
지역적 정보를 이용한 장면 전환 검출

신성윤* · 진찬용** · 이양원***

Scene Change Detection Using Local Information

Seong-Yoon Shin* · Chan-Yong Jin** · Yang-Won Rhee***

요 약

본 논문은 지역의 의사 결정 트리와 클러스터링을 사용하여 장면 전환 검출 방법을 제시한다. 지역 의사 결정 트리는 클러스터 경계 검출 장면과 그 인접 프레임의 사이의 차이 값을 시간 유사 분포를 비교하기 위해 같은 방식으로 검출하고, 그리고 클러스터 단위로 차이 값의 유사성과 그룹 프레임의 끊어지지 않는 시퀀스를 감지한다.

ABSTRACT

This paper proposes a Scene Change Detection method using the local decision tree and clustering. The local decision tree detects cluster boundaries wherein local scenes occur, in such a way as to compare time similarity distributions among the difference values between detected scenes and their adjacent frames, and group an unbroken sequence of frames with similarities in difference value into a cluster unit.

키워드

의사결정 트리, 클러스터링, 장면 전환 검출, 프레임

Key word

Decision Tree, Clustering, Scene Change Detection, Frame

* 중신회원 : 군산대학교 컴퓨터정보공학과(제1저자, s3397220@kunsan.ac.kr) 접수일자 : 2012. 06. 01
** 정회원 : 원광대학교 정보·전자상거래학부 심사완료일자 : 2012. 06. 01
*** 정회원 : 군산대학교 컴퓨터정보공학과(교신저자)

I. 서 론

비디오 분할은 장면 전환 검출이라고도 하는데, 비디오를 계층적이고 구조적인 형태로 표현하기 위하여 영상, 문자, 오디오와 같은 매체 속에 포함되어 있는 내용들을 특징별로 분석하여 계층별로 분류하는 작업을 말한다. 최근의 많은 장면 전환 검출 방법들이 이미 수십 년 전부터 제시되어 왔다[1][2]. 가장 보편적인 장면 전환 검출 방법은 연속적인 프레임들 사이의 주어진 특징에 의해 표현되는 차이값을 계산하여 얻는 방법이다. 비록 정확률이 어느 정도 높다 해도 이들 알고리즘들의 강건성에 대한 제약들은 아직도 문제점으로 남아있다[3]. 다음으로 많이 활용하는 장면 전환 검출 방법은 히스토그램 비교 방법이다. 이 방법은 장면의 경계 검출을 위하여 사용되는 가장 보편화된 방법이다. Tono 등[4][5][6]은 그레이-레벨의 히스토그램 비교를 통하여 임계치를 기준으로 샷 경계를 추출하는 가장 간단한 방법을 제안하였다. Ueda 등[7]은 장면의 경계를 검출하기 위해서 컬러 히스토그램의 변화 비율을 사용했고, Naga 등[6]은 그레이 레벨과 컬러 히스토그램을 기반으로 한 몇 개의 간단한 통계학적 비교를 수행하였다. Zhan 등[4]은 픽셀 차이, 통계 차이, 그리고 몇 개의 히스토그램 방법을 비교하여 히스토그램 방법이 정확성과 속도사이의 좋은 교환요소를 발견하였다. Naga 등[6]은 두 프레임 사이의 차이 값을 강조할 뿐만 아니라 카메라나 객체의 움직임을 강조할 수 있는 X^2 -test를 제안하였다. 그러나 X^2 -test는 Tono 등[5]이 제안한 선형 히스토그램 비교방법보다 전체적인 성능이 더 좋지 않았으며, 계산량이 증가하는 단점을 갖는다. Garg 등[8]은 6개의 다른 컬러 좌표시스템(RGB, HSV, YIQ, $L^*a^*b^*$, $L^*u^*v^*$, Munsell)을 이용하여 3개의 히스토그램 기반방법에 대하여 성능을 평가하였다. Zhan 등[4]은 두 개의 임계치를 이용하여 점진적 장면변화를 검출할 수 있는 twin-comparison 방법을 제안하였으며, Bore 등[9]은 실시간 비디오에 대해서 비디오 분할기술을 비교한 결과 twin-comparison이 간단하며 매우 잘 적용될 수 있음을 발견하였다. 이 방법들은 간단하고 카메라나 객체의 움직임에 민감하지 않은 장점을 가지고 있지만 공간정보를 가지고 있지 않기 때문에 비교적 유사한 두 개의 다른 영상이 주어지면 샷 경계 검출에 실패하는 경우가 발생한다.

