

동적 임계값과 컷 프레임 차를 이용한 점진적 전환 검출 기법

염 성 주[†] · 김 우 생^{††}

요 약

내용기반 검색을 위한 비디오 데이터 장면전환 검출에서 점진적인 전환을 검출하는 것은 갑작스런 전환을 검출하는 것에 비해 일반적으로 어려운 문제로 알려져 있다. 본 논문에서는 가변형 동적 임계값과 가장 최근에 검출된 컷 프레임과 현재 프레임간의 특징값 차이인 컷 프레임 차를 이용하여 갑작스런 전환과 점진적인 전환을 찾아내는 기법을 제안한다. 이를 위하여 본 논문에서는 점진적인 전환이 갖는 특성과 수학적 모델을 제시하고 컷 프레임 차를 이용하여 점진적인 전환을 검출할 수 있음을 보인다. 그리고 이를 바탕으로 갑작스런 전환과 점진적인 전환을 함께 검출할 수 있는 방법을 제시한다. 실제로 동영상 데이터를 대상으로 한 실험을 통해 제안하는 기법이 점진적인 전환 효과의 종류에 종속적이지 않으며 적은 연산 비용으로 쉽게 점진적인 전환 유무를 검출할 수 있음을 보인다.

Gradual scene change detection using Cut frame difference and Dynamic threshold

Sungju Youm[†] · WooSaeng Kim^{††}

ABSTRACT

Gradual scene change detection is known as more difficult problem than abrupt scene change detection on video data analysis for contents based retrieval. In this paper, we present a new method for scene change detection for both abrupt and gradual change using the variable dynamic threshold and cut frame difference (CFD). For this, we present the characteristics and mathematical models of gradual transitions and then, how can be detected by the CFD. And also we present new scene change detection algorithm based on cut frame difference. By the experimental result using real world video data indicate that the proposed method detect various scene changes both abrupt and gradual change efficiently without time-consuming computation and any dependency on a kind of gradual change effects.

키워드 : 비디오 분석(Video Analysis), 장면전환검출(Scene Change detection), 내용기반검색(Content-Based Search)

1. 서 론

장면전환 지점 검출(scene change detection)은 동영상 데이터의 내용기반 검색을 위해 필요한 기초작업으로써 스트림(stream) 형태의 동영상 데이터로부터 장면들을 구분하기 위해 컷(cut)이라 불리는 전환 지점들을 자동으로 찾아내는 작업을 말하며 이를 바탕으로 동영상 데이터에 대한 색인과 검색이 이루어지게 되므로 효과적인 검출 방안이 반드시 필요하다.

지금까지 장면전환 지점 검출을 위한 기존의 연구에서는 동영상 데이터의 압축/비압축 도메인을 대상으로 매우 다양한 방법들이 제안되고 있다[1]. 대부분의 연구에서는 동

영상 데이터의 각 프레임을 대표할 수 있는 특징값(feature)을 선정하여 인접 프레임간의 특징값의 차이를 비교하는 형식을 취하고 있다. 가장 일반적으로 널리 알려진 방법은 그레이레벨(gray-level) 히스토그램을 사용하는 방법[2]으로 인접 프레임간의 히스토그램 차이를 비교하여 정해진 임계값을 초과하는 경우 이를 컷으로 검출하는 방법이다. 히스토그램 기반의 방법들은 구현이 쉬우며 적은 연산 비용으로 비교적 효과적인 장점을 갖는 반면 명암의 변화에 민감하기 때문에 컷이 아님에도 이를 컷으로 검출해내는 오 검출(fault cut)이 자주 발생하는 문제점을 갖는다. 이러한 문제점을 보완하기 위해 움직임 벡터(motion vector)[3], 에지 정보[4, 5], 화소 값들간의 차이[6]등과 같이 전혀 다른 특징들을 비교대상으로 삼는 다양한 방법들이 제안된 바 있다. 그러나 이러한 기법들은 사용함으로써 카메라의 시점이 변하는 갑작스런 전환(abrupt change)을 비교적 쉽게 검출 할

* 본 논문은 광운대학교 2002년도 교내연구비에 의해 지원 되었음.

† 정 회 원 : 광운대학교 대학원 컴퓨터학과

†† 종신회원 : 광운대학교 컴퓨터공학부 교수

논문접수 : 2002년 1월 25일, 심사완료 : 2002년 5월 31일

수 있지만 여러 가지 영상 처리효과를 통해 한 장면에서 다른 장면으로 서서히 변화하는 점진적 전환(gradual change)을 검출하는 방안에 대한 연구는 아직 미흡한 실정이다.

점진적인 전환을 검출하기 위한 기존의 연구로 가장 널리 알려진 이중 비교(twin comparison)기법은 서로 다른 두개의 임계값을 써서 점진적 전환을 검출하는 방법이다[7]. 그러나 이 방법은 두개의 임계값을 효과적으로 선정하기가 매우 어렵고 실세계의 다양한 상황에 대해 잘 대처하지 못하는 문제점을 갖는다. [8-10]등의 연구에서는 동영상 데이터의 특정 부분을 시공간 이미지(spatio-temporal image)로 재구성하여 점진적인 전환이 일어날 때 나타나는 패턴을 인식하는 방법을 통해 점진적인 전환을 검출하는 방안을 제안했다. [11]의 연구에서는 각 프레임들로부터 추출되는 특징값들을 통계적인 분석을 통해 재구성하여 점진적인 전환이 일어나는 경우의 나타날 수 있는 변화를 검출하는 방법을 통해 점진적인 전환을 검출하는 방안을 제안했다. 이 밖에도 블록 비교와 움직임 패턴 추적을 통한 와이프 검출 방법[12, 13], 영상의 대조(contrast)를 특징으로 사용하는 방법[14], 변화된 부분만을 기록하는 축적 프레임을 사용한 방법[15] 그리고 에지 객체를 기반으로 한 방법[16] 등이 제안된 바 있다. 그러나 이러한 연구들은 대부분 점진적인 전환이 일어나는 상황 그 자체를 인식하여 검출하려고 시도하기 때문에 전환 효과에 대해 종속적이며 연산비용이 매우 높은 반면 실세계의 동영상 데이터에서 나타나는 다양한 형태의 전환을 검출하기에는 아직 미흡한 면이 많은 실정이다.

