

개선된 적응적 임계값 결정 알고리즘과 후처리 기법을 적용한 동영상 분할 방법

원인수[†], 이준우^{**}, 임대규^{***}, 정동석^{****}

요 약

급격하게 증가하고 있는 동영상의 관리 도구로써 동영상을 계층적이고 구조적으로 구분하는 동영상 분할은 색인, 요약, 검색 등 다양한 응용 분야에서 공통적으로 적용될 수 있는 핵심 기술이라 할 수 있다. 기존의 적응적 임계값을 사용하는 동영상 분할 방법은 연속되는 프레임 간의 차이 값과 일정 간격의 크기를 갖는 윈도우에서 임계값을 계산하여 동영상 분할을 수행하였다. 그러나 이 경우, 윈도우의 크기보다 전환점의 발생 간격이 짧거나, 주변의 차이 값이 변동이 많으면, 정확한 검출을 하지 못한다. 상기 문제점을 개선하기 위하여 본 논문에서는 동영상의 포맷에 따라 윈도우의 크기를 결정하고, 윈도우 안에서 가중치를 사용하여 주변 값의 변화에 민감하게 반응하는 개선된 적응적 임계값 결정 알고리즘을 제안한다. 또한 카메라 불빛과 큰 물체의 빠른 움직임 등에 의한 오검출을 줄이기 위해 후처리 기법을 적용하였다. 실험을 통해서 제안된 방법은 기존 방법과 비교하여 3.7%의 성능 향상을 보이며, 변형된 동영상에서 95.5%의 재현성을 갖는 것을 확인하였다. 따라서 제안된 동영상 분할 방법은 기존 방법과 비교하여 정확성이 높고, 다양한 변형에도 재현성을 가지므로 동영상 관리 도구로써 많은 응용 분야에 적용할 수 있다.

Video Segmentation Method using Improved Adaptive Threshold Algorithm and Post-processing

In Su Won[†], Jun Woo Lee^{**}, Dae Kyu Lim^{***}, Dong Seok Jeong^{****}

ABSTRACT

As a tool used for video maintenance, Video segmentation divides videos in hierarchical and structural manner. This technique can be considered as a core technique that can be applied commonly for various applications such as indexing, abstraction or retrieval. Conventional video segmentation used adaptive threshold to split video by calculating difference between consecutive frames and threshold value in window with fixed size. In this case, if the time difference between occurrences of cuts is less than the size of a window or there is much difference in neighbor feature, accurate detection is impossible. In this paper, Improved Adaptive threshold algorithm which enables determination of window size according to video format and reacts sensitively on change in neighbor feature is proposed to solve the problems above. Post-Processing method for decrement in error caused by camera flash and fast movement of large objects is applied. Evaluation result showed that there is 3.7% improvement in performance of detection compared to conventional method. In case of application of this method on modified video, the result showed 95.5% of reproducibility. Therefore, the proposed method is more accurate compared to conventional method and having reproducibility even in case of various modification of videos, it is applicable in various area as a video maintenance tool.

Key words: Video Segmentation(동영상 분할), Cut Detection(전환점 검출), Adaptive Threshold(적응적 임계값), Post-processing(후처리 기법)

* 교신저자(Corresponding Author): 정동석, 주소: 인천시 남구 용현동 253(402-751), 전화: 032)860-7415, FAX: 032)868-3654, E-mail: dsjeong@inha.ac.kr
접수일: 2009년 11월 18일, 수정일: 2010년 2월 17일
완료일: 2010년 4월 17일

[†] 정회원, 인하대학교 전자공학과 박사과정
(E-mail: woninsu@inha.edu)

^{**} 정회원, 인하대학교 전자공학과 박사과정
(E-mail: junwoo@inha.edu)

^{***} 정회원, 인하대학교 전자공학과 박사과정
(E-mail: neptune2009@inha.edu)

^{****} 인하대학교 전자공학과 교수

※ 이 논문은 인하대학교 교내 학술연구비의 지원을 받아 수행된 연구임.

1. 서 론

최근 방송 콘텐츠와 인터넷의 동영상 서비스의 증가, 그리고 다양한 멀티미디어 기기들의 유통으로 인해 동영상은 다양한 분야에서 생산되며 활용되고 있다. 또한 이러한 환경에서 일반인들도 미디어를 언제든 손쉽게 접할 수 있는 환경이 제공되었으며, 이에 따른 무분별한 복제와 불법 유통으로 인해 심각한 사회적 문제가 발생하고 있다. 이렇게 무분별하게 급증하고 있는 동영상에 대한 효율적인 관리 방법에 대한 연구가 오랜 시간 연구되어 왔고, 동영상의 특성상 방대한 자료의 처리 문제, 다양한 형식의 갖는 동영상의 관리 방법에 대한 문제가 제기되고 있다.

동영상을 구성하는 기본 단위는 프레임(Frame)이라 부르는 영상 정보이나, 이를 모두 처리하기엔 어려움이 많다. 따라서 동영상을 계층적이고 구조적으로 구분하기 위해서 색인, 요약, 검색 및 식별 등 다양한 동영상 응용처리에 적용할 수 있는 동영상 분할은 필수적인 작업이라 할 수 있다[1]. 동영상 분할에 대한 방법은 샷 기반(Shot based), 장면 기반(Scene based), 객체 기반(Object based), 클러스터링 기반(Clustering) 등 다양한 방법이 있으나, 동영상의 특징을 고려하였을 때, 빠르고 직관적으로 관찰이 가능한 샷 기반의 동영상 분할이 가장 적합하다고 할 수 있다. 샷 기반의 동영상 분할 방법에서 샷의 전환점을 검출하는 작업을 전환점 검출(Cut Detection)이라고 하며, 전환점은 크게 급작스런 전환(Abrupt Transition)과 점진적인 전환(Gradual Transition)으로 구분된다. 그러나 점진적인 전환은 그 기준이 명확하지 않으며, 현재의 기술로는 검출 능력이 한계를 보이고 있어, 동영상 관리 도구로 사용되기에는 제약 사항이 많다. 따라서 본 논문은 점진적인 전환을 이용한 동영상 분할은 하지 않고, 명확히 구분이 되는 급작스런 전환점 검출만을 수행하며, 다양한 변형에서도 일관성과 재현성을 갖는 동영상 분할 방법에 대해 제안한다.

