

Article

Shot Boundary Detection Based on Global Features and the Target Features

Qiuling Li ¹, Xiao Chen ¹, Bingbing Wang ¹, Jing Liu ^{1,2}, Guofeng Zhang ^{1,*}  and Bin Feng ^{1,*}¹ School of Information Science and Technology, Tai Shan University, Taian 271021, China² Shenyang Institute of Computing Technology, University of Chinese Academy of Sciences, Shenyang 110168, China

* Correspondence: zhangguofeng@tsu.edu.cn (G.Z.); binfeng@tsu.edu.cn (B.F.)

Abstract: Video processing plays an important role in the intelligent monitoring and management system of agricultural information. Video shot boundary detection is the basic symmetry step underlying video processing techniques. According to the current shot boundary detection algorithm, the feature changes between gradual transition frames are difficult to detect, and the misdetection situation is caused by ignoring the attention of the target feature during the feature extraction. A novel symmetry multi-step comparative scheme of shot boundary detection algorithm based on global features and target features is proposed. First, the RGB color histogram features of the video frame are extracted. Second, foreground object detection for the video frames is performed using the Gaussian Mixture Model (GMM), and the scale-invariant features transformation (SIFT) of the foreground targets is extracted. Finally, global features and target features fusion through weights, calculating the difference between adjacent frames across multiple steps, generate a pattern distance map. The pattern distance map of the gradual transition and the cut detection is different; we can judge the gradual transition and the cut detection according to the pattern distance map. Experiments show that the proposed symmetry method improves by about 2% in recall and accuracy compared to other algorithms.

Keywords: shot boundary detection; RGB; target feature; GMM; SIFT

Citation: Li, Q.; Chen, X.; Wang, B.; Liu, J.; Zhang, G.; Feng, B. Shot Boundary Detection Based on Global Features and the Target Features. *Symmetry* **2023**, *15*, 565. <https://doi.org/10.3390/sym15030565>

Academic Editors: Jeng-Shyang Pan and Sergei D. Odintsov

Received: 1 December 2022

Revised: 15 January 2023

Accepted: 14 February 2023

Published: 21 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As the primary industry in the world, agriculture occupies a very important position in social development. Video processing plays an important role in the intelligent monitoring and management system of agricultural information. It is also the basic symmetry step in video processing. With the rapid development of information technology, video data have been growing rapidly. Multimedia information is increasingly rich, and content-based video retrieval [1], video annotation, video index, video summary, and other technologies have emerged. Shot boundary detection technology is the basis of video-structured analysis, and it is also an important step in video retrieval and other related work. The quality of the shot boundary detection algorithm directly affects the performance of video retrieval and other subsequent work, so it is particularly important to find a more efficient shot boundary detection algorithm. Although traditional shot boundary detection algorithms, such as color histogram, edge comparison, and frame difference methods, can detect cut detection to a certain extent, they cannot detect gradual transition correctly. Traditional algorithms are not sensitive to interference information such as illumination and motion [2]. The clustering-based shot boundary detection algorithm needs to set the number of clusters, which partly limits the accuracy of the algorithm results and increases some randomness [3]. The sub-block-based shot boundary detection algorithm is better for motion and noise compared to the general algorithms, but it is highly computational and susceptible to illumination [4]. With the development of deep learning, shot boundary detection using Convolutional

Neural Networks (CNN) emerges. For example, Michale Gygli recently proposed a Neural Network with 3D Convolution for shot boundary detection. Although the accuracy and speed of the algorithm exceed the existing algorithm due to the increased dimension, it also significantly increases the complexity of computation and the requirements for computer hardware [5,6]. An article published in 2022 was based on primary segments using the adaptive threshold of color-based histogram differences and candidate segments using the SURF feature descriptor. The field of shot boundary detection is still studying traditional methods [7].

RGB color histogram features describe the surface properties of the image that are widely adopted in many fields of digital image processing. It is a traditional and efficient way of feature extraction. RGB color histogram features describe the proportion of different colors in the whole image and do not care about the spatial position of each color. The characteristic of this feature is that it can generally control the global changes in the video frame overall but ignores the internal detail changes in the video frame and the changes in the main target features.

Background modeling algorithms are widely used in the field of target detection, and the representative ones are Gaussian Mixture Model and the Codebook model. In practical application, the Codebook model consumes large memory and affects real-time performance [8]. However, the GMM algorithm is an efficient algorithm for background modeling, and the GMM algorithm can adapt to the scene changes; in this paper, the foreground object detection for the video frames using GMM.

