

확신도를 이용한 동영상의 화면변환 감지

장석우^{1*}

¹안양대학교 디지털미디어학과

Detecting Shot Boundaries of Dynamic Images Using Certainty Factors

Seok-Woo Jang^{1*}

¹Department of Digital Media, Anyang University

요약 본 논문에서는 비디오 데이터에서 확신도 추론을 이용하여 급진적인 화면변환과 점진적인 화면변환을 동시에 검출하는 새로운 방법을 제안한다. 제안된 방법에서 급진적인 화면변환은 컷을 의미하고, 점진적인 화면변환은 페이드 인, 페이드 아웃, 디졸브, 수평 와이프, 수직 와이프, 반도어(Barn Doors), 아이리스 라운드(Iris Round)를 의미한다. 먼저, 각 화면변환을 대표하는 특징을 추출한 후 확신도를 이용하여 특징들을 효과적으로 통합하면서 발생한 화면변환을 추출한다. 실험결과에서는 제안된 방법이 기존의 방법에 비해 여러 가지 종류의 화면변환을 보다 정확하게 추출함을 다양한 비디오 데이터를 이용한 실험을 통해 보여준다.

Abstract In this paper, we propose a new method to detect abrupt and gradual shot transitions of video data by using certainty factors. The abrupt transitions denotes cuts and the gradual transitions fade in, fade out, dissolve, horizontal wipes, vertical wipes, Barn Doors, and Iris Rounds. The suggested method first extracts representative features for each shot transition and determines corresponding shot transitions by integrating all the extracted features and inferring adequate transitions. To verify the performance of the proposed shot transition method, experimental results show that the suggested method can detect shot transitions more accurately than existing methods.

Key Words : Feature Extraction, Abrupt Transition, Gradual Transition, Wipes, Certainty factor

1. 서론

디지털 카메라, 테블릿 PC, 스마트 폰 등의 저렴한 동영상 캡처 기기, 그리고 초고속 유무선 네트워크 기술 등의 빠른 발달로 동영상은 점점 더 대중화되고 있으며, 관련된 동영상 데이터의 양도 기하급수적으로 증가하고 있다. 따라서 이런 동영상 데이터를 효과적으로 관리하는 것이 필요한데, 이를 위해서는 먼저 동영상의 장면 사이의 경계를 나타내는 샷 전환(shot transition)을 검출하는 것이 중요하다[1].

일반적으로, 샷 전환은 크게 급진적인 샷 전환과 점진적인 샷 전환으로 분류될 수 있다[2]. 급진적인 샷 전환에는 컷(cut)이 있으며, 점진적인 샷 전환에는 페이드 인과

아웃(fade in/out), 디졸브(dissolve), 와이프(wipe) 등이 있다. 그리고 와이프는 이동 경계선이 수평 방향으로 이동하는 수평 와이프, 수직 방향으로 이동하는 수직 와이프, 영상의 중앙에서 마치 커튼을 양쪽으로 걷듯이 이동하는 반도어(Barn Door) 와이프, 이동 경계선이 원형의 모양을 가지는 아이리스 라운드(Iris Round) 와이프 등이 있다.

문헌에서는 동영상으로부터 샷 전환을 검출하는 여러 가지 방법들이 소개되어 있다[3-6]. [3]에서는 명암값 또는 색상 히스토그램을 이용하여 컷 위주의 샷 전환을 검출하는 방법을 제안하였다. [4]에서는 에지를 추출하고 새롭게 출현하는 에지와 기존에 존재하다가 사라지는 에지의 개수를 비교하여 샷 전환을 검출하는 방법을 제안하였다. [5]에서는 블록 단위의 칼라와 에지 특징을 추출하고, 이

*교신저자 : 장석우(swjang@anyang.ac.kr)

접수일 11년 11월 08일

수정일 11년 11월 30일

게재확정일 11년 12월 13일

를 통합하여 샷 전환을 검출하는 방법을 제안하였다. [6]에서는 동영상으로부터 차영상을 추출하고, 추출된 차영상으로부터 무게중심을 계산한 후, 이 중심위치의 궤적정보를 분석하여 이동 경계선이 선형인 와이프를 검출하였다. 그리고 이런 방법들 이외의 다른 방법들도 계속해서 소개되고 있다[7].