갑작스런 전환과 점진적인 전환은 일어나는 형태와 전환에 걸리는 시간만 다를 뿐 둘 다 영상제작자에 의해 두개의 서로 다른 장면을 전환시키는 효과라는 공통점이 있다. 따라서 어떠한 형태의 점진적인 전환이라 할지라도 전환이 일어나기 전과 후의 장면에 속한 프레임들은 갑작스런 장면전환의 경우와 마찬가지로 서로 다르다. 본 논문에서는 이러한 특징에 착안하여 하나의 임계값과 컷 프레임차(cut frame difference)를 사용하여 점진적인 전환의 유무를 쉽게 검출할 수 있는 새로운 기법은 제안한다. 컷 프레임차란 가장 최근에 갑작스런 컷으로 검출된 프레임과 그 이후 프레임들간의 차이를 말하며 점진적이 전환이 일어나는 경우 거의 변화가 없는 인접 프레임 차와는 다르게 컷 프레임차는 매우 높게 나타나는 특성을 갖는다. 제안 하는 기법은 먼저 인접 프레임간의 특징값 차이와 가변형 동적 임계값을 사용해 갑작스런 컷을 먼저 검출하고 검출된 컷과 컷 사이에 존재할 수 있는 점진적인 전환을 유무를 컷 프레임차와 임계값을 비교하여 찾아내는 방식이다. 제안하는 기법은 점진적인 전환을 찾아내기 위한 부가적인 연산이 매우 적고 갑작스런 전환을 검출하기 위한 하나의 임계값을 그대로 사용하며 점진적인 전환의 종류에 종속적이지 않는

장점을 갖는다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 점진적인 전환의 특성과 모델을 소개하고 3장에서는 갑작스런 전환과 점진적인 전환을 가변형 동적 임계값과 컷 프레임차를 이용해 검출하는 방안과 알고리즘을 기술한다. 4장에서는 제안하는 기법을 실세계 동영상 데이터를 대상으로 실험한 결과를 보이며 마지막으로 5장에서 결론과 향후 과제를 제시한다.

2. 점진적 전환의 특성과 모델링

이번 장에서는 점진적인 전환의 일반적인 특성을 소개하고 점진적인 전환의 가장 일반적인 패턴인 디졸브 페이드인/아웃 그리고 와이프 등의 수학적 모델을 정의한다.

2.1 점진적 전환의 특성

점진적인 전환은 한 장면에서 다른 장면으로 다양한 효과를 사용하여 서서히 변하는 것으로 인접 프레임간의 특징 비교만을 가지고는 이를 발견해 내기가 대단히 어렵다. 일반적으로 점진적인 전환은 약 30프레임 정도에 걸쳐 일어나는 것으로 알려져 있으며 상황에 따라 이보다 적은 프레임수로 빠르게 진행 되기도 한다. 또한 점진적인 전환의 특성상 전환이 일어나는 동안에는 움직임이 많게 되므로 전환이 시작되기 전이나 후에 대부분 비교적 정적인 장면들이 위치하게 된다. 따라서 본 논문에서는 점진적인 전환이 일어난 전후의 장면들에서는 또 다른 점진적인 전환이 일어나지 않는다고 가정한다.

점진적인 전환 효과의 종류로는 빈 화면에서 서서히 장면이 삽입되는 페이드인(fade-in), 반대로 빈 화면으로 서서히 사라지는 페이드아웃(fade-out), 두 장면이 천천히 오버랩(overlap)되면서 바뀌는 디졸브(dissolve), 한 장면이 다른 장면을 밀어내는 식인 와이프(wipe), 카메라의 시점이 일정한 속도로 이동하며 발생하는 패닝(panning)과 줌인/아웃(zoom-in/out)등이 있다. 이 밖에도 아이리스(iris), 모핑(morphing) 그리고 최근 시도되는 다양한 3D그래픽 효과 등도 서로 다른 장면을 연결하는 점진적인 전환의 한 형태로 볼 수 있다.

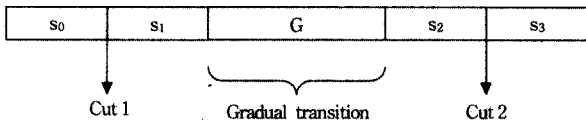
본 논문에서는 점진적인 전환 효과들을 크게 디졸브 계열과 패닝 계열로 분류하여 다루기로 한다. 디졸브 계열은 이미지 처리(image processing)효과를 사용한 전환으로써 디졸브, 페이드인/아웃, 와이프 등이 여기 속한다. 패닝 계열은 카메라의 이동에 의해서 발생하는 전환으로 패닝과 줌인/아웃 등이 여기 속한다. 디졸브 계열 효과들은 전환 효과가 일어나는 동안 서로 다른 장면의 일부가 동시에 화면에 나타나는 특성을 갖는다. 또한 대체적으로 갑작스런 전환의 단조로움을 피하기 위해서 영상 제작자에 의해 의

도적으로 삽입된 것들이므로 전환이 일어나는데 일정 시간 동안이 필요하다는 점을 제외하면 갑작스러운 전환 효과와 거의 같은 특성을 갖게 된다.

패닝 계열 효과들은 디졸브 계열과는 약간 다른 특성을 갖는다. 왜냐하면 이들은 이미지 처리 효과를 사용하는 것이 아닌 카메라의 이동에 의해 발생하기 때문이다. 따라서 카메라가 어떤 의도를 가지고 움직여 졌는가에 따라 그 의미가 달라진다. 카메라가 빠른 속도로 이동을 하는 경우는 대체적으로 컷과 같은 장면 전환의 일환으로 볼 수 있다. 예를 들어 생방송 뉴스 현장 카메라의 경우 인위적인 영상 편집이 불가능 하기 때문에 카메라의 시점을 빠르게 전환 하는 것은 갑작스러운 전환과 비슷한 의미를 갖게 된다. 이와는 다르게 경치를 파노라마 식으로 보여준다든지 할 때에는 장면 전환으로써의 의미가 없다고 할 수 있겠다. 특별히 본 논문에서는 장면전환의 의미로써 사용된 패닝 계열의 효과들의 경우 점진적인 전환의 검출 대상으로 삼는다.

2.2 점진적 장면 전환 모델

점진적인 전환 모델링 하기 위하여 먼저 동영상 데이터에서 나타나는 점진적인 전환의 일반적인 형태를 살펴보면 (그림 1) 과 같다.



