

## 비디오 영상에서 점진적 장면전환 검출

이광국, 김형준, 김희율

한양대학교 전자통신전파공학과

전화 02-2299-8402 / 핸드폰 011-9737-9694

### Detecting Gradual Transitions in Video Sequences

Gwang-Gook Lee, Hyoung-Joon Kim, Whoi-Yul Kim

Division of Electrical and Computer Engineering, Hanyang University

nohoho@vision.hanyang.ac.kr

#### Abstract

Automated video segmentation is important as the first step of video indexing, video retrieval and other uses. Unlike abrupt changes that are relatively easy to detect, gradual transitions like dissolve, fade-in and fade-out are rather difficult to detect. In this paper, we propose a method for detecting gradual transitions based on local statistics and less dependent to a given threshold level. Experimental results show that the proposed method detected about 85% of gradual transitions.

#### 1. 서론

비디오 데이터의 장면전환 검출은 동영상 검색, 색인 등의 응용을 위한 첫 단계로서 매우 중요하다. 지금까지 제안된 장면전환 검출 방법들에는 연속된 두 프레임간에 픽셀의 차이 비교 [1], 히스토그램의 차이 비교 [2], 모션 벡터의 비교 [3], 그리고 DCT 압축 계수 비교 [4] 등의 방법이 있다.

장면전환 검출을 위한 기본적인 방법은 연속된 프레임 간의 특징값의 차이가 미리 설정된 임계값 이상이면 장면전환이 발생한 것으로 판단하는 것이다. 이런 방법으로는 급격한 장면전

환은 비교적 쉽게 검출할 수 있으나, 점진적인 장면전환은 검출하기 어렵다. 임계값이 너무 높게 설정되면 점진적인 장면전환을 찾아낼 수 없고, 반대로 임계값이 너무 낮게 설정되면 하나의 점진적 장면전환에서 너무 많은 장면전환을 찾아내게 되거나, 물체의 빠른 움직임 등으로 인한 영상의 변화를 장면전환으로 잘못 검출하게 된다.

본 논문에서는 점진적 장면전환 구간 내와 구간 전후의 특징값의 변화를 관찰하여 점진적 장면전환을 검출하는 방법을 제안한다. 제안된 방법은 지역적 통계에 기반하므로 특징값의 일시적인 변화에 둔감하며, 서로 다른 영상에 대해서도 적응적으로 대응하는 장점이 있다.

본 논문에서 사용한 특징량으로는 HSV 컬러 히스토그램을 사용하였으며, 영상의 작은 변화에 덜 민감하도록 히스토그램의 각 성분이 4 bit로 양자화 되었다. 프레임간의 차이(distance)를 구하기 위해서는  $\chi^2$  방법이 선택되었다 [5].

본 논문의 구성은 다음과 같다. 2장에서는 장면전환 검출을 위한 기존의 방법을 소개하며, 3장에서는 본 논문에서 제안된 방법을 소개한다. 4장에서는 실험 결과를 보이고, 마지막으로 5장에서 결론을 맺는다.

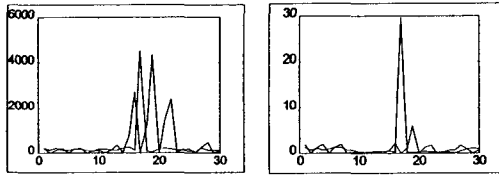
## 2. 기존의 방법

### 2.1. 적응적 임계치의 사용

임계값에 기반한 장면전환 검출방법은 사용된 임계값에 따라 결과가 크게 좌우된다. 이러한 임계값에 대한 의존도를 낮추기 위해 적응적 임계값의 사용이 필요하다. 적응적 임계값을 설정하기 위해 [6]에서는 현재 프레임의 앞, 뒤로 길이  $N$ 인 로컬 윈도우를 씌워 윈도우 내의 특징값의 평균을 이용하여 임계값을 설정하는 방법을 제안하였다.  $m_T$ 는 선택된 적응적 임계값이며,  $T_p$ 는 미리 선택되는 임계값,  $\mu_N$ 은 로컬 윈도우 내의 평균값이다.

$$m_T = T_p \mu_N$$

현재 프레임의 특징값이 위와 같이 얻어진 적응적 임계값보다 크면서, 윈도우 내의 최대값일 경우에 장면전환이 발생한 것으로 판단한다.



(a) distance (b) distance/local mean

그림 1. (a) 고정 임계값을 사용, (b) 적응적 임계값을 사용했을 때 장면전환 부분(실선)과 빠른 움직임이 있는 부분(점선)에서의 특징값의 변화

그림 1은 고정된 임계값을 사용했을 때와 적응적 임계값을 사용했을 때 특징값의 변화를 보여준다. (a)는  $\chi^2$  테스트를 통해 얻어진 특징값의 변화를 나타낸 것이고, (b)는 그 특징값을 윈도우내의 평균값으로 나누었을 때의 변화이다. 실선은 급격한 장면전환시의 특징값의 변화이며, 점선은 물체의 큰 운동이 있는 부분에서의 특징값의 변화를 나타낸다. 고정된 임계값 사용 시 물체의 큰 움직임이 나타나는 구간을 장면전환과 구분하기 어렵지만, 적응적 임계값의 사용으로 이를 쉽게 구분할 수 있음을 볼 수 있다.

