

동영상 컷 검출을 위한 가변형 동적 임계값 기법

정회원 엄 성 주*, 김 우 생*

Variable Dynamic Threshold Method for Video Cut Detection

Sung-ju Youm*, Woo-saeng Kim** *Regular Members*

요 약

컷 검출은 내용기반 검색에 필요한 인덱싱을 위해 수행되어야 하는 기초 작업으로 이를 위한 매우 다양한 기법들이 제안된바 있다. 그러나 기존의 연구에서는 대부분 고정된 하나의 임계값을 사용하기 때문에 동영상의 종류나 내용에 따라 최적의 임계값을 정해야만 하는 문제점을 갖는다. 본 논문에서는 컷 검출 간격의 확률적인 분포에 따라 임계값을 조절하며 컷이 발생하면 이전 컷과의 간격과 특징값 차이를 다음 컷 검출을 위한 임계값 설정에 반영하는 가변형 동적 임계값 방법을 제안한다. 이를 위해 임계값 조절에 필요한 인자 값들을 실행시간에 구하는 방법과 이를 사용한 컷 검출 알고리즘을 제시한다. 또한 실험을 통해 제안하는 방법이 기존의 방법에 비해 오 검출율을 줄일 수 있어 효율적임을 보인다.

ABSTRACT

Video scene segmentation is fundamental role for content based video analysis and many kinds of scene segmentation schemes have been proposed in previous researches. However, there is a problem, which is to find optimal threshold value according to various kinds of movies and its content because only fixed single threshold value usually used for cut detection. In this paper, we proposed the variable dynamic threshold method, which change the threshold value by a probability distribution of cut detection interval and information of frame feature differences and cut detection interval in previous cut detection is used to determine the next cut detection. For this, we present a cut detection algorithm and a parameter generation method to change the threshold value in runtime. We also show the proposed method, which can minimize fault alarm rate than the existing methods efficiently by experimental results.

1. 서 론

컴퓨터와 통신 그리고 데이터 압축기술의 발달로 VOD(video on demand)와 같이 동영상 데이터를 이용한 다양한 서비스가 가능하게 되었다. 영상, 음성, 문자 정보 등 다양한 정보를 갖으면서 데이터 양이 방대한 동영상으로부터 사용자가 필요로 하는 정보를 찾기 위해서는 내용에 기반한 색인 기법이 필요하며, 일반적으로 비디오 색인을 하기 위한 출발점은 컷 검출(cut detection)이라고 볼 수 있다.

압축/비압축 도메인을 대상으로 한 기존의 컷 검출에 관한 연구에서는 매우 다양한 형태의 컷 검출 기법들이 제안된 바 있다¹⁾. 기존에 제안된 대부분의 방법들은 인접 프레임 비교를 통해 프레임간의 변화를 측정하기 위해 프레임의 픽셀, 히스토그램, 에지 성분, 텍스트, 모션, DCT 정보 등 다양한 특징들을 사용하였다. 이 중에서 가장 보편적으로 사용되는 히스토그램 기반의 방법은 컬러 분포를 이용하는 방법으로써 구현이 쉬우며 적은 연산 비용으로 갑작스러운 장면전환(abrupt change)에 의한 컷을 효과적으로 찾아낼 수 있다는 장점을 갖는다.

* 광운대학교 컴퓨터공학부 멀티미디어 연구실(sjyoum@cs.kwangwoon.ac.kr)
논문번호 : 010395-1214, 접수일자 : 2002년 1월 8일

그러나 명암의 변화에 민감하기 때문에 컷이 아님에도 이를 컷으로 검출해 내는 오 검출(fault cut)이 발생한다는 단점을 갖는다. 이를 보완하기 위한 연구에서는 형태 정보를 추가적으로 사용하는 기법들을 제안한 바 있다^{[2][3]}. 히스토그램 기반의 방법과는 전혀 다른 시도를 한 연구들^{[4][5][6]}에서는 프레임들을 구성하는 각각의 화소값의 차이, 에지 정보 그리고 움직임 벡터 등을 사용하여 인접 프레임간에 연속성을 비교하는 방법으로 컷을 찾아내려고 시도했다. 이들 연구에서는 프레임간의 컬러 정보를 거의 사용하지 않음으로써 히스토그램 기반의 방법들에서 발생하는 문제점을 많이 보완할 수 있었지만 움직임과 잡음에는 민감한 단점을 갖는다.

