

SURF 기반의 장면 전환 검출 방법

오현주^{1*}, 박지용¹, 홍석민¹, 강현민¹
¹구미전자정보기술원 실감미디어연구센터
 *hjoh@geri.re.kr

Scene Change Detection Based on SURF

Hyunju Oh¹, Jiyong Park¹, Seokmin Hong¹, Hyunmin Kang¹
¹Realistic Media Research Center,
 Gumi Electronics & Information Technology Research Institute

요 약

장면 전환 검출을 위해 히스토그램 등 단일 특성만 고려할 경우 움직임이 많은 영상과 조명의 변화가 다양하거나 단조로운 색상으로의 장면 전환이 일어나는 영상에 검출이 어렵다는 단점이 있다. 이를 개선하기 위해 프레임 간 색상 히스토그램을 비교 후, SURF를 적용하는 방법을 제안한다.

1. 서론

최근 영상 정보의 색인 및 검색, 탐색 및 데이터베이스 구축 등 다양한 멀티미디어 시스템이 등장하고 있다. 일반적으로 영상 데이터의 분할 기준이 되는 것은 프레임, 장면 등이다. 프레임은 영상 구조의 가장 하위 단위로 장면을 이루는 기본 단위이며, 장면은 같은 시간, 같은 장소에서 같은 내용을 내포하는 연속된 프레임의 집합이다[1]. 장면 전환 검출 기술은 크게 압축된 영역에서의 장면 전환 검출과 비압축된 영역에서의 장면 전환 검출 방법으로 구분할 수 있다.

2. 기존의 장면 전환 검출 방법

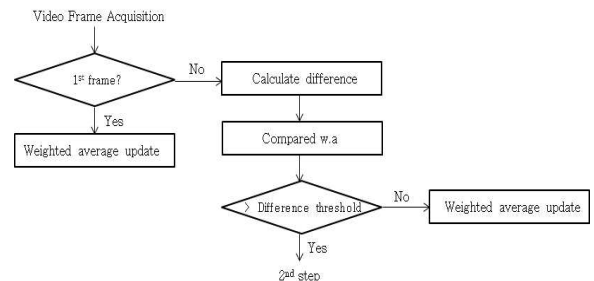
압축된 영역에서의 장면 전환 검출 방법은 8*8 DCT(Discrete Cosine Transform) 블록의 DC 성분들을 추출해 만든 DC 이미지를 사용한다. DCT는 입력된 아날로그 원 영상을 DCT라는 과정을 거쳐 저주파와 고주파의 주파수 성분으로 분해하는 변환 기술이다. 화소의 관계를 주파수의 관계로 변환하고, 이를 통해 원 영상을 압축하기 쉽게 하는 것이다. 인간의 시각 특성은 저주파 성분에 민감해 정보의 제거를 잘 인지하고, 고주파 성분에 둔감하여 정보의 제거를 잘 인지하지 못한다. DCT는 이러한 점을 사용해 압축을 용이하게 하지만, DC 영상을 만들기 위해 완전 압축 과정이 필요하다는 단점이 있다[2].

비 압축 영역에서는 디코딩해서 얻어진 이미지 정

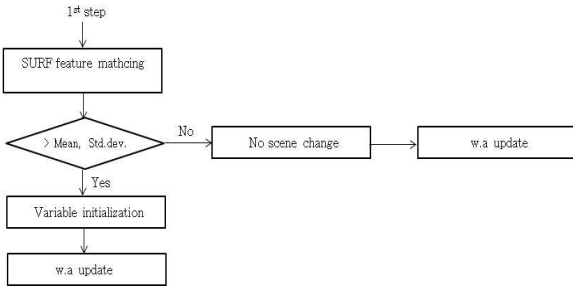
보인 히스토그램을 이용하거나, 화소 비교를 이용하는 방법으로 검출한다. 히스토그램을 이용한 방법은 먼저 프레임에서의 다수 픽셀에 의해 얻어진 히스토그램을 이용한다. 모든 프레임을 주변 프레임의 히스토그램 분배와 비교한다. 히스토그램을 이용한 방법은 각 프레임의 밝기나 색상 분배만 이용하기 때문에 공간 정보가 없다는 것이 단점이다[3].

3. SURF 기반의 장면 전환 검출 방법

장면 전환 검출은 일반적으로 프레임 간 밝기 차이나 색상 정보를 이용해 일정 크기 이상의 차이가 존재할 경우 장면 전환으로 판단하는 히스토그램 기법들이 있다. 하지만 단일 특성만 고려할 경우 속도는 개선되나, 움직임이 많은 영상, 조명 변화가 다양한 영상 등에서 검출이 어렵다는 단점이 있다. 본 논문에서 제안한 방법은 색상 히스토그램 분석과 특징점 추출 알고리즘인 SURF를 적용하는 두 단계를 통해 검출 정확도를 제고하고자 하였다.



