전역적 정보를 이용한 샷 경계 검출

신성윤* · 신광성* · 이현창** · 진찬용** · 이양원*

*군산대학교, **원광대학교

Shot Boundary Detection Using Global Information

Seong-Yoon Shin* · Kwang-Sung Shin* · Hyun-Chang Lee** · Chan-Yong Jin** · Yang-Won Rhee*

*Kunsan National University, **Wonkwang University

E-mail : {s3397220, waver, ywee}@kunsan.ac.kr, hclglory@wku.ac.kr

요 약

본 논문은 프레임 차이 값을 카메라에서 브레이크로 인해 발생하는 높은 유사 콘텐츠의 경계 추출을 허용하는 글로벌 의사 결정 트리를 기반으로 샷 경계 검출 방법을 제공합니다. 처음엔 프레임 사이의 차이 값은 지역 X^2-히스토그램과 정규화를 통해 계산됩니다. 다음으로, 차이 값 사이의 거리는 정규화를 통해 계산됩니다.

ABSTRACT

This paper presents a shot boundary detection method based on the global decision tree that allows for extraction of boundaries of high variations occurring due to camera breaks from frame difference values. For a start, difference values between frames are calculated through local X2-histogram and normalization. Next, the distances between difference values are calculated through normalization.

키워드

의사 결정 트리(Decision Tree), 클러스터링(Clustering)

Ⅰ. Introduction

Processing of video is increasing rapidly in Ubiquitous environment. Videos are non-structured data, and they need to be structured according to a certain rule to provide ease of access. To ensure its structurization, therefore, a video needs to be semantically classified.

Ⅱ. Normalization

In this paper, the X2-histograms in [1][2] were used to detect scene changes. The Equation (1) below can be used to detect scene changes.

$$d(f_i, f_{i-1}) = \sum_{k=1}^{b} d_{\chi^2}(f_i, f_{i-1}, k) \tag{1}$$

$$d_{\chi^2}(f_i, f_{i-1}, k) = \sum_{j=1}^{n} \left(\frac{(H_i^r(j) - H_{i-1}^r(j))^2}{\max(H_i^r(j), H_{i-1}^r(j))} \times \alpha + \frac{(H_i^g(j) - H_{i-1}^g(j))^2}{\max(H_i^g(j), H_{i-1}^g(j))} \times \beta + \frac{(H_i^b(j) - H_{i-1}^b(j))^2}{\max(H_i^b(j), H_{i-1}^b(j))} \times \gamma \right.$$

In this context, bl refers to the total number of blocks, and Hir(k) refers to the difference value in the red-channel histogram between the block bl of the ith frame and the gray-level k. In addition, α, β, and ɣ are the constants used to convert grey-level intensity according to the NTSC standards where α=0.299, β=0.587, and ɣ=0.114.

As a means of normalization, the log functions and constants used in image processing to increase grey-level image values were modified and applied to the difference values. The Equation (2) below can apply to the proposed method.

$$d_{\log} = c \times \log(1 + d^2) \qquad (2)$$

$$c = \frac{\max(d_{\log})}{\max(\log(1 + d^2))}$$

The log function d from above is the difference value extracted from the Equation (1), and the constant c is a multiplier constant obtained from d.

Ⅲ. Global Shot Boundary Detection

Figure 1 illustrates the overall numerical expression that allows you to decide on a global scene. First off, d(fi, fi-1) is the expression for the X2-histogram in Expression (1) obtained as a result of the first shot boundary detection. In addition, dlog(i) refers to the normalized difference value, which corresponds to the ith frame in the (fi, fi-1) frames that are arranged in sequence as a result of the first shot boundary detection. Also, dlog(i-1) refers to the normalized difference value for the (i-1)th frame; bdlog(i) refers to the difference between itself and the difference value for the preceding frame (|dlog(i)-dlog(i-1)|); and fdlog(i) refers to the difference between itself and the difference value for the subsequent frame (|dlog(i)-dlog(i+1)).

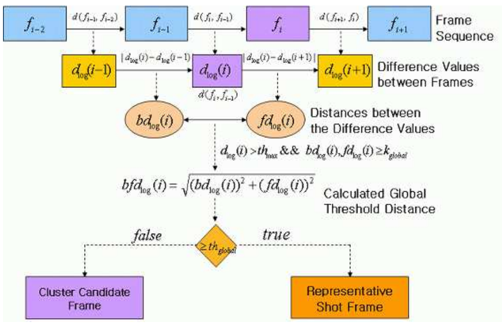

Fig. 1. The Global Decision Tree Structure

Ⅲ. Experiment

If a video has a lot of scene changes as in advertisement or entertainment videos, the distance between adjacent frames is relatively long. On the other hand, a news video that has few abrupt scene changes and more gradual scene changes features a relatively shorter distance between adjacent frames. As a result of application of the thresholds available for a news video to an advertisement or entertainment video, there were few problems with the global shot detection. In this paper, therefore, the global thresholds were obtained from the distance between the two adjacent difference values, and the thresholds showing no sensitivity to video types were determined.

Ⅳ. Conclusion

This paper demonstrated how abrupt scene changes like camera breaks were detected using the global decision tree. Shot frames were extracted in a way that scene changes were detected using the X^2-histogram from the inputted video, and that the difference values between adjacent frames were met through normalization. In addition, the proposed global decision tree allowed for extraction of scene changes, including the abrupt movements of an object and/or a camera, and flashlights.

Acknowledgement

"This research is partially supported by Institute of Information and Telecommunication Technology of KNU"

Reference

[1] S. Y. Shin, S. B. Pyo. "Video Browsing Using An Efficient Scene Change Detection in Telematics," Journal of the Korea Society of Computer and Information, Vol. 11, No. 4, pp. 147-154, 2006

[2] Y. L. Kim, Y. W. Rhee., "Scene Change Detection Using Local x^2-Test," Journal of the Korea Society of Computer and Information, Vol. 11, No. 3, 193-201, 2006