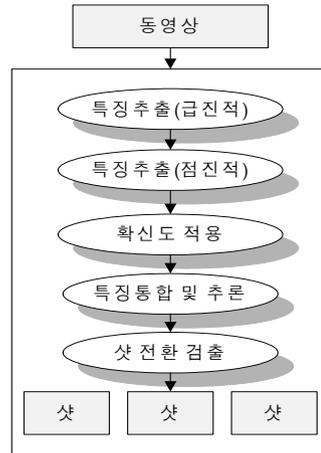
위와 같이 샷 전환 검출과 관련된 많은 기존의 방법들이 소개되어 있으나, 이들 대부분은 여러 샷 전환 중에서 특정한 종류만을 그 검출 대상으로 하고 있다. 즉, 대부분의 기존의 방법들은 급진적인 샷 전환인 컷 검출을 위주로 수행되고 있다. 또는 점진적인 전환 중에서 특정 전환, 예를 들어 페이드나 디졸브, 아니면 와이프 중에서도 이동 경계선이 직선인 와이프만을 주 대상으로 하고 있다.

따라서 본 논문에서는 기존 방법의 이런 제한점을 해결하기 위해서 확신도(certainty factors) 추론을 이용하여 표 1과 같이 점진적인 전환인 컷과 급진적인 전환인 페이드 인과 아웃, 디졸브, 수평 와이프, 수직 와이프, 반 도어, 아이리스 라운드 와이프를 모두 검출하는 새로운 방법을 소개한다.

[표 1] 샷 전환 검출 대상
[Table 1] Shot transitions to detect

샷 전환		설명
점진적 전환	컷	장면이 순간적으로 변함
급진적 전환	페이드 인	고정영상이 점진적으로 장면화 됨
	페이드 아웃	영상이 점진적으로 고정영상으로 변함
	디졸브	페이드 인과 페이드 아웃이 교차해서 발생
	수평 와이프	이동 경계선이 직선이고, 수평 방향으로 이동
	수직 와이프	이동 경계선이 직선이고, 수직 방향으로 이동
	반 도어	두 이동 경계선이 직선이고, 양쪽 방향으로 이동
아이리스 라운드	원형인 이동 경계선이 점점 크게 변함	

그리고 그림 1은 본 논문에서 제안하는 확신도를 이용한 샷 전환 검출 알고리즘의 전체적인 구조도를 보여준다. 그림 1에서 확인할 수 있듯이 제안된 방법은 먼저 연속적으로 입력되는 동영상을 받아들인 후 급진적인 샷 전환과 점진적인 샷 전환을 대표하는 각각의 특징들을 추출한다. 그런 다음, 확신도를 적용하여 여러 특징들을 통합하면서 해당되는 샷 전환의 발생 유무 및 그 종류를 효과적으로 추론한다.



[그림 1] 전체 구조도
[Fig. 1] Overall flow diagram

2. 급진적인 샷 전환의 특징

본 논문에서는 샷 전환의 특징을 추출하기 위해 먼저 입력된 영상의 R, G, B 칼라공간을 인간의 시각 시스템과 가장 유사한 모델이라고 알려진 색상, 채도, 명도로 표현되는 HIS 칼라공간으로 변형하여 사용한다[8]. 식 (1)은 R, G, B 칼라공간을 H, I, S 칼라공간으로 변환하는 식을 보여준다.

$$H = \cos^{-1} \left(\frac{[(R-G) + (R-B)]}{2 \sqrt{(R-G)^2 + (R-B)(G-B)}} \right) \quad (1)$$

$$I = \frac{1}{3} \times (R + G + B)$$

$$S = 1 - \frac{3}{(R + G + B)} \times [\min(R, G, B)]$$

그런 다음, 영상을 모양과 크기가 동일한 N×N 화소로 구성된 사각형의 블록으로 분할한 후 각 샷 전환의 특징을 추출한다. 화소 단위가 아니라 블록 단위로 장면의 특징을 추출하는 이유는 화소 단위의 지역적인 잡음의 영향을 최소화하기 위함이다.

여러 가지 샷 전환 중에서 급진적인 샷 전환인 컷은 인접한 영상 사이에서 장면이 순간적으로 변화하는 전환이다. 따라서 컷을 기준으로 두 인접 영상 사이의 칼라 유사도는 매우 낮다. 본 논문에서는 이런 특성을 고려하기 위해서 식 (2)와 같이 두 인접 영상 사이의 블록 단위의 색상과 명암값의 상관관계(correlation)를 이용하여 컷 특징을 정의한다.

