

# 카메라의 동작을 보정한 장면전환 검출

장 석 우<sup>†</sup> · 최 형 일<sup>\*\*</sup>

## 요 약

본 논문에서는 비디오 데이터로부터 장면 사이의 경계를 검출하고, 이들을 그 종류별로 분류하는 장면전환 검출 방법을 제안한다. 제안한 장면전환 검출 방법은 급진적인 장면전환인 컷(cut)과 점진적인 장면전환인 페이드(fade) 및 디졸브(dissolve)를 검출한다. 본 논문에서는 영상 내에 포함된 카메라의 동작 정보를 이용하여 영상을 보정하고, 보정된 영상으로부터 특징을 추출하여 장면전환을 검출한다. 따라서 카메라의 동작으로 인해 기인하는 여러 가지 오 검출을 방지한다. 또한, 영상을 보정하는 과정에서 지역적인 이동 물체의 동작을 제거하므로 이동 물체의 동작으로 인해 기인하는 장면전환의 오 검출도 방지한다. 실험에서는 다양한 비디오 데이터를 입력 받아 기존의 장면전환 검출 방법들과 제안한 방법의 성능을 비교 분석함으로써 제안한 방법의 우수함을 보인다.

키워드 : 장면전환 검출, 동작벡터, 동작 보정, 특징 추출

## Shot Transition Detection by Compensating Camera Operations

Seok-Woo Jang<sup>†</sup> · Hyung-Il Choi<sup>\*\*</sup>

### ABSTRACT

In this paper, we propose an effective method for detecting and classifying shot transitions in video sequences. The proposed method detects and classifies shot transitions including cuts, fades and dissolves by compensating camera operations in video sequences, so that our method prevents false positives resulting from camera operations. Also, our method eliminates local moving objects in the process of compensating camera operations, so that our method prevents errors resulting from moving objects. In the experiments, we show that our shot transition approach can work as a promising solution by comparing the proposed method with previously known methods in terms of performance.

Key Words : Shot Transition Detection, Motion Vector, Motion Compensation, Feature Extraction

### 1. 서 론

정보 서비스의 주역으로 떠오르고 있는 디지털 비디오 데이터는 동영상, 오디오 및 텍스트 정보를 함께 포함하고 있는 복합적 성격의 멀티미디어 데이터로서 영상 매체의 발달과 더불어 매우 급속히 증가하고 있으며, 정보의 중요한 교환 수단으로 발전하고 있다. 그러나 비디오 데이터는 자료의 양이 방대하고 그 형태 또한 다양하므로 비디오의 내용을 자동으로 분석하고 효과적으로 브라우징 하기 위해서 비디오를 보다 작은 단위의 장면(shot)으로 분할하는 장면전환 검출이 필요하다[1].

장면전환은 장면 사이의 전이가 발생하는 부분, 즉 장면 사이의 경계를 의미한다. 따라서 장면전환이 발생하기 전과

후의 영상은 서로 다른 장면으로 구성된다. 장면전환은 비디오 편집으로 인해 생성되는 것으로서 급진적인 장면전환인 컷(cut)과 점진적인 장면전환인 페이드(fade)와 디졸브(dissolve)로 구성된다[2]. 컷은 인접한 영상 사이에서 장면이 급변하는 전환이다. 페이드는 고정 영상(constant image)이 점진적으로 장면화되는 전환인 페이드 인(fade in)과 장면이 점진적으로 고정 영상으로 변화하는 전환인 페이드 아웃(fade out)이 있다. 그리고 디졸브는 페이드 아웃과 페이드 인이 교차해서 발생하는 장면전환이다. 일반적으로, 디졸브에서 페이드 아웃과 페이드 인은 동시에 발생하며 페이드 되는 속도는 일정하다.

장면전환을 검출하는 기존의 연구 방법으로는 특징값의 히스토그램을 이용하는 방법, 에지를 이용하는 방법, 규칙 생성을 이용하는 방법 등이 있다[3, 4, 5]. 특징값의 히스토그램을 이용하는 방법에서는 이중의 임계치를 사용하여 추출된 특징을 이중으로 비교하여 장면전환을 검출한다. 규칙 생성을 이용하는 방법은 여러 가지 특징을 추출한 후 규칙

※ 이 논문은 2004년도 학술진흥재단의 지원에 의하여 연구되었음.  
(KRF-2004-005-D00198)

† 정 회 원 : 한국건설기술연구원 건설정보화연구부 선임연구원

\*\* 중 심 회 원 : 숭실대학교 미디어학부 교수

논문접수 : 2005년 2월 24일, 심사완료 : 2005년 6월 22일

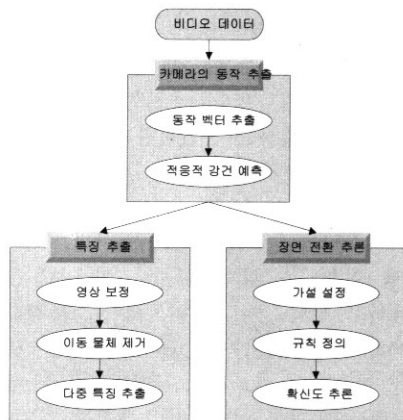
을 이용하여 장면전환을 추론한다. 에지를 이용하는 방법은 입력 영상으로부터 에지를 추출하고, 추출된 에지 중에서 새롭게 출현하는 에지인 인(in) 에지와 존재하다가 사라지는 에지인 아웃(out) 에지의 수를 비교하여 장면전환을 검출한다. 이런 기존의 장면전환 검출 방법들의 단점은 크게 세 가지로 요약할 수 있다. 첫째, 기존의 방법들은 급진적인 장면전환인 컷 검출을 위주로 하고 있으며 점진적인 장면전환 검출은 매우 미약한 실정이다. 둘째, 영상에 카메라의 동작이 포함될 경우 카메라의 동작을 보정하지 않은 상태에서 장면전환을 검출하므로 많은 오 검출을 유발한다. 셋째, 영상 내에 이동하는 물체가 포함될 경우 물체의 움직임으로 인한 많은 오 검출을 유발한다.

본 논문에서는 위에서 기술한 문제점들을 해결하기 위해 새로운 장면전환 검출 방법을 제안한다. 제안하는 장면전환 검출 방법은 여러 가지 특징들을 추출한 후 규칙을 이용하여 장면전환을 추론함으로써 점진적인 장면전환과 급진적인 장면전환을 모두 검출한다. 그리고 영상 내에 포함된 카메라의 동작 정보를 이용하여 영상을 보정하고, 보정한 영상으로부터 특징을 추출하여 장면전환을 검출한다. 따라서 카메라의 동작으로 인해 기인하는 여러 가지 오 검출을 방지한다. 또한, 영상을 보정하는 과정에서 지역적인 이동 물체의 동작을 제거하므로 이동 물체의 동작으로 인해 기인하는 장면전환의 오 검출도 방지한다.

1장에서는 연구 동기와 전체적인 개요를 기술하였다. 2장에서는 확장적 블록 정합에 대해 기술하고, 3장에서는 적응적 강건 예측에 대해 기술하며, 4장에서는 특징을 추출하는 방법에 대해 기술한다. 그리고 5장에서는 확신도를 이용하여 장면 전환을 추론하는 방법에 대해 기술하고, 6장에서는 실험결과 및 결론을 기술한다.

