

논문-98-3-2-05

동영상내에서의 장면 전환 검출 간격의 최적화

하명환*, 나운정*, 이상길*

Optimization of Temporal Skip Factor for Scene Change Detection in Video Sequence

Myung-Hwan Ha*, Yoon-Jeong Na*, and Sang-Gil Lee*

요 약

시간적 표본화를 채택한 다중경로방법에 의한 장면전환 검출방법은 전 동영상에 대해 매 프레임마다 순차적으로 검색하는 방법에 비해서 빠르며, 동일한 정확성을 갖고 있다. 그러나, 검출 시간을 최소화하기 위한 적절한 검출 간격을 선택하는 어떠한 기준이나 방법도 제시되지 않았으며 검출 간격을 경험에 의해 선택할 수밖에 없었다. 이 논문에서는 장면 전환 검출 시간, 검출 간격과 실제 장면 전환 간격의 통계적 특성 사이의 관계를 수식으로 유도하고 실험에 의해 확인하였다. 또한, 주어진 평균 장면 전환 간격으로부터 검출 시간을 최소화하는 최적 검출 간격을 얻을 수 있음을 보였다. 평균 장면 전환 간격이 알려져 있지 않은 동영상에 대해서 최적 검출 간격을 추정할 수 있는 알고리즘을 제안하였다.

Abstract

The previous method that performs scene change detection by means of a multi-pass approach using temporal sampling is fast and has the same order of accuracy as the one-pass approach that passes through the entire video. However, any criterion or method to select appropriate skip factor for minimum processing time has not been suggested. So one should select skip factor only empirically. In this paper, the relation among the processing time, skip factor, and the statistics of scene change distance is derived theoretically and the result is verified through experiment. In addition we show that the optimal skip factor for minimum processing time can be chosen from the given mean scene change distance. For the video of which mean scene change distance is not known, we propose an algorithm to estimate the optimal skip factor.

I. 서 론

영상자료를 디지털화한 다음 네트워크에 분산 저장하고, 브라우저를 사용하여 필요한 자료를 검색할 수 있는 멀티미디어 데이터 베이스의 구축은 현재 세계적으로 큰 관심거리이다. 방송국에서 이러한 동영상검색 시스템이 구축되면 개인 PC를 이용하여 방송 자료에 쉽게 접근하는

것이 가능하며, 프로그램 제작 면에서도 획기적인 발전이 있을 것이라 기대한다.

그림 1은 이러한 동영상검색 시스템의 한 예를 보인 것이다. 사용자는 동영상의 내용을 요약해 놓은 조각 그림 리스트를 통해서 정보를 한 눈에 파악할 수 있으며, 필요하면 비디오를 원하는 지점부터 직접 재생할 수 있으므로 편리하게 원하는 장면을 찾을 수 있다. 이러한 동영상 검색 시스템을 구현하기 위해서는 동영상 파일을 프레임 단위로 검색하면서 자동으로 인덱싱 정보를 만드는 장면 전환 검출 알고리즘이 필요하게 되며, 이에 관한 많은 연구 결과가 소개된 바 있다^[1].

* KBS 기술연구소
Korean Broadcasting System Technical Research Institute

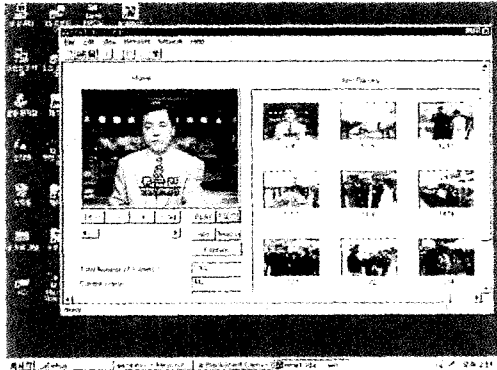


그림 1. 동영상 검색 시스템의 예
Fig. 1. An example of video retrieval system

장면 전환 검출 알고리즘은 매 프레임마다 프레임간의 밝기차나 히스토그램차이를 계산하므로 계산량이 매우 많으며, 검출 속도 또한 아주 느리게 된다. Zhang^[2]은 검출 속도의 향상을 위해 시간적 표본화 방법을 제안하였으나, 검출 간격(skip factor)을 실험적으로만 결정할 수 있었으며, 검출 간격과 검출 속도의 관계에 대하여 명확하게 설명하지 않고 있다.

본 논문에서는 동영상의 통계적 특성, 검출 간격, 검출 속도와의 관계를 수식적으로 유도하고 이것을 실험으로 확인하였으며, 최적의 검출 간격은 동영상의 평균 장면 전환 간격과 관계함을 보였다. 또한, 평균 장면 전환 간격이 알려져 있지 않은 동영상에 대해서 최적 검출 간격을 추정할 수 있는 알고리즘을 제안하였다. 제II장에서는 동영상의 통계적 특성을 고찰하고, 최적의 검출 간격을 유도한다. 제III 장에서는 실제 동영상으로부터 유도된 검출 속도 식의 타당성을 입증하고, 제안된 최적 장면 전환 검출 알고리즘의 실험 결과를 소개하며, 제IV 장에서 결론을 맺고자 한다.