(그림 1) 점진적인 전환의 예

(그림 1)은 디졸브나 페이드 같은 점진적 장면이 포함된 경우를 보여준다. 구간 S_0, S_1, S_2, S_3 는 자기 다른 일반적인 장면이며 구간 G 는 점진적인 전환이 일어나는 구간 즉, '점진적인 장면'이라고 가정한다. 페이드 같은 점진적 장면 전환의 경우는 전환 구간 전후의 장면들에 따라 일반 검출 기법으로 검출될 수 있는지가 결정된다. 본 연구에서는 구간 G 가 일반적인 컷 검출 기법으로는 검출 할 수 없는 경우만을 고려하기 때문에 (그림 1)의 비디오에 일반적인 컷 검출 방법을 적용하면 Cut_1 과 Cut_2 지점에서 컷을 검출 하면서 3 개의 장면으로 구분하게 된다. 그러나 (그림 1)의 두 번째 장면 S_2 에서 점진적 전환 구간 G 를 제외 하면 S_1 과 S_2 는 서로 다른 장면이기 때문에 갑작스러운 전환과 유사한 특성을 갖게 된다. 즉 S_2 에 속하는 임의의 프레임과 S_1 에 속하는 임의의 프레임을 비교하면 기존의 갑작스러운 장면 전환 여부를 알아내는 방법으로도 장면이 전환 되었음을 알아낼 수 있다. 따라서 본 연구에서는 현재 프레임을 전 프레임과 비교 하여 갑작스런 컷을 검출 할 뿐 아니라

현재 프레임을 가장 최근에 검출된 컷 프레임과도 비교를 하여 점진적인 장면 전환도 함께 검출하고자 한다.

[11]의 연구를 바탕으로 점진적인 장면 전환에 대한 수학적인 모델을 정의 하기 위해 구간 S_1 에 있는 프레임을 $p(x, y, t)$ 구간 S_2 에 있는 프레임을 $q(x, y, t)$ 라고 가정할 때, 디졸브는 식 (1)과 같이 표현 할 수 있다. 이때 $\alpha(t)$ 는 감소 함수이고 $\alpha(0)=1$ 과 $\alpha(T)=0$ 의 값을 갖으며, $\beta(t)$ 는 증가 함수로써 각각 $\beta(0)=0$ 과 $\beta(T)=1$ 의 값을 갖는다. 또한 디졸브의 전환 조건으로는 기존 연구에서 많이 사용하는 piece-wise linear function의 일반화된 $\alpha(t)+\beta(t)=1$ 의 조건을 가정한다.

$$f(x, y, t) = \alpha(t)p(x, y, t) + \beta(t)q(x, y, t) \quad (0 \leq t \leq T) \quad (1)$$

페이드인과 페이드아웃은 디졸브의 특수한 경우이기 때문에 C 를 단일색 프레임이라고 정의 할 때, 식 (1)에서 $p(x, y, t)=C$ 로 하면 식 (2)와 같은 페이드인이 되며,

$$f(x, y, t) = \alpha(t)C + \beta(t)q(x, y, t) \quad (0 \leq t \leq T) \quad (2)$$

식 (1)에서 $p(x, y, t)=C$ 로 하면 식 (3)과 같은 페이드아웃이 된다.

$$f(x, y, t) = \alpha(t)p(x, y, t) + \beta(t)C \quad (0 \leq t \leq T) \quad (3)$$

반면 와이프는 M 을 원소가 1이나 0으로 구성된 와이프 패턴을 표현하는 행렬이라고 가정하고 \otimes 를 행렬간의 원소 대 원소의 곱이라고 가정할 때 식 (4)처럼 정의할 수 있다.

$$f(x, y, t) = M(t) \otimes p(x, y, t) + \overline{M(t)} \otimes q(x, y, t) \quad (0 \leq t \leq T) \quad (4)$$

본 논문에서는 편의상 프레임 $f(x, y, t)$ 를 f_t 로, 프레임 $f(x, y, t)$ 의 특징값인 $V(f(x, y, t))$ 를 $V(f_t)$ 로 간단히 표현한다. 또한 (그림 1)의 구간 S_2 의 시작 프레임을 q_0 라고 가정하며, 구간 S_1 의 시작 프레임은 p_0 이며 동시에 컷 프레임이라고 가정한다.

3. 동적 임계값과 컷 프레임 차를 이용한 점진적인 전환 검출 기법

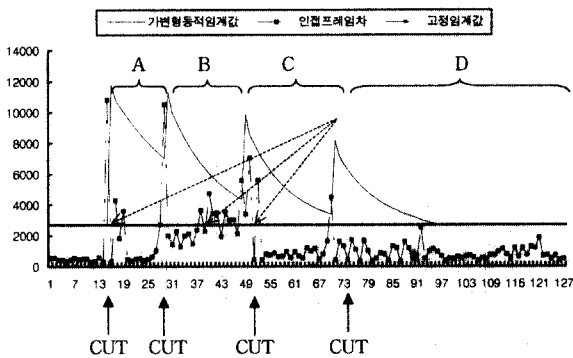
본 논문에서 제안하는 점진적인 전환 검출 기법은 먼저 갑작스런 컷을 검출하고 컷과 컷 사이에 점진적인 전환이 있었는가를 점검하는 방식으로 진행된다. 이를 위해 이번 장에서는 가변형 동적 임계값에 의해 갑작스런 컷을 검출하는 기법을 먼저 소개하고 이를 통해 점진적인 전환을 검출하는 기법을 기술한다.

3.1 가변형 동적 임계값을 이용한 컷 검출

임계값을 사용하는 컷 검출 기법들은 일반적으로 임계값을 어떻게 설정하느냐에 따라 효율에 큰 차이가 난다. 예를 들어 임계값을 높게 설정하면 컷임에도 불구하고 이를 놓치게 되는 미검출(miss cut) 오류가 발생하고 반대로 낮게 설정하면 미 검출은 줄어들지만 대신 컷이 아님에도 불구하고 컷으로 검출해내는 오검출(fault cut)이 많아지는 문제점을 갖고 있다. 동적 임계값 기법은 이러한 문제점을 해결하기 위해 컷이 발생하는 경우 임계값을 높이고 일정 시간 간격 동안 컷이 발생 하지 않으면 임계값을 서서히 낮추어 컷 검출 직후 발생하는 오 검출을 크게 줄일 수 있는 기법이다[17]. 특히 가변형 동적 임계값 기법은 적절한 임계값의 자동적인 변화를 위해 컷이 발생하는 시간 간격에 따른 분포와 런 타임시 발생하는 정보들을 다음 컷 검출을 위한 설정에 반영하는 방법으로 최적의 임계값을 유지할 수 있도록 개선한 기법이다. 가변형 동적 임계값 T_{dyn} 는 식 (5)와 같이 표현된다. 여기서 T_{low} 는 동영상이 진행되는 시간 t 의 함수 $f(t)$ 에 의해 감소되는 임계값의 최하점을 의미하고 T_{high} 는 컷이 검출된 직후 증가되는 임계값으로 증가되는 양은 그때까지 실시간으로 분석된 동영상의 정보를 사용하여 가변적으로 정해진다.

$$T_{dyn} = T_{low} + (T_{high} - T_{low})f(t) \quad (\text{단, } T_{high} \geq T_{low}) \quad (5)$$

(그림 2)는 함수 $f(t)$ 를 지수 감소함수로 사용했을 경우의 가변형 동적 임계값 기법의 수행 예를 보여준다.