## 2. 확장적 블록 정합

(그림 1)은 본 논문에서 제안하는 장면전환 검출의 전체적인 구조를 보여준다. 그림에서 보는 바와 같이 제안한 장면전환 검출은 크게 카메라의 동작 추출 단계, 특징 추출 단계, 장면전환 추론 단계의 세 단계로 구성된다.



(그림 1) 장면전환 검출의 전체 구조

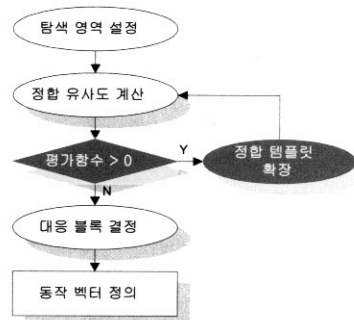
카메라의 동작 추출 단계에서는 인접한 두 영상으로부터 블록 정합을 이용하여 동작 벡터를 구하고, 이를 강건 예측(robust estimation)에 적용하여 배경 영역의 동작 정보를 나타내는 카메라의 동작 인수를 구한다. 본 논문에서는 기존의 블록 정합을 개선한 확장적 블록 정합을 이용하여 동작 벡터를 보다 정확하게 구하며, 기존의 강건 예측을 개선한 적응적 강건 예측을 이용하여 입력 데이터에 포함된 오류 자료를 최대한 제거하면서 카메라의 동작 인수를 구한다.

특징 추출 단계에서는 추출된 카메라의 동작 정보를 이용하여 영상을 보정하고, 보정된 영상으로부터 장면전환을 대표하는 특징들을 블록 단위로 추출한다. 본 논문에서는 영상을 보정하는 과정에서 지역적인 이동 물체의 움직임과 잡음(noise) 영역을 제거한다. 그리고 장면전환을 대표하는 컷 특징, 페이드 특징, 디졸브 특징과 더불어 카메라 동작의 일관성을 나타내는 특징도 추출한다.

장면전환 검출 단계에서는 불확실성 이론인 확신도(certainty factor)를 이용하여 장면전환을 추론한다. 이 단계에서는 각 장면전환에 대한 가설을 설정하고, 장면전환을 나타내는 영상 특징과 카메라 동작의 일관성 특징을 통합하여 규칙을 생성한다. 그리고 생성된 규칙들로부터 확신도 추론 과정을 통해 장면전환의 종류를 최종적으로 판명한다.

본 논문에서는 보다 정확하게 동작벡터를 추출하기 위해 확장적 블록 정합을 사용한다[6]. 블록 정합은 영상을 모양과 크기가 동일한 사각형의 블록으로 분할한 후 블록 단위의 특징을 정의하고, 정의된 특징에 의해 구성되는 정합척도를 일정한 탐색 영역 내의 블록들에 적용하여 가장 높은 정합척도를 가지는 블록을 찾는다. 그리고 기준 블록과 정합된 블록 사이의 중심 위치의 변위를 동작벡터로 정의한다.

확장적 블록 정합은 정합의 구분력이 저하되고 오 정합이 발생하는 기존 방법들의 문제점을 해결하기 위해, 구분력 있는 최대 정합 유사도를 추출할 때까지 정합 템플릿의 크기를 확장하면서 블록 사이의 정합을 반복한다. 블록 정합의 반복 여부는 평가 함수를 통해 결정한다. 평가 함수는 현재의 탐색 영역에서 구한 정합 유사도 분포 중 최대 정합 유사도를 가지는 블록의 정합 구분력 유무를 판단하여 블록 정합을 반복할 것인지를 결정한다. (그림 2)는 활동적 블록 정합 방법의 전체 구조를 보여준다. (그림 2)에서 확인할 수 있듯이 평가 함수가 양의 값을 가질 경우에는 정합 템플릿



(그림 2) 확장적 블록 정합의 구조

을 확장하면서 블록 정합을 반복하고, 평가 함수가 음의 값을 가질 경우에는 정합된 대응 블록을 이용하여 동작벡터를 결정한다.

본 논문에서 사용하는 정합 유사도(DBS: displaced block similarity)는 (식 1)과 같다.

$$DBS(i, j; u, v) = \left( 1 - \frac{1}{n^2} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} \left| \frac{I_n(i+x, j+y) - I_{n-1}(i+u+x, j+v+y)}{I_{max}} \right| \right) \times 100 \quad (1)$$

(식 1)에서 n은 블록의 크기를 의미하고, (i, j)는 영상의 i, j에 위치하는 블록의 시작 좌표를 의미하며, (u, v)는 기준 블록과 정합되는 블록 사이의 변위를 의미한다. 그리고  $I_n, I_{n-1}$ 은 현재 시점과 이전 시점 영상의 명암값을 의미하고,  $I_{max}$ 는 명암값의 최대값 255를 의미한다. 그리고 정합 유사도 DBS는 0에서 100사이의 범위를 갖도록 정규화된다.

본 논문에서 제안하는 확장적 블록 정합은 구분력 있는 최대 정합 유사도를 추출할 때까지 정합 템플릿의 크기를 확장하면서 블록 사이의 정합을 반복적으로 수행한다. 이를 위해서 본 논문에서는 두 가지 평가 기준을 설정하여 사용한다. 첫째, 정합 템플릿의 확장 여부에 대한 기준을 설정한다. 즉, 정해진 탐색 영역에서 추출한 정합 유사도의 분포 중 최대 정합 유사도가 구분력이 있는가를 평가한다. 이는 최대 정합 유사도  $DBS_{max}$ 를 가지는 블록의 위치를 중심으로 인접 블록과의 DBS의 차이를 비교하여 평가한다. 본 논문에서는  $DBS_{max}$ 를 가지는 블록과 인접 블록의 DBS의 차이 중 최소값을 경사도(peakness)로 정의하고, 경사도가 실험적으로 결정된 임계치 TH 이하일 경우에만 블록을 확장한다. 둘째, 정합 템플릿의 크기를 확장하면서 정합 유사도를 반복적으로 계산할 경우 정합 템플릿의 확장 지속 여부에 대한 기준을 설정한다. 본 논문에서는 정합 템플릿을 확장함에 따라 경사도의 1차 미분값이 +에서 -로 변하는 변곡점을 발견할 때까지 블록 정합을 반복한다.

위에서 기술한 두 가지 평가 기준을 바탕으로 본 논문에서 사용하는 평가 함수를 수식으로 정의하면 (식 2)와 같다. (식 2)에서  $(x(n), y(n))$ 은 정합 템플릿의 크기가  $N \times N$ 일 때의 최대 정합 유사도  $DBS_{max}$ 를 가지는 블록의 중심 위치를 의미하고,  $i^*$ 와  $j^*$ 는 위치가  $(x(n), y(n))$ 인 블록의 인접 블록의 중심 위치를 의미한다.

$$\Phi(x(n), y(n)) = \max \left[ \begin{matrix} e_1 + TH(x(n), y(n)) \\ e_2 + GD(x(n), y(n)) \end{matrix} \right] \times \frac{e_1 + TH(x(n), y(n))}{e_2 + GD(x(n), y(n))} \quad (2)$$

$$TH(x(n), y(n)) = T_{PK} - PK(x(n), y(n))$$

$$GD(x(n), y(n)) = PK(x(n), y(n)) - PK(x(n-1), y(n-1))$$

$$PK(x(n), y(n)) = \min_{-1 \leq i', j' \leq 1} [DBS(x(n), y(n)) - DBS(x(n) + i', y(n) + j')].$$

$$\text{where } \begin{cases} e_1 : \text{infinitesimal positive number} \\ e_2 : \text{infinitesimal negative number} \\ T_{PK} : \text{threshold value of peakness} \end{cases}$$

### 3. 적응적 강건 예측

본 논문에서는 확장적 블록 정합으로 구한 동작벡터를 적응적 강건 예측[7]에 적용하여 카메라의 동작을 구한다. 강건 예측은 입력 자료에 포함된 오류 자료(outlier)를 제거하면서 정상 자료(nonoutlier)만으로 모델의 인수를 구하는 통계적인 방법이다[8]. 본 논문에서는 카메라의 동작 인수를 구하기 위해 강건 예측을 사용하고, 강건 예측의 입력으로 동작벡터를 사용한다. 따라서 본 논문에서 의미하는 정상 자료는 순수하게 카메라의 동작만을 포함하고 있는 배경 영역의 동작벡터를 의미하고, 오류 자료는 배경 영역의 동작벡터를 제외한 동작벡터, 즉 물체의 이동이나 그 외의 잡음으로 형성된 동작벡터를 의미한다.