인접 프레임 비교를 통한 컷 검출 방법들에서 컷을 판단하는 기본적인 원리는 인접 프레임간의 특징값 차이가 미리 정해진 임계값(threshold)을 초과하는 경우 이를 컷으로 판단하는 것이다. 따라서 효과적으로 컷을 검출하기 위해서는 적절한 임계값을 선택해야 하는 것이 중요한 것임에도 불구하고 기존의 연구에서는 대부분 특정 데이터들을 대상으로 한 실험을 통해 경험적으로 찾아낸 고정된 임계값을 주로 사용해 왔다. 그러나 이러한 경험적인 접근 방법은 다양한 형태의 실세계 동영상들에 일반적으로 적용하기 매우 곤란하다. 임계값 설정에 관한 앞선 연구에서는 이러한 문제점을 해결하기 위해 프레임간의 변화 형태나^{[7][8][9]} 컷이 발생하는 통계적 간격에^[10] 기반을 둔 컷 검출 기법들을 제안한 바 있다. 또한 동영상 전체를 작은 구간들로 나누어 각 구간마다 프레임간의 평균 변화율에 대한 다음 프레임의 차이를 이용해 지역 윈도우에서 적절한 임계값(adaptive local window threshold)을 설정하는 방법과^[7] 프레임간의 특징값 변화를 2차 미분하여 변화율의 차이가 없는 경우 이들을 컷으로 검출하지 않는 방법을 통해 연속적으로 발생할 수 있는 오 검출을 줄일 수 있는 기법 등이^[8] 제안된 바 있다. 또 다른 연구에서는 프레임들의 평균 컬러를 구하고 이를 임계값 설정에 반영하는 방법을 통해 고정된 임계값의 단점을 해결하려 시도했다^[9].

기존의 컷 검출 방식에서 주로 사용되었던 고정된 임계값 방식은 히스토그램 분포도와 같은 특징들이 동영상마다 다르기 때문에 문제가 될 수도 있지만, 또한 하나의 영상이라도 많은 장면과 색상 변화가 있기 때문에 같은 임계값을 지속적으로 적용하는 것은 문제가 될 수 있다. 이러한 문제점을 해결하여 컷 검출의 효율성을 향상시키기 위해서

컷 검출 과정 중 임계값이 동적으로 변화하는 동적 임계값 기법을 제안된 바 있다^[10]. 그러나 이 연구에서는 임계값의 변화를 위해 여전히 관측을 통해 얻어진 컷간의 간격을 사용했으며 프레임간의 특징값 변화 형태를 고려하지 않은 문제점이 있었다. 따라서 본 연구에서는 검출되는 컷의 시간적 분포와 프레임간의 특징값 변화 추이 그리고 이미 검출된 컷들에 대한 정보를 이용하여 임계값 설정을 위한 인자들이 가변적으로 변하며 자동으로 설정되는 새로운 가변형 동적 임계값 기법을 제안한다. 이를 위해 본 논문에서는 가변형 동적 임계값을 구하기 위해 필요한 인자 값들을 런타임 때 구하는 방법과 이를 사용한 컷 검출 알고리즘을 제시한다. 제안하는 가변형 동적 임계값 기법은 임계값을 사용하는 대부분의 컷 검출 기법에 적용이 가능하며 동영상에 대한 분석이 진행되면서 발생하는 컷들의 정보에 의해 자동으로 검출 간격과 임계값이 정해진다. 그러므로 동영상의 종류에 구애받지 않고 적용할 수 있으며 빠른 움직임이나 카메라 후래쉬 등과 같이 연속적인 특징값의 변화에 의한 오 검출을 줄일 수 있는 장점을 갖는다.

제안하는 기법은 실세계의 다양한 동영상 데이터를 사용한 실험에서도 고정 임계값을 사용하는 경우 보다 효과적으로 컷을 검출해냄을 알 수 있었다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 가변형 동적 임계값 기법에 대하여 설명한다. 3장에서는 실험 및 결과 분석을 하고 마지막으로 4장에서는 결론을 내린다.

II. 가변형 동적 임계값을 이용한 컷 검출 기법

이번 장에서는 먼저 동영상 데이터에서 컷이 발생하는 시간적인 분포와 특징값 변화에 따른 컷의 특성에 대해서 살펴본다. 그리고 이를 바탕으로 가변형 동적 임계값을 사용한 컷 검출 기법에 대해 기술한다.

1. 컷 발생과 변화에 따른 특성

동영상 데이터의 변화에 따른 특성은 동영상의 종류 및 내용과 관련된다. 본 논문에서는 이러한 특성을 시간적인 특성과 특징값 변화에 대한 특성으로 고찰해 보기로 한다. 동영상 데이터에서 발생하는 컷은 영상의 종류에 따라 발생하는 빈도 수나 형태가 매우 다양하게 나타나지만 공통적으로 대부분의 컷들 사이의 간격이 최소한의 프레임 수를 초

과하는 것을 알 수 있다. 그림 1은 각종 동영상 데이터에서 컷이 발생한 직후 다음 컷이 검출되는 시간을 통계적 조사로 나타낸 것으로, 뉴스나 드라마와 같이 컷의 간격이 상대적으로 일정하게 나타나는 동영상들의 경우 대부분 컷이 발생한 후 약 40~100 프레임 이내에 다음 컷이 발생할 확률이 가장 높으며 변화가 심한 광고 영상의 경우는 대략 20 프레임 이내에 다음 컷이 발생할 확률이 높음을 알 수 있다. 따라서 실제 동영상들에서 컷이 발생한 후 대략 10프레임 이내에 컷이 발생할 확률은 거의 없는 것으로 관측되었다. 이러한 관측을 통해 동영상의 종류에 따라 정도의 차이는 있지만 일반적으로 컷이 발생한 후 바로 다음 몇 개의 프레임 내에서 컷이 다시 발생할 확률은 매우 낮고 일정 시간이 지나면서 컷이 발생할 확률이 점점 높아지다가 일정한 시간이 지나면 다시 확률이 서서히 낮아짐을 알 수 있다.

