Applied Mathematics & Information Sciences
An International Journal

Effective Scene Change Detection by Using Statistical Analysis of Optical Flows

Jung Lee[1], Sun-Jeong Kim[2], and Chan Seob Lee[3]
[1]College of Information & Communication, Korea University, Korea
[2]Department of Ubiquitous Computing, Hallym University, Korea
[3]Department of Logistics & Distribution Information, Hyechon University, Korea
Email Address: airjung@gmail.com

Received: Received May 02, 2011; Revised July 25, 2011; Accepted September 12, 2011
Published online: 1 January 2012

Abstract: We present a novel method that exploits the statistical properties of optical flows to find representative video frames that contain scene change moments in video contents. For effective scene change detection, we first divide the optical flows into background and foreground groups. Optical flow is a useful and effective method for tracking object motion between consecutive video frames. By analyzing the variation of optical flows, we can detect rapid scene change between consecutive frames. A scene change probability for each frame is computed by applying some basic statistical methods, such as average and standard deviation. Starting from the selected frames with high probability, we find a clear image that contains no overlapping contents by inspecting the moment that optical flow values changes slowly and steadily. Experimental results show the robustness and effectiveness of our method.

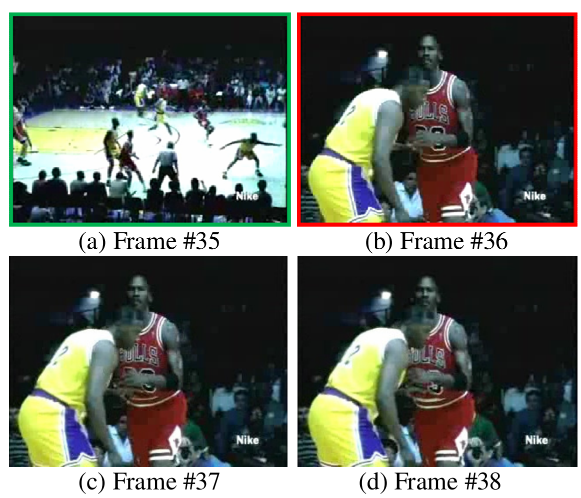
Keywords: Optical Flow, Scene Change Detection

1 Introduction

With the recent advances in multimedia creation and propagation technology, and the growth of network bandwidth, video content with high quality and large memory requirements is increasing explosively. It is therefore necessary to see a summary of a video's content before downloading it. In addition, many people want to skip parts of a video. There is thus a need for a framework that summarizes the video content and provides accurate scene change frames.

To achieve that goal, various methods have been proposed such as tag insertion and video frame analysis. We statistically analyze the change of optical flow that has normally been used for tracking objects in video scenes. Based on our analysis, we detect content-switching moments (called seed frames) and find scene-representative images (see Fig.1).

Extracted scene change frames can provide a simple summary of a video's content and can also be used to skip intermediate content and jump to the part we want to see.



(a) Frame #35 (b) Frame #36
(c) Frame #37 (d) Frame #38

Fig.1 An example of a scene change frame (b) extracted from consecutive video frames.

2 Related Work

2.1 Optical Flow

The goal of this paper is to extract scene change frames by tracking optical flows and analyzing them statistically. To do this, we first classify the optical flows into two groups, background and foreground. Optical flow means the motion patterns of the objects in the video or consecutive images. After its first introduction [3], many methods have been invented for its exact measurement and used in various image- and video-related areas such as computer vision and graphics.

Because the optical flow is essential to our approach, it is important to track the optical flows accurately and robustly. There are two representative optical flow calculation methods, Horn–Schunck and Lucas–Kanade. [5, 6]

Horn and Schunck introduced a global constraint to overcome this local problem [5]. By providing high-density optical flow, it is easy to see flow information inside objects that cannot be considered in the previous approaches. However, their method is weaker against noise than the local calculation methods.