최근의 연구에서 보면 컬러 히스토그램과 X^2 -히스토

그램을 합성한 방법[10]을 이용하거나 급진적 장면부터 점진적 장면까지 모두 검출하는 강건하고 복합적인 지역적 X^2 -테스트를 이용한 방법[11]들이 연구되고 있다.

II. 장면 전환 검출과 정규화

컬러 히스토그램과 χ^2 히스토그램의 장점을 합성한 방법 1은 다음 식 (1)과 같이 정의하여 사용한다. 이는 각 컬러공간에 대한 명암도 등급변환을 위한 가중치를 적용하여 사용한 것으로 α, β, γ 는 NTSC 표준에 따른 명암도 등급변환을 위한 상수를 나타내며 본 논문에서는 $\alpha=0.299, \beta=0.587, \gamma=0.114$ 로 정의하여 사용하였다.

$$\begin{aligned} A &= H'_r(k) - H'_g(k) \\ B &= H'_s(k) - H'_b(k) \\ C &= H'_i(k) - H'_j(k) \\ d(f_i, f_j) &= \sum_{k=0}^{N-1} (A^* \alpha + B^* \beta + C^* \gamma) \end{aligned} \quad (1)$$

컬러 히스토그램과 χ^2 히스토그램의 장점을 합성한 방법 2는 다음 식 (2)와 같이 정의하여 사용한다.

$$\begin{aligned} A &= \frac{(H'_r(k) - H'_g(k))^2}{\max(H'_r(k), H'_g(k))} \\ B &= \frac{(H'_s(k) - H'_b(k))^2}{\max(H'_s(k), H'_b(k))} \\ C &= \frac{(H'_i(k) - H'_j(k))^2}{\max(H'_i(k), H'_j(k))} \\ d(f_i, f_j) &= \sum_{k=0}^{N-1} (A^* \alpha + B^* \beta + C^* \gamma) \end{aligned} \quad (2)$$

본 논문에서는 다음 식 (3)와 같이 A, B, C값을 제공하여 컬러 히스토그램과 χ^2 히스토그램의 장점을 극대화한 새로운 방법의 장면 전환 검출을 제시한다.

$$\begin{aligned} A &= \left(\frac{(H'_r(k) - H'_g(k))^2}{\max(H'_r(k), H'_g(k))} \right)^2 \\ B &= \left(\frac{(H'_s(k) - H'_b(k))^2}{\max(H'_s(k), H'_b(k))} \right)^2 \\ C &= \left(\frac{(H'_i(k) - H'_j(k))^2}{\max(H'_i(k), H'_j(k))} \right)^2 \\ d(f_i, f_j) &= \sum_{k=0}^{N-1} (A^* \alpha + B^* \beta + C^* \gamma) \end{aligned} \quad (3)$$

여기서 α, β, γ 는 방법 1과 방법 2와 똑같다. 또한 정규화를 방법은 영상처리에서 영상의 명암 값 향상을 위하여 사용되는 로그함수와 상수를 변형하여 차이 값에 적용하였다. 제안된 방식은 다음 식 (2)와 같다.

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

$$c = \frac{\max(d_{log})}{\max(\log(1 + d^2))} \quad (4)$$

여기서 로그함수 d 는 식 (3)으로부터 추출된 프레임 차이값이고 상수 c 는 d 로부터 계산된 배율상수이다.

III. 지역적 정보를 이용한 장면 전환 검출

본 논문에서 지역적 결정트리를 이용한 샷 바운더리 검출에서는 지역적 X^2 -히스토그램과 정규화 식을 이용하였다. 지역적 컷은 전역적 컷이 발생한 후, 발생되어지는 컷으로 샷이 검출된 후에 발생하는 다양한 장면의 변화를 찾기 위한 과정이다. 지역적 컷은 현재의 샷이 끝나는 시점, 즉 새로운 샷이 다시 검출할 때 까지 현재의 샷에 종속되어 수행되어진다. 또한 지역적 장면은 많은 수의 장면의 변화를 가지고 있기 때문에 지역적 임계치를 가지고 각각의 차이 값을 클러스터로 묶는 작업이 필요하다. (그림 1)에서는 이러한 과정을 진행하기 위하여 사용된 전체적인 수식구조를 그래프로 보여준다.