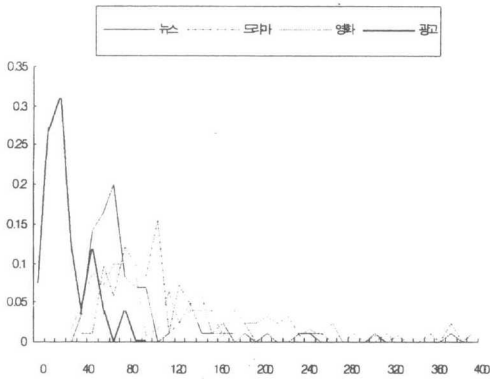
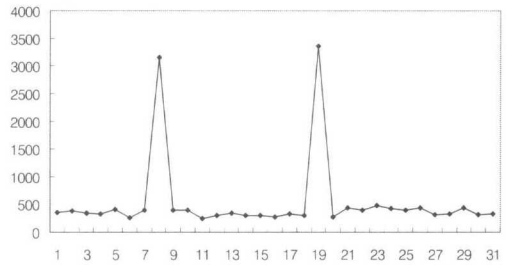
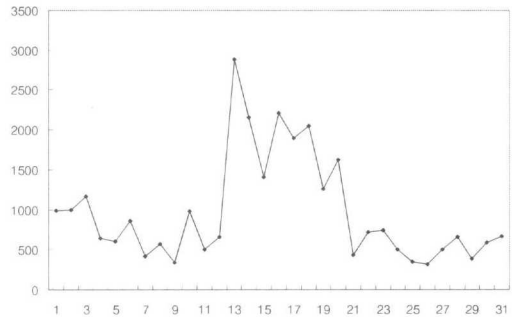


그림 1. 컷이 발생한 후 시간이 지남에 따라 다음 컷이 발생하는 확률

컷이 발생하는 전후의 프레임 간 특징값 변화를 살펴보면 영상의 변화에 따라 특징값의 변화하는 형태가 서로 다른 것을 알 수 있다. 그림 2는 특징값 변화에 따른 특성을 보여주는 것으로 x축은 각 프레임, y축은 각 프레임과 이전 프레임과의 특징값 차이를 나타내 주고 있는 그림이다. 그림 2의 (a)의 경우는 컷이 발생하기 전이나 발생 한 직후 프레임들 간의 특징 값의 차이가 작은 경우로 고정된 임계값을 사용해서 대부분의 갑작스러운 컷을 검출해내기 용이하다. (b)의 경우는 화면상의 객체들의 변화가 많아 컷이 발생한 후에도 프레임들 간의 특징값의 변화가 심한 경우를 보여준다. 이는 빠른 움직임이나 조명등으로 인한 변화로 볼 수 있다.



(a) 안정적인 경우



(b) 변화가 심한 경우

그림 2. 화면의 종류에 따른 특징 값 차이의 변화

(b)의 경우 고정된 임계값을 적용하면 연속적인 오검출(fault cut) 상황이 발생하게 된다. 따라서 임계값에 의해 컷 검출을 시도하는 기법들의 경우 화면들의 변화가 안정적인 경우와 움직임이 심할 경우를 고려해야 효과적으로 컷을 검출해 낼 수 있을 것이다.

고정된 임계값을 사용한 기존의 컷 검출 기법들은 이러한 점을 고려하지 않기 때문에 컷이 발생한 후 몇 프레임이 지나지 않아서 컷이 아닌데도 불구하고 잡음이나 빛, 카메라 동작 등의 영향으로 인하여 컷으로 잘못 검출하는 경우가 많다. 일반적으로 고정된 임계치를 낮게 잡으면 제대로 된 컷을 발견할 확률이 높으나 잘못된 컷을 발견할 확률 또한 높아진다.

반면에 고정된 임계치를 너무 높게 잡으면 잘못된 컷을 발견하는 경우는 적어지나 제대로 된 컷을 발견하는 경우도 줄어들게 된다. 따라서 본 연구는 이러한 두 가지 특성을 고려하여 컷이 발생할 가능성이 많은 구역에서는 임계값을 낮추고, 컷이 일어날 가능성이 적은 구역에서는 임계값을 높임으로써 실제의 컷은 검출하되 잘못된 컷은 가급적 줄이자는 데 목적이 있다.