The Lucas–Kanade technique is widely used to calculate optical flows for computer vision [6]. This technique performs basic optical flow calculation based on a least-squares method for the pixels within some defined distance. By combining information about pixel color and pattern distribution, it considerably reduces the correspondence ambiguity and noise sensitivity of the previous pointwise approaches.

In addition to the methods mentioned above, some methods using PDE(Partial Differential Equation) or nonlinear diffusion have been proposed [1, 2, 9, 10, 11]. In this paper, we use the Lucas–Kanade method for the optical flow calculation and extract scene change frames based on that calculation.

2.2 Scene-representative Frame Extraction

To summarize a video's contents and access some parts of them quickly, some video player or UCC sites use a simple method that divides the video in equal parts and use the frames at the moments of the divisions. However, the method is not guaranteed to represent the whole video contents properly.

The most intuitive and clear method among the video analysis methods is when the video creator attaches text annotations that identify and describe the scenes. However, it is very difficult to provide reliable information because of mismatches of viewpoint by subjective characteristics when video indexing. Much time and money is consumed because the database creation process cannot be performed automatically. These problems are critical factors for the enormous volume of newly created video contents.

To perform the video analysis process automatically, some researchers have proposed using multimedia properties such as color, shape, and texture. For automatic analysis methods, the most important thing is to extract characteristics that express the video properties well.

For effective structure analysis and characteristics extraction of video contents, Mandal et al. proposed a color moment-based indexing method applied by wavelet transform [7]. Smith and Chang classified contents by finding similar areas in consecutive frames based on the color distributions [13].

Hampapur and Shaharay established a mathematical model to judge the continuity and discontinuity between consecutive frames and thereby find the scene change moment [4, 12].

Yeo and Liu determined the motion continuity of the objects in a video scene by dividing every frame into constant-sized blocks and comparing corresponding blocks in consecutive frames [14]. This approach is not sensitive to noise, but slow scene changes are not properly extracted.

Nagasaka and Tanaka proposed a pixel comparison method that extracts scene change moments from an uncompressed video stream [8]. Most pixel comparison methods are weak when there is quick motion in the video because they consider only the pixels inside a user-specified range.

This paper uses a pixel-based method. We calculate optical flows first, and they are classified into the two groups - background and foreground. Then the separated optical flow groups are analyzed statistically to find scene discontinuities. Finally, the user-specified number of scene change frames is extracted by considering the adjacency of each scene change frame.

3 Optical Flow Segmentation

As mentioned earlier, we use the Lucas–Kanade method to estimate the optical flows in the input video. The estimated optical flows track both foreground objects (such as people and vehicles) and background objects (such as buildings and trees). However, the motion of foreground objects is often very different from that of background objects in direction and size. Thus, the estimated optical flows themselves do not help in analyzing the scene contents.

The foreground and background of the optical flow vectors were isolated from each of these targets to detect scene change moments. For the segmentation, we considered a number of videos to assess what would be essential for the segmentation in magnitude or direction. We found that the size of the movement was the dominant factor to distinguish them in most cases.

Based on the above findings, we perform a filtering process that separates the foreground and background by using the size of the optical flow vectors in each frame.

First, we calculate the average magnitude of all the M optical flow vectors and classify those vectors with magnitudes within the scope of the difference threshold compared with the average of the foreground. We set the threshold to the standard deviation. If the data follow a normal distribution, the portion of data from (average – standard deviation) to (average + standard deviation) will normally include approximately 68% of the values. We concluded that this was sufficient to separate the foreground optical flow vectors from the remainder. Assuming that the motion of the background optical flow vector is relatively small, the vectors having magnitudes below the threshold (average – standard deviation) are classified as background. The vectors above the threshold (average + standard deviation) are considered as noise data and excluded from the analysis process explained later. Fig.2 shows the robustness of our segmentation process.