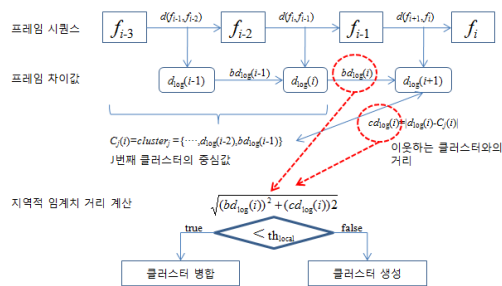


그림 1. 지역적 결정트리 구조
Fig. 1 Local Decision Tree Architecture

지역적 결정트리에서 이웃하는 프레임들 (f_i, f_{i-1})간의 i 번째 차이 값은 $d_{log}(i)$ 이며, $bd_{log}(i)$ 는 i 번째와 $i-1$ 번째

의 차이 값에 대한 절대차이($|d_{log}(i)-d_{log}(i-1)|$)를 나타낸다. $c_j(i)$ 는 m 번째 샷의 j 번째 클러스터를 나타내며, $c_j(i-1)$ 는 프레임 (f_{i-1})까지의 중심값을 나타낸다. $cd_{log}(i)$ 는 현재 프레임(f_i)과 j 번째 클러스터의 중심값 $c_j(i-1)$ 과의 절대차이를 계산하며, 지역적 임계치 거리 (th_{local})는 계산된 두 값($bd_{log}(i), cd_{log}(i)$)의 거리를 계산하여 얻을 수 있다.

지역적 결정트리는 먼저, 현재 프레임과 이웃하는 프레임간의 차이 값($bd_{log}(i)$)이 임의의 지역 상수 값(k_{local})보다 클 경우에는 프레임간의 장면변화가 있는 경우로 예측하여 새로운 클러스터를 생성하며, 프레임을 할당한다. 그러나 작을 경우에는 기존의 클러스터와의 중심값에 대한 거리($cd_{log}(i)$)와 현재 프레임의 이웃하는 프레임에 대한 차이($bd_{log}(i)$)의 크기를 계산하여 두 값의 거리 ($\sqrt{(bd_{log}(i))^2 + (cd_{log}(i))^2}$)가 임계치(th_{local})보다 작을 경우에는 기존 프레임을 클러스터($c_j(i)$)에 병합하며, 클 경우에는 새로운 클러스터($c_{j+1}(i)$)를 생성하여 할당한다.

IV. 실험

실험을 위하여 펜티엄 4-2.53GHz 컴퓨터와 Windows Xp professional에서 Visual C++6.0과 SDK로 DirectX8.1을 이용하였다. 또한 비디오 데이터로서 실험 비디오로서 15개의 비디오를 이용하였고 초당 프레임의 수는 30f/s, 비디오 사이즈는 320×240 으로 이용하였다.

제안된 비디오 정규화 방법은 추출된 차이 값의 범위를 일정한 범위로 압축하여 차이 값들을 보다 세분화하고, 군집화 과정을 보다 쉽게 해줄 수 있는 장점을 가지고 있다.

다음 표 1은 정규화 적용 전의 d 값과 후의 $d_{log}(i)$ 값에 대한 차이 값의 분포비교를 나타낸다.

표 1. 정규화 적용 전후의 값 비교
Table. 1 Value Comparison before/after Normalization

	최대 값		최소 값		평균	
	d	dlog	d	dlog	d	dlog
광고	213357.7	463.2	2136.3	236.5	8901.9	445.6
뉴스	222368.6	467.3	812.6	255.3	3874.6	338.6
스포츠	237086.2	423.2	3456.3	365.6	27482.9	387.6

그림 2는 정규화 적용 전과 후의 뉴스의 차이값 분포도를 나타낸다.

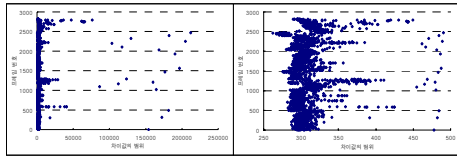


그림 2. 정규화 적용 전후의 차이값의 분포(뉴스)
Fig. 2 Distribution of Difference Value before/after Normalization(News)

그림 3에서는 차이 값들의 일부 구간에 대한 클러스터의 분포(C21-C27)와 그룹(G1-G5)의 구성을 보여준다.

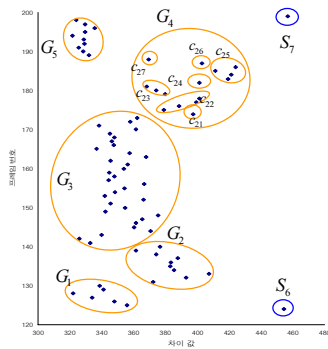


그림 3. 클러스터의 분포
Fig. 3 Cluster Distribution

그림 4는 일부 구간의 장면 그룹에 대한 클러스터의 구성과 대표 프레임들을 보여주기 위한 그림이다.

