

# 뉴스 데이터에서 자막프레임 검출방법

남윤성\*, 배종식\*\*, 최형진\*

\*강원대학교 컴퓨터학과

\*\*한림정보산업대학교 전산정보처리과

e-mail:nyspeter@dreamwiz.com

## Detecting Method of Video Caption Frame on News Data

Yun-Seong Nam\*, Jong-sik Bae\*\*, Hyung-Jin Choi\*

\*Dept. of Computer Science, Kang-Won University

\*\*Dept. of Computer Aided Information Processing, Hallym College

### 요 약

디지털 영상 자료가 대중화됨에 따라 방대한 양의 자료를 효과적으로 이용 및 검색하기 하기 위해 영상 데이터의 색인과정이 필수적이다. 뉴스 데이터에서 자막 프레임은 뉴스의 내용을 한 눈에 파악할 수 있는 중요한 정보이다. 따라서 본 논문에서는 뉴스 데이터에서 색인과정을 위해 우선 자막 프레임을 검출하는 기법을 제안하고자 한다. 자막이 있는 프레임을 검출하기 위해 가변길이 프레임 생략법을 이용하여 키프레임을 검출한다. 영상보정을 위한 전처리 작업으로 BC(Brightness & Contrast) 필터기법을 제안하고 자막영역을 대상으로 IT(Invers & Threshold) 기법을 적용하여 자막프레임을 검출하는 방법을 제안한다.

### 1. 서론

정보의 홍수속에 살아가고 있는 현재, 원하는 정보를 찾아서 가공하는 일은 매우 중요한 일로 대두되고 있다. 매체가 점차 영상을 포함하는 멀티미디어 데이터로 변경되고, 디지털화되어 저장, 전송 및 처리되고 있다. 따라서 방대한 디지털 영상을 어떻게 효율적으로 전송, 저장, 처리하는가 하는 문제가 대두되고 있으며, 최근에는 수 많은 영상을 활용하는 측면을 생각하게 되었다. 즉 수 많은 영상을 데이터베이스화하여 미래에 다시 사용할 수 있도록 하려고 하는 것이다.

이미 인터넷 방송국과 같이 디지털화 된 영상 데이터를 생산해내는 업체가 상당수에 이르며 각 대학 방송국에서도 새 시대에 발맞추어 디지털 방송으로 방향을 전환하고 있다.

본 논문에서는 자막프레임이 중요한 영향을 미치는 뉴스데이터에서 자막프레임을 검출하여 보다 효

율적인 색인화 및 검색을 하도록 한다. 자막이 많이 나오는 뉴스데이터에서 가변길이 프레임 생략법을 이용해 키프레임을 검출한 후 전처리 과정으로 BC(Brightness & Contrast) 필터기법을 통해 프레임을 강화한 후 자막이 빈번하게 발생하는 자막영역을 추출해 IT(Inverse & Threshold)기법으로 자막의 여부를 판단하고 자막이 있는 장면을 검출하는 방법을 제안한다.

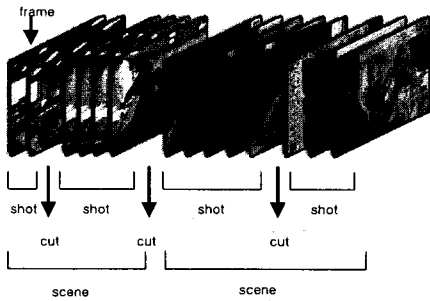
### 2. 관련 연구

#### 2. 1 영상의 구조

비디오나 영상은 제작 단계에서 물리적인 샷(shot)과 논리적인 장면(scene)으로 구성된다. 영상의 구조는 크게 4종류의 계층적 모임으로 구성되는데, 최하위 계층은 개개의 정지영상을 가르키는 프레임(frame)이고, 한 카메라의 연속적인 촬영에 의한 프레임의 모임을 샷이라 한다. 그리고 내용상 관련이 있는 인접한 샷의 집합을 장면이라 하고, 한 주제에 대한 연속적인 장면들의 집합을 영상의 최상위 계층인 시퀀스(sequence)라 한다. 샷이란 촬영

※ 이 연구는 강원대학교 기초 과학연구소의 연구비 지원으로 수행되었습니다.