샷을 기반으로 하는 동영상 분할을 위해 다양한 방법들이 연구되어 왔다[2,3]. 가장 보편적인 전환점 검출 방법은 연속적인 프레임들 간에 공간적인 특징에 의해 나타나는 차이 값을 계산하여 판단하는 것이다. 그러나 이러한 방법들은 정확성은 보장이 되지만, 다양한 크기와 형식을 가진 동영상에서의 일관성 문제와 변형된 동영상에서의 재현성에 대한 문제가 있다 [4]. 또한 뉴스 등과 같은 특정 동영상에서 빈번하게

발생되는 카메라 불빛 등으로 인한 오검출 문제가 발생한다. 따라서 본 논문에서는 동영상의 시공간적 특성과 주변 특징 값들의 통계학적 분석을 이용한 전환점 검출 기법을 제시하여 높은 정확성과 변형된 동영상에서도 재현성을 갖는 것에 그 주안점을 두었다. 그 방법으로 연속되는 프레임에서 χ^2 검정을 이용하여 컬러 히스토그램 차이 값을 계산하고 주변 특징들의 변동에 민감한 개선된 적응적 임계값 결정 방법으로 후보 전환점을 선정한다. 그리고 카메라 불빛 등으로 인한 오검출과 동영상 특성의 변형 등으로 인해 발생하는 오검출을 해결하기 위해 전환점 주변 동영상의 공간적인 순차성을 이용하여 후보 전환점의 유효성 검사를 하여 선정된 후보 전환점의 정확성을 판단한다.

본 논문의 VI장으로 구성되며, II장에서는 기존의 동영상 분할 방법을 소개를 하고, III장에서는 동영상 분할을 위한 특징 추출 방법, 개선된 적응적 임계값 결정 알고리즘, 그리고 후처리 기법을 제안한다. IV장에서는 제안된 방법의 성능 평가와 기존 방법과 비교한 실험 결과를 분석한다. 그리고 마지막으로 V장에서는 결론 및 향후 과제에 대해 기술한다.

2. 기존의 동영상 분할 방법

동영상은 연속되는 유사한 프레임들로 하나의 샷이 구성되며, 연속된 샷들은 하나의 장면으로 구성된다. 샷 간의 급작스런 전환점을 검출하는 것을 전환점 검출이라고 한다[5]. 그림 1은 동영상의 기본적인 구조를 나타낸다.

기존에 발표된 전환점 검출 방법은 연속된 두 프레임간의 공간적인 영상 정보 차이를 이용하는 방법이 대부분을 이루고 있으며 대표적인 방법은 다음과 같다.

가장 빠르고 간단한 방법은 차이 값이 특정 임계

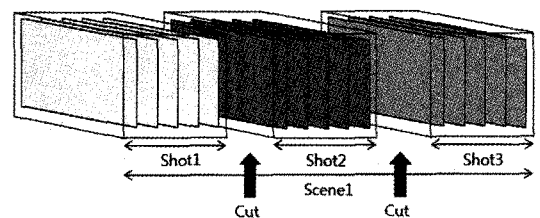


그림 1. 동영상의 기본 구조

값 이상으로 변하는 모든 화소의 수를 계산하여 비교 (Pixel Comparison)하는 방법이다[6,7]. 그러나 이 방법은 카메라와 객체의 움직임, 명암 변화 등에 매우 민감하여 검출 성능이 떨어진다. 블록 기반 비교 방법(Block-based Comparison)[8]은 카메라와 객체의 움직임에 민감하지 않도록 공간적 특성을 이용한 방법으로, 각 프레임을 블록으로 나누고, 연속되는 두 프레임 간의 일치하는 블록들에 대한 화소수를 비교하여 변화를 계산하는 방법으로 화소 비교 방법과 비교하여 작은 움직임에는 좋은 성능을 보이지만, 큰 객체의 움직임이나 갑작스런 명암 변화 등에서 오검출을 하게 된다.

히스토그램 비교 방법(Histogram comparison)은 최근 전환점 검출을 위해 가장 많이 사용되는 방법이다. Tonomura 등[9]은 회색 음영(Grey-level) 히스토그램 비교를 통하여 특정 임계값을 기준으로 전환점을 검출하는 간단한 방법을 제시하였고, Ueda 등[10]은 전환점 검출을 위해 컬러 히스토그램의 변화율을 판단 기준으로 사용하였다. Naghade 등[11]은 히스토그램을 이용한 통계학적 비교를 수행하였고, H.J. Zhang 등[6]은 앞선 여러 가지 방법과 히스토그램 방법을 비교하여 히스토그램 방법이 정확성과 속도 사이의 좋은 요소임을 실험으로 증명하였다. Garg 등[12]은 6개의 다른 컬러 모델들을 이용하여 성능을

평가하였으며, Tsamura 등[13]은 컬러 밀집 벡터 (Color Coherence Vector)를 이용하는 방법을 제시하였다. Naghade 등[11]은 히스토그램 교차 비교 (Intersection) 방법으로 두 특징 정보의 거리를 계산하고, 다항식 접근 방식으로 정합이 된 히스토그램을 압축하여 사용하였다.

앞선 방법들은 간단하고 카메라의 느린 이동과 작은 객체의 움직임에는 민감하지 않는 장점을 가지고 있지만, 전환점 판단을 위해 고정된 임계값을 사용함으로써, 다양한 동영상의 형식에 민감한 결과를 보이며, 비교적 유사한 두 개의 다른 영상이 주어지면 전환점 검출에 실패하거나, 갑작스런 명암 변형(flash light)나 큰 객체의 움직임, 카메라의 빠른 이동 등에 의해 오검출이 많아지는 경우가 빈번하게 발생한다. 또한 동일한 동영상이라 할지라도 동영상 복제에 의해 특징 정보가 변형이 되어 재현성이 낮은 결과를 가져오게 된다. 이런 문제들을 해결하기 위해 많은 연구가 이루어졌으며, 특히 D. Zhang 등[14]은 특정한 고정 임계값을 사용하지 않고, 슬라이딩 윈도우를 이용한 적응적인 임계값 결정 알고리즘을 제안하여 다양한 형식의 동영상에서 차이 값의 분포를 이용하여 직접적이고 적응적으로 임계값을 결정하는 동영상 분할 방법을 제시하였다. 그림 2는 D. Zhang 등이 제안한 적응적 임계값 결정 방법을 나타내고 있다. 슬

