

논문 2006-43CI-6-10

플래시라이트에 강건한 장면전환 검출 알고리즘

(Robust Scene Change Detection Algorithm for Flashlight)

고 경 철*, 최 형 일*, 이 양 원**

(Kyong Cheol Ko, Hyung Il Choi, and Yang Weon Rhee)

요 약

비디오에 삽입된 플래시라이트는 연속된 프레임사이의 차이 값을 높게 형성하여 장면전환 검출을 위한 임계값 결정에 많은 어려움을 주며, 특히 장면전환 지점으로 잘못 검출되는 문제점을 가지고 있다. 따라서 본 논문에서는 이러한 플래시라이트를 연속된 프레임으로부터 효율적으로 제거하여 신뢰할 수 있는 장면전환 지점을 검출하는 강건한 장면전환 검출 알고리즘을 제안한다. 제안된 방법은 먼저, 연속된 프레임사이의 차이 값 추출을 위하여 객체나 카메라의 움직임에 덜 민감하고 프레임의 공간정보를 이용하는 지역 히스토그램 비교에 의한 추출 식을 사용하며, 추출된 차이 값들의 큰 변이 폭에 의한 임계 값 결정의 문제점을 해결하기 위하여 차이 값들의 동적 압축에 의한 정규화 작업을 수행한다. 또한 추출된 차이 값들의 시간적 연속성의 변이에 따라 플래시라이트가 가지는 특징을 이용하여 플래시라이트와 장면전환 검출을 따로 구분하여 추출할 수 있는 새로운 장면전환 검출 알고리즘을 제안한다. 제안된 방법은 플래시라이트가 들어있는 다양한 비디오 타입으로부터 실험되어졌으며, 실험결과 플래시라이트 검출에 높은 신뢰성과 효율성을 보여주었다.

Abstract

Flashlights in video has many problem to detect the scene change because of high difference values from successive frames. In this paper propose the reliable scene change detection algorithms by extracting the flashlights. This paper proposes a robust scene change detection technique that uses the weighted chi-square test and the automated threshold-decision algorithms. The weighted chi-square test can subdivide the difference values of individual color channels by calculating the color intensities according to NTSC standard, and it can detect the scene change by joining the weighted color intensities to the predefined chi-square test which emphasize the comparative color difference values. The automated threshold-decision algorithm uses the difference values of frame-to-frame that was obtained by the weighted chi-square test. At first, The Average of total difference values is calculated and then, another average value is calculated using the previous average value from the difference values, finally the most appropriate mid-average value is searched and considered the threshold value. Experimental results show that the proposed algorithms are effective and outperform the previous approaches.

Keywords: 플래시라이트, 장면전환검출, 차이 값, 지역 히스토그램 비교, 임계값, 정규화

I. 서 론

비디오는 실세계의 다양한 내용들을 함축적으로 담

아낼 수 있는 유용한 도구로서 영상, 문자, 오디오와 같은 복합매체들로 구성된다. 특히 최근에는 컴퓨터와 통신, 그리고 데이터 압축기술 및 표준기술의 발달로 비디오를 이용한 다양한 서비스가 가능하게 되었으며, 초고속 통신망의 보급과 멀티미디어 처리 기술로 인해 비디오는 다양한 분야에서 폭 넓게 사용되고 있다^[1,2,9].

비디오 분할은 비디오 데이터베이스의 구축을 위하여 필요한 첫 번째 단계로서 비디오 데이터를 보다 체계적이고 구조적인 형태로 재구성하기 위하여 필요한 필수적인 기초 작업이다. 따라서 비디오 시퀀스를 구성하는 연속된 프레임들로부터 비디오의 기본구성단위인

* 정희원, 숭실대학교 정보미디어 기술연구소
(Information Media Technology Research Institute,
Soongsil University)

** 정희원, 군산대학교 컴퓨터정보과학과
(Department of Computer Information Science,
Kunsan National University)

※ 본 논문은 2006년 정부(교육인적자원부)의 재원으로
한국 학술 진흥 재단의 지원을 받아 수행된 연구임
(KRF-2006-005-J03801)

접수일자: 2006년9월12일, 수정완료일: 2006년10월30일

장면전환 지점을 검출하고, 검출된 장면으로부터 장면을 대표하는 대표프레임을 추출하는 작업들로 이루어진다. 이러한 장면전환 검출은 비디오의 분할, 색인, 검색, 브라우징, 요약과 같은 작업을 수행하기 위한 기반 기술이다.

기존의 장면전환 검출은 크게 압축된 영역에서의 장면전환 검출과 비 압축된 영역에서의 장면전환 검출로 구분할 수 있다.

압축된 영역에서의 장면전환 검출은 압축된 영역으로부터 부분적으로 복원된 데이터만을 이용하여 장면전환 지점을 찾는 방법이며, DCT 계수를 이용한 방법과 움직임 벡터를 이용한 방법 등이 대표적으로 사용되고 있다^[3]. 그러나 이 방법은 부분적 데이터만을 이용하여 장면전환 지점을 검출하기 때문에 데이터 손실에 따른 검출의 신뢰성을 떨어뜨리는 문제점을 갖는다.

비 압축된 영역에서의 장면전환 검출은 전체 영역으로부터 픽셀 비교에 의한 방법, 히스토그램의 비교에 의한 방법 등이 사용되며^[10], 이중 히스토그램에 의한 비교 방법이 일반적으로 가장 널리 사용되는 방법이다.

장면전환은 연속된 두 인접한 프레임사이의 차이 값에 따른 임계값을 이용하여 검출되어지며, 크게 급진적 장면전환과 점진적 장면으로 구분지어 설명할 수 있다.

