

---

# MSE를 이용한 동영상데이터의 컷 검출시스템에 관한 연구

김단환\* · 정기봉\* · 오무송\*\*

## A Study on the Cut Detection System of Video Data using MSE

Dan-Hwan Kim\* · Ki-Bong Joung\* · Moo-Song Oh\*\*

### 요 약

컴퓨터의 발전과 정보통신기술의 발달은 멀티미디어 기술 확산과 대용량의 멀티미디어 동영상데이터 사용을 증가시켰다. 동영상데이터는 전체적인 데이터를 파악하고, 원하는 동영상을 바로 재생 가능할 뿐만 아니라 동영상 데이터의 정보가 요약된 리스트제공이 필요하다. 그리고 효과적인 동영상 검색을 위해서는 동영상 데이터의 색인과정이 필수적이며 꼭 필요한 기술이다.

따라서 본 연구는 동영상 데이터 내용기반 색인에 기초가 될 프레임의 컷 검출의 효과적인 방법을 제안한다. 제안된 방법은 동영상데이터 각각의 프레임을 대각선 방향으로 나누어 일정한 픽셀 색상 정보 값을 추출하여 각각 프레임에서 추출된 칼라색상의 픽셀값은  $A(i, j)$ 행렬로  $j$ 는 프레임 수,  $i$ 는 프레임의 영상 높이로 저장한다. 저장되어진 픽셀값은 MSE(Mean Square Error)을 이용하여 프레임간의 특정 값의 차이를 임계 값보다 클 경우 빠르고 정확하게 컷을 검출하였다. 프레임 컷 검출에 대한 실험을 포괄적으로 하기 위해 여러 종류의 동영상 데이터를 실험 대상으로 하여 컷 검출 시스템의 성능을 비교 분석하였다.

### ABSTRACT

The development of computer technology and the advancement of the technology of information and communications spread the technology of multimedia and increased the use of multimedia data with large capacity. Users can grasp the overall video data and they are able to play wanted video back. To grasp the overall video data it is necessary to offer the list of summarized video data information. In order to search video efficiently an index process of video data is essential and it is also indispensable skill. Therefore, this thesis suggested the effective method about the cut detection of frames which will become a basis of an index based on contents of video image data. This suggested method was detected as the unchanging pixel color intelligence value, classified into diagonal direction. Pixel value of color detected in each frame of video data is stored as  $A(i, j)$  matrix -  $i$  is the number of frames,  $j$  is an image height of frame. By using the stored pixel value as the method of MSE(Mean Square Error) I calculated a specified value difference between frames and detected cut quickly and exactly in case it is bigger than threshold value set in advance. To carry out an experiment on the cut detection of frames comprehensively, I experimented on many kinds of video, analyzing and comparing efficiency of the cut detection system.

### 키워드

MSE(Mean Square Error), Cut Detection, Video Data, threshold value

## 1. 서 론

정보통신 기술의 비약적인 발전은 멀티미디어 데이터의 활용을 다양한 분야에서 적극적으로 요구하게 되었다. 이러한 멀티미디어 정보 중에서도 동영상 데이터는 인간에게 많은 정보를 자연스럽게 제공해 줌으로서 관심의 대상이 되고 있다. 동영상 자료검색, 전송 그리고 데이터를 처리하는 환경의 수요가 해마다 증가하고 있다. 멀티미디어 자료 중에서 동영상 자료는 데이터의 양이 방대하며, 정보 전달 매체로 통신 네트워크를 이용하는 경우 빠른 시간 내에 정확한 정보검색과 처리 할 수 있는 기술이 매우 중요하다.

기존의 멀티미디어 데이터 검색 방법은 영상의 내용(content)을 기반[1]으로 하는 여러 방법들이 연구되고 있다. 이러한 방법은 기존의 텍스트 기반의 단점을 보완하고자 물체의 모양, 질감, 색상과 같은 영상의 특징을 이용하여 검색하는 방법이다 [2]. 동영상의 내용기반 검색을 위해서는 동영상 인덱싱이 중요한 기술이다. 이러한 방법중의 하나로 히스토그램 방법이 제안되었지만, 히스토그램의 임계값이 영상값의 분포도에 따라 변경되는 문제점이 있다. 이러한 문제점을 해결하기 위해 컷 검출기법을 제안한다. 이 기법의 기본 원리는 프레임 간 MSE를 측정해서 이 값이 주어진 임계값을 초과하면 장면의 변화가 일어나는 곳으로 검출하는 것이다.

