

효율적인 비디오 브라우징 제공을 위한 강건한 장면전환 검출 기법의 제안

이해균*, 이양원

군산대학교 컴퓨터정보공학과

Robust Scene Change Detection Technique for the Efficient Video Browsing Service

Lee Hae Gun, Rhee Yang Won

Kunsan National University

E-mail : haegun@krf.or.kr, ywrhee@kunsan.ac.kr

요 약

본 논문에서는 사용자에게 보다 효율적이고 직관적인 비디오 브라우징 서비스를 제공하기 위하여 비디오 데이터의 종류와 특성에 제한받지 않고 강건하게 적용될 수 있는 장면전환 검출 알고리즘을 제안하고자 한다. 제한된 알고리즘은 명암 값의 급 변화나 객체의 빠른 움직임, 영상의 왜곡 등에 의한 장면전환 검출의 오류를 제거할 수 있으며, 특히 연속된 프레임사이의 강건한 차이 값 추출을 위한 개선된 식을 제안하고, 추출된 차이 값들로부터 변화패턴을 학습하고 특징을 추출함으로써 자동 임계치 결정에 활용하였다. 제안된 방법은 급진적인 장면변화가 많고 플래시라이트와 같은 조명의 변화가 많은 다양한 비디오 데이터를 가지고 실험되었으며, 실험결과 기존의 방법에 비교하여 효율적이고 신뢰할 수 있는 결과 값을 보여주었다.

1. 서론

비디오는 실세계의 다양한 내용들을 함축적으로 담아낼 수 있는 유용한 도구로서 영상, 문자, 오디오와 같은 복합매체들로 구성된다. 특히 최근에는 컴퓨터와 통신, 그리고 데이터 압축 기술 및 표준 기술의 발달로 비디오를 이용한 다양한 서비스가 가능하게 되었으며, 초고속 통신망의 보급과 멀티미디어 처리 기술로 인해 비디오는 다양한 분야에서 폭 넓게 사용되고 있다[1]. 기존의 장면전환 검출은 크게 압축된 영역에서의 장면전환 검출과 비 압축된 영역에서의 장면전환 검출로 구분할 수 있다. 압축된 영역에서의 장면전환 검출은 압축된 영역으로부터 부분적으로 복원된 데이터만을 이용하여 장면전환 지점을 찾는 방법이며, DCT 계수를 이용한 방법[2]과 움직임 벡터를 이용한 방법[3] 등이 대표적으로 사용되고 있다. 그러나 이 방법은 부분적 데이터만을 이용하여 장

면전환 지점을 검출하기 때문에 데이터 손실에 따른 검출의 신뢰성을 떨어뜨리는 문제점을 갖는다. 비 압축된 영역에서의 장면전환 검출은 전체 영역으로부터 픽셀 비교에 의한 방법, 히스토그램의 비교에 의한 방법 등이 사용되며, 이중 히스토그램에 의한 비교 방법이 일반적으로 가장 널리 사용되는 방법이다[1].

본 논문에서는 플래시라이트를 제거하고 주어진 비디오로부터 효율적인 장면전환 추출을 하기 위하여 다음과 같이 제안한다. 먼저 연속된 프레임으로부터 차이 값 추출을 위하여 지역 히스토그램에 의한 비교 방법을 사용하고 있다. 이것은 객체나 카메라의 움직임에 민감하지 않고 프레임간의 공간 정보를 활용할 수 있는 장점이 있다. 추출된 차이 값은 변이 폭이 넓게 형성되는 특징이 있으며, 이것은 장면전환 검출을 위한 임계 값 결정에 많은 어려움을 준다. 본 논문에서는 이러한 변이 폭이 넓게 형성되는 차이 값의 폭을 줄이고 임계값 결정의 효율성

을 높이기 위하여 로그 식을 이용한 차이 값의 압축에 의한 정규화 작업을 수행한다. 정규화 된 차이 값은 일정한 크기로 압축되어 임계값 결정에 많은 장점을 준다. 제안된 장면전환 검출 알고리즘은 정규화 된 차이 값들로부터 차이 값들의 시간적 연속성에 대한 크기와 연결정보를 고려하여 플래시라이트를 제거하며 효과적으로 장면전환 지점을 검출할 수 있다.

2. 본론

1. 지역 히스토그램 비교에 의한 차이 값 추출

본 논문에서는 연속된 프레임 사이의 강건한 차이 값 추출을 위하여 기존의 히스토그램 기반 알고리즘과 비교하여 좋은 성능을 보여주는 χ^2 -테스트를 변형하여 이용하였으며, 컬러 값의 세분화 작업에 따른 검출효과를 높이기 위하여 명암도 등급에 따른 가중치를 적용한 컬러 히스토그램 비교를 적용하였다. [huan01, ko06, Garg95].