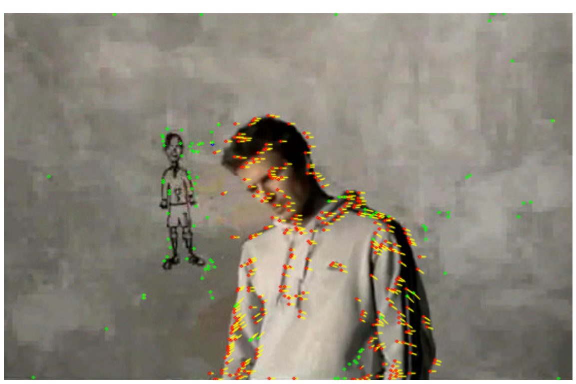


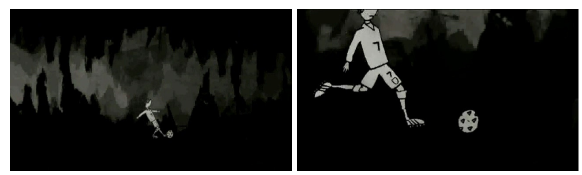
Fig.2 Optical flow segmentation example. (red dot: foreground, green dot: background, yellow arrow: optical flow)

4 Statistical Analysis of Optical Flows

4.1. Case Study : Finding a Scene Change Moment

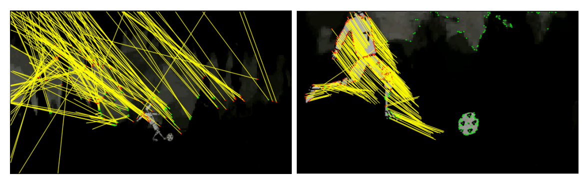
In the previous section, we separated the optical flow vectors of each frame in the foreground and background. For each group, we examine the changes between consecutive frames and extract a scene change moment based on the amount of change.

We divided scene change moments into two main cases. In the first case, rapid changes in both background and foreground indicate that the video changes to a new scene or that part of the contents is enlarged or reduced (see Fig.3). Most scene transitions occur in this way, and the objects and features between consecutive frames do not match properly. The magnitude and the direction of the computed optical flow vectors are therefore scattered randomly as shown in Fig.4.



(a) (b)

Fig.3 An example of rapid scene change resulting from the enlargement of contents.



(a) (b)

Fig.4 Optical flows corresponding to the frames of Fig.3 (red dot: foreground point, green dot: background point, yellow arrow: optical flow).

In the second scene change case, only the foreground objects disappear quickly and new ones are seen. In this case, the background optical vectors still have the consistent gradient pattern but the foreground ones are randomly scattered.

We can see that these two scene-changing cases show randomly distributed foreground optical flow vectors with random size. Therefore, the foreground vectors may be important for extracting scene change moments. We will compare these two types of optical flow vectors by average and standard deviation for each frame in Section 4.2 and 4.3.

4.2 Average of Optical Flow Vectors

In this section, we analyze the pattern of average magnitude changes of optical flow vectors. Fig.7 and 11-(a) show the average magnitude of optical flow vectors for two different video sources. We can see some sharp peaks where the value increases rapidly. They indicate the possible occurrence of scene change moments in those frames.

The values for the background vectors are smaller than those of the foreground vectors, but their patterns of change are very similar. There are thus no big differences between the foreground and background vectors to be used for finding scene change moments.

4.3 Standard Deviation of Optical Flow Vectors

We compute the standard deviations of the magnitudes of optical flow vectors from the source videos. Similar patterns to those of average analysis are also shown in this case. In Fig.11-(a), there is little noticeable change in the first half of the video. This is because the input video does not have much movement in either foreground or background. On the contrary, Fig.11-(b) clearly shows some sharp peaks (see the green squared areas) for the same video. This shows why the standard deviation represents the amount of distribution better than the average. From this observation, we confirm that the foreground optical flows react to scene changes more than the background ones. The foreground optical flow vectors will therefore be used for scene change detection.

