International Journal of Engineering & Technology, 7 (4.19) (2018) 636-641

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Video Shot Boundary Detection Method Using Modified Artificial Bee Colony and Fast Accelerated Segment Test Algorithm

Isra'a Hadi*[1], Hikmat Z. Neima[2]

[1]*Professor, University of Babylon, College of IT*
[2]*Ph.D. Student, University of Babylon, College of IT*
**Corresponding Author Email: Israa_hadi1968@yahoo.com*

Abstract

Over recent years, video has become one of the most important means of communication. This is due to the rapid growth in network bandwidth and device storage capacity. The management of video databases therefore requires efficient video analysis methods. Video shot boundary detection (SBD) is considered an essential step in a wide range of video analysis processes such as video browsing, indexing, retrieval and summarization. Many SBD methods have been proposed; however, accuracy in the detection of shot boundaries is still unsatisfactory. In this paper, an efficient SBD method is proposed. The proposed method consists of two stages: identification and verification. In the identification stage, candidate shot boundaries are identified based on a modified Artificial Bee Colony (ABC) algorithm. In the verification stage, the global and local features of frames are extracted to verify the shot boundaries. Fuzzy histograms are utilized as a global feature while features from Fast Accelerated Segment Test (FAST) corner detection is implemented to extract interest points as a local feature. The proposed method has shown highly accurate results.

Keywords: *Shot boundary detection, Artificial Bee Colony, Fast Accelerated Segment Test, Fuzzy Histogram.*

1. Introduction

The rapid growth of network bandwidth and device storage capacity has led to an explosion in the use of video data, which are now widely used in many applications such as video on demand, video conferencing and distance learning. This growth has motivated the development of efficient tools for video indexing, search and retrieval [1].

The basic unit in video structure is called a shot. A video shot can be defined as a sequence of consecutive frames taken from a single camera. Thus, in order to analyse video structure, video shots must first be obtained by temporally segmenting a video sequence. The process of detecting video shots concentrates on locating the boundaries between shots. This process is known as video SBD [2].

Depending on the type of transition from one shot to another, shot boundaries can be categorized into two types. The first type is a cut transition (CT), in which the transition between two successive frames occurs due to an abrupt change in terms of specific features. A gradual transition (GT), in contrast, proceeds by steps or degrees. GT can be further classified into dissolve, fade and wipe. In dissolve, the last frames of the current shot are temporally overlapped with the first frames of the next shot. Fade occurs as a fluid change in the brightness of the frame that continues until the frame is black (fade out), whereafter the black frame smoothly turns into the next shot (fade in). Finally, in wipe, the appearing and disappearing shots are interchanged in the transitional frames until the appearing shot completely replaces the disappearing one [3-4].

Artificial intelligence (AI) is one of the best-known computer science research fields. AI is defined as the study of intelligent machines that perceive environmental conditions and take appropriate decisions to maximize their chances of achieving a solution to a given problem. The ABC algorithm has been proven as a successful optimizer for numerical problem-solving, which has motivated researchers to apply ABC to a range of other problems, including integer programming problems, binary optimization, combinatorial optimization and machine learning [5].

The basic idea behind SBD is to locate discontinuities in the visual content of a frame. SBD has attracted considerable research attention of late, and a number of methods have been presented. Most of these extract specific features from frames and compare the extracted features in order to locate the presence of shot boundaries [6].

Due to existing of various video types beside to other factors such as change in video illumination, video analysis is a challenging task. As SBD is a critical step in most video analysis processes, accurate SBD approaches are still required. In this paper, an efficient method of shot boundary detection is proposed. The method makes use of a modified ABC algorithm to identify candidate boundaries. Afterward, a combination of the FAST algorithm and fuzzy histograms is implemented in order to verify candidate boundaries.

The remainder of this paper is organized as follows: Section 2 introduces the background and related concepts. Section 3 reviews the existing body of literature. In Section 4, the proposed method is described. This is followed by a discussion of the experimental results in Section 5. Finally, Section 6 sets out the conclusion.

2. Background

This section presents the concepts that are used in this paper.