급진적 장면전환은 프레임사이의 차이 값 변화가 매우 심하게 나타나는 프레임들의 변화를 말하며, 장면들의 갑작스런 변화가 발생하는 프레임들을 이야기할 수 있다.

점진적 장면전환은 페이드 인/아웃, 디졸브와 같은 카메라의 특수효과에 의해 발생하는 현상으로 장면들의 점차적인 변화가 발생하는 프레임들을 이야기 한다.

급진적 장면전환이나 점진적인 장면전환 모두 두 인접한 프레임들의 차이 값에 의해 검출될 수 있으며, 차이 값들로부터 장면전환을 추출하기 위하여 차이 값에 대한 임계치를 설정하는 것이 가장 중요하다. 그러나 이러한 임계치의 설정은 값의 크기에 따라 장면전환 검출의 수가 많이 달라지는 문제점이 있으며, 갑작스런 객체의 출현이나 프레임의 왜곡으로 인한 차이 값의 변화에 능동적으로 대처하지 못하는 단점을 가지고 있다.

특히 플래시라이트와 같은 경우에는 갑작스런 명암 값의 변화에 의해 높은 차이 값을 형성하여 임계치의 설정에 의한 장면전환 검출이 매우 어려우며, 기존의 임계치 설정에 의한 장면전환 검출 방법에 의해서는 효과적으로 장면을 검출하지 못하는 단점을 가지고 있다.

기존의 논문에서는 이러한 플래시라이트를 제거하기

위한 방법으로 플래시라이트 모델을 이용하여 플래시의 특징을 구분하고, 이웃하는 프레임들의 구간을 설정하여 구간간의 일정한 비율을 고려한 임계치를 설정하여 플래시라이트를 검출하는 방법들을 제안하였다^[4,10].

그러나 위 방법은 프레임간의 구간에 대한 적절한 길이를 설정하는 것이 어렵고, 또한 연속적으로 플래시라이트가 긴 구간에 걸쳐 삽입되었을 때 검출비율이 떨어지는 문제점, 장면전환과 플래시라이트 지점의 구분을 위한 임계치 결정에 관한 문제점 등을 갖고 있다.

본 논문에서는 플래시라이트가 삽입된 프레임간의 차이 값들이 갖는 고유한 특징을 이용하여 보다 빠르고 효율적으로 장면전환 지점과 플래시라이트 지점을 구분하여, 강건하게 장면전환 지점만을 추출할 수 있는 방법을 다음과 같이 제안한다.

먼저 연속된 프레임으로부터 차이 값 추출을 위하여 지역 히스토그램에 의한 비교 방법을 사용한다. 이것은 객체나 카메라의 움직임에 민감하지 않고 프레임간의 공간 정보를 활용할 수 있는 장점이 있다. 추출된 차이 값은 변이 폭이 넓게 형성되는 특징이 있으며, 이것은 장면전환 검출을 위한 임계 값 결정에 많은 어려움을 준다. 본 논문에서는 이러한 변이 폭이 넓게 형성되는 차이 값의 폭을 줄이고 임계값 결정의 효율성을 높이기 위하여 로그 식을 이용한 차이 값의 동적 압축에 의한 정규화 작업을 수행한다. 정규화 된 차이 값은 일정한 크기로 압축되어 임계값 결정에 많은 장점을 준다.

제안된 장면전환 검출 알고리즘은 정규화 된 차이 값들로부터 차이 값들의 시간적 연속성에 대한 크기와 연결정보를 고려하여 플래시라이트를 제거하며 효과적으로 장면전환 지점을 검출할 수 있다.

본 논문은 구성은 다음과 같다. II장 본문에서는 연속된 프레임에서의 차이 값 추출식과 추출된 차이 값의 정규화 과정에 대해서 이야기 한다. 그리고 차이 값들을 이용한 강건한 장면전환 추출 알고리즘에 대해서 설명한다. III장에서는 실험을 통하여 제안된 내용의 효율성을 검증하며, IV장 결론에서는 제안된 방법의 결론과 향후 연구과제에 대해서 언급한다.

II. 장면 전환 검출 알고리즘

1. 지역 히스토그램 비교에 의한 차이 값 추출

본 논문에서는 연속된 프레임 사이의 강건한 차이 값 추출을 위하여 기존의 히스토그램 기반 알고리즘과 비교하여 좋은 성능을 보여주는 χ^2 -테스트를 변형하여

이용하였으며^[1], 컬러 값의 세분화 작업에 따른 검출효과를 높이기 위하여 명암도 등급에 따른 가중치를 적용한 컬러 히스토그램 비교를 적용하였다^[5]. 또한 공간정보에 대한 손실을 줄이고, 다른 두 프레임이 유사한 히스토그램을 갖는 문제점을 해결하기 위하여 지역 히스토그램 비교를 이용하였다^[6].

컬러 히스토그램 비교($d_{r,g,b}(f_i, f_j)$)는 인접한 두 프레임(f_i, f_j)의 각 컬러공간에 대한 히스토그램 비교를 통하여 계산되어지며 식(1)과 같이 정의하여 사용한다.

$$d_{r,g,b}(f_i, f_j) = \sum_{k=0}^{N-1} (|H_i^r(k) - H_j^r(k)| + |H_i^g(k) - H_j^g(k)| + |H_i^b(k) - H_j^b(k)|) \quad (1)$$

$H_i^r(k), H_i^g(k), H_i^b(k)$ 는 i 번째 프레임(f_i)에서의 각 컬러 공간(r, g, b)에 대한 빈(k)의 수(N)를 나타낸다.