II. 장면 전환 검출 알고리즘의 최적화

1. 동영상 검색 알고리즘

1.1 장면 전환 검출 알고리즘

비디오를 촬영하는 기본적인 도구는 카메라로서 여러 개의 카메라 장면들을 편집하여 연속적으로 보여줌으로써 하나의 비디오를 구성하게 된다. 디지털 비디오에서 이러한 하나의 카메라에 의한 연속적인 장면을 샷(shot)이라 표현하고, 샷 사이의 장면의 전환이 이루어지는 부분을 컷(cut)이라고 한다. 하나의 샷은 연속적인 비디오 프레임들로 구성되어 있으므로 이러한 샷을 비디오의 최소 단위로 규정할 수 있고, 샷 단위의 검색은 프레임 단위의 검색에

비해 정보의 양을 훨씬 줄일 수 있으므로 효율이 높다고 할 수 있다^[3]. 샷 단위의 검색을 위해서는 샷 단위로 경계를 구분 지어 인덱스의 단위를 만들어 주는 장면 전환 검출(scene change detection) 과정이 필요하게 되며, 순간적으로 장면이 바뀌는 컷 검출(cut detection)과 특수효과(디졸브, 페이드 인, 페이드 아웃, 와이프)에 의한 장면 전환의 경우에 사용되는 점진적 장면 전환 검출(gradual scene transition detection)이 있다.

컷 검출 과정에서는 일반적으로 식 (1), (2)와 같이 인접한 프레임간의 밝기차(luminance difference)나 히스토그램의 차이를 계산하여 임계치 이상이면 컷으로 판단하므로 비교적 간단하나, 점진적 장면 전환의 경우는 더 복잡한 알고리즘이 사용된다^[2].

$$DP_i = \sum_{k=0}^M \sum_{l=0}^N | Y_i(k, l) - Y_{i+1}(k, l) | \quad (1)$$

DP_i : i 번째 프레임의 프레임간 밝기차

$Y_i(k, l)$: (k, l) 화소의 휘도값

M : 수직방향의 총 화소수

N : 수평방향의 총 화소수

$$SD_i = \sum_{j=0}^G | H_i(j) - H_{i+1}(j) | \quad (2)$$

SD_i : i 번째 프레임의 히스토그램 차이

G : 휘도 또는 컬러 레벨의 총 개수

$H_i(j)$: j 휘도 성분의 히스토그램 값

1.2 검출 속도 향상을 위한 표본화

매 프레임마다 화소단위로 밝기차나 히스토그램차이를 계산하는 경우, 많은 계산량때문에 검색 속도가 현저히 떨어진다. 더구나, 비디오 데이터가 MPEG과 같이 압축되어 있는 경우는 디코딩과정이 필요하므로 검색 시간이 더욱 더 길어지게 된다. 이 경우 속도 향상을 위해서 시간적/공간적으로 표본화하는 방법이 제안되었다^{[4][5]}. 즉, MPEG으로 압축된 데이터의 경우, I 픽처만을 뽑은 다음, 매 DCT 블록 중에서 첫 번째 계수(DC 계수)만을 사용하여 프레임간 유사도를 측정하여 장면 전환 검출에 활용하는 방법이다.

보통의 MPEG 스트림 경우에 I 픽처는 수 프레임마다 한번꼴로 있으므로 장면 전환 검출에 사용하기 적당하다. 또한, DC 계수는 매 블록의 평균을 대표하는 값이므로 공간적으로도 1/8의 표본화된 영상을 사용한 셈이다. (그림 2)

위 방법들의 단점은 검출속도는 향상되었으나, 시간적으로 표본화된 프레임을 사용함으로써 정확한 장면 전환 지점을 찾을 수 없거나, 공간적인 정보를 희생함으로써 검출 성능이 떨어지는 단점이 있었다.

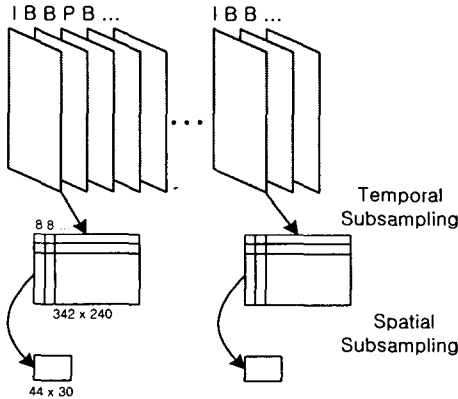


그림 2. MPEG 스트림에서 시공간적인 표본화
Fig. 2. Spatio-temporally scaling in MPEG stream

Zhang^[2]은 검출 속도는 향상시키면서도, 매 프레임을 검색하는 것과 같은 성능을 가지는 다중경로방법(multi-pass approach)을 제안하였다.(그림 3) 첫 번째 단계에서 일정한 간격으로 장면 전환 검출을 수행하고, 두 번째 단계에서는 처음 단계의 결과로 나온 후보 구간에 대해서만 프레임 단위로 정밀 검색을 수행하여 정확한 장면 전환 지점을 찾는 방법이다. 이 경우에 검출 속도는 얼마나 많은 프레임간 비교를 수행하는가가 중요한 요소이며, 꼭 필요한 프레임간의 비교만을 하는 적절한 검출 간격을 선택함으로써 검출 시간을 최소화할 수 있을 것이다.