$$F_{cut} = \alpha \times Corr(BIM_{t-1}, BIM_t) + \beta \times Corr(BHM_{t-1}, BHM_t) \quad (2)$$

where $0 < \alpha \leq \beta < 1, \alpha + \beta = 1$

식 (2)에서 컷 특징 F_{cut} 은 현재 시점과 이전 시점 영상 사이의 색상과 명암값의 상관관계[9]를 나타내는 $Corr(BHM_{t-1}, BHM_t)$ 과 $Corr(BIM_{t-1}, BIM_t)$ 으로 정의되며, F_{cut} 은 0에서 1사이의 범위를 갖도록 정규화된다. 즉, F_{cut} 이 0에 근접한 값을 가진다면 인접한 두 영상 사이의 칼라 값의 상관관계는 낮다는 의미이며, F_{cut} 이 1에 근접한 값을 가진다면 두 영상 사이의 칼라 값의 상관관계는 높다는 의미이다. 따라서, 컷 특징 F_{cut} 은 0에 근접할수록 컷일 확률이 높고, 반대로 1에 근접할수록 컷일 확률이 낮다. 그리고 BHM_{t-1} 과 BIM_t 는 현재 시점의 i 번째 블록의 색상 평균과 명암값 평균을 의미한다. 그리고 α 와 β 는 색상과 명암값에 대한 가중치로 사용하는데 본 논문에서는 색상에 보다 높은 가중치를 부여한다.

3. 점진적인 샷 전환의 특징

3.1 페이드 특징

페이드 변환은 고정영상이 점차 밝아지면서 장면으로 전환되는 페이드 인과 장면이 점차 고정영상으로 변화하는 페이드 아웃 변환이 있다. 보통, 페이드가 발생하는 동안 영상의 명암은 일정한 값 δ 만큼 곱해진다. 페이드 인은 δ 를 0에서 1로 증가시키며, 반면에 페이드 아웃은 δ 를 1에서 0으로 감소시킨다. 상수 δ 가 변하는 속도에 따라 페이드 전환의 속도도 조절된다. 즉, 페이드가 발생하는 동안 영상의 전체 명암값은 일정한 상수로 점진적으로 변화된다. 본 논문에서는 그런 변화를 검출하기 위해서 식 (3)과 같이 전체 명암 변화의 비율을 정의한다. 식 (3)은 연속적인 영상의 명암값들이 시간의 경과에 따라 어떻게 변화하는지를 평가하는 것으로, I 는 명암값을 의미하고, I_{max} 는 명암값의 최대값을 의미한다. 그리고 M 과 N 은 영상의 가로 및 세로의 크기를 의미한다.

$$F_{fade} = \frac{D_{diff}}{D_{Adiff}} \quad (3)$$

$$D_{diff} = \frac{\sum_{i=1}^M \sum_{j=1}^N (I(i,j,t) - I(i,j,t-1))}{I_{max} \cdot M \cdot N}$$

$$D_{Adiff} = \frac{\sum_{i=1}^M \sum_{j=1}^N |I(i,j,t) - I(i,j,t-1)|}{I_{max} \cdot M \cdot N}$$

위의 명암 비율은 -1에서 1사이의 범위의 값을 가지는데 프레임이 밝아지는지 또는 어두워지는지를 나타낸다. 이것은 페이드 아웃 동안 음의 값을 가지고 페이드 인 동안 양의 값을 가지며, 페이드가 발생하는 동안 1에 근접한 크기를 가진다.

3.2 디졸브 특징

디졸브 샷 변환은 한 화면이 사라짐과 동시에 다른 화면이 점차로 나타나는 장면 전환으로 페이드 인과 페이드 아웃이 교차하는 화면전환을 의미한다. 그리고 디졸브 화면전환이 발생하는 동영상에서는 장면 사이의 색상, 명암값, 채도가 점진적으로 변화하는 현상이 발생한다. 따라서 본 논문에서는 색상, 명암값, 채도의 차이를 모두 이용하여 디졸브 특징을 식 (4)와 같이 정의하는데, 식 (4)에서 H 와 S 는 색상값과 채도값을 의미하고, H_{max} 와 S_{max} 는 색상값과 채도값의 최대값을 의미한다.

$$F_{dis} = \alpha \cdot H_{Adiff} + \beta \cdot I_{Adiff} + \gamma \cdot S_{Adiff} \quad (4)$$

$$H_{Adiff} = \frac{\sum_{i=1}^M \sum_{j=1}^N |H(i,j,t) - H(i,j,t-1)|}{H_{max} \cdot M \cdot N}$$

$$S_{Adiff} = \frac{\sum_{i=1}^M \sum_{j=1}^N |S(i,j,t) - S(i,j,t-1)|}{S_{max} \cdot M \cdot N}$$

where $0 < \beta \leq \alpha, \gamma < 1, \alpha + \beta + \gamma = 1$

식 (4)에서 디졸브 특징 F_{dis} 는 현재 시점과 이전 시점 영상 사이의 색상의 절대값 차이, 명암값의 절대값 차이, 그리고 채도의 절대값 차이의 합으로 정의되며, 0에서 1사이의 범위를 가지도록 정규화된다. 그리고 디졸브 특징 F_{dis} 이 1에 근접할수록 디졸브일 확률이 높으며, 반대로 0에 근접할수록 디졸브일 확률이 낮다.

