination effects of object motions. Also as the study suggest there are two broad classed in detection type as abrupt transitions and gradual transitions. The abrupt transitions are easier to locate but the gradual transitions are difficult to detect due to the special editing effects embedded in the videos. This paper elaborates the details about the work done in the detection of shot boundary by various researchers.

Index Terms—SBD, WHT, GT, AT

I. INTRODUCTION

Recent years have seen a notable increase in the field of multimedia's innovations in technology. Video data on the internet has undergone a significant evolution and appears to be growing daily. Video is the most often applied data type on the internet since it is the most popular medium for exchanging information in a variety of sectors, including science, sports, entertainment, and education. Due to the excessively high increase in the quantity of films and the size of the repository, ongoing monitoring is important for the efficient arrangement and operation of this material [1].

As of February 2020, data show that over 500 hours of video are uploaded every minute. It could be challenging to select the ideal video clip from such a big collection because there are so many videos available. The manual method of searching through videos for pertinent content and evaluating it takes a lot of time and effort [2]. This issue prompted research efforts to create a video retrieval solution [3]. Partitioning a video into its constituent parts is a crucial first step in understanding its composition. This is composed of many levels, including shots, frames, and scenes. The basic framework of a video is shown in Figure 1.
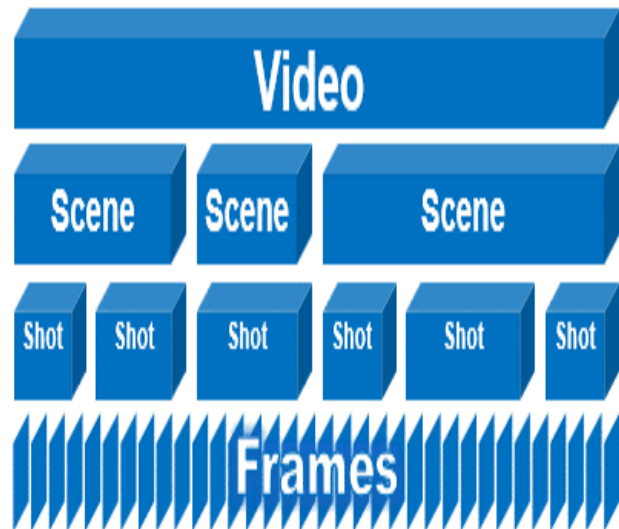

Fig. 1. Basic Video structure

Frames are the fundamental building blocks of a video stream. Scenes are assembled together from several shots, which are created by joining multiple frames together. These scenes make up the whole video [4]. A sequence of connected successive images taken with a single camera that show continuous motion in both space and time is referred to as a single shot [5]. An essential step in carrying out tasks like browsing, indexing, and video retrieval is shot boundary detection (SBD). There are two different kinds of shot boundaries: gradual transition (GT) and abrupt transition (AT). Whereas a gradual transition allows for gradual frame changes with varying duration and velocity according to the needs of the video creators, an abrupt transition in a video shows abrupt and absolute changes in frames. Special editing effects like dissolve, wipe, fade-out, and fade-in are also displayed in the slow transition. The method of detecting shot boundaries involves multiple stages, including the extraction of features, the creation of continuity functions, and the identification and categorization of shot boundaries. The SBD procedure concept is shown in Figure 2. Preprocessing extracted frames may be part of the detection process. The most favored qualities from the state of the art are colour, edge, and texture; these features are then employed as input for the continuity signal's development. This provides a measure of the degree of resemblance or dissimilarity between the frames, and threshold can be used to detect shots.

The literature that was already in existence recommended using motion, edge-based algorithms, histogram differencing, and pixel differencing for SBD. The easiest way to determine the degree of dissimilarity is to use the pixel difference

approach. The method involves computing and comparing the pixel difference between consecutive frames. However, its efficiency at detecting gradual transitions is shown to be lower. The edges of the objects in the frames are located using the edge-based method. The goal of the histogram-based techniques is to lessen sensitivity to movements of objects or cameras. Motion vectors are used in the motion-based method.
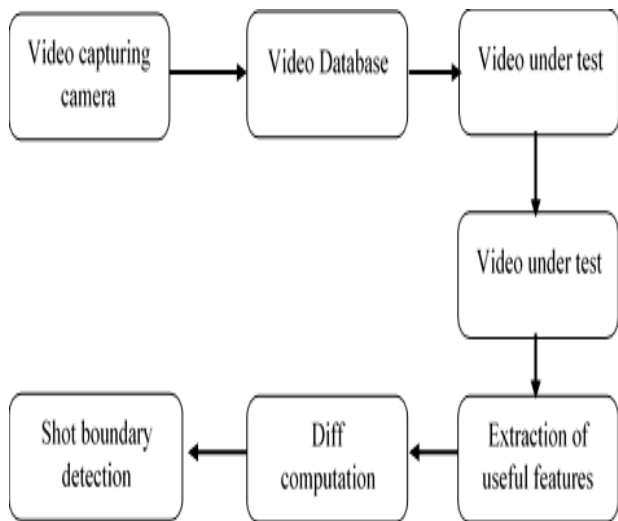

Fig. 2. Basic steps in Video Shot Boundary Detection

II. REVIEW OF LITERATURE

The existing study demonstrated the significance of accurate identification of shot boundary. As there are challenges such as object / camera motion, illumination effects, flicker and flashes. The literature demonstrated the details about remedy on the mentioned challenges and the detection of abrupt and gradual transitions.

