

이웃 장면들의 특성을 이용한 비디오 디졸브 검출

(Video-Dissolve Detection using Characteristics of Neighboring Scenes)

원 종 운 [†] 최 재 각 ^{**} 박 철 현 ^{***} 김 범 수 ^{****}
(Jong-Un Won) (Jae-Gark Choi) (Chul-Hyun Park) (Bum-Soo Kim)

곽 등 민 ^{***} 오 상 근 ^{***} 박 길 흠 ^{*****}
(Dong-Min Kwak) (Sang-Keun Oh) (Kil-Houm Park)

요 약 본 논문에서는 디졸브 모델링 오차를 이용한 디졸브 검출 방법을 제안한다. 디졸브 모델링 오차는 디졸브 구간을 구성하는 두 장면 사이에 상관성이 없는 이상적인 디졸브 모델과 상관성이 존재하는 실제 디졸브 사이의 차이로 정의 된다. 제안된 방법은 두 단계로 구성된다. 첫 번째 단계에서는 분산 곡선에서 나타나는 디졸브 구간의 특성인 아래로 볼록한 포물선을 검출하여 디졸브 후보 구간을 설정한다. 두 번째 단계에서는 선택된 후보 구간 각각에 대해 디졸브 모델링 오차를 정의하고 모델링 오차가 기준치와 비교한다. 기준치는 목표 모델링 오차로 각 후보 구간의 분산과 사용자에 의해 주어지는 목표 상관성에 대한 함수로 표현됨으로 제안된 기준치, 목표 모델링 오차는 각 후보 구간의 분산 변화에 적응적이다. 제안한 디졸브 검출 방법은 디졸브 구간을 구성하는 두 장면의 상관성을 고려하므로, 의미론적 디졸브 검출 방법이 될 수 있다. 제안된 방법을 다양한 동영상에 적용하여 성능을 평가하였다. 실험 결과 제안한 방법이 다양한 분산의 변화에도 불구하고 기존의 방법보다 정확하고 신뢰성 높은 디졸브 검출 결과를 나타내었다.

키워드 : 디졸브 검출, 장면 변화 검출

Abstract In this paper, we propose a new adaptive dissolve detection method based on the analysis of a dissolve modeling error which is the difference between an ideally modeled dissolve curve with no correlation and an actual dissolve curve including a correlation. The proposed dissolve detection method consists of two steps. First, candidate dissolve regions are extracted using the characteristics of a downward convex parabola, then each candidate region is verified based on the dissolve modeling error. If the dissolve modeling error for a candidate region is less than a threshold defined by the target modeling error with a target correlation, the candidate region is determined as a dissolve region with a lower correlation than the target correlation. The threshold is adaptively determined based on the variances between the candidate regions and the target correlation. By considering the correlation between neighbor scenes, the proposed method is able to be a semantic scene-change detector. The proposed method was tested on various types of data and its performance proved to be more accurate and reliable regardless of variation of variance of test sequences when compared with other commonly used methods.

Key words : dissolve detection, scene change detection, video parsing, video browsing system video indexing

[†] 비 회 원 : 한국전자통신연구원 우정기술연구센터 연구원
bestwon@etri.re.kr
^{**} 비 회 원 : 동의대학교 컴퓨터공학과 조교수
cjg@dongeui.ac.kr
^{***} 비 회 원 : 경북대학교 전자공학과
nagne@palgong.knu.ac.kr
imis@palgong.knu.ac.kr

teasa@palgong.knu.ac.kr
^{****} 정 회 원 : 상주대학교 전자전기공학부 교수
bskim@palgong.knu.ac.kr
^{*****} 비 회 원 : 경북대학교 전자공학과 교수
khpark@ee.knu.ac.kr
논문접수 : 2002년 5월 3일
심사완료 : 2003년 4월 21일

1. 서론

디지털 기술의 발전과 더불어 디지털 비디오 기술 또한 급속히 발전하고 있다. 특히 VOD(video-on-demand), 영상 회의, 전자 도서관, 멀티미디어 교육 분야에서 디지털 비디오에 대한 요구가 날로 증가되고 있다. 그러나 디지털 비디오에 대한 폭넓은 이용이 제한되는 이유는 디지털 콘텐츠를 사용하는 거대한 양의 데이터 베이스로부터 원하는 정보를 추출하기가 쉽지 않기 때문이다. 이러한 문제는 비디오 동영상을 적절한 방법으로 색인을 하여 해결할 수 있다. 비디오 동영상이 색인된다면, 멀티미디어 데이터 베이스에서 원하는 비디오 클립에 빠르게 접근할 수 있을 것이다. 이와 같이 비디오 색인을 하기 위해서는 무엇보다 먼저 동영상을 시간에 대해서 적절하게 분할하여야 한다. 이러한 것은 동영상을 구성하는 장면과 장면의 변화 부분을 검출하는 장면 변화 검출(scene change detection)을 통해 이루어진다. 일반적으로 장면 변화는 갑작스러운 장면 변화(abrupt scene change)와 점진적 장면 변화(gradual scene change)로 나눌 수 있다. 그리고 점진적 장면 변화에는 fade-in, fade-out, 디졸브(dissolve) 등을 포함한다.

