

# 비디오 장면변환 검출 알고리즘

## New Algorithm of Video Cuts-Detection

이동섭, 김재원, 배석찬, 이양원  
군산대학교 컴퓨터 과학과

Dong-Seop Lee, Jae-Weon Kim, Seok-Chan Bae, Yang-Weon Lee  
Dept. of Computer Science, Kunsan National Univ.  
Tel and Fax : 0654-469-4552, E-mail : leeds@ai.kunsan.ac.kr

### ABSTRACT

본 논문은 뉴스 비디오 데이터베이스를 구축하기 위한 장면 분할 기법 중 Color Histogram에서 각각의 RGB를 따로 계산하여 차이값을 세분화하는 장점과,  $\chi^2$  Histogram에서 차이값을 강조하는 장점을 이용하여 NTSC 표준에 따른 가중치를 적용한 새로운 장면 분할 방법을 제안하였다.

제안 알고리즘의 성능 평가를 위한 실험 도메인은 국내 KBS, MBC, SBS 방송의 뉴스 비디오와 국외 CNN, NHK의 뉴스 비디오를 택하였다. 주어진 환경내에서 제안한 방법을 기존의 Color Histogram,  $\chi^2$  Histogram, 그리고 Bin to bin difference (B2B)과의 실험결과와 비교하여 효율적임을 보였으며, 임계치의 자동선택 가능성을 제시 하였다.

## 1. 서 론

최근 멀티미디어 기술이 발전하면서 다양한 데이터를 저장, 관리, 검색하는 데이터베이스 기술이 요구되고 있다. 특히 동영상과 같은 멀티미디어 데이터는 인간에게 많은 정보를 매우 자연스럽게 제공해 줌으로써 많은 관심의 대상이 되고 있으며, 특정한 문제 영역에 맞도록 저장하고 이용하는 내용 기반 멀티미디어 데이터베이스의 구축에 대한 연구가 활발히 진행 중이다.

이러한 동영상 자료 중에서도 뉴스 비디오는 사회 전반의 중요한 정보를 함축적으로 짧은 시간대에 표현하는 것으로, 그 시대를 분석하는 귀중한 자료가 되고 구조적인 형태를 가지

므로 내용 기반에 의한 비디오 검색의 주요 연구 대상이 되어 왔다.

본 논문에서는 뉴스 비디오 데이터베이스에서 검색과 결과의 대상이 되는 장면을 분류하기 위한 여러 기법들을 살펴보고 이러한 기법들을 이용하여 보다 나은 결과를 위해 실험 결과를 비교 분석하여 새로운 기법을 제시한다.

본 논문의 구성은 다음과 같다. 2장에는 현재까지 진행되어 온 관련 연구에 대해서 고찰하였고, 3장에서는 본 논문에서 제안한 방법을 설명한다. 4장에서는 현재의 뉴스 도메인을 실험 도메인으로 정하고, 기존의 방법과 제안한 방법을 적용한 후의 실험 결과에 대해서 비교 분석하였다. 5장에서는 본 논문의 결론 및 향후 연구 방향에 대하여 설명하였다.

## 2. 기존 연구

장면을 분할하기 위한 방법에는 여러 가지가 있다. 프레임간의 차이를 이용하는 방법과 칼라 히스토그램을 이용하는 방법, 객체의 이동을 계산하여 사용하는 방법, 에지를 추출하여 이용하는 방법과 전체적인 웨이블릿(wavelet)을 측정하여 이용하는 방법 등 많은 방법이 사용되었다. 이러한 방법 중 뉴스 도메인처럼 갑작스런 장면의 검출에는 주로 히스토그램 차이를 계산하여 사용하는 방법이 많이 사용되었다.

일반적으로 히스토그램 차이 값을 적용시키는 방법에 사용된 식은 다음과 같다[1][2][3][4].

$$d(I_i, I_{i+1}) = \sum_{j=0}^{N-1} |H_i(j) - H_{i+1}(j)| \quad (\text{식 1})$$

$d(I_i, I_{i+1})$ 는 이웃 프레임과의 유사성을 측정하는 함수로,  $H_i(j)$ 는  $i$ 번째 프레임의  $j$ 번째 히스토그램 요소 값이라는 의미이다. 즉, 앵커 장면에서는 배경이 일정하므로 에지 히스토그램의 모양이 거의 비슷하다. 그러므로 차이값이 크게 나타나는 곳은 앵커 장면에서 사건 장면으로 넘어가는 장면이라고 할 수 있다. 그리고 사건 장면에서는 일반적으로 장면의 변화가 많아 이웃한 프레임과의 차이가 크게 나타나는 경향이 있는데, 이럴 경우 첫 프레임과 끝 프레임의 키 프레임으로 선택한다[1][2][4].