일반적으로 기존의 강건 예측은 크게 두 단계로 구성된다. 첫째, 입력 자료로부터 오류 자료를 제거하는 단계이다. 오류 자료가 포함될 경우 모델을 최적화시키는 인수가 부정확하게 수렴하므로 입력 자료의 통계적인 분포를 분석하여 오류 자료를 제거한다. 둘째, 오류 자료를 제거한 정상 자료만을 이용하여 모델의 인수를 구한다. 모델의 인수를 구하는 과정에서는 최소화 기법을 사용한다. 그러나 기존의 강건 예측은 정상 자료와 오류 자료를 구분하기 어려운 최소화 기법의 초기 단계에서도 임계치를 이용하여 정상 자료와 오류 자료를 분리한다. 따라서 정상 자료와 오류 자료의 부정확한 분리로 최소화 기법의 단계가 반복됨에 따라 모델의 인수가 잘못 갱신된다. 또한, 최소화 기법의 각 단계마다 가중치 사이에 존재하는 연관성을 무시하고 가중치 함수를 매번 새롭게 계산하므로 가중치의 불규칙적인 진동이 발생한다. 따라서 본 논문에서는 이런 문제점을 개선한 적응적 강건 예측을 사용한다.

본 논문에서 카메라의 동작을 측정하기 위해 사용하는 측정 모델은 (식 3)과 같다. 일반적으로, (식 3)은 어파인(affine) 모델이라고 불린다.

$$y = y(x, y, a) \quad (3)$$

$$y(x, y, a) = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix} = \begin{bmatrix} a_1x + a_2y + a_3 \\ a_4x + a_5y + a_6 \end{bmatrix}$$

(식 3)에서 인수 벡터 a는 본 논문에서 추출하고자 하는 카메라의 동작 인수로서 기존의 방법을 이용하여 쉽게 상각(tilt angle), 좌우각(pan angle), 경사각(swing angle), 초점 거리(focal length)와 같은 조망인자(viewing parameter)로 변환할 수 있다[9]. 그리고  $u(x, y)$ 와  $v(x, y)$ 는 적응적 강건 예측의 입력으로 사용되는 동작벡터의 수평 성분과 수직 성분을 의미한다. (식 3)의 측정 모델을 기초로 목적 함수를 정의하면 (식 4)와 같다. (식 4)에서  $w_i$ 는 i번째 입력 자료의 가중치를 의미하고,  $x_i'$ 와  $y_i'$ 은 i번째 입력 자료의 수평 성분과 수직 성분을 의미하며,  $\hat{x}_i(x_i, y_i, a)$ 와  $\hat{y}_i(x_i, y_i, a)$ 는 측정 모델의 수평 성분과 수직 성분을 의미한다.

$$\begin{aligned} \chi^2(\mathbf{a}) &= \sum_{i=1}^N w_i \left[ \frac{y_i' - y(x_i, y_i, \mathbf{a})}{\sigma_i} \right]^2 \\ &= \sum_{i=1}^N w_i \left[ \frac{(x_i' - \hat{x}_i(x_i, y_i, \mathbf{a}))^2}{\sigma_{x_i}^2} + \frac{(y_i' - \hat{y}_i(x_i, y_i, \mathbf{a}))^2}{\sigma_{y_i}^2} \right] \\ &= \sum_{i=1}^N w_i \| \mathbf{r}_i \|^2 \end{aligned} \quad (4)$$

본 논문에서 사용하는 적응적 강건 예측은 이진 가중치 함수를 사용하는 기존의 강건 예측과는 다르게 연속적인 가중치 함수를 사용한다. 연속적인 가중치 함수로는 시그모이드 함수(sigmoid function)를 이용한다. 본 논문에서 사용하는 시그모이드 가중치 함수는 이진 가중치 함수보다 정상 자료와 오류 자료의 소속 정도를 보다 효과적으로 표현하므로 최소화 기법의 반복 초기 단계에서 범할 수 있는 에러를 최소화할 수 있다. 본 논문에서 사용하는 가중치 함수는 (식 5)와 같이 정의한다.

$$\mathbf{W}^k = (w_1^k, w_2^k, \dots, w_j^k, \dots, w_N^k) \quad (5)$$

$$w_j^k = \alpha \cdot w_j^{k-1} + \beta \cdot (1 - \text{sig}^k(x = j; a^k, c^k))$$

$$\text{sig}^k(x; a^k, c^k) = \frac{1}{1 + e^{-a^k(x - c^k)}}$$

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

(식 5)에서  $\mathbf{W}^k$ 는 최소화 단계의 k번째 반복에서 사용하는 가중치 벡터로서  $w_i^k$ 는 0에서 1사이의 범위를 가지는 i번째 입력 자료의 가중치를 나타낸다. 그리고 x는 시그모이드 함수의 입력 벡터를 나타내고,  $c^k$ 는 시그모이드 함수를 수평축을 기준으로 이동시키는 바이어스(bias)를 나타내며,  $a^k$ 는  $x=c$  위치에서 시그모이드 함수의 기울기를 조절하는 역할을 수행한다. 본 논문에서는 에러를 오름차순으로 정렬한 누적 에러를 사용하므로 일반적인 시그모이드 함수를 사용하지 않고  $x=c$  위치에서 y축을 중심으로 회전 이동한 시그모이드 함수를 가중치 함수로 사용한다.

본 논문에서 사용하는 적응적 강건 예측은 최소화 기법의 반복 과정에서 현재 단계의 가중치 계산 시 이전 단계의 가중치를 반영하도록 정의된다. 또한, 최소화 기법이 반복되면서 에러가 감소함에 따라 시그모이드 함수를 하드 리미트(hard limit) 형태로 조율함으로써 정상 자료와 오류 자료를 보다 유연하게 분리한다. 본 논문에서의 가중치 조율은 시그모이드 함수의 두 인수  $c_k$ 와  $a_k$ 를 조율하여 수행하는데, 가중치 조율에 대한 자세한 사항은 참고문헌 [7]에서 확인할 수 있다.

#### 4. 영상 보정에 의한 특징 추출

본 장에서는 카메라의 동작을 나타내는 동작벡터의 수직 및 수평성분을 이용하여 영상을 보정하고, 보정된 영상으로부터 장면전환을 나타내는 4가지의 특징, 즉 컷 특징, 페이

드 특징, 디졸브 특징, 카메라 동작의 일관성 특징을 추출하는 방법에 대해 기술한다.