본 연구에서는 동영상 데이터를 시각적으로 컷을 한 눈에 파악 할 수 되도록 각 프레임 별로 특정영역을 샘플링 하였다. 샘플링된 자료는 소용량 정지영상으로 변환한 것이며, 또한 컷이 발생하는 프레임임을 확인할 수 있도록 하였다. 급격한 장면이 전환되는 동영상데이터와 점진적으로 장면이 전환되는 동영상데이터의 컷을 검출한다. 이 컷 검출 알고리즘은 픽셀값을 추출하여 칼라색상의 픽셀값을  $A(i, j)$ 행렬로  $j$ 는 프레임 수,  $i$ 는 프레임의 영상 높이로 저장하고, MSE(Mean Square Error)를 이용하여 전환프레임간의 특정 값의 차이를 구해 미리 정해진 임계값보다 클 경우 컷을 검출한다. 검출된 컷은 실험을 포괄적으로 하기 위해 여러 종류의 동영상 데이터를 실험 대상으로 하여 성능을 비교 분석한다.

## II. 관련연구

### 1. 동영상의 구조

동영상 데이터의 구조는 다단계 계층적 구조로 표현하여 프레임(frame)은 동영상 데이터를 구성하는 최소 단위로 각각의 픽셀로 이루어진 필름 한 장에 해당하는 하나의 정지영상이다. 샷(shot)은 동영상을 구분 할 때의 기본단위로서 필름이 끊기지 않고 연결된 프레임 집합의 구성이다. 장면은 하나의 촬영대상이 연속된 샷들의 집합으로 이루어진 공간상, 시간상의 의미적 분리의 구성이다. 컷은 샷과 샷 사이의 경계이며 장면이 변화된 프레임이다.

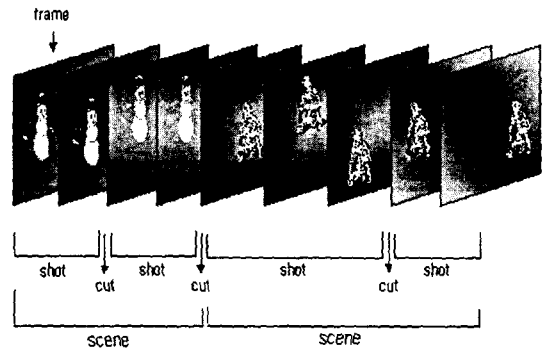


그림 1. 동영상 데이터의 구조  
Fig. 1 Structure of Video Data

### 2. 일반적인 컷 검출 기법

#### ① 픽셀차이를 이용한 방법[3,4,5]

동영상 컷 검출을 위해서 각 프레임의 픽셀값 차이를 이용하는 방법으로 픽셀값의 차이는 매우 낮은 값을 갖는다. 그러나 컷이 있는 프레임은 차이가 두드러지게 커지므로 이를 이용하여 검출할 수 있다. 이러한 방법은 프레임의 잡음(noise)에 민감하고 특히 점진적인 컷 검출은 거의 되지 않는 취약점이 있다.

#### ② 명도 히스토그램을 이용한 방법[3,4]

이 방법은 가장 널리 사용되는 컷 검출 중의 하나로 프레임의 명도 히스토그램의 차이를 비교하여 일정한 임계값 이상이면 장면전환으로 인식하는 기법으로 픽셀차이를 이용하는 방법보다는 잡음에 강한 성질을 갖는다. 그러나 점진적 컷 검출에는 취약점을 보인다.