식(1)로부터 각 컬러공간에 대한 명암도 등급 변환을 위한 가중치를 적용하여 식(2)와 같이 재정의하여 사용할 수 있다.

$$d_{wr,wg,wb}(f_i, f_j) = \sum_{k=0}^{N-1} (|H_i^r(k) - H_j^r(k)| \times \alpha + |H_i^g(k) - H_j^g(k)| \times \beta + |H_i^b(k) - H_j^b(k)| \times \gamma) \quad (2)$$

α, β, γ 는 NTSC 표준에 따른 명암도 등급변환을 위한 상수를 나타내며 본 논문에서는 $\alpha = 0.299, \beta = 0.587, \gamma = 0.114$ 로 정의하여 사용하였다.

두 프레임사이의 차이를 강조하기 위한 통계 분석방법 중에서 χ^2 -테스트 비교($d_{w\chi^2}(f_i, f_j)$)는 히스토그램의 비교변화를 통하여 장면전환을 검출할 수 있는 효율적인 방법이며, 식(3)과 같이 정의하여 사용한다.

$$d_{w\chi^2}(f_i, f_j) = \begin{cases} \sum_{k=0}^{N-1} \frac{(H_i(k) - H_j(k))^2}{\max(H_i^r(k), H_j^r(k))} & \text{if } (H_{i,j} \neq 0) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

히스토그램 기반 방법은 공간정보를 적용하지 않아 유사한 컬러분포를 갖는 다른 두 영상인 경우 같은 영상으로 간주하여 검출하는 문제점이 발생할 수 있다. 이러한 문제점은 프레임의 영역을 분할하여 지역에 대한 히스토그램 분포를 비교하는 방법으로 해결할 수 있다.

식(4)에서는 영역의 분할에 따른 각 영역의 컬러 히스토그램 비교와 누적을 통한 차이 값 생성

을 나타내고 있다.

$$d(f_i, f_j) = \sum_{bl=1}^m DP(f_i, f_j, bl) \quad (4)$$

$$DP(f_i, f_j, bl) = \sum_{k=1}^{N-1} |H_i(k, bl) - H_j(k, bl)|$$

$H_i(k, bl)$ 는 프레임(f_i)에 대한 블록(bl)의 k 번째 히스토그램의 분포를 나타내며, m 은 전체 블록의 수를 나타낸다.

위의 식(2)에와 같이 각 컬러공간에 대한 가중치를 적용한 세분화된 히스토그램 비교와 식(3)의 통계방식을 이용한 차이 값의 확장 및 식(4)의 지역히스토그램에 의한 프레임의 공간정보에 대한 활용의 장점을 이용하여 본 논문에서는 차이 값 추출에 대한 강건함과 신뢰성을 위하여 다음 식(5)와 같이 위의 식들을 결합한 지역 히스토그램 비교에 의한 차이 값 추출식을 다음과 같이 적용하여 사용한다.

$$d(f_i, f_j) = \sum_{bl=1}^m d_{\chi^2}(f_i, f_j, bl)$$

$$d_{\chi^2}(f_i, f_j, bl) = \sum_{k=1}^{N-1} \left(\frac{(H_i^r(k) - H_j^r(k))^2}{\max(H_i^r(k), H_j^r(k))} \times \alpha + \frac{(H_i^g(k) - H_j^g(k))^2}{\max(H_i^g(k), H_j^g(k))} \times \beta + \frac{(H_i^b(k) - H_j^b(k))^2}{\max(H_i^b(k), H_j^b(k))} \times \gamma \right) \quad (5)$$

위의 식(5)에서 $H_i^r(k), H_i^g(k), H_i^b(k)$ 는 i 번째 프레임(f_i)이 갖는 각 컬러 공간(r, g, b)에 대한 히스토그램의 분포를 나타내며, N 은 빈(k)의 전체 수, m 은 블록(bl)의 전체 수를 나타낸다.

본 논문에서는 제안된 식(5)으로부터 프레임을 같은 영역의 블록으로 분할한 후, 각 블록들에 대한 히스토그램 비교를 통하여 차이 값을 생성하였다. 생성된 차이 값은 급진적 장면의 변화나 점진적 장면의 변화 모두에 적용할 수 있는 강건한 차이 값의 추출을 보여준다.

2. 차이 값의 정규화

제안된 식(식(1))에 의하여 추출된 차이 값은 프레임간의 특징정보에 따라 편차가 크게 나타나며, 프레임간의 연속적인 연결정보를 얻기가 매우 어렵다. 특히 급진적 장면전환이나 점진적인 장면전환을 추출하기 위한 임계치의 결정은 각 차이 값들의 변화에 능동적으로 대

처하여 결정해야 하는 문제점을 갖는다.

따라서 차이 값의 편차를 줄이고 시간적으로 연결된 차이 값들의 변이를 쉽게 파악하고 정보를 얻기 위한 방법이 필요하다.

기존의 프레임에 대한 전체 픽셀수를 이용한 정규화 방법은 차이 값들에 대한 폭을 일정한 영역으로 줄여 사용하였으나 차이 값들의 시간적 연속성과 연관성에 대한 정보는 제공해 주지 못하는 단점을 가지고 있다.

본 논문에서는 차이 값들을 일정한 값의 범위 안에 동적으로 압축하여 차이 값들의 시간적 연속성과 연관성을 쉽게 파악하여 차이 값들로부터 급격한 장면변화를 추출할 수 있는 차이 값의 정규화 방법을 제안하여 사용한다.

제안된 방법은 영상처리에서 영상의 명암 값 향상을 위하여 사용되는 로그함수와 배율상수를 변형하여 차이 값에 적용하였다^[7].

$$d_{\log} = c \times \log(1 + d^2),$$

$$c = \frac{\max(d_{\log})}{\max(\log(1 + d^2))} \quad (6)$$

c는 배율상수이고, d는 차이 값 추출 식(식(5))에 의하여 추출된 차이 값을 나타낸다. 차이 값의 제곱은 동적범위로 차이 값을 나타내기 위하여 수행되어진다.