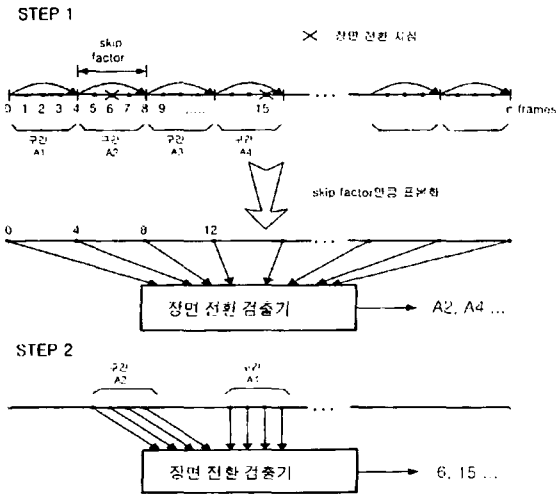


그림 3. 다중 경로 방법
Fig. 3. Multi-pass approach

그림 4는 실제 검출 간격에 따른 검출 시간을 실험하여 그래프로 나타낸 예로, 5분 분량의 뉴스의 경우, 검출 간

격이 10정도에서 가장 시간이 작게 걸리는 것을 알 수 있다. 검출 간격이 너무 크면 검색 구간 내에 장면 전환 지점이 있을 확률이 아주 높으며, 결국, 정밀 검색을 수행하는 구간이 많아지므로 검출 시간이 길어지고, 너무 작으면 동영상의 시간적 중복성을 고려하지 않은 불필요한 프레임간 비교가 이루어지므로 역시 검출 시간이 늘어나게 된다. 검출시간을 최소화하기 위해서는 동영상내의 장면전환 지점의 분포를 고려한 검출 간격으로 선택해야 하며, 다음 절에서 이를 설명한다.

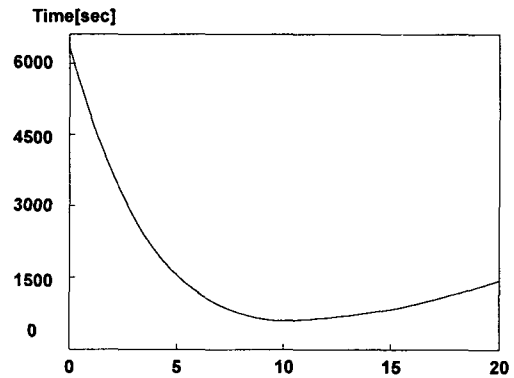


그림 4. 검출 간격과 검출 시간의 관계
Fig. 4. Skip factor vs detection time

2. 동영상 장면 전환의 통계적 특성

동영상의 종류로는 몇 분 가량의 광고, 하루 분량의 뉴스 등 다양하지만, 동영상 검색 시스템에서는 인덱싱 작업이 필요할 만큼 충분히 길이가 긴 동영상을 취급할 것이므로 다음을 가정한다.

가정 : n 프레임 길이의 동영상 내에서 장면 전환 지점의 개수를 k 라 하면, 충분히 큰 n 에 대하여 $\frac{k}{n}$ 는 일정하다.

표 1에서 나타낸 것처럼 동영상을 제작할 때 특별한 의도가 없는 이상, 이 가정은 비교적 잘 성립한다. 이때, 평균 장면 전환 발생률 θ ($0 \leq \theta \leq 1$), 평균 장면 전환 간격 d ($1 \leq d \leq n$)를 다음처럼 정의한다.

$$\theta = \frac{k}{n} = \frac{1}{d} \tag{3}$$

평균 장면 전환 발생률은 동영상의 전체적인 특성을 나타내는 것으로, 일반적으로, 짧은 시간에 많은 정보를 전달하고자 제작된 광고의 경우가 드라마나 뉴스의 경우보다 높게 나타난다. 표 1의 a), b)에 각각 5분 분량의 뉴스

와 광고의 경우에 평균 장면 전환 발생율을 구한 것이다. n 이 커짐에 비교적 일정한 값을 나타냄을 알 수 있다.