#### ③ DCT계수를 이용한 방법[5,6]

압축영상(MPEG)의 일반적 방법인 DCT변환의 계수를 이용하는 방법이다. 이는 압축영상의 스트림에 존재하는 계수의 차이를 이용하는 방법으로

압축영역에서 복원에 필요한 많은 계산을 거치지 않고 바로 컷 검출하는 특징을 갖고 있다.

표 1. 동영상 데이터의 MSE 추출 알고리즘  
Table. 1 MSE detect Algorithm of Video data

### III. 제안한 컷 검출 시스템

#### 1. 동영상 데이터의 특정영역 표본추출

동영상 데이터는 장시간에 연속적인 프레임들이 서로 유사한 배경 구성과 이미지들의 변형으로 이루어짐으로 연속성이 끊어지는 부분이 본 연구에서 검색 대상이 되며 그 영상이 컷이 된다. 각 프레임 별로 동영상 데이터의 대각선 방향으로 특정영역을 표본추출 하여 대용량 동영상 데이터를 소용량의 정지영상으로 전환과 A(i, j)행렬로 j는 프레임 수, i는 프레임의 영상 높이로 저장한 내용을 그림으로 도식화하면 그림 2와 같다.

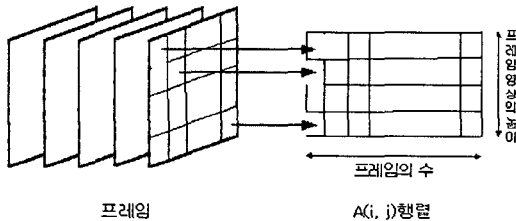


그림 2. 프레임에서 추출된 픽셀값의 저장구조  
Fig. 2 Store structure of pixel value detected in frame

본 연구에서는 일반적인 장면 전환의 프레임의 컷 검출과 장면 전환이 점진적으로 변화는 장면 전환 프레임을 검출하는데 중점을 둔다. 이를 위하여 동영상 데이터 A를 정의하는데 A(i, j)행렬로 j는 프레임의 수, i는 프레임 영상의 높이)로 나타낸다. 각 프레임의 픽셀의 RGB값들을 대각선으로 하여서 대표 값들로 표본추출된 행렬 A의 열 ai과 ai+1의 차이가 큰 i열이 장면 전환이 일어나는 프레임의 컷으로 검출한다.

#### 2. 프레임의 MSE(Mean Square Error)추출 알고리즘

동영상의 각 프레임들의 RGB값을 대각선 방향으로 추출한 값들은 행렬 A의 열 ai와 ai+1의 차이가 큰 i열이 장면 전환이 일어나는 곳이다. 행렬 A의 열 ai과 ai+1의 차이를 정의하기 위하여 다음 알고리즘을 이용한다.

```

;동영상데이터 추출
for( i = 1 to frame x 축출)
    store a[i][j] = Image[i][j]; //j frame의 대각선방향의 데이터 값
    Pict_Image[i][j] = Image[i][j]; // 프레임의 대각선 방향의 데이터 표본 추출
;추출 데이터의 SSD구함
for( f = 1 to frame_end) {
    for( g = 1 to i) {
        hap = hap + (Store a[g][f] - Store a[g][f+1])2 ;
    }
    SSD_Tab[f] = SQRT(hap);
    hap = 0;
};
;임계값과 비교
k = 0; //컷 frame 수 카운터
for( f = 1 to frame_end) {
    Chang[f] = SSD_Tab[f] - SSD_Tab[f+1];
    if (Chang[f] >= t) {
        Chang_Image[] = f; //frame 데이터 저장
        k = k + 1;
    }
};
    
```

$$E_1 = \sqrt{(a_{11} - a_{12})^2 + (a_{21} - a_{22})^2 + \dots + (a_{i1} - a_{i2})^2} \dots\dots\dots(1)$$

식 (1)에서 E1은 영상 데이터 행렬 A의 1열 a1과 2열 a2의 MSE를 의미한다.

$$E_k = \sqrt{(a_{1(k-1)} - a_{1k})^2 + (a_{2(k-1)} - a_{2k})^2 + \dots + (a_{k(k-1)} - a_{k2})^2} \dots\dots\dots(2)$$

