

비디오 프레임의 최소 화소를 이용한 장면 전환 검출 기술

김원희*, 문광석*, 김종남*

*부경대학교 전자컴퓨터정보통신공학부
e-mail:whkim07@pknu.ac.kr

Shot Change Detection Algorithm Using Minimum Pixels of Video Frame

Won-Hee Kim*, Kwang-Seok Moon*, Jong-Nam Kim*

*Dept. of Computer Multimedia Engineering, PuKyong Nat'l Univ.

요 약

장면 전환 검출은 대용량 비디오 데이터의 효과적인 관리를 위해서 사용되는 기술로서 현재까지 비디오 프레임의 크기를 대폭 축소시킨 환경에서의 연구는 미미하다. 따라서 본 논문에서는 비디오 프레임의 최소 화소를 이용한 장면 전환 검출 기술을 제안한다. 장면 전환 검출을 위한 특징값 추출 요소로 가중치 분산을 사용하였고, 가변 구간 참조를 통한 적응적인 임계값을 설정하였다. 실험을 통해서 기존의 방법들보다 precision에서 2~20.4%, recall에서 3~18.2%, F1에서 1.1~19.3% 향상된 것을 확인하였고, 비디오 데이터 화소수를 1/256로 축소하여 실험한 결과 기존의 방법들보다 검출률이 월등하게 향상된 것을 알 수 있었다. 제안하는 방법은 계산량 감소를 통한 고속 처리를 가능하게 하여 다양한 소프트웨어 및 하드웨어 플랫폼에서의 고속 장면 전환 검출에 유용하게 사용될 수 있다.

1. 서론

장면 전환 검출 기술은 비디오 데이터의 장면 변화가 발생하는 경계를 검출하는 기술이다. 여기에서 장면의 정의는 카메라를 통한 녹화에서의 물리적인 시작과 정지로 정의되며, 하나의 의미적인 단위로 구분되어지기도 한다. 장면 변화에는 연속되는 장면이 급격하게 변화하는 급진적인 변화와 서서히 변화하는 점진적인 변화로 구분된다. 급진적인 변화는 두 장면 사이의 단순한 연결로 구성되고, 점진적 변화는 두 장면 사이의 일정한 프레임들이 영상처리 기술에 의해서 중복되어 있다. 장면 전환 검출 기술은 효과적인 비디오 데이터를 관리하기 위한 기반 기술로 사용되는데 비디오 데이터의 분할, 색인, 검색, 요약과 같은 작업을 수행하기 위해서 필수적으로 사용된다[1].

장면 전환 검출 기술에 관한 기존의 연구들에서는 일반적으로 사용되는 SD급 동영상 또는 CIF급 영상을 실험 데이터로 사용한다. 이런 동영상은 영상의 크기에 따른 충분한 화소가 보장되므로 비교적 정확한 장면 전환 검출이 가능하다. 하지만 비디오 데이터의 크기가 HD급 이상의 해상도로 제작되면서 이러 데이터들을 그대로 처리하는 것은 장면 전환 검출 기술의 계산량에서 문제점으로 작용한다. 즉 충분한 표본들로부터 정확한 계산이 가능할 수 있지만 계산복잡도가 상당히 증가하게 된다. 따라서 계산복잡도를 줄이기 위해서 비디오 데이터를 축소시키거나 일부 표본들만 이용한다.

본 논문에서는 일반적인 크기의 비디오 데이터에서도 정확한 장면 전환 검출이 가능하며, 최소 화소만 추출하여

생성된 비디오 데이터에 대한 장면 전환 검출에서도 정확도의 손실이 거의 없는 장면 전환 검출 기술 개발을 목표로 한다. 이를 위해서 제안하는 방법은 장면 간 차이를 적절하게 표현할 수 있는 특징값으로 차분 프레임의 가중치 분산을 사용하며, 적응적인 임계값 설정을 위해서 가변 참조 구간에서의 평균 특징값을 이용한다. 제안한 방법은 동일한 영상에 대한 실험을 통해서 기존의 방법들보다 검출률이 향상된 것을 확인하였고, 1/256로 축소된 비디오 데이터를 이용한 실험에서도 기존의 방법들보다 검출 정확도가 월등하게 향상된 것을 알 수 있었다.

