

# MSE를 이용한 영상의 점진적 장면전환 검출에 관한 연구

°김단환\*, 김형균\*, 정기봉\*, 신중호\*, 김재석\*, 오무송\*\*  
조선대학교 컴퓨터공학과

## A Study on the Gradual Scene Change Detection of Image Using MSE

°Dan-Hwan Kim\*, Hyeng-Gyun Kim\*, Ki-Bong Joung\*, Choong-Ho Sin\*, Jae-Seog Kim\*, Moo-Song Oh\*\*  
Dept. of Computer Eng. Chosun Univ.\*

### 요약

대용량의 동영상 데이터 이용에 있어 효과적인 동영상 검색을 위해서는 동영상 데이터의 색인과정이 필요하다. 본 논문은 AVI영상에서 동영상 장면 전환점 검출에 관한 효과적인 방법을 제안한다. 제안된 방법은 프레임을 대각선 방향의 픽셀 값을 추출하고 정지영상으로 변환하여 동영상의 전체 구조를 파악할 수 있으며, 장면 전환점을 한눈에 확인할 수 있다. 각각의 프레임에서 추출한 픽셀의 칼라 값을 A행렬  $i$ (프레임 수) $\times$  $j$ (프레임 영상 높이)로 저장하고, MSE(Mean Square Error)를 용해서 일정의 임계값에 도달하지 못할 경우 이 프레임을 장면 전환점으로 검출한다.

### 1. 서론

멀티미디어 기술의 발전으로 현 사회의 전반에 정보 전달이 동영상으로 보편화되고 있는 것으로 보면 많은 정보검색기술이 계속적으로 발전하고 있다. 또한 정보전달 매체의 사용도 대폭 증가되고 있어 그 정보에 대한 효율적인 관리의 중요성이 증대되고 있다.

MPEG 동영상은 효과적인 정보전달의 매체로서 데이터베이스 구축 시 사용자가 편리하게 사용할 수 있는 내용기반 색인이 필요하다. 그러므로 효과적인 동영상의 색인을 하기 위해서 동영상의 의미 있는 장면으로 나누는 장면전환 검출이 필요하다.

본 연구에서는 장시간에 걸친 동영상 데이터를 한눈에 알아볼 수 있도록 각 프레임 별로 영상데이터의 특정영역을 샘플링하여 대용량 동영상 데이터를 소용량의 정지영상으로 전환하는 것으로 연속된 각각의 프레임들을 대각선 방향의 픽셀 값을 추출하여 정지영상을 만들어 낸다. 각각의 프레임에서 추출한 픽셀

의 칼라 값은 행렬A에  $i \times j$ 행렬로  $i$ 는 프레임 수,  $j$ 는 프레임의 영상 높이로 저장하여, MSE(Mean Square Error)를 이용해서 일정의 임계값을 넘지 못하면 다음 프레임을 장면 전환점으로 검출하게 된다.

### 2. 관련연구

동영상의 일반적인 데이터의 구조를 살펴보면, 동영상의 데이터는 다단계의 계층적 구조로 표현될 수 있다. 동영상 데이터의 상위 단계는 스토리(story) 계층으로 나눌 수 있고, 스토리 계층의 바로 아래 단계는 "씬(Scene)"이라는 동영상의 이벤트(event) 단위로 나눌 수 있다. 동영상을 씬 단위로 나누는 것이 가장 이상적인 방법이지만, 씬이라는 것은 의미그룹으로서 자동으로 씬을 검출하는 것이 현재 매우 어려운 작업이라고 할 수 있다. 따라서 물리적 단위인 장면 전환점

을 검출하는 작업이 수행된다. 장면전환의 종류로는 순간적으로 장면이 바뀌는 급격한 장면전환과 여러 프레임이 서서히 바뀌는 점진적인 장면전환 등으로 나눌 수 있다.

동영상 데이터의 구조를 살펴보면, 다단계 계층적 구조로 표현될 수 있다.

- Frame : 동영상 데이터를 구성하는 최소단위로 각각의 픽셀로 이루어진 영상.
- Shot : 동영상을 구분 할 때의 기본단위로서 필름이 끊기지 않고 연결된 프레임들의 집합
- Scene : Shot들의 집합으로 이루어진 공간상이나 시간상의 의미적 분리로 구성
- Cut : Shot과 Shot사이의 경계에 해당하는 장면전환점

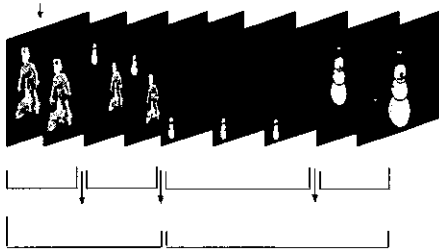


그림 1. 동영상의 구조

일반적인 장면 전환지점 검출 기법에는 다음과 같은 방법이 있다.