위의 방법은 일반적인 방법으로 데이터의 성격에 따라 다음과 같은 여러 가지 방법들이 제시되어 사용되었다.

### Template Matching

이 방법은 동일한 위치의 두 영상의 픽셀을 비교하는 방법으로 방법이 매우 간단하고 수행 시간이 적게 드는 장점이 있으나 객체의 이동이나 잡음에 매우 민감하여 불만족스러운 결과를 산출할 수도 있다[4].

$$d(I_i, I_j) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |I_i(x, y) - I_j(x, y)| \quad (\text{식 2})$$

### Color Histogram

칼라 히스토그램은 일반적으로 (식 1)번 방법이 많이 쓰이지만 (식 3)번과 같이 RGB 칼라

공간대로 따로 계산하여 사용할 수도 있다.

$$d_{RGB}(I_i, I_j) = \sum_k (|H_i^r(k) - H_j^r(k)| + |H_i^g(k) - H_j^g(k)| + |H_i^b(k) - H_j^b(k)|) \quad (\text{식 3})$$

### $\chi^2$ Histogram

이 방법은 많은 연구에서 사용되며, 칼라 히스토그램이나 template matching보다도 좋은 결과를 나타내는 방법이다[4].

$$d(I_i, I_j) = \sum_{k=1}^n \frac{(H_{i(k)} - H_{j(k)})^2}{H_{j(k)}} \quad (\text{식 4})$$

위의 방법에서 변형된 방법으로는 아래의 (식 5)가 있다.

$$d(I_i, I_j) = \sum_{k=1}^n \sqrt{H_{i(k)}^2 - H_{j(k)}^2} \quad (\text{식 5})$$

### Bin to bin difference (B2B)

$$fd_{bb}(h_1, h_2) = \frac{1}{N^2} \sum_i |h_1[i] - h_2[i]| \quad (\text{식 6})$$

## 3. 비디오 키프레임 추출

본 논문에서는 기존의 히스토그램 차이값을 계산하는 알고리즘 중에서 Color Histogram과  $\chi^2$  Histogram를 이용하여 프레임간의 차이값을 계산하는 식을 제안한다[5]. 두 알고리즘의 장점을 복합하여 변형한 (식 7)은 Color Histogram의 각각 RGB를 따로 계산하여 차이값을 좀더 세분화 하였으며,  $\chi^2$  Histogram 방법의 차이값을 강조하는 장점을 이용하여 프레임간의 차이값 계산이 쉽게 하여 분할을 용이하게 할 수 있게 한다.

이렇게 해서 얻어진 히스토그램 차이값이 임계치 값보다 크면 키 프레임으로 선택되는 것이다. 임계치를 선택하는데 있어서는 주변의 히스토그램 차이값을 참조하는 동적(dynamic) 임계치를 이용하였다.

$$d(I_i, I_j) = \sum_{k=1}^n \left( \frac{(H_i^r(k) - H_j^r(k))^2}{H_i^r(k)} \times 0.299 + \frac{(H_i^g(k) - H_j^g(k))^2}{H_i^g(k)} \times 0.587 + \frac{(H_i^b(k) - H_j^b(k))^2}{H_i^b(k)} \times 0.114 \right) / 3 \quad (\text{식 7})$$

(식 7)에서 곱해진 가중치 값은 영상을 RGB 칼라에서 명암도 등급으로 변환하기 위한 것으로 NTSC 표준에서 비롯된 것이다.

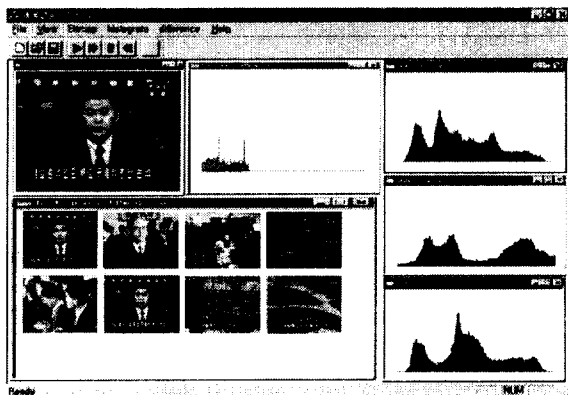
#### 4. 실험

본 논문의 실험은 Windows 98 환경에서 visual C++5.0을 이용하여 구현하였으며, 실험에 사용한 비디오 데이터들은 오스카II를 이용하여 AVI 압축 형태로 저장한 국내 KBS, MBC, SBS 방송의 뉴스 비디오와 국외 CNN, NHK의 뉴스 비디오를 <표 1>과 같이 실험 비디오로 택하였으며, 1998년 5월 29에서 31일 사이의 뉴스를 기본으로 하였다.