3.3 와이프 특징

본 논문에서는 여러 가지 와이프 화면전환 중에서 이동 경계선(moving border)의 형태가 선형 라인의 모양을 가지는 수평 와이프, 수직 와이프, 반도어 와이프, 그리고 이동 경계선의 형태가 원의 모양을 가지는 아이리스 라운

드 와이프를 검출한다.

제한된 방법에서는 와이프를 검출하기 위해서 먼저 인접 영상 사이에 차연산을 적용하고 이를 이진화하여 차영상을 구한다. 그런 다음, 이진화된 차영상에 허프 변환(Hough transform)을 적용하여 영상에 존재하는 라인과 원을 검출한다. 그리고 원형 이동 경계선을 검출하는 경우에는 이 경계선의 에지는 완전한 원형으로 표현되지 않고 일부만 나타나는 경우가 많기 때문에 일반적인 허프 변환을 사용하지 않고 원형의 에지가 일부만 나타나도 원의 형상을 검출할 수 있는 개선된 허프 변환을 사용한다.

보통 와이프는 한 프레임의 크기가 $M \times N$ 인 k 개의 프레임에 걸쳐 점진적으로 발생한다. 만일, k 개의 프레임에 걸쳐 검출된 이동 경계선의 무게중심의 좌표를 $\{(C_{x_1}, C_{y_1}), (C_{x_2}, C_{y_2}), \dots, (C_{x_k}, C_{y_k})\}$, 그리고 원형 이동 경계선이 존재할 경우 원의 반지름의 길이를 $\{r_1, r_2, r_3, \dots, r_k\}$ 라 할 때 본 논문에서는 수평 와이프의 특징을 식 (5)와 같이 정의한다.

$$F_{hwipe} = 1/k \cdot \sum_{i=1}^k |C_{dx}(i)| - M/k \quad (5)$$

$$C_{dx}(i) = C_{x_i} - C_{x_{i-1}}$$

식 (5)의 F_{hwipe} 는 0에서 1사이의 값으로 정규화된다. 즉, F_{hwipe} 의 값이 0보다 작으면 0으로, 그리고 1보다 크면 1의 값으로 설정된다. 일반적으로, k 개의 프레임에 걸쳐 발생하는 이상적인 양의 방향 수평 와이프의 경우 검출된 라인의 x 좌표가 시간에 따라 M/k 만큼 이동해야 한다. 따라서 F_{hwipe} 가 0에 가까울수록 수평 와이프가 발생할 확률이 높다.

수평 와이프의 특징과 유사하게 수직 와이프 특징은 이동 경계선의 무게중심의 y 좌표를 이용하여 식 (6)과 같이 정의하며, 0에서 1사이의 값을 가진다.

$$F_{vwipe} = 1/k \cdot \sum_{i=1}^k |C_{dy}(i)| - N/k \quad (6)$$

$$C_{dy}(i) = C_{y_i} - C_{y_{i-1}}$$

반도어 와이프의 경우에는 라인 형태의 이동 경계선이 2개이며, 중앙으로부터 좌우 방향으로 각 라인이 동일한 속도로 이동한다. 따라서 이런 특성을 충분히 이용하여 반도어 와이프의 특징을 식 (7)과 같이 정의할 수 있다. 식 (7)에서 $C_x^L(i)$ 은 두 개의 라인 중에서 왼쪽에 위치한 라인을 의미하며, $C_x^R(i)$ 은 오른쪽에 위치한 라인을 의미

한다. 식 (7)의 F_{barn} 도 0에서 1사이의 값으로 정규화되며, 0에 가까울수록 반도어 와이프가 발생할 확률이 높다.

$$F_{barn} = \sum_{i=1}^{k/2} |C_x^L(i) + C_x^R(i)| \quad (7)$$

이동 경계선의 형태가 원형인 아이리스 라운드 와이프가 발생하는 동안에는 원형 이동 경계선의 중심은 중앙에 위치하고, 원의 반지름 r 이 시간에 따라 $M/2k$ 크기씩 일정한 속도로 이동한다. 이런 특성을 반영하기 위해 본 논문에서는 아이리스 라운드의 특징을 식 (8)과 같이 정의한다.

$$F_{iris} = 1/2k \cdot \sum_{i=1}^k |r_{dx}(i)| - M/2k \quad (8)$$

$$r_{dx}(i) = r_{x_i} - r_{x_{i-1}}$$

식 (8)의 F_{iris} 도 0에서 1사이의 값으로 정규화되는데, F_{iris} 가 0에 가까울수록 아이리스 라운드 와이프가 발생할 확률이 높다.