정규화 된 후의 차이 값들의 분포는 이전의 차이 값들의 분포에 비하여 차이 값을 일정한 범위내로 동적 압축하여 보다 세분화되고 정밀한 장면전환 검출을 유도할 수 있는 분포를 보여준다. 특히 차이 값들 간의 편차를 줄이고 차이 값들의 연결성을 강조함으로써 본 논문에서 제안된 장면전환 검출 알고리즘에 의하여 보다 효율적으로 장면전환 지점을 추출할 수 있는 장점을 가진다.

(그림 1)에서는 임의의 비디오로부터 정규화 되기 전의 차이 값과 정규화 후의 차이 값들에 대한 분포와 정규분포를 비교하여 나타내고 있다.

3. 플래시라이트에 강건한 장면전환 검출 알고리즘

3.1 플래시라이트(flashlight)의 특징

플래시라이트는 뉴스나 연예 비디오에서 출현되는 인물이나 사물의 강조를 위하여 반짝 빛을 내는 섬광현상(0.02/second)을 말하며, 짧은 빛의 출현으로 인하여 일반적인 비디오(25~30frame/second)의 경우 독립된 한 개의 플래시라이트는 하나의 프레임에만 독립적으로 삽입되는 특징을 갖는다. 이러한 특징은 여러 개의 플래시라이트가 동시에 나타날 때에도 적용되어지며, 각

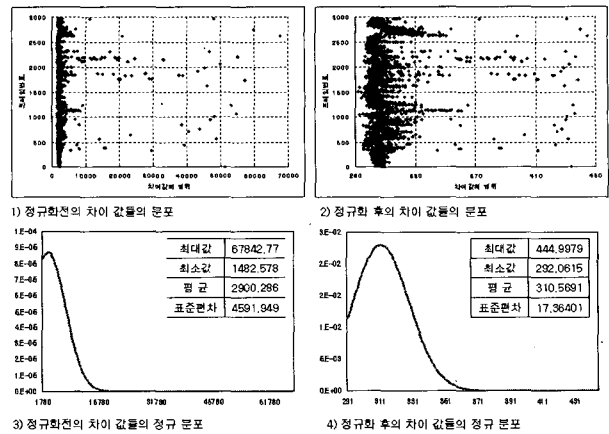


그림 1. 정규화되기 전의 전체 차이 값들에 대한 분포 및 정규분포

Fig. 1. Normal distribution of normalized difference values and original difference values.

표 1. 플래시라이트의 종류와 특징
Table 1. Feature and kind of flashlight.

종 류	특 징
직접 플래시라이트	<ul style="list-style-type: none"> 정면이나 가까운 장소에서 발생한 강한 섬광 높은 명암 값 형성
간접 플래시라이트	<ul style="list-style-type: none"> 먼 거리나 이웃하는 장소에서 발생한 약한 섬광 중간, 낮은 명암 값 형성

각의 플래시라이트가 서로간의 위치와 거리, 시간에 의해 여러 연속된 프레임에 걸쳐 독립적으로 서로 다른 명암 값을 가지고 삽입되어 프레임간의 차이 값에 영향을 준다.

본 논문에서는 플래시라이트의 종류를 구분하고 특징을 설명하기 위하여 표 1과 같이 정의하여 사용하였다.

직접 플래시라이트는 하나의 독립된 섬광으로부터 하나의 프레임에 직접 삽입되어 강한 명암 값을 형성하는 특징을 가지며, 간접 플래시라이트는 직접 섬광이 나타난 후에 곧바로 삽입되어지는 간접적인 섬광을 말한다. 다양한 플래시라이트가 삽입된 비디오는 이러한 직, 간접 플래시라이트에 의해 연속된 여러 개의 프레임들이 하나의 플래시라이트 구간을 이루고 있는 것이 일반적이다.

그림 2는 임의의 뉴스비디오로부터 삽입된 직, 간접 플래시라이트의 영향으로 생성된 프레임내의 구간들에 대한 차이 값들의 분포와 특징을 나타내기 위한 그림이다.

표시된 각 구간(A,B,C,D)은 일정한 임계치를 기준으로 플래시라이트가 1개 이상 연속적으로 삽입된 프레임

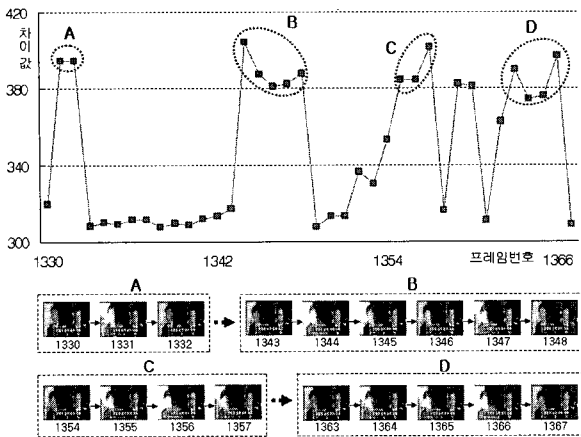


그림 2. 플래시라이트가 연속적으로 삽입된 프레임 구역에 대한 차이 값의 분포

Fig. 2. Distribution of difference values which has many consecutive flashlights

들의 차이 값들에 대한 영역을 표시한 것이다.