식 (2)에서 Ek는 영상 데이터 행렬 A의 k-1열 ak-1과 k열 ak의 MSE이다. 이렇게 MSE를 이용하여 식(2)에서 얻어진 Ek를 이용하여 프레임의 컷 f를 찾을 수가 있다.

$$|E_{f-1} - E_f| \geq t \dots\dots\dots(3)$$

식(3)에서 t는 프레임 컷을 찾는 임계값[4]이고, Ef는 전환이 일어나는 프레임 번호를 나타낸다. 표본추출된 정지영상 비디오 스트림에 대한 E값의 변화가 크기 때문에 Cut 전환점보다 높게 나올 수가 있다. 따라서 본 논문에서는 표본추출된 정지영상 데이터의 수직 에지 성분을 두드러지게 함으로써 Cut 전환점의 E값을 올려주어 그 효율을 높여준다. 여기서 사용하는 에지 추출방법으로 소벨오퍼레이터를 사용한다.

소벨오퍼레이터는 이미지에서 거리가 가까운 픽셀의 값이 원 픽셀의 값과 근사한 값이면 감소시키고, 거리가 가까운 픽셀의 값이 원 픽셀의 값과 차이가 많이 많으면 값을 증가시킨다. 이러한 원리를 이용하는 소벨오퍼레이터[7,8] 값들은 다음과 같다.

표 2 소벨오퍼레이터의 값들  
Table. 2 Values of Sobeloperator

-1	0	1
-2	0	2
-1	0	1

소벨오퍼레이터를 통과한 정지영상데이터로 MSE 추출 알고리즘을 이용하여 Cut 장면은 물론 점진적인 장면전환점까지 일어나는 부분을 판단할 수 있다.

#### IV. 실험 및 결과

본 연구의 실험은 영상처리시간의 단축을 위해서도 사용자의 간섭 없이 특징의 추출 전 과정이 빠르게 진행하기 위해 Windows 2000 Server환경에서 RAM은 512MB, PentiumIII 700Mhz, GeForce2MX, Visual C++ 6.0을 이용하였다. 실험에 사용한 동영상데이터는 일상에서 접할 수 있는 기회는 무수히 많다. 이렇게 많은 동영상 중에서 광고, 뉴스, 영화, 드라마 네 가지 데이터를 동일한 환경에서 AVI 파일 포맷에 맞춰 성능 측정을 하였다.

표 3. 컷 검출에 이용한 동영상 데이터의 포맷  
Table. 3 Format of Video data using Cut detect

동영상종류	프레임크기	프레임수	컷의수	파일포맷
광고	320x320	480	95	AVI
뉴스	320x320	625	196	
영화	320x240	805	217	
드라마	320x240	98	13	

동영상데이터의 컷 검출 시스템의 초기화면은 화면 좌측에 검색에 이용될 동영상데이터를 로드되고 임계값을 입력한 후 검색 시작버튼을 누르게 되면 컷 검출이 시작된다. 이렇게 검출된 프레임은 중간화면에 프레임의 정보와 리스트를 출력한다. 또한 임계값은 검출된 컷의 수가 바뀌므로, 장르에 따라 최적화된 임계값으로 변환시켜 사용하면 더 나은 결과를 얻을 수 있다. 실험에서 사용된 임계값은 최대 크기를 1로 정규화 시킨 수치이다.

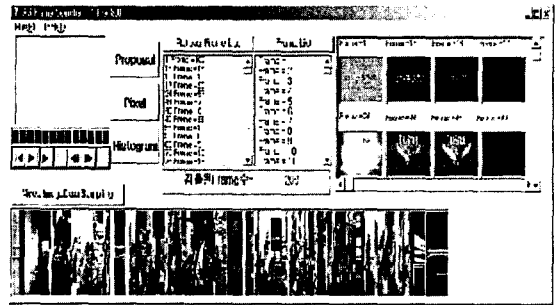


그림 3. 제안한 알고리즘에 의한 컷 검출 결과  
Fig. 3 Cut detect result by proposition algorithm

실험에 사용한 광고, 뉴스, 드라마, 영화가 각기 화면구성에 특성을 지니고 있음에도 불구하고 본 방법이 모든 유형의 동영상에 효율적인지 판단하기 위하여 다양한 데이터를 이용하였다. 먼저 각각의 동영상에 대한 컷을 검출해 정확도와 재현율을 비교한다.