SIFT feature is characterized by maintaining the invariance of rotation, scale scaling, and brightness change, and it is a stable local feature. The characteristic of this feature is that it can care about internal detail changes in the video frame.

Therefore, this paper presents a shot boundary detection algorithm based on global features and target features. The RGB color histogram features of the video frame are extracted. Then foreground object detection for the video frames using GMM and extracting the SIFT features of the foreground targets. In this paper, the effect is verified through experiments and opens up ideas for the study of the shot boundary detection algorithm.

2. Related Works

2.1. RGB Color Histogram Features

Color is most widely used in visual features of the image and is found effective in many retrieval algorithms as it is independent of the size and orientation of the image.

Color histograms are widely used in the field of computer vision as a simple and effective feature. Due to a large amount of video frame data, it is particularly important to extract the global features of video frames simply and effectively.

RGB color space is an additive color system based on tri-chromatic theory. RGB is easy to implement; it is very common and is used in virtually every computer system as well as television, video, etc. [9].

2.2. Gaussian Mixture Model

GMM is suitable for modeling videos with more complex backgrounds, which are mainly used for target detection. The first step of GMM to detect moving objects is to calculate the sample value probability density of the pixels in a frame in a certain period of time and then to determine whether each pixel is the background content of the video frame according to the common principles in the statistical difference [10].

The specific process of GMM for background modeling is the following [11]:

First, assuming that each pixel in the video frame is independent of the other, the value of each pixel can be generated by a mixed distribution. This distribution is composed

of k independent Gaussian distributions. In this paper, the value of k is 5, the pixel value is set as x , and the probability of the pixel value at time t is:

$$p(x_j) = \sum_{i=1}^k w_{i,t} \eta(x_t, \mu_{i,t}, \Sigma_{i,t})$$

In the formula: $w_{i,t}$, $\mu_{i,t}$ and $\Sigma_{i,t}$ are the weights, means, and covariance of the i th Gaussian distribution at time t , respectively. Assuming that the three RGB channels of the pixels are independent of each other and have the same variance, then:

$$\Sigma_{i,t} = \sigma_k^2 I$$

If satisfied:

$$|I_t - \mu_{i,t}| < \gamma \sigma_{i,t}$$

It means that x and this Gaussian model is matched, then the parameters of each model are updated, and the weights are normalized. The first B Gaussian distributions were used to establish the background model:

$$B = \operatorname{argmin}_b \left(\sum_{k=1}^b w_k > T \right)$$

In the formula, the threshold T value is 0.7, and the first B Gaussian distribution in the sorted K Gaussian distribution is the best description of this background pixel. Check the matching relationship of each pixel value with its resulting first B Gaussian distribution, and if the pixel value matches one of the first B Gaussian distributions, the pixel point is a background point or a foreground point otherwise.

2.3. SIFT Features

The SIFT features were proposed by Canadian Professor David G. Lowe. The SIFT features are characterized by maintaining the invariance of rotation, scale scaling, and brightness change and are stable local features [12]. In this paper, the SIFT features were extracted from the images containing the moving targets. The acquisition of the scale space used in the SIFT feature point extraction algorithm requires Gaussian fuzzy, and the two-dimensional Gaussian fuzzy function is defined as follows [13].

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} \exp\left\{-\frac{(x^2 + y^2)}{2\sigma^2}\right\}$$

In the formula, σ represents the standard deviation of the normal distribution. Then, the scale-space representation of a 2-dimensional image at different scales can be obtained from the image and the Gaussian convolution kernel:

$$L(x, y, \sigma) = G(x, y, \sigma) I(x, y)$$

In the formula, (x, y) represents the pixel position of the image, σ is the scale-space factor, and $L(x, y, \sigma)$ represents the scale space of the image.