(a) 1단계 : 색상 히스토그램 분포 비교를 통한 장면 전환 후보 결정 과정



(b) 2단계 : SURF를 이용한 장면 전환 판단 과정
(그림 1) 제안한 방법의 흐름도

3.1. 색상 히스토그램 차 분석

첫 번째 단계는 색상 히스토그램 분석 단계로, 제안하는 방법은 현재 및 이전 프레임 간의 색상 히스토그램 분포 차를 이용한다. 식 (3.1)에 따라 유사도를 측정한다. $H_i(n)$ 과 $H_j(n)$ 은 각각 이전 프레임의 히스토그램과 현재 프레임의 히스토그램을 나타낸다. 이후 식 (3.2)에 해당하는 임계치 Difference threshold인 *Weighted average*와 비교한다. α 값은 실험 결과에 따라 0.6으로 사용하였다.

$$w(i, j) = \sum_{n=0}^{N-1} |H_i(n) - H_j(n)| \quad (3.1)$$

$$W_{avg} = W_{avg} \cdot \alpha + cur \cdot (1 - \alpha) \quad (3.2)$$

*Weighted average*를 사용하는 이유는 먼저 변화량이 큰 요소에 영향을 덜 받기 위해 정규화하고, 두 번째로는 영상의 특성을 고려하지 않고 고정된 임계치를 사용할 경우 임계치의 크기에 따라 너무 많거나 적게 장면 전환이 일어나기 때문이다. 따라서 적응적으로 임계치를 두어 정확한 장면 전환 검출이 가능하도록 하였다. 속도 개선을 위해 임계치 이상일 경우에만 두 번째 단계인 SURF 기반의 장면 전환 검출 과정을 진행한다.

3.2. SURF 기반의 장면 전환 검출 방법

색상 히스토그램 분포 차를 이용한 프레임 간 유사도 판별 후 장면 전환 후보들이 결정되고, 이 후보들에 한하여 SURF 기반의 장면 전환 검출 단계가 진행된다. 방법은 SURF는 feature 공간에서 가장 짧은 거리를 갖는 특징점 간 매칭을 수행하는데 이 거리는 식 (3.3)로 표현되는 유클리드 거리에 해당한다. 이 거리를 평균과 표준 편차 임계치와 비교하는데, 임계치 이상인 경우 장면 전환이 일어났다고 판단한다. 이 과정이 끝나면 다음 장면 전환 검출을 위해 각종 수치를 초기화하고 3.1절, 3.2절 과정을 반복한다.

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \\ = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (3.3)$$

4. 실험결과 및 분석

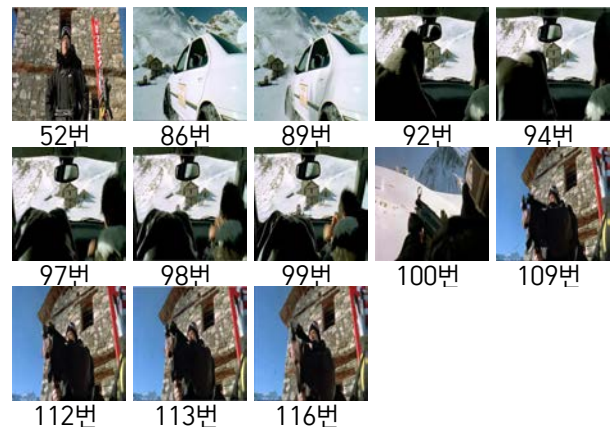
본 논문은 정확한 장면 전환 검출을 위해 색상 히스토그램과 특징점 추출 알고리즘을 이용한 방법을 제안하였으며, 단일 특성만 고려하는 경우와 제안하는 방법 간의 실제 성능을 실험하였다. 개발 언어로는 MATLAB을 사용하였으며, 실험은 CIF(352×288) 크기의 동영상 압축 표준 실험 영상 중 사람, 사물, 풍경 등의 정보를 내포한 7개의 영상을 사용하였다.

<표 1> 실험에 사용한 영상 정보

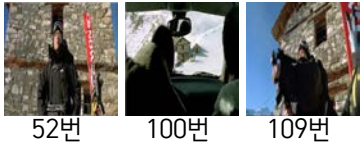
| 분류 | 이름 | 재생 시간(초) |
|-------------|-------------------------|----------|
| Documentary | Chionin | 70 |
| | Journey of Amazing Cave | 150 |
| Movie | Taxi 3 | 200 |
| | Terminator 2 | 180 |
| Nature | Amazon | 240 |
| | Coral | 240 |
| Sports | Step Into Liquid | 120 |

4.1. 실험 결과 분석

기존 방법과 제안 방법의 비교를 위해 Taxi 3 영상의 일부분인 120 프레임을 비교해 보았다. 총 5번의 장면 전환이 일어나는 이 영상의 경우, 제안한 방법은 정확하게 검출하는 것을 알 수 있었다. 하지만 기존 방법 중 색상 히스토그램 단일 특성만 고려하는 경우 전환이 일어나지 않은 영상임에도 불구하고 장면 전환으로 검출되는 문제가 있었다. 에지 탐지를 이용한 방법의 경우에도 실제로 장면 전환이 발생한 경우를 찾지 못하는 등 기존 방법들의 한계가 있음을 확인할 수 있었다.