컬러 히스토그램 비교($d_{r,g,b}(f_i, f_j)$)는 인접한 두 프레임(f_i, f_j)의 각 컬러공간에 대한 히스토그램 비교를 통하여 계산되어지며 식(1)과 같이 정의하여 사용한다.

$$d_{r,g,b}(f_i, f_j) = \sum_{k=0}^{N-1} (|H_i^r(k) - H_j^r(k)| + |H_i^g(k) - H_j^g(k)| + |H_i^b(k) - H_j^b(k)|) \quad \text{식(1)}$$

$H_i^r(k), H_i^g(k), H_i^b(k)$ 는 i 번째 프레임(f_i)에서의 각 컬러 공간(r, g, b)에 대한 빈(k)의 수(N)를 나타낸다.

식(1)로부터 각 컬러공간에 대한 명암도 등급변환을 위한 가중치를 적용하여 식(2)와 같이 재정의하여 사용할 수 있다.

$$d_{wr,wg,wb}(f_i, f_j) = \sum_{k=0}^{N-1} (|H_i^r(k) - H_j^r(k)| * \alpha + |H_i^g(k) - H_j^g(k)| * \beta + |H_i^b(k) - H_j^b(k)| * \gamma) \quad \text{식(2)}$$

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.

두 프레임사이의 차이를 강조하기 위한 통계 분석방법 중에서 χ^2 -테스트 비교($d_{w\chi^2}(f_i, f_j)$)는 히스토그램의 비교변화를 통하여 장면전환을 검출할

수 있는 효율적인 방법이며, 식(3)과 같이 정의하여 사용한다.

$$d_{w\chi^2}(f_i, f_j) = \begin{cases} \sum_{k=0}^{N-1} \frac{(H_i(k) - H_j(k))^2}{\max(H_i^r(k), H_j^r(k))} & \text{if } (H_{i,j} \neq 0) \\ 0 & \text{otherwise} \end{cases} \quad \text{식(3)}$$

히스토그램 기반 방법은 공간정보를 적용하지 않아 유사한 컬러분포를 갖는 다른 두 영상인 경우 같은 영상으로 간주하여 검출하는 문제점이 발생할 수 있다. 이러한 문제점은 프레임의 영역을 분할하여 지역에 대한 히스토그램 분포를 비교하는 방법으로 해결할 수 있다.

식(4)에서는 영역의 분할에 따른 각 영역의 컬러 히스토그램 비교와 누적을 통한 차이 값 생성을 나타내고 있다.

$$d(f_i, f_j) = \sum_{bl=1}^m DP(f_i, f_j, bl) \quad \text{식(4)}$$

$$DP(f_i, f_j, bl) = \sum_{k=1}^{N-1} |H_i(k, bl) - H_j(k, bl)|$$

$H_i(k, bl)$ 는 프레임(f_i)에 대한 블록(bl)의 k 번째 히스토그램의 분포를 나타내며, m 은 전체 블록의 수를 나타낸다.

위의 식(2)에와 같이 각 컬러공간에 대한 가중치를 적용한 세분화된 히스토그램 비교와 식(3)의 통계방식을 이용한 차이 값의 확장 및 식(4)의 지역히스토그램에 의한 프레임의 공간정보에 대한 활용의 장점을 이용하여 본 논문에서는 차이 값 추출에 대한 강건함과 신뢰성을 위하여 다음 식(5)와 같이 위의 식들을 결합한 지역 히스토그램 비교에 의한 차이 값 추출식을 다음과 같이 제안한다.

$$d(f_i, f_j) = \sum_{bl=1}^m d_{\chi^2}(f_i, f_j, bl) \quad \text{식(5)}$$

$$d_{\chi^2}(f_i, f_j, bl) = \sum_{k=1}^{N-1} \left(\frac{(H_i^r(k) - H_j^r(k))^2}{\max(H_i^r(k), H_j^r(k))} \times \alpha + \frac{(H_i^g(k) - H_j^g(k))^2}{\max(H_i^g(k), H_j^g(k))} \times \beta + \frac{(H_i^b(k) - H_j^b(k))^2}{\max(H_i^b(k), H_j^b(k))} \times \gamma \right)$$

위의 식(5)에서 $H_i^r(k), H_i^g(k), H_i^b(k)$ 는 i 번째 프레임(f_i)이 갖는 각 컬러 공간(r, g, b)에 대한 히스토그램의 분포를 나타내며, N 은 빈(k)의 전체 수, m 은 블록(bl)의 전체 수를 나타낸다.