The SIFT feature point extraction algorithm simultaneously detects the local maxima in the image 2D plane space and the difference of gaussian (DOG) scale space as the feature points. The DOG operator is shown below:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$

The overall generation of an image SIFT feature vector consists of four steps [14]:

1. Scale-space detection is performed to initially determine the key point location and scale;
2. The 3-dimensional quadratic function is pseudomatched to accurately determine the position and scale of the key points, removing the key points with low contrast and the unstable edge response points;
3. The gradient orientation distribution properties of the key neighborhood pixels are used to specify the orientation parameters for each key;

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \alpha \tan 2\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right)$$

The above formula refers to the modular value and direction formula of the gradient at (x, y) , respectively. At this time, the key detection is completed, and the scale used by L is each key point;

4. Generate the SIFT feature vectors.

3. Fusion Feature Algorithm

In this algorithm, we extract color histogram features for video frames, using three color channels: R, G, and B as feature vectors. Each color channel was quantified using 8 bins to obtain the 8-dimensional histogram features for each channel, and the three channels were quantified with the 512-dimensional eigenvectors describing each frame. The following formula represents the color histogram of the frame.

$$R(n, i), 0 \leq i \leq M - 1$$

In the formula, n and i represent the frame number in the video sequence and the bin number in the histogram, respectively, and $M = 512$.

Foreground object detection for the video frames using GMM and extracting the SIFT features of the foreground targets. Each key point has three pieces of information: location, scale, and direction. A standard setting subregion of 4 by 4 for each key is used, and direction histograms using eight intervals such that a key point produces 128 data, forming a 128-dimensional SIFT eigenvector.

In the end, 8 bins of each image and each channel were quantified from the features extracted by the RGB color histogram, and 512-dimensional features were fused with 128-dimensional features quantified from each key point of the moving target in each image extracted by SIFT. The flow chart of the algorithm is shown in Figure 1.

3.1. Multi-Step Comparison Scheme

In the multi-step comparison scheme, a step-length l is set first, where l refers to the distance between two frames. The compared color histogram differences are $n - l$ and $n + 1 + l$, respectively. When l equals 0, it represents the difference between two adjacent frames. The distance between the two frames in the multiple steps is [15]:

$$\text{sigma}(n, l) = \frac{100}{6 \cdot W \cdot H} \sum_{i=0}^{M-1} |h(n-l, i) - h(n+1+l, i)|$$

In the formula, $\text{sigma}(n, l)$ represents the histogram difference between frames $h(n-l, i)$ and $h(n+1+l, i)$, and the W and H represent the width and height of the frame. The algorithm is performed by calculating the differences between frames across multiple steps; their changes were detected by analyzing their patterns in the distance map. The pattern distance map of the gradual transition detection is shown in Figure 2.

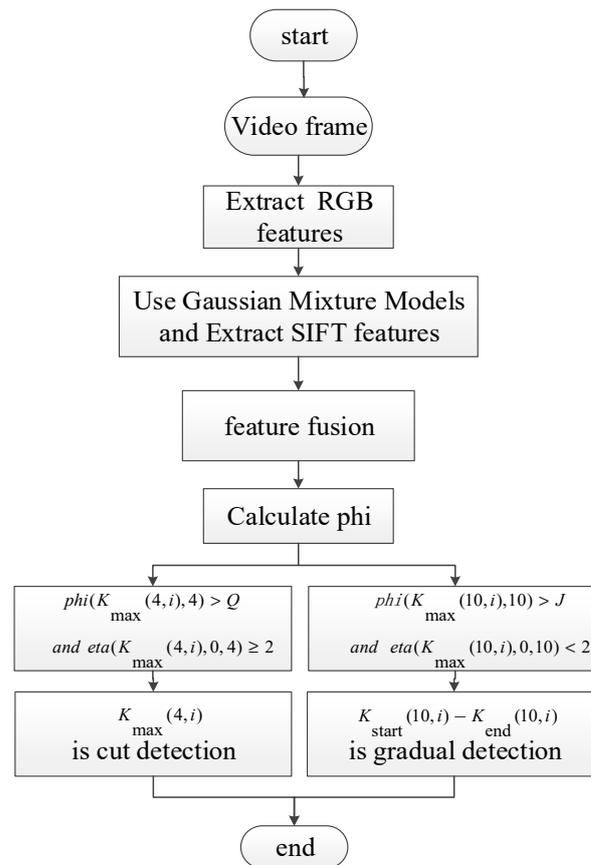


Figure 1. The flow chart of the algorithm.

