

컬러 히스토그램과 픽셀 강도를 이용한 동영상 Shot 검출 방법

최준호*, 이홍렬**, 김판구**

*조선대학교 문화콘텐츠기술연구소

**조선대학교 컴퓨터공학과

e-mail : xdman@paran.com

A Method of Video Shot Detection Using Color Histogram and Pixel Intensity

Jun-Ho Choi*, Hong-Ryul Lee**, Pan-Koo Kim**

*Culture Technology Institute, Chosun University

**Dept of Computer Engineering, Chosun University

요 약

멀티미디어 데이터 중 가장 많은 정보를 내포하고 있는 동영상의 정보를 효과적으로 검색하기 위해서는 동영상 스트림의 내용이 장면별로 인덱스 되고 관리되어야 한다. 이를 위해 동영상 데이터의 가장 기본적인 단위인 Shot 추출의 자동화 처리는 동영상을 이용한 시스템의 가장 중요한 작업이다. 이에 본 논문에서는 기존의 Shot 검출 방법을 보완하기 위해 동영상의 특징 요소인 픽셀값, 컬러 히스토그램, 경계면 픽셀 수 등을 추출하여 이에 대한 유사도 측정을 통한 Shot 검출 방법을 제안하고자 한다.

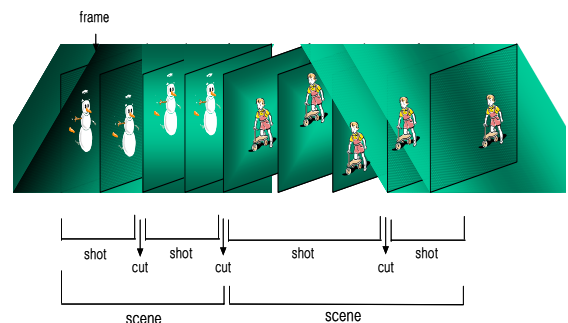
1. 서론

최근 멀티미디어 관련 기술의 발전과 인터넷 사용의 증가로 동영상의 사용이 증가함에 따라 동영상을 효과적으로 검색하는 방법이 주요 관심분야로 대두되고 있다. 멀티미디어 데이터 중 1)가장 많은 정보를 내포하고 있는 동영상의 정보를 효과적으로 검색하기 위해서는 동영상 스트림의 내용이 장면별로 인덱스 되고 관리되어야 한다. 이를 위해 동영상 데이터의 가장 기본적인 단위인 Shot 추출의 자동화 처리는 동영상을 이용한 시스템의 가장 중요한 작업이다. 이에 본 논문에서는 기존의 Shot 검출 방법을 보완하기 위해 동영상의 특징 요소인 픽셀값, 컬러 히스토그램, 경계면 픽셀 수 등을 추출하여 이에 대한 유사도 측정을 통한 Shot 검출 방법을 제안하고자 한다.

할 수 있다. 따라서 일반적으로 물리적 단위인 컷(Cut)을 검출하는 작업이 수행된다. 동영상 데이터의 구조를 살펴보면 다단계 계층적 구조로 표현될 수 있다. 동영상 데이터는 그 양이 방대할 뿐 아니라, 순차적인 특성으로 인하여 정보의 검색 및 브라우징이 대단히 곤란하다는 문제점을 가지고 있다. 일반적인 동영상 브라우저를 생각할 경우, 동영상의 필요한 부분에 접근하기 위해서는 수많은 고속검색을 반복하여야 하며, 이는 시간적인 면에서 동영상 브라우징을 크게 저해하는 요소가 되고 있다. 이와 같은 문제점을 해결하고 동영상의 접근성을 개선하기 위한 방법으로 (그림 1)과 같은 계층적인 데이터 모델이 제시되어 있다. 전체 동영상을 시퀀스(sequence), 신(scene), 샷(shot) 단위로 분할한다.

2. 동영상의 Shot 추출 방법

동영상 데이터는 Scene 단위의 비디오의 이벤트(event) 단위로 나눌 수 있다. 비디오를 Scene 단위로 나누는 것이 가장 이상적인 방법이지만 Scene은 의미그룹으로서 자동으로 Scene을 검출하는 것은 매우 어려운 작업이라고



(그림 1) 동영상의 구조

1) 본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소(CT)육성사업의 연구결과로 수행되었음.

일반적인 동영상의 Shot 추출 방법은 다음과 같다.

① 픽셀차(Pixel Difference)를 이용한 방법

비디오 장면 전환 검출을 위해서 각 프레임의 픽셀값을 이용하여 프레임간의 픽셀값의 차이는 매우 낮은 값을 갖는다. 그러나 장면 전환이 있는 곳은 픽셀값들의 차이가 두드러지게 커지므로 이를 이용하면 장면 전환점을 찾아낼 수가 있다. 이러한 방법은 프레임의 잡음(noise)에 민감하고 특히 점진적인 장면 전환 기법은 검출이 거의 되지 않는 취약점을 보인다.

