

통계적 특성을 이용한 비디오의 분할 및 대표 프레임의 추출방법

조완현*, 박순영**, 박종현**

전남대학교 통계학과, 목포대학교 전자공학과

whcho@chonnam.ac.kr, {sympark, jhpark}@chungkye.mokpo.ac.kr

Wan-Hyun Cho*, Soon-Young Park*, Jong-Hyun Park**

*Dept. of Statistics, Chonnam National University

**Dept. of Electronics Engineering, Mokpo National University

요약

본 논문에서는 각 프레임으로부터 추출된 통계적 특성을 이용하여 동영상의 분할방법과 분할된 각 장면에 대한 대표프레임을 추출하는 방법을 제안한다. 제안된 방법은 동영상의 각 프레임에 대하여 다해상도 분석을 실시하여 저주파 부 밴드로부터 히스토그램 특징을 추출하여 급격한 장면전환을 분할하는데 이용하였으며 또한 점진적인 장면전환을 검출하기 위해서는 고주파 부 밴드로부터 계산되는 화소의 분산치를 계산하여 특징벡터로 사용하였다. 실험의 결과를 통하여 제안된 동영상의 분할방법과 대표프레임 추출에 대한 알고리즘들이 동영상 자료의 분석 및 색인을 위한 효율적인 동영상 분할을 가능하게 하며, 차후 내용기반 영상과 비디오의 색인 및 검색을 위한 전처리의 단계로 사용되어질 수 있는 매우 유용한 방법임을 보였다.

1. 서론

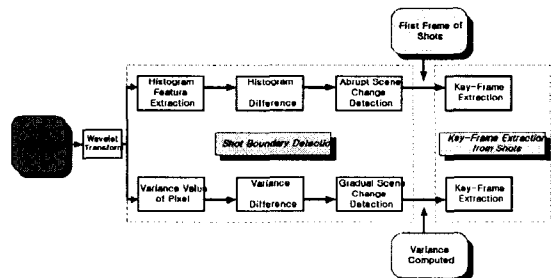
동영상에 대한 분할은 단일 프레임으로 구성된 정지영상에 대한 공간적인 분할과 연속 프레임으로 구성된 동영상에 대한 시간적인 분할이 있다. 먼저 공간적인 분할은 정지영상을 동질의 특성을 갖는 균일한 영역으로 구분해 내는 과정으로서 영상인식, 영상이해, 영상검색 등에서 필수적인 과정 중의 하나로 사용되어지고 있으며 또한 동영상에 대한 시간적 분할은 동일한 내용을 표현하는 하나의 장면인 샷(shot)과 두 개의 독립된 장면의 경계인 컷(cut) 그리고 점진적인 장면전환에 속하는 페이드(fade)와 디졸브(dissolve) 등으로 주어진 영상을 순서적으로 분할하는 것을 말한다.

본 논문에서는 동영상에 대한 시간적인 분할로써 통계적인 특징을 이용한 컷 경계 검출 및 대표프레임의 추출을 목적으로 한다. 먼저 급격한 장면전환인 컷 검출을 위해서는 주어진 각 영상에 대하여 다해상도 분석을 실시하고, 저주파 부 밴드로부터 히스토그램 특징을 추출하여 컷 검출에 사용하였으며 점진적인 장면전환인 페이드와 디졸브를 검출하기 위해서는 고주파 부 밴드에 위치한 웨이블릿 계수들은 수직, 수평, 대각 방향의 에지 정보로부터 계산되는 화소값의 분산의 특징벡터를 이용하여 효율적으로 검출하였다.

2. 통계적 특성을 이용한 동영상의 분할

2.1 제안된 알고리즘의 구조

본 논문에서 제안된 동영상의 장면전환에서 발생하는 샷의 경계검출과 각 샷에 대한 대표프레임의 추출에 대한 알고리즘의 구성도는 다음 그림과 같다. 먼저 동영상의 각 프레임에 대하여 웨이블릿 변환을 적용하여 다해상도분석을 실시하고, 저대역 부 밴드로부터 히스토그램 특징을 추출하고, 고대역 부 밴드로부터는 화소 값들의 분산특징을 추출한다. 다음으로 인접한 두 프레임들에 대하여 히스토그램 특징 값을 차와 분산특징 값들의 차를 계산하여 급격한 장면전환과 점진적인 장면전환을 탐지한다. 그리고 분할된 각 샷으로부터 대표프레임을 추출한다.