갑작스런 장면 변화 검출 방법에 대한 연구는 많이 이루어져왔다[1-5]. 또한 그 결과도 우수하다. 반면 비디오 동영상에는 많은 점진적 장면 변화가 포함되어 있음에도 불구하고 점진적 장면 변화 검출에 대한 연구는 비교적 미미하여[6], 물체 또는 카메라 움직임 등의 영향으로 디졸브 검출 성능이 갑작스런 장면 변화에 비해서 낮은 수준이다[5-11]. 따라서 잡음에 강인한 점진적 장면 변화 검출에 대한 연구가 필요한 실정이다. Zhang et al.[5]은 연속된 프레임들의 히스토그램 차를 이용하는 twin-comparison 방법을 이용하여 갑작스런 장면 변화와 점진적 장면 변화를 검출하였다. 이 방법은 비교적 좋은 결과를 나타내지만, 응용 대상에 따라 적절한 임계값을 설정해야만 한다. Alattar[7] and Meng et al.[8]은 점진적 장면 변화의 하나인 디졸브 검출을 위해 동영상의 각 프레임의 분산으로부터 만들어지는 분산 곡선의 특징을 이용하였다. 분산 곡선은 디졸브 구간에서 아래로 볼록한 포물선 모양을 나타내게 되는데, Alattar는 디졸브 구간의 시작과 끝에서 분산 곡선의 2차 미분값이 음의 최소값으로 나타남을 증명하고 이 음의 최소값이 일정한 임계값을 넘는 구간에 대해 그 구간의 평균이 또 다른 임계값 보다 크고 그 구간의 길이가 일정한 길이보다 큰 구간을 디졸브 구간으로 정의

하였다. 또한 Meng은 분산 곡선의 1차 미분에서 인접한 최대값과 최소값의 차이를 기준으로 디졸브 구간을 찾았다. 그러나 이들 방법은 1, 2차 미분에서 최대값과 최소값 또는 다른 조건을 만족할 수 있도록 임계값을 설정해야 하는데, 대상 동영상에 따라 사용자의 경험에 의해 적절한 임계값을 설정해야 한다.

본 논문에서는 디졸브 구간을 구성하는 두 장면 사이의 상관성과 분산을 고려하여 얻어지는 디졸브 모델링 오차(dissolve modeling error)를 이용한 디졸브를 검출하는 방법을 제안한다. 디졸브 모델 오차는 두 장면 사이의 상관성을 포함하고 있는 실제 분산 곡선으로 주어지는 실제 디졸브와 두 장면 사이의 상관성이 없을 경우 추정되는 이상적인 디졸브 모델 사이의 차이로 주어진다. 제안된 방법은 두 단계로 이루어진다. 첫 번째 단계에서는 분산 곡선의 1, 2차 미분을 이용하여 디졸브 후보 구간(candidate dissolve region)을 검출한다. 이렇게 얻어지는 후보 영역들은 아래로 볼록한 포물선 구간이 된다. 따라서 디졸브 후보 구간은 진짜 디졸브 구간과 의사 디졸브 구간(pseudo dissolve)을 모두 포함하고 있는데, 두 번째 단계에서는 후보 영역들에서 디졸브 모델 오차를 해석하여 디졸브를 판별한다. 특정 후보 영역에서 측정되는 디졸브 모델 오차의 최대값이 사용자가 검출하고자 하는 목표 상관성(target correlation)을 가지는 디졸브에서 발생할 수 있는 최대 디졸브 모델 오차 보다 작은 경우, 후보 영역은 디졸브 영역으로 판별된다. 여기서 목표 상관성은 비디오 분할을 제어할 수 있는 요소가 된다.

본 논문의 나머지는 다음과 같이 구성되어 있다. 2장에서는 디졸브 모델과 특징을 설명하고, 3장에서는 제안된 디졸브 검출 과정을 소개한다. 4장에서 제안된 디졸브 검출 방법과 기존의 방법을 실험하여 그 결과를 분석하고, 5장에서 결론을 맺는다.

2. 디졸브 모델과 특징

2.1 디졸브 모델

두 장면을 부드럽게 변화시키기 위해서 디졸브, fade-in 그리고 fade-out과 같은 점진적 장면 변화를 비디오 편집 기법으로 종종 사용하고 있다. 디졸브 영역은 이웃한 두 장면의 선형적인 합으로 생성된다. 한 장면의 밝기를 점진적으로 감소시키면서 동시에 다른 장면의 밝기를 점진적으로 증가시키는 것이다. 이때 한 장면의 밝기를 감소시켜 어둡게 만드는 것을 "fade-out"이라 하고 역으로 밝기가 균일한 장면으로부터 어떤 장면으로 점점 변화하는 것을 "fade-in"이라 한다. 따라서 fade-out

의 끝과 fade-in의 시작에서 분산은 영이 된다. 그러므로 fade-in과 fade-out 검출은 디졸브 검출에 비해 아주 쉽다. 따라서 본 논문에서는 디졸브 검출에 초점을 두고 설명한다.

일반적으로 디졸브 모델은 다음과 같이 모델링 된다. $p(x, y, t)$ 를 변환전 장면, $q(x, y, t)$ 는 변환 후 장면 그리고 $f(x, y, t)$ 를 $p(x, y, t)$ 와 $q(x, y, t)$ 의 결합에 의해 생성되는 천이 프레임, 즉 디졸브 영역의 프레임이라고 하자. $p(x, y, t)$ 와 $q(x, y, t)$ 가 엘고딕 랜덤 프로세서(ergodic random processes)라고 가정한다면, 이들은 각각 $p(x, y)$ 와 $q(x, y)$ 로 나타낼 수 있다. 그러므로 이들 장면에서 분산 $\sigma_p^2(t)$ 와 $\sigma_q^2(t)$ 는 일정한 상수 σ_p 와 σ_q 가 된다. 디졸브의 시작과 끝을 t_1 과 t_2 라 한다면, 디졸브 모델은 식 (1)로 주어진다.

$$f(x, y, t) = \begin{cases} p(x, y) & t < t_1 \\ \alpha(t)p(x, y) + \beta(t)q(x, y) & t_1 \leq t \leq t_2 \\ q(x, y) & t > t_2 \end{cases} \quad (1)$$