② 명도 히스토그램(Histogram)을 이용한 방법

이 방법은 가장 널리 사용되는 장면 전환 기법 중의 하나로 프레임의 명도 히스토그램을 찾아 프레임간의 히스토그램의 차이를 비교하여 일정한 임계값 이상이면 장면 전환으로 인식하는 기법으로 픽셀차이를 이용하는 방법보다는 잡음(noise)에 강한 성질을 갖는다. 하지만 이 방법 또한 점진적 장면 전환을 찾는 것에는 취약점을 보인다.

③ DCT 계수를 이용한 방법

압축영상(MPEG)의 일반적 방법인 DCT 변환의 계수를 이용하는 방법이 있다. 이는 압축영상의 스트림에 존재하는 계수의 차이를 이용하는 방법으로 압축영역에서 복원에 필요한 많은 계산을 거치지 않고 바로 장면 전환을 검출하는 특징을 갖고 있다.

3. Shot 추출을 위한 특징 요소 유사도 측정

동영상에서 연속적인 프레임은 서로 유사한 배경과 영상의 변형으로 이루어진다. 영상 정보의 연속성의 편차가 큰 부분이 추출의 대상이 되는 Cut 프레임일 가능성이 높다. 연속되는 두 프레임에서 영상을 블록화하여 각 블록에 대한 히스토그램을 추출한 후, 히스토그램의 평균값을 구하고 그에 대응하는 다음 프레임의 블록과 비교하여 유사도를 측정하게 된다. 동영상의 장면 유사도의 평균을 구하여 일정 오차범위에 해당되면 이를 Cut 프레임으로 판단하게 된다. 연속적인 장면 전환점에서의 프레임 추출 오류를 해결하기 위해 연속된 프레임에 오차범위를 주어 최적의 프레임 추출 결과를 얻을 수 있다.

(1) 동영상의 특징 요소 유사도

본 연구에서 제시하는 Shot 추출 방법은 다음과 같이 픽셀의 밝기, 히스토그램, 경계면 픽셀 수의 차이를 이용한다. 각 특징 요소별 계산 방법은 다음과 같이 정의하였다.

가) 픽셀값 유사도

$$D_p = \frac{\sum (I_m(x,y) - I_{m+1}(x,y))}{N} \quad (\text{수식 1})$$

(N : 픽셀수, I : 프레임 좌표)

나) 히스토그램 유사도

$$D_h = \frac{\sum_{i=1}^k |H_m(i) - H_{m+1}(i)|}{K} \quad (\text{수식 2})$$

(H : 히스토그램, K : 픽셀값 범위)

다) 경계면 픽셀 수

$$D_e = \frac{|E_m - E_{m+1}|}{N} \quad (\text{수식 3})$$

(Em : m번째 픽셀 수, N : 픽셀수)

본 연구에서는 두 개의 연속되는 프레임 사이의 특징 요소의 차이를 구하고, 이를 각 프레임 별로 유사도 측정을 수행한 후 특징값을 각 프레임의 픽셀의 분산값을 구한 후 처리한다.

$$V = \frac{\sum_{i=1}^k H(i) \cdot (i - M)^2}{N} \quad (\text{수식 4})$$

(H : 히스토그램, N : 픽셀수)

키 프레임 추출을 위해 전 프레임과 현재 프레임을 비교하여 유사도 오차범위 내에서 영상의 변화가 크면 이를 Shot으로 인식한다. 이전 프레임과 현재 프레임의 두 영상의 유사도 측정으로 차영상을 얻게 되는데, 추출된 차영상에서는 모든 픽셀의 값을 합하여 오차 범위를 구하는데, 차 값이 클수록 영상의 변화폭이 큰 것으로 간주되고, 임의의 오차범위를 정하여 그 이상이 되면 키 프레임으로 간주하게 된다.