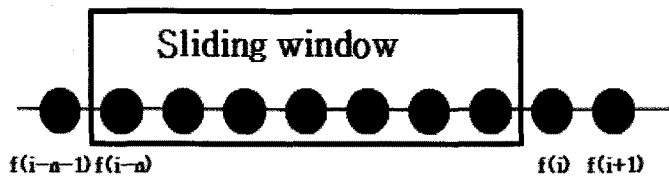
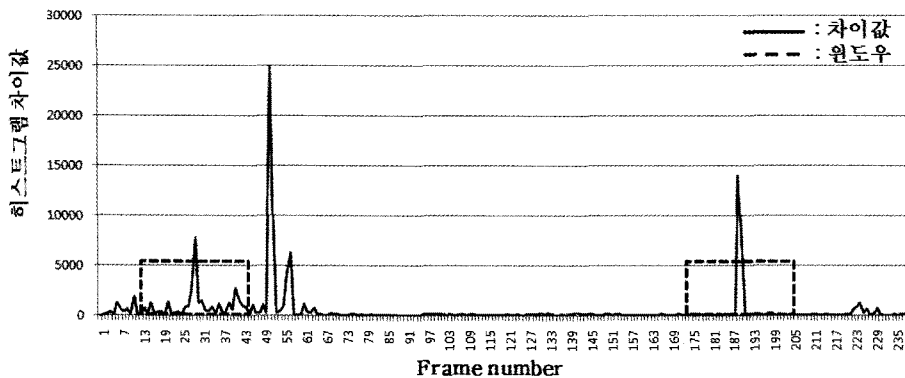


그림 2. D. Zhang 등[14]이 제안한 적응적 임계값 결정 알고리즘

라이딩 윈도우 내의 히스토그램 차이 값들의 평균과 편차를 계산하여 임계값을 결정하고 현재 프레임의 차이 값이 계산된 임계값보다 높으면 전환점이라고 판단한다. 또한 현재 프레임을 기준으로 좌, 우로 나누어 밝기 값의 평균을 비교하여 카메라 불빛을 검출하는 후처리 기법을 사용하였다.

그러나 D. Zhang의 방법은 고정된 윈도우 크기를 사용함으로써, 동영상의 변형이 가해지면 임계값이 변하게 되며, 윈도우 내의 모든 차이 값을 동일하게 사용함으로써 주변 값들의 변동에 대해 취약하게 된다. 즉, 광고와 뮤직비디오 등 특정 동영상에서 슬라이딩 윈도우의 크기보다 짧은 간격으로 전환점이 일어나는 경우 임계값이 급격하게 늘어나 전환점을 검출하지 못하거나, 한 개의 전환점만 검출해 내는 결과를 가져온다. 또한, 좌, 우 프레임의 밝기 값의 평균을 비교하여 카메라 불빛을 판단하는 후처리 기법은 전환점을 중심으로 구분된 샷과 샷이 유사한 밝기로 구성되었을 때는 정확한 전환점을 검출하지 못하는 결과를 보여줄 수 있다. 실제로 이런 경우는 매우 빈번하게 발생한다. 제안한 동영상 분할 방법은 이런 상황들을 고려하여 적응적 임계값 결정시 고정 크기의 윈도우를 사용하지 않고 동영상의 재생 시간 정보를 받아 사용함으로써 동영상의 변형되었을 때도 동일하게 적용이 가능하며, 윈도우 내의 차이 값을 이용할 때, 현재의 프레임과의 가중치를 적용하여 급격한 차이 값의 변동에 민감하게 동작하게 된다. 또한 주변 프레임의 공간적인 변동까지 고려한 후처리 기법은 카메라 불빛 뿐 아니라 큰 물체의 움직임에 의한 오검출도 판단이 가능하다.

3. 제안한 동영상 분할 방법

그림 3은 본 논문에서 제안하고 있는 전체적인 동영상 분할 방법을 나타내고 있다. 동영상이 입력되면 크기에 상관없는 결과를 위해 CIF 크기의 균등화를 적용한다. 동영상의 특징 점으로 히스토그램을 사용하며, 차이 값을 비교하기 위해 χ^2 검정을 적용하였다. 또한 특정한 임계값을 사용하지 않고, 주변 값들의 유사성을 통계적으로 분석하는 개선된 적응적 임계값 결정 방법과 후보 전환점의 유효성 판단을 위해 동영상의 순차성을 이용한 후처리 기법을 혼합하는 동영상 분할 방법을 제안한다.

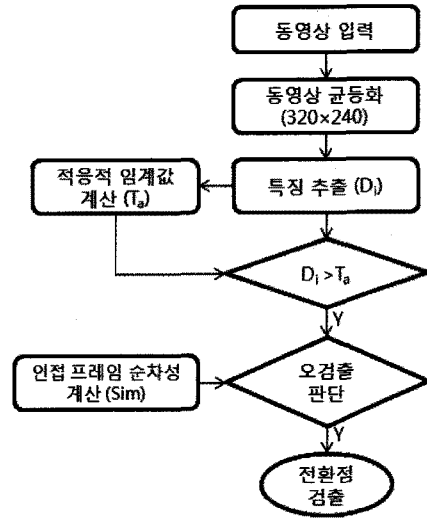


그림 3. 제안한 동영상 분할 방법

3.1 전환점 검출을 위한 특징 추출 방법

디지털 동영상은 RGB 컬러 모델로 표현되는데, 본 논문에서는 PAL(Phase Alternating Line), NTSC(National Television Standards Committee) 등의 표준에 사용되는 YUV 컬러 모델로 변환하여 사용하였다. YUV 컬러 모델은 사람의 눈이 색차보다는 밝기에 민감하다는 사실에 착안한 방식으로, 밝기(Luminance)인 Y와 색차(Chrominance)인 U와 V로 구분한다. YUV 컬러 모델은 일반적으로 컬러 히스토그램을 전환점 검출 방법에 좋은 성능을 보이고 있다고 알려져 있다[12]. 본 논문에서 사용한 YUV의 변환식은 다음 식 (1)과 같다.

$$\begin{aligned}
 Y &= 0.39R + 0.59G + 0.11B \\
 U &= (B - Y) \times 0.493 \\
 V &= (R - Y) \times 0.877
 \end{aligned} \tag{1}$$