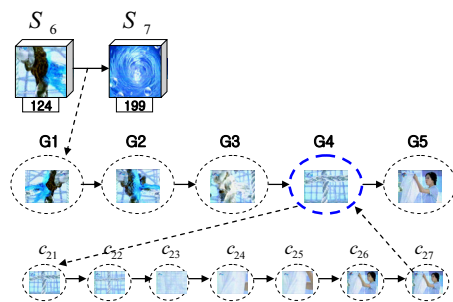


그림 4. 그룹에 대한 클러스터 내부 구성
Fig. 4 Interior Composition of Cluster for Group

V. 결 론

본 논문에서는 장면 전환 검출을 위하여 컬러 히스토그램과 지역적 X^2 -히스토그램을 합성한 방법을 이용하고, 이를 정규화 한다. 장면영역인 클러스터를 그룹화하기 위하여 먼저 지역적 결정트리를 이용하여 샷의 경계를 검출한다. 검출된 샷의 경계점은 하나의 장면이 시작되는 시작점이며, 지역적 결정트리는 이 장면의 시작점을 기준으로 프레임들의 차이 값들에 대한 유사성과 이웃하는 클러스터의 중심값을 기준으로 클러스터들을 구성하게 된다. 클러스터의 구성은 이어지는 또 다른 새로운 샷의 경계점이 검출될 때까지 샷에 종속되어 계속되며, 이러한 과정의 반복에 의하여 하나의 비디오로부터 여러 개의 샷과 그 샷에 종속된 많은 수의 클러스터를 생성하게 된다.

참고문헌

- [1] Koprinska and S. Carrato.: Temporal Video Segmentation: A Survey. Signal Processing Image Communication, Elsevier Science, 2001
- [2] G. Ananger, T.D.C. Little.: A survey of technologies for parsing and indexing digital video. Journal of Visual Communication and Image Representation, pp. 28-4, 1996
- [3] U. Gargi, R. Kasturi, and S. H. Strayer.: Performance Characterization of Video-Shot-Change Detection Methods. IEEE transaction on circuits and systems for video technology, Vol. 10, No. 1, Feb., 2000
- [4] Zhang, H. J., Kankanhalli, A., and Smoliar, S. W.: Automatic Partitioning of Full-motion Video. Multimedia Systems, Vol. 1, No. 1, pp. 10-28, 1993
- [5] Y. Tonomura.: Video handling based on structured information for hypermedia systems. in Proc. ACM Int. Conf. Multimedia Information Systems, pp. 333-344, 1991
- [6] Nagasaka, A. and Tanaka, Y.: Automatic Video Indexing and Full-Video Search for Object Appearances. in Visual Database Systems II, Knuth,

E., Wegner, L., Editors, Elsevier Science Publishers, pp. 113-127, 1992

[7] Ueda, H., Miyatake, T., and Yoshizawa, S.: IMPACT: An Interactive Natural-motion-picture Dedicated Multimedia Authoring System. in proceedings of CHI, 1991 ACM, pp. 343-350, 1991

[8] U. Gragi, R. Kasturi, S. Antani.: Evaluation of video sequence indexing and hierarchical video indexing. in Proc. SPIE Conf. Storage and Retrieval in Image and Video Databases, pp. 1522-1530, 1995

[9] J. S. Boreczky, L. A. Rowe.: Comparison of Video Scene Change Detection Techniques. in Proc, SPIE Conf. Storage and Retrieval in Image and Video Databases, pp. 170-179, 1996

[10] 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

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

저자소개



신성윤(Seong-Yoon Shin)

2003년 : 군산대학교
컴퓨터과학과(이학박사)
2006년 ~현재 : 군산대학교
컴퓨터정보공학과 교수

2009년~2010 : 한국정보통신학회 멀티미디어 및
컴퓨터그래픽스 분과위원장

2011년~현재 : 한국정보통신학회 재무상임이사

※ 관심분야 : Image Processing, Multimedia, Computer
Vision



진찬용(Kwang-Seong Shin)

고려대학교(학사), KAIST(석사)
서남대(경영학박사)
전 한국정보전략학회 회장
현재 : 원광대학교 정보전자
상거래학부/경영학부 교수

※ 관심분야 : 정보통신정책, MIS



이양원(Yang-Won Rhee)

1994년 : 숭실대학교
전자계산학과(공학박사)
1986년 ~현재 : 군산대학교
컴퓨터정보공학과 교수

※ 관심분야 : Telematics, Fuzzy Theory, Image
Processing