### 2.2. Twin Comparison

점진적인 장면전환은 급격한 장면전환에 비해 그림 2와 같이 상대적으로 낮은 특징값이 여러 프레임에 걸쳐 지속적으로 발생하는 특징이 있다.

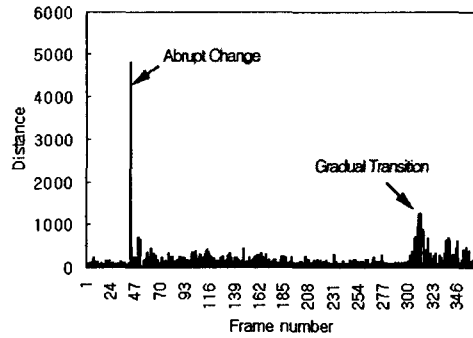


그림 2. 점진적 장면전환의 특징

점진적인 장면전환의 검출을 위해 [7]에서는 twin comparison이 소개되었다. 이 방법은 두 개의 임계값을 설정하여 특징값이 높은 임계값 이상이면 급격한 장면전환으로 판단하고, 낮은 임계값 이상이면 낮은 임계값 이상의 값을 갖는 구간에서 특징값을 누적하여 이 누적한 특징값이 주어진 임계값보다 클 경우에 점진적 장면전환으로 판단한다.

## 3. 제안된 방법

2장에서 소개된 twin comparison 방법은 점진적인 장면전환 구간에서의 특징값의 변화에 기반하여 점진적 장면전환 검출의 기본적인 방법을 제시한다. 하지만, 임계값이 정확하게 선택되지 못했을 경우에는 장면전환 구간을 정확하게 찾아내지 못하며 이는 잘못된 판단으로 이어지게 된다. 또 장면전환 구간 내에서 특징값이 일시적으로 임계값 아래로 떨어지는 등의 순간적인 변화에 민감하게 반응하며, 이 또한 잘못된 결과로 이어지게 된다.

본 논문에서는 이런 문제를 해결하기 위해 다음과 같은 가정을 두고 적응적 임계값과 twin

comparison을 이용한 장면전환 검출 방법을 제시한다.

1. 점진적 장면 전환이 일어나는 구간에서 특징값은 그림 3 (a)와 같은 모양으로 변화한다.
2. 1과 같은 변화 전과 후에는 장면이 전환된다.

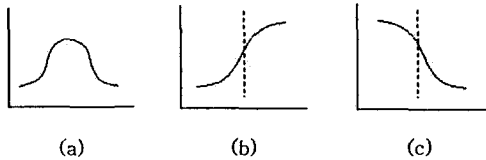


그림 3. (a) 점진적 장면전환, (b) 점진적 장면전환 시작 지점, (c) 점진적 장면전환 끝 지점에서의 특징값의 변화

첫 번째 가정은 실험을 통하여 정도의 차이는 있지만 대부분의 점진적 장면전환에서 만족함을 확인할 수 있었다. 두 번째 가정은 장면전환이 당연히 갖게 되는 특징이다.

점진적인 장면전환의 시작부분에서는 그림 3의 (b)와 같은 특징값의 변화가 일어나기 때문에, 로컬 윈도우의 오른쪽 부분의 평균은 왼쪽 부분의 평균보다 크다. 장면전환의 끝 부분에서는 마찬가지로 그림 3의 (c)와 같은 변화가 발생한다. 따라서, 다음과 같은 조건으로 장면전환의 구간을 판단할 수 있다.

$$\text{Start of transition if : } \alpha\mu_{left} < \mu_{right}$$

$$\text{End of transition if : } \mu_{left} > \alpha\mu_{right}$$

$$\text{where, } \mu_{left} = \frac{1}{N} \sum_{i=0}^{N-1} d(i), \mu_{right} = \frac{1}{N} \sum_{i=N+1}^{2N-1} d(i)$$

$\mu_{left}$  와  $\mu_{right}$  는 각각 윈도우 내의 특징값의 왼쪽 부분과 오른쪽 부분의 평균이며,  $\alpha$  는 미리 설정된 임계값이다.  $d(i)$  는 길이  $2N+1$  인 윈도우 내의  $i$  번째 프레임의 특징값을 말한다. 이러한 장면전환 구간 판단 방법은 윈도우 내의

평균값을 이용하기 때문에 특징값의 일시적인 변화에 덜 민감하게 동작한다.

이렇게 결정된 점진적 장면전환 구간 내에서 누적된 특징값이 주어진 임계값보다 클 경우 점진적 장면전환이 발생하였다고 판단하였다. 이때의 임계값은 장면전환 시작 부분에서의 특징값의 평균을 이용하여 적응적으로 선택되었다.