4. 확신도 추론

특징을 추출한 후에는 이 특징을 이용해 각 화면전환을 효과적으로 나타낼 수 있는 규칙을 생성하고 불확실성이론인 확신도(certainty factor)를 이용해 장면전환을 추론한다. 확신도를 이용한 불확실성 처리 방법은 스탠포드 대학에서 1970년대 초부터 약 10년간에 걸쳐 개발한 시스템인 MYCIN을 개발하는 과정에서 제안된 방법이다[10].

본 논문에서 사용하는 확신도는 증거(evidence)에 대한 가설(hypothesis)의 확신 정도로서 식 (9)와 같이 증거에 대한 가설의 확신과 증거에 대한 가설의 불확신의 차이로서 정의된다.

$$CF[h, e] = MB[h, e] - MD[h, e] \quad (9)$$

식 (9)에서 CF 는 증거 e 가 주어진 경우 가설 h 에 대한 확신도를 나타내며, -1과 1 사이의 범위를 가진다. MB (measure of belief)는 증거 e 가 주어진 경우 가설 h 에 대한 확신의 정도로서 가설에 대한 증거의 지지 정도를 나타낸다. MD (measure of disbelief)는 증거 e 가 주어진 경우 가설 h 에 대한 불신의 정도로서 가설에 대한 증거의 부정 정도를 나타낸다. 그리고 MB 와 MD 는 모두 0에서 1사이

의 범위를 가진다.

식 (10)은 확신도를 규칙에 적용한 예를 보여준다. 식 (10)에서 규칙의 조건부를 증거로 판단하고, 규칙의 결론부를 가설이라고 해석한다.

$$\text{IF } \langle \text{evidence} \rangle \text{ THEN } \langle \text{hypothesis} \rangle \text{ CF}[h,e] \quad (10)$$

식 (10)에서 증거 e에 대한 확신도가 주어진 경우 증거에 대한 가설의 확신도, 즉 규칙에 대한 확신도를 CF[h, e]로 표현하고 있다. 일반적으로, 확신도를 규칙에 적용한 경우 결론부의 확신도는 조건부의 확신도와 규칙에 대한 확신도의 곱으로 계산된다. 본 논문에서는 확신도를 이용하여 장면전환을 추론하기 위해 아래와 같이 9가지의 가설을 설정한다.

- (1) 가설 H1: 인접한 영상이 유사하다(STAY).
- (2) 가설 H2: 컷이 발생한다(CUT).
- (3) 가설 H3: 페이드 인이 발생한다(FADE IN).
- (4) 가설 H4: 페이드 아웃이 발생한다(FADE OUT).
- (5) 가설 H5: 디졸브가 발생한다(DISSOLVE).
- (6) 가설 H6: 수평 와이프가 발생한다(HWIPE).
- (7) 가설 H7: 수직 와이프가 발생한다(VWIPE).
- (8) 가설 H8: 반도어가 발생한다(BARN).
- (9) 가설 H9: 아이리스 라운드가 발생한다(IRIS).

특징들을 통합하여 화면변환에 대한 규칙을 생성하려면 각 특징의 확신도를 결정해야 한다. 본 논문에서는 실험적인 결과를 이용해 각 특징의 확신도를 그림 2와 같이 정의한다. 컷, 디졸브, 와이프 특징에 대한 확신도는 두 가지 레벨 Low, High로 정의하며, 페이드 특징의 확신도는 세 가지 레벨 Low, Medium, High로 정의한다. 그리고 추출한 특징들을 통합하여 생성하는 가설의 규칙은 식 (11)과 같다.

$$\begin{aligned} &\text{IF High}(F_{\text{corr}}) \text{ AND Low}(F_{\text{dis}}) \\ &\text{Then STAY } \text{CF} = 0.5 \end{aligned} \quad (11)$$

$$\begin{aligned} &\text{IF Low}(F_{\text{corr}}) \text{ AND Low}(F_{\text{dis}}) \\ &\text{Then CUT } \text{CF} = 0.8 \end{aligned}$$

$$\begin{aligned} &\text{IF High}(F_{\text{corr}}) \text{ AND High}(F_{\text{fade}}) \\ &\text{Then FADE IN } \text{CF} = 0.95 \end{aligned}$$

$$\begin{aligned} &\text{IF High}(F_{\text{corr}}) \text{ AND Low}(F_{\text{ratio}}) \\ &\text{Then FADE OUT } \text{CF} = 0.95 \end{aligned}$$

$$\begin{aligned} &\text{IF High}(F_{\text{corr}}) \text{ AND High}(F_{\text{dis}}) \\ &\text{Then DISSOLVE } \text{CF} = 0.8 \end{aligned}$$