표 1. 동영상 길이와 평균 장면 전환 발생율의 관계
Table 1. Video length vs mean scene change frequency

n	k/n	θ
1800	1/1800	1/1800
2700	4/2700	1/675
3600	11/3600	1/327
4500	18/4500	1/250
5400	52/5400	1/104
6300	62/6300	1/102
7200	66/7200	1/109
8100	71/8100	1/114
9000	82/9000	1/109

a) 뉴스
a) News

n	k/n	θ
1800	36/1800	1/50.0
2700	52/2700	1/51.9
3600	86/3600	1/41.9
4500	123/4500	1/36.5
5400	154/5400	1/35.0
6300	159/6300	1/39.6
7200	175/7200	1/41.4
8100	209/8100	1/38.7
9000	246/9000	1/36.5

b) 광고
b) Commercial

3. 검출 간격의 최적화 유도

장면 전환 검출 간격을 L , 서로 다른 두 프레임을 비교하는데 걸리는 계산 시간을 t_c , 전체 n 프레임의 동영상

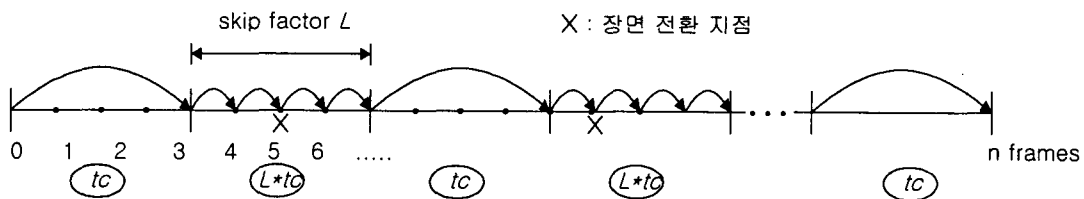


그림 5. 검색 간격 L로 장면 전환 검출하는 형태
Fig. 5. A figure of scene change detection with skip factor L

내에 장면 전환 지점의 개수를 k 라 하자. t_c 는 프레임간의 히스토그램 차이를 계산하거나, 화소의 밝기차를 계산하는 등, 실제 프레임들 비교하는데 걸리는 시간이다. 전체 동영상을 검색하는데 걸리는 시간을 그림 5에서 설명하였다. (X 표는 장면 전환 지점을 나타낸다.)

만약 L 구간 내에 장면 전환 지점이 없으면 L 구간의 처음과 끝, 두 프레임만을 비교하므로, t_c 의 시간이 걸리고, 구간 내에 장면 전환 지점이 있으면, 구간 전체에 대하여 검색을 수행해야 하므로 $L \times t_c$ 의 시간이 걸린다. n 프레임 내에 k 개의 장면 전환 지점이 있다면, 총 검색 시간 T 는 식 (4)와 같다.

$$T = L \times t_c \times k + t_c \left(\frac{n}{L} - k \right) \quad (4)$$

여기서, $\frac{n}{L}$ 은 구간의 갯수이며, 일반적으로 특정구간에 장면 전환 지점이 집중적으로 분포하지 않을 것이므로, k 개 구간에는 장면 전환 지점이 있으며, $\frac{n}{L} - k$ 개 구간에는 장면 전환이 없을 것이다. n 이 충분히 크면 식 (3)을 사용하여 식 (4)를 다음처럼 전개할 수 있다.

$$T = L \times t_c \times \frac{n}{d} + t_c \left(\frac{n}{L} - \frac{n}{d} \right) \quad (5)$$

$$\frac{T}{n} = t_c \left[\frac{L}{d} + \frac{1}{L} - \frac{1}{d} \right] \quad (6)$$

식 (6)의 좌변은 단위 프레임을 검색하는데 걸리는 평균 시간(sec/frame)이다. 그림 6은 식 (6)의 좌변을 프레임간 비교 시간 t_c 로 정규화해서 L 에 대하여 도식한 것이며, 그림 4에 나타난 실험 결과를 설명하는 유용한 식이다. d 는 검색하고자 하는 비디오가 결정되면 일정한 상수이다.

이때, 최소의 검색 시간이 걸리는 L 은 다음 수식에 의해 구해진다.

$$\frac{d}{dL} \left(\frac{T}{n} \right) = t_c \left[\frac{1}{d} - \frac{1}{L^2} \right] = 0 \quad (7)$$

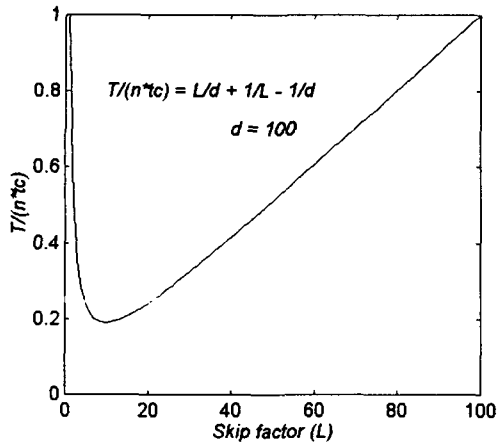


그림 6. $T/(n \times tc) = L/d + 1/L - 1/d$
 Fig. 6. $T/(n \times tc) = L/d + 1/L - 1/d$

$$\therefore L = \sqrt{d} \tag{8}$$

식 (8)은 검색 속도를 최소화하는 검출 간격은 평균 장면 전환 간격(d)에 루트를 취해서 정량적으로 계산할 수 있음을 보인다. 5분 분량의 뉴스의 경우(표 1) d 는 100근처이며, 실험에 의하면(그림 4), 검출 간격이 10근처에서 검색 속도가 최소화되는 것을 볼 때 실험 결과와 아주 잘 일치한다. 식 (6)은 영상의 평균 장면 전환 간격과 검색 속도, 검출 간격과의 관계를 규명하는 식으로 실제 동영상의 경우에 식이 잘 성립하려면, 다음의 두 가지 조건이 만족되어야 한다.