여기서 x, y 그리고 t 는 각각 수평, 수직 그리고 시간에 대한 연속 변수이다. 변수 $\alpha(t)$ 는 $(t_2 - t)/(t_2 - t_1)$ 으로 점진적 장면 변화 구간에서 감소 함수가 된다. 한편 $\beta(t)$ 는 점진적 장면 변화 구간에서 증가함수가 되며 $(t - t_1)/(t_2 - t_1)$ 으로 표현된다. 두 변수 $\alpha(t)$ 와 $\beta(t)$ 의 합은 항상 1이 된다. 디졸브 구간내의 프레임 $f(x, y, t)$ 의 분산은 식 (2)로 표현된다.

$$\begin{aligned} \sigma_f^2(t) &= E[(f - \bar{f})^2] \\ &= E[(\alpha(t)p(x, y) + \beta(t)q(x, y) - \alpha(t)\bar{p}(x, y) - \beta(t)\bar{q}(x, y))^2] \\ &= \alpha^2(t)\sigma_p^2 + \beta^2(t)\sigma_q^2 + 2\alpha(t)\beta(t)E[(p(x, y) - \bar{p}(x, y))(q(x, y) - \bar{q}(x, y))] \end{aligned} \quad (2)$$

$p(x, y)$ 와 $q(x, y)$ 가 통계적으로 독립이라 가정한다면, 두 장면 사이의 공분산이 영, 즉 상관성이 없음으로 식 (2)는 식 (3)으로 근사화되어 간단히 표현된다.

$$\begin{aligned} \sigma_f^2(t) &= \alpha^2(t)\sigma_p^2 + \beta^2(t)\sigma_q^2 \\ &= [\sigma_p^2 + \sigma_q^2]\alpha^2(t) - 2\sigma_p^2\alpha(t) + \sigma_q^2 \end{aligned} \quad (3)$$

식 (3)으로 주어지는 상관성이 없는 디졸브 모델을 이상적인 디졸브 모델이라 한다. 디졸브 구간은 분산 곡선에서 아래로 볼록한 포물선 모양으로 나타난다. 이와 같이 기존의 방법들은 식(1)-(3)의 디졸브 모델을 기반으로 디졸브 구간을 검출하였다[7,8].

2.2 디졸브 구간의 특징

식 (3)의 이상적인 디졸브 모델은 식 (2)의 세 번째 항이 의미하는 디졸브 구간에서 두 장면 사이의 상관성

의 영향을 무시하여 얻어진 것으로 상관성의 영향으로 발생하는 모델링 오차를 포함하고 있다. 기존의 모델 기반 디졸브 검출 방법[6-11]에서는 두 장면 사이의 상관성이 없다고 가정하여 이러한 모델링 오차를 무시하였다. 본 절에서는 두 장면 사이의 상관성에 따른 디졸브 구간의 특징을 설명한다.

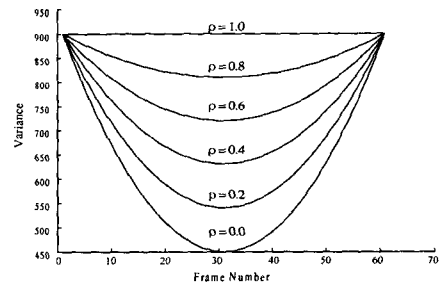
디졸브 모델링 오차는 식 (2)의 세 번째 항으로 식 (4)로 주어진다.

$$\begin{aligned} D_{pq}(t) &= 2\alpha(t)\beta(t)E[(p - \bar{p})(q - \bar{q})] \\ &= 2\alpha(t)\beta(t)\sigma_p\sigma_q\rho_{pq} \end{aligned} \quad (4)$$

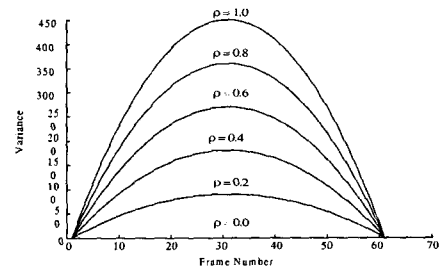
여기서, ρ_{pq} 는 σ_p 와 σ_q 에 의해 정규화된 공분산이고, $0 \leq |\rho_{pq}| \leq 1$ 이다. 디졸브 모델링 오차의 최대값은 디졸브의 중심인 $\alpha(t) = 0.5$ 에서 발생한다. 이때 발생하는 디졸브 모델링 오차의 최대값 D_{max} 는 식 (5)로 주어진다.

$$D_{max} = \frac{\sigma_p\sigma_q\rho_{pq}}{2} \quad (5)$$

그림 1은 디졸브 구간에서 상관성 변화에 따른 실제 분산 곡선과 모델링 오차의 변화를 나타내고 있다. 공분산 ρ_{pq} 가 1에 가까울수록 즉 상관성이 높을수록, 분산 곡선은 직선에 가까워지며 디졸브 특징으로 나타나는



(a) 실제 분산 곡선



(b) 디졸브 모델링 오차

그림 1 실제 분산과 상관성에 따른 디졸브 모델링 오차

포물선의 골의 깊이가 얕아지고 모델링 오차는 증가한다. 반대로 공분산 ρ_{pq} 가 영에 가까울수록 이상적인 디졸브 모델과 같아지고 포물선의 골의 깊이가 깊어지고 모델링 오차는 감소한다. 특히 디졸브 구간의 중심에서 모델링 오차가 최대가 됨을 알 수 있다. 구간 $[p, q]$ 에서 두 장면 사이의 상관성이 C 라면 이때 발생할 수 있는 최대 디졸브 모델링 오차 $D_{max, c}$ 는 식 (6)으로 나타낼 수 있다.