본 논문의 구성은 2장에서는 기존 연구에 대해서 기술하였고, 3장에서는 제안하는 방법에 대하여 설명하였다. 3장에서는 실험결과 및 분석을 기술하였고, 마지막으로 5장에서는 결론을 맺는다.

2. 관련 연구

장면 전환 검출 알고리즘은 공간영역과 변환영역에서 다양한 연구가 진행되어왔다. 공간영역에서는 화소 기반의 방법, 블록 기반의 방법, 히스토그램 기반의 방법 등이 있으며, 변환영역에서는 DC계수 기반의 방법과 매크로블록 기반의 방법들이 있다.

먼저 화소 기반의 방법은 연속되는 프레임에서 대응되는 화소의 밝기 차이를 이용해서 장면 전환 지점을 검출하는 방법이다. 이 방법은 작은 영역에서의 큰 변화나 큰 영역에서의 작은 변화를 구별할 수 없는 단점이 있다. 카메라와 객체의 동작, 명암등에 매우 민감하게 반응하여 오

검출 발생 확률이 높다[2].

블록 기반의 방법은 카메라와 객체의 동작에 민감하지 않도록 지역적인 특성을 강조한 방법이다. Zhang 등이 제안한 유사율 비교법은 프레임 간에 대응되는 블록의 유사율을 계산하여 임계값과의 비교를 통해서 장면 전환 유무를 결정하는 방법이다. Shahrray는 프레임을 12개의 겹치지 않는 블록으로 나눈 후에 명암 값을 이용하여 이전의 영상과 가장 잘 일치하는 블록을 찾는 방법을 제안하였다. 또 Xiong 등은 영상의 일부 영역들만 계산하여 장면 경계를 검출하는 망 비교 방법을 제안하였고, 시간과 공간정보를 활용한 단계변수 알고리즘을 소개하였다. 하지만 이런 방법들은 객체와 카메라의 움직임에 민감하게 반응하는 단점이 있다[3].

히스토그램 비교법은 현재까지도 장면 전환 검출에서 가장 일반적으로 사용되는 방법이며 보편적으로 가장 좋은 성능을 나타낸다. Tonomura는 연속되는 두 프레임의 밝기 히스토그램의 차이의 합을 특정 임계값과 비교하여 장면 전환을 판단하는 방법을 소개하였다. Nagasaka 등은 두 프레임 사이의 차이 값을 강조할 뿐만 아니라 카메라나 객체의 움직임을 강조할 수 있는 χ^2 -test 방법을 제안하였다. 이 방법은 기존의 히스토그램 방법에 비하여 정확성이 낮고 계산량이 많은 단점이 있다. Ueda 등은 컬러 히스토그램의 변화 비율을 이용한 장면 전환 검출 기술을 제안하였다. Gragi 등은 6개의 다른 컬러 좌표 시스템을 이용한 3개의 히스토그램 기반 방법의 성능평가를 하였다. 이외에도 지역적 χ^2 -test와 정규화를 이용한 장면 전환 검출 기술도 연구되었고, 누적 히스토그램 비교법, 블록 히스토그램 비교법, 히스토그램 교차법 등의 다양한 방법들이 연구되었다[4].

변환 영역에서도 장면 전환을 판단하는 다양한 기법들이 연구되었다. Meng 등은 DC 계수와 모션 벡터의 분산을 이용해서 장면 변화를 분석하였다. Zhang 등은 P픽처와 B픽처의 유효 모션 벡터의 수를 이용한 장면 전환 검출 기술을 제안하고, Yeo 등은 DC 영상의 시퀀스를 이용해서 장면 전환을 분석하였다. 이런 접근 방법들은 연속된 프레임들의 DC 영상의 히스토그램의 차이 또는 화소 차이에 기반한 방법들이다. Fernando 등은 B 픽처의 매크로블럭 예측의 통계적 특성을 이용해서 장면 전환을 검출하는 알고리즘을 제안하였고, Seong 등은 Fernando의 방법을 디지털 위성 수신기가 삽입된 하드디스크에 적용하였다. 이런 방법들은 I픽처와 P픽처로만 구성된 고품질 비디오 스트림에는 적용할 수 없는 단점이 있다[5].