(그림 2) 가변형 동적 임계값을 통한 급진적 장면 검출 기법

일반적인 고정 임계값을 사용하는 경우 (그림 2)의 구간 A, B, C에서 인접 프레임간의 특징값 차이가 연속해서 임계값을 넘기 때문에 많은 오검출이 발생하게 된다. 그러나 가변형 동적 임계값의 경우 컷이 발생한 직후에 임계값이 높아지기 때문에 이러한 오류를 발생하지 않으며 시간이 지남에 따라 임계값을 서서히 낮추며 컷을 검출하고 구간 D와 같이 일정 시간 내에 컷이 검출되지 않는 경우 고정 임계값과 동일하게 진행 된다.

3.2 컷 프레임 차를 사용한 점진적 장면 전환 검출 방법

본 논문에서는 현재 프레임과 가장 최근에 검출된 컷 프레임과의 비교를 통해 점진적인 전환이 일어날 경우에도 특징값의 변화가 감지되어 점진적 장면 전환 구간을 검출하는 방법을 제안한다. 특정 구간 G 에 속한 임의의 프레임을 f_i 라 하고 전에 발생한 컷들 중 가장 최근의 컷 프레임을 $f_j(i < j)$ 라 할 때 구간 A 에 속한 프레임과 가장 최근의 컷 프레임 f_i 간의 특징값 비교인 CFD (cut frame difference)는 식 (6)와 같이 정의된다.

$$CFD(A) = |V(f_i) - V(f_j)| \quad (6)$$

구간 S_1 에서의 CFD 는 식 (7)과 같이 표현되며, p_0 와 p_i 는 같은 장면에 속한 프레임들이기 때문에 δ_1 은 0에 가까운 값이 된다.

$$CFD(S_1) = |V(p_0) - V(p_i)| = \delta_1 \approx 0 \quad (7)$$

반면 구간 S_2 에서의 CFD 는 식 (8)과 같이 표현되며, p_0 와 q_i 는 서로 다른 장면의 프레임이기 때문에 δ_2 는 큰 상수값이 된다.

$$CFD(S_2) = |V(p_0) - V(q_i)| = \delta_2 \quad (8)$$

구간 G 가 디졸브의 경우 CFD 는 식 (9)과 같이 표현이 되고, $\beta(t)$ 는 증가함수 이기 때문에 CFD 값은 처음에는 0에 가까운 값을 유지하다가 시간이 지나면서 비율 β 로 증가하다 최종적으로는 큰 상수 값을 유지하게 된다.

$$\begin{aligned} CFD(dissolve) &= |V(p_0) - \alpha(t)V(p_i) - \beta(t)V(q_i)| \\ &\approx |(1 - \alpha(t))V(p_0) - \beta(t)V(q_i)| \quad (9) \\ &= \beta(t)|V(p_0) - V(q_i)| \end{aligned}$$

본 논문에서 가정하는 페이드 인의 경우는 페이드 시작전의 다른 장면들의 특징값이 페이드인의 시작하는 단색의 특징값과 비슷한 경우이기 때문에 $V(C) \approx V(p_T) \approx V(p_0)$ 의 조건에 의해 구간 G 가 페이드인의 경우는 다음과 같이 표현된다.

$$\begin{aligned} CFD(fade-in) &= |V(p_0) - \alpha(t)V(C) - \beta(t)V(q_i)| \\ &\approx |(1 - \alpha(t))V(p_0) - \beta(t)V(q_i)| \quad (10) \\ &= \beta(t)|V(p_0) - V(q_i)| \end{aligned}$$

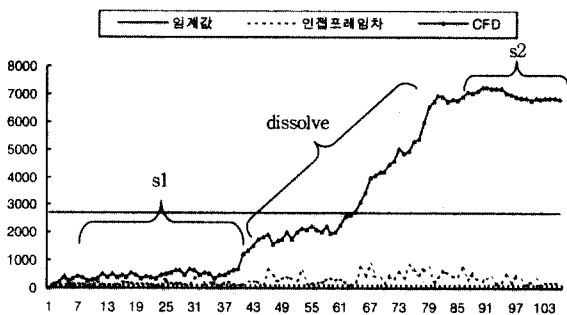
또한 구간 G 가 페이드아웃의 경우는 페이드아웃의 끝나가는 단색의 특징값이 다음 장면들의 특징값과 비슷한 경우이기 때문에 $V(C) \approx V(q_0)$ 조건에 의해 구간 G 가 페이드아웃의 경우는 식 (11)과 같이 표현된다.

$$\begin{aligned}
 CFD(fade-out) &= |V(p_0) - \alpha(t)V(p_t) - \beta(t)V(C)| \\
 &\approx |(1 - \alpha(t))V(p_0) - \beta(t)V(q_0)| \quad (11) \\
 &= \beta(t)|(V(p_0) - V(q_0))|
 \end{aligned}$$

와이프의 장면들의 경우는 전 장면의 일정한 패턴의 픽셀들은 점차 제거되며 그 위치에 다음 장면의 픽셀들이 나타나게 된다. 따라서 p_t^w 와 $p_t^{\bar{w}}$ 를 전장면 프레임 q_t 와 다음 장면 프레임 q_t 가 이러한 패턴에 의해서 변화된 프레임이라고 가정할 때 구간 G 가 와이프의 경우는 다음과 같이 표현된다.

$$\begin{aligned}
 CFD(wipe) &= |V(p_0) - V(M(t) \otimes p_t) - V(\overline{M(t)} \otimes q_t)| \\
 &= |V(p_0) - V(p_t^w) - V(q_t^{\bar{w}})| \\
 &\approx |V(q_0) - V(\alpha(t)p_t) - V(\beta(t)q_t)| \\
 &\approx |(1 - \alpha(t))V(p_0) - \beta(t)V(q_t)| \\
 &= \beta(t)|(V(p_0) - V(q_t))| \quad (12)
 \end{aligned}$$