2. 가변형 동적 임계값

제안하는 가변형 동적 임계값은 컷의 발생 유무와 프레임간의 특징값 변화의 특성 두 가지를 고려해서 결정된다. 먼저 컷의 발생을 고려한 가변형 동적 임계값 T_{dyn} 은 식 (1)과 같이 정의된다. 여기서 T_{low} 란 기존의 고정된 임계값과 동일한 것으로 실험을 통하여 적절한 값을 구할 수 있으며 T_{high} 란 컷이 일어난 직후 동적으로 높아진 임계값을 말한다. 또한 t 란 컷이 발생한 후 경과되는 시간(프레임 수)을 의미하며 0부터 t_{max} 까지의 값을 갖는다. 즉 가변형 동적 임계값은 평상시에는 T_{low} 의 고정된 임계값을 유지하다가 컷이 발생한 직후는 임계값이 T_{high} 값으로 높아진 후 중간에 새로운 컷이 발생하지 않는 한 시간이 경과되면서 임계값이 낮아져 최대 t_{max} 시간 후에는 T_{low} 값과 같아지게 된다.

$$T_{dyn} = T_{low} + (T_{high} - T_{low}) \cdot f(t) \quad (1)$$

(단, $T_{high} \geq T_{low}$)

따라서 T_{dyn} 은 그림 1에서 나타난 컷 발생 확률 분포와 동일하게 컷 발생 직후 급격히 증대한 후 서서히 값이 감소하게 되며 감소 추세는 식 (1)에서 함수 $f(t)$ 를 어떤 식으로 사용 하느냐에 따라 여러 가지 형태를 만들 수 있다. 함수 $f(t)$ 는 선형 함수와 지수형 함수를 사용 할 경우 각각 식 (2)와 (3)처럼 정의된다.

$$f(t) = 1 - \frac{t}{t_{max}} \quad (\text{단, } 0 \leq t \leq t_{max}) \quad (2)$$

$$f(t) = \alpha^{-\frac{t}{t_{max}}} \quad (\text{단, } \alpha \geq 1, 0 \leq t \leq t_{max}) \quad (3)$$

앞의 식들에서 T_{high} 와 t_{max} 값은 런 타임시 구해지는 가변 값으로써 본 연구에서는 앞 절에서 언급한 프레임간의 특징값 변화 특성과 현재 컷 이전에 앞서 검출된 컷들의 정보를 사용하여 구하는 방법을 제시한다. 먼저 T_{high} 는 기본적으로 컷으로 판단되는 프레임과 바로 이전 프레임간의 차이와 과거 컷 발생 후의 프레임들의 평균 변화율을 더해서 구한다. n 번째 컷으로 판단된 컷 프레임을 F_n , F_n 의 바로 이전 프레임을 F'_n 라 하고 함수 $V(\alpha)$ 를 프레임 α 의 특징값이라 하면 n 번째 컷 프레임에서의

T_{high} 값인 T_{high}^n 은 식 (4)와 같이 정의된다.

$$T_{high}^n = \frac{\sum_{k=0}^{n-1} |V(F_k) - V(F'_k)|}{n-1} \quad (4)$$

가변형 동적 임계값을 적용한 컷 검출 기법은 특징값의 변화가 T_{dyn} 를 초과하는 프레임을 컷으로 검출해내는 것을 기본으로 한다. 그러나 그림 2의 (b)와 같이 특징값의 변화가 연속적으로 임계값을 초과하는 경우에는 연속적인 오 검출 현상이 일어날 수 있다. 이러한 경우 컷이 발생한 프레임의 바로 이전 프레임에서의 특징 값 변화가 어떠한 형태였는가 살펴보는 것이 중요하다. 만약 이전 프레임에서의 특징값 차이도 높았다면 이는 연속적인 프레임 변화의 일부로 볼 수 있다. 따라서 본 논문에서는 특징값의 변화가 T_{dyn} 를 초과하는 프레임의 바로 이전 프레임에서의 특징값 차이가 T_{low} 를 초과했는가를 점검하여 이전 프레임의 특징값 차이가 T_{low} 를 초과했다면 이를 t_{max} 시간동안 새로운 컷으로 인정하지 않는 방법을 사용한다.

t_{max} 의 값은 이전에 발생한 컷들을 대상으로 컷 발생 후 다음 컷이 발생 될 때까지 걸린 시간들의 평균으로 구한다. t_{max}^n 를 n 번째 컷이 발생했을 경우 $n-1$ 번째 컷이 발생한 후 걸린 시간 즉, 프레임 수로 정의할 때 이는 식 (5)처럼 표현된다. 단 최초의 컷일 경우는 다른 동영상들로부터 실험을 통해 얻어진 적절한 값을 적용한다.

$$t_{max}^n = \frac{\sum_{k=0}^{n-1} t_{max}^k}{n-1} \quad (5)$$

가변형 동적 임계값의 장점은 동영상의 종류에 따라 또는 같은 동영상이라도 장면 전환의 빈도와 프레임간의 변화에 따라 임계값이 달라진다는 것이다. 예를 들어 장면 전환이 느리며 프레임간의 특징값 차이가 큰 동영상의 경우, 자동으로 임계값이 높아지고 이를 적용하는 기간도 길게 설정된다. 이런 점들을 고려한 가변형 동적 임계값을 사용한 컷 검출 알고리즘은 다음과 같다.