시에 멈춤 없이 한번에 기록한 연속적인 동작을 의미하고 장면은 논리적으로 서로 밀접한 관련이 있는 연속적인 샷의 묶음(cluster)이다[1].

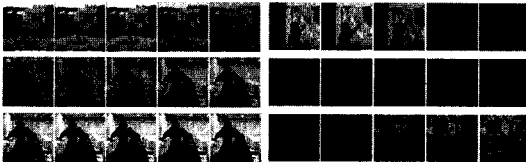


(그림 1) 영상의 구조

### 2. 2 장면 전환 검출

장면 전환의 검출은 비디오 분할 작업에 해당하며, 내용 기반 검색(content-based retrieval) 및 대표 영상(representative image) 추출에서 반드시 필요하고 중요한 기술이다[2].

장면 전환이란 연속적인 영상 이미지의 어떤 한 장면에서 다른 장면으로 전이하는 것을 일컫는 말로 전이되는 특성에 따라 통상 컷(cut), 페이드(fade), 디졸브(dissolve), 와이프(wipe), 등으로 분류한다.



(a) 디졸브 (b) 페이드  
(그림 2) 장면전환의 종류

#### 1) 픽셀 차이를 이용한 검출

픽셀 단위의 검출 방법은 연속된 두 프레임에서 상응하는 픽셀 사이의 특징차를 구하여, 그 차이가 임계값 이상이 되면 컷으로 검출하는 방법으로 주로 픽셀의 농도값을 특징으로 사용한다. 연속된 프레임에서 같은 좌표에 위치한 픽셀간 명암값의 차이가 일정크기보다 큰 픽셀들의 개수가 임계치 이상인 프레임을 컷으로 검출한다[3].

$$D_i(x, y) = \begin{cases} 1 & |F_{i+1}(x, y) - F_i(x, y)| > T \\ 0 & \text{otherwise} \end{cases}$$

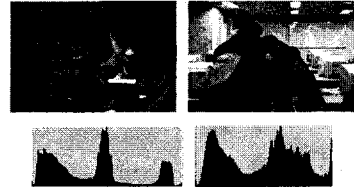
$$\frac{\sum_{x, y=1}^{M \cdot N} D_i(x, y)}{M \cdot N} \times 100 > T \quad (\text{식 1})$$

카메라의 움직임과 물체의 움직임에 매우 민감하다는 단점을 가지고 있어 컷 검출에 어려움이 있다.

#### 2) 히스토그램을 이용한 방법

히스토그램 비교법은 비디오 영상에서 이웃하는 프레임들 사이의 히스토그램 차로서 장면 전환의 발생 여부를 판단하는 방법이다[4]. 히스토그램 비교법의 장점은 픽셀 비교법보다 카메라 또는 물체의 움직임에 덜 민감하여 장면전환 검출에 정확성과 속도 면에서 좋은 결과를 가져올 수 있는 방법이다.

$$D_i(x, y) = \sum_{j=1}^G |H_i(j) - H_{i-1}(j)| > T \quad (\text{식 2})$$



(그림 3) 히스토그램 차이

### 2. 3 영상 텍스트

비디오 영상에 존재하는 텍스트는 다음과 같이 두 가지 종류이다.

- 씬 텍스트(scene text) : 카메라에 찍혀 나타나는 텍스트로서 비디오 영상에서 거리에 있는 간판이나 도로의 번호판과 같은 것을 말한다.
- 편집 텍스트(editing text) : 특정한 내용을 설명하기 위하여 인위적으로 화면에 삽입하는 텍스트이다. 뉴스의 자막이나 비디오의 자막 등이 이에 해당된다.

뉴스에서 자막이 가지는 의미는 이미 앞에서 말한 바와 같이 매우 중요하다. 그리고 자막과 함께 나오는 장면 또한 자막 못지 않게 상당히 중요하다. 특히 뉴스 데이터에서 앵커오른쪽 위에 사진과 자막은 뉴스내용을 한눈에 파악 할 수 있는 매우 중요한 프레임이다.