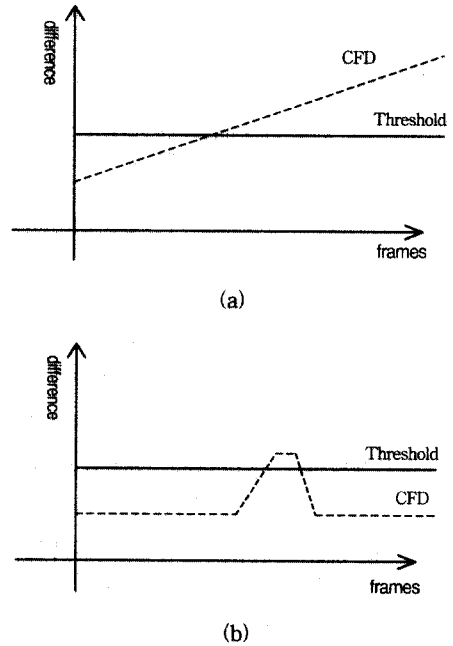
따라서 이들 점진적인 전환 구간은 모두 처음에는 작은 CFD값을 유지하다가 시간이 지나면서 CFD값이 점점 증가하다가 최종적으로는 큰 CFD값을 갖게 된다. (그림 3)는 실제 비디오 데이터에서 두개의 장면이 디졸브에 의해서 전환되는 상황이고 이때의 인접 프레임 차와 CFD 변화를 보여준다. 이 동영상의 전체 프레임 수는 110이고 그 중 46 프레임에서 76프레임에 걸쳐서 디졸브가 일어났다. (그림 3)에서 볼 수 있듯이 인접 프레임차는 거의 없어 이것만을 가지고는 점진적인 장면 전환을 판단하기 불가능하다. 그러나 CFD의 변화는 임계값 보다 높은 구간과 낮은 구간으로 확연히 구분될 수 있음을 알 수 있다. 따라서 CFD가 주어진 어떤 임계값을 초과하는 경우 적어도 점진적인 전환이 일어났을 가능성이 존재한다는 것을 알려준다. 또한 (그림 3)의 CFD 그래프 형태는 식 (8)과 일치하는 것을 알 수 있다.



(그림 3) 디졸브 시의 인접프레임 차와 CFD의 변화 비교

CFD의 변화가 임계값을 초과하는 상황은 점진적인 변화가 일어날 때 이외에도 몇 가지 다른 경우가 존재할 수 있다. 가장 대표적인 경우는 패닝과 같은 카메라 움직임의 효

과가 있는 경우이다. 예를 들어 어떤 물체를 아주 천천히 줌인(zoom-in)하는 장면이 있다고 가정하면 장면의 시작 프레임부터 끝까지 일정한 양으로 계속 변하게 되므로 CFD의 변화는 임계값을 넘게 된다. 또 한가지 경우는 한 장면 내에서 물체가 갑자기 움직일 때 나타나는 현상이다. 예를 들어 아나운서가 몸을 갑자기 숙였다가 일으킨다든지 하면 첫 프레임과의 차이가 몸을 숙이는 부분에서 갑자기 커져 CFD의 변화는 임계값을 넘길 가능성이 있다. 그러나 이러한 경우들은 점진적 장면이 일어났을 때의 CFD 변화 형태와 다르게 나타난다. (그림 4)의 (a)는 장면전환으로써의 의미를 갖지 않는 일반적인 패닝 계열의 효과가 일어날 때의 CFD 변화를 보여준다. 이러한 효과일 경우에는 일정한 양으로 계속해서 카메라의 위치나 시점이 변화하게 되므로 첫 프레임으로부터 일정한 비율로 차이가 증가하게 되고 CFD가 임계값을 넘는 지점을 중심으로 전/후의 CFD 변화는 지속적인 증가 또는 감소의 형태를 보여준다. (b)는 갑작스러운 움직임에 의한 변화를 보여준다. 이런 경우 CFD의 변화는 객체의 변화가 일어난 부분에서 순간적으로 임계값을 초과하게 되나 곧 원래의 값으로 돌아오게 된다.



(그림 4) 카메라 움직임 등의 CFD의 변화 형태

따라서 본 논문에서는 CFD를 통하여 점진적인 전환 구간을 찾기 위해 CFD와 임계값이 만나는 지점을 전후로 CFD값이 어떠한 형태로 변화 하는지를 파악한다. 연속된 구간 S_1 , G , S_2 에서 CFD의 변화가 다음의 조건을 만족할 경우 점진적인 전환 구간으로 판단한다.

$$\begin{aligned}
 CFD(S_1) = \delta_1 < \theta \text{ and } \delta_1 \leq CFD(G) \leq \delta_2 \\
 \text{and } CFD(S_2) = \delta_2 > \theta \quad (13)
 \end{aligned}$$

3.3 장면전환 검출 알고리즘

인접 프레임 비교와는 달리 CFD의 변화는 전환이 일어나는 구간과 그렇지 않은 구간에서 차이가 극명하게 나타난다. 이러한 차이는 동일한 내용의 장면이 진행되는 경우에는 특정한 값으로 계속 유지되고 전환등의 변화가 발생하면 일정한 비율로 내려가거나 또는 올라가게 된다. 따라서 본 논문에서는 점진적인 전환이 일어난 가능성이 있는 장면에서 시간 t에 대한 CFD의 미분을 통해 변화 형태를 파악 한다. 먼저 구간 S₁이 시간 t₀부터 t₁까지 진행되고 구간 S₂가 시간 t₂부터 t₃까지 진행한다고 가정하면 점진 구간 G는 시간 t₁부터 t₂까지로 표현할 수 있다 (t₀ < t₁ < t₂ < t₃). 시간 t를 점진구간 G가 일어나는 동안 즉, t₁ < t < t₂ 이라 할 때 이 구간에서의 CFD측정값을 μ(t)라 표현하면 이에 대한 변화율은 μ'(t) = $\frac{d\mu(t)}{dt}$ 로 표현할 수 있다. 디졸브와 와이프 그리고 페이드 인/아웃 등이 일어나는 점진 구간 G에서 CFD의 변화는 3.2절의 정의에 따라 식 (14)로 표현되고

$$CFD(G) = \beta(t) |V(p_0) - V(q_t)| \quad (14)$$

점진 구간 G에서 μ'(t)는 다음과 같이 정리 할 수 있다.