① 픽셀차이(pixel difference)를 이용한 방법

동영상 장면 전환 검출을 위해서 각 프레임의 픽셀 값을 이용하여 프레임간의 픽셀 값의 차이는 매우 낮은 값을 갖는다. 그러나 장면 전환이 있는 곳은 픽셀 값들의 차이가 두드러지게 커지므로 이를 이용하면 장면 전환점을 찾아 낼 수가 있다. 이러한 방법은 프레임의 잡음(noise)에 민감하고 특히 점진적인 장면 전환 기법은 검출이 거의 되지 않는 취약점을 보인다.

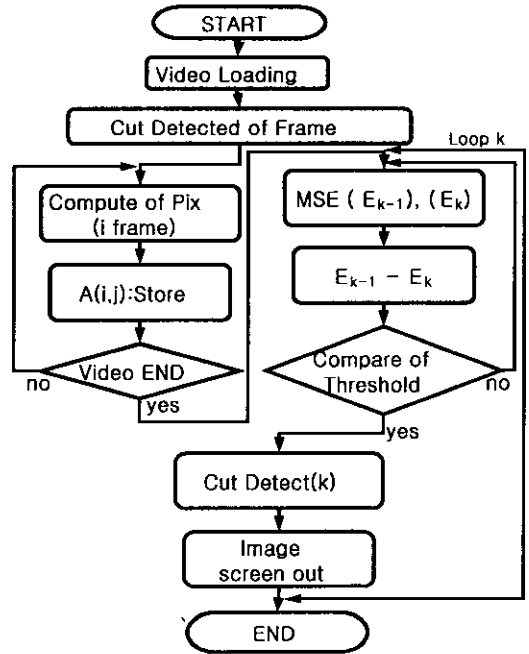
② 명도히스토그램(histogram)을 이용한 방법

이 방법은 가장 널리 사용되는 장면 전환 기법 중의 하나로 프레임의 명도 히스토그램을 찾아 프레임간의 히스토그램의 차이를 비교하여 일정한 임계값 이상이면 장면 전환으로 인식하는 기법으로 픽셀차이를 이용하는 방법보다는 잡음(noise)에 강한 성질을 갖는다. 하지만 이 방법 또한 점진적 장면 전환을 찾는 것에는 취약점을 보인다.

③ DCT 계수를 이용한 방법

압축영상(MPEG)의 일반적 방법인 DCT 변환의 계수를 이용하는 방법이 있다. 이는 압축영상의 스트림에 존재하는 계수의 차이를 이용하는 방법으로 압축영역에서 복원에 필요한 많은 계산을 거치지 않고 바로 장면 전환을 검출하는 특징을 갖고 있다

3. 제안된 장면 전환 검색 시스템



동영상에서 시간적으로 연속적인 프레임들은 서로 유사한 배경 구성과 이미지들의 변형으로 이루어진다. 따라서, 동영상의 데이터정보는 연속성이 끊어지는 부분이 검색 대상이 되며, 그 영상이 장면 전환점이 될 수 있다. 본 연구에서 제안한 점진적인 장면 전환 검출은 동영상 데이터의 각 프레임들을 그림 3과 같이 대각선 방향으로 추출한 픽셀의 칼라 값을 행렬A(i,j)로 만든다. i는 프레임 수, j는 프레임의 영상 높이로 저장한다.

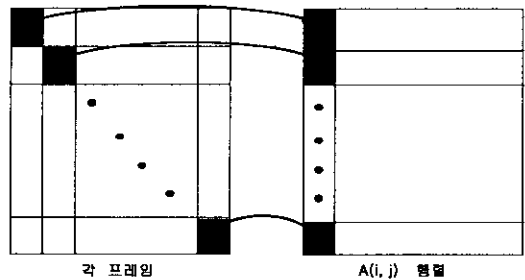


그림 3. 프레임의 추출한 픽셀값 저장

동영상의 각 프레임에서 추출된 픽셀 값은 소용량의 정지영상으로 샘플링한다. 그림 4에서와 같이 샘플링된 영상에서 장면 전환점은 수직선상 픽셀 값들의 확연한 변화가 나타나게 되며, 점진적인 장면전환점은 픽셀 값들의 점진적인 변화로 나타나게 된다. 이러한 영상 데이터는 여러 프레임간의 간격으로 샘플링함으로써 그 크기를 줄일 수가 있다. 또한 동영상 데이터의 전체 구조를 파악할 수 있다는 특징을 가진다.



그림 4. 대각선 방향의 픽셀 값을 샘플링한 영상

동영상의 각 프레임들에 대한 대각선 방향으로 추출한 픽셀의 칼라 값을 A행렬에 저장하고, 장면 전환이 일어나는 곳은 MSE(Mean Square Error)를 사용하여 검출한다.