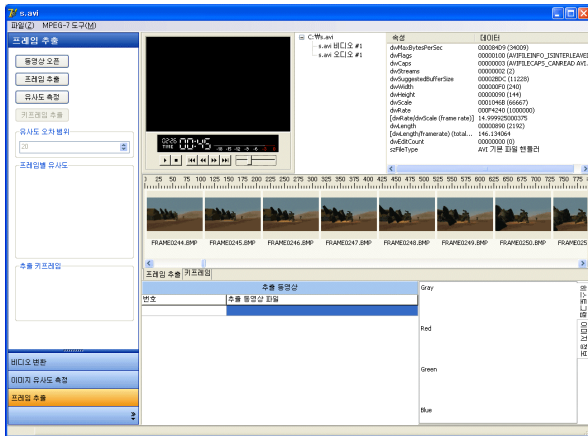


(그림 2) 특정 요소 유사도 측정 결과

(2) 특정 요소 유사도를 이용한 Shot 추출

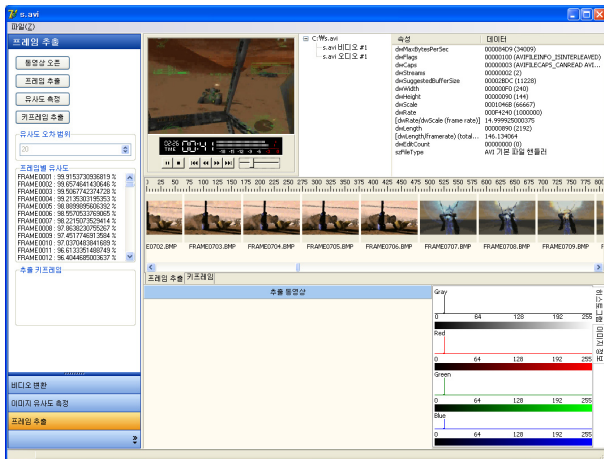
동영상의 Shot 추출은 동영상의 분석의 첫 단계이고, 이는 동영상의 내용 인식 과정에서 가장 중요한 부분이다. Shot 추출을 위해서는 먼저 객체의 큰 움직임이나 카메라의 움직임, 밝기의 변화 등에 큰 오차를 보이지 않는 신뢰성이 확보되어야 하고, 연속 프레임의 비교로 구현할 수

없는 연속적인 Shot 경계를 추출한다.



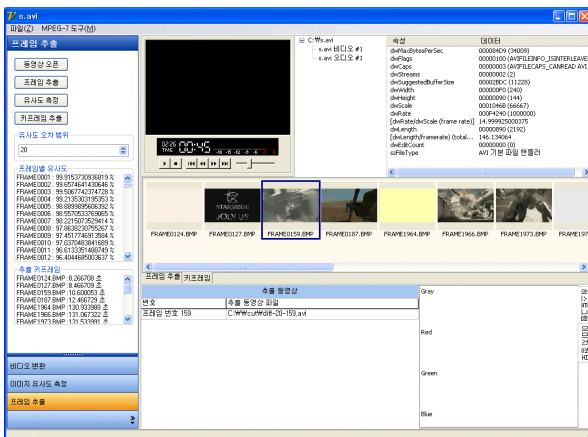
(그림 3) 동영상의 프레임 추출 결과

(그림 3)에서는 입력된 동영상에 비디오와 오디오 부분으로 분리된 후 비디오를 프레임별로 분리되는 과정을 보여주고 있다.



(그림 4) 프레임 유사도 측정 결과

(그림 4)는 분리된 프레임별로 본 논문에서 제안한 유사도를 측정하는 과정이다. 측정결과를 통해 설정된 오차범위내의 프레임이 키 프레임으로 선정된다.



(그림 5) 키프레임 추출 결과

(그림 5)는 추출된 키 프레임을 기준으로 동영상에 대한 Shot 검출되어 동영상으로 생성되는 과정이다. 본 연구에서는 이와 같은 작업을 자동화함으로써 보다 효율적인 동영상의 Shot 검출 작업이 되도록 하였다.

4. 결론

본 연구에서는 동영상의 Shot 추출을 위해 픽셀값, 히스토그램, 경계면 픽셀수의 유사도 측정 방법을 제안하였다. 이를 통해 동영상의 연속적인 경계를 단순 색상 분포를 갖는 프레임과 복잡한 색상 분포를 갖는 프레임으로 분류하여 단순 색상 분포 프레임의 경우 각 프레임의 픽셀의 분산값을 참고하여 연속적인 경계를 검출하는 방법을 제안하였다. 이와 같은 방법은 추출된 프레임 중 대표 프레임을 선정하는 데 중요한 요소로 활용될 수 있다.

향후 연구 방향은 동영상의 전체 프레임에 대한 블록화를 수행하면 시간적 효율성이 떨어지는 현상이 있으므로. 효율적으로 블록의 크기, 위치 등을 설정하는 방안이 필요하다.

참고문헌

- [1] Pei-Yi Chen, Arbee L.P. Chen, "Video Retrieval Based on Video Motion Tracks of Moving Objects", Proceedings of SPIE Volume 5307, pp.550-558, 2003.
- [2] A.C.Aragon, O.Starostenko, "Image Retrieval by Ontological Description of Shapes (IRONS), Early Results,"Proceedings of the First Canadian Conference on Computer and Robot Vision, pp. 341-346, 2004.
- [3] L.Huayong, Z.Hui, "A Content-Based Broadcasted Sports Video Retrieval System Multiple Modalities: SportBR," International Conference on Computer and Information Technology, CIT, 2005.
- [4] Li, Z.Nian, Zhong, Xian and Drew, Mark S., "Spatial-temporal joint probability Images for video segmentation", Pattern Recognition, vol35, pp. 1847-1867, 2002.
- [5] M. Andrea Rodriguez, Max J. Egenhofer, Address D. Blaser, "Query Pre-processing of topological Constraints:Comparing a Composition-based with Neighborhood-Based Approach", SSTD 2003, LNCS 2750, pp. 362-379, 2003.
- [6] Fatih Porikli, Huifang Sun, "Compressed Domain Video Object", IEEE Transactions on Circuits and Systems for Video Technology, may 2005.