<표 1> 실험에 사용된 뉴스 비디오

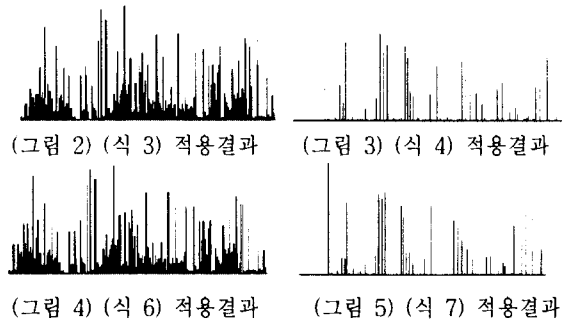
	방송사	길이	frame/sec	size
A	KBS	300초	5	400×300
B	KBS	278초	5	400×300
C	MBC	300초	5	400×300
D	MBC	210초	5	400×300
E	SBS	280초	5	400×300
F	SBS	260초	5	400×300
G	CNN	300초	5	400×300
H	NHK	200초	5	400×300

(그림 1)은 구현된 키 프레임 추출 프로그램을 보여준다. 즉, 동영상으로부터 얻어진 한 프레임에 대한 칼라 히스토그램을 보여주며, 이웃한 프레임과의 차이값을 나타낸다. 그리고 임계치를 적용하여 얻어진 키 프레임을 화면내에 작은 그림으로 나타낸다.



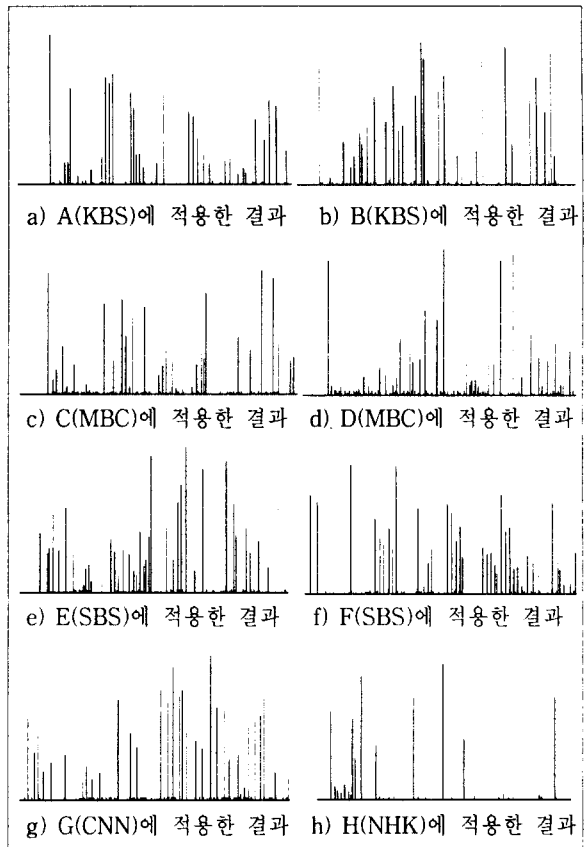
(그림 1) 키 프레임 추출의 전체화면

(그림 2)에서 (그림 5)까지는 실험 데이터 A에 적용하여 얻어진 히스토그램 값의 실례이다.



(식 3)과 (식 6)의 결과로 나타나는 히스토그램은 임계치를 설정하는데 있어 매우 민감하게 반응한다. 즉 임계치의 작은 변화에도 그 결과 값은 크게 달라질 수 있다. 그러나 (식 4)와 (식 7)을 이용한 장면변환 값은 대부분 히스토그램 위에 위치하여 임계치 선택을 쉽게 할 수 있는 장점을 가진다.

(그림 6)은 각각의 실험 데이터에 대하여 본 논문에서 제안한 식을 적용하여 얻은 히스토그램의 예를 나타낸 것이다.