히스토그램 비교를 통한 차이 값은 χ^2 검정을 이용하여 계산한다. χ^2 검정은 연속된 프레임 사이의 히스토그램 차이 값을 강조할 뿐만 아니라 카메라와 객체의 움직임을 강조할 수 있는 특징이 있으며[11], 계산식은 식 (2)과 같다.

$$diff_k(k, k-1) = \sum_{i=1}^M \left(\frac{h_k(i) - h_{k-1}(i)}{h_k(i) + h_{k-1}(i)} \right) \tag{2}$$

$diff_k(k, k-1)$ 는 k번째 프레임과 k-1번째 프레임 간의 컬러 히스토그램 차이 값을 나타내며, $h_k(i)$ 는 k번째 프레임에서의 히스토그램을 계산한 것이며,

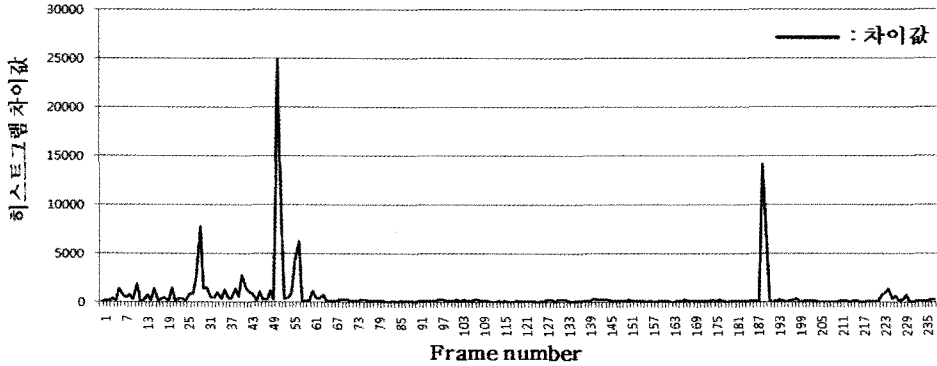


그림 4 연속된 프레임에서 χ^2 검정을 이용한 컬러 히스토그램 차이 값 그래프

M 은 히스토그램의 빈 수를 의미한다. 그림 4는 연속된 프레임에서 χ^2 검정을 이용한 컬러 히스토그램 차이 값 그래프를 나타낸 것이다. 결과를 보면 전환점과 전환점이 아닌 부분이 명확하게 구분됨을 알 수 있다.

3.2 개선된 적응적 임계값 결정 알고리즘

연속된 프레임에서의 컬러 히스토그램 차이 값은 각 동영상의 장르나 속성에 따라 편차가 크게 나타나게 된다. 이를 극복하기 위한 적응적 임계값 결정 알고리즘으로 동영상의 재생 시간 정보를 이용한 가변적인 크기를 갖는 대칭 윈도우를 사용하고, 대칭 윈도우에서 중심 프레임과의 간격에 따른 가중치를 적용하여 주변의 인접한 차이 값의 변동에 민감하게 반응하는 개선된 방법을 제안한다.

현재 프레임 번호가 k 라면, $[k, k \pm 0.5 s]$ 의 크기를 갖는 가변적인 슬라이딩 윈도우를 사용한다. 슬라이딩 윈도우 내에서 현재 프레임을 기준으로 좌우 0.5 초의 간격 안의 컬러 히스토그램 차이 값의 평균과 표준 편차를 계산하는데, 현재 프레임과 거리에 따라 차이 값에 가중치를 적용하여 평균과 표준 편차를 계산하게 된다. 가중치는 그림 5와 같이 멀수록 작은

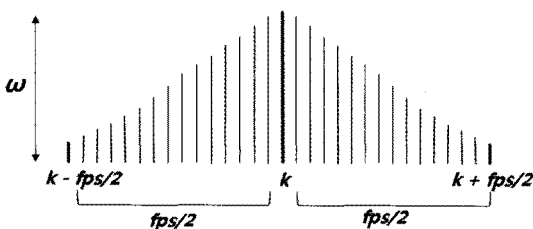


그림 5. 제안하는 슬라이딩 윈도우의 예

값을 갖도록 한다.

현재 프레임을 기준으로 왼쪽 윈도우에 속해 있는 차이 값들의 평균(μ_L)과 표준편차(σ_L)를 식 (4)와 (5)에 의해 구한다.

$$\mu_L = \frac{1}{N} \sum_{k=k-N}^{k-1} (0.1 \times k) \times diff_k(k, k-1) \quad (4)$$

$$\sigma_L = \sqrt{\frac{1}{N} \sum_{k=k-N}^{k-1} (0.1 \times k) \times (diff_k(k, k-1) - \mu_L)^2} \quad (5)$$

같은 방식으로 오른쪽 윈도우에 속해 있는 차이 값들의 평균(μ_R)과 표준편차(σ_R)를 식 (6)과 (7)에 의해 구한다.

$$\mu_R = \frac{1}{N} \sum_{k=k+1}^{k+N} (1 - 0.1 \times k) \times diff_k(k, k-1) \quad (6)$$

$$\sigma_R = \sqrt{\frac{1}{N} \sum_{k=k+1}^{k+N} (1 - 0.1 \times k) \times (diff_k(k, k-1) - \mu_R)^2} \quad (7)$$

구해진 평균과 표준편차를 이용하여 왼쪽 윈도우의 임계값(T_{left})과 오른쪽 윈도우의 임계값(T_{right}) 계산하기 위한 식 (8)와 식 (9)를 정의한다.

$$T_{Left} = \mu_{Left} + h \times \sigma_{Left} + \alpha \quad (8)$$

$$T_{Right} = \mu_{Right} + h \times \sigma_{Right} + \alpha \quad (9)$$

식 (8), (9)에서 h 는 표준 편차에 대한 가중치이고, α 는 임계값의 최소 허용을 나타내는 상수로써 각각의 값이 작게 되면 잘못된 컷이 검출될 수가 있고 (*false*), 값이 크게 되면 컷이 발생함에도 검출하지 못하는 경우(*miss*)가 발생한다. 최종적으로 기존의 방법과 동일하게 다음 두 가지의 조건을 만족할 때, 후보 전환점으로 결정한다.