- 1) 프레임 카운터 i , T_{low} 그리고 t_{max} 를 초기화한다.
- 2) i 를 증가시키며 임계값 T_{low} 를 적용하여 컷을 검출한다.

- 3) 컷이 검출되면 새로운 T_{high} 와 t_{max} 값을 계산한다.
- 4) F_i 가 마지막 프레임이면 종료하고, 아닐 경우는 i 를 증가시킨 후 임계값 T_{dyn} 를 적용한다.
- 5) $T_{dyn} < V(F_i)$ 일 때 $V(F_{i-1}) < T_{low}$ 를 만족하면 F_i 를 컷으로 판정하고 3)로 진행하며, 아닐 경우는 4)로 진행한다.

그림 3은 제안하는 가변형 동적 임계값 기법과 기존의 고정 임계값 기법에서 컷을 검출해내는 상황을 실제 동영상 데이터를 대상으로 수행해 본 것으로 함수 $f(t)$ 는 식 (3)을 사용한 경우이다. 여기서 x축은 프레임 수를 y축은 특징값의 변화를 나타내며 실제 컷이 일어나는 지점은 cut으로 표시했다. 고정 임계값 방법이나 가변형 동적 임계값 방법은 모두 첫 번째 컷을 검출하는 방식은 동일하다. 그러나 그 이후의 처리과정은 서로 다른 형태를 보여준다. 고정 임계값 방법에서는 구간 A에서 첫 번째 컷을 검출한 후 15~19 프레임 사이의 값들이 또 다시 임계값을 넘기 때문에 이들을 컷으로 검출하는 오류가 발생한다. 그러나 가변형 동적 임계값 방법은 첫 번째 컷이 발생한 직후 임계값이 T_{dyn} 연산에 의해 높아지기 때문에 이를 컷으로 검출하지 않게 된다. 구간 B와 C에서는 프레임간의 특징값 차이가 대체적으로 크기 때문에 고정 임계값을 사용할 경우 연속적인 오 검출 현상이 일어나게 된다. 그러나 가변형 동적 임계값의 경우 앞서 검출된 컷에 의해 T_{high} 의 평균값이 충분히 높아지게 되므로 28과 49 프레임을 만을 컷으로 검출하여 고정 임계값에서 나타날 수 있는 오 검출을 피할 수 있다. 마지막으로 구간 D에서는 컷 검출 후 계속적으로 높은 프레임 변화에 의해 연속적인 오 검출이 나타나게 되는

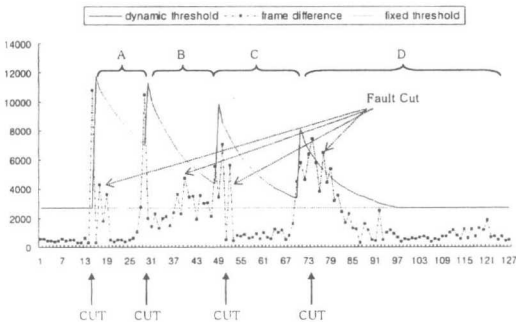


그림 3. 가변형 동적 임계값과 고정 임계값의 컷 검출 방법 비교

상황이다. 이 구간에서는 특징값의 변화가 T_{dyn} 를 초과하는 프레임의 바로 이전 프레임들의 특징값 변화가 모두 T_{low} 를 초과하고 있으므로 제한한 알고리즘에 의해 이를 연속적인 변화로 간주하여 70 프레임만을 컷으로 검출하며 시간이 지나면서 T_{dyn} 값은 감소하여 T_{low} 값에 근접하게 된다.