구간 A는 직접 플래시라이트 하나가 삽입되어 생성된 두 개의 높은 차이 값의 분포를 보여주고 있다. 이것은 같은 유사정보를 갖는 연속된 프레임간의 구역에서 갑작스런 명암 값의 변화가 인접한 두 프레임내의 차이 값을 높게 형성하기 때문에 발생한다.

구간 B, C, D에서는 직, 간접 플래시라이트가 2개 이상 연속된 프레임에 걸쳐 나타났을 때의 구간 영역들에 대한 차이 값의 분포를 나타낸다.

B와 같은 구간에서는 직접 플래시라이트가 먼저 프레임에 삽입된 후, 간접 플래시라이트가 연속적으로 서로 다르게 두 프레임에 걸쳐 삽입된 후, 다시 직접 플래시라이트가 삽입되어, 결과적으로 이웃하는 프레임간의 5개 차이 값이 서로 높게 형성되는 특징을 보여준다. C와 같은 경우에는 직접 플래시라이트가 먼저 프레임에 삽입된 이후, 다음 프레임에는 간접 플래시라이트가 삽입되어 플래시라이트가 두 프레임에 걸쳐 연속적으로 나타나지만 프레임간의 차이 값은 서로 높게 형성되는 특징을 보여준다. D의 구간에서도 직접 플래시라이트와 간접 플래시라이트가 연속적으로 프레임 내에 삽입되며 만들어 낸 차이 값의 분포를 보여주고 있다.

실제 플래시라이트가 연속적으로 삽입되는 경우에는 하나의 독립된 직접 플래시라이트에 의해 연속적으로 삽입되는 경우는 어렵고, 이웃하는 여러 개의 간접 플래시라이트에 의해 서로 다르게 삽입되는 경우가 일반적이다. 이러한 경우에는 플래시라이트가 서로 간섭되어 많은 편차를 갖는 차이 값을 형성하는 특징을 보여준다.

따라서 이러한 특징을 이용하여 플래시라이트에 의해 장면이 잘못 검출되어지는 문제점을 해결하고 장면 전환 지점만을 검출할 수 있는 새로운 장면전환 검출 기법을 제안할 수 있다.

3.2 제안된 장면전환 검출 알고리즘

본 논문에서는 하나의 카메라에 의해서 연속적으로 촬영된 프레임들의 구역을 장면으로 정의하며, 장면전환 검출은 이러한 구역들을 구별하는 나누는 것을 목적으로 한다.

제안된 방법은 연속되는 프레임들로부터 차이 값을 구하고, 정규화 작업을 실시한 후에, 입력되는 차이 값들의 특징으로부터 플래시라이트를 제거한 후, 장면전환의 존재여부를 판단하게 된다.

플래시라이트에 의한 장면전환의 잘못된 검출의 문제점을 해결하기 위하여 본 논문에서는 인접한 프레임과의 차이 값들에 대한 연속성 여부를 측정하여 사용한다.

차이 값의 연속성이란 차이 값들이 일정한 거리를 두고 서로 유사하게 연속적으로 분포하는 특성을 말하며 이것은 프레임들이 서로 유사한 정보들로 구성되는 것을 말한다. 일반적으로 장면전환으로 검출되는 지점은 프레임의 변화가 급변하게 일어나는 지점이므로 이러한 차이 값의 연속성이 끊어지고, 하나의 차이 값만이 특정 위치에서 높게 형성하는 특징을 갖으며, 이후의 차이 값은 다시 연속성을 갖는다. 그러나 플래시라이트는 이러한 차이 값의 연속성을 높은 차이 값의 범위에서 다시 이어가는 독특한 성질을 갖는다. 따라서 플래시라이트와 장면전환은 이러한 특징을 이용하여 쉽게 구별하여 추출할 수 있다.

다음 제안되는 장면전환 검출 알고리즘은 이러한 플래시라이트의 특성을 이용하여 보다 강건하게 장면전환 추출을 할 수 있도록 제안된 방법이다.

제안된 장면전환 검출 알고리즘

- Step 1. 주어진 비디오로부터 차이 값 추출식(식(5))을 이용하여 프레임((f_i, f_j))의 j 번째 차이 값($d(j)$)을 계산한다.
- Step 2. 추출된 차이 값($d(j)$)으로부터 정규화식(식(6))을 적용하여 정규화 된 차이 값($d_{log}(j)$) 생성한다.
- Step 3. 정규화 된 차이 값($d_{log}(j)$)으로부터 다음 조건을 비교하여 장면전환 검출 지점을 계산한다.

$$\text{조건 (1)} \quad d_{\log}(j) \geq th_{\max}$$

정규화 된 차이 값은 전역적 임계치(th_{\max})를 만족해야 한다. 전역적 임계치는 장면전환지점으로 추출된 차이 값들의 평균값으로부터 계산되었다.

$$\text{조건 (2)} \quad \begin{aligned} bd_{\log}(j) &\geq th_{\min} \\ (bd_{\log}(j) &= |d_{\log}(j) - d_{\log}(i)|) \end{aligned}$$

이전과 현재의 차이 값($bd_{\log}(j)$)은 지역적 임계치(th_{\min})를 만족해야 한다.

$$\text{조건 (3)} \quad \begin{aligned} fd_{\log}(j) &\geq th_{\min} \\ (bd_{\log}(j) &= |d_{\log}(j) - d_{\log}(j+1)|) \end{aligned}$$

현재와 이후의 차이 값($fd_{\log}(j)$)은 지역적 임계치(th_{\min})를 만족해야 한다.

$$\text{조건 (4)} \quad bfd_{\log} = \sqrt{(bd_{\log}(i))^2 + (fd_{\log}(i))^2} \geq th_{\text{global}}$$

차이 값 간의 유동성 있는 거리 값의 조정을 위하여 두 차이 값들 간의 거리는 일정한 임계치(th_{global}) 이상을 만족해야 한다.