2.1. Artificial Bee Colony (ABC) Optimization

Optimization is a critical issue in the field of artificial intelligence. Optimization problems are widely present in control engineering, information science, industrial design, and other fields. Their prevalence has translated into a great deal of attention being paid to optimization algorithms. Such algorithms are classified into population-based algorithms, deterministic algorithms and non-deterministic (stochastic) algorithms [7] [8].
In population-based algorithms, each population member cooperates with other members in the population in order to improve solutions. Population-based algorithms are subdivided into evolutionary algorithms or swarm intelligence-based algorithms. Swarm intelligence-based algorithms are inspired by the intelligence of swarms, which can organize themselves systematically and intelligently. Swarm intelligence-based algorithms such as the genetic (GA), particle swarm optimization (PSO), ant colony optimization (ACO) and artificial bee colony (ABC) algorithms are simple, easy to implement, scalable, fault-tolerant and effective [9].
Among these algorithms, ABC, which was proposed by Karaboga in 2005, has proven to be effective at achieving global optimization. ABC employs a small number of control parameters, which advantage has led to the recent application of the algorithm to the solution of several problems, such as image segmentation. The ABC algorithm is inspired by the behaviour of swarms of bees, and is based on the means by which bees search for the optimal food source. In the context of this study, the location of the food source is analogous to the location in the search space, while the quantity of nectar in the location is analogous to the fitness of the candidate solution [10].
ABC simulates three types of bees: foragers, onlookers, and scouts. In ABC, each forager modifies its position in the search space. Onlookers randomly modify the position of foragers and move to a new position if any such random modification improves the forager. Scout bees replace foragers if there is no improvement after a number of modifications [9].
ABC has been found to excel at exploration, but be less effective at exploitation. In order to improve the ABC algorithm, a number of modifications on the original algorithm have been presented. The improved discrete ABC (DABC) algorithm is proposed in [10]. The resultant algorithm has been used to solve hybrid flexible flowshop problems with dynamic operation skipping. In DABC, the local search strategy is modified to improve the exploitation ability of standard ABC. The first step in DABC is initialization of the population, followed by the evaluation of each solution within the initialized population to identify the optimal solution. The third step is the employed exploitation task using a hybrid neighbourhood structure. The fourth and fifth steps are onlooker and scout phases, respectively. In the sixth step, the proposed local search is applied around the food sources that have been selected thus far. The heart of the modification in DABC lies in the process of exploitation. The enhanced local search depends on randomly generating new positions for the current solution based on a skipping neighbourhood structure, which ensures good exploitation and exploration.
In [11], the authors adopt the concept of particle swarm optimization (PSO) to present a method known as gbest-guided ABC. The method incorporates the global best information into the solution search Equation, which improves the exploitation of the original ABC algorithm. The authors argue that gbest-guided ABC outperforms the original ABC algorithm. Another modification of ABC is proposed in [12], in which a quick ABC (qABC) algorithm modifies the behaviour of onlooker bees in order to improve the local search of the original ABC. In qABC, onlooker bees choose food sources in a manner that differs from that used by forager bees to determine the best food sources, while, in the original ABC algorithm, both onlooker and forager bees determine food sources in the same manner. The proposed modification to the way in which onlookers select food sources depends on search-in-the-neighbourhood solutions. The neighbourhood area is defined in relation to the structure similarity measurement using the mean Euclidean distance.
In [13], a multi-swarm ABC (MABC) is proposed for dynamic optimization problems. As the optimal solution in dynamic problems changes over time, and the standard ABC lacks the ability to detect that change, the author proposes a strategy to detect such changes in the environment. The detection strategy calculates the best bee fitness in every generation, and uses that fitness as a baseline. Thereafter, changes are detected if differences become apparent between newly detected fitness and the baseline, indicating that the previous best fitness has been degraded. The second modification to standard ABC is to divide the population into several sub-swarms during the initialization phase. A further modification proposed by MABC relates to the strategy of the current solution in the current sub-swarm. MABC is claimed to be superior to standard ABC in relation to the speed of convergence.

2.2. FAST Corner Detection

Interest points (or, interchangeably, corners) in an image are pixels that occupy certain positions and can be repeatedly detected. The detection of interest points is the process of locating the positions of these points. Interest point or corner detection is utilized to extract important features from an image; these features are beneficial in several applications such as image matching and object detection in video sequences [14].
Several well-known algorithms are used to detect interest points, including the Moravec, Harris, SUSAN, and FAST corner detection algorithms. FAST was proposed by Rosten and Drummond, and has as its principal advantage its excellent computational efficiency compared to other interest point algorithms, such as Harris and SUSAN [15].