본 논문에서는 장면전환을 나타내는 특징을 추출하기 위해 HIS 칼라 공간을 사용한다[10]. HIS 칼라 공간은 인간의 시각 시스템과 유사한 모델로서 칼라를 표현하기 위해 색상(hue), 명암값(intensity), 채도(saturation)를 사용한다. HIS 칼라 공간은 각 구성요소 간의 상호 관계가 적은 칼라 모델로서 한 가지 구성 요소의 값의 변화가 다른 구성요소의 값의 변화에 영향을 미치지 않는다.

장면전환 중 컷은 한 장면에서 다른 장면으로 급전하는 변환이므로 컷을 기준으로 이전 장면과 다음 장면 사이의 유사성은 매우 낮다. 따라서 장면 사이의 색상, 명암값, 채도의 변화는 매우 심하게 나타난다. 특히, 서로 다른 장면일수록 명암값이나 채도에 비해 색상의 변화가 심하다. 본 논문에서는 이런 특성을 기반으로 컷 장면전환에 대한 특징을 (식 6)과 같이 정의한다.

$$\mathbf{F}_{corr} = \alpha \times H_{corr} + \beta \times I_{corr} \quad (6)$$

$$H_{corr} = \frac{1}{2} \left( 1 + \frac{\sum_{i=1}^N (BHM_{i-1}^t - HM_{i-1}^t)(BHM_i^t - HM_i^t)}{\sqrt{\sum_{i=1}^N (BHM_{i-1}^t - HM_{i-1}^t)^2} \sqrt{\sum_{i=1}^N (BHM_i^t - HM_i^t)^2}} \right)$$

$$I_{corr} = \frac{1}{2} \left( 1 + \frac{\sum_{i=1}^N (BIM_{i-1}^t - IM_{i-1}^t)(BIM_i^t - IM_i^t)}{\sqrt{\sum_{i=1}^N (BIM_{i-1}^t - IM_{i-1}^t)^2} \sqrt{\sum_{i=1}^N (BIM_i^t - IM_i^t)^2}} \right)$$

where  $\alpha, \beta$ : weights ( $0 < \alpha \leq \beta < 1, \alpha + \beta = 1$ )

(식 6)에서 컷 특징  $\mathbf{F}_{corr}$ 은 현재 시점과 이전 시점 영상 사이의 색상과 명암값의 상관 관계를 나타내는  $H_{corr}$ 과  $I_{corr}$ 으로 정의하며, 0에서 1사이의 범위를 갖도록 정규화한다. 컷 특징  $\mathbf{F}_{corr}$ 이 0에 근접할수록 컷일 확률이 높고, 반대로 1에 근접할수록 컷일 확률이 낮다. (식 6)에서  $BHM_i^t$ 와  $BIM_i^t$ 는 현재 시점의 i번째 블록의 색상 평균과 명암값 평균을 의미하며,  $HM_i^t$ 과  $IM_i^t$ 는 현재 시점 영상의 색상 평균과 명암값 평균을 의미한다. 그리고  $\alpha$ 와  $\beta$ 는 색상과 명암값에 대한 가중치로 사용하는데 본 논문에서는 색상에 보다 높은 가중치를 부여한다. (식 6)에서 사용하는 색상 평균과 명암값 평균은 (식 7)과 같이 정의한다.

$$BHM_i^t = \frac{w_i \cdot \left( \sum_{k=1}^n \sum_{l=1}^n BH_i^t(x+k, y+l) \right)}{n^2} \quad (7)$$

$$BHM_{i-1}^t = \frac{w_i \cdot \left( \sum_{k=1}^n \sum_{l=1}^n BH_{i-1}^t(x+u, y+v+l) \right)}{n^2}$$

$$BIM_i^t = \frac{w_i \cdot \left( \sum_{k=1}^n \sum_{l=1}^n BI_i^t(x+k, y+l) \right)}{n^2}$$

$$BIM_{i-1}^t = \frac{w_i \cdot \left( \sum_{k=1}^n \sum_{l=1}^n BI_{i-1}^t(x+u, y+v+l) \right)}{n^2}$$

$$HM_i^i = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n BH_i^i(x+k, y+l) \right)}{n^2 \cdot N}$$

$$HM_{i-1}^i = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n BH_{i-1}^i(x+u_i+k, y+v_i+l) \right)}{n^2 \cdot N}$$

$$IM_i^i = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n BI_i^i(x+k, y+l) \right)}{n^2 \cdot N}$$

$$IM_{i-1}^i = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n BI_{i-1}^i(x+u_i+k, y+v_i+l) \right)}{n^2 \cdot N}$$

(식 7)에서  $BH_i^i(x, y)$ 와  $BI_i^i(x, y)$ 는  $i$ 번째 블록의 시작 좌표에서의 색상과 명암값을 의미하고,  $n$ 은 블록의 크기를 의미하며,  $N$ 은 전체 블록의 개수를 의미한다. 그리고  $w_i$ 는 각 블록에 대한 가중치로서 적응적 강건 예측에서 가중치 조율 과정을 통해 구한 하드 리미트 형태의 가중치이다. 즉, 오류 자료에 해당하는 블록은 0인 가중치를 가지며, 정상 자료에 해당하는 블록은 1인 가중치를 가진다. 또한,  $u_i$ 와  $v_i$ 는  $i$ 번째 블록으로부터 추출한 동작 벡터의 수평 성분과 수직 성분으로서 이 성분만큼 보정한 이전 영상으로부터 특징을 추출하여 현 시점의 영상 특징과 비교한다.

장면전환 중 페이드 인은 명암값을 증가하면서 고정 영상을 점진적으로 장면화시키는 전환이다. 그리고 페이드 아웃은 명암값을 감소하면서 장면을 고정 영상으로 점진적으로 변화시키는 장면전환이다. 따라서 페이드 장면전환이 발생할 경우에는 인접한 영상 사이의 색상이나 채도의 변화보다는 명암값의 변화가 현저히 발생한다. 본 논문에서는 이런 사실을 근거로 (식 8)과 같은 명암값의 차이의 비율을 이용하여 페이드 인과 페이드 아웃의 특징을 정의한다.

$$F_{ratio} = \frac{I_{diff}}{I_{Adiff}} \quad (8)$$

$$I_{diff} = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n BI_i^i(x+k, y+l) - BI_{i-1}^i(x+u_i+k, y+v_i+l) \right)}{n^2 \cdot N \cdot I_{max}}$$

$$I_{Adiff} = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n |BI_i^i(x+k, y+l) - BI_{i-1}^i(x+u_i+k, y+v_i+l)| \right)}{n^2 \cdot N \cdot I_{max}}$$

(식 8)에서 페이드 특징  $F_{ratio}$ 는 현재 시점과 이전 시점 영상 사이의 명암값의 차이와 명암값의 절대값의 차이의 비율로 정의되며, -1에서 1사이의 값을 가진다. 페이드 특징  $F_{ratio}$ 가 -1에 근접할수록 페이드 아웃일 확률이 높으며, 반대로  $F_{ratio}$ 가 1에 근접할수록 페이드 인일 확률이 높다. (식 8)에서  $I_{max}$ 는 명암값의 최대값인 255를 의미한다.