Step 4. Step 3의 조건(1) ~ 조건(4)를 모두 만족하는 프레임은 장면전환 지점으로 검출되며, 조건중 하나라도 만족하지 못할 경우에는 Step 1로 넘어가 연속된 다음 프레임으로부터 새로운 차이 값을 생성한다.
}

III. 실험

본 논문에서는 제안 알고리즘의 실험평가를 위하여 개발의 편리성과 확장성이 용이한 Win API와 C++를 병행하여 사용하였으며, 윈도우 멀티미디어를 다루기 위한 기본요소인 DirectX SDK8.1을 실험에 이용하였다. 실험에 사용된 비디오의 형식은 표 2와 같다.

실험에 사용된 비디오는 하이라이트가 상대적으로 많이 들어있는 뉴스와 연예 비디오를 선별하여 사용하였다.

표 2. 실험에 사용된 비디오 형식
Table 2. Type of experimental Video.

비디오	설명
실험 비디오	뉴스(5개), 연예(5개)
비디오 저장 도구	Sigma TV Capture Board
비디오 저장 형식	MPG(MPEG-V1), AVI(MS-CRAM)
초당 프레임의 수	30f / s
비디오 사이즈	320 × 240

표 3. 실험에 사용된 비디오의 종류와 특징

Table 3. Feature and kind of experimental video type.

비디오 종류	프레임 수	급진적 장면전환	플래시라이트	
뉴스	mbc_1	2772	26	31
	mbc_2	1132	14	4
	mbc_3	2975	21	14
	kbs_1	2665	17	55
	kbs_2	2167	22	6
연예	movie	1175	11	25
	rain_1	2578	15	62
	lee_1	960	9	58
	photo_1	1028	10	15
	photo_2	595	5	6

(표 3)은 실험에 사용된 비디오의 종류와 특징에 대하여 나타내고 있다.

급진적 장면전환은 카메라 브레이크가 발생한 지점을 임의로 설정하여 검출한 지점의 수를 말하며, 플래시라이트의 수는 일정한 임계를 넘어가는 차이 값 중 플래시라이트가 들어간 프레임의 개수를 이용하여 결정하였다.

실험결과에 대한 평가는 일반적으로 가장 많이 사용되는 Rcall과 Precision을 이용한 방법을 사용하였으며 정의는 다음과 같다.

$$Recall = \frac{N_c}{N_c + N_m} \times 100\%$$

$$Precision = \frac{N_c}{N_c + N_f} \times 100\%$$

N_c 는 검출된 장면프레임의 수, N_m 은 검출되지 못한 장면프레임의 수를 나타내며, N_f 는 잘못 검출된 장면프레임의 수를 나타낸다.

표 4에서는 실험에 사용된 비디오의 정확성과 리콜에 대한 평가를 보여주고 있다.

표 4. 검출된 장면의 정확도 및 리콜

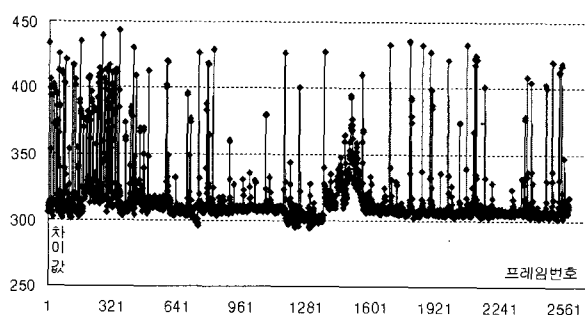
Table 4. Precision and recall of detected scene.

비디오 종류	급진적 장면 수	검출된 장면 수	precision (%)	recall (%)	
뉴스	mbc_1	26	26	92	96
	mbc_2	14	14	100	100
	mbc_3	21	20	100	95
	kbs_1	17	16	94	100
	kbs_2	22	22	100	100
연예	movie	11	9	100	82
	rain_1	15	15	100	100
	lee_1	9	9	100	100
	photo_1	10	8	80	100
	photo_2	5	4	100	80

실험에 사용된 비디오의 전체 정확성은 96.6%로 나왔으며, 리콜은 95.3%로 나타났다. 정확성이 떨어지는 비디오의 경우는 많은 수의 플래시라이트가 동시에 프레임에 연속적으로 삽입되어 특정 위치에서의 플래시라이트에 의한 차이 값의 특징이 갑작스런 장면전환 지점에서의 차이 값의 특징과 유사한 한 분포를 형성하여 장면전환이 잘 못 검출된 문제이며, 리콜에서 문제점이 발생하는 비디오는 장면전환이 발생한 지점에서의 차이 값들에 대한 전역적 임계치와 지역적 임계치의 설정에 대한 민감성의 문제에 의해 나타났다. 이러한 민감성은 유사 프레임 구간의 정보왜곡에 의한 차이 값의 변형이 가장 큰 원인이며, 향후 계속 연구가 진행되어야 할 중요한 문제이다.

그림 4와 그림 5는 실험에 사용된 비디오 중에서 연예 비디오(rain_1)와 뉴스 비디오(kbs_1)에 대한 전체 차이 값들의 분포 및 장면전환 검출 알고리즘에 의하여 추출된 각 장면들에 대한 이미지와 프레임번호, 차이 값의 분포, 차이 값 간의 이웃거리 등에 대한 정보를 보여주고 있다.