2.3. Fuzzy Histogram

Image histogram is simply a representation of the occurrence of grey level frequencies in an image. It is considered a global image feature, and is primarily utilized in order to obtain a global description of an image. Image histograms are widely used in image processing operations such as image enhancement [16].
In 1965, Lotfi Zadeh introduced the concept of fuzzy sets. A fuzzy set is a generalization of a classical (crisp) set. Since the introduction of the concept, it has been widely incorporated in multiple research fields in order to improve results. Its ability to do so lies in the fact that it allows an element to be a member of a set with a certain degree using a specific membership function. Fuzzy histograms are therefore able to introduce more accurate results than classical histograms [17]. In summary, the fuzzy histogram is a generalization of the crisp histogram and replaces it in many applications.

2.4. Evaluation Criteria

In order to evaluate shot boundary detection methods, two basic metrics are usually utilized: precision and recall. In general, precision is defined as the proportion of relevant information returned by the system as in Equation (1). Recall, on the other hand, is defined as the proportion of all the relevant information that is returned by the system as in Equation (2) [18].

$$Recall = \frac{N_C}{N_C + N_M} \tag{1}$$

$$Precision = \frac{N_C}{N_C + N_F} \quad (2)$$

Where
N_C : the number of shot boundaries that are correctly detected.
N_M : the number of shot boundaries that are missed in the detection process.
N_F : the number of shot boundaries that are falsely detected.

3. Related Work

This Section presents several approaches found in literature that are related to the proposed method. In [19], an SBD method based on performance analysis of multiple colour histograms is proposed. In the proposed method, shot boundaries are identified by computing the histogram differences between successive frames in several colour spaces using assorted difference measures. The authors conclude that the histogram intersection technique performs better than other histogram techniques. Another method for SBD is presented in [20], and employs the hidden Markov model (HMM) to discover shot boundaries. It categorizes video frames into abrupt transition region, gradual transition region and non-transition region. This method consists of two steps: training data to train the parameters of HMM, and testing video based on HMM.
In [21], the proposed SBD method relies on a binary k-NN classifier to identify shot boundaries. The method, however, suffers from computational complexity. In [13], the authors propose a method in which a support vector machine (SVM) algorithm is utilized to classify video frames into normal frames, cut transitional frames and gradual transitional frames. This method requires vast amounts of memory and computing power due to the presence of numerous training samples. A new fast framework that strives for a balance between detection accuracy and performance is proposed in [22]. In this method, the first step detects non-boundary segments and focuses on segments containing shot transitions, which are preserved for further detection as a second step.
In [23], the authors propose a method based on a hybrid feature set, which they use as a model for the frame. The hybrid feature set contains frame transition parameters and estimation errors based on global and local frame statistics. The global statistic here is the colour histogram, while the local statistics include edge strength scatter matrix and motion matrix. This method has been found to be efficient and works well on a variety of benchmark videos. Birinci and Kiranyaz have suggested a strong method for fast and accurate shot boundary detection where the entire design depends on human perceptual instructions and the familiar 'visual information-seeking mantra' [24]. The proposed method can be used for discovering all types of piecemeal transitions as well as acute changes. Furthermore, the proposed method is fully generic, and can be utilized with any video content without the need for advance training or tuning.
A technique combining local and global feature descriptors to obtain the temporal characteristics of videos is proposed in [25]. In this technique, an adaptive threshold is used. Moreover, the candidate segment selection and transition pattern analysis are executed by the variation score between video frames. The experimental result refers to good computational efficiency. Finally, Fanan and Khobragade in 2016 proposed a video shot boundary detection method that achieves the condition of real time applications. This method uses a pre-processing scheme prior to shot detection, which includes SVD and candidate segment selection [26].
There are many SBD approaches have been proposed in literature; however, the accuracy of SBD still needs to be improved. In the following Section, the proposed method is elaborated.

4. Proposed Method

As is clear from Figure 1 below, which shows the block diagram of the proposed method, there are two major stages: identification and verification. After extracting the frames from the input video, the identification stage commences, which is followed by the verification stage. The following subsections present the details of each stage.