$$\begin{aligned} \mu'(t) &= \frac{d(\beta(t) |V(p_0) - V(q_t)|)}{dt} \\ &= \delta_2 \frac{\beta(t_2) - \beta(t_1)}{t_2 - t_1} = \delta_2 \cdot \beta'(t) \end{aligned} \quad (15)$$

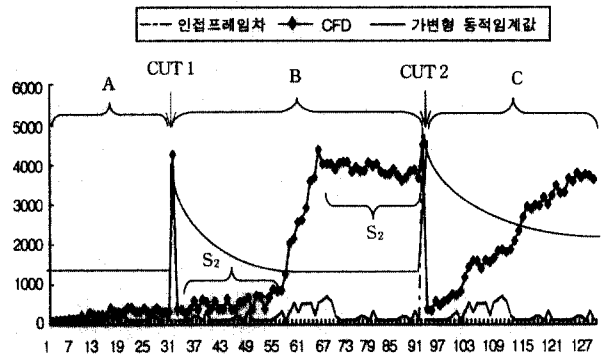
그러므로 식 (15)에 의해 점진 구간 G에서 CFD 미분은 증가 함수 β(t)의 미분 β'(t)에 비례하게 되고 점진구간 전과 후의 구간 즉, S₁과 S₂에서 CFD 변화는 상수 값을 갖게 되므로 μ'(t)는 0이 된다는 사실을 알 수 있다. 따라서 본 논문에서 제안하는 알고리즘은 점진적인 장면전환이 포함된 가능성이 있는 장면의 시작과 끝에서 S₁과 S₂구간이 존재 하는가를 찾아내는 방법을 통해 점진적인 전환 효과의 종류에 관계없이 전환이 일어난 경우 각 장면들을 검출하는 방식을 취한다. 전체적인 알고리즘은 다음과 같다

- 1) 프레임 카운터 fc와 컷 카운터 cc를 0으로 초기화 한다.
- 2) fc를 증가시키며 가변형 동적 임계값에 의해 갑작스런 장면전환을 찾는다.
- 3) fc가 마지막 프레임이면 종료하고 아니면 4)로 진행한다.
- 4) 컷이 발견되면 5)로 진행하고 아니면 2)로 진행한다.
- 5) 프레임 cc에서 부터 fc-1 동안 CFD가 동적 임계값을 초과했다면 6)로 진행, 그렇지 않다면 cc=fc로 하고

2)로 진행

6) CFD가 식 (13)을 만족하면 S₁과 S₂를 서로 다른 장면으로 구분한 후 cc=fc로 하고 2)로 진행 한다.

특별히 S₁과 S₂ 구간을 찾기 위해 본 논문에서는 장면 S의 첫 프레임과 마지막 프레임에서 서로 역 방향으로 CFD를 조사하며 각각 μ(t) ≠ 0이 되는 지점까지의 구간을 S₁, S₂로 정한다. 단 실제 동영상의 경우 동일 장면 내라 할지라도 프레임간의 특징값 차이는 완전히 0이 되지 않으므로 적절한 임계값 σ로 대신하고 (σ > 0) S₁과 S₂는 사용자가 장면으로 인식할 수 있는 최소의 시간 T이상 유지되어야만 장면으로 인정 한다. 이때 σ는 실제 동영상 데이터에서 장면전환이 일어나지 않는 구간의 특징값 변화 범위를 측정함으로써 얻어지는 경험적인 값이다. (그림 5)는 실제계의 동영상 데이터를 대상으로 한 실험에서 본 논문이 제시한 알고리즘을 통해 급진적인 전환과 점진적인 전환을 찾는 예를 보여준다.



(그림 5) 알고리즘 수행의 예

(그림 5)에서 A구간은 장면전환이 없는 구간이고 CUT1에 의해 장면 B로 전환된다. 동적 임계값은 CUT1에 의해 증가 되었다가 시간이 지남에 따라 감소하게 된다. 그러나 구간 B에서 인접 프레임의 변화가 크지 않으므로 다음 컷인 CUT2가 나올때 까지 별다른 장면전환을 검출하지 못하게 된다. CUT2가 발견되면 CUT1과 CUT2사이의 구간 B에 대한 CFD를 조사 한다. CFD는 61프레임 부근에서 동적 임계값과 만나게 되고 이 지점을 θ로 삼는다. 이때 CFD는 θ 이하의 구간에서는 동적 임계값 보다 작고 θ이상의 구간에서는 동적 임계값 보다 큰 형태를 보이므로 식 (12)를 만족하여 점진적인 전환일 가능성이 있다. 이제 CUT1에서부터 θ로 프레임 카운터를 증가시키며 μ'(t) < σ를 만족하는 프레임들을 S₁으로 한다. 같은 방식으로 CUT2로부터 프레임 카운터를 θ쪽으로 감소시키며 μ'(t) < σ를 만족하는 프레임들을 S₂로 한다. 따라서 구간 B는 식 (13)을 만족하게 되므로 S₁과 S₂를 서로 다른 장면으로 선언한다. 그러나 구간 C의 경우는 θ 전과 후의 부분에서 μ'(t) < σ를 만족할 만한

부분이 없으므로 식 (13)을 만족시키지 못하게 된다. 따라서 구간 C는 점진적인 전환이 아닌 것으로 판단 한다.

4. 구현 및 실험 결과

본 연구에서 제안하는 점진적인 전환 검출 기법을 시험하기 위한 프로토타입은 MS WINDOWS 2000 환경에서 MS Visual C++ 6.0과 Directshow 라이브러리를 통해 MPEG-1 데이터를 대상으로 실험 하였다. 각 데이터는 29 frames/sec의 프레임율과 1152000 bits/sec의 비트율을 가지며 352×240의 해상도를 갖는다. 동영상 분석을 위해 추출되는 각 프레임들은 빠른 처리를 위해 80×80으로 정규화 시키고 [18]에서 제안한 방법에 의해 64-bin의 그레이 이미지로 변환하여 사용했다. 프레임간의 비교대상이 되는 특징값으로는 가장 일반적인 그레이 레벨 히스토그램을 사용했다. 전체 필터 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 (16)과 같이 정의 된다.

$$\sum_{j=0}^n |H_{i-1}(j) - H_i(j)| \quad (16)$$

일반적으로 점진적인 전환이 일어나는 횟수는 갑작스런 전환에 비해 매우 적은 편이다. 따라서 본 논문에서는 특별히 실험을 위하여 점진적인 전환의 대표적인 형태인 디졸브와 와이프 그리고 아이리스(iris) 효과를 인위적으로 삽입한 시험용 동영상(TEST sample)을 제작하여 먼저 실험하였다. 시험용 동영상 데이터는 뉴스 동영상으로부터 2개의 장면을 추출하여 Adobe Premiere 6.0에서 제공하는 디졸브, 와이프, 그리고 아이리스에 해당하는 29개의 전환 효과[19]를 차례로 사용하여 전환시키는 방식으로 제작 되었다.