장면전환 중 디졸브는 페이드 인과 페이드 아웃이 점진적으로 교차해서 발생하는 변환이므로 장면 사이의 색상, 명암값, 채도가 점진적으로 변화한다. 따라서 본 논문에서는 색상, 명암값, 채도의 차이를 모두 이용하여 디졸브 특징을 (식 9)와 같이 정의한다.

$$F_{diff} = \alpha \cdot H_{Adiff} + \beta \cdot I_{Adiff} + \gamma \cdot S_{Adiff} \quad (9)$$

$$H_{Adiff} = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n |BH_i^i(x+k, y+l) - BH_{i-1}^i(x+u_i+k, y+v_i+l)| \right)}{n^2 \cdot N \cdot H_{max}}$$

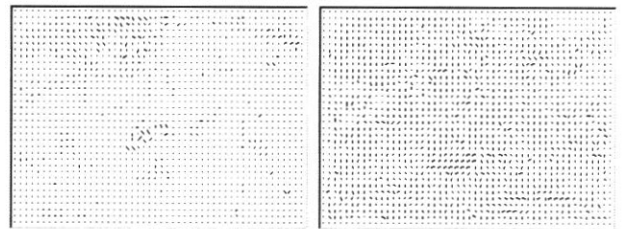
$$I_{Adiff} = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n |BI_i^i(x+k, y+l) - BI_{i-1}^i(x+u_i+k, y+v_i+l)| \right)}{n^2 \cdot N \cdot I_{max}}$$

$$S_{Adiff} = \frac{\sum_{k=1}^N w_i \cdot \left( \sum_{x=1}^n \sum_{y=1}^n |BS_i^i(x+k, y+l) - BS_{i-1}^i(x+u_i+k, y+v_i+l)| \right)}{n^2 \cdot N \cdot S_{max}}$$

where  $\alpha, \beta, \gamma$  : weights ( $0 < \beta \leq \alpha, \gamma < 1, \alpha + \beta + \gamma = 1$ )

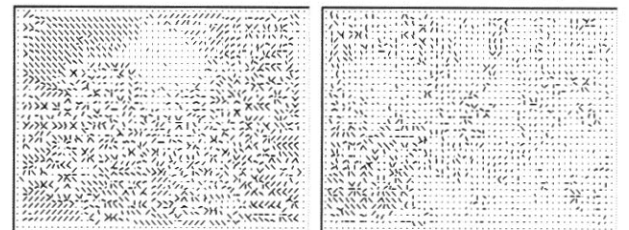
(식 9)에서 디졸브 특징  $F_{diff}$ 는 현재 시점과 이전 시점의 영상 사이의 색상의 절대값 차이, 명암값의 절대값 차이, 채도의 절대값 차이의 합으로 정의되며, 0에서 1사이의 범위를 갖도록 정규화된다. 디졸브 특징  $F_{diff}$ 이 1에 근접할수록 디졸브일 확률이 높으며, 반대로 0에 근접할수록 디졸브일 확률이 낮다. (식 9)에서  $BS_i^i(x, y)$ 는  $i$ 번째 블록의 시작 좌표에서의 채도값을 의미한다.

본 논문에서는 장면전환에 대한 특징을 추출하기 위해 인접한 영상 사이의 칼라 특징과 더불어 카메라 동작의 일관성을 나타내는 특징을 추출하여 사용한다. 일반적으로, 카메라의 동작이 발생하는 영상으로부터 추출한 동작벡터의 구조는 (그림 3)과 같이 카메라의 서로 다른 동작에 따라 그 형태는 다르지만 일관성 있는 형태를 유지한다. 이에 비해 장면전환이 발생한 영상에서는 인접 영상 간의 칼라 특징 사이의 유사성이 없으므로 추출된 동작벡터의 구조도 (그림 4)와 같이 일관적이지 않고 불규칙적인 형태를 나타낸다.



(a) 카메라 정지 동작의 동작벡터 (b) 카메라 틸트 동작의 동작벡터

(그림 3) 카메라의 동작에 따른 동작벡터의 형태



(a) 컷 장면전환에서의 동작벡터 (b) 디졸브 장면전환에서의 동작벡터

(그림 4) 장면전환에 따른 동작벡터의 형태

따라서 본 논문에서는 장면전환을 검출하기 위해 세 가지 칼라 특징과 더불어 카메라 동작의 일관성 특징도 활용한다.

본 논문에서 사용하는 카메라 동작의 일관성 특징은 (식 10)과 같이 적응적 강건 예측에서 오류 자료를 제거하고, 최소화 기법을 통해 최종적으로 추출한 카메라 동작의 에러값을 사용한다.

$$F_{camera} = \sum_{i=1}^N w_i \cdot [y_i' - y(x_i, y_i, a)]^2 \quad (10)$$

(식 10)에서 정의한 카메라 동작의 일관성 특징이 큰 값을 가질수록 장면전환이 발생할 확률이 높으며, 반대로 카메라 동작의 일관성 특징이 낮은 값을 가질수록 장면전환이 발생할 확률이 낮다.

### 5. 확신도를 이용한 추론

본 장에서는 장면전환에 대한 특징들을 이용하여 규칙을 생성하고, 불확실성 이론인 확신도(certainty factor)를 이용하여 장면전환을 추론하는 방법에 대해 기술한다. 확신도에 의한 불확실성 처리 방법은 화농성 질환의 감염 여부를 진단하여 처방을 내려주는 의료 진단용 전문가 시스템이며, 미국의 스탠포드 대학에서 1970년대 초부터 약 10년간에 걸쳐 개발한 시스템인 MYCIN을 개발하는 과정에서 제안된 방법이다[11].

확신도는 증거(evidence)에 대한 가설(hypothesis)의 확신 정도로서 (식 11)과 같이 증거에 대한 가설의 확신과 증거에 대한 가설의 불확신의 차로써 정의한다.

$$CF[h, e] = MB[h, e] - MD[h, e] \quad (11)$$

(식 11)에서 CF는 증거 e가 주어진 경우 가설 h에 대한 확신도를 나타내며 -1에서 1사이의 범위를 가진다. MB(Measure of Belief)는 증거 e가 주어진 경우 가설 h에 대한 확신의 정도로서 가설에 대한 증거의 지지 정도를 나타낸다. MD(Measure of Disbelief)는 증거 e가 주어진 경우 가설 h에 대한 불신도로서 가설에 대한 증거의 부정 정도를 나타낸다. 그리고 MB와 MD는 모두 0에서 1사이의 범위를 가진다. 또한, 확신도 CF를 결정할 경우 MB가 0이거나 MD가 0의 값을 가지면 확신도 CF는 MD나 MB만으로 정의된다. 본 논문에서도 확신도를 결정할 때 MD는 고려하지 않고 MB만을 고려하여 확신도를 결정한다.

(식 12)는 확신도를 규칙에 적용한 예를 보여준다. (식 12)에서 규칙의 조건부를 증거로 해석하고 규칙의 결론부를 가설이라고 해석한다.

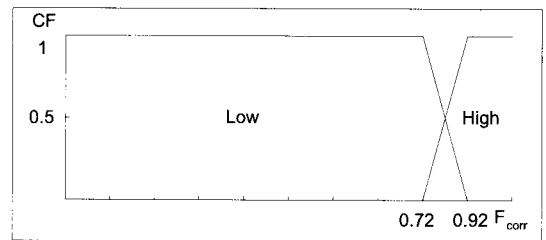
$$\text{IF } \langle \text{evidence} \rangle \text{ THEN } \langle \text{hypothesis} \rangle \quad CF[h,e] \quad (12)$$

(식 12)에서 증거 e에 대한 확신도가 주어질 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 계산한다.

