

# DWT영역에서 에지 성분을 이용한 효과적인 Dissolve 검출

김 운, 이 배 호

전남대학교 컴퓨터공학과

전화 : (062) 530-0719 / 팩스 : (062) 530-1809

## The Efficient Dissolve Detection using Edge Elements on DWT Domain

Woon Kim, Bae-Ho Lee

Department of Computer Engineering, Chonnam National University

E-mail : woonida@netian.co.kr

### Abstract

There are many problems such as low detection ratio, velocity and increase of false hit ratio on the detection of gradual scene changes with the previous shot transition detection algorithms. In this paper, we propose an improved dissolve detection method using color information on low-frequency subband and edge elements on high-frequency subband.

The possible dissolve transition are found by analyzing the edge change ratio in the high-frequency subband with edge elements of each direction. Using the double chromatic difference on the lowest frequency subband, we have the improvement of the dissolve detection ratio. The simulation results show that the performance of the proposed algorithm is better than the conventional one for dissolve detection on a diverse set of uncompressed video sequences.

### I. 서론

최근 컴퓨터와 방송 등 정보 서비스의 발달과 통합된 통신망 및 대용량의 저장 매체 등의 발달로 인하여 디지털 멀티미디어 정보가 매일 엄청난 양으로 증가되고 있으며 데이터의 양 또한 과거의 텍스트 정보와 비교 할

수 없을 정도이다. 이런 멀티미디어 자료 중에서 각종 비디오 데이터는 동영상과 함께 오디오 및 텍스트 정보를 포함하고 있는 복합적 성격의 데이터로서 영상 매체의 발달과 더불어 그 중요성이 점차 증가하고 있다.

비디오 데이터를 효과적으로 관리하고 검색해 내기 위한 가장 기본이 되는 과정은 서로 유사한 데이터들의 집합을 찾아내고 분리해 내는 세그멘테이션 작업이다. 이 작업을 위해 장면 전환을 정확하게 검출할 필요성이 있으며 이에 대한 방법에는 여러 개의 부분적인 알고리즘이 적용된다. 이는 장면 전환의 지속성에 따라 급격한 장면 전환과 점진적인 장면 전환으로 나뉜다. 급격한 장면 전환은 비연속성을 지닌 컷(cut)을 나타내고, 점진적인 장면 전환에는 페이드(fade), 디졸브(dissolve), 와이프(wipe) 변화를 포함하는 연속적인 장면 전환을 일컫는다.

장면 전환에 대한 정의는 다음과 같다.

- Cut : 갑작스런 장면 변화(scene change)나 급격한 장면변화(abrupt shot change)
- Dissolve : 장면이 선형적으로 바뀌는 장면 변화
- Fade : 장면이 서서히 밝아지거나, 반대로 어두워지는 장면 변화
- Wipe : 한 장면이 사라지기 전에 다른 장면이 점진적으로 가로질러 들어오는 장면 변화

좋은 장면 전환검출 알고리즘은 급격한 장면 전환이나 점진적인 장면전환을 포함한 어떠한 종류의 장면 전환에도 민감해야한다. 그리고 장면 전환을 제외한 다른 변

환(비디오 내의 물체나 카메라의 움직임)에서는 반응이 없어야 한다.

이에 본 논문은 비디오 영상의 프레임을 다중해상도 부대역으로 분해하기 위해 전처리 과정으로 3-레벨 웨이브릿 변환을 이용하여 최저주파 영역의 색상 정보와 고주파 영역의 에지 성분을 이용한 장면 전환 검출 기법 중 개선된 디졸브 검출 방법을 제안한다.

본 논문의 구성은 II장에서 기존의 장면 전환 검출 알고리즘에 대해서 살펴보고, III장에서는 다해상도에 의한 장면검출 결과를 알아본 후, IV장에서는 7개의 디졸브를 포함한 영상 시퀀스에 대한 실험 결과를 보이고, 마지막으로 V장에서는 제안된 방법에 대한 결론을 맺는다.

## II. 기존의 장면 전환 검출기법

최근 장면 전환을 검출에 대한 논문을 살펴보면 검출 방식에 있어서 크게 두 부류로 나뉜다. 복원된 영상 시퀀스의 장면전환 추출방법과 압축된 영상 시퀀스의 장면전환 추출방법으로 나누어진다.

복원된 영상 시퀀스에서 장면전환을 추출하는 대표적인 방법으로는 주로 화소단위 비교법, 유사율 측정법과 히스토그램 비교법 등이 있다. Zabih[1]는 cut, fade, dissolve, wipe를 포함한 다양한 장면 전환을 검출하기 위해 에지 추적방법을 도입하였다. 이 논문에서 제시한 에지 추적방법은 우수하지만 점진적인 전환의 16%만 검출이 가능하고 움직임 추정의 과정이 필요하며 강한 움직임에는 영향을 받는 단점이 있다. Hampapur[2]는 공간적 변화를 검출하기 위해 model-driven 접근법을 제시하였지만 이는 픽셀의 밝기가 변하지 않는다는 가정 아래 적용된 알고리즘이다.