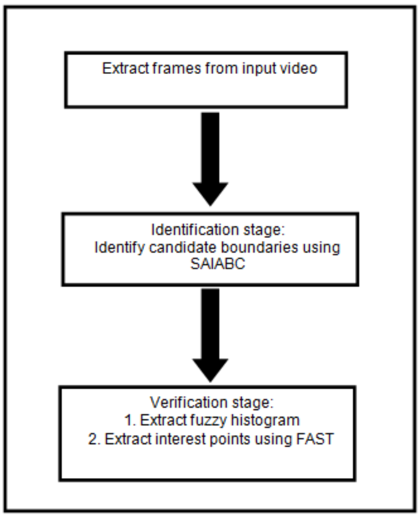

Fig. 1: Block diagram of the proposed method

4.1. Identification Stage

In this stage, candidate shot boundaries are identified using a modified ABC optimization algorithm. Standard ABC has been proven to be competitive in comparison to other swarm-intelligence algorithms; however, despite performing well in terms of exploration, it is less effective at exploitation [11]. This occasionally results in ABC getting stuck at a local optimal solution. Hence, the proposed modified ABC adopts the idea of accepting a worse solution, which originated in the well-known combinatorial optimization algorithm, simulated annealing (SA). The SA-Inspired ABC (SAIABC) performs the local search that is used by SA – rather than greedy search – to enhance the exploitative capabilities of the original ABC. SA is a metaheuristic technique used to find the optimal global solution to a given problem. In SA, the risk of accepting a worse solution with a probability given by Boltzman distribution [27]. Algorithm 1 demonstrates the pseudo code of accepting the worse solution in a certain probability. Algorithm 2 shows the pseudo code of SAIABC. Algorithm 2 invokes Algorithm 1 whenever the obtained solution is worse than the old one.

$$v_{ij} = x_{ij} + rand\,(x_{ij} - x_{kj}) \quad (3)$$

$$P_i = \frac{Fit_i}{\sum_{i=1}^{N} Fit_i} \quad (4)$$

$$Fit_i = \begin{cases} \frac{1}{1+f_i} & if\ f_i \geq 0 \\ 1 + abs\,(f_i) & if\ f_i < 0 \end{cases} \quad (5)$$

$$p_a = p_o \,\frac{1 + \cos\,((Iter/Maxiter)\,\pi)}{2} \quad (6)$$

Where
v_{ij} : new solution that is found by employ bees by local search around i^{th} food source , i = 1, 2, ..., SN (source food number).
x_{ij} : upper bound of dimension *j*.
x_{kj} : lower bound of dimension *j*.
P_i : represents the amount of food source nectar.
Fit_i : fitness value of i^{th} solution.
p_a : acceptance probability.
p_o : initial probability.

In canonical ABC, a new solution is randomly obtained by Equation (3) before being evaluated by Equation (5). Standard ABC applies a greedy algorithm to find a better solution, meaning that it accepts a new solution as the best solution if it is better than the old one, regarding of fitness function, and keeps the old solution if the new one is worse. This methodology holds a considerable risk that the algorithm will stick to a local optimum solution. On the other hand, in SAIABC, instead of making use of a greedy algorithm, a worse solution is accepted within a certain probability. This acceptance probability is computed by Equation (6), as shown in Algorithm 1.

In identification stage, video frames are read and SAIABC is applied in order to obtain frames (candidate frames) that have a high probability of being shot boundaries. The identified frames need to be verified to be either real boundaries or identified because of illumination or camera motion effects.

Algorithm 1: Pseudo code of accepting of worse solution

```
Input : x_i, v_i, trial_i, p_o

Output : x_i

1-  If v_i is better than x_i
        i.   x_i = v_i
        ii.  trial_i = trial_i + 1
2-  Else
        i.   Calculate p_a using Equation (6)
        ii.  If r < p_a then  // r is a random real no. [0-1]
                 1.  x_i = v_i
                 2.  trial_i = trial_i + 1
        iii. Else
                 1.  trial_i = trial_i + 1
        iv.  Endif
3-  Endif
```

Algorithm 2: Pseudo code of SAIABC

```
Input : N, Maxiter

Output : Best Solution

1-   Generate initialize population
2-   Evaluate fitness of population using Equation (5)
3-   Iter = 1
4-   While Iter ≤ Maxiter do
        i.    Send employed bees to produce new solution as in Equation (3)
        ii.   Evaluate new solution
        iii.  If new solution is worse then
                  Apply algorithm 1
        iv.   Calculate probability using Equation (4)
        v.    Send onlooker bees to produce new solution depending on
              probability
        vi.   Evaluate new solution
        vii.  If new solution is worse then
                  Apply algorithm 1
        viii. Memorize the best solution found so far
        ix.   Iter = Iter + 1
5-   End While
6-   Return the best solution
```