$$D_{max, c} = \frac{\sigma_p \sigma_q c}{2} \tag{6}$$

실제 분산 곡선과 이상적인 디졸브 모델 사이에 발생할 수 있는 최대 오차 D_{max} 가 $D_{max, c}$ 보다 작다면 구간은 상관성이 C 보다 작은 디졸브 구간으로 생각할 수 있다. 따라서 상관성이 C 일 때 최대 디졸브 모델링 오차 $D_{max, c}$ 는 상관성이 C 보다 작은 디졸브 구간을 검출하기 위한 임계값이 되고, 이때 상관성 C 는 목표 상관성(target correlation)이 된다. 목표 상관성은 사용자가 정의하는 것으로 목표 상관성이 영에 가까울수록 최대 모델링 오차는 영에 가까워져 이상적인 모델에 가까운 디졸브를 검출하고, 1에 가까울수록 상관성이 큰 디졸브들도 검출한다. 식 (6)에 나타나듯이 임계값은 구간의 지역적 특징인 분산의 함수로 표현됨으로 적응성을 포함한다. 목표 상관성에 의해 주어지는 최대 모델링 오차를 간단히 목표 모델링 오차(target modeling error)라고 한다.

식 (3)으로 표현되는 이상적인 디졸브 모델 $\sigma_f^2(t)$ 에서 디졸브 구간의 중심($\alpha(t) = 0.5$)에서 분산 $\tilde{\sigma}_{center}^2$ 는 식 (7)로 주어진다.

$$\tilde{\sigma}_{center}^2 = \frac{\sigma_p^2 + \sigma_q^2}{4} \tag{7}$$

식 (3)으로 표현되는 이상적인 디졸브 모델 $\sigma_f^2(t)$ 와 디졸브 구간의 특징의 하나인 식 (7)로 표현되는 디졸브 중심의 분산 값은 모두 디졸브 구간을 구성하는 두 장면의 분산 σ_p 와 σ_q 에 영향을 받는다. 또한 이상적인 디졸브 모델은 식 (2)의 세 번째 항에 해당하는 디졸브 모델 근사화 오차에 해당하는 디졸브 모델링 오차를 포함하고 있다. 그러므로 정확한 디졸브 검출을 위해서는 디졸브를 구성하는 두 장면의 분산의 영향뿐만 아니라 두 장면 사이의 상관성에 대한 고려가 있어야 한다.

3. 디졸브 검출

본 논문에서는 디졸브 모델링 오차를 이용하여 디졸브를 검출하기 위해 그림 2에 나타나듯이 두 단계로 나누어 디졸브 구간을 검출한다. 첫 번째 단계로 디졸브 구간이 될 가능성이 높은 후보 구간을 검출한다. 후보 구간은 디졸브 구간이 분산 곡선에서 포물선 모양으로 나타남으로 분산 곡선의 1, 2차 미분의 특징을 이용하여 아래로 볼록한 포물선을 찾아 이를 후보 구간으로 설정한다. 그리고 검출된 후보 구간에 대해서 두 번째 단계로 최대 디졸브 모델링 오차를 기준으로 디졸브 구간을 판별한다.

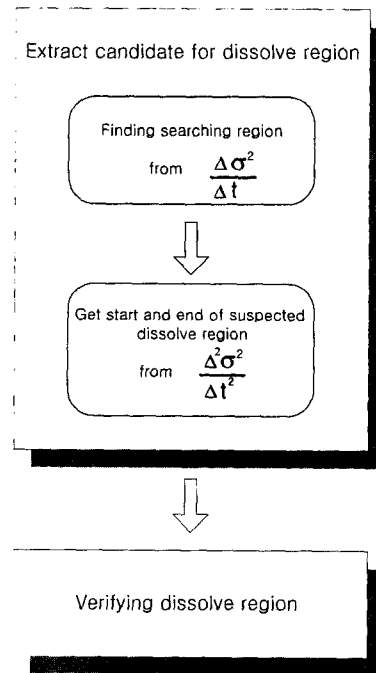


그림 2 디졸브 검출 블록 선도

3.1 디졸브 후보 검출

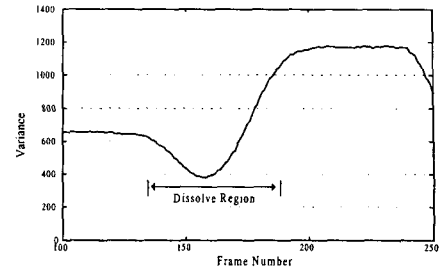
후보 구간은 디졸브 구간이 분산 곡선에서 아래로 볼록한 포물선 모양으로 나타나는 특징을 이용하여 검출된다. 분산 곡선에서 같은 장면인 구간은 분산 곡선의 변화가 없이 평평한 상태로 나타난다. 반면 디졸브 구간은 아래로 볼록한 포물선 모양으로 나타난다. 따라서 디졸브 구간은 분산 곡선의 1차 미분에서 항상 음에서 양으로 변하는 영 교차점(zero crossing point)을 포함한다. 또한 디졸브 구간에 대응하는 포물선 구간에서는 분

산 곡선의 변화가 있음으로 분산 곡선의 1차 미분은 영보다 크거나 작은 값을 가진다. 하지만 같은 장면인 구간에서는 분산 곡선의 변화가 거의 없으므로 1차 미분이 영이 된다. 따라서 분산 곡선의 1차 미분에서 음에서 양으로 변하는 영 교차점을 먼저 찾고 영 교차점을 기준으로 좌, 우로 탐색하여 1차 미분이 처음으로 영이 되는 지점을 후보 구간의 시작과 끝으로 설정할 수 있다. 일반적으로 카메라 또는 물체의 움직임 그리고 잡음의 영향으로 후보 구간은 실제 디졸브 구간과 일치하지 않고, 대체로 더 넓은 구간이 된다. 따라서 분산 곡선의 1차 미분의 특징으로 얻어지는 후보 구간을 탐색 구간으로 정의하고, 탐색 구간 내에서 정확한 디졸브 구간에 대응하는 후보 구간을 정의한다.