압축된 영상 시퀀스에서 접근하는 방법, Yeo와 Liu [3] 방법은 각 프레임의 DC성분 계수 값만으로 시퀀스를 만들어 각 해당 블록의 DC값의 SAD(Sum of Absolute Difference)를 비교하여 프레임간의 유사성을 비교하는 방법, 컬러 히스토그램 비교법, pixel-level의 차이를 비교하는 방법, 정규화 된 내부 벡터를 사용하는 방법들이 제시되었다. 이는 급격한 장면 변화 검출에서 계산적으로 빠르지만, 많은 동영상에서는 (특히 영화) 많은 수의 fade-in/out, dissolve와 같은 점진적인 전환을 포함하고 있어서 아직까지는 점진적인 장면 전환을 검출하는 데는 어려움이 따른다는 단점이 있다.

이렇게 점진적인 장면 전환은 몇 가지의 다른 장면 변화와 같이 존재하는데 이로 인해 검출 시 검출율이 떨어지고, 점진적인 장면전환이 아닌 구간을 점진적인 장면 전환으로 오인하여 검출하는 경우 또한 발생한다. 이러한 단점과 한계를 극복하기 위해 본 논문에서는 최근 영

상처리에 새로운 변환 기법인 웨이브릿 변환(WT: Wavelet Transform)을 적용하였다. 그림 1은 본 논문에서 제안한 시스템 구성도이다.

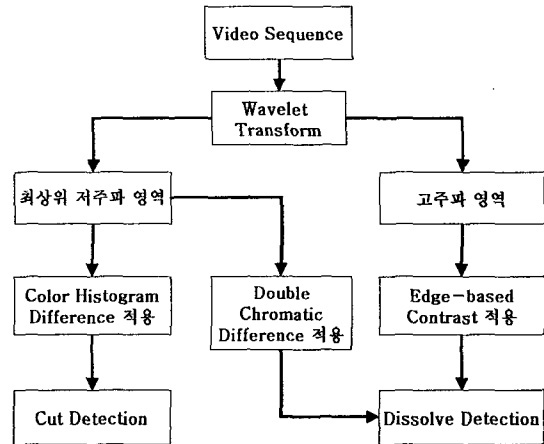


그림 1. 제안한 시스템 구성도

## III. 다해상도 장면 전환 검출기법

WT은 Grossmann과 Morlet에 의하여 발표[4]되었으며, Mallat은 WT에 의한 다해상도(Multiresolution) 피라미드(pyramid) 영상과, 다중 스케일(multiscale) 에지 정보 결합에 의한 에지 검출에 대하여 연구하였다.[5][6] 그림 2은 2차원 영상에 대한 웨이브릿 변환을 나타낸 것이다.

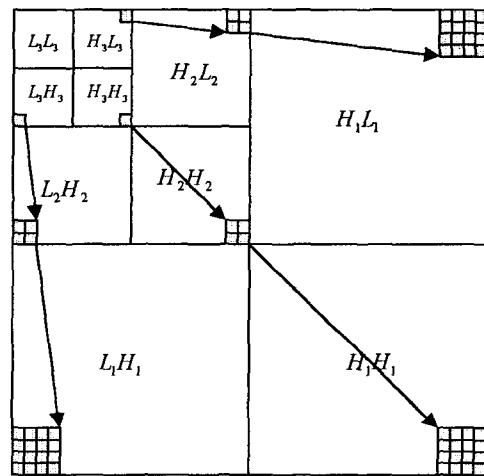


그림 2. 2차원 영상에 대한 웨이브릿 변환

WT는 영상을 피라미드 구조로 표현하는 기존의 방식인 다해상도 분석(multi-resolution analysis)과 대역분할

부호화를 하나로 통합한 변환방식이다. 이는 고주파 대역에서는 시간 분해능을 높이고, 저주파 대역에서는 주파수 분해능을 높이는 옥타브 대역 분할을 통해 시간과 주파수에 대한 국부성을 가지고 신호를 표현하므로 저주파 성분이 많은 일반 영상신호의 분석에 유리하다. 또한 이 표현 방법이 인간의 시각 체계와 유사하다는 장점으로 인해 최근 영상처리분야에서 각광을 받고 있고, 이를 이용하여 보다 적절하게 작업을 수행 할 수 있다.

WT는 필요한 각 특징 계수를 추출하기 위해 저주파수와 고주파수 정보로 분해한다. 본 논문에서는 처음에 각 동영상 프레임에 계산이 가장 빠르고 구현하기에 매우 간단한 2D Haar DWT을 이용하여 부대역으로 3-레벨 분해한다.