[그림1] 제안된 컷 경계검출 및 키 프레임 추출 알고리즘 구성도

2.2 동영상의 급격한 장면전환 검출

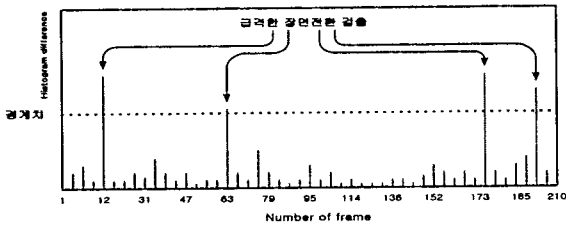
본 논문에서는 다양한 동영상 자료들을 대상으로 급격한 장면전환의 검출을 위하여 히스토그램의 차이에 의한 탐색방법을 사용하였다. 동영상의 자료를 다해상도 분석을 실시한 후 저 대역 부 밴드로부터 히스토그램의

차이에 대한 특징 값은 다음의 수식에 의하여 계산된다..

$$D_h = \sum_{i=1}^K |H_{m+1}(i) - H_m(i)|$$

여기서 H_m 은 m 번째 프레임의 히스토그램특징이고, H_{m+1} 은 $m+1$ 번째 프레임의 히스토그램특징이다.

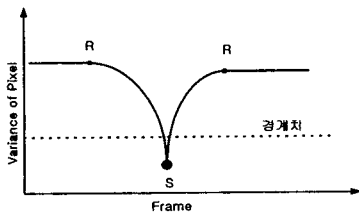
이 특징 값의 성격은 동일한 셋 내의 유사한 프레임들 사이에서는 히스토그램의 차이 값이 적게 나타나고, 이질적인 셋들의 경계에 해당하는 부분의 인접한 프레임들 사이에서는 히스토그램의 차이 값이 높게 나타나는 특징을 가지고 있다. 따라서 이점을 이용하여 연속적인 프레임들 사이에서 셋의 경계를 검출할 수 있다. 다음은 급진적인 장면전환에서 히스토그램특징 값들의 차이를 보여주는 그림이다.



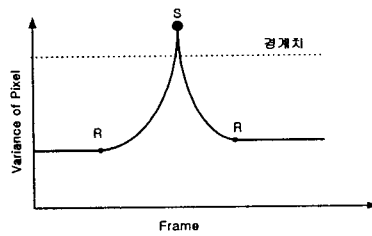
[그림2] 급진적인 장면 전환에서 셋 경계분할

2.3 동영상의 점진적인 장면전환 검출

점진적 장면전환 검출을 위한 특징 값으로 주어진 동영상 자료에 대하여 웨이블릿 변환을 실시하고, 고대역 부 밴드에 위치한 웨이블릿 계수들에 대하여 수직, 수평, 대각 방향의 화소값들에 대한 분산특징을 계산한다. 이때 분산특징 값들은 다음 [그림3]에서 표시한 것과 같이 페이드가 일어나는 경우에 프레임들의 수열이 페이드 아웃(fade out)되고 있을 경우, 화소값의 분산은 급속하게 감소하게 되며, 페이드(fade in) 안이 시작되면 그 값은 다시 증가하게 된다. 반대로 다음 [그림4]에서와 같이 디졸브가 일어나는 경우에는 두 장면이 겹치기 시작하는 부분에서부터 화소값의 분산은 점점 증가하게 되며, 완전히 겹치는 부분에서 최고점을 지나 다시 분산값이 감소하게 된다. 따라서 셋 경계 검출의 방법은 [그림3]와 [그림4]에서 표시된 점 S를 검출하는 것이 된다.