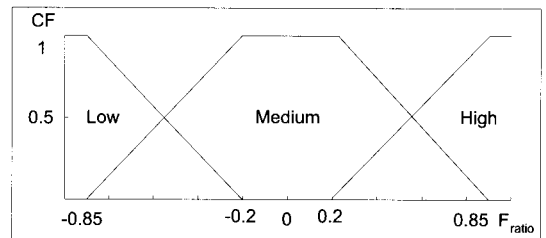
본 논문에서는 확신도를 이용하여 장면전환을 추론하기 위해서 먼저 아래와 같이 5가지 가설을 설정한다.

- (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).

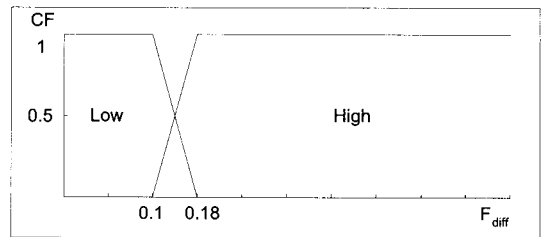
특징들을 통합하여 장면전환에 대한 규칙을 생성하려면 각 특징에 대한 확신도를 결정해야 한다. 본 논문에서는 실험적인 결과를 이용하여 각 특징의 확신도를 (그림 5)와 같이 정의한다. 컷, 디졸브, 카메라 동작의 일관성 특징에 대한 확신도는 두 가지 레벨 Low, High로 정의하며, 페이드 특징에 대한 확신도는 세 가지 레벨 Low, Medium, High로 정의한다. 그리고, 추출한 특징들을 통합하여 가설에 대한 규칙은 (식 13)과 같이 생성한다.



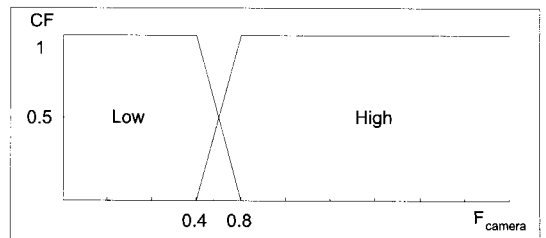
(a) F<sub>corr</sub>의 확신도



(b) F<sub>ratio</sub>의 확신도



(c) F<sub>diff</sub>의 확신도



(d) F<sub>camera</sub>의 확신도

(그림 5) 특징의 확신도

- IF High( $F_{corr}$ ) AND Low( $F_{camera}$ ) Then STAY CF = 0.8
  - IF High( $F_{corr}$ ) AND Low( $F_{diff}$ ) Then STAY CF = 0.5
  - IF Low( $F_{corr}$ ) AND High( $F_{camera}$ ) Then CUT CF = 0.8
  - IF High( $F_{corr}$ ) AND Low( $F_{ratio}$ ) Then FADE IN CF = 0.95
  - IF High( $F_{corr}$ ) AND Low( $F_{ratio}$ ) Then FADE OUT CF = 0.95
  - IF High( $F_{corr}$ ) AND High( $F_{diff}$ ) AND Then DISSOLVE CF = 0.8
  - IF High( $F_{diff}$ ) AND High( $F_{camera}$ ) AND Then DISSOLVE CF = 0.5
- (13)

위의 규칙 중 페이드에 대한 규칙에서는 카메라 동작의 일관성 특징을 사용하지 않았다. 그 이유는 페이드 인과 페이드 아웃 장면전환이 발생하는 동안 카메라 동작의 에러값 특징이 페이드의 발생 유무에 관계없이 불규칙적인 형태를 나타내기 때문이다.

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

### 6. 실험결과

본 장에서는 제안한 장면전환 검출 방법을 다양한 비디오 데이터에 적용하여 실험한 결과에 대해 기술한다. 실험에서는 뉴스, 스포츠, 뮤직 비디오, 애니메이션, 다큐멘터리, 영화, 광고, 드라마 등과 같은 다양한 비디오 데이터를 입력 데이터로 사용하였다.

본 논문에서는 기존의 장면전환 검출 방법 중 히스토그램을 이용한 방법[3], 에지를 이용한 방법[4], 규칙 생성을 이용한 방법[5]과 제안한 방법의 성능을 비교 분석하였다. 실험에서는 다양한 비디오 데이터를 분석하여 중요하다고 판단되는 장면전환을 포함한 영상들만을 선별하여 장면전환 검출을 시도하였다. 이 비디오 데이터는 39개의 컷, 8개의 페이드 인과 페이드 아웃, 8개의 디졸브를 포함한다. <표 1>은 각 방법으로 검출한 장면전환 추출의 정확도를 보여준다. <표 1>에서  $N_c$ 는 정확하게 검출한 장면전환의 개수를 의미하고,  $N_f$ 는 잘못 검출한 장면전환의 개수를 의미하며,  $N_m$ 은 검출하지 못한 장면전환의 개수를 의미한다.

<표 1> 장면전환 검출의 정확도

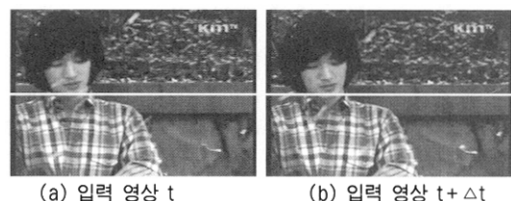
Method	cut			fade in(out)			dissolve		
	$N_c$	$N_m$	$N_f$	$N_c$	$N_m$	$N_f$	$N_c$	$N_m$	$N_f$
intensity histogram	36	3	14	0(0)	8(8)	0(0)	3	5	2
edge-based	34	5	8	4(6)	4(2)	2(8)	5	3	3
rule-based	35	4	4	6(6)	2(2)	1(1)	4	4	8
proposed method	37	2	2	6(6)	2(2)	1(0)	6	2	3

컷 장면전환의 경우 기존의 방법과 제안한 방법 모두 존재하는 컷을 정확하게 검출하는 편이다. 그러나 제안한 방법은 기존의 방법보다 장면전환이 발생하지 않는 부분을 컷으로 오 검출하는 비율이 현저히 낮다. 기존의 방법 중 히스토그램을 이용하는 방법은 카메라의 동작이나 물체의 움직임에는 덜 민감한 반면 영상의 밝기 변화에 매우 민감하므로 오 검출한 컷의 비율이 높다. 그리고 에지를 이용한 방법과 규칙 생성을 이용한 방법은 영상의 밝기 변화에는 덜 민감한 반면 카메라의 동작과 물체의 움직임으로 인한 오 검출이 발생한다. 이에 비해, 제안한 방법은 영상의 밝기 변화, 카메라의 동작, 이동 물체의 움직임에 덜 민감하므로 잘못 검출한 컷의 비율이 현저히 낮다. 그러나 기존의 방법과 제안한 방법 모두 플래시와 같이 조명이 급격히 변하는 영상에서는 컷으로 오 검출하는 부분이 존재한다.

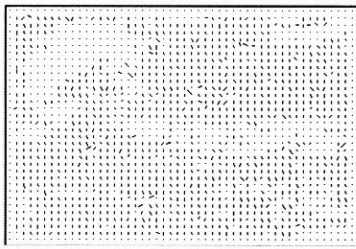
페이드 장면전환의 경우 히스토그램을 이용하는 방법은 페이드를 전혀 검출하지 못하고 대부분 컷으로 오 검출한다. 에지를 이용하는 방법은 카메라의 동작이 포함된 페이드가 발생하는 경우를 제외하고는 페이드를 어느 정도 올바르게 검출한다. 그러나 장면전환이 발생하지 않는 동일한 장면에서 페이드로 오 검출하는 비율이 높다. 규칙 생성을 이용하는 방법과 제안한 방법은 좋은 페이드 특징을 사용하므로 페이드를 올바르게 검출하는 편이다.