5 Experimental Results

In the previous sections, we have analyzed the patterns of the average and standard deviation of optical vector magnitudes. Both patterns have sharp peaks at some frames. We consider a sharp peak as a seed for finding a scene change moment. A frame followed by a seed frame cannot be a scene change moment because the two frames have some overlapping images when connecting the different contents (see Fig.5-(b)), which are not appropriate to represent the video contents.

(a) (b) (c)

Fig.5 An example of an overlapping image (b) in consecutive frames.

We therefore first select a frame with a sharp peak as a seed and then find a scene change frame from the frames following the seed frame. To avoid the overlap problem, the scene change moment is picked when the change of metric value, average or standard deviation in this paper, has stabilized.

5.1 Scene Change Detection Based on Averages

(a) Frame #34, 35, 36

(b) Frame #254, 255, 256

(c) Frame #360, 361, 364

(d) Frame #550, 551, 552

(e) Frame #628, 629, 630

Fig.6 Scene change frames (red border) found based on the average magnitude of the foreground optical flow vectors. (green border: seed frame from green dots in Fig.7) Video source : http://www.youtube.com/watch?v=NseKug63naM

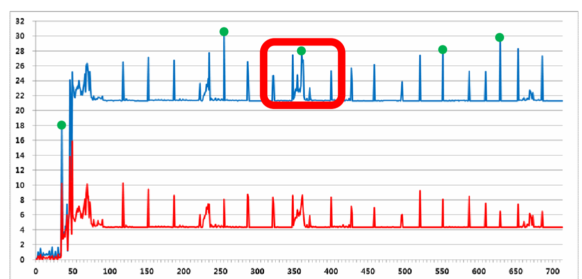
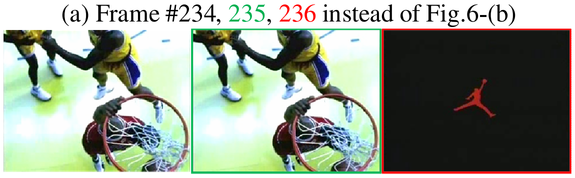


Fig.7 Averages of optical flow vector magnitudes in each frame (blue: foreground, red: background, x-axis: frame number)

Fig.6 shows the seed frames and the scene change frames by the average magnitude analysis. As seen in Fig.6-(c), the seed frame and its scene change frame may not be connected continuously. This is because the frames followed by the seed frame have overlapping contents, and the variation of the average magnitude is not stable, so the scene change frame is chosen when the variation is stabilized.

5.2 Scene Change Detection Based on Standard Deviations

(a) Frame #234, 235, 236 instead of Fig.6-(b)



(b) Frame #685, 686, 687 instead of Fig.6-(d)

Fig.8 Scene change frames (red border) newly found based on the graph of Fig.9. (green border: seed frame from green dots in Fig.9)

Fig.8 shows the results based on the standard deviation analysis of input video. As mentioned in Section 5.1, the seed frames are selected as the frames with sharp peaks, and the scene change frames are determined when the variation of the standard deviation value has stabilized. Some of these scene change frames were identified in Fig.6 ((a), (c), (e)), but the others are newly added.

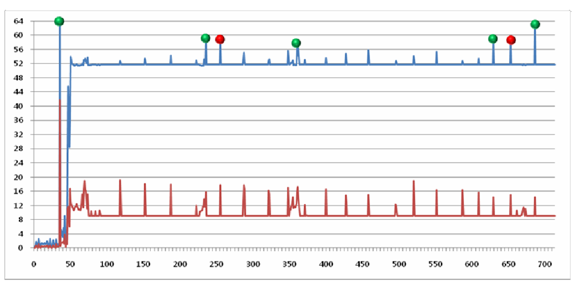
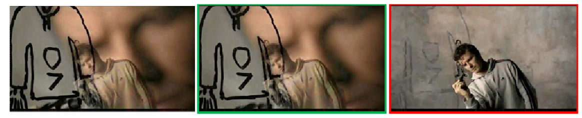