1. 슬라이딩 윈도우 안에서 현재 프레임 k 의 히스토그램 차이 값이 최대일 때,
2. 현재 프레임 k 의 히스토그램 차이 값이 $T_{cut} = MAX(T_{left}, T_{right})$ 보다 클 때.

그림 6은 제안한 방법을 이용해 차이 값과 임계값을 추출해낸 그래프를 나타낸다. 그림을 보면 실제 컷(○ 표시)은 정확히 검출하고, 실제로 컷이 아닌 부분(× 표시)은 검출하지 않은 결과를 나타하고 있다.

3.3 동영상의 공간적인 순차성을 이용한 후처리 기법

연속된 히스토그램의 차이 값을 이용한 방법은 갑작스런 밝기의 변화나 큰 물체의 빠른 움직임에 매우 민감하여 오검출이 발생하는 문제점이 있다. 특히 카

메라 불빛 등으로 인해 밝기 값이 순간적으로 크게 변하였을 때 전환점이라고 판단하는 오류가 나타난다. 이를 해결하기 위해, 앞선 방법으로 선정된 후보 전환점의 유효성을 판단하기 위한 후처리 기법을 적용한다. 제안하는 후처리 기법은 동일한 샷 내의 프레임을 공간적으로 나누어 밝기 값의 평균을 구하고, 이를 큰 값부터 서열로 나타낼 때, 유사성을 가질 것이라는 특성을 이용한다. 이는 일반적으로 오디널 방법[1]이라고도 부르며, 동영상 복사 검출 방법의 특징으로 사용하기도 한다. 즉, 오검출된 후보 전환점의 주변 프레임은 동일한 샷에 포함되므로 공간적인 순차성이 유사할 것이라는 특징을 이용하여 오검출을 판단하는 방법이다.

그림 7은 수족관에서 찍은 홈 비디오의 예이다.

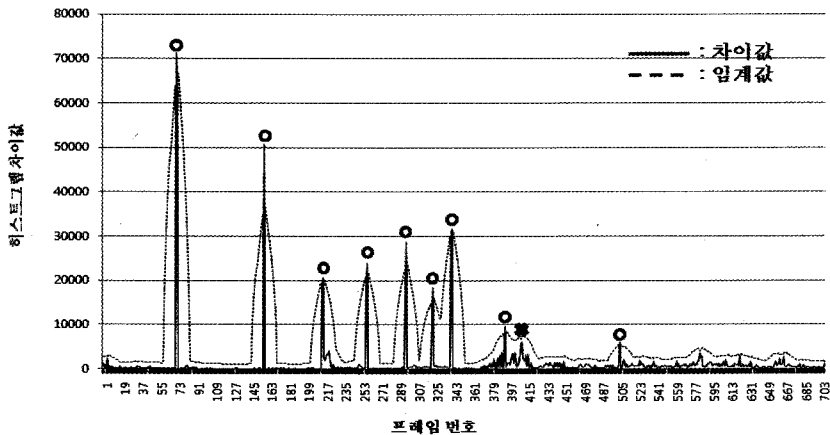
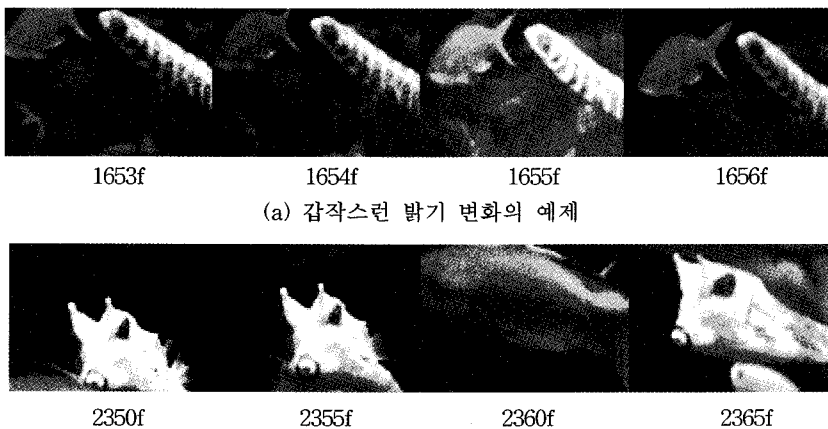


그림 6. 개선된 적응적 임계값 결정 알고리즘을 이용한 결과 그래프



(a) 갑작스런 밝기 변화의 예제
(b) 짧은 시간 간격에서 큰 물체의 움직임

그림 7. 오검출된 전환점의 예

그림 7(a)는 1655번 프레임에서 카메라 불빛(camera light)에 의해 급격한 밝기 변화가 발생하여 오검출이 발생한 예이며, (b)는 큰 물체의 빠른 움직임으로 인해 발생하는 오검출의 예이다. 이를 해결하기 위해 후보 전환점으로 선정된 프레임의 바로 앞과 뒤 주변 프레임들을 블록으로 나누어 밝기 값의 평균을 구하고 밝기 순으로 정렬하여 각 블록에 해당되는 서열을 할당한다. 그리고 각 프레임에서 동일한 위치의 블록 간에 서열의 차이를 계산하고 그 차이가 0인 블록의 횟수를 계산한다. 차이가 0인 블록이 4보다 많으면 유사한 영상이라 판단하여 컷을 검출하지 않고, 4 이하이면 정상적인 컷으로 판단한다. 다음 그림 8은 수행 절차의 예를 나타낸다.

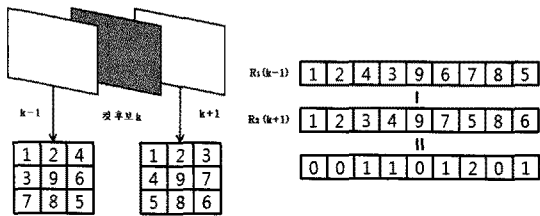


그림 8. 공간적인 순차성을 이용한 후처리 기법

4. 실험 결과 및 분석

4.1 실험 동영상 구성

본 논문의 성능 평가를 위해 제안된 동영상 분할 방법의 전환점 검출 실험과 변형 동영상에서의 재현성 실험을 수행하였다. 동영상 분할 성능 평가를 위해 사용한 실험 동영상은 지상파 방송의 TV 프로그램과 UCC 등과 같은 인터넷 동영상으로 드라마, 스포츠, 연예오락, 시사교양, 뉴스, 애니메이션, 홈비디오 등의 7개의 장르로 구성된 315개의 편으로서, 총 7시간 52분에 해당하는 분량이다. 표 1은 실험 동영상의 구성과 올바른 전환점의 수(Ground Truth), 그리고 전환점의 발생 빈도를 나타낸다. 또한 원본 동영상에서 100편의 동영상을 선정하여 각 편마다 표 2의 6개의 변형 유형을 적용하여 약 15시간에 해당하는 변형 동영상을 구성하고, 이를 이용한 실험을 통하여 변형 동영상에서의 재현성을 평가하였다. 표 2의 편집 유형은 일반적으로 동영상의 온라인 전송 및 배포용으로 적용되는 변형의 기본적인 형태로서, 재부호화(re-encoding)으로 인한 밝기 등의 동영상 특