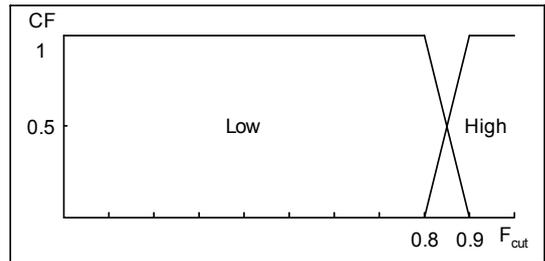
$$\begin{aligned} &\text{IF High}(F_{\text{corr}}) \text{ AND Low}(F_{\text{hwipe}}) \\ &\text{Then HWIPE } \text{CF} = 0.85 \end{aligned}$$

$$\begin{aligned} &\text{IF High}(F_{\text{corr}}) \text{ AND Low}(F_{\text{vwipe}}) \\ &\text{Then VWIPE } \text{CF} = 0.85 \end{aligned}$$

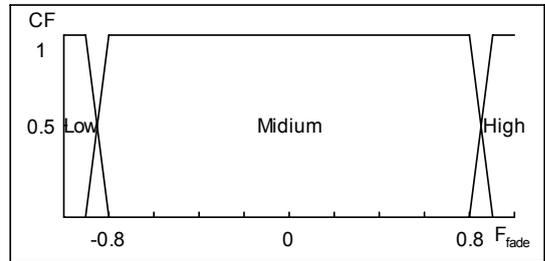
$$\begin{aligned} &\text{IF Low}(F_{\text{corr}}) \text{ AND Low}(F_{\text{barn}}) \\ &\text{Then BARN } \text{CF} = 0.8 \end{aligned}$$

$$\begin{aligned} &\text{IF Low}(F_{\text{corr}}) \text{ AND Low}(F_{\text{iris}}) \\ &\text{Then IRIS } \text{CF} = 0.8 \end{aligned}$$

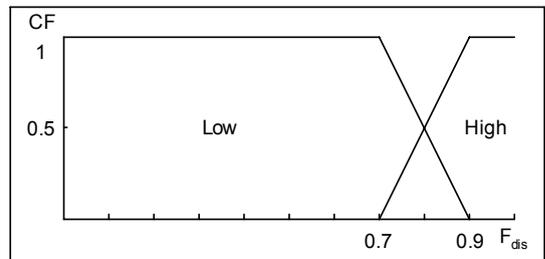
위에서 정의한 규칙의 확신도는 조건부에 대한 결론부의 신뢰도로서 실험적으로 생성한다 그리고 장면전환에 대한 최종적인 추론은 결론부에 대한 확신도를 계산한 후 최대값의 확신도를 가지는 규칙의 조건부를 최종적인 추론 결과로 선택한다.