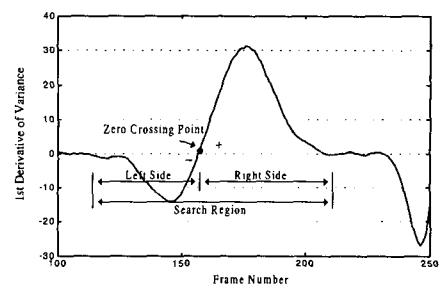
Alatter는 분산 곡선의 2차 미분에서 디졸브에 대응하는 포물선의 시작과 끝에서 음의 최소값이 나타남을 증명하였고[7], 이러한 특징을 이용하여 2차 미분이 어떤 임계값 보다 작을 경우 그 점들을 디졸브 후보 구간의 시작과 끝으로 정의하여 디졸브 구간을 검출하였다. 하지만 음의 최소값은 디졸브 구간을 형성하는 두 장면의 분산 뿐만 아니라 특히 이들 분산의 편차가 크게 영향을 받는다. 두 장면의 분산의 편차가 큰 경우 분산이 작은 장면에서 발생하는 음의 최소값의 크기는 분산이 큰 장면 쪽에서 발생하는 음의 최소값 보다 아주 작다. 하지만 이들 분산이 작거나 분산의 편차가 크더라도 디졸브 구간에서 음의 최소값은 항상 디졸브의 시작과 끝에서 발생한다. 따라서 본 논문에서는 임계값을 이용하여 디졸브의 시작과 끝을 정의하지 않고, 분산 곡선의 1차 미분 특징을 이용하여 설정된 탐색 구간 내에서 음의 최대값을 찾아 디졸브 구간의 시작과 끝으로 설정한다.

그림 3은 후보 구간 설정 과정을 설명하고 있다. 그림 3(a)는 디졸브 구간의 분산 곡선이고, 그림 3(b)는 분산 곡선의 1차 미분에서 탐색 구간 설정에 대한 설명을 하고 있다. 분산 곡선의 1차 미분에서 우선 음에서 양으로 변하는 영 교차점을 찾는다. 주어진 영 교차점을 기준으로 왼쪽으로 탐색하여 처음으로 1차 미분이 영이 되는 위치를 탐색 구간의 시작 위치로 설정하고, 영 교차점을 중심으로 오른쪽으로 탐색하여 분산 곡선의 1차 미분이 처음으로 영이 되는 위치가 탐색 구간의 끝 위치로 설정 된다. 이때 영 교차점과 탐색 구간의 시작 위치 사이 구간을 탐색 구간의 왼쪽 영역(left side)이라 하고, 마찬가지로 영 교차점과 탐색 구간의 끝 위치 사이의 구간을 오른쪽 영역(right-side)이라고 한다.

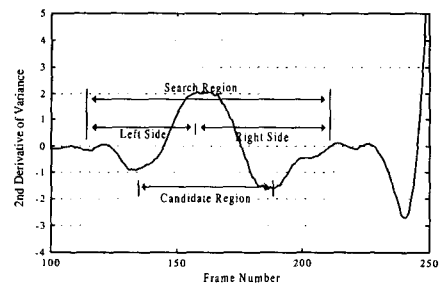
그림 3(c)는 탐색 구간 내에서 분산 곡선의 2차 미분



(a) 분산 곡선



(b) 탐색 구간



(c) 후보 구간

그림 3 디졸브 후보 구간 설정

특징을 이용하여 디졸브 후보 구간을 설정하는 것을 나타내고 있다. 탐색 구간의 왼쪽 영역과 오른쪽 영역에서 분산 곡선에 대한 2차 미분의 음의 최소값의 위치를 각각 디졸브 후보 구간의 시작과 끝의 위치로 설정한다.

3.2 디졸브 구간의 판별

디졸브 후보 구간들은 분산 곡선에서 디졸브 구간에 대응하는 아래로 볼록한 포물선을 검출하기 위해 분산 곡선의 1, 2차 미분 특징만을 이용하여 검출되었다. 따라서 앞서 추출된 디졸브 후보 구간들은 디졸브에 대응하는 포물선과 카메라 움직임 또는 물체의 움직임 등으로 인해 발생하는 포물선을 모두 포함하고 있다. 따라서 디졸브의 다른 특징을 이용하여 디졸브 후보 구간에서 진짜 디졸브 구간을 판별하여야 한다.

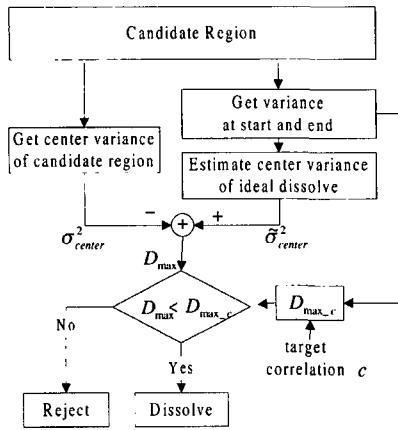


그림 4 디졸브 구간 판별을 위한 흐름도

그림 4는 디졸브 구간을 판별하기 위한 흐름도를 나타내고 있다. 설정된 목표 상관성과 각 후보 구간에서 주어지는 후보 구간의 시작과 끝의 분산을 식 (6)에 대입하여 디졸브 판별을 위한 기준값, 목표 모델링 오차 D_{max_c} 를 계산한다. 이때 목표 상관성 C 는 사용자가 정의하는 것으로 목표 상관성이 영에 가까울수록 엄격한 디졸브 검출을, 1에 가까울수록 대략적인 디졸브 검출 결과를 얻을 수 있다.