이 외에도 Sethi 등은 연속되는 I 픽처의 DC 계수의 밝기 히스토그램의 차이를 사용했다. 이 방법은 고속의 장면 검출이 가능하지만 정확도가 저하되는 단점이 있다. Kim 등은 MPEG 압축 영역에서 인트라 픽처의 축소 영상을 이용한 장면 전환 검출 알고리즘을 PVR에 적용하였다[6].

3. 가중치 분산을 이용한 장면 전환 검출 방법

장면 전환 검출을 위해서는 비디오 데이터의 각 프레임의 특성을 나타내는 특징값과 이 특징값을 통해서 장면 전환 유무를 결정하는 임계값이 필요하다. 본 논문에서는 프레임의 특성을 나타내는 특징값으로 차분 프레임의 분산값을 이용한다. 분산값을 이용하는 것은 일반적으로 화소값을 이용하는 방법보다 검출 정확도가 높으며, 히스토그램 기반의 방법들보다 부표본화에 강인하고, 블록 기반의 방법들보다 계산량이 적기 때문이다.

본 논문에서는 분산값을 이용하면서 장면 전환점에서 특징을 더욱 명확하게 구분하기 위해서 가중치를 적용한다. 평균과 분산을 계산하는 과정에서 가중치를 적용하는 것은 장면 전환이 없는 프레임들과 장면 전환이 발생하는 프레임에서의 차이를 증가시키는 것이 목적이다. 식 (1)과 식 (2)에서는 일반적인 평균과 분산을 구하는 계산식을 나타내었다.

$$m = \frac{\sum_{x,y} (f_i(x,y) - f_{i+1}(x,y))}{W \times H} \quad (1)$$

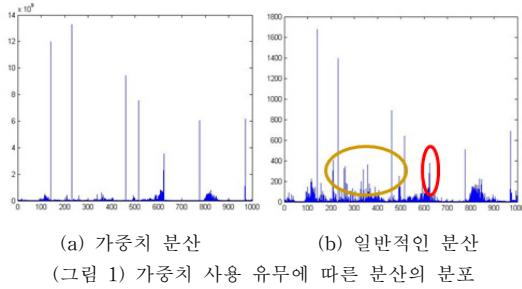
$$v = \frac{\sum_{x,y} (d_i(x,y) - m)^2}{W \times H} \quad (2)$$

여기에서 m 은 차분 프레임의 평균, v 는 차분 프레임의 분산, f_i 는 현재 프레임, f_{i+1} 는 다음 프레임을 각각 나타낸다. d_i 는 차분 프레임, W 는 프레임의 가로 크기, H 는 프레임의 세로 크기를 나타낸다. 제안하는 가중치가 적용된 평균과 분산은 식 (3)과 식 (4)에서 나타내었다.

$$m' = \frac{\sum_{x,y} (f_i(x,y) - f_{i+1}(x,y))}{(W/w_d) \times (H/w_d)} \quad (3)$$