III. 실험 및 결과 분석

본 연구에서 제안하는 가변형 동적 임계값 방법을 시험하기 위한 프로토타입은 MS-WINDOWS 2000 환경에서 MS Visual C++ 6.0 과 Directshow 라이브러리를 통해 제작되었다. 시험용 데이터는 TV 방송에서 얻은 영화와 뉴스 등으로 실시간에 인코딩 된 MPEG-1 데이터이다. 각 데이터는 29 frames/sec의 프레임 율과 1152000 bits/sec의 비트 율을 가지며 352 x 240의 해상도를 갖는다. 동영상 분석을 위해 추출되는 각 프레임들은 빠른 처리를 위해 80 x 80으로 정규화 시키고64-bin의 그레이 이미지로 변환하여 사용했다^[11]. 초기 설정을 위해 사용한 값은 다양한 동영상 데이터로부터 실험을 통해 관측된 적절한 값으로, 최초 컷이 발생하기 전까지 $t_{max}=25$, $T_{low}=2000$ 으로 설정했다. 제안하는 동적 임계값 기법에서 T_{low} 는 일반적인 히스토그램 비교 시에 사용하는 고정 임계값 보다 다소 낮게 사용한다. 왜냐하면 초기에 발견되는 컷의 정보를 임계값 설정에 빠르게 반영하고 T_{dyn} 에 의해 오 검출에 대한 부담이 적어지므로 상대적으로 미 검출 (miss detection)을 억제할 수 있기 때문이다.

제안하는 가변형 동적 임계값 기법의 성능 비교를 위해서 컷 검출을 위해 사용되는 가장 일반적인 그레이 레벨 히스토그램 기법을 사용했다. 전체 컬러 수를 n 이라 하고 $H_i(j)$ 를 프레임 F_i 에서 그레이 레벨 j 의 히스토그램 값이라 할 때 인접한 두 프레임 F_i 와 F_{i-1} 간의 히스토그램 차이는 식 (7)과 같이 정의된다.

$$\sum_{j=1}^n |H_{i-1}(j) - H_i(j)| \quad (7)$$

1. 가변형 동적 임계값 기법의 실험 결과

실제 컷의 위치와 검출된 컷의 비교를 위해 우선 사람의 눈에 의해 각 비디오 데이터에서 정확한 컷의 위치를 발견하였으며 이 데이터를 실제 컷의 근

거로 삼았다. 사용된 데이터 중 뉴스는 장면의 전환이 비교적 확실한 반면 영화는 줌, 패닝 등의 카메라 기법을 통한 점진적인 장면 변화 등으로 장면 전환이 확실하지 않은 경우가 많았다. 표 1은 같은 동영상 데이터를 대상으로 고정 임계값과 가변형 동적 임계값을 적용해 얻어진 결과로써 지수형 임계값 평가에는 실험을 통해 가장 효율적으로 판단된 $\alpha=6.0$ 을 사용하였다.

표 1. 컷 검출 결과 비교

D:Detected M:Missed F:False Alarm

Video Data	실제 CUT	Fixed			Linear			Exponential		
		D	M	F	D	M	F	D	M	F
CF 1850 frames	66	60	5	62	55	10	13	59	6	10
Drama 5784 frames	62	54	8	5	54	8	0	54	8	0
News 30072 frames	76	68	8	4	66	10	0	68	8	0

실험 결과 전체적으로 고정 임계값에 비해 가변형 동적 임계값을 사용할 경우 오 검출율이 줄어들게 되는 것을 알 수 있었다. 특히 컷간의 간격이 짧고 빠른 움직임이 많이 포함되어 오 검출 발생 빈도가 대단히 높은 광고 영상의 경우에 동적 임계값 방법이 매우 효율적임을 알 수 있었다. 뉴스와 드라마와 같이 움직임이 적고 비교적 컷의 구분이 확연하게 나타나는 영상들의 경우에도 가변형 동적 임계값을 적용함으로써 보다 효과적으로 컷을 검출할 수 있었다. 가변형 동적 임계값 방법 중 지수형 함수를 사용하는 경우 컷이 일어나는 형태에 따라 부분적으로 우수한 성능을 나타내기도 하였지만 표 1에서 알 수 있듯이 가변형 동적 임계값 방법은 감소 함수의 종류에 상관없이 대체적으로 고정 임계값 방법보다 효율적임을 알 수 있다.

정확도 평가는 일반적으로 많이 사용되는 recall과 precision으로 나타냈으며 이들은 식 (8)와 같이 정의된다.

$$\begin{aligned}
 recall &= \frac{detected}{detected+missed} \\
 precision &= \frac{detected}{detected+false\ alarm}
 \end{aligned}
 \tag{8}$$

각 동영상에 대한 전체적인 결과를 recall과 precision으로 나타내 보면 표 2와 같다. 가변형 동적 임계값의 경우 컷 발생 직후에 높아지는 임계값

에 의해 오 검출을 크게 줄일 수 있기 때문에 전체적으로 precision이 향상된 것을 알 수 있다. 가변형 동적 임계값의 최소 값인 T_{low} 는 근본적으로 고정 임계값과 같은 의미를 갖기 때문에 recall의 경우는 세 가지 방법 모두 비슷한 수준을 나타내고 있다.