한편 화면 내에서 커다란 물체의 움직임 등에 의한 변화도 그림 3의 (a)와 같은 모양을 갖기 때문에, 이런 변화가 장면전환으로 잘못 판단될 수도 있다. 그러나, 이런 경우에는 실제적인 장면전환이 일어나지 않으며, 두 번째 가정을 만족하지 못한다. 따라서, 장면전환의 시작 부분과 끝 부분의 영상의 특징량의 차이가 주어진 임계값 이상이 될 때에만 장면전환으로 판단하면 이러한 오류를 줄일 수 있다. 이때의 임계값은 급격한 장면전환 검출시의 임계값과 같은 방법으로 윈도우 내의 특징값의 평균을 이용한 적응적인 임계값이 선택되었다.

#### 4. 실험 및 결과

실험 영상으로는 영화 예고편과 뮤직 비디오 영상이 선택되었다. 이들은 빠른 움직임을 많이 포함하고 있어 장면전환 검출에 어려움이 많지만, 다양한 점진적 장면전환을 포함하고 있기 때문에 실험 영상으로 선택되었다.

표 1은 실험 영상의 총 프레임 수, 존재하는 장면변환의 개수 등 사용한 영상들의 정보를 나타내고 있다. 표 2에는 실험 결과가 나타나있다. 실험결과에서 recall과 precision은 아래와 같이 정의된다.

$$\text{recall} = \frac{N_c}{N_c + N_m} \times 100\%$$

$$\text{precision} = \frac{N_c}{N_c + N_f} \times 100\%$$

$N_c$  는 검출된 장면전환 중 올바른 장면전환의

개수이며,  $N_m$  은 검출하지 못한 장면전환의 개수,  $N_f$  는 잘못 검출된 장면전환의 개수를 나타낸다.

표 1. 실험에 사용된 영상의 정보

No.	No. of frames	No. of Abrupt changes	No. of Gradual transitions
1	4005	43	29
2	3616	55	22
3	5552	0	34
4	6858	78	37

표 2. 장면전환 검출 결과

No.	Abrupt change		Gradual transition	
	recall	precision	recall	precision
1	93%	98%	81%	88%
2	94%	98%	88%	65%
3	-	-	87%	77%
4	95%	97%	81%	89%

실험 결과 제안된 방법은 급격한 장면전환 및 점진적 장면전환 모두에서 만족할만한 성능을 보여주었다. 잘못된 검출 결과는 주로 카메라의 움직임에 의해 화면에 전체적인 변화가 발생한 경우, 점진적 장면전환이 천천히 진행되어 장면 전환의 시작지점을 판단하지 못한 경우, 화면상에 강한 조명의 변화가 발생한 경우 등에 의해 발생하였다. 실험 영상 2에서 특히 낮은 정확도를 보인 것은 해당 영상이 다수의 급격한 움직임과 조명변화를 포함하고 있기 때문이다.

## 5. 결론

본 논문에서는 지역적 통계치에 기반하여 적응적으로 동작하는 점진적 장면전환 검출 방법을 제안하였다. 실험을 통해서 제안한 방법이 급격한 장면전환 뿐 아니라 점진적인 장면전환에 대해서도 잘 동작함을 보였다. 본 논문에서 제안한 방법은 로컬 윈도우 내의 평균값에 기반하므로 다양한 영상에 대하여 적응적으로 대응하며,

노이즈와 같은 일시적인 변화에 둔감한 장점을 갖는다.

카메라의 움직임에 의해 화면이 전체적으로 변한 경우와 점진적 장면전환이 너무 천천히 진행된 경우 등에는 장면전환을 제대로 검출하지 못했다. 잘못된 결과를 줄이기 위해 카메라 움직임의 검출, 작은 변화에도 반응할 수 있는 에지와 같은 특징량의 사용 등에 대한 연구가 필요하다.

## References

- [1] B. Furht and S. W. Smoliar, *Video and image processing in multimedia systems*, Kluwer Academic Publishers, 1995.
- [2] H. J. Zhang, A. Kankanhali, and S. W. Smoliar, "Automatic Partitioning of Full-motion Video," *Multimedia Systems*, Vol. 1, 1993, pp. 10-28.
- [3] F. Arman, A. Hsu, and M. Y. Chiu, "Image Processing on Encoded Video Sequences," *Multimedia Systems*, Vol. 1, 1994, pp. 211-219.
- [4] B. Shahraray, "Scene Change Detection and Content-Based Sampling of Video Sequences," *Digital Video Compression : Algorithms and Technologies*, SPIE 2419, 1995, pp. 2-13.
- [5] G. Lupatini, C. Saraceno, and R. Leonardi, "Scene break detection : a comparison," *Image Processing*, Vol. 2, 2001, pp. 394-397.
- [6] Y. Yusoff, W. Christmas, and J. Kittler, "Video Shot Cut Detection Using Adaptive Thresholding," *British Machine Vision Conference*, 2000, pp. 362-372.
- [7] S. W. Smoliar and H. J. Zhang, "Content-Based Video Indexing and Retrieval," *IEEE Multimedia*, Vol. 1, 1994, pp. 62-72.