$$v' = \frac{\sum_{x,y} (d_i(x,y) - m')^2}{(W/w_d) \times (H/w_d)} \quad (4)$$

여기에서 m' 은 가중치를 적용한 평균, v' 은 가중치를 적용한 분산, w_d 는 분산 가중치를 각각 나타낸다. 분모에 가중치를 나누어줌으로써 가중치가 적용된 평균 m' 은 $w_d^2 \times m$ 이 된다. 이 m' 을 이용한 분산 v' 은 v 보다 큰 값으로 계산된다. 여기에서 중요한 것은 v' 의 값이 v 의 값에서 선형적으로 증가하는 것이 아니라 장면 전환 지점에서만 비약적으로 높은 값을 나타낸다. 이런 값의 분포로 인해서 장면 전환 지점을 쉽게 구분할 수 있다. (그림 1)을 통해서 동일한 비디오 시퀀스에 대한 가중치가 적용된 분산과 일반적인 분산에 대한 분포를 비교하였다. (그림 1.a)에서 알 수 있듯이 제안한 가중치 분산의 분포는 장면 전환 지점을 쉽게 인지할 수 있다. 반면 (그림 1.b)와 같이 일반적인 분산에서는 장면 전환에 대한 판단이 모호한 지점들이 발생하는 것을 알 수 있다.



다음으로 고려되어야 하는 요소는 임계값이다. 임계값은 장면 전환 검출에서 특징값으로부터 장면 전환을 판단하는 역할을 한다. 따라서 적합한 임계값 설정이 정확한 장면 전환 검출 알고리즘을 결정하게 된다. 기존의 고정임계값을 사용하는 방법이나 비디오 시퀀스 전체를 분석하는 방법은 영상의 변화에 강인하게 적용하지 못하며 계산 시간이 많이 걸리는 단점이 있다. 따라서 적응적인 임계값 설정이 요구된다.

본 논문에서는 실시간 처리가 가능하며 계산복잡도가 낮으며 영상의 변화에 강인한 적응적 임계값 설정 방법을 개발한다. 위의 3가지 조건을 만족하는 임계값 설정을 위해서 입력 영상 자체의 정보를 이용하고, 현재 프레임 이전의 일정 구간의 정보만 이용하며, 이미 계산되어진 값을 사용하는 적응적 임계값 설정 방법을 제안한다. 제안하는 임계값 설정 방법을 식 (5)를 통해서 나타내었다.

$$TH = \frac{\sum_{i=S_c+1}^{S_c-1} v'(i)}{S_c - S_p + 1} \times w_i \quad (5)$$

여기에서 S_c 는 현재 프레임 번호, S_p 는 가장 최근 발생한 장면 전환 프레임 번호, w_i 는 임계값 가중치를 각각 나타낸다. 즉 제안하는 적응적 임계값 TH 는 가장 최근 발생한 장면 전환 프레임의 다음부터 현재 프레임 이전까지 차분 프레임의 가중치 분산값의 평균에 임계값 가중치 w_i 를 곱한 값으로 정의된다. 현재 프레임에서의 가중치 분산값 $v'(S_c)$ 가 적응적 임계값 TH 보다 큰 경우에 장면 전환 프레임으로 판단한다. 여기서 임계값 가중치 w_i 를 사용하는 것은 가중치가 없는 TH 값은 일정 구간에서의 분산값들의 평균값이 되므로 현재 프레임에서의 분산값과 유사한 값이 되는 경우가 있다. 이런 경우에 실제 장면 전환이 아니지만 장면 전환으로 판단되는 오검출이 발생하게 된다. 따라서 이런 오검출을 제거하여 임계값의 신뢰도를 높이기 위해서 임계값 가중치가 사용된다. 실험을 통해서 w_i 의 값은 10~15가 적절한 것을 알 수 있었고, 이 구간 내에서의 값의 변화는 검출 결과에 크게 영향을 주지 않는다. 제안한 임계값 설정 방법은 추가적인 사용자의 개입없이 동작하므로 실제 응용 환경에서의 사용이 가능하다.

4. 실험 결과 및 분석

제안한 방법의 성능 평가를 위해서 다음 조건에서 실험하였다. PC의 환경은 CPU Core Duo 2.4GHz, RAM 2G이다. 장면 전환 검출 프로그램은 VC++ .net 7.1을 이용해서 구현하였고, 검출 결과와 육안으로 구분한 장면 전환 결과와의 비교는 MATLAB 7.3을 통해서 이루어졌다. 실험에 사용한 비디오 데이터는 4가지 장르의 10개 영상을 이용했다. 전체 영상에서 5분 동안의 시퀀스를 추출하여 176×144 크기의 YUV 4:2:0 포맷의 9000프레임 비디오 데이터를 생성하였다. <표 1>에서는 실험에 사용된 영상의 종류와 영상별 장면 전환 프레임의 개수를 나타내었고, <표 2>에서는 장면 전환의 구성과 비율을 나타내었다. 실험에서 사용된 가중치 요소로써 w_d 는 8, w_s 는 11을 사용하였다.