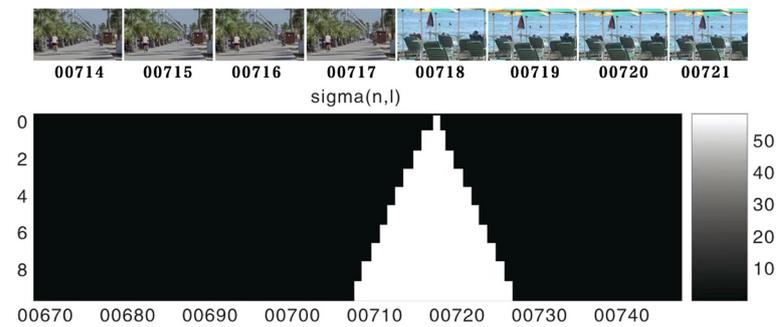


Figure 2. The pattern distance map of the cut detection.

The pattern distance map of the cut detection is shown in Figure 3.

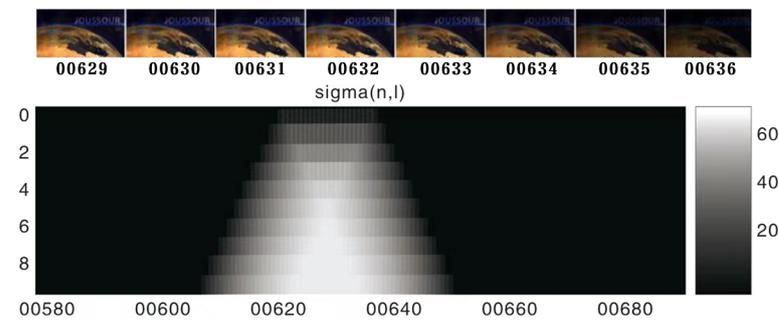


Figure 3. The pattern distance map of the gradual transition.

The sum of all possible steps can be described as:

$$\phi(n, L) = \sum_{l=0}^{L-1} \eta(n, l, L)$$

In the formula, L is the maximum step length, and $\eta(n, l, L)$ represents the difference between $\sigma(n, l)$ and the local temporal mean caused by limiting object motion or camera motion. If the frame number K is the potential peak starting point, and the detection starting point satisfies the following formula:

$$\eta(K_{\text{start}}(L) - 1, L) < 0 \cup \eta(K_{\text{start}}(L) + 1, L) > 0$$

The detection endpoint meets the following formula:

$$\eta(K_{\text{end}}(L) - 1, L) > 0 \cup \eta(K_{\text{end}}(L) + 1, L) < 0$$

The maximum number of video frames in the potential peak area can be defined as:

$$\begin{aligned} & \phi(K_{\text{max}}(L, i), L) \\ & = \text{Max}(\phi(K_{\text{start}}(L, i), L), \dots, \phi(K_{\text{end}}(L, i), L)) \end{aligned}$$

In the formula, $K_{\text{start}}(L, i)$ indicates the number of frames at the starting point, $K_{\text{end}}(L, i)$ indicates the endpoint, and $K_{\text{max}}(L, i)$ indicates the maximum frame number.

The steps of the multi-step comparative scheme of the shot boundary detection algorithm are as follows:

The RGB color histogram features and SIFT features were, respectively, extracted, and 512 dimensional features were fused with 128 dimensional features quantified from each key point of the moving target.

Calculate the histogram difference value according to set steps L , and calculate the sum of all possible steps ϕ for each frame according to the defined formula.

Different steps L were set, and the cut detection and gradual transition were judged according to the formulas ϕ and η , respectively.

The specific detection process of the gradual transition and the cut detection are shown in Sections 3.2 and 3.3.

3.2. Cut Detection

Experiments show that the maximum step size in cut detection set to 4 is the most appropriate. We set the maximum step size to 4. If the following formula is met:

$$\phi(K_{\text{max}}(4, i), 4) > Q \cup \eta(K_{\text{max}}(4, i), 0, 4) \geq 2$$

$K_{\text{max}}(4, i)$ is retained as a result of cut detection. In the formula, the Q refers to the cut detection threshold.

3.3. Gradual Detection

Experiments show that the maximum step size in gradual detection set to 10 is the most appropriate. We set the maximum step size to 10. If the following formula is met:

$$\phi(K_{\text{max}}(10, i), 10) > J \cup \eta(K_{\text{max}}(10, i), 0, 10) \geq 2$$

From $K_{\text{start}}(10, i)$ to $K_{\text{end}}(10, i)$, is retained as a result of gradual detection. The J refers to the cut detection threshold.