1. 특정구간에 장면 전환 지점이 집중적으로 분포하지는 않는다.
2. 검출 간격 L 로 장면 전환 검출을 실시할 때에도 검출 오류가 발생하지 않을 정도로, 동일한 샷내에서는 프레임간 움직임이 적다.

조건 2가 필요한 이유는, 움직임이 큰 동영상을 검출 간격만큼 떨어진 프레임과 비교하여 장면 전환 검출을 수행하는 경우, 프레임의 내용에 많은 변화가 있어서 알고리즘에 따라서는 장면 전환 지점으로 잘못 판단하는 경우가 발생하여 결국 정밀 검사를 자주 수행하게 되어 검색 시간이 더 길어지기 때문이다.

두 가지 조건이 정확하게 만족되는 가상의 동영상을 생각하여 이것을 이상적인 동영상이라 하겠다. n 프레임 중에서 k 개의 장면 전환 지점이 있으며, 장면 전환 지점의 위치를 임의로 설정하여 조건 1을 만족하게 하였다. 그리고 하나의 샷내에서는 어떤 움직임도 없이 정지영상의 시퀀스만으로 이루어진 동영상을 생성하여, 조건 2를 만족하게 하였다. (그림 7)

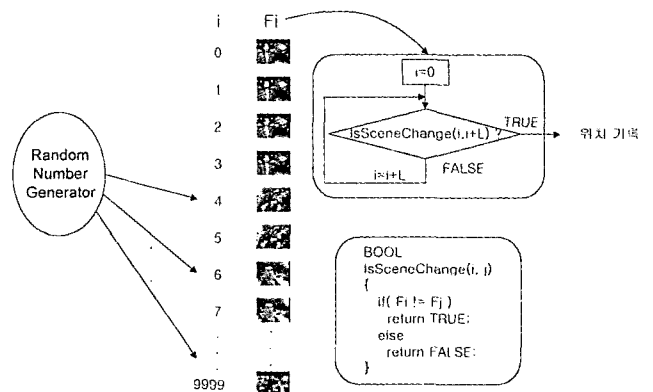


그림 7. 이상적인 동영상
 Fig. 7. Ideal video

그림 8의 실선은 $n = 10,000, d = 100$ 인 이상적인 동영상의 경우에 검출 간격 L 에 대하여, 두 프레임을 비교하는 모듈 IsSceneChange가 호출되는 횟수를 기록한 것이다. 검출 간격이 1인 경우는 10,000번 호출되며, 10인 경우에는 2,000번 정도만 호출됨을 알 수 있다. IsSceneChange 모듈이 한 번 수행되는 데 걸리는 시간에 호출되는 횟수를 곱하면 10,000 프레임을 검색하는데 걸리는 시간을 계산할 수 있겠다.

점선은 식 (8)을 표시한 것으로, 검출 간격이 큰 경우에 이상적인 동영상의 경우와 차이를 보이는데, 이 경우는 임의로 설정된 장면 전환 지점이 검출 간격 내에 한 개 이상 존재하기 때문이다.

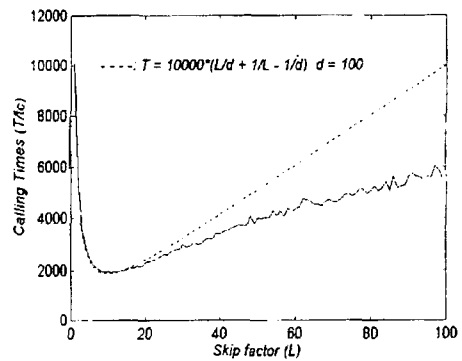


그림 8. 이상적인 동영상의 경우 검색 시간
 Fig. 8. Computation time for ideal video

4. 최적 검출 간격에 의한 장면 전환 검출

이상적인 동영상에서는 평균 장면 전환 간격 d 를 정확하게 줄 수 있었으나, 실제 동영상의 경우는 이것이 알려져 있지 않다. 여기에서는 검색을 진행해 나가면서, 평균 장면 전환 간격을 계산하여, 스스로 검출 간격을 결정하는