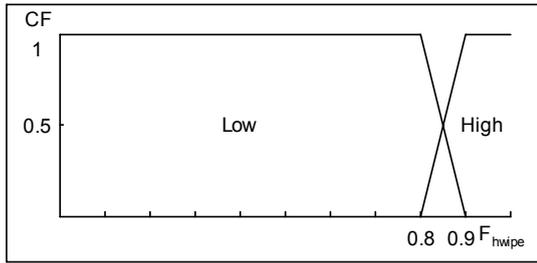
(a) F_{cut} 의 확신도



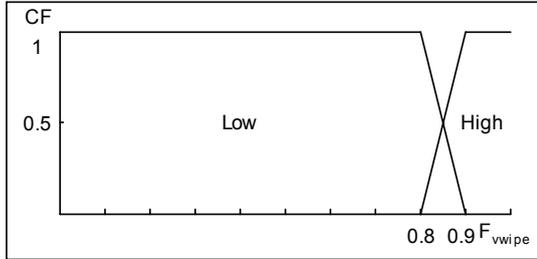
(b) F_{fade} 의 확신도



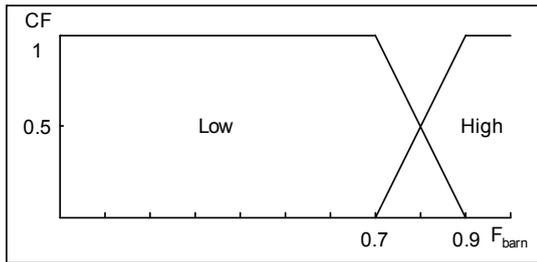
(c) F_{dis} 의 확신도



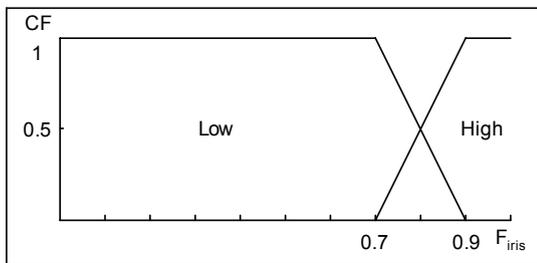
(d) F_{hwipe} 의 확신도



(e) F_{vwipe} 의 확신도



(f) F_{barn} 의 확신도



(g) F_{iris} 의 확신도

[그림 2] 특장의 확신도
[Fig. 2] Certainty factors of features

5. 실험결과

본 논문에서는 Visual C++ 2008을 이용하여 장면전환 검출 알고리즘을 구현하였다. 그리고 제안된 알고리즘의 성능을 비교 평가하기 위해서 사용한 입력 비디오 데이터로는 뉴스, 영화, 광고, 뮤직비디오, 스포츠, 애니메이션,

드라마 등의 여러 종류의 비디오 데이터를 활용하였다. 이 디지털 비디오 데이터 중에는 50개의 컷, 20개의 페이드 인, 20개의 페이드 아웃, 30개의 디졸브, 그리고 40개의 와이프 전환의 총 160개의 장면전환 효과가 포함되어 있다. 그림 3은 본 논문에서 사용한 장면전환 영상의 예를 보여준다.



(a) 페이드 인



(b) 페이드 아웃



(c) 디졸브

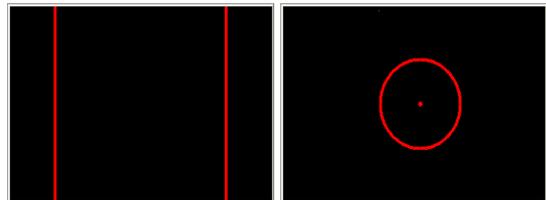


(d) 반도어



(e) 아이리스 라운드

[그림 3] 장면전환의 예
[Fig. 3] Examples of shot transitions



[그림 4] 라인과 원 검출
[Fig. 4] Extraction of lines and circles

그림 4는 제안된 방법에서 와이프를 검출할 때 추출한 두 개의 라인과 원 검출의 예를 보여준다. 그리고 본 논문

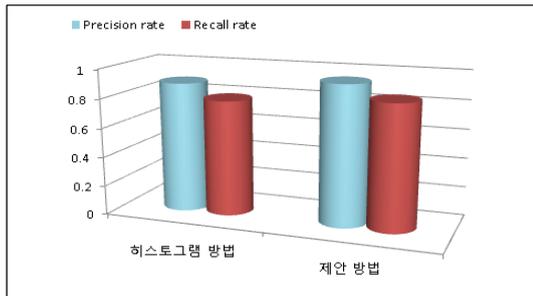
에서는 제안된 방법의 성능을 비교 평가하기 위해서 기존의 히스토그램을 이용한 방법과 허프 변환을 이용한 방법, 그리고 제안된 방법을 비교 실험하였다. 식 (12)는 장면전환 검출의 성능을 평가하기 위해 사용한 정확도 척도이다[11].

$$R_{precision} = \frac{N_c}{N_c + N_f} \quad (12)$$

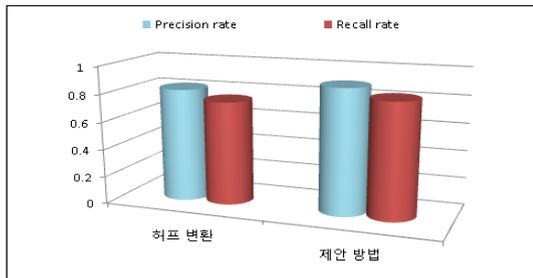
$$R_{recall} = \frac{N_c}{N_c + N_m}$$

식 (12)에서 N_c 는 정확하게 검출한 장면전환의 개수를 의미하고, N_f 는 잘못 검출한 전환의 개수를 의미하며, N_m 은 검출하지 못한 전환의 개수를 의미한다. 그리고 $R_{precision}$ 은 영상에서 검출된 전체 장면전환 중에서 정확하게 검출된 전환의 비율을 의미하고, R_{recall} 은 영상에 존재하는 전체 장면전환 중에서 정확하게 검출된 전환의 비율을 의미한다.

기존의 방법 중에서 본 논문에서 시도하는 모든 장면전환을 검출하는 방법은 찾기 어려우므로, 컷, 페이드, 디졸브에 대한 실험과 와이프에 대한 실험으로 구분하여 수행하였다. 즉, 컷, 페이드, 디졸브 검출은 기존의 히스토그램을 이용한 방법과 제안한 방법을 비교하였고, 와이프 검출은 기존의 허프변환을 이용한 방법과 제안한 방법을 비교하였다.