4. Results and Discussion

In this paper, the experiments were tested in the RAI dataset, the Open-Source Video dataset, and multiple sports videos. The evaluation criteria for the detection results of the data sets are recall, accuracy, and a comprehensive index F_1 . The formula is as follows:

$$R_p = \frac{N_c}{N_c + N_f} \times 100\%$$

$$R_r = \frac{N_c}{N_c + N_l} \times 100\%$$

$$F_1 = \frac{2 * R_r * R_p}{R_r + R_p}$$

In the formula, R_p means accuracy, R_r is recall, F_1 is the comprehensive index of accuracy and recall, N_c is the number of correctly detected boundaries, N_f is the number of missed boundaries, and N_l is the number of missed boundaries.

4.1. Experimental Results and the Comparison

This experiment focuses on the test results of 10 videos and the RAI data set. The experimental results show that the present algorithm detects the sixth video and the seventh video of the RAI dataset much better; the detection results of sports videos are also good,

The video frame using the Gaussian Mixture Model diagram is shown on the right of Figure 4:



Figure 4. Video frame using Gaussian Mixture Model diagram.

This experiment also had good detection results for the 5th, 8th, 9th, and 10th videos, but the detection effect for the 1st, 2nd, and 3rd videos was relatively general. The main reason is that the GMM in the proposed fusion algorithm can better detect the moving objects in the video; when the video frame is cut detection and gradual transition, the characteristics of the motor target will change greatly, so it is not easy to produce the error and missed detection.

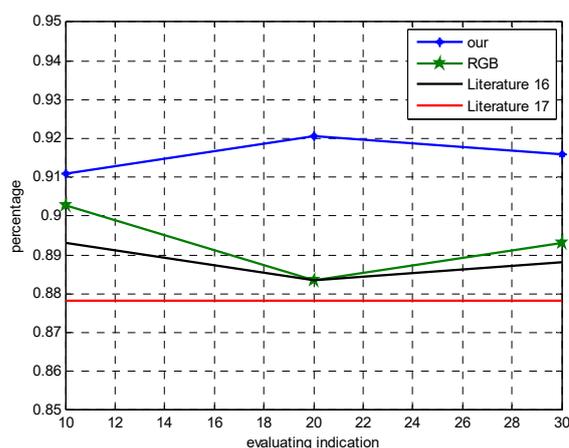
Ref. [16] proposed the shot boundary detection algorithm based on the genetic algorithm and fuzzy logic method, and Ref. [17] proposed detecting mutation and gradient using the visual similarity of adjacent video frames.

In order to verify the effectiveness of this algorithm for object detection, our algorithm was compared with the two mentioned above and the method of only extracting the features of the RGB color histogram. We selected representative videos with motion targets in the RAI dataset, Open-Source Video dataset, and multiple sports videos. The total number of test data boundaries in the table is 189. The results are shown in Table 1.

The experimental data in Table 1 can describe the experimental results more comprehensively, but the comparison of the four methods is not obvious enough. In order to make the contrast results clear, the detection results of the shot boundary detection by the four methods in Table 1 are shown in Figure 5.

Table 1. The results of the comparison.

Method	N_l	N_f	N_c	R_p	R_r	F1
our	15	17	174	91.10%	92.06%	91.58%
RGB	22	18	167	90.27%	88.36%	89.30%
Ref. [16]	22	20	167	89.30%	88.36%	88.83%
Ref. [17]	23	23	166	87.83%	87.83%	87.83%

**Figure 5.** The results of the comparison [16,17].

In order to verify the applicability of this algorithm to general videos, our algorithm was compared with the two mentioned above and the method of only extracting the features of the RGB color histogram. The results are shown in Table 2. We randomly selected videos in the RAI dataset, Open-Source Video dataset, and multiple sports videos. The total number of test data boundaries in the table is 294.

Table 2. The results of the comparison.

Method	N_l	N_f	N_c	R_p	R_r	F1
our	27	24	267	91.75%	90.82%	91.28%
RGB	36	28	258	90.21%	87.76%	88.97%
Ref. [16]	35	30	259	89.62%	88.10%	88.85%
Ref. [17]	37	35	257	88.01%	87.41%	87.71%

The experimental data in Table 2 can describe the experimental results more comprehensively, but the comparison of the four methods is not obvious enough. In order to make the contrast results clear, the detection results of the shot boundary detection by the four methods in Table 1 are shown in Figure 6.