본 논문에서는 제안된 식(5)으로부터 프레임을 같은 영역의 블록으로 분할한 후, 각 블록들에 대한 히스토그램 비교를 통하여 차이 값을 생성하였다. 생성된 차이 값은 급진적 장면의 변화나 점진적 장면의 변화 모두에 적용할 수 있는 강건한 차이 값의 추출을 보여준다.

2. 차이 값의 정규화

제안된 식(5)에 의하여 추출된 차이 값은 프레임간의 특징정보에 따라 편차가 크게 나타나며, 프레임간의 연속적인 연결정보를 얻기가 매우 어렵다. 특히 급진적 장면전환이나 점진적인 장면전환을 추출하기 위한 임계치의 결정은 각 차이 값들의 변화에 능동적으로 대처하여 결정해야 하는 문제점을 갖는다. 따라서 차이 값의 편차를 줄이고 시간적으로 연결된 차이 값들의 변이를 쉽게 파악하고 정보를 얻기 위한 방법이 필요하다. 기존의 프레임에 대한 전체 픽셀수를 이용한 정규화 방법은 차이 값들에 대한 폭을 일정한 영역으로 줄여 사용하였으나 차이 값들의 시간적 연속성과 연관성에 대한 정보는 제공해 주지 못하는 단점을 가지고 있다. 본 논문에서는 차이 값들을 일정한 값의 범위 안에 동적으로 압축하여 차이 값들로부터 급격한 장면변화를 추출할 수 있는 강건한 임계치 결정의 지원은 물론 차이 값들의 시간적 연속성과 연관성을 쉽게 파악하여 점진적인 장면의 변화도 추출할 수 있는 차이 값의 정규화 방법을 제안하고자 한다. 제안된 방법은 영상처리에서 영상의 명암 값 향상을 위하여 사용되는 로그함수와 배율상수를 변형하여 차이 값에 적용하였다.

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

$$c = \frac{\max(d_{\log})}{\max(\log(1 + d^2))} \quad \text{식(6)}$$

c 는 배율상수이고, d 는 차이 값 추출 식(5)에 의하여 추출된 차이 값을 나타낸다. 차이 값의 제곱은 동적 범위로 차이 값을 나타내기 위하여 수행되어진다.

3. 제안된 장면전환 검출 알고리즘

본 논문에서는 하나의 카메라에 의해서 연속적으로 촬영된 프레임들의 구역을 장면으로 정의하며, 장면전환 검출은 이러한 구역을 구별하는 나누는 것을 목적으로

한다. 제안된 방법은 연속되는 프레임들로부터 차이 값을 구하고, 정규화 작업을 실시한 후에, 입력되는 차이 값들의 특징으로부터 플래시라이트를 제거한 후, 장면전환의 존재여부를 판단하게 된다. 플래시라이트에 의한 장면전환의 잘못된 검출의 문제점을 해결하기 위하여 본 논문에서는 인접한 프레임과의 차이 값들에 대한 연속성 여부를 측정하여 사용한다. 차이 값의 연속성이란 차이 값들이 일정한 거리를 두고 서로 유사하게 연속적으로 분포하는 특성을 말하며 이것은 프레임들이 서로 유사한 정보들로 구성되는 것을 말한다.

제안된 장면전환 검출 알고리즘

Step 1. 주어진 비디오로부터 차이 값 추출식(5)을 이용하여 프레임(f_i, f_j)의 j 번째 차이 값($d(j)$)을 계산한다.

Step 2. 추출된 차이 값($d(j)$)으로부터 정규화식(6)을 적용하여 정규화 된 차이 값($d_{\log}(j)$) 생성한다.

Step 3. 정규화 된 차이 값($d_{\log}(j)$)으로부터 다음 조건을 비교하여 장면전환 검출 지점을 계산한다.

$$\text{조건 (1)} \quad d_{\log}(j) \geq th_{\max}$$

정규화 된 차이 값은 전역적 임계치(th_{\max})를 만족해야 한다. 전역적 임계치는 장면전환지점으로 추출된 차이 값들의 평균값으로부터 계산되었다.

$$\text{조건 (2)} \quad \begin{cases} bd_{\log}(j) \geq th_{\min} \\ (bd_{\log}(j) = |d_{\log}(j) - d_{\log}(i)|) \end{cases}$$

이전과 현재의 차이 값($bd_{\log}(j)$)은 지역적 임계치(th_{\min})를 만족해야 한다.

$$\text{조건 (3)} \quad \begin{cases} fd_{\log}(j) \geq th_{\min} \\ (bd_{\log}(j) = |d_{\log}(j) - d_{\log}(j+1)|) \end{cases}$$

현재와 이후의 차이 값($fd_{\log}(j)$)은 지역적 임계치(th_{\min})를 만족해야 한다.

$$\text{조건 (4)} \quad bfd_{\log} = \sqrt{(bd_{\log}(i))^2 + (fd_{\log}(i))^2} \geq th_{\text{global}}$$

차이 값 간의 유동성 있는 거리 값의 조정을 위하여 두 차이 값들 간의 거리는 일정한 임계치(th_{global}) 이상을 만족해야 한다.