(a) 컷, 페이드, 디졸브



(b) 와이프

[그림 5] 성능 비교
[Fig. 5] Performance comparison

그림 5는 각 방법에 대한 $R_{precision}$ 과 R_{recall} 의 비교 결과를 그래프로 보여주고 있다. 그림 5에서 확인할 수 있듯이 히스토그램을 이용한 방법과 허프변환을 이용한 방법보다 제안된 방법이 보다 정확하게 여러 장면전환을 검출함을 확인할 수 있었다.

6. 결론

본 논문에서는 급진적인 장면전환인 컷과 점진적인 장면전환인 페이드 인, 페이드 아웃, 디졸브, 그리고 와이프 장면전환을 확신도 추론을 통해 검출하는 새로운 방법을 제안하였다. 제안된 방법에서는 각 장면전환을 가장 잘 나타내는 특징을 추출하고, 확신도를 이용하여 이 특징을 통합하면서 해당되는 장면전환의 발생 유무 및 그 종류를 판단하였다.

향후에는 가설 규칙에 사용된 확신도 수치 계산을 자동화할 계획이며, 동영상으로부터 여러 감성을 인식하기 위한 전처리 단계로서의 장면전환을 검출하는 방향으로 새로운 연구를 시도할 예정이다.

References

- [1] L.-H. Chen, Y.-C. Lai, and H.-Y. M. Liao, "Movie Scene Segmentation Using Background Information" Pattern Recognition, Vol. 41, No. 3, pp. 1056-1065, 2008.
- [2] S. Li and M.-C. Lee, "Effective Detection of Various Wipe Transitions," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 6, pp. 663-673, 2007.
- [3] R. Joyce and B. Liu, "Temporal Segmentation of Video Using Frame and Histogram Space," IEEE Transactions on Multimedia, Vol. 8, No. 1, pp. 130-140, 2006.
- [4] R. Zabih, J. Miller, and K. Mai, "A Feature-based Algorithm for Detecting and Classifying Production Effects," Multimedia Systems, Vol. 7, No. 2, pp. 119-128, 1999.
- [5] C. A. Dhawalel and S. Jain, "Motion Compensated Video Shot Detection Using Multiple Feature Experts," International Journal of Graphics, Vision, and Image Processing, Vol. 8, No. 5, pp. 1-11, 2009.
- [6] P. Campisi, A. Neri, and L. Sorigi, "Wipe Effect Detection for Video Sequences," In Proc. of the IEEE Workshop on Multimedia Signal Processing, pp. 161-164, 2002.
- [7] A. F. Smeaton, P. Over, and A. R. Doherty, "Video Shot Boundary Detection: Seven Years of TRECVID Activity,"

- Computer Vision and Image Understanding, Vol. 114, No. 4, pp. 411-418, 2010.
- [8] K. Bunte, M. Biehl, M. F. Jonkman, and N. Petkov, "Learning Effective Color Features for Content Based Image Retrieval in Dermatology," Pattern Recognition, Vol. 44, No. 9, pp. 1892- 1902, 2011.
- [9] F. Y. Shih, "Image Processing and Pattern Recognition: Fundamentals and Techniques," John Wiley & Sons, 2010.
- [10] J. Giarratano, "Expert Systems: Principles and Programming," Third Edition, PWS Publishing Company, 1998.
- [11] D. Lelescu, D. Schonfeld, "Statistical Sequential Analysis for Real-Time Video Scene Change Detection on Compressed Multimedia Bitstream," IEEE Transactions on Multimedia, Vol. 5, No. 1, pp. 106-117, 2003.
- [12] B. T. Jeon, "A Study on Analyzing Caption Characteristic for Recovering Original Images of Caption Region in TV Scene", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 10, No 4, pp. 177~182, 2010.
- [13] Y. K. Jung, "Decision of Gaussian Function Threshold for Image Segmentation", Journal of The Institute of Webcasting, Internet and Telecommunication, Vol 9, No 5, pp. 163~168, 2009.

장 석 우(Seok-Woo Jang)

[정회원]



- 1995년 2월 : 숭실대학교 전자계학과 (공학사)
- 1997년 2월 : 숭실대학교 컴퓨터학과 (공학석사)
- 2000년 8월 : 숭실대학교 컴퓨터학과 (공학박사)
- 2009년 3월 ~ 현재 : 안양대학교 디지털미디어학과 조교수

<관심분야>

로봇비전, 증강현실, HCI, 비디오 색인 및 검색, 등