표 1. 실험 동영상의 구성

장르	전체 프레임 수	올바른 전환점 수	전환점 발생주기(s)
드라마	97218	756	4.29
스포츠	48609	162	10.00
연예오락	145818	594	8.18
시사교양	48609	198	8.18
뉴스	97218	837	3.87
애니메이션	194427	1467	4.42
홈비디오	219141	360	20.29
총계	851040	4374	6.49

표 2. 재현성 평가를 위한 변형 동영상 유형

변형 효과	변형 설명
Brightness Change (BC)	전체 밝기 10% 증가
Frame-rate Reduction (FR)	프레임 올 감소 (15fps)
Monochrome (MONO)	흑백 영상 전환 (B/W)
Resolution Reduction (RR)	CIF (176×144)
Text and logo overlay (TLO)	10% 내에서의 로고 및 문자 삽입
Severe Compression (SC)	521kbps

징 변화(Brightness Change), 프레임 감소(Frame-rate Reduction), 흑백 전환(Mono), 해상도 변화(Resolution Reduction), 문자 및 로고 등의 추가(Text and Logo Overlay), 비트율의 변화(Severe Compression) 등 MPEG-7에서 정의하고 있는 변형 유형의 일부분을 채택하여 사용하였다.

4.2 동영상 분할 성능 평가

전환점 검출에 대한 실험 평가의 기준으로 Recall과 Precision, 그리고 이 둘을 혼합하여 전체적인 성능을 나타내는 $F_{measure}$ 를 사용하였으며, 그에 대한 정의는 식 (10), (11), (12)와 같다.

$$Recall = \frac{D}{D + D_{miss}} \times 100 \tag{10}$$

$$Precision = \frac{D}{D + D_{false}} \times 100 \tag{11}$$

$$F_{measure} = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{12}$$

D는 올바른 전환점의 개수(The number of ground

truth), D_{miss} 는 전환점임에도 검출하지 못한 개수(The number of miss detection), D_{false} 는 잘못된 전환점을 검출된 개수(The number of false detection)를 나타낸다. *Recall*은 올바른 전환점을 정확하게 검출한 비율을 나타내는 것으로, D_{miss} 가 한개도 없으면 1을 나타내고 D_{miss} 가 많아질수록 그 수치는 0에 가까워진다. *Precision*은 올바른 전환점이 아닌 위치를 잘못 검출한 비율을 나타내는 것으로, D_{false} 가 한개도 없으면 1을 나타내고, D_{false} 가 많아질수록 그 수치는 0에 가까워진다. 즉 두 평가 기준 모두 1에 가까울수록 성능이 좋음을 나타낸다.

본 실험을 위해 사용한 컬러 히스토그램은 64개의 빈을 갖도록 양자화 하였고, 적응적 임계값 결정 알고리즘에서 슬라이딩 윈도우의 크기 N 은 1초, h 는 2, α 는 1000을 사용하였다. 표 3은 제안된 방법을 이용한 컷 검출 결과를 나타낸다.

제안된 컷 검출 방법의 실험 결과로 *Recall*과 *Precision*은 각각 95.3%와 94.9%를 보이고 있다. 전체적인 전환점 검출 성능이 우수한 것을 알 수 있으나, 홈비디오의 경우 편집이 안 된 의미 없는 영상들로 구성되어 있어, 검출 결과에서 *false*가 많이 발생하는 특별한 경우로 볼 수 있다.

그림 9는 카메라 불빛이나 큰 물체의 급격한 움직임은 판단하여 오검출을 줄이기 위한 후처리 기법의 적용 전과 후의 결과를 나타낸다. 결과를 보면 후처리 방법을 사용함으로써, 잘못된 전환점의 개수가 줄어들어 *Precision*이 높아짐을 알 수 있다. 그러나 *Recall*이 0.7% 가량 감소하는 현상을 볼 수 있는데, 이는 샷과 샷 사이의 특징 차이가 작아 오검출 방법을 적용하여도 판단하지 못한 특별한 경우가 종종 존재하기 때문이다. 그러나 오검출 판단 방법을 적용

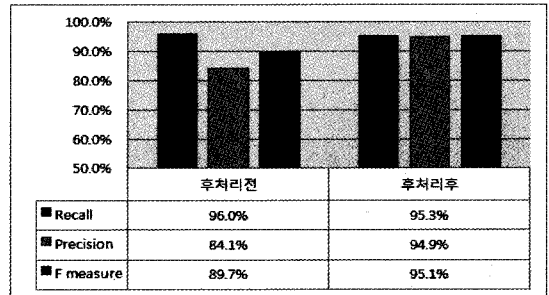


그림 9. 후처리 기법을 사용하기 전과 후의 비교

하는 것이 *Recall*은 다소 줄어들지만, 이에 비해 *Precision*이 약 10%의 성능 향상을 보임으로써, 전체적인 전환점 검출 성능이 높아짐을 나타낸다.

4.3 성능 비교 실험

제안한 방법의 객관적인 평가를 위해 기존 방법들과의 성능 비교 실험을 수행하였다. 컬러 히스토그램을 기반으로 하는 방법 중 고정 임계값을 사용한 HJ. Zhang[6]의 방법과 이에 대한 단점을 극복하기 위해 적응적 임계값 결정 방법을 제안한 D. Zhang[14]의 방법과 성능을 비교하여 평가하였다. HJ. Zhang의 방법은 고정된 임계값을 사용하여 전환점을 검출하는 히스토그램의 차이 값을 이용한 전환점 검출의 기본적인 방법이고, D. Zhang의 방법은 현 프레임 위치에서 슬라이딩 윈도우를 이용하여 주변 차이 값의 통계학적 분석을 통해 자동으로 임계값을 결정하며, 카메라 불빛 등으로 인한 오검출을 막기 위해 후처리 기법을 적용한 방법이다. 그림 10은 기존 방법과 제안한 방법과의 비교 결과를 나타낸 그림이다.