Hato, Eman Abdulmunem, Matheel E.et al.[6] suggested the use of Speeded-Up-Robust-Features (SURF) features which extracted from half number of frames of video file to reduce the execution time. The similarity has been computed using a number of matching features and then compared to a predefined global threshold for the abrupt shot detection. This approach claimed a significant performance metrics values of precision, recall rate and F1 score.

Gao, Yin, Lai, Yi, Liu, Ying et al. [7] proposed an approach of visual perception for the shot boundary detection. This also claimed the significant values of performance metrics.

Zhou, Shangbo, Wu, Xia, Qi, Ying, Luo, Shuyue, Xie, Xianzhong et al.[8] explain the SBD approach uses the the image color feature, and local descriptors and combine a kind of motion area extraction algorithm to achieve shot boundary detection. This alos uses the color histogram and SURF features for cut detection , and the gradual transiton detection taken place usong the motion arera extraction, SIFT and slice matching.

Rashmi, B S, Nagendraswamy, H S et al.[9] uses the approach extracts block based Mean Cumulative Sum Histogram (MCSH) from each edge gradient fuzzified frame as a combination of local and global feature. The relative standard deviation (RSD) statistical measure has been applied on the

obtained MCSH to detect abrupt and gradual shots in the video.

Chavate, Shrikant, Mishra, Ravi et al. [10] used the transform method and the deep learning approaches for the detection of cut and the gradual transitions, also provided the remedy over illumination effects.

Chakraborty, Dipanita, Chiracharit, Werapon, Chamnongthai, Kosin, Charoenpong, Theekapun et al.[11] explain the detection technique for wipe transition which is a type of gradual transition, here the video is first segmented and further application of 3D convolutional neural network for RGB stream and optical flow velocity for motion stream network were used. Then classification of wipe and no-motion transitions were done.

Comparative analysis of different methods for video shot boundary detection:

From the available literature, there are several techniques that claimed the accurate detection of abrupt and gradual transitions, also in many cases the effects of illumination and the camera motion also considered during the detection procedure.

Mostly in recent literatures the deep learning approaches were mentioned that found effective as far as the classification of gradual transitions types are concerned.

The following table shows the comparative analysis for the several methods mentioned in the literatures.

Table 1: Comparison of different SBD approaches

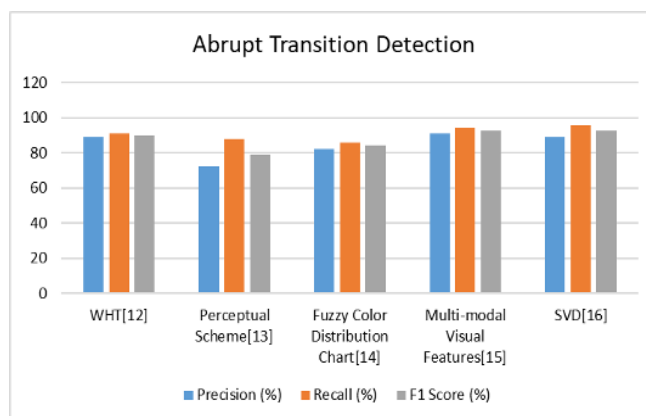| Methods	Abrupt Transition			Gradual Transition		
	P (%)	R (%)	F1 Score (%)	P (%)	R (%)	F1 Score (%)
WHT[12]	89	91	90	87	85	86
Perceptual Scheme[13]	72.1	88	79.2	72.3	63.3	67.5
Fuzzy Color Distribution Chart[14]	82.2	86.1	84.1	62.4	76.5	68.7
Multi-modal Visual Features[15]	91.1	94.2	92.8	73	75.1	74
SVD[16]	89.2	95.9	92.8	68.9	78.9	73.6

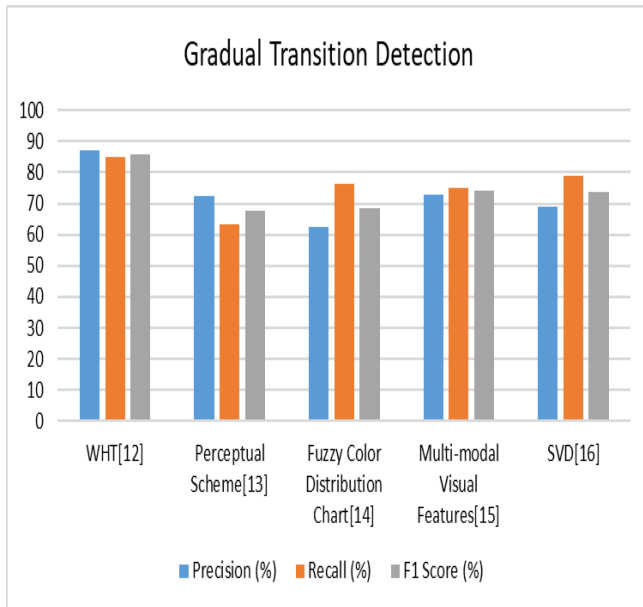**Fig. 3. Comparison of abrupt transition detection using different approaches**

Fig. 4. Comparison of Gradual transition detection using different approaches

III. CONCLUSION

From the available literature and the recent research in the area of video shot boundary detection, we concluded that there are several challenges like illumination and object motion are leading to faulty results. The Gradual transitions are still difficult to detect accurately. The feature extraction and classification techniques are responsible for precise detection. Deep learning approaches also helped to identify the correct gradual transition effects.