실세계 데이터는 9시 뉴스와 연속극으로부터 추출했고 광고 동영상의 경우 대부분 길이가 짧은 것을 감안하여 다수의 광고를 연결하여 데이터로 사용했다. 시험용 동영상 데이터와 실세계 동영상 데이터를 대상으로 한 실험결과는 <표 1>과 같다.

본 논문에서 제안하는 점진적인 전환 검출 방법은 검출해야 할 대상이 전환효과 그 자체가 아닌 전환이 일어나기 전과 후의 장면이 되므로 어떠한 형태의 점진적인 전환이

라 해도 전 후의 장면에서 추출되는 특징값의 변화가 안정적이라면 두개의 장면을 확실하게 구별해 내는 장점을 보였다. 실세계 동영상 데이터에 대해서도 제안하는 기법은 점진적인 변화를 검출하는데 대체적으로 효율적이었다. 제대로 검출하지 못한 부분의 동영상을 직접 확인해보면 대부분 빠른 카메라 워크가 동반되거나 모핑과 같은 특수효과와 같이 전환 효과가 연이어 또는 함께 사용되는 경우와 광고의 경우 일부 장면들의 길이가 너무 짧아서 일어나는 오류가 대부분이었다. 오류가 발생하는 대부분의 경우 식 (13)에서 S_2 구간을 검출해 내는데 실패하기 때문에 점진적인 전환이 일어난 것으로 검출하지 못하게 된다. 그러나 해당 구간들을 직접 확인해보면 하나의 장면으로 분류해 내기는 매우 모호한 경우가 대부분으로 제안하는 기법은 전체적으로 고른 효율을 보인다고 할 수 있겠다. 실험결과의 정확도 평가를 위해서 일반적으로 많이 사용하는 recall과 precision을 사용하였으며 이들의 정의는 식 (17)과 같다. 그리고 전체적인 recall과 precision의 비교는 <표 2>와 같다.

$$recall = \frac{detected}{detected + missed}$$

$$precision = \frac{detected}{detected + false\ alarm} \quad (17)$$

<표 2> 전체적인 recall과 precision 평가

Video data	갑작스런 전환 검출		점진적인 전환 검출	
	Recall	Precision	Recall	Precision
TEST sample (10400)	1	1	1	1
CF (1850 frames)	0.90	0.85	0.71	1
Drama (10704 frames)	0.92	0.97	0.83	1
News (30072 frames)	0.89	1	0.96	0.96
Total	0.91	0.95	0.93	0.97

본 논문에서 제안하는 기법은 동작 임계값에 의해 먼저 갑작스런 컷을 검출하고 컷으로 구분된 구간에서 점진적인 전환을 찾는 방식으로 수행되나 갑작스런 컷 검출에 따른 오류가 점진적인 전환 검출에 영향을 미치지 않음을 알 수 있다. 특히 제안하는 기법은 precision이 대단히 높은 것을 알 수 있는데 갑작스런 전환 검출에서 오 검출이 일어나는 경우에도 CFD의 변화는 식 (13)의 조건을 만족시키기까지 매우 힘들기 때문이다. 오히려 제안하는 기법은 동적 임계

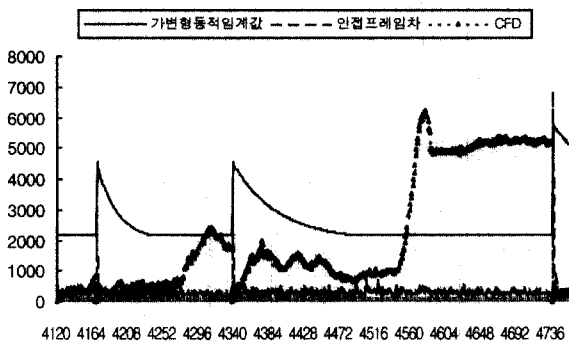
<표 1> 다양한 동영상 데이터를 통한 실험 결과

Video data	갑작스런 전환	점진적인 전환	갑작스런 전환 검출			점진적인 전환 검출		
			Detected	Missed	False Alarm	Detected	Missed	False Alarm
TEST sample (10400)	29	29	29	0	0	29	0	0
CF (1850 frames)	65	7	59	8	10	5	2	0
Drama (10704 frames)	121	12	112	9	3	10	2	0
News (30072 frames)	76	52	68	8	0	49	2	2

(그림 6) 뮤직 비디오 동영상 분석 사례

값이 컷을 검출하지 못한 미 검출 상황에서 장면전환이 일어났음을 알려주는 장점을 보였다.

(그림 6)은 실제 뮤직 비디오 동영상 분석 결과의 일부를 보여준다. 여기서 갑작스런 컷은 4170, 4337, 4728 프레임으로 모두 3 지점이며 4291 프레임을 중심으로 줌인 효과가 나타나고 약 4549 프레임부터 4573 프레임에 걸쳐 디졸브 효과가 사용되고 있다.



(그림 7) 뮤직비디오에서의 CFD 변화

(그림 7)은 (그림 6)의 뮤직 비디오 동영상에 대한 CFD 분석 상태를 보여준다. 갑작스런 전환의 경우 모두 가변형 동적 임계값에 의해 쉽게 발견됨을 알 수 있었다. 4170부터 4336 프레임 동안의 구간은 카메라가 주인공을 서서히 줌인 하는 내용으로 구성되어 있다. 이 구간에서의 CFD변화는 순간적으로 가변형 동적 임계값을 초과 하게 되지만 식 (13)을 만족하지 못하므로 점진적인 전환이 일어난 것으로 판단하지 않는다. (그림 7)의 4549 프레임은 디졸브 효과가 일어나는 장면으로 (그림 6)에서 해당 구간의 인접 프레임차는 거의 변화가 없는 것에 비해 CFD 변화는 가변형 동적 임계값을 중심으로 확연히 구분되므로 2개의 장면으로 분할 하게 된다. 따라서 제안하는 기법이 갑작스런 전환과 점진적인 전환에 모두에 대해 효과적으로 장면을 분할 할 수 있음을 알 수 있었다.