그림 10에서 보는 바와 같이 제안된 방법은 고정 임계값을 사용하는 H. Zhang의 방법보다 $F_{measure}$ 가

표 3. 제안된 방법을 이용한 컷 검출 결과

장 르	D	D_{miss}	D_{false}	<i>Recall</i>	<i>Precision</i>	$F_{measure}$
드 라 마	756	45	18	94.4%	97.7%	96.0%
스 포 츠	162	18	12	90.0%	93.1%	91.5%
연예오락	594	20	26	96.7%	95.8%	96.3%
시사교양	198	15	14	93.0%	93.4%	93.2%
뉴 스	837	27	34	96.9%	96.1%	96.5%
만화영화	1467	54	53	96.4%	96.5%	96.5%
홈비디오	360	36	77	90.9%	82.4%	86.4%
총 계	4374	215	234	95.3%	94.9%	95.1%

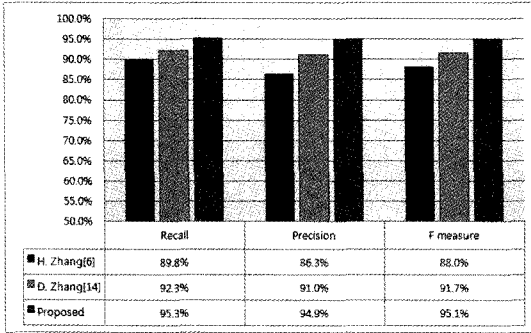


그림 10. 기존 방법과 제안된 방법의 검출 결과 비교 결과

7.1%, 적응적 임계값 결정 방법을 사용하는 D. Zhang의 방법에 비해 3.4% 우수한 것을 알 수 있다. 성능 평가의 분석 결과로 전환점이 매우 빠른 간격으로 재발생되었을 때, D. Zhang의 방법은 이를 검출하지 못하는 경우가 많지만, 제안된 방법은 이를 검출해 내었다. 또한 앞 뒤 샷의 밝기 값을 비교하여 오검출 판단을 하는 D. Zhang의 방법은 동영상의 특징상 전환점으로 나누어진 앞 뒤 샷들의 밝기 값 또한 유사한 경우가 매우 많다. 그러나 제안된 방법은 단순한 밝기 값의 비교가 아닌 공간적인 순차성을 이용하여, 샷들의 유사도를 판단하기 때문에, 좀 더 효과적인 오검출 판단을 할 수 있었다고 생각된다.

4.2 재현성 평가 결과

변형 동영상에 대한 재현성 평가를 위해 제안된 전환점 검출 방법으로 원본과 변형된 동영상에서 전환점을 각각 검출하였다. 표 4는 제안된 전환점 검출 방법의 재현성 실험 결과를 나타내며, D' 는 원본 동영상과 동일한 위치의 전환점을 검출한 개수, D'_{miss} 는 원본 동영상에서 검출한 전환점을 검출하지 못한 개수, D'_{false} 는 원본 동영상에서 검출하지 않은 전

표 4. 제안된 전환점 검출 방법의 재현성 실험 결과

변형 유형	D'	D'_{miss}	D'_{false}	Recall	Precision	$F_{measure}$
BC	1592	27	37	98.3%	97.7%	98.0%
FR	1599	90	142	94.7%	91.8%	93.2%
MONO	1608	49	40	97.0%	97.6%	97.3%
RR	1607	87	57	94.9%	96.6%	95.7%
SC	1604	90	65	94.7%	96.1%	95.4%
TLO	1554	129	92	92.3%	94.4%	93.4%
총 계	9564	472	433	95.3%	95.7%	95.5%

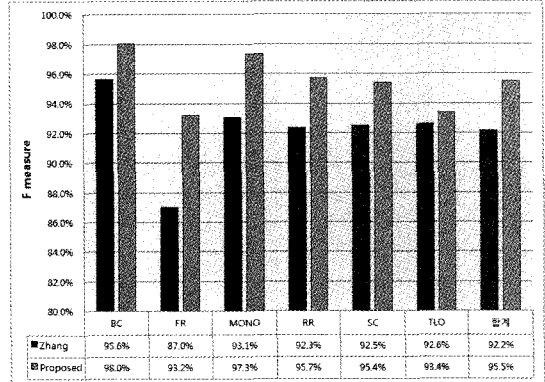


그림 11. 변형 유형에 따른 비교 실험 결과

환점을 검출한 개수를 나타낸다.

그림 11은 각각 변형 유형에 따른 재현성 평가 실험 결과와 동영상 장르에 따른 재현성 평가 실험 결과를 나타낸다. 실험 결과에서 제안된 방법의 $F_{measure}$ 는 95.5%의 성능을 보여주며, 이는 제안된 방법은 변형된 동영상에도 재현성이 높음을 알 수 있다. 또한 D. Zhang의 방법과 비교하여 재현성이 약 3.3%정도 높음을 보여준다. D. Zhang의 방법은 전체적으로 보아도 제안된 방법보다 재현성이 낮으며, 특히 FR에서 성능 차이가 두드러지는데, 이는 고정 크기의 윈도우를 사용하고 윈도우 내에서 동일한 차이 값의 비교를 통한 임계값 결정 방법을 사용하였기 때문이다. 이는 앞서 도출했던 것과 마찬가지로 동영상의 특성이 변경되어 전환점의 시간적 발생 간격이 변동되거나, 윈도우 내에서의 차이 값의 변동에 민감하게 반응하지 못한 결과라고 볼 수 있다.

5. 결론 및 향후 과제

본 논문은 대용량의 동영상을 효과적으로 관리하

기 위해 기존의 방법보다 성능이 좋고 변형에도 재현성이 높은 동영상 분할 방법에 대해 기술하였다. 동영상의 여러 특성 중 전환점은 동영상의 변형에도 강인하며, 검색 및 식별의 도구로서 동영상의 특징을 가장 잘 나타내는 특성이라 할 수 있다.