주어진 영상을 일정한 계층으로 분할 한 후, 저주파 성분은 원 영상의 대부분의 색상정보를 가지고 있으므로 장면전환의 특성을 가지고 있다. 이와 같은 특징을 이용하여 분할 된 영역 중 최저주파수 영역( $L_3, L_3$ )에 화소 단위 비교법 중의 컬러 히스토그램 비교법(CHD: Color Histogram Difference) 식(1)을 적용하여 컷을 검출한다.

$$CHD_i = \frac{1}{N} \sum_{r=0}^{2^B-1} \sum_{g=0}^{2^B-1} \sum_{b=0}^{2^B-1} |P_i(r, g, b) - P_{i-1}(r, g, b)| \quad (1)$$

여기서  $P_i(r, g, b)$ 은 N 개의 픽셀을 가진  $f_i$  프레임에서의 컬러 픽셀 수를 나타내고,  $r, g, b \in [0, 2^B-1]$ 이며 B값은 노이즈에 대한 민감성을 줄이기 위해 3의 값을 갖는다.

이렇게 검출된 영역에서 디졸브를 검출하는데 여기에서 이용하는 방법은 웨이브렛 변환대역에서의 중요계수는 원 영상에서의 에지 성분과 밀접한 연관이 있다는 사실에 근거[8]를 두고 에지 특성을 이용하였다. 에지 정보는 영상 내에서 변화 정도가 일정 크기 이상인 화소점 또는 대역에서 에지를 찾아내며, 이러한 에지는 실제로 웨이브렛 대역에서 큰 계수값을 갖게 되며, 영상의 의미있는 정보를 제공할 뿐 아니라 시각적으로도 매우 중요한 역할을 한다. 공간영역에서의 에지는 웨이브렛 변환 후, 중요계수로 검출될 뿐 아니라 이 계수들은 같은 방향(수직, 수평, 대각)을 따라 나타난다는 사실에 기반을 두고 있다.

그리하여 분할 영역 중 고주파 성분(High Frequency Band : HH, LH, HL)에서 각 방향의 에지 성분을 이용하여 프레임간 Contrast 비교법(EC: Edge-based Contrast) 식(2)을 적용하여 디졸브 가능 영역을 찾는다.

$$EC(K) = 1 + \frac{s(K) - w(K) - 1}{s(K) + w(K) + 1}, EC(K) \in [0, 2] \quad (3)$$

$$w(K) = \sum_{x,y} W_k(x, y) \quad \text{and} \quad s(K) = \sum_{x,y} S_k(x, y)$$

$$W_k(x, y) = \begin{cases} K(x, y) & \text{if } \theta_w \leq K(x, y) < \theta_s \\ 0 & \text{else} \end{cases}$$

$$S_k(x, y) = \begin{cases} K(x, y) & \text{if } \theta_s \leq K(x, y) \\ 0 & \text{else} \end{cases}$$

$K(x, y, t)$  : 프레임  $f_i$ 에 대한 edge map

$\theta_w, \theta_s$  : weak, strong threshold

$w(K), s(K)$  : weak, strong edges

하지만, 디졸브 검출 시 나타나는 다른 문제는 비디오 내의 큰 물체나 카메라의 움직임이 잘못된 검출을 야기시키는 것이다. 여기서 본 논문은 Yu[9]가 제안한 이중색채 대조방법(DCD: Double Contrast Difference) 식(3)을 최저주파 성분에 적용함으로써 에러요인을 제거, 에러율을 낮추었다.

$$\begin{aligned} DCD(t) &= \sum_{x,y} F \left( \left| \frac{f(x, y, t_0) + f(x, y, t_N)}{2} - f(x, y, t) \right| \right) \\ &= \sum_{x,y} \left| \frac{\alpha(t_0) + \alpha(t_N)}{2} - \alpha(t) \right| \times |p(x, y) - q(x, y)| \\ &= \left| \frac{\alpha(t_0) + \alpha(t_N)}{2} - \alpha(t) \right| \times \sum_{x,y} |p(x, y) - q(x, y)| \quad (3) \end{aligned}$$

여기서  $f(x, y, t) = \alpha(t)p(x, y, t) + \beta(t)q(x, y, t)$ 이고 디졸브 검출을 향상시키기 위해 다음과 같은 가정을 따른다.

$$\alpha(t) + \beta(t) = 1$$

$$p(x, y, t) = p(x, y) \quad \text{and} \quad q(x, y, t) = q(x, y)$$

( $\alpha(t)$  : 감소함수,  $\beta(t)$ : 증가함수)

#### IV. 실험결과

실험을 위해, 복원된 상태(비 압축영역)의 비디오 시퀀스에서 7개의 디졸브를 갖고 있는 영상클립에서 테스트 되었으며, 실제 디졸브가 존재하는 프레임 수와 제안한 알고리즘을 이용하여 디졸브를 검출한 결과는 다음 표 1과 같다. 그림 3은 알고리즘을 영상클립에 적용하여 얻은 결과를 프레임간 Contrast 비교법(EC)을 적용하여 그래프로 나타내었다.