<표 1> 실험에 사용된 영상의 종류와 장면전환개수

구분	제 목(장면전환개수)
드라마	커피프린스(93), 프리즌브레이크(88) 천의전쟁(52)
영화	프리덤라이터스(111), 올드미스다이어리(52) 복면달호(32)
뉴스	KBS 뉴스(69), SBS 뉴스추적(63)
쇼오락	무한도전(98), 개그콘서트(78)

<표 2> 장면 전환 프레임의 유형과 분포

구분	장면전환프레임	급진적변화	점진적변화
개수(개)	736	723	13
비율(%)	100	98.2	1.8

제안한 방법의 성능 평가 척도로는 일반적으로 사용되는 precision, recall, F1을 사용하였다[7]. 각 척도의 정의는 다음의 식 (6) ~ 식 (8)에 나타내었다. 여기에서 N_c 는 정확하게 검출한 장면 전환 개수, N_f 는 잘못 검출한 장면 전환 개수, N_m 은 검출하지 못한 장면 전환 개수를 각각 나타낸다.

$$Precision = \frac{N_c}{N_c + N_f} \quad (6)$$

$$Recall = \frac{N_c}{N_c + N_m} \quad (7)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

첫 번째 실험에서는 176×144 크기의 비디오 데이터를 대상으로 기존의 방법들과 성능을 비교하였다. <표 3>에서 알 수 있듯이 제안한 방법이 precision에서 2~20.4%, recall에서 3~18.2%, F1에서 1.1~19.3% 향상된 것을 알 수 있다. 기존의 방법들 중에서 제안한 방법과 비슷한 성

능을 보이는 방법들도 있지만 기존의 방법들은 고정 임계값을 사용하기 때문에 실제 응용환경에 적용할 수 없는 단점이 있다. 기존의 방법들은 최고의 검출 결과가 나올 때까지 임계값을 수정하면서 반복실험한 결과이다.

<표 3> 기존 방법과의 검출 정확도 비교

구 분	Precision	Recall	F1
C0 화소비교[2]	84.5	84.5	84.5
C1 유사율비교[3]	95.9	95.8	95.9
C2 MSE비교[2]	90.0	90.2	90.1
C3 히스토그램비교[4]	93.5	93.2	93.3
C4 컬러히스토그램비교[4]	96.5	96.6	96.5
C5 χ^2 -test비교[4]	95.5	95.7	95.6
C6 분산값비교법[3]	88.4	88.0	88.2
C7 차분히스토그램분산값비교[2]	83.4	83.4	83.4
C8 차분프레임분산값비교[3]	87.0	87.0	87.0
C9 DC계수비교[6]	75.1	75.0	75.1
P 제안한방법	95.5	93.2	94.4

두 번째 실험에서는 11×9로 부표본화한 비디오에서 검출 정확도를 비교하였다. <표 4>에서와 같이 제안한 방법의 검출 정확도가 비교 방법들보다 월등히 높은 것을 알 수 있다. 기존의 방법들은 부표본화에 따른 15~50% 내외의 검출 정확도의 손실이 발생하였다. 따라서 제안한 방법은 최소의 화소 표본을 이용한 환경에서도 정확한 검출이 가능한 것을 알 수 있다.