다양한 환경에서도 좋은 성능을 갖는 동영상 분할 방법을 위해 χ^2 검정을 이용하여 연속되는 프레임 간의 컬러 히스토그램 차이 값을 추출한다. 그리고 현재 프레임을 중심으로 가변적인 슬라이딩 윈도우를 사용하여 프레임간의 간격에 따라 가중치를 적용하여 차이 값의 평균과 표준 편차를 계산하는 개선된 적용적인 임계값 결정 알고리즘으로 후보 전환점을 선정한다. 그러나 컬러 히스토그램을 이용한 방법은 갑작스런 명암의 변경이나 큰 물체의 움직임 등으로 인해 오검출이 발생한다. 이러한 문제를 해결하기 위해 공간적인 순차적 방법을 이용하여 선정된 후보 전환점의 앞/뒤 프레임 간에 유사 판단을 하고, 유사도가 높으면 잘못된 검출 결과로 판단하는 후처리 기법을 적용하였다.

제안된 전환점 검출 방법의 검증하기 위해서 기존의 적용적 임계값 결정 알고리즘과 성능 비교 실험을 하였으며 기존의 알고리즘보다 약 3.4%의 성능 향상을 보이고 있음을 실험을 통하여 입증하였다. 또한 변형된 동영상에서 전환점을 검출하여 원본 동영상에서 검출한 전환점과 비교하였을 때의 성능은 95.5%를 보여줌으로써 제안된 전환점 검출 방법이 변형된 동영상에 대해서도 재현성이 높다는 것을 나타내었으며, 기존의 방법 대비 3.3%의 성능 향상을 나타내었다. 즉, 동영상 관리를 위한 전처리 단계로서 제안된 동영상 분할 방법은 효과적이라 할 수 있다.

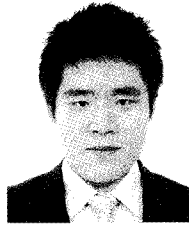
그러나 제안된 방법은 동영상이 임의의 편집에 의해 시퀀스 중간에 새로운 동영상이 삽입되거나 삭제가 되는 경우에는 전환점이 변경이 되는 문제가 발생한다. 또한 동영상의 특징 정보로 대표 프레임을 추출함에 있어 변형에도 재현성이 더 높은 방법에 대한 연구가 강구되어야 할 것이다.

참 고 문 헌

- [1] N. Bhat and K. Nayar, "Ordinal measures for image correspondence," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 20, No.4, pp. 415-423, 1998.
- [2] I. Koprinska and S. Carrato, "Temporal Video Segmentation: A Survey," *Signal Processing; Image Communication*, Vol.16, pp. 477-500, 2001.
- [3] G. Ananger and T.D.C. Little, "A survey of technologies for parsing and indexing digital video," *Journal of Visual Communication and Image Representation*, pp. 28-43, 1996
- [4] U. Gargi, R. Kasturi, and S. H. Strayer, "Performance Characterization of Video Shot Change Detection Methods," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.10, No.1, 2000.
- [5] C. Cotsaces, N. Nikolaidis, and I. Pitas, "Video shot detection and condensed representation: a review," *IEEE Signal Processing Magazine*, Vol.23, No.2, pp. 28-37, 2006.
- [6] HJ. Zhang, A. Kankanhalli, and S. W. Smoliar, "Automatic Partitioning of Full-motion Video," *Proceeding of the second ACM on Multimedia Systems*, Vol.1, No.1, pp. 10-28, 1993.
- [7] A. Hampapur, R. Jain, and T. Weymouth, "Digital Video Segmentation," *Proceeding of the second ACM on Multimedia*, pp. 357-364, 1994.
- [8] R. Kasturi, and R. Jain, "Dynamic Vision," *IEEE Computer Society Press in Computer Vision: Principles*, pp. 469-480, 1991.
- [9] Y. Tonomura, "Video handing based on structured information for hypermedia systems in," *International Conference Multimedia Information Systems*, pp. 333-344, 1991.
- [10] H. Ueda, T. Miyatake, and S. Yoshizawa, "IMPACT: An Interactive Natural-motion-picture Dedicated Multimedia Authoring System," *Proceeding of the ACM on Human Factors in Computing Systems*, pp. 343-350, 1991.
- [11] A. Nagasaka, and Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances," *Proceeding of the IFIP*

on Visual Database Systems II, pp. 113-127, 1992.

- [12] U. Gragi, R. Kasturi, and S. Antani, "Evaluation of video sequence indexing and hierarchical video indexing," Proceeding of the SPIE on Storage and Retrieval in Image and Video Databases, pp. 1522-1530, 1995,
- [13] E. Tsamoura, "Video Shot Meta-segmentation based on Multiple Criteria for Gradual Transition Detection," International Workshop on Content-based Multimedia Indexing, pp. 51-57, 2008
- [14] D. Zhang and HJ. Zhang, "A New Shot Boundary Detection Algorithm" Proceeding of the SPIE on Advances in Multimedia Information Processing, Vol.2195, pp. 63-70, 2001



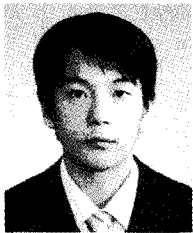
이 준 우

2001년 3월~2006년 2월 인하대학교 전자공학과 공학사
 2006년 3월~2008년 2월 인하대학교 전자공학과 공학석사
 2008년 3월~현재 인하대학교 전자공학과 박사과정
 관심분야: 비디오 코덱, 비디오 카피 디텍션



임 대 규

1998년 3월~2002년 8월 인하대학교 전자공학과 공학사
 2002년 9월~2004년 8월 인하대학교 전자공학과 공학석사
 2009년 3월~현재 인하대학교 전자공학과 박사과정
 관심분야: 비디오 코덱, 컴퓨터 비전



원 인 수

1999년 3월~2006년 2월 인하대학교 전자공학과 공학사
 2007년 3월~2009년 2월 인하대학교 전자공학과 공학석사
 2009년 3월~현재 인하대학교 전자공학과 박사과정
 관심분야: 비디오 코덱, 비디오 카피 디텍션



정 동 석

1977년 서울대학교 학사
 1985년 Virginia Tech 공학석사
 1988년 Virginia Tech 공학박사
 1988년~현재 인하대학교 전자공학부 교수
 2000년~2004년 정보전자공동연구소 소장
 관심분야: 컴퓨터 비전, 영상 처리, 포렌식 워터마킹