디졸브 장면전환의 경우 히스토그램을 이용하는 방법은 임계값을 사용하므로 영상에 민감하게 반응한다. 즉, 디졸브를 검출하기 위해 설정한 임계값 이하에서도 디졸브가 발생하는 경우가 많으므로 디졸브를 장면전환이 발생하지 않는다고 오 검출하는 경우가 많다. 에지를 이용하는 방법과 규칙 생성을 이용하는 방법은 물체가 이동하면서 디졸브 장면전환이 발생하는 경우를 제외하고는 디졸브를 잘 검출하는 편이다. 제안한 방법 역시 디졸브 장면전환이 천천히 발생하는 경우를 제외하면 디졸브를 잘 검출하는 편이다.

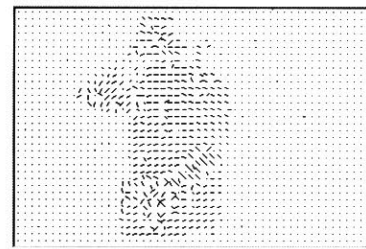
본 논문에서는 카메라의 동작과 이동 물체가 포함된 영상에 대한 실험결과만을 예로 보여준다. (그림 6)은 수직 방향으로 상승하는 카메라의 틸팅 동작이 발생하는 입력 영상을 보여주고, (그림 7)은 입력 영상으로부터 추출한 동작벡터를 보여준다. 그리고 (그림 8)은 동작벡터를 이용하여 카메라의 동작을 보정한 영상을 보여준다. (그림 8)에서 검은색의 블록으로 표시된 영역은 적응적 강건 예측에서 제거된 오류 자료를 나타낸다. (그림 6)과 (그림 8)의 영상 내에 표시된 흰색의 수평선은 영상의 변화를 보다 쉽게 시각적으로 보여준다. (그림 8)의 두 영상은 카메라의 동작을 보정한 영상이므로 거의 유사함을 확인할 수 있다.



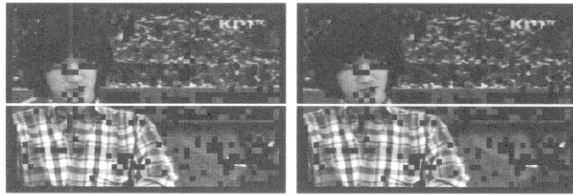
(그림 6) 카메라의 틸팅 동작이 발생하는 입력 영상



(그림 7) 동작 벡터

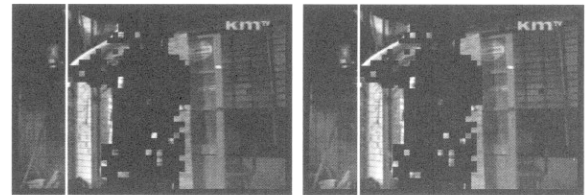


(그림 11) 동작 벡터



(a) 보정된 입력 영상 t (b) 보정된 입력 영상 t + Δt

(그림 8) 카메라의 동작을 보정한 입력 영상

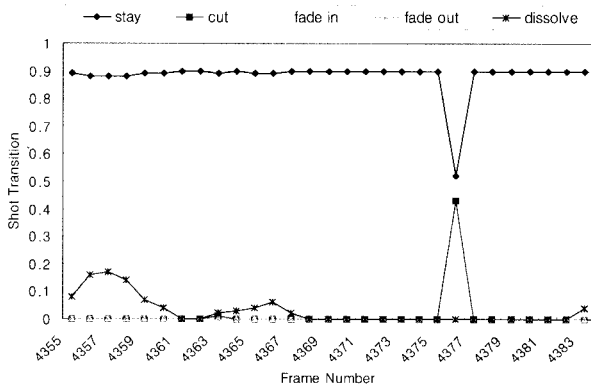


(a) 보정된 입력 영상 t (b) 보정된 입력 영상 t + Δt

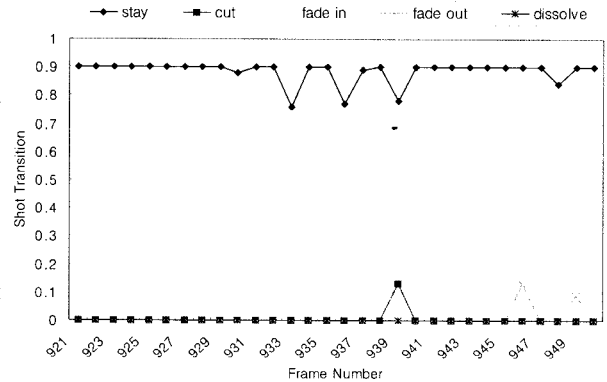
(그림 12) 카메라의 동작을 보정한 입력 영상

기존의 방법은 카메라의 동작으로 인해 장면전환이 발생한다고 오 검출하는 반면 제안된 방법은 영상을 보정하여 장면전환을 검출하므로 (그림 9)와 같이 오 검출이 발생하지 않는다.

기존의 방법들은 (그림 10)의 영상에 대해 이동하는 물체로 인해 장면전환이 발생한다고 오 검출하는 반면 제안된 방법은 영상을 보정하여 장면전환을 검출하므로 (그림 13)에서와 같이 오 검출이 발생하지 않는다.



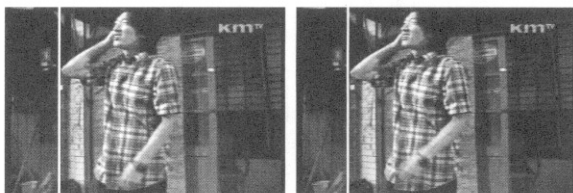
(그림 9) 장면전환의 확신도



(그림 13) 장면전환의 확신도

(그림 10)은 왼쪽으로 이동하는 물체가 포함된 입력 영상을 보여주고, (그림 11)은 입력 영상으로부터 추출한 동작 벡터를 보여준다. 그리고 (그림 12)는 동작벡터를 이용하여 카메라의 동작을 보정한 영상을 보여준다.

본 논문에서는 본론 부분에서 언급한 임계치들을 실험적으로 결정하였다. 그리고 (식 14)와 같은 정확도 척도를 사용하여 장면전환의 성능도 평가하였다[12, 13].



(a) 입력 영상 t (b) 입력 영상 t + Δt

(그림 10) 이동하는 물체가 포함된 입력 영상

$$R_{precision} = \frac{N_c}{N_c + N_f} \tag{14}$$

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

(식 14)에서  $R_{precision}$ 은 영상에서 검출된 전체 장면전환 중에서 정확하게 검출된 장면전환의 비율을 의미하고,  $R_{recall}$ 은 영상에 존재하는 전체 장면전환 중에서 정확하게 검출된 장면 전환의 비율을 의미한다. 그리고 <표 2>와 <표 3>은



<표 1>로부터 (식 14)를 이용하여 각 장면전환에 대해 추출한  $R_{precision}$ 과  $R_{recall}$ 을 보여준다.