(a) 색상 히스토그램을 이용한 장면 전환 검출



(b) 에지를 이용한 장면 전환 검출



(c) 제안한 방법을 이용한 장면 전환 검출

(그림 2) 기존 방법과 제안 방법의 장면 전환 검출

4.1.1. 프레임 간 색상 히스토그램 분포 차 추출

CIF 크기의 압축 표준 영상들을 이용해 프레임 간 색상 히스토그램 분포 차를 계산하였고, 영상의 밝기, 색상 분포에 따라 히스토그램 분포 차가 달라지는 것을 확인할 수 있었다.

4.1.2. SURF 특징점 매칭 과정

두 번째 과정은 색상 히스토그램 분포 차를 이용해 장면 전환 후보들을 대상으로 수행되는 과정이다. 성능평가 척도로는 Recall, Precision, F-Measure를 사용하였고 정의는 식 (4.1)과 같다.

$$\begin{aligned}
 Recall &= \frac{N_c}{N_c + N_m} \\
 Precision &= \frac{N_c}{N_c + N_f} \\
 F-Measure &= \frac{2 \times Recall \times Precision}{Recall + Precision}
 \end{aligned}
 \tag{4.1}$$

여기서 N_c , N_m 그리고 N_f 는 각각 검출된 장면 전환 수, 검출하지 못한 장면 전환 수, 잘못 검출된 장면 전환 수를 의미한다. 실험 데이터를 통해 Recall, Precision을 구하고 이를 이용해 조화 평균인 F-Measure를 구한다. 식 (4.1)에 따라 N_m 와 N_f 이 작을수록 좋은 성능을 나타낸다고 할 수 있다.

정확도에 대한 손득 수치 비교는 표 2로 나타내었다. 기존 방법들과 비교했을 때 점진적 장면 전환의 경우 외에 일반적으로 제안 방법의 평균이 향상됨을 확인하였다. 색상 히스토그램 방법에 비해 Recall은 11.61%, Precision은 75.17%, F-Measure는 44.97%, 에지 검출 방법과 비교했을 때에는 각각 19.78%, 17.08%, 19.21% 향상되었다.

<표 2> 정확도 손득(기존 방법과의 비교)

| 분류 | 이름 | 색상 히스토그램과의 비교 | | | 에지 검출과의 비교 | | |
|-------------|------------|---------------|---------------|---------------|------------|---------------|---------------|
| | | Recall (%) | Precision (%) | F-Measure (%) | Recall (%) | Precision (%) | F-Measure (%) |
| Documentary | Chionin | 0 | 0 | 0 | 0 | 0 | 0 |
| | Journey | 0 | 15.38 | 7.69 | 0 | 11.05 | 5.53 |
| Movie | Taxi | 1.29 | 246.92 | 130.22 | 1.33 | 1.06 | 1.18 |
| | Terminator | 34.11 | 40.4 | 37.24 | 20.7 | 27.8 | 24.22 |
| Nature | Amazon | 11.8 | 9.81 | 10.82 | 5.47 | 4.46 | 4.97 |
| | Coral | -9.26 | 102.48 | 49.91 | -7.54 | 40.16 | 17.71 |
| Sports | Step | 31.69 | 36.04 | 33.91 | 118.48 | 45.05 | 80.86 |
| 평균 | | 11.61 | 75.17 | 44.97 | 19.78 | 17.08 | 19.21 |

5. 결론

본 논문은 정확한 장면 전환 검출을 위해 색상 히스토그램 분포 차이 분석과 SURF를 이용한 두 단계로 이루어진 방법을 제안하였다. 단일 특성을 사용하는 경우 야기되는 정확도 감소의 문제를 해결하기 위해 두 단계 과정을 통한 검출 방법을 제안하였으며, 영상의 특성을 고려한 적응적 임계치와 영상의 환경 요인에 강인한 특징점 검출을 적용하였다. 제안하는 방법의 성능 평가를 위해 Recall, Precision, F-Measure의 정확도 평가 척도를 이용하였으며, 단일 특성을 사용한 경우보다 일반적으로 정확도가 향상됨을 확인할 수 있었다.

사사

본 연구는 과학기술정보통신부 ‘홀로그램핵심기술개발(R&D) 사업’의 일환으로 수행하였음.[2022-0-00552, 홀로그램 기반 문화재 복원 및 가시화 서비스 사업화 실증]

참고문헌

[1] C. Cotsaces, N. Nikolaidis, and I. Pitas, “Video Shot Boundary Detection and Condensed Representation: A Review,” IEEE Signal Processing Magazine, Vol. 23, pp. 28–37, 2006.

[2] Sung Min Kim, Ju Wan Byun, Chee Sun Won, “A Scene Change Detection in H.264/AVC Compression Domain,” Lecture Notes in Computer Science, Vol. 3768, pp. 1072–1082, 2005.

[3] Min Woo Park, Young Jin Nam, Sung-Ryul Kim, Dae-Wha Seo, Soon Ki Jung, “Scene-Change Detection in MPEG2 using B Frame Size and GOP Length,” HCI, pp. 629–634, February 2008.