4.2. Verification Stage

After identification of candidate frames, the second stage, verification, is performed. In the verification stage, the candidate frames are double-checked. In the first check, each candidate frame (F_i) is compared with the previous frame (F_{i-1}). The comparison process is performed in terms of corners detected from both frames being processed. FAST corner detection algorithm is applied for the purpose of extracting corners. The major advantage of FAST corner detection is its computational efficiency compared to other interest point algorithms such as Harris and SUSAN, making it particularly well suited to online video processing. This makes it optimal for deployment in the proposed method.

To increase the accuracy of the results of the detection process, in addition to corner detection, the fuzzy histograms of both frames are also compared. Fuzzy histogram is utilized in this paper due to its ability to introduce more accurate results than classical histogram.

The candidate frames that pass the first evaluation step are then further evaluated in the second step. Candidate frames that fail the initial evaluation step are discarded. Each passed frame (F_i) is compared with the next frame (F_{i+1}) in the second evaluation step. In the same way as the first evaluation step, the second evaluation step discards those frames that fail the comparison process. The passed frames then compose the final result of the shot boundary detection, which represent accurate shot boundaries.

5. Experimental Results and Discussion

This section introduces and discusses the results obtained by the proposed method. The proposed method was applied to six standard video sequences that were downloaded from an open video website. These video sequences were recognized by TRECVid organization. As the process of downloading the raw video sequences along with their descriptions from the TRECVid website was complex, three of the six video sequences are manually described. Table 1 sets out the results of the six videos. The video sequences chosen were deliberately diverse both in terms of their duration as well as the number of shots they contain.

Table 1: Recall and precision of proposed method

Video Sequence	Duration	No. of Frames	Detected CT	Actual CT	Cut	
					R	P
Ani05	6.19 m	11363	38	38	100%	100%
Ani09	6.50 m	12306	41	39	100%	95.12 %
Bor03	26.56 m	48450	234	231	98.26 %	97.42 %
Nad57	6.57 m	12510	45	45	100%	100%
Senses109	4.7 m	7401	13	13	100%	100%
Senses115	4.7 m	74 19	12	12	100%	100%

As explained in Section 1, transitions may take one of two forms: cut transition (CT) or gradual transition (GT). This paper focuses only on CT. Table 1 shows that the proposed method results in a very high degree of accuracy in regard to both recall and precision. In four of the six video sequences, actual shot boundaries were obtained with 100 per cent recall and precision, while the remaining two video sequences were obtained with a high degree of accuracy. These results indicate the robustness of the proposed method.

6. Conclusion

Shot boundary detection is a fundamental and essential aspect of most video analysis processes, including video retrieval and video summarization. In this paper, an efficient method of video shot boundary detection is proposed. This method makes three contributions. First, it modifies ABC based on idea inspired by SA (SAIABC). Second, it utilises SAIABC in shot boundary identification. Third, it extracts both global and local features in order to verify the identified boundaries. In conclusion, using SAIABC to identify candidate shot boundaries requires no threshold value, which is the main advantage of utilizing SAIABC in detecting shot boundaries. This is because the threshold value is critical and determination of its value has a significant effect on the process outcome. The use of an optimization algorithm in shot identification grants the proposed method significance, while verification by means of global and local features makes the method more accurate. As a consequence, the proposed method yields excellent results.