5. 결 론

일반적으로 점진적인 전환은 갑작스런 전환에 비해 발생

빈도가 상대적으로 낮기 때문에 점진적인 전환 검출만을 위하여 특별히 복잡한 알고리즘을 사용하는 것은 비효율적이라 할 수 있다. 또한 점진적인 전환은 전환이 일어나는 그 자체 보다는 전환이 일어나기 전과 후의 장면이 사용자에게 중요하다. 따라서 점진적인 전환도 갑작스런 전환과 같이 장면이 전환됐다는 사실을 쉽게 파악할 수 있는 방법이 필요하다고 하겠다.

본 논문은 가변형 동적 임계값과 컷 프레임 차를 사용하여 갑작스런 전환 및 점진적인 전환을 쉽게 검출할 수 있는 기법을 제안했다. 이를 위해 본 논문에서는 점진적인 전환 모델을 수학적으로 정의하고 CFD를 통해 각 장면이 구분될 수 있음을 보였으며 이에 따른 알고리즘을 제시했다. 제안한 기법은 점진적인 장면전환을 검출하기 위한 부가적인 연산이 매우 적으며 갑작스런 장면전환 검출을 위해 사용되는 가변형 동적 임계값을 같이 사용하므로 구현이 쉽고 빠른 처리가 가능한 장점을 갖는다. 제안한 기법은 기존의 연구와 달리 점진적인 전환이 일어나는 구간을 검출하려고 하지 않고 전환이 일어나기 전과 후의 장면을 검출해냄으로써 전환이 일어났다는 사실을 알아내는 접근 방법을 사용했다. 따라서 장면전환 효과의 종류나 전환이 일어나는 속도에 종속적이지 않으면서 장면전환 사실을 효과적으로 검출할 수 있었다. 실세계의 동영상 데이터틀 대상으로 한 실험에서도 제안한 기법을 통해 실제로 갑작스런 전환 및 점진적인 전환을 효율적으로 검출할 수 있음을 보였다. 앞으로 2개 이상의 점진적인 전환이 연속적으로 발생 할 때의 처리 방안과 제안한 방법에 가장 효율적으로 사용될 수 있는 특징값들을 추출 하는데 관한 연구를 진행할 예정이다.

참 고 문 헌

- [1] John S. Boreczky, Lawrence A. Rowe, "Comparison of video shot boundary detection techniques," in *Storage & Retrieval for Image and Video Database IV*, Proc. of SPIE 2670, pp.170-179, 1996.
- [2] Yoshinobu Tonomura, "Video handling based on structured information for hypermedia systems," *ACM Proc. of international conference on Multimedia Information Systems*, 91, pp.333-344, 1991.
- [3] Jurgen Stauder, Henri Nicolas, "Motion-Based Video Indexing Evaluating Object Shading," In *Proceedings of ICIP*, Kobe, October, 1999.
- [4] 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.
- [5] Ramin Zabih, Justin Miller, Kevin Mai, "A Feature-Based Algorithm for Detecting and Classifying Scene Breaks," *ACM Multimedia*, pp.189-200, 1995.
- [6] Kiyotaka Otsuji, Yoshinobu Tonomura, Yuji Ohba, "Video Browsing using brightness data," *Proc. of SPIE. VCIP'91*, pp.980-989, 1991.
- [7] HJ. Zhang, A. Kankanhalli, S. W. Smoliar, "Automatic partitioning of full-motion video," *Multimedia Systems*, Vol.1, No.1, pp.10-28, 1993.
- [8] C. W. Ngo, T. C. Pong, R. T. Chin, "Detection of Gradual Transitions through Temporal Slice Analysis," *Proc. IEEE CVPR99*, pp.36-41, June, 1999.
- [9] Chong-Wah Ngo, Ting-Chuen Pong, Roland T. Chin, "Video Partitioning by Temporal Slice Coherency," *IEEE Transactions on Circuit and Systems for Video Technology*, Vol.11, No.8, Aug., 2001.
- [10] Min Gyo Chung, Hyeokman Kim, S. Moon-Ho Song, "A Scene Boundary Detection Method," *Image Processing, 2000. Proceedings. 2000 International Conference*, Vol.3, pp.933-936, 2000.
- [11] W. A. C. Fernando, C. N. Canagarajah, D. R. Bull, "A Unified Approach to Scene Change Detection in Uncompressed and Compressed Video," *Consumer Electronics, IEEE Transactions on*, Vol.46, Issue : 3, pp.769-779, Aug., 2000.
- [12] Dalong Li, Hanqing Lu, "Model Based Video Segmentation," *The IEEE Workshop on Signal Processing Systems*, pp.120-129, 2000.
- [13] W. A. C. Fernando, C. N. Canagarajah, D. R. Bull, "Wipe Scene Change Detection in Video Sequences," *ICIP 99. Proceedings. 1999 International Conference on Image Processing*, Vol.3, pp.294-298, 1999.
- [14] Rainer Lienhart, Andre Zaccarin, "A System for Reliable Dissolve Detection in Videos," *Proceedings. 2001 International Conference on Image Processing*, Vol.3, pp.406-409, 2001.
- [15] 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*, May, 2001.
- [16] H. Yu, G. Bozdagi, S. Harrington, "Feature-based Hierarchical Video Segmentation," *Proceedings of International Conference on Image Processing*, pp.498-501, Vol.2, 1997.
- [17] 염성주, 김우생, "동영상 컷 검출을 위한 가변형 동적 임계값 기법", *한국통신학회논문집*, 제27권 제4호, 2002.
- [18] 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.
- [19] Adobe, *Adobe Premiere 6.0 online help document*.

열 성 주

e-mail : sjyoum@kyungbok.ac.kr

1991년 광운대학교 전자계산학과 졸업
(학사)

1993년 광운대학교 대학원 전자계산학과
(이학석사)

1993년~현재 광운대학교 컴퓨터과학과
박사과정 수료

1998년~현재 경북대학 인터넷정보과 조교수

관심분야 : 멀티미디어 시스템, 내용기반 검색, 비디오 분석, 이
미지 프로세싱

김 우 생

e-mail : woosaeng@cs.kwangwoon.ac.kr

1982년 서울대학교 수료

1985년 University of Texas at Austin
졸업(학사)

1987년 University of Minnesota(이학
석사)

1987년~1988년 현대전자 Zeus Computer 과장

1991년 University of Minnesota(이학박사)

1992년~현재 광운대학교 컴퓨터공학부 교수

관심분야 : 멀티미디어, 영상/비디오 처리, 데이터베이스