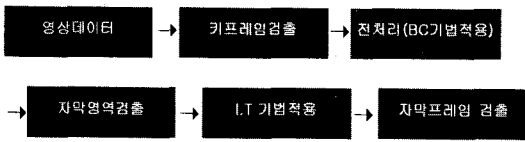


(그림 3) 앵커와 같이 나오는 자막 프레임

영상에 삽입되는 텍스트들의 특징은 대부분 단색이고, 크기나 모양의 변화가 거의 없으며, 제한된 크기를 가지며, 배경과 대조를 이루고, 배경에 상관없이 다수의 프레임에 나타나며, 윤곽선 또는 경계를 가지고 있고, 자막의 위치가 대부분 정해져 있다

### 3. 자막 프레임 검출방법 제안

자막 프레임 검출방법을 각 단계별로 설명하며, 전체적인 흐름도는 (그림 4)와 같다.



(그림 4) 자막프레임 검출과정

#### 3. 1 키프레임 추출

키프레임 추출은 비디오에서 장면 전환을 검출하는 방법이다. 키 프레임을 검출하는 이유는 자막이 있는 프레임은 극히 일부이기 때문에 모든 프레임을 검출하여 비교한다는 것은 비효율적이다. 최근 압축기술이 눈부시게 발전하고 저장매체로서 DVD가 좋은 선두적 역할을 하고 있지만 영상의 경우 아무리 좋은 압축을 하더라도 압축을 통한 손상되는 부분은 원본으로서의 역할을 하지 못한다. 영상을 다시 사용할 목적이려면 압축된 영상은 재사용하기에는 영상의 질이 떨어진다. 따라서 본 논문에서는 압축되지 않은 비디오 데이터에서 검출 방법을 제안한다.

본 논문에서는 뉴스영상데이터를 중심으로 하기 때문에 뉴스데이터의 특성상 점진적인 장면전환은 거의 사용하지 않으며 주로 급진적인 컷을 사용하기 때문에 장면전환검출로는 히스토그램 비교법을 사용하는 것이 시간적인 면과 성능면에서 좋은 결과를 얻을 수 있다. 따라서 본 논문에서는 정확도와 속도를 고려하여 가장 적당한 방법으로 가변길이 프레임 생략법을 이용한 검출 방법을 적용하였다.

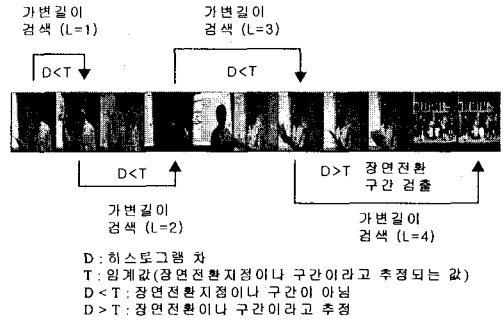
가변길이 프레임 생략법은 가중치를 조절하여 장면전환이 자주 일어나는 경우 생략하는 프레임 수를 줄이고 장면전환이 거의 발생하지 않으면 생략하는 프레임수를 많이 해서 히스토그램 비교횟수를 줄이는 방법이다[5].

$$Length = W \cdot Min\_Unit \cdot Frame\_per\_sec \quad (식 3)$$

- 단, Length: 생략하는프레임수
- Min\_Unit: 장면전환 검출이 가능한 최소 프레임 수
- Frame\_per\_sec: 초당 재생 프레임 수
- W: 가중치
- 장면전환 미검출:  $W = W + 1$
- 장면전환 검출:  $W = 1/2$

Min\_Unit과 Frame\_per\_sec 값을 각각 1로 두어 가변길이 초기값을 1로 하여 프레임을 생략하며 히

스토그램 차(D)를 비교한다. 임계값(T)보다 히스토그램 차(D)가 작으면 가중치(W)를 증가시켜 생략할 프레임수를 증가하여 히스토그램 차를 비교한다.  $D > T$  인 영역이 발견되면 장면전환 지점이 있으므로 예상하여 지정된 프레임을 이진 검색을 이용하여 장면전환 지점으로 검출한다.