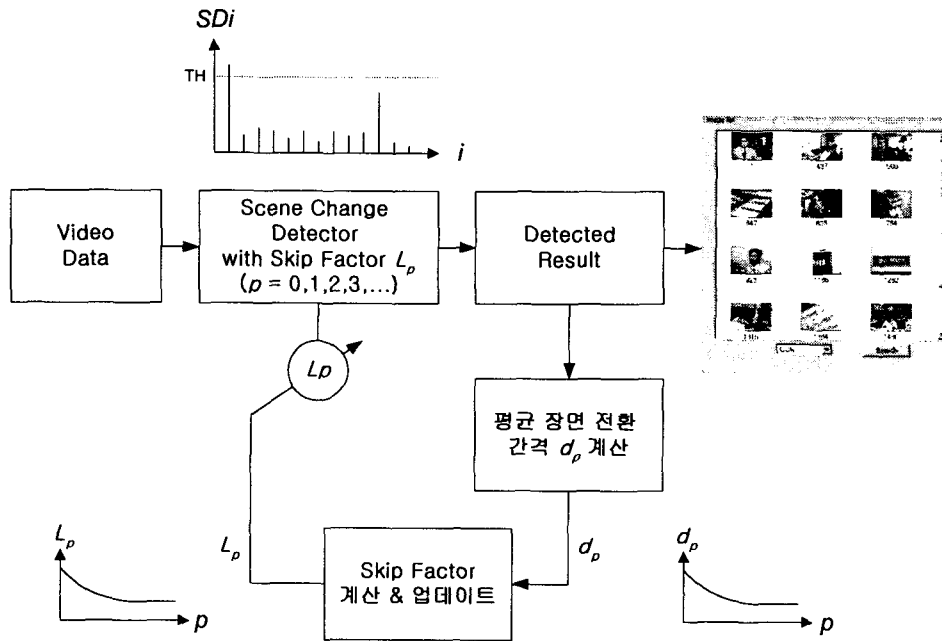


그림 9. 최적 검출 간격 장면 전환 검출
Fig. 9. Scene change detection with optimal skip factor

최적 검출 간격 장면 전환 검출 알고리즘을 제안하고자 한다. 그림 9는 알고리즘 설명도이다. 검출 간격을 먼저 초기치 L_0 로 한다. 일정 기간 U 가 지난 후 검출 결과로부터 평균 장면 전환 간격 d_1 을 계산한다. d_1 로부터 $L_1 (= \sqrt{d_1})$ 를 계산해서, 검출 간격을 갱신한다. 변경된 검출 간격 L_1 을 사용하여 다음 검색을 계속한다. 장면 전환 비율이 알려져 있지 않은 임의의 동영상에 대해서도 2절에서 언급한 장면 전환의 통계적 특성으로 인해 검색이 진행되는 동안 스스로 적절한 검출 간격으로 수렴하게 된다. 검출 간격을 업데이트하는 간격 U 가 짧을수록 수렴 속도가 빠르게 되나, 너무 작은 경우 불필요한 계산이 되므로 적절하게 선택한다.

III. 실험 및 결과

1. 실험 시스템 구성

모든 프로그램 개발은 Visual C++ 5.0 환경 하에서 개발하였으며, MPEG 스트림 재생을 위해서 멀티미디어 데이터를 제어하고 처리하는 구조인 ActiveMovie SDK를 사용하였다^[6]. 그림 10은 전체 장면 전환 검출 시스템의 흐름도이다. ActiveMovie를 사용하여 소프트웨어 MPEG 디코딩을 수행한 다음, 영상의 화소값을 구하기 위해 화면 캡처 방법을 사용하였다. 이것은 마이크로소프트 윈도우에서

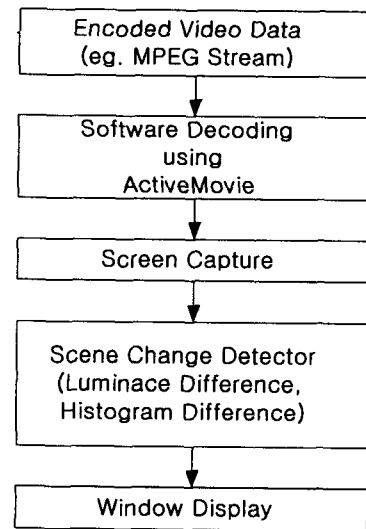


그림 10. 실험 흐름도
Fig. 10. Flow diagram

제공하는 GDI 함수를 사용하여 비트맵을 만들고, 화면상의 특정 영역에서 생성된 비트맵으로 비트 전송을 시켜 가능하였다. 65,536 컬러 VGA의 경우 1 WORD 길이의 화소값이 얻어지며 이것을 R, G, B로 나눈 다음 장면 전환 검출 알고리즘의 입력으로 사용하였다. 장면 전환 검출

모듈에서는 프레임간 밝기차와 히스토그램 차이를 동시에 고려하여 장면 전환 검출 여부를 판단하였다. 검색이 끝난 후에 장면 전환 지점은 원도 상에 조각 그림으로 표시된다.

2. 동영상 검색 소요시간

여러 종류의 MPEG 압축된 동영상에 대하여 검출 간격을 변경하면서 전체 동영상을 검색하는데 걸리는 시간을 조사하였다. 실험에 사용한 영상은 뉴스, 광고, 드라마 등이다. 5분(9,000 프레임) 분량의 뉴스를 여러 가지 검출 간격을 사용하여 실험한 다음, 검출 간격과 검출 시간을 실험으로 측정하여 표 2에 표시하였다. 9,000 프레임 내에 82개의 장면 전환 지점이 존재하였다. 표 2에서 알 수 있는 것처럼 검출 간격이 11에서 검출 시간이 가장 적게 걸렸다.