4.2. Discussion

In the first experiment, we selected representative videos with motion targets in the RAI dataset, Open-Source Video dataset, and multiple sports videos. Experimental data are presented in Table 1. From Table 1, we can see that our algorithm was compared with Refs. [16,17], and the method of only extracting the features of the RGB color histogram. From the experimental results, our algorithm is effective for shot boundary detection. In our algorithm, the foreground object detection for the video frames using GMM and extracting the scale-invariant features transformation (SIFT) of the foreground targets; therefore, the misdetection situation caused by ignoring the attention of the target feature during the feature extraction is solved to some extent.

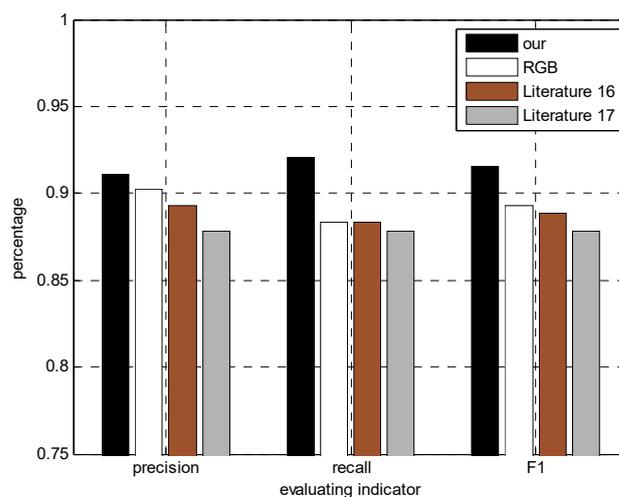


Figure 6. The results of the comparison [16,17].

In the second experiment, we randomly selected videos from the RAI dataset, the Open-Source Video dataset, and multiple sports videos. Experimental data are presented in Table 2. From Table 2, we can see that our algorithm was compared with the single feature algorithm, which is better than the single feature algorithm in both the recall accuracy and the recall rate because the features fusion algorithm can compensate for the shortcomings of a single feature. Secondly, this paper's algorithm was compared with Refs. [16,17]. The algorithm proposed in this paper is well improved compared with both Refs. [16,17]. Because the multi-step comparison scheme can calculate frame differences in multiple steps by setting the step, this scheme benefits from the detection of gradual transition.

5. Conclusions

In this paper, we proposed a multi-step comparison scheme shot boundary detection algorithm based on global features and target features. According to the experimental results, the accuracy and recall improved compared with other algorithms. The main reason is that the algorithm can solve the misdetection and missing detection due to the neglect of the target feature during feature extraction.

It can be found by comparison with the other literature that our algorithm develops the shot boundary detection algorithm, which solves the problem of ignoring the target features, and the recall and accuracy of the algorithm are improved compared with other algorithms. The innovation points are as follows:

1. In our method, by extracting the RGB color histogram global features of video frames and the scale-invariant feature transform (SIFT) target features. It can not only compensate for the misdetection of shot boundary detection caused by extracting only the global features while ignoring the detailed features but also compensate for the misdetection of shot boundary detection caused by extracting only the local features while ignoring the global changes.
2. We combined the Gaussian Mixed Model (GMM) algorithm to the field of shot boundary detection and then extracted the scale-invariant feature transform (SIFT) features and further improved the misdetection situation caused by ignoring the attention to the target features.

However, because of the limitations of the algorithm itself, the algorithm will have a better detection effect for a specific video, and finding a better shot boundary detection algorithm will further reduce the number of errors and omissions.

Author Contributions: Conceptualization, Q.L. and B.W.; methodology, Q.L.; software, J.L.; validation, Q.L., G.Z., X.C. and B.F.; formal analysis, Q.L.; investigation, X.C.; resources, B.F.; data curation, B.W.; writing—original draft preparation, Q.L.; writing—review and editing, G.Z.; visualization, B.W.; supervision, B.F.; project administration, Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Project of Shandong key R & D Program (Soft Science Project) (Grant No. 2021RK02002), Shandong Federation of Social Sciences under (Grant No. 2021-YYGL-32), Shandong Provincial Natural Science Foundation (Grant No. ZR2021QF056), National Natural Science Foundation of China (Grant No. 62071320).