다음으로 각 후보 구간에서 발생하는 실제 모델링 오차의 최대값을 구한다. 최대 모델링 오차는 디졸브 구간의 중심에서 발생하므로, 후보 구간의 중심에서 분산 σ_{center}^2 와 식 (7)로 주어지는 이상적인 디졸브 모델의 중심에서의 분산 $\hat{\sigma}_{center}^2$ 의 차이가 각 후보 구간에서 발생하는 디졸브 모델링 오차의 최대값 D_{max} 가 된다. 따라서 각 후보 구간에서 발생하는 디졸브 모델링 오차의 최대값 D_{max} 가 목표 모델링 오차 D_{max_c} 보다 작다면 그 후보 구간은 디졸브 구간이 된다. 그렇지 않다면 그 후보 구간은 디졸브 구간이 아니거나 디졸브 구간의 두 장면 사이의 상관성이 C 보다 큰 경우가 된다.

4. 실험 및 결과

제안된 디졸브 검출 방법을 다양한 비디오 동영상에 대상으로 실험하였다. 실험에 사용된 동영상은 SIF (220x240) 해상도의 디졸브 구간이 90-75 프레임이다. 디졸브 구간에서 두 장면 사이의 상관성은 0.01~0.42이다.

그림 5는 "autumn" 동영상의 분산 곡선을 보여 주고 이어서 2개의 디졸브 구간이 표시된 프레임을 나타내

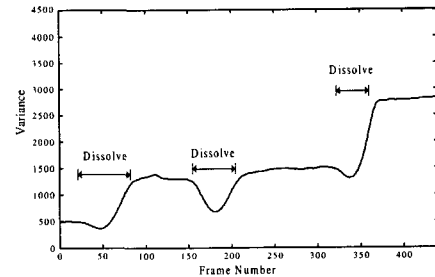


그림 5 "autumn" 동영상의 분산 곡선

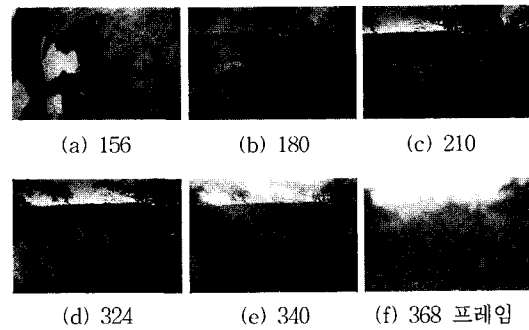


그림 6 "autumn" 동영상의 디졸브 구간의 프레임

다. 그림 6은 "autumn" 동영상의 두 번째 디졸브와 세 번째 디졸브 구간의 처음, 가운데 그리고 마지막 프레임을 각각 나타내고 있다. "autumn" 동영상의 두 번째 디졸브 구간에서 두 장면 사이의 상관성은 0.09이고, 세 번째 디졸브 구간에서 두 장면 사이의 상관성은 0.42이다. 세 번째 디졸브 구간의 처음과 끝 두 프레임이 아주 유사함을 알 수 있다. 목표 상관성이 0.35일 경우 두 번째 디졸브 구간은 검출되었지만 세 번째 디졸브 구간은 검출되지 못하였다. 하지만 목표 상관성을 0.45로 둘 경우 두 디졸브 구간이 모두 검출되었다. 그림 6에 보이듯이 324번째 프레임과 368번째 프레임의 유사성이 156번째 프레임과 210번째 프레임의 유사성 보다 상당히 높음을 알 수 있다. 표 1에 실험 영상에 대해서 목표 상관성의 변화에 따른 디졸브 검출 결과를 나타내었다. 일반적으로 같은 장면에서 두 프레임 사이의 상관성은 0.8~1.0로 주어지고, 반면 서로 다른 장면의 두 프레임 사이의 상관성은 보통 0.5보다 작게 나타났다. 그러므로 목표 상관성을 0.5정도 설정할 경우 비교적 정확한 디졸브 검출 결과를 얻을 수 있다.

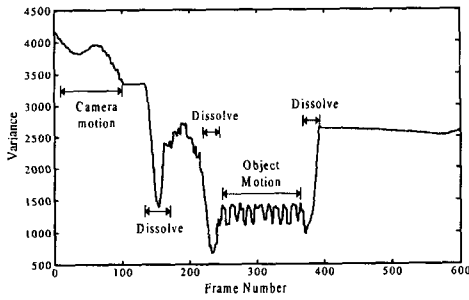
실험 영상에 대한 실험 과정은 다음과 같이 진행되었다. 제안된 방법은 목표 상관성을 0.5로 고정하여 전체 실험 영상에 적용한 결과이고, 기존의 방법들은 여러 개

표 1 목표 상관성에 따른 디졸브 검출 결과
(detection : O, failure : X)

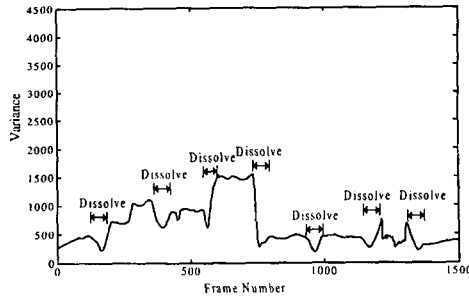
Seq. Name	Dissolve		Target correlation C			
	Region	Correlation ρ	0.15	0.25	0.35	0.45
Autumn	27 - 82nd	-0.01	O	O	O	O
	156 - 210th	0.09	O	O	O	O
	324 - 368th	0.42	X	X	X	O
Sunrise	12 - 87th	-0.11	O	O	O	O
	135 - 186th	-0.02	O	O	O	O
	242 - 270th	0.16	X	O	O	O
Toy	370 - 396th	-0.09	O	O	O	O
	137 - 164th	0.06	O	O	O	O
	216 - 249th	0.18	X	O	O	O
	366 - 393rd	0.32	X	X	O	O