$$Precision\ rate = \frac{number\ of\ detection\ cuts}{number\ of\ total\ frame\ cuts}$$

표 4. 정확율의 비교  
Table. 4 Compare of Precision rate

Precision rate	pixel	histogram	DDS
광고	63.16 (60/95)	61.05 (58/95)	84.21 (80/95)
뉴스	80.61 (158/196)	88.78 (174/196)	91.84 (180/196)
영화	82.49 (179/217)	91.24 (198/217)	96.31 (209/217)
드라마	76.92 (10/13)	84.62 (11/13)	84.62 (11/13)

$$Recall\ rate = \frac{number\ of\ precision\ cuts}{number\ of\ detection\ cuts}$$

표 5. 재현율의 비교  
Table. 5 Compare of Recall rate

Recall rate	pixel	histogram	DDS
광고	81.67 (49/60)	77.59 (45/58)	88.75 (71/80)
뉴스	89.24 (141/158)	77.01 (134/174)	97.22 (175/180)
영화	84.36 (151/179)	81.82 (162/198)	90.43 (189/209)
드라마	60.00 (6/10)	63.64 (7/11)	81.82 (9/11)

현재 비교의 대상으로 선정된 픽셀차에 의한 검출, 히스토그램에 의한 검출은 컷 검출에 보편적으로 이용하는 방법이므로 제안한 알고리즘과의 비교를 표 3과 표 4와 같이 확인할 수 있다. 광고영상에서는 다른 방법에 비해 7%정도 우수한 검출율을 보여주고 있다. 검출율의 비교에서 드라마처럼 장면의 변화가 적은 동영상데이터에서 장면변화가 많은 광고 데이터까지 14%에서 21%의 검출 결과 차이를 보였다. 검출율은 precision rate를 전체 검출된 프레임수로 나누어진 비율을 의미하며, 전체적으로 본 연구에서 제안한 알고리즘이 19% 정도의 효율적인 검출결과를 보였다.

## V. 결 론

멀티미디어 기술과 멀티미디어데이터를 이용한 자료이용이 보편화되고 있는 현재의 정보화 환경에서 동영상으로 이루어진 데이터의 정보를 효과적으로 이용하고 관리하기 위해서 내용기반으로 한 동영상데이터의 요약된 정보와 검색 그리고 인덱싱하는 과정에서 기초가 되는 프레임의 컷 검출이 필수적이다. 본 연구에서는 동영상데이터의 컷 검출에 대한 새로운 알고리즘을 제안하였고, 제안한 알고리즘은 프레임의 픽셀의 색상차를 이용한 방법을 적용하여 각각의 프레임을 대각선 방향으로 영상 데이터를 검출하여 대용량의 동영상 데이터를 정지영상으로 표본 추출하여 동영상의 전체 구조를 파악할 수 있도록 하였으며, 각각의 프레임에서 추출한 픽셀의 칼라 값은  $A(i, j)$ 행렬로  $j$ 는 프레임 수로  $i$ 는 프레임의 영상 높이로 저장하여 MSE를 이용하여 장면변화가 발생하는 컷을 신속하고 정확한 컷 검출이 이루어지도록 컷 검출 시스템을 구현하였다.