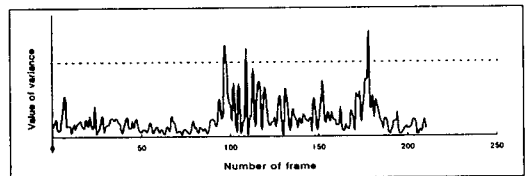


[그림3] 페이드에서 셋 경계의 검출



[그림4] 디졸브에서 셋 경계의 검출

먼저 화소값의 분산을 감시하다가, 페이드인 경우 가장 최소값을 가지는 점을 셋의 경계로 인식하고 디졸브의 경우 가장 최대값을 가지는 점을 셋의 경계로 인식하게 된다. 다음은 디졸브가 일어나는 경우에 분산 값의 변화를 도식한 그림이다.

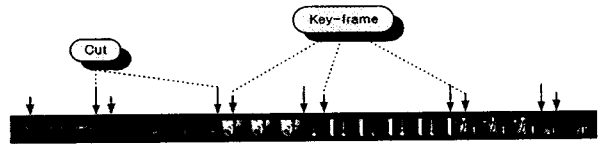


[그림5] 디졸브의 경우에 화소들에 대한 분산 값들의 변화

2.4 동영상의 대표프레임의 추출

2.4.1 급격한 장면전환에서 대표프레임 추출

급격한 장면 전환은 셋의 경계 부분에서 인접한 프레임들 사이에 히스토그램의 차이가 현저하게 드러나는 특징을 가지고 있다. 이는 컷의 특성상 페이드, 디졸브 등과 같이 셋 내의 인접한 프레임들 사이에 어떤 특수한 영상에 대한 처리가 없기 때문에 급격한 장면 전환시 셋에서의 키 프레임은 어느 것을 선택하나 큰 차이가 없게 된다. 따라서 셋 내의 첫 번째 프레임을 대표프레임으로 선정을 하였다.

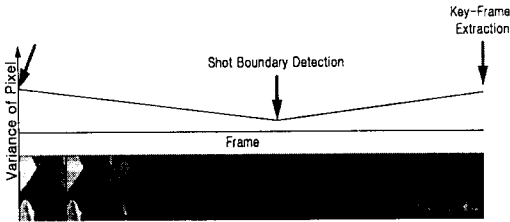


[그림6] 컷에서의 키 프레임 추출

2.4.2 점진적인 장면전환에서 대표프레임 추출

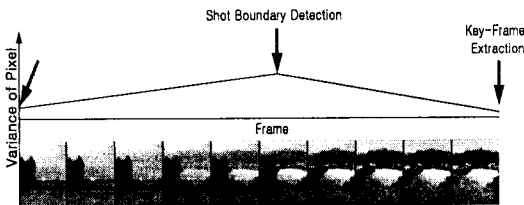
먼저 페이드가 일어나는 경우에는 [그림7]와 같이 페이딩 과정에서 분산값이 변화는 경우, 셋 경계 검출에서 분산값이 가장 적은 곳을 셋 경계로 선택하였다. 이와 반대로 대표프레임 추출은 페이딩 과정이 끝나는 지점에서 원 영상의 프레임이 나타나게 되며, 이때의 분산값이 점점 상승하여 최고점이 된다. 따라서 이때의 프레임을 대표프레임으로 선택하게 된다. 즉, 프레임에서 화소값 분산이 가장 큰 부분의 프레임을 대표프레임으로 선정하게 되며, 이 경우 자연스럽게 페이딩 중간 과정의 프레

임이 대표프레임으로 선정되는 것을 막을 수 있다.