표 2. 검출 시간의 실험치와 이론치 비교 (5분 분량의 뉴스)
Table 2. Comparison of experimental and theoretical detection time (5 minutes news)

검출 간격	실험 [sec]	식(6) [sec]
1	6395	6395
2	3246	3256
3	2248	2248
4	1778	1774
5	1512	1512
6	1355	1357
7	1263	1263
8	1203	1207
9	1176	1177
10	1169	1164
11	1164	1164
12	1187	1174
13	1191	1191
14	1250	1214
15	1246	1242
16	1251	1274
17	1318	1308
18	1289	1346
19	1296	1385
20	1344	1427

이제 식 (4)가 얼마나 실험 결과를 잘 설명하는지 조사해보자. 실험 결과로부터 t_c 를 계산한다. t_c 는 서로 다른 두 프레임을 비교하는데 걸리는 시간이며, 전체 9,000 프레임을 비교하는데 6,395초 걸렸으므로(검출 간격이 1인 경우), 단위 프레임 비교 시간 t_c 는 $6,395/9,000[\text{sec}/\text{frame}]$ 이다.

이 시간은 ActiveMovie가 MPEG 스트림을 디코딩하는 시간, 특정 프레임으로 랜덤 액세스하는 시간, 프레임간의 밝기차와 히스토그램차이를 계산하고 장면 전환 지점 여부를 판단하는 시간, 장면 전환 지점이면 조각그림으로 만들어 원도에 디스플레이하는 시간 등이 포함된다. 장면 전환 지점간 평균 거리 d 는 $9,000/82[\text{frame}]$ 이다. 식 (6)에 의해서,

$$T = t_c \left[\frac{n}{d} \times L + \frac{n}{L} - \frac{n}{d} \right] \tag{8}$$

$$T = \left(\frac{6,395}{9,000} \right) \left[82 \times L + \frac{9000}{L} - 82 \right]$$

표 2에 검출 간격 L 에 따른 식 (8)의 값을 표시하였다. 실험치와 이론치가 비교적 잘 일치함을 알 수 있다. 검출 시간 T 가 최소가 되는 L 의 이론치는 $10.4(=\sqrt{9,000/82})$ 이며 실험 결과로 나온 11과 근사함을 알 수 있다.

표 3은 4,914 프레임(약 3분) 길이의 광고의 경우에 실험치와 이론치를 비교한 것이다. 실험 결과, 광고의 경우는 검출 간격이 5에서 검출 시간이 최소가 되었다. 이론값은 6에서 나왔는데, 영상의 길이가 비교적 짧아 약간의 오차를 보였다.

표 4는 장면 전환 지점간 평균 거리가 비교적 긴 드라마의 경우이며, 실제 검출 시간이 이론치보다 더 길게 걸렸으며, 최적의 검출 간격도 이론치와 많은 오차를 보였다.

표 3. 검출 시간의 실험치와 이론치 비교 (3분 분량의 광고)
Table 3. Comparison of experimental and theoretic detection time (3 minutes commercial)

검출 간격	실험 [sec]	이론 [sec]
1	3599	3599
3	1483	1395
5	1239	1110
6	1259	1087
7	1308	1099
9	1453	1180
11	1603	1302
13	1742	1447
15	1932	1605

3. 장면 전환 지점간 평균 거리가 큰 경우

비교적 장면 전환 지점간 평균 거리가 긴 비디오의 경우는 실험치와 이론치가 잘 맞지 않는데, 이것은 평균 거리가 길어질수록, 이상적인 동영상의 조건 2가 만족되지 않기 때문이다. 즉, 평균 거리가 길어질수록, 현존하는 장면 전환 검출 알고리즘으로는 그 구간 내에 장면 전환 지

점이 존재하지 않음에도 불구하고, 장면 전환 지점으로 판단하는 경우가 발생되기 때문이다. 실제로 실험에 사용한 5분 분량의 드라마의 경우, 장면 전환 지점간 평균거리가 거의 200에 가깝게 관측되었으며, 이론적으로는 검출 간격이 14에서 최소의 검색 시간이 걸려야 하지만, 실험적으로는 9로 조사되었다. 이것은 비교적 큰 검출 간격으로 장면 전환 검출 시에는 검출 간격이 1인 경우보다, 밝기차나 히스토그램차이의 임계치를 움직임을 고려한 만큼 크게 주어야 함을 나타낸다.