그림 4의 연예비디오로부터 전체 차이 값들의 분포



1) 정규화 된 전체 차이 값들의 분포

프레임번호	1shot	2shot	3shot	4shot	5shot	6shot	7shot	8shot
$d_{\text{rain}}(0)$	433.7	425.5	406.9	439.6	443.7	430.7	409.3	412.1
$bd_{\text{rain}}(0)$	128.5	52.6	74.7	113.7	118.7	100.4	90.3	64.4
$fd_{\text{rain}}(0)$	433.7	121.9	20.8	123.7	58.9	122.5	100.9	106.9
$hfd_{\text{rain}}(0)$	452.4	122.8	80.8	168.0	132.6	156.4	135.5	124.6

프레임번호	9shot	10shot	11shot	12shot	13shot	14shot	15shot
$d_{\text{rain}}(0)$	419.2	427.2	429.9	426.9	400.1	428.2	410.2
$bd_{\text{rain}}(0)$	85.5	95.6	105.6	97.4	79.1	86.6	48.2
$fd_{\text{rain}}(0)$	97.4	114.9	124.4	119.6	104.5	124.1	92.5
$hfd_{\text{rain}}(0)$	130.3	149.6	163.2	154.4	130.4	152.5	102.1

2) 장면전환 검출 알고리즘에 의해 추출된 각 장면의 정보

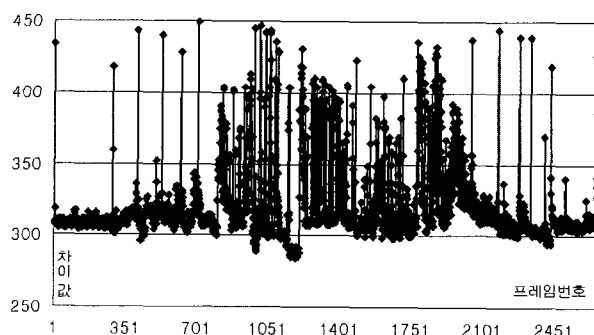
그림 4. 연예비디오(rain_1)로부터 추출된 전체 차이 값들의 분포 및 장면 구성 정보

Fig. 4. Structural Information of extracted total difference values from Rain_1 video.

에서 많은 수의 플래시라이트가 시작부분에 높은 차이 값을 형성하며 삽입되어 밀접하게 분포되어 있는 것을 확인할 수 있으며, 비디오 중간, 중간에도 직, 간접 플래시라이트가 간헐적으로 삽입되어 높은 차이 값을 형성하고 있는 그래프를 확인할 수 있다.

그림 5의 뉴스비디오에서는 전체 차이 값의 분포로부터 중간 지점에서 많은 수의 플래시라이트가 높은 차이 값을 형성하며 중점적으로 삽입되어 있는 것을 확인할 수 있다.

플래시라이트가 많이 삽입된 비디오는 기존의 임계치를 통한 장면전환 검출방법으로는 장면전환 지점의 추출이 매우 어렵고 효율성과 정확성도 많이 떨어지는 문제점을 갖는다. 그러나 제안된 장면전환 검출 알고리즘은 이러한 플래시라이트를 효과적으로 제거할 뿐만 아니라 장면전환 지점이 발생한 프레임도 정확하게 검출할 수 있다. 그림 4와 그림 5에 있는 표와 이미지는 주어진 비디오의 검출 가능한 장면들이 거의 모두 검출되었으며, 플래시라이트에 의해 잘못 검출된 장면이 없음을 보여준다. 특히 비디오의 시작부분이나 중간부분



1) 정규화 된 전체 차이 값들의 분포

프레임번호	1shot	2shot	3shot	4shot	5shot	6shot	7shot	8shot
$d_{\text{kbs}}(0)$	433	418	442.9	439.6	428.6	449.1	428.2	404.9
$bd_{\text{kbs}}(0)$	126.1	58.7	118.7	110.4	103.4	120.8	85.4	39.8
$fd_{\text{kbs}}(0)$	433.7	111.9	129.9	124.7	106.5	124.9	112.9	93.5
$hfd_{\text{kbs}}(0)$	451.6	126.4	176	166.6	148.4	173.8	141.6	101.6

프레임번호	9shot	10shot	11shot	12shot	13shot	14shot	15shot	16shot
$d_{\text{kbs}}(0)$	410.3	436.1	409.9	437.1	444.1	439.5	439.2	418.7
$bd_{\text{kbs}}(0)$	59.1	70.5	77.0	81.2	130.2	110.7	134.7	77.9
$fd_{\text{kbs}}(0)$	88.5	129.1	50.0	88.8	138.3	140.0	133.4	121.1
$hfd_{\text{kbs}}(0)$	104.6	150.1	91.8	120.3	167.8	178.5	199.6	143.9

2) 장면전환 검출 알고리즘에 의해 추출된 각 장면의 정보

그림 5. 뉴스비디오(kbs_1)로부터 추출된 장면들에 대한 구성정보

Fig. 5. Structural Information of extracted total difference values from Kbs_1 video.

에 삽입된 연속된 많은 수의 플래시라이트가 효과적으로 제거되었으며, 플래시사이트 사이에 들어 있는 장면 전환 지점도 정확하게 검출되었음을 검출된 장면들의 구성정보로부터 확인할 수 있다.

이것은 제안된 장면전환 검출 알고리즘이 플래시라이트에 대한 특징을 효과적으로 분석하여 제거하였음은 물론, 장면전환 지점도 제안된 방법에 의하여 효율적으로 검출하였음을 보여준다.