표 2. 전체적인 recall과 precision 비교

Total	Fixed		Linear		Exponential	
	Recall	Precision	Recall	Precision	Recall	Precision
Video 37706 Frame	0.89	0.72	0.86	0.93	0.89	0.95

서로 다른 형태의 동영상 데이터에 대한 실험에서는 데이터의 종류에 따라 컷 발생 간격이 틀려지므로 t_{max} 값이 결과에 영향을 미치게 된다. 동영상의 종류에 따라 컷 발생간격은 차이가 있지만 제안하는 기법의 경우 약 5~10번째 컷이 출현한 이후부터는 최적의 컷 간격에 t_{max} 가 근접함을 알 수 있었다.

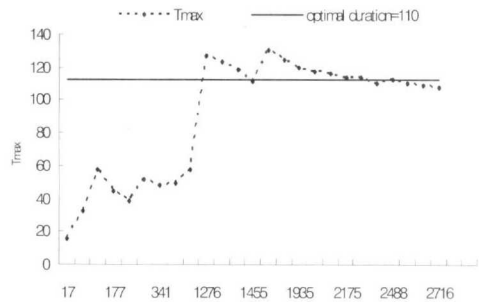


그림 4. 경험적인 최적치에 근접하는 의 변화

그림 4는 뉴스 동영상에서 가변적으로 구해지는 t_{max} 의 변화 추이를 나타내는 것으로 본 논문에서 제안한 가변 t_{max} 값을 사용했을 경우에도 경험적으로 얻어진 최적의 값에 효과적으로 근접하고 있음을 알 수 있었다.

제안하는 기법을 통해 실제동영상에서 컷으로 검출된 프레임들을 살펴보면 그림 5와 같다. 그림 5는 광고 동영상에서 조명이 의해 오 검출 될 확률이 높은 상황을 보여준다. 이 부분은 공연장의 모습으로 무대 조명과 카메라의 빠른 줌 효과가 사용되는 장면이다. 68과 87번째 프레임은 컷으로 판단된 프레임이므로 68~86번째 프레임까지가 하나의 장면을

구성한다. 그리고 87번째 프레임부터는 다른 장면으로 진행된다.



그림 5. 선평에 의한 오 검출 가능성이 있는 영상의 예

이 광고 동영상에서 장면 중간에 무대 조명의 선평으로 인해 고정 임계값을 사용할 경우 선평 부분에 해당하는 프레임들을 오 검출하게 되어 73에서 75 번째 프레임들을 모두 컷으로 판정하게 될 가능성이 높다. 그러나 제안한 방법의 경우 충분히 높아진 임계값에 의해 선평부분에서 오 검출하지 않고

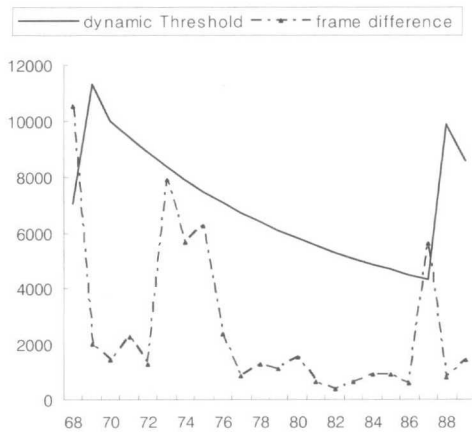


그림 6. 실제 동영상에서의 적용 예

앞선 컷에 의해 조절된 t_{max} 에 의해 임계값은 낮아지게 되어 실제 컷인 87번째 프레임은 검출할 수 있게 된다. 따라서 제안하는 가변형 동적 임계값 방법에서는 68번째 프레임과 87번째 프레임만을 정확히 컷으로 인식하는 장점을 보였다. 그림 5에 대한 분석 결과는 그림 6과 같다.

IV. 결론

인접 프레임간의 특징값을 임계값과 비교하는 컷 검출 기법들은 임계값을 어떻게 설정하느냐에 따라 성능이 크게 변하게 된다. 따라서 적절한 임계값을 설정하는 것은 매우 중요한데 다양한 형태의 실제계의 동영상 데이터로부터 고정된 하나의 임계값을 추출한다는 것은 어려운 문제가 아닐 수 없다. 일반적으로 고정된 임계값을 낮게 잡으면 제대로 된 컷을 발견할 확률이 높으나 잘못된 컷을 발견할 확률 또한 높아진다. 반면에 고정된 임계값을 너무 높게 잡으면 잘못된 컷을 발견하는 경우는 적어지나 제대로 된 컷을 발견하는 경우도 줄어들게 된다.