표 4. 검출 시간의 실험치와 이론치 비교 (5분 분량의 드라마)
Table 4. Comparison of experimental and theoretical detection time (5 minutes drama)

검출 간격	실험 [sec]	식(6) [sec]
1	6359	6359
2	3230	3211
3	2215	2182
4	1749	1683
5	1487	1396
6	1329	1215
7	1244	1095
8	1174	1012
9	1158	955
10	1175	916
11	1210	889
12	1211	872
13	1256	862
14	1329	858
15	1340	859
16	1510	864
17	1472	871
18	1504	882
19	1576	894
20	1667	909

4. 최적 검출 간격 장면 전환 검출

제 II장에서 설명한 최적 검출 간격 장면 전환 검출 알고리즘을 실험하였다. 검출 간격을 900[frame]마다 갱신하면서 실험한 결과, 뉴스의 경우는 초기에 긴 앵커샷이 있어서 5,400[frame]($p = 6$)이후에 수렴하였으나, 비교적 평균 장면 전환 간격이 작은 광고의 경우는 900[frame]내에 충분히 많은 장면 전환 지점이 존재하여 뉴스보다 빨리 수렴하는 것으로 관찰되었다. 그림 11은 뉴스와 광고의 경우 알고리즘이 진행하면서 검출 간격 L_p 가 최적의 값으로 수렴하는 모양을 보여준다.

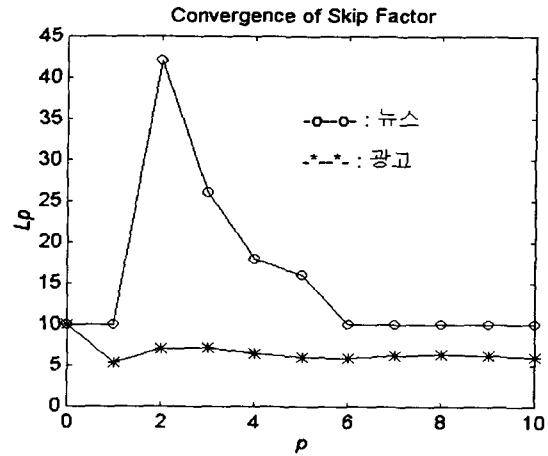


그림 11. 검출 간격의 수렴 형태
Fig. 11. Convergence of skip factor

IV. 결 론

본 연구에서는 장면 전환 검출 시간과 검출 간격, 영상의 장면 전환 특성과의 관계를 수식적으로 유도하고 검출 시간을 최소화하는 최적 검출 간격을 유도된 식으로부터 계산하여 영상의 평균 장면 전환 간격과 관계함을 증명하였다. 정량적인 방법으로 최적의 검출 간격을 구할 수 있었으며, 실험을 통해 확인하였다. 영상의 평균 장면 전환 간격이 알려져 있지 않은 비디오의 경우는 제안된 최적 검출 간격 장면 전환 검출 알고리즘을 사용함으로써 뉴스, 광고, 드라마 등 다양한 종류의 동영상에 대하여 효율적으로 장면 전환 검출이 가능하였다.

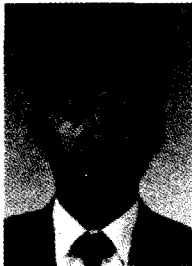
참 고 문 헌

- [1] G. Ahanger and T. D. Little, "A survey of technologies for parsing and indexing digital video," *Journal of Visual Communication & Image Representation*, vol. 7, no. 1, pp. 28-43, 1996.
- [2] H. J. Zhang, A. Kankanhalli and S. W. Smoliar, "Automatic partitioning of full-motion video," *Proc. ACM Multimedia System*, vol. 1, no. 1, pp. 10-28, 1993.
- [3] S. W. Smoliar and H. J. Zhang, "Content-based video indexing and retrieval," *IEEE Multimedia*, vol. 1, no. 2, pp. 62-72, 1994.
- [4] I. K. Sethi and N. Patel, "A statistical approach to scene change detection," *Proc. SPIE*, vol. 2420, pp. 329-338, Mar. 1995.
- [5] Y. Nakajima, "A video browsing using fast scene

cut detection for and efficient networked video database access," *IEICE Trans. Inf & Syst.*, vol. E-77-D, no. 12, pp. 1355-1364, Dec. 1994.

[6] Microsoft Corporation, *ActiveMovie 1.0 SDK Developer's Guide*, 1997.

저 자 소 개



하 명 환

1971년 2월 27일생.
 1993년 2월 경북대학교 공과대학 전자공학과 졸업(공학사).
 1995년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사).
 1995년 2월 ~ 현재 한국방송공사 기술연구소 연구원.
 주관심분야 : 멀티미디어 방송제작, 비디오 인덱싱, 영상 처리.



나 윤 정

1969년 1월 14일생.
 1992년 2월 서울대학교 공과대학 전기공학과 졸업(공학사).
 1994년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사).
 1994년 2월 ~ 현재 한국방송공사 기술연구소 연구원.
 주관심분야 : 멀티미디어 방송제작, 비디오 인덱싱, 영상 처리.



이 상 길

1959년 8월 28일생.
 1981년 2월 서울대학교 공과대학 전자공학과 졸업(공학사).
 1983년 2월 서울대학교 대학원 전자공학과 졸업(공학석사).
 1994년 2월 서울대학교 대학원 전자공학과 졸업(공학박사).
 1995년 NHK 방송기술연구소 객원연구원.
 1983년 2월 ~ 현재 한국방송공사 기술연구소 근무.
 주관심분야 : 멀티미디어 방송, 영상 부호화 및 신호 처리.