Fig.9 Standard deviations of the optical flow vector magnitudes in each frame

5.3 Discussion

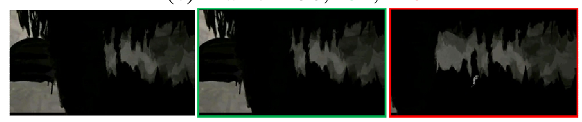
Fig.10 shows the scene change detection results for the graph of Fig.11-(b). In the input video, there are many frames with overlapping contents due to the slight movement between consecutive frames, so we skip them until we find a stable standard deviation value and pick that frame as the scene change frame.

Comparing the results of Fig.6 and 8, it is hard to say whether the average or the standard deviation is the better metric. Both detect some reasonable seed frames and find scene change frames, and the scene changes summarize the whole video contents effectively.

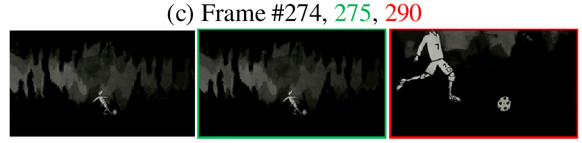


(a) Frame #63, 64, 79

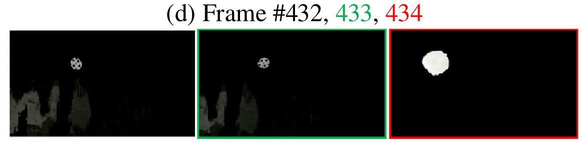
(b) Frame #130, 131, 146



(c) Frame #274, 275, 290

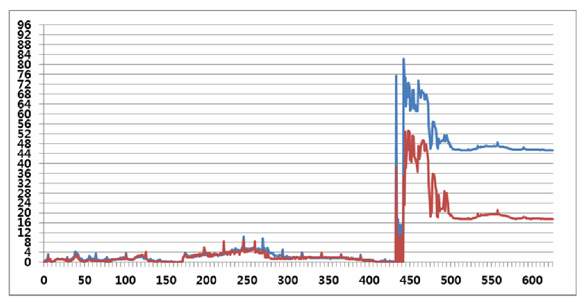


(d) Frame #432, 433, 434

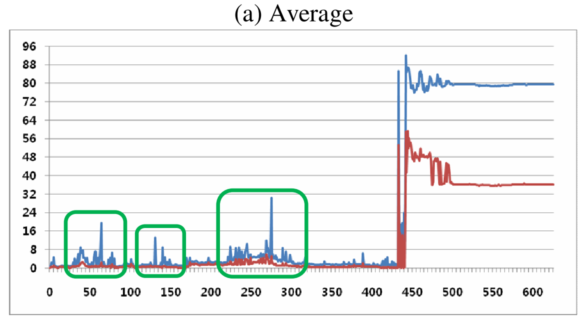


(e) Frame #471, 472, 486

Fig.10 Scene change frames (red border) found based on the graph of Fig.11-(b) (green border: seed frame).



(a) Average



(b) Standard deviation

Fig.11 Analysis of optical flow vector magnitudes in each frame. Video source : http://www.youtube.com/watch?v=25NK-dG5pBg

However, for videos with little movement, the standard deviation was more effective for detecting seed frames than the average (see the green square areas in Fig.11-(b)). Given the observation in Section 4.1, we can expect the standard deviation metric of foreground optical flow vectors to give us seed frames and scene change frames clearly and robustly regardless of the amount of motion in input videos.

In Fig.9, the frames with red dots have higher standard deviations than frame #360, which is picked as a seed frame, but they are not chosen. This is because they are so close to the seed frames with higher values that they might have contents similar to those of seed frames already selected. To avoid the problem of repeated contents, we set the least distance for adjacent seed frames as [(total number of video frames) / (4 * user-defined number of seed frames)]. With this limitation, seed frames are distributed at proper distances.