References

[1] Hannane, R., A. Elboushaki, and K. Afdel. "Efficient video summarization based on motion SIFT-Distribution histogram". IEEE 13th International Conference in Computer Graphics, Imaging and Visualization (CGiV), 2016.
[2] Koprinska, I. and S. Carrato, "Temporal video segmentation: A survey", 2001, 16(5): p. 477-500.
[3] Cotsaces, C., N. Nikolaidis, and I. Pitas, "Video shot detection and condensed representation: a review". IEEE signal processing magazine, 2006. 23(2): p. 28-37.
[4] Boreczky, J.S. and L.A. Rowe, "Comparison of video shot boundary detection techniques", Journal of Electronic Imaging, 1996. 5(2): p. 122-129.
[5] Karaboga, D., et al., "A comprehensive survey: artificial bee colony (ABC) algorithm and applications", Artificial Intelligence Review, 2014. 42(1): p. 21-57.
[6] Huo, Y., Y. Wang, and H. Hu, "Effective algorithms for video shot and scene boundaries detection", IEEE 15th International Conference on Computer and Information Science (ICIS), 2016 IEEE/ACIS.
[7] He, Z.-a., et al., "A modified artificial bee colony algorithm based on search space division and disruptive selection strategy", Mathematical Problems in Engineering, 2014.
[8] Hadi, I. and M. Sabah, "Upgrade Video Tracking Technique Using Enhanced Hybrid Cat Swarm Optimization Based on Multi Target Model and Accumulated Histogram", Journal of Computational and Theoretical Nanoscience, 2015, 12(11): p. 4017-4027.
[9] Mao, W., H.-y. Lan, and H.-r. Li, "A new modified artificial bee colony algorithm with exponential function adaptive steps. Computational intelligence and neuroscience", 2016: p. 23.
[10] Li, J.-q., Q.-k. Pan, and P.-y. Duan, "An improved artificial bee colony algorithm for solving hybrid flexible flowshop with dynamic operation skipping", IEEE transactions on cybernetics, 2016. 46(6): p. 1311-1324.
[11] Zhu, G. and S. Kwong, "Gbest-guided artificial bee colony algorithm for numerical function optimization", Applied mathematics and computation, 2010. 217(7): p. 3166-3173.
[12] Karaboga, D. and B. Gorkemli, "A quick artificial bee colony-qABC-algorithm for optimization problems", IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), 2012.
[13] Jia, D.-L., S.-X. Qu, and L.-Y. Li., "A Multi-swarm Artificial Bee Colony Algorithm for Dynamic Optimization Problems", IEEE International Conference on Information System and Artificial Intelligence (ISAI), 2016.
[14] Yi-bo, L. and L. Jun-Jun, "Harris corner detection algorithm based on improved contourlet transform", Procedia Engineering, 2011. 15: p. 2239-2243.
[15] Rosten, E. and T. Drummond," Fusing points and lines for high performance tracking", Tenth IEEE International Conference on Computer Vision, ICCV 2005.
[16] Acharya, T. and A.K. Ray, "Image processing: principles and applications", 2005: John Wiley & Sons.
[17] Gera, Z. and J. Dombi, "Fuzzy reasoning models and fuzzy truth value based inference". Department of Computer, 2009.
[18] Xu, J., L. Song, and R. Xie, "Shot boundary detection using convolutional neural networks". IEEE Conference on Visual Communications and Image Processing (VCIP), 2016.
[19] Gargi, U., R. Kasturi, and S.H. Strayer," Performance characterization of video-shot-change detection methods", IEEE transactions on circuits and systems for video technology, 2000. 10(1): p. 1-13.
[20] Bae, T.M., S.H. Jin, and Y.M. Ro, "Video segmentation using hidden Markov model with multimodal features", International Conference on Image and Video Retrieval, 2004: Springer.
[21] Cooper, M., T. Liu, and E. Rieffel, "Video segmentation via temporal pattern classification", IEEE Transactions on Multimedia, 2007. 9(3): p. 610-618.
[22] Li, Y.-N., Z.-M. Lu, and X.-M. Niu, "Fast video shot boundary detection framework employing pre-processing techniques", IET image processing, 2009. 3(3): p. 121-134.
[23] Mohanta, P.P., S.K. Saha, and B. Chanda," A model-based shot boundary detection technique using frame transition parameters", IEEE Transactions on Multimedia, 2012. 14(1): p. 223-233.
[24] Birinci, M. and S. Kiranyaz, "A perceptual scheme for fully automatic video shot boundary detection", signal processing: image communication, 2014. 29(3): p. 410-423.
[25] Tippaya, S., et al., "Video shot boundary detection based on candidate segment selection and transition pattern analysis", IEEE International Conference on Digital Signal Processing (DSP), 2015.
[26] Fanan, N.G. and A.S. Khobragade, "A fast robust technique for video shot boundary detection", IEEE Online International Conference on Green Engineering and Technologies (IC-GET), 2016.

[27] Du, K.-L. and M.N. Swamy, "Neural networks and statistical learning", 2013: Springer Science & Business Media.