Step 4. Step 3의 조건(1) ~ 조건(4)를 모두 만족하는 프레임은 장면전환 지점으로 검출되며, 조건중 하나라도 만족하지 못할 경우에는 Step 1로 넘어가 연속된 다음 프레임으로부터 새로운 차이 값을 생성한다.

3. 결론

본 논문에서는 주어진 비디오로부터 연속된 프레임사이의 차이 값 추출을 위하여 객체나 카메라의 움직임에 덜 민감하고 프레임의 공간정보를 이용하는 지역 히스토그램 비교에 의한 추출 방식을 사용한다. 또한 추출된 차이 값의 큰 변이 폭에 의한 임계 값 결정의 문제점을 해결하기 위하여 차이 값들의 동적 압축에 의한 정규화 작업을 수행한다. 제안된 장면전환 검출 알고리즘은 정규화 된 차이 값들의 시간적 연속성의 변이에 따른 플래시라이트의 특징을 이용하여 플래시라이트와 장면전환 검출을 따로 구분하여 추출할 수 있는 새로운 장면전환 검출 알고리즘이다. 제안된 방법은 기존의 장면전환 검출에서 문제가 되었던 여러 부분들을 효과적으로 해결하고 있다. 먼저 기존의 카이-테스트와 비교하여 차이 값 추출에 좋은 성능을 보여주고 있으며, 다양한 비디오 장르에 따른 임계치 결정의 문제점을 정규화 된 차이 값으로부터 일정한 범위내의 차이 값을 이용하여 임계치를 설정함으로써 해결하고 있으며, 플래시라이트에 의하여 발생할 수 있는 잘못된 장면전환 검출의 문제점도 해결하고 있다. 따라서 제안된 방법은 실시간으로 입력되는 비디오를 분석하여 장면전환 검출을 보다 강건하게 지원함으로써 비디오 자동 분류를 위한 기반작업의 효율성을 높여줄 수 있음은 물론 비디오의 일반적 특성을 통한 색인을 지원함으로써 다양한 검색을 요구하는 사용자의 의도에 보다 적절한 대응을 시도할 수 있는 시스템 개발에 적절하게 응용될 수 있다.

[참고문헌]

- [1] C. L. Huang and B. Y. Liao, "A Robust Scene Change Detection Method for Video Segmentation," IEEE Trans on CSVT, Vol. 11, No. 12, pp. 1281-1288, December 2001.
- [2] U. Gargi, R. Kasturi, and S. H. Strayer, "Performance Characterization of Video Shot Change Detection Methods," IEEE Trans on CSVT, Vol. 10, No. 1, pp. 0001-0013, February 2000.
- [3] H. Zhang, A. Kankamhalli, and S. Smoliar, "Automatic partitioning of full-motion video," ACM Multimedia Systems, New York: ACM Press, Vol. 1 1993, pp. 10-28.
- [4] C. L. Huang and B. Y. Liao, "A Robust Scene-Change Detection Method for Video Segmentation," IEEE Trans. Circuits System. Video Technology, Vol.11, No.12, December 2001.
- [5] K. C. Ko and Y. W. Lee, "Scene Change Detection using the Chi-test and Automated Threshold Decision Algorithm," ICCSA06, Vol. 3983 2006, pp. 1060-1069.
- [6] 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, 1995, pp. 1522-1530.
- [7] Gonzalez, "Digital Image Processing 2/E (S/C)," Prentice-Hall, 2002.
- [8] Ralph M. Ford, Craig Robson, Daniel Temple, Michael Gerlach, "Metrics for shot boundary detection in digital video sequences," Multimedia Systems 8: 37-46(2000).
- [9] A. Ekin, A. M. Tekalp, and R. Mehrotra, "Automatic soccer video analysis and summarization," IEEE Trans. on Image Processing, Vol. 12, No. 7, pp. 796-807, July 2003.
- [10] I. Koprinska and S. Carrato, "Temporal Video Segmentation: A Survey," Signal Processing Image Communication, Elsevier Science 2001.