IV. 결 론

본 논문에서는 주어진 비디오로부터 연속된 프레임 사이의 차이 값 추출을 위하여 객체나 카메라의 움직임에 덜 민감하고 프레임의 공간정보를 이용하는 지역 히스토그램 비교에 의한 추출식을 사용한다. 또한 추출된 차이 값의 큰 변이 폭에 의한 임계 값 결정의 문제점을 해결하기 위하여 차이 값들의 동적 압축에 의한 정규화 작업을 수행한다. 제안된 장면전환 검출 알고리즘은 정규화 된 차이 값들의 시간적 연속성의 변이에 따른 플래시라이트의 특징을 이용하여 플래시라이트와 장면전환 검출을 따로 구분하여 추출할 수 있는 새로운 장면전환 검출 알고리즘이다.

제안된 방법은 기존의 장면전환 검출에서 문제가 되었던 여러 부분들을 효과적으로 해결하고 있다. 먼저 기존의 카이-테스트와 비교하여 차이 값 추출에 좋은 성능을 보여주고 있으며, 다양한 비디오 장르에 따른 임계치 결정의 문제점을 추출된 차이 값으로부터 일정한 범위내의 동적압축을 통한 정규화 작업을 수행하여 차이 값의 범위를 줄이고, 전역 임계치를 설정함으로써 해결하고 있으며, 플래시라이트에 의하여 발생할 수 있는 잘못된 장면전환 검출의 문제점도 제안된 장면전환 검출 알고리즘을 통하여 효율적으로 해결하고 있다.

따라서 제안된 방법은 실시간으로 입력되는 비디오를 분석하여 장면전환 검출을 보다 강건하게 지원함으로써 비디오 자동 분류를 위한 기반작업의 효율성을 높여줄 수 있음은 물론 비디오의 일반적 특성을 통한 색인을 지원함으로써 다양한 검색을 요구하는 사용자의 의도에 보다 적절한 대응을 시도할 수 있는 시스템 개발에 적절하게 응용될 수 있다.

그러나 본 논문에서 제안된 방법은 급진적 장면전환만을 모델로 하고 있어 점진적인 장면에서의 다양한 변화를 충분히 수용하지 못하고 있으며, 전역적 임계치와 지역적 임계치에 대한 보

다 세부적인 정의에 대한 연구가 요구되어진다. 특히 플래시라이트가 짧은 시간 안에 복잡하게 삽입된 비디오와 같은 경우에는 플래시라이트와 장면전환 지점간의 특징차이가 유사해지는 경우가 발생하여 장면전환 지점의 검출에 실패하는 경우가 많다. 따라서 보다 다양한 환경을 가지는 비디오 데이터에서의 실험이 요구된다.

참 고 문 헌

- [1] C. L. Huang and B. Y. Liao, "A Robust Scene Change Detection Method for Video Segmentation," *IEEE Trans on CSVT*, Vol. 11, No. 12, pp. 1281-1288, December 2001.
- [2] U. Gargi, R. Kasturi, and S. H. Strayer, "Performance Characterization of Video Shot Change Detection Methods," *IEEE Trans on CSVT*, Vol. 10, No. 1, pp. 0001-0013, February 2000.
- [3] H. Zhang, A. Kankamhalli, and S. Smoliar, "Automatic partitioning of full-motion video," *ACM Multimedia Systems*, New York: ACM Press, Vol. 1 1993, pp. 10-28.
- [4] D. Zhang, W. Qi, H. J. Zhang, "A New Shot Boundary Detection Algorithm," *IEEE Pacific Rim Conference on Multimedia*, pp.63-70, 2001.
- [5] K. C. Ko and Y. W. Lee, "Scene Change Detection using the Chi-test and Automated Threshold Decision Algorithm," *ICCSA06*, Vol. 3983 2006, pp. 1060-1069.
- [6] U. Gragi, R. Kasturi, S. Antani, "Evaluation of video sequence indexing and hierarchical video indexing," in: *Proc. SPIE Conf. Storage and Retrieval in Image and Video Databases*, 1995, pp. 1522-1530.
- [7] Gonzalez, "Digital Image Processing 2/E (S/C)," Prentice-Hall, 2002.
- [8] Ralph M. Ford, Craig Robson, Daniel Temple, Michael Gerlach, "Metrics for shot boundary detection in digital video sequences," *Multimedia Systems* 8: 37-46(2000).
- [9] A. Ekin, A. M. Tekalp, and R. Mehrotra, "Automatic soccer video analysis and summarization," *IEEE Trans. on Image Processing*, Vol. 12, No. 7, pp. 796-807, July 2003.
- [10] I. Koprinska and S. Carrato, "Temporal Video Segmentation: A Survey," *Signal Processing Image Communication*, Elsevier Science 2001.
- [11] C. L. Huang and B. Y. Liao, "A Robust

Scene-Change Detection Method for Video Segmentation," IEEE Trans. Circuits System. Video Technology, Vol.11, No.12, December 2001.

저 자 소 개



고 경 철(정회원)
2001년 군산대학교 컴퓨터과학과 졸업(이학석사)
2006년 군산대학교 컴퓨터정보과학과 졸업(이학박사)
2006년~현재 송실대학교 정보미디어연구소 연구원

<주관심분야 : 지능형 에이전트, 멀티미디어, 컴퓨터비전>



최 형 일(정회원)
1979년 연세대학교 전자공학과 졸업(공학사)
1982년 미시간대학교 전산공학과 졸업(공학석사)
1987년 미시간대학교 전산공학과 졸업(공학박사)

1987년~현재 송실대학교 미디어학부 교수
<주관심분야 : 컴퓨터비전, 패턴인식, 퍼지이론, 비디오검색, 인터페이스 에이전트 등>



이 양 원(정회원)
1994년 송실대학교 전자계산학과 공학박사
1986년~현재 군산대학교 컴퓨터정보과학과 교수

<주관심분야 : 멀티미디어, 컴퓨터비전, 인공지능, 가상현실>