Data Availability Statement: <https://aimagelab.ing.unimore.it/imagelab/researchActivity.asp?idActivity=19> (accessed on 15 April 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shanshan, L. *Improved Algorithm for Shot Mutation Detection Based on SIFT Feature Points*; Wuhan Polytechnic University: Wuhan, China, 2019.
2. Chakraborty, S.; Singh, A.; Thounaojam, D.M.J. A novel bifold-stage shot boundary detection algorithm: Invariant to motion and illumination. *Vis. Comput.* **2021**, *38*, 445–456. [[CrossRef](#)]
3. Xu, W.; Xu, L. Shot Boundary Detection Algorithm Based on Clustering. *Comput. Eng.* **2010**, *36*, 230–237.
4. Xi, C. *A Shot Boundary Detection Algorithm of MPEG-2 Video Sequence Based on Chi-Square Detection and Macroblocktype Statistics*; Shanghai Jiao Tong University: Shanghai, China, 2009; pp. 1–70.
5. Gygli, M. Ridiculously Fast Shot Boundary Detection with Fully Convolutional Networks. In Proceedings of the 2018 International Conference on Content-Based Multimedia Indexing, CBMI2018, La Rochelle, France, 4–6 September 2018; pp. 1–4. [[CrossRef](#)]
6. Souek, T.; Loko, J. TransNet V2: An effective deep network architecture for fast shot transition detection. *arXiv* **2020**, arXiv:2008.04838.
7. Raja Suguna, M.; Kalaivani, A.; Anusuya, S. The Detection of Video Shot Transitions Based on Primary Segments Using the Adaptive Threshold of Colour-Based Histogram Differences and Candidate Segments Using the SURF Feature Descriptor. *Symmetry* **2022**, *14*, 2041. [[CrossRef](#)]
8. Li, J.; Shao, C.F.; Yang, L.Y. Pedestrian detection based on improved Gaussian mixture model. *Jilin Daxue Xuebao (Gongxueban)/J. Jilin Univ. (Engine-Ering Technol. Ed.)* **2011**, *41*, 41–45.
9. Kekre, H.B.; Sonawane, K. Comparative study of color histogram based bins approach in RGB, XYZ, Kekre's LXY and L'X'Y' color spaces. In Proceedings of the International Conference on Circuits, San Francisco, CA, USA, 22–24 October 2014; IEEE: Piscatvie, NJ, USA, 2014.
10. Lihua, T.; Mi, Z.; Chen, L. Key frame extraction algorithm based on feature of moving target. *Appl. Res. Comput.* **2019**, *10*, 3138–3186.
11. Kailliang, G.; Tuanfa, Q.; Yuebo, C.; Kan, C. Detection of moving objects using pixel classification based on Gaussian mixture model. *J. Nan Jing Univ. (Nat. Sci.)* **2011**, *47*, 195–200.
12. Hannane, R.; Elboushaki, A.; Afdelk, K.; Naghabhushan, P.; Javed, M. An efficient method for video shot boundary detection and key frame extraction using SIFT-point distribution histogram. *Int. J. Multimed. Info-Rmation Retr.* **2016**, *5*, 89–104. [[CrossRef](#)]
13. Zonggui, C.; Xiaojun, D.; Lingrong, Z.; Yingjun, Z. Application of Improved SIFT Algorithm in Medical Image Registration. *Comput. Technol. Dev.* **2022**, *32*, 71–75.
14. Xuelong, H.; Yingcheng, T.; Zhenghua, Z. Video object matching based on sift algorithm. In Proceedings of the Conference Neural Networks and Signal Processing, Nanjing, China, 7–11 June 2008; IEEE: Piscatvie, NJ, USA, 2008; pp. 412–415.
15. Cai, C.; Lam, K.M.; Tan, Z. Shot Boundary Detection Based on a Multi-Step, Comparison Scheme. In Proceedings of the TRECVID 2005, Gaithersburg, MD, USA, 14–15 November 2005. Experiments in The Hong Kong Polytechnic University.
16. Meitei, T.D.; Thongam, K.; Manglem, S.K.; Roy, S. A genetic algorithm and fuzzy logic approach for video shot boundary detection. *Comput. Intell. Neurosci.* **2016**, *2016*, 8469428.
17. Apostolidis, E.; Mezaris, V. Fast shot segmentation combining global and local visual descriptors. In Proceedings of the IEEE International Conference on Acoustics, Florence, Italy, 4–9 May 2014; IEEE: Piscatvie, NJ, USA, 2014; pp. 6583–6587.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.