<표 4> 11×9로 부표본화한 비디오에서 검출정확도 비교

구 분	Precision	Recall	F1
C0 화소비교[2]	76.0	75.4	75.7
C1 유사율비교[3]	-	-	-
C2 MSE비교[2]	79.1	79.2	79.2
C3 히스토그램비교[4]	35.1	34.8	34.9
C4 컬러히스토그램비교[4]	39.6	36.8	38.1
C5 χ^2 -test비교[4]	37.4	36.8	37.1
C6 분산값비교법[3]	72.0	72.3	72.1
C7 차분히스토그램분산값비교[2]	43.0	43.2	43.1
C8 차분프레임분산값비교[3]	72.7	72.4	72.6
C9 DC계수비교[6]	-	-	-
P 제안한방법	92.9	91.8	92.3

두 가지 실험을 토대로 제안한 방법이 비디오 시퀀스의 크기에 상관없이 좋은 검출 결과를 보이는 것을 알 수 있었다. 대부분의 비교 방법들이 부표본화한 비디오에서 손실이 큰 것에 비해서 제안한 방법은 손실이 2% 내외로 적었다. 실험을 통해서 제안한 방법이 기존 방법들에 비해서 훨씬 낮은 계산량으로 정확한 장면 전환 검출이 가능한 것을 알 수 있었다. 또한 임계값에 대한 설정이 적응적이기 때문에 실제 응용 환경에서도 사용 가능하다.

5. 결론

본 논문에서는 일반적인 크기의 비디오 데이터에서도 정확한 장면 전환 검출이 가능하며, 최소 화소만 추출하여 생성된 비디오 데이터에 대한 장면 전환 검출에서도 정확도의 손실이 거의 없는 장면 전환 검출 기술을 제안하였다. 제안한 방법은 가중치 분산을 이용해서 장면 전환이 발생하는 프레임에서의 특징값을 정확하게 추출하였고, 가변 참조 구간의 평균 특징값을 이용해서 영상에 적응적인 임계값을 설정하였다. 제안한 방법은 동일한 영상에 대한 실험을 통해서 기존의 방법들보다 precision에서 2~20.4%, recall에서 3~19.2%, F1에서 1.1~19.3% 향상된 것을 확인하였고, 비디오 데이터를 1/256로 축소하여 실험한 결과도 기존의 방법들보다 검출 정확도가 월등하게 높았다. 제안한 방법은 계산량 감소를 통한 고속 처리가 가능하므로 다양한 소프트웨어 및 하드웨어 플랫폼에서의 고속 장면 전환 검출에 유용하게 사용될 수 있다.

감사의 글

본 논문은 중소기업청의 산학연공동기술개발지원사업(선도형), 한국산업기술재단의 지역혁신인력양성사업의 지원으로 수행되었음.

참고문헌

- [1] C. Cotsaces, N. Nikolaidis, and I. Pitas, "Video shot detection and condensed representation," *IEEE Signal Processing Magazine*, Vol. 23, pp. 28-37, 2006.
- [2] S. W. Smoliar and H. J. Zhang, "Content-based video indexing and retrieval," *IEEE Multimedia*, Vol. 1, No. 2, pp. 62-72, 2006.
- [3] J. Bescos, G. Cisneros, J. M. Menendez, and J. Cabrera, "A unified model for techniques on video-shot transition detection," *IEEE transaction on Multimedia*, Vol. 7, pp. 293-307, 2005.
- [4] S. Y. Shin, G. R. Sheng, and K. H. Park, "A scene change detection scheme using local χ^2 -test on telematics," *International Conference on Hybrid Information Technology*, Vol. 1, pp. 588-592, 2006.
- [5] W. A. C. Fernando, C. N. Canagarajah, and D. R. Bull, "Scene change detection algorithms for content-based video indexing and retrieval," *Electronics and Communication Journal*, Vol. 13, No. 3, pp. 117-126, 2001.
- [6] J. R. Kim, S. J. Suh, and S. H. Sull, "Fast scene change detection for personal video recorder," *IEEE Transaction on Consumer Electronics*, Vol. 49, pp. 683-688, 2003.
- [7] G. Boccignone, A. Chinaese, V. Moscato, and A. Picariello, "Foveated shot detection for video segmentation," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, pp. 365-377, 2005.