6 Conclusion

We have presented a novel method that detects scene change frames robustly by using statistical analysis of optical flows in the input video. Based on the case studies on optical flows and scene change metrics, we concluded that the standard deviation of the foreground optical flow vectors is suitable for scene change detection. The extracted scene change frames also summarize the contents of the input video well, and we can access each part of the video contents easily.

Our future research includes developing an advanced metric for scene change detection and introducing a novel method to express the flows of the video scene other than the optical flow.

Acknowledgements

This work was supported by the Second Brain Korea 21 Project.

This work was partly supported by the IT R&D program of MKE/KEIT[10035619-2011-02, Research and Development of Technologies for Digital Textile based Around Body Computing]

References

[1] Alvarez L., Weickert J., and Sanchez J., “Reliable Estimation of Dense Optical Flow Fields with Large Displacement,” Int. J. Computer Vision, vol.39, no.1, pp.41-56, 2000.

[2] Ben-Ari R. and Sochen N., “A General Framework and New Alignment Criterion for Dense Optical Flow,” IEEE Conf. Computer Vision and Pattern Recognition, vol.1, pp.529-536, 2006.

[3] Burton A. and Radford J., Thinking in Perspective : Critical Essays in the Study of Thought Processes, Routledge, 1978.

[4] Hampapur A., Jain R., and Weymouth T., “Digital Video Segmentation,” Proc. ACM Multimedia 94, pp.357-364, 1994.

[5] Horn B.K.P. and Schunck B., “Determining Optical Flow,” Artificial Intelligence, vol.17, pp.185-203, 1981.

[6] Lucas B.D. and Kanade T., “An Iterative Image Registration Technique with an Application to Stereo Vision,” Proc. Imaging Understanding Workshop, pp.121-130, 1981.

[7] Mandal M.K., Aboulnasr T., and Panchanathan S., “Image Indexing Using Moments and Wavelets,” IEEE Trans. Consumer Electronics, vol.42, no.3, pp.557-565, 1996.

[8] Nagasaka A. and Tanaka Y., “Automatic Video Indexing and Full Motion Search for Object Appearance,” Visual Database System, pp.113-127, 1999.

[9] Papenberg N., Bruhn A., Brox T., Didas S., and Weickert J., “Highly Accurate Optic Flow Computation with Theoretically Justified Warping,” Int. J. Computer Vision, vol.67, no.2, pp.141-158, 2006.

[10] Proesmans M., van Gool L., Pauwels E., and Oosterlinck A., “Determination of Optical Flow and Its Discontinuities Using Non-linear Diffusion,” In European Conf. Computer Vision, vol.801, pp.295-304, 1994.

[11] Scharr H. and Spies H., “Accurate Optical Flow in Noisy Image Sequences Using Flow Adapted Anisotropic Diffusion,” Signal Processing : Image Communication, vol.20, no.6, pp.537-553, 2005.

[12] Shahraray B., “Scene Change Detection and Content-Based Sampling of Video Sequences,” Proc. SPIE, vol.2419, pp.2-13, 1995.

[13] Smith J.R. and Chang S.F., "VisualSEEK : A Fully Automated Content-based Image Query System," ACM Multimedia 96, pp.87-98, 1996.

[14] Yeo B.L. and Liu B., “Rapid Scene Anlysis on Compressed Video,” IEEE Trans. CSVT, vol.5, no.6, pp.533-544, 1995.

Jung Lee is a research professor in the department of information & communication in Korea University. His research interests include Computer Graphics, Image Processing, and Image-based Modeling/Rendering.

Sun-Jeong Kim is an associate professor in the department of ubiquitous computing in Hallym University. Her research interests include Geometric Modeling, Scientific Visualization, Virtual Reality and Augmented Reality.

Chan Seob Lee is a professor in the department of logistics & distribution information in Hyechon University. His research interests include Media Information and Logistics.