표 2 검출된 디졸브 구간

Seq. Name	Actual Dissolve Region	Detected Dissolve Region	
		Proposed method	Alattar's method
Autumn	27 - 82nd	27 - 82nd	24 - 77th
	156 - 210th	156 - 210th	153 - 206th
	324 - 368th	324 - 368th	323 - 364th
Sunrise	12 - 87th	12 - 87th	12 - 87th
	135 - 186th	135 - 186th	131 - 183rd
	242 - 270th	242 - 270th	239 - 267th
	370 - 396th	370 - 396th	366 - 393rd
Toy	137 - 164th	137 - 164th	134 - 161st
	216 - 249th	216 - 249th	214 - 247th
	366 - 393rd	366 - 393rd	365 - 390th
Music_1	150 - 199th	150 - 199th	147 - 195th
	353 - 427th	353 - 427th	353 - 426th
	549 - 581st	549 - 581st	546 - 577th
	736 - 767th	736 - 767th	732 - 764th
	942 - 990th	942 - 990th	938 - 985th
	1136 - 1211th	1136 - 1211st	1135 - 1186th
Music_2	1307 - 1368th	1307 - 1368th	-
	544 - 561st	544 - 561st	538 - 563rd
	626 - 667th	626 - 667th	-
	675 - 709th	675 - 709th	-
	818 - 854th	818 - 854th	818 - 851st
	933 - 971st	933 - 971st	-



(a) "toy" 동영상의 분산 곡선



(b) "music" 동영상의 분산 곡선

그림 7 실험 동영상의 분산 곡선

의 변수의 값을 바꾸어가면서 recall이 증가하도록 하고, precision값이 감소되기 시작할 때의 실험 결과를 나타내었다. 이와 같이 얻은 결과를 표 2와 3에 나타내었다. 그림 7에 나타나는 "toy"와 "Music_1" 동영상의 분산 곡선에서 알 수 있듯이 이들 동영상은 다양한 카메라와 물체 움직임을 포함하고 있다. 또한 "Music_1" 동영상의 경우 매우 낮은 분산을 가지고 있다. 하지만 목표 상관성을 0.5로 설정하고 이들 동영상에 적용한 결과 모든 디졸브 구간이 정확히 검출되었다. 비록 카메라 또는 물

표 3 디졸브 검출 결과 비교

	Proposed method	Alattar's Method	Meng's method
Total	22	22	22
Detected	26	23	33
Correct	22	18	20
Recall	100%	81.8%	90.9%
Precision	84.5%	78.3%	60.6%

체의 움직임에 의해 발생한 다양한 포물선이 있지만 이들 포물선에서 추정된 디졸브 모델링 오차의 최대값 D_{max} 가 목표 모델링 오차 D_{max} 보다 컸기 때문에 이들이 디졸브 구간으로 검출되지 않았다. 그러나 비교적 어두운 영상에서 밝아졌다 어두워지는 조명과 물체의 움직임 작용으로 인해 분산 곡선에서 디졸브와 유사한 포물선을 형성하여 디졸브 구간으로 잘못 인식되는 경우가 있었다. 기존의 방법에서는 낮은 분산에서 형성되는 디졸브 구간을 검출하지 못하면서 조명 또는 물체의 움직임으로 발생한 포물선을 디졸브 구간으로 잘못

검출하였다. 특히 그림 7(a)의 "toy" 동영상 분산 곡선에서 첫번째 포물선은 카메라 움직임에 의해 발생한 것으로 Alattar의 방법에서는 디졸브 구간으로 잘못 인식하였다. "Music_1"과 "Music_2" 동영상의 경우 낮은 분산에서 형성되는 디졸브 구간이 많아 Alattar 방법에서는 디졸브 검출에 실패하였다.

본 논문에서는 디졸브 구간이 분산 곡선에서 포물선 모양으로 나타난다는 특징을 이용하여 디졸브 구간을 검출하였기 때문에 디졸브와 유사한 포물선을 형성하는 현상을 디졸브와 구별하기 어렵다. 이러한 false detection을 제거하기 위해서는 새로운 후처리의 연구가 필요하다. 실험 동영상에 포함된 모든 디졸브 구간을 검출할 수 있었다.

표 3은 식 (8)과 (9)에 표시된 Recall과 Precision을 기준으로 디졸브 검출 성능을 기존의 방법들과 비교하여 나타내었다. 실험에 사용된 동영상에 포함된 디졸브 개수(Total)는 17개이다. 표 3에서 검출된 디졸브 수를 Detected, Detected 가운데 진짜 디졸브의 수를 Correct로 나타내었다.

$$Recall = \frac{Correct}{Total} \times 100(\%) \quad (8)$$

$$Precision = \frac{Correct}{Detected} \times 100(\%) \quad (9)$$

표 3에 나타나듯이 제안한 방법의 경우, 분산 변화에 상관없이 모든 디졸브 구간을 검출하여 100%의 Recall을 나타냄과 동시에 정밀도를 나타내는 Precision에서 가장 높은 84.5%를 나타내었다.