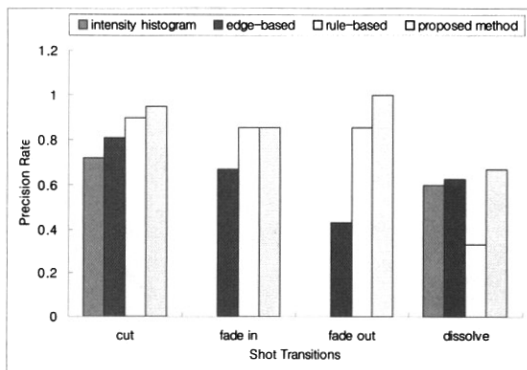
<표 2>  $R_{precision}$

Method	cut	fade in	fade out	dissolve
intensity histogram	0.720000	0.000000	0.000000	0.600000
edge-based	0.809524	0.666667	0.666667	0.625000
rule-based	0.897436	0.857143	0.857143	0.333333
proposed method	0.948718	0.857143	0.857143	0.666667

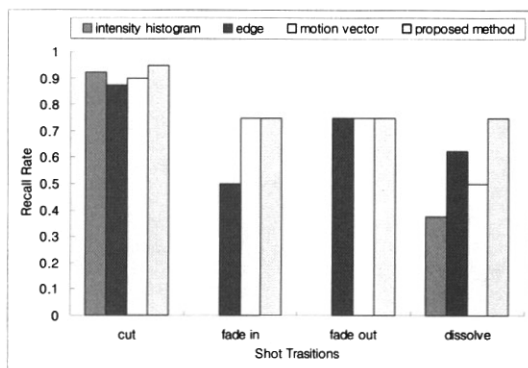
<표 3>  $R_{recall}$

Method	cut	fade in	fade out	dissolve
intensity histogram	0.923077	0.000000	0.000000	0.375000
edge-based	0.871795	0.500000	0.750000	0.625000
rule-based	0.897436	0.750000	0.750000	0.500000
proposed method	0.948718	0.750000	0.750000	0.750000

(그림 14)와 (그림 15)는  $R_{precision}$ 과  $R_{recall}$ 의 추출 결과를 도식적으로 보여준다. (그림 14)와 (그림 15)에서 히스토그램을 이용한 방법은 페이드를 전혀 검출하지 못하므로 페이드 인과 페이드 아웃에 대한 정확도의 값이 0이 되므로 그래프에는 나타나지 않는다.



(그림 14)  $R_{precision}$  그래프



(그림 15)  $R_{recall}$  그래프

## 7. 결 론

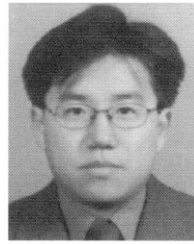
본 논문에서는 비디오 테이터를 효과적으로 분할하기 위해 카메라의 동작을 보정하여 장면 사이의 경계를 검출하는 장면전환 검출 방법을 제안하였다. 본 논문에서는 영상 내에 포함된 카메라의 동작 정보를 이용하여 영상을 보정하고, 보정된 영상으로부터 특징을 추출하여 장면전환을 검출하였다. 따라서 카메라의 동작으로 인해 기인하는 여러 가지 오검출을 방지한다. 또한, 영상을 보정하는 과정에서 지역적인 이동 물체의 동작을 제거하므로 이동 물체의 동작으로 인해 기인하는 장면전환의 오 검출도 방지한다. 실험에서는 다양한 비디오 테이터를 입력 받아 기존의 장면전환 검출 방법들과 제안한 방법의 성능을 비교 분석함으로써 제안한 방법의 우수함을 보였다.

향후에는 장면전환에 대한 각 특징의 확신도 추출 및 규칙 생성을 자동화할 예정이다. 또한 본 논문에서 검출한 컷, 페이드 인, 페이드 아웃, 디졸브 장면전환, 그리고 줌과 같은 카메라의 동작 뿐만 아니라 slide, wipe, Iris 등의 장면전환 검출도 시도할 예정이다.

## 참 고 문 헌

- [1] M. S. Lee, Y. M. Yang, S. W. Lee, "Automatic Video Parsing Using Shot Boundary Detection and Camera Operation Analysis," *Pattern Recognition*, Vol.34, No.3, pp.711-725, 2001.
- [2] Hong Heather Yu, Wyne Wolf, "Multi-Resolution Video Segmentation Using Wavelet Transformation," *Storage and Retrieval for Image and Video Databases*, SPIE Vol.3312, pp.176-187, 1998.
- [3] HongJiang Zhang, Atreyi Kankanhalli, Stephen W. Smoliar, "Automatic Partitioning of Full-Motion Video," *Multimedia Systems*, Vol.1, No.1, pp.10-28, 1993.
- [4] Ramin Zabih, Justin Miller, Kevin Mai, "A Feature-Based Algorithm for Detecting and Classifying Production Effects," *Multimedia Systems*, Vol.7, No.2, pp.119-128, 1999.
- [5] Seok-Woo Jang, Cheol-Ho Moon, Hyung-Il Choi, "Shot Transition Detection with FAM-Based Fuzzy Inference," *IEEE International Conference of Fuzzy Systems*, Vol.2, pp. 869-874, 1999.
- [6] 장석우, 김봉근, 김계영, 최형일, "동작 정보 추출을 위한 가변적 탐색 영역과 블록 크기의 정합," 정보처리학회논문지B, 제 10-B권, 제3호, pp.321-328, 2003.
- [7] 장석우, 김진욱, 이근수, 김계영 "적응적 가중치 함수를 이용한 모션 벡터의 필터링," 한국정보과학회논문지, 제31권, 제11호, pp.1474-1482, 2004.
- [8] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, *Numerical Recipes in C*, Second Edition, Cambridge University Press, 1992.

- [9] Jong-Il Park, Choong-Woong Lee, "Robust Estimation of Camera Parameters from Image Sequence for Video Composition," *Signal Processing*, Vol.9, pp.43-53, 1996.
- [10] Frank Perez and Christof Koch, "Toward Color Image Segmentation in Analog VLSI : Algorithm and Hardware," *International Journal of Computer Vision*, Vol.12, No.1, pp. 17-42, 1994.
- [11] Joseph Giarratano, *Expert Systems : Principles and Programming*, Third Edition, PWS Publishing Company, 1998.
- [12] Nilesh V. Patel, Ishwar K. Sethi, "Video Shot Detection and Characterization for Video Databases," *Pattern Recognition*, Vol.30, No.4, pp.583-592, 1997.
- [13] Dan Lelescu, Dan 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.

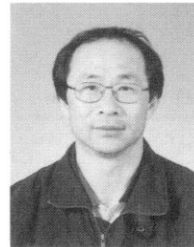


### 장석우

e-mail : swjang@kict.re.kr

1995년 숭실대학교 전자계산학과(공학사)  
1997년 숭실대학교 전자계산학과(공학석사)  
2000년 숭실대학교 컴퓨터학과(공학박사)  
2003년~현재 한국건설기술연구원 건설정보화연구부 선임연구원

관심분야: 컴퓨터 비전, 영상 처리, 동작 이해, 비디오 검색 등



### 최형일

e-mail : hic@computing.ssu.ac.kr

1979년 연세대학교 전자공학과(공학사)  
1982년 미시간대학교 전산공학과(공학석사)  
1987년 미시간대학교 전산공학과(공학박사)  
1987년~현재 숭실대학교 미디어학부 교수

관심분야: 컴퓨터 비전, 패턴 인식, 퍼지 이론, 비디오 검색 등