MSE를 이용한 행렬 A의 열성분  $a_{k-1}$  와  $a_k$ 의 차이는 식(1)과 같다.

$$E_k = \sqrt{(a_{1(k-1)} - a_{1k})^2 + \dots + (a_{i(k-1)} - a_{ik})^2} \quad \text{--- (1)}$$

식(1)에서  $E_k$ 를 이용하여 식(2)와 같이 장면 전환점 프레임 번호를 검색하고 화면에 프레임을 출력한다.

$$|E_{f-1} - E_f| \geq \text{임계값}, f: \text{프레임 번호} \quad \text{---- (2)}$$

#### 4. 실험 환경

본 연구는 Windows 98 환경에서 Visual C++ 6.0을 이용해 구현하였으며, 실험에 사용한 자료는 애니메이션 영상으로 320\*240 RGB 영상이다. 총 frame 수는 900, 플레이 타임은 65sec이다.

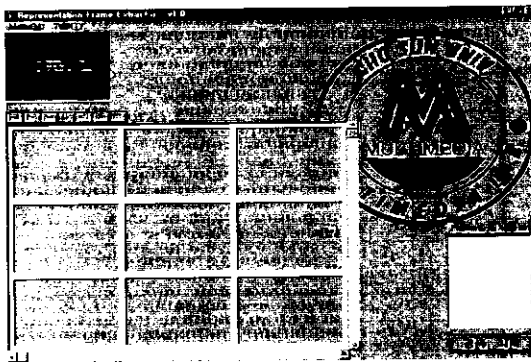


그림 5 장면 전환 검색 시스템 초기화면

장면전환 검색 시스템의 초기화면은 그림 5와 같다. 먼저 장면전환점을 검색할 동영상의 데이터를 받아들이고 검색시작을 선택하면 시스템은 실행된다. 장면전환점으로 추출한 프레임은 왼쪽하단에 출력하며, 그 영상을 선택할 경우, 그 프레임부터 동영상 데이터를 확인할 수 있도록 하였다.

장면전환점으로 검색된 프레임들은 그림 6과 같은 결과를 얻을 수 있었다.

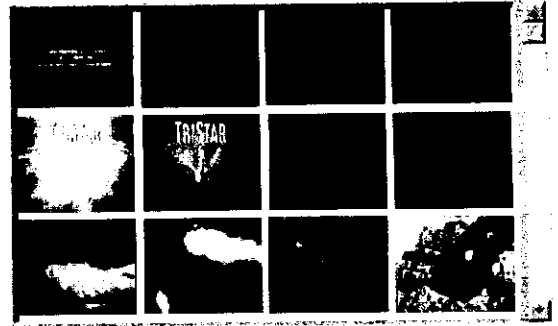


그림 6. 장면 전환점 검출 리스트

#### 4. 결 론

본 연구에서는 동영상 정보의 효율적인 검출을 위한 색인 방법으로 AVI 동영상에서 MSE를 이용한 장면 전환점 검출방법에 대해 제안하였다. 각각의 비교 프레임은 대각선 방향의 픽셀 값을 추출하여 행렬에 저장하여 정지영상으로 변환하게 된다. 이렇게 만들어진 정지영상은 한눈에 장면전환점을 파악할 수 있으며, 저장된 픽셀 값을 MSE를 이용하여 프레임간의 차이를 구해 임계값보다 크거나 같을 경우 장면 전환점으로 검출하였다.

실험에 의해서 총 프레임 수 900, 장면전환점의 프레임 수 148, 인어진 프레임 143, 검출율 97%를 보여 주고 있으며, 히스토그램을 이용하여 검출한 장면 전환점은 검출율 약 92%정도로 보다 좋은 검출 결과를 확인할 수 있었다.

향후 연구 방향은 삼차원 영상에서 장면전환 검출과 감시 카메라를 이용한 물체의 움직임을 검출하는 방법에 대한 연구가 필요할 것으로 본다.

#### 참고문헌

- [1] H. Zhang, et, al., "Automatic partitioning of full motion video", Multimedia System, Vol.1, pp.10-28, 1993.
- [2] W. Niblack et al. The Qbic project Querying Images, by content using color, texture, and shape. In SPIE 1908, Storage and Retrieval for Image and Video Databases, Feb. 1993.
- [3] Hampapur, A., Jain, R., and Weymouth, T., "Digital Video Segmentation", Proc. ACM Multimedia 94, San Francisco, CA, pp.357-364, October, 1994.
- [4] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full Motion Search for Object Appearance", Proc. of IFIP on Visual Database System, pp. 113-127, Sep. 1991.
- [5] Boon-Lock Y and Bede Liu, "Rapid scene analysis on compressed video", IEEE Tr. on CSVT, Vol.5, no.6, pp.533-544, Dec. 1995.
- [6] William I, Grosky, Peter Neo, Rajiv Mehrotra, "A Pictorial Index Mechanism for Model based Matching", Data and Knowledge Engineering, vol.8, pp.309-327, 1992.
- [7] 김성철, 오일근, 장종환, "동영상의 고속 장면전환 검출을 위한 이진검색 알고리즘에 관한 연구", 제12회 정보처리학회 추계학술대회, 1999.
- [8] 이명주, 김형균, 정기봉, 오우송, 김태성, "칼라임계값을 이용한 동영상의 컷 검출에 관한 연구", 제3권 1호, 한국멀티미디어학회 춘계학술대회, 2000.