5. 결론

본 논문에서는 디졸브 구간의 두 장면 사이의 상관성에 비례하는 디졸브 모델링 오차에 의한 디졸브 검출 방법을 제안하였다. 디졸브 모델링 오차는 디졸브 구간의 상관성과 분산에 대한 함수로 주어짐으로 적용적인 디졸브 검출이 가능하였다. 특히 목표 상관성을 조절함으로써 사용자가 원하는 수준의 디졸브 검출이 가능하였다. 기존의 방법과 비교하였을 때 제안된 방법이 각 구간의 분산과 상관성의 특징을 잘 반영하여 우수한 디졸브 검출 성능을 나타냄을 실험을 통하여 확인하였다. 또한 제안된 방법은 실제 디졸브 모델을 잘 반영하기 때문에 카메라 또는 물체 움직임 등으로 인해 발생하는 디졸브 검출의 오류를 줄이면서 디졸브 검출의 신뢰성을 높일 수 있었다. 다양한 디졸브를 포함하는 동영상에 대한 실험 결과 제안된 방법이 디졸브이 길이와 분산의 변화에 상관없이 디졸브 구간을 신뢰성 있게 검출하였

다. 하지만 디졸브와 유사한 포물선을 형성하는 조명의 영향 또는 물체의 움직임은 구별할 수 없었다. 향후 이러한 false detection을 제거하기 위한 연구가 계속되어야 할 것이다.

참고 문헌

- [1] J.Boreczky, L.Rowe, "Comparision of video shot boundary detection techniques," *SPIE*, **2670**: 122-128, 1996.
- [2] R. Dugad, K.Ratakonda, N. Ahuja, "Robust video shot change detection," *IEEE Second Workshop on Multimedia Signal Processing*, 1998.
- [3] S.M. Bhandarkar, A.A. Khombhadia, "Motion-based parsing of compressed video," *Proceedings of 1998 IEEE International Workshop on Multimedia Database Management Systems*, 80-87, 1996.
- [4] T.C.T. Kuo, Y.B.Lin, A.L.P. Chen, S.C.Chen, C.Y.Ni, "Efficient Shot change detection on compressed video data," *Proceedings of 1996 IEEE International Workshop on Multimedia Database Management Systems*, 101-108, 1996.
- [5] H. J. Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic partitioning of full motion video," *ACM Multimedia Systems*, **1**: pp.10-28, 1993.
- [6] H.B.Lu Y.J.Zhang, Y.R.Yao, "Robust Gradual Scene Change Detection," *ICIP 99. Proceedings*, **3**: 304-308, 1999.
- [7] 2. A. M. Alattar, "Detecting and compressing dissolve regions in video sequences with DVI multimedia image compression algorithm," *ISCAS*, 13-16, 1993.
- [8] J. Meng, Y. Juan, S. F. Chang, "Scene change detection in a MPEG compressed video sequence," *IS&T/SPIE Symposium. Proceedings*, vol. 2419, Feb. 1995.
- [9] Boon Lock Yeo, Bede Liu, "Rapid Scene Analysis on Compressed Video," *IEEE Trans. on Circuits and Systems for Video Technology*, **5**: 533-544, 1995.
- [10] R. Lienhart, "Reliable Dissolve Detection," *Processing SPIE* **4315**: 219-230, 2001.
- [11] B.T. Truong, C. Dorai, S. Venkatesh, "Improved fade and dissolve detection for reliable video segmentation," *Proceedings of 2000 IEEE International Conference on Image Processing*, **3**: 961-964, 2000.



원 중 운

1996년 2월 한국해양대학교 제어계측공학과 졸업(공학사) 1998년 2월 한국해양대학교 제어계측공학과 졸업(공학석사) 1999년 3월~현재 경북대학교 전자공학과 박사과정. 2002년 1월~현재 한국전자통신연구원 우정기술연구센터 연구원

관심분야는 영상신호처리, 영상 및 멀티미디어 통신, 영상 내용기반 검색, 워터마킹 등



오 상 근

1996년 2월 경북대학교 전자공학과 졸업 1998년 2월 경북대학교 전자공학과 석사 2002년 2월 경북대학교 전자공학과 박사

관심분야는 영상처리, 영상인식, 생체인식



최 재 각

1984년 2월 경북대학교 전자공학과 졸업(공학사). 1987년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사). 1997년 8월 한국과학기술원 전기 및 전자공학과 졸업(공학박사) 1987년 2월~1998년 2월 한국전자통신연구원 선임연구원

1998년 3월~2001년 8월 경일대학교 제어계측공학과 조교수. 2001년 9월~현재 동의대학교 컴퓨터공학과 조교수. 관심분야는 영상처리, 영상 및 멀티미디어 통신, 워터마킹 등



박 길 홈

1982년 경북대학교(공학사-전자공학) 1984년 한국과학기술원(공학석사-전자공학) 1990년 한국과학기술원(공학박사-전자공학). 1991년~현재 경북대학교 전자공학과 교수. 관심분야는 의료 영상 인식, 초음파 영상 검색, 영상 압축, 지문

인식, 얼굴 영상 인식, 디지털 워터마킹



박 철 현

1996년 2월 경북대학교 전자공학과 졸업 1999년 2월 경북대학교 전자공학과 석사 1999년 3월~현재 경북대학교 전자공학과 박사과정. 관심분야는 Biometrics, 영상처리, 영상인식



김 범 수

1997년 2월 경북대학교 전자공학과 졸업(공학사). 1999년 2월 경북대학교 전자공학과 졸업(공학석사). 2003년 2월 경북대학교 전자공학과 졸업(공학박사). 2003년 3월 현재 상주대학교 전자전기공학부 강사. 관심분야는 영상 내용기반 검색, 영상

분할, 영상 워터마킹 등



곽 동 민

1997년 2월 경북대학교 전자공학과 졸업(공학사). 1999년 2월 경북대학교 전자공학과 졸업(공학석사) 1999년 3월~현재 경북대학교 전자공학과 박사과정. 관심분야는 영상 내용기반 검색, 영상 분할, 영상 워터마킹 등