표 1과 그림 3에서와 같이 총 7개의 디졸브가 존재하는데, 프레임간 Contrast 비교법을 적용하여 검출한 결과, 4개 밖에 검출하지 못하였다. 이는 디졸브 구간에서 카메라의 움직임(pan, zoom)이 있었고, 화면 내에서 움직이는 사람들이 많았기 때문에 이 구간의 그래프를 보면 plateau 가 있음을 알 수 있다. 이러한 구간을 제대로 검출하지 못하였는데 DCD방법을 이용하여 다시 검

디졸브	실제구간	EC	DCD
1	83-97	X	X
2	170-190	X	O
3	290-310	O	O
4	362-385	O	O
5	550-570	O	O
6	730-752	X	O
7	915-935	O	O

표 1. 디졸브 실제 위치와 검출 결과  
(O : 검출, X : 비검출)

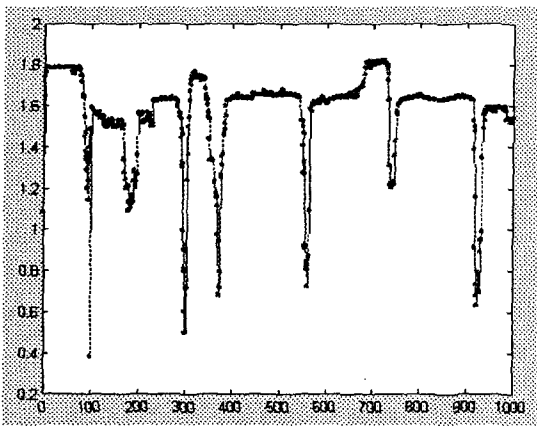


그림 3. 제안한 알고리즘을 적용한 결과

출한 결과, 한곳을 제외한 모든 곳을 검출해 냈다. 실제로 85프레임 구간에서는 디졸브 후 카메라의 움직임이 있었으며 곧바로 급격한 장면 전환이 일어난 것이었다.

### V. 결론

본 논문에서는 웨이블릿 변환 영역에서 동영상에 대한 장면 전환 검출기법을 제시하였다. 이 방법은 공간 영상에서 영상의 중요한 특성을 제공하는 에지 정보가 웨이블릿 변환 영역에서 중요 계수로 나타난다는 사실에 근거를 두고 있다. 제안한 방식은 알고리즘이 단순할 뿐 아니라 기존의 방식에 비하여 정확한 검출 결과를 얻을 수 있었다.

다양한 데이터를 통한 실험으로 제안한 알고리즘에 대한 평가와 이에 대한 좀더 효율적인 알고리즘에 대해 연구할 것이다.

### 참고문헌

- [1] R. Zabih, J. Miller, and K. Mai, "A feature-based algorithm for detecting and classifying scene breaks", Proceedings of ACM Multimedia'95, San Francisco, CA, pp. 189-200, 1993.
- [2] A. Hampapur, R. Jain, and T. Weymouth, "Digital video segmentation", Proceedings of ACM Multimedia'94, pp. 357-364, 1994.
- [3] B. L. Yeo and B. Liu, "Rapid scene analysis on compressed video", IEEE Trans. on circuit and systems for video technology, vol. 5, no. 6, pp. 533-544, Dec. 1995.
- [4] A. Grossmann and J. Morlet, "Decomposition of hardy-functions into square integrable wavelets of contrast shape", SIAM J. Math., vol. 15, pp. 723-736, 1984.
- [5] Stephane Mallat, "Wavelets for a Vision", Proc. of the IEEE, vol. 84, no. 4, pp. 604-614, 1996.
- [6] Ruskai, et al., *Wavelets and their applications*, Jones and Bartlett, pp. 67-104, 1992.
- [7] B. Zhu, A. H. Tewfik, O. N. Gerek, "Image coding with mixed representations and visual masking", ICASSP95, pp. 2327-2330, 1995.
- [8] Martin Vetterli and Jelena Kovacevic, *Wavelets and Subband Coding*, Prentice-Hall Inc., London, 1995.
- [9] H. Yu, G. Bozdage, S. Harrington, "Feature-based video segmentation", ICIP'97, pp. 498-501, 1997.
- [10] Rainer Lienhart, "Comparison of automatic shot boundary detection algorithms", SPIE Vol. 3656, pp. 290-301, 1999.
- [11] H. B. Lu, Y. J. Zhang, "Detection abrupt scene change using neural network", Proceeding of Third International Conference, VISUAL'99, pp. 291-298, 1999.