ACKNOWLEDGMENT

REFERENCES

[1] H. Shao, Y. Qu, and W. Cui, "Shot Boundary Detection Algorithm Based on HSV Histogram and HOG Feature," 2015.

[2] A. K. Prabavathy and J. D. Shree, "Histogram difference with Fuzzy rule base modeling for gradual shot boundary detection in video cloud applications," Cluster Comput., 2017, doi: 10.1007/s10586-017-1201-0.

[3] M. Thilagam and K. Arunish, "Content-based image retrieval techniques: A review," Proc. IEEE Int. Conf. Intell. Comput. Commun. Smart World, I2C2SW 2018, pp. 106–110, 2018, doi: 10.1109/I2C2SW45816.2018.8997161.

[4] S. Chavate and R. Mishra, "Efficient Detection of Abrupt Transitions Using Statistical Methods," ECS Trans., vol. 107, no. 1, pp. 6541–6552, 2022, doi: 10.1149/10701.6541ecst.

[5] W. Yan, H. She, and Z. Yuan, "Robust Registration of Remote Sensing Image Based on SURF and KCCA," J. Indian Soc. Remote Sens., vol. 42, no. 2, pp. 291–299, 2014, doi: 10.1007/s12524-013-0324-x.

[6] E. Hato and M. E. Abdulmunem, "Fast Algorithm for Video Shot Boundary Detection Using SURF features," SCCS 2019 - 2019 2nd Sci. Conf. Comput. Sci., pp. 81–86, 2019, doi: 10.1109/SCCS.2019.8852603.

[7] Y. Gao, Y. Lai, and Y. Liu, "Fast Video Shot Boundary Detection Based on Visual Perception," 2019 IEEE Int. Conf. Consum. Electron. ICCE 2019, pp. 1–4, 2019, doi: 10.1109/ICCE.2019.8662083.

[8] S. Zhou, X. Wu, Y. Qi, S. Luo, and X. Xie, "Video shot boundary detection based on multi-level features collaboration," Signal, Image Video Process., vol. 15, no. 3, pp. 627–635, 2021, doi: 10.1007/s11760-020-01785-2.

[9] B. S. Rashmi and H. S. Nagendraswamy, "Video shot boundary detection using block based cumulative approach," Multimed. Tools Appl., vol. 80, no. 1, pp. 641–664, 2021, doi: 10.1007/s11042-020-09697-6.

[10] S. Chavate and R. Mishra, "An Efficient Approach for Shot Boundary Detection in Presence of Illumination Effects using Fusion of Transforms," Int. J. Eng. Trends Technol., vol. 70, no. 4, pp. 418–432, 2022, doi: 10.14445/22315381/IJETT-V70I4P236.

[11] D. Chakraborty, W. Chiracharit, K. Chamnongthai, and T. Charoenpong, "Wipe Scene Change Detection in Object-Camera Motion Based on Linear Regression and an Inflated Spatial-Motion Neural Network," IEEE Access, vol. 11, no. April, pp. 33080–33099, 2023, doi: 10.1109/ACCESS.2023.3262796.

[12] P. G. G. Lakshmi and S. Domnic, "Walsh-hadamard transform kernel-based feature vector for shot boundary detection," IEEE Trans. Image Process., vol. 23, no. 12, pp. 5187–5197, Dec. 2014, doi: 10.1109/TIP.2014.2362652.

[13] M. Birinci and S. Kiranyaz, "A perceptual scheme for fully automatic video shot boundary detection," Signal Process. Image Commun., vol. 29, no. 3, pp. 410–423, 2014, doi: 10.1016/j.image.2013.12.003.

[14] J. Fan, S. Zhou, and M. A. Siddique, "Fuzzy color distribution chart -based shot boundary detection," Multimed. Tools Appl., vol. 76, no. 7, pp. 10169–10190, 2017, doi: 10.1007/s11042-016-3604-y.

[15] [M. M. Khan, K. Chamnongthai, and S. Member, "Multi-modal Visual Features Based Video Shot Boundary Detection," vol. 3536, no. c, pp. 1–13, 2017, doi: 10.1109/ACCESS.2017.2717998.

[16] Y. Bendraou, F. Essannouni, D. Aboutajdine, and A. Salam, "Shot boundary detection via adaptive low rank and svd-updating," Comput. Vis. Image Underst., vol. 161, pp. 20–28, 2017, doi: 10.1016/j.cviu.2017.06.003.