본 연구에서는 이러한 문제를 보완하기 위하여 컷들간의 간격과 프레임간의 특징값 변화를 바탕으로 임계값이 변화하는 가변형 동적 임계값 기법을 제안하였다. 제안된 방법은 컷이 발생할 가능성이 많은 구역에서는 임계값을 낮추고, 컷이 일어날 가능성이 적은 구역에서는 임계값을 높임으로써 오 검출을 줄일 수 있는 방법이다. 가변형 동적 임계값 방법에서는 특히 컷간의 간격과 컷 프레임에서의 특징값 차이들을 이용해 다음 컷을 검출하는데 필요한 인자들을 자동으로 생성해 냄으로써 컷 발생 직후에 높아지는 임계값의 정도와 시간이 지남에 따라 감소 함수에 의해 임계값이 낮아지는 정도를 효율적으로 유지할 수 있도록 하였다. 또한 컷이 발생하는 지점에서의 프레임 변화를 임계값 설정에 반영할 수 있도록 하여 지역적인 인접 프레임 변화에 의해 오 검출되는 상황을 최소화 할 수 있도록 하였다. 구현과 실험을 통하여 제안하는 가변형 동적 임계값 기법은 기존의 고정 임계값을 사용할 때 보다 오 검출을 줄이면서 효율적으로 컷을 검출해 낼 수 있음을 확인 할 수 있었으며 가변적으로 설정되는 인자 값들이 동영상상이 진행됨에 따라 최적의 값에 근접함을 알 수 있었다.

앞으로 인접프레임 비교와 컷으로 판정되는 프레임과의 비교 등의 변화를 가변형 동적 임계값 기법에 적용하여 점진적인 전환의 유무를 검출하는 연

구와 검출되는 컷들의 가중치를 고려한 최적 컷 선출 기법들에 대한 연구가 필요하다고 본다.

참 고 문 헌

[1] John S. Boreczky, Lawrence A. Rowe, "Comparision of video shot boundary detection techniques", in Storage & Retrieval for Image and Video Database IV, Proc. of SPIE 2670, pp. 170-179, 1996.

[2] Akio Nagasaka, Yuzuru Tanaka, "Automatic Video indexing and full video search for object appearance", 2nd Working Conference on Visual Database Systems, Oct., 1991.

[3] Hirotada Ueda, Takafumi Miyatake, Satoshi Toshizawa, "IMPACT: An Interactive natural motion picture dedicated Multimedia Authoring system", CHI'91 Conference Proceedings, pp. 343-350, 1991.

[4] Jurgen Stauder, Henri Nicolas, "Motion-Based Video Indexing Evaluating Object Shading", In Proceedings of ICIP, Kobe, October 1999.

[5] W.J. Heng, K. N. Ngam, "The Implementation of Object-Based Shot Boundary Detection using Edge Tracing and Tracking", IEEE International Conference on Image Processing (ICIP '99), Japan, 1999.

[6] Shahraray. B., Scene change detection and content-based sampling of video sequences, in Digital Video Compression : Algorithms and Technologies, Proc. SPIE 2419, pp 2-13, February, 1995.

[7] Jianhao Meng, Yujen Juan, Shih-Fu Chang, Scene Change Detection in a MPEG Compressed Video Sequence, IS&T/SPIE Symposium Proceedings, Vol. 2419, Feb., 1995.

[8] Xinying Wang, Zhengke Weng, Scene Abrupt Change Detection, Electrical and Computer Engineering, 2000 Canadian Conference on, Vol. 2 , pp 880-883, 2000.

[9] Lu Tong, P.N. Suganthan, An Adaptive Cumulation Algorithm for Video Shot Detection, Proceedings of 2001 International Symposium on Intelligent Multimedia, Video

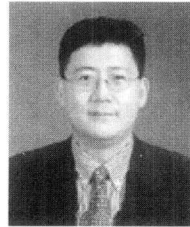
and Speech Processing, pp 296-299, May, 2001.

[10] 윤인구, 김우생, 동적 임계값을 이용한 컷 검출, 한국정보처리학회 논문집, 7월, 1999.

[11] Makoto Miyahara, Yasuhiro Yoshida, "Mathematical transform of (R, G, B) color data to Munsell (H, V, C) color data," SPIE Visual Communication and Image Processing '88, Vol.1001, pp.650-657, 1988.

염 성 주(Sung-Ju Youm)

정회원

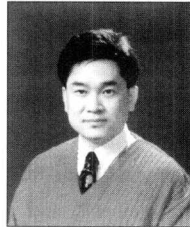


1991년 : 광운대학교 전자계산학과 졸업 (학사)
 1993년 : 광운대학교 대학원 전자계산학과(이학석사)
 1993년~현재 : 광운대학교 컴퓨터 공학부 박사수로 경북 대학 인터넷 정보과 조교수

<주관심 분야> 멀티미디어 시스템, 내용기반 검색, 비디오 분석, 이미지 프로세싱

김 우 생(Woo-saeng Kim)

정회원



1982년 : 서울대학교 수료
 1985년 : University of Texas at Austin (학사)
 1987년 : University of Minnesota (이학석사)
 1987년~1988년 : 현대전자. Zeus Computer 과장

1991년 : University of Minnesota(이학박사)
 2002년 : UC 버클리 대학 교환 교수
 1992년~현재 : 광운대학교 컴퓨터공학부 교수
 <주관심 분야> 멀티미디어, 영상/비디오 처리, 데이터베이스