본 연구에서는 각각의 동영상의 내용에 따라 표현하는 방법이 다른 네 가지 종류의 동영상 데이터를 비교대상으로 사용하였다. 기존의 방법에서 많이 이용되고 있는 픽셀값에 의한 방법, 히스토그램에 의한 검출법과 비교한 결과에서 detection rate에서 20%정도의 효과적인 검출을 하였으며, precision rate면에서는 10%정도 향상됨을 확인하였다. 기존의 방법들은 점진적인 장면 변화가 있는 부분에서는 민감하지 반응하지 못하였으나 제안한 알고리즘에서는 대체로 효율적이었다. 따라서 본 연구에서 제안한 시스템이 의료기관의 의료 영상 데이터처럼 프레임의 변화가 미세한 부분에 이용하였을 때 매우 효과적이다. 앞으로 영상데이터를 인덱싱하기 위해서는 단순한 프레임 전환뿐만 아니라 특수효과를 내는 카메라 등의 이미지와 같은 복잡한 프레임 변화에도 검출되도록 하는 연구가 계속되어야 할 것이다.

## 참고문헌

- [1] R. John, et al. "Integrated Spatial and Feature Image Systems : Retrieval, Analysis and Compression", Submitted in partial fulfillment of the requirements for the degree of Doctor of philosophy in the Graduate School of Arts and Sciences, Columbia University, 1997.
- [2] S. Berretti, A. Del Bimbo, P. Pala, "Sensations and Psychological Effects in Color Image Database", IEEE International Conference on Images Processing, NewYork USA July, pp.560-563, 1997.
- [3] B. L. Yeo, b. Liu, "Rapid Scene Analysis on Compressed Video", IEEE Tran. on Circuit and System for Video Technology, Vol.5, No.6. pp.533-544, December 1995.
- [4] C. W. Neo, T. C. Pong, R. T. Chin, "Video Partioning by Temporal Slice Coherency", IEEE Tran. on Circuit and System for Video Technology, Vol.11, No.8, Aug., 2001.
- [5] F. Arman, A. Hsu, and M.-Y. Chiu, "Image Processing on Compressed Data for Large Video Databases", Proc. 1st ACM Intl. Conf. on Multimedia, Anaheim CA, pp. 267-272, August 1993.
- [6] J. Meng, Y. Juan, S. F. Chang, "Scene

Change Detection in a MPEG Compressed Video Sequence", Proc. IS&T/SPIE 2419, February 1995.

- [7] M. G. Chung, H. K. Kim, M. H. Song, "A Scene Boundary Dection Method", Image Processing 2000, Proceeding 2000 International Conference, Japan, July. pp.933-936, 2000.
- [8] R. Lienhart, A. Zaccarin, "A System for Reliable Dissolve Detection in Videos", Proceeding. 2001 International Conference on Image Processing, Vol.3, Japan, pp. 406-409, August, 2001.

저자소개

김단환(Dan-Hwan Kim)



1995년 3월 광주대학교 전자계산학과 공학사  
 1999년 8월 조선대학교 산업대학원 컴퓨터공학과 공학석사  
 2004년 2월 조선대학교 대학원 컴퓨터공학과 공학박사 수료

※ 관심분야 : 멀티미디어, 영상처리, 영상통신

정기봉(Ki-Bong Joung)



1994년 3월 광주대학교 전자계산학과 공학사  
 1997년 8월 조선대학교 산업대학원 컴퓨터공학과 공학석사  
 2003년 2월 조선대학교 대학원 컴퓨터공학과 공학박사

※ 관심분야 : 멀티미디어, 영상처리, 영상통신

오무송(Moo-Song Oh)



1965년 2월 조선대학교 전기공학과 공학사  
 1968년 9월 조선대학교 전기공학과 공학석사  
 2001년 2월 전남대학교 전기공학과 공학박사

1994년 3월~1994년 7월 조선대학교 교수협의회 의장 역임  
 1994년 1월~1995년 12월 한국정보과학회 호남제주지부장 역임  
 1996년 12월~1997년 2월 일본 오사카 부립대학 객원교수  
 2000년 3월~2004년 10월 한국전기전자재료학회 평의원  
 2000년 12월~현재 광주교도소 교화위원  
 2003년 2월~2003년 8월 호주 티스마니아 대학 객원교수  
 2003년~현재 한국과학기술단체 총연합 광주전남지회 부회장(현)  
 2004년 11월~현재 한국전기전자재료학회 부회장(현)

※ 관심분야 : 멀티미디어, 영상처리, 영상통신