[그림7] 페이딩 과정의 화소값 분산의 변화

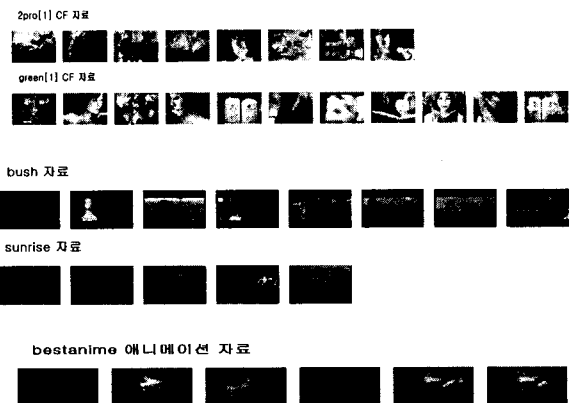
다음으로 디졸브가 일어나는 경우에는 [그림8]와 같이 디졸브 과정에서 분산값이 변화는 경우, 셋 경계 검출에서 분산값이 가장 높은 곳은 셋 경계로 선택한 것과 반대로 분산값이 떨어지기 시작하여 가장 작은 값을 가질 때 그 셋의 원 영상의 프레임이 나타나며, 이때의 프레임을 대표프레임으로 선택하게 한다. 즉, 프레임내의 화소값 분산이 가장 작은 부분의 프레임을 대표프레임으로 선정하게 되며, 이럴 경우 자연스럽게 디졸브 중간 과정인 두 장면이 겹치는 부분을 피해서 대표프레임이 선택되게 된다.



[그림8] 디졸브 과정의 화소값 분산의 변화

3.2 대표프레임 추출

대표프레임의 선정은 영상내용에 의한 검색의 정확성에 직접적으로 영향을 미칠 수 있는 중요한 문제이다. 따라서 본 논문의 컷, 페이딩, 디졸브의 경우 히스토그램 정보 및 셋 경계 검출에서 이미 추출된 화소 값 차이의 변화 패턴을 이용하여 대표프레임을 선정하는 방법을 소개하였다. 이 방식은 이미 계산되어 있는 특징량을 사용하므로 추가적인 처리부담이 거의 없이 기존의 단순한 방식에 비하여 훨씬 합리적인 프레임의 선정이 가능하다는 장점을 가지고 있다. 다음은 실험 자료에 대한 대표프레임 추출에 관한 실험 결과를 보여주고 있다.



[그림9] 대표프레임의 추출 결과

3. 실험 및 결과

3.1 동영상의 셋 경계 검출

실험데이터는 MPEG포맷의 디지털 동영상으로 500~600 프레임으로 구성되어 있으며, 이 데이터에는 주로 간략한 TV_CF, CNN, TV 드라마, 애니메이션, 자연영상 등을 대상으로 실험을 실시하였으며, 각 프레임의 크기는 352x240 화소이다. 사용된 동영상은 다양한 형태의 컷, 페이드, 디졸브를 포함하고 있으며 정확한 장면전환 검출을 위하여 주관적으로 판단 가능한 장면전환을 포함하는 동영상을 사용하였다. 다음의 [표1]은 본 논문에서 제안된 장면전환 검출방법에 대한 5가지 동영상에 대한 실험결과를 보여주고 있다.

[표1] 장면전환 검출 결과

	Frames	Detected	Missing Shots	False Alarms	Recall	Precision
2pro[1]	450	8	0	0	100%	100%
green[1]	602	11	0	1	100%	91%
bush	863	9	2	2	81%	81%
sunrise	450	5	1	1	83%	83%
bestanime	551	5	0	1	100%	83%

4. 결론 및 토의

본 논문에서는 통계적인 특징에 기반을 둔 동영상의 셋 경계 검출 및 대표 프레임 추출에 관하여 연구하였다. 컷인 경우에는 인접한 프레임간에 히스토그램의 차이를 이용하였으며, 페이드, 디졸브인 경우에는 화소에 의한 분산치를 이용하여 셋 경계를 분할하게 하였다. 셋 경계검출에 있어서 컷, 페이드, 디졸브등의 처리를 따로 하여 보다 더 정확한 셋 경계를 검출할 수 있었다.

참고문헌

- [1] 박중현, 박순영, 조완현, "색상특징과 웨이블릿 기반 특징을 이용한 영상검색", 대한전자공학회 추계 종합 학술대회 논문집, pp 487-498, Nov. 1999.
- [2] 박중현, 박순영, 조완현, "계층적 Hidden Markov Model을 이용한 비디오 장면전환 검출", 제 13회 영상처리 및 이해에 관한 워크샵, pp. 196-201, Jan. 2001.
- [3] Li, Ze-Nian, Zhong, Xiang and Drew, Mark S., "Spatial-temporal joint probability images for video segmentation", Pattern Recognition, vol35, pp 1847 - 1867, 2002.