(그림 6) 실험 데이터에 (식 7)을 적용한 결과

<표 2>는 각각의 입력 데이터에 따른 실험 결과를 비교한 것으로, 현재 결과는 모든 키 프레임을 찾을 수 있는 임계치를 적용할 때의 오류를 비교한 것이다.

$N_{total}$ 은 실험 데이터에서 나타난 키 프레임의 전체 수를 의미하고,  $N_f$ 는 추출되지 않아야 할 정보가 추출된 경우의 수를 의미한다.

<표 2> 데이터 A~H의 실험결과 비교

	키 프레임 추출기법	$N_{total}$	$N_f$
A (KBS)	Color Histogram	38	14
	$\chi^2$ Histogram	38	13
	Bin to bin difference	38	14
	(식 7) 적용 결과	38	10
B (KBS)	Color Histogram	33	20
	$\chi^2$ Histogram	33	13
	Bin to bin difference	33	18
	(식 7) 적용 결과	33	13
C (MBC)	Color Histogram	34	16
	$\chi^2$ Histogram	34	13
	Bin to bin difference	34	17
	(식 7) 적용 결과	34	12
D (MBC)	Color Histogram	39	19
	$\chi^2$ Histogram	39	15
	Bin to bin difference	39	21
	(식 7) 적용 결과	39	13
E (SBS)	Color Histogram	40	21
	$\chi^2$ Histogram	40	13
	Bin to bin difference	40	24
	(식 7) 적용 결과	40	13
F (SBS)	Color Histogram	40	14
	$\chi^2$ Histogram	40	8
	Bin to bin difference	40	13
	(식 7) 적용 결과	40	6
G (CNN)	Color Histogram	36	19
	$\chi^2$ Histogram	36	11
	Bin to bin difference	36	18
	(식 7) 적용 결과	36	6
H (NHK)	Color Histogram	18	6
	$\chi^2$ Histogram	18	4
	Bin to bin difference	18	5
	(식 7) 적용 결과	18	4

실험결과  $\chi^2$  Histogram와 본 논문에서 제안한 (식 7)의 경우에는 장면의 변화가 발생하는 프레임이 출현하는 경우에 히스토그램 차이값이 높게 나타났으며 반면에 변화가 심하지 않는 경우에는 낮게 나타나 임계치 선택을 쉽게 할 수 있었다.

Color Histogram과 Bin to Bin difference를 적용한 경우에는 앵커 프레임의 차이값이 다른 키 프레임의 차이값과 거의 비슷했으며 성공률이 높은 대신 키 프레임이 아닌 프레임을 키 프레임으로 추출하는 오류가 많았다.

## 5. 결론 및 향후 연구방향

본 논문은 뉴스 비디오 데이터베이스를 구축하기 위한 장면 분할 기법 중 Color Histogram에서 각각의 RGB를 따로 계산하여 차이값을 세분화하는 장점과,  $\chi^2$  Histogram에서 차이값을 강조하는 장점을 이용하여 NTSC 표준에 따른 가중치를 적용한 새로운 장면 분할 방법을 제안하였다.

본 논문에서 제안한 방법에서의 임계치 자동 선택은 기존 방법에 비해 보다 높은 자동화 할 가능성을 제시하였다. 그러나  $N_f$ 를 제거하기 위한 방법이 계속적으로 연구되어야 할 것이다.

## 참고 문헌

- [1] E. Ardizzone, M. L. Cascia, "Automatic Video Database Indexing and Retrieval," Multimedia Tools and Applications, Vol.4, Num.1, pp.29-56, January 1997.
- [2] J. C. LEE, Q. LI, W. XIONG, "VIMS : A Video Information Management System", Multimedia Tools and Applications, Vol.4, Num.1, pp.7-28, January 1997.
- [3] B. Furht, S. W. Smoliar, H. J. Zhang, "Video and Image Processing in Multimedia System," Kluwer academic Publishers, pp. 335-356, 1995.
- [4] S.J. Dennis, R. Kasturi, U.Gargi, S.Antani, "An Evaluation of Color Histogram Based Methods in Video Indexing", Research Progress Report CSE-96-053 for the contract MDA 904-95-C 2263, 1995
- [5] 이동섭, 이양원, "뉴스 비디오의 장면 분할을 위한 키 프레임 추출 기법". 제 25회 한국정보과학회 춘계학술발표회 pp.613-615, Vol.25 No 1, 1998