(그림 5) 가변길이 프레임 생략법

#### 3. 2 전처리 과정

키 프레임으로 추출된 영상에서 자막의 유무를 검출한다. 영상 데이터의 특성상 영상에서 이미지를 추출할 경우 이미지는 스캔받은 이미지에 비교하여 해상도가 상당히 떨어진다. 그러므로 영상강화작업으로 전처리가 필요하다. 본 논문에서는 좀 더 선명한 자막프레임을 검출하기 위하여 키 프레임으로 검출된 비디오 이미지를 전처리로 BC 필터기법을 적용한다.

BC 필터기법이란 먼저 영상의 전체적인 밝기를 조절하는 밝기교정(brightness correction)과 대비(contrast)를 높여서 자막의 특성을 보다 선명하게 나타나도록 하는 것이다.

#### 3. 3 자막프레임 검출

뉴스자막의 특성상 자막이 나오는 영역은 대부분 지정되어 있다. 전처리 과정 후 뉴스에서 자주 나타나는 자막영역을 임의로 지정하여 이 영역만을 대상으로 IT 기법을 적용하여 자막프레임을 검출한다.



(그림 6) 자막영역

IT 기법이란 먼저 전처리 작업 된 이미지에 대해서 반전(inverse)기법으로 자막과 배경이 분리되도록

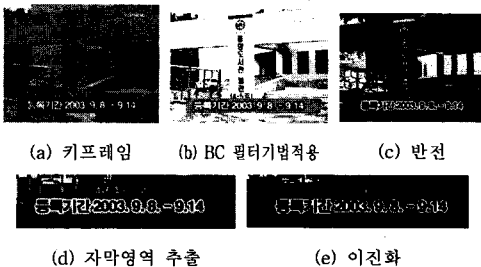
한 후 이진화(threshold)기법으로 자막의 특성이 보다 선명하게 출력되도록 하여 자막프레임을 검출하는 방법이다.

반전기법이란 미리 지정된 함수를 기반으로 하여 이전 화소를 역으로 변환하는 처리이다. 프레임의 명도가 반전되어 검은색은 흰색으로 흰색은 검은색으로 만들어 색상은 보색으로 만들어 진다.

이진화기법은 프레임을 2가지 색으로 단순화 한다. 명도단계인 0 - 255에서 임계치를 기준으로 그보다 높은 픽셀을 흰색, 낮으면 검은색으로 표현하여 순흰색과 순검정점으로만 프레임을 만들어 낸다. 장면전환점이라고 판단되는 키프레임과 키프레임을 기점으로 15번째 프레임과 30번째 프레임을 검출한 후 3장의 프레임을 비교한다. 이것은 자막의 특성상 몇 초 동안 지속된다는 특성을 이용한 것이다. 인접한 세 프레임에서 똑같은 경계영역이 존재하는지를 확인하고, 앞선 프레임에서의 경계 영역이 다음 프레임의 경계 영역과 70%이상 겹치면 이를 자막 프레임으로 검출한다.

4. 실험 및 결과

본 논문은 강원대학교 방송국(KUBS)에서 제작한 뉴스영상 5회분(24분 20초)을 가지고 실험을 하였다. 실험 장비로는 편집용 컴퓨터 IBM6274-56K와 편집보드 RTX-100를 사용하여 실험하였다. 이진화 임계치는 200을 지정하였고, 그 이유는 뉴스데이터에서 삽입한 자막의 경우는 대부분 대비(Contrast)가 높기 때문이다. 뉴스 데이터 중 약 80%의 검출율을 보였다. 검출하지 못한 프레임은 배경영상과 차이가 나지 않는 자막의 경우와 그래픽 처리가 이루어진 자막프레임의 경우에 검출을 하지 못했다.



(그림 7) 자막프레임 검출과정

5. 결론 및 향후 연구 방향

본 논문에서는 효율적인 영상데이터의 색인작업을

위해 뉴스 데이터에서 자막 프레임을 검출하는 방법을 제시하였다. 키 프레임 추출시 가변길이 프레임 생략법을 이용함으로써 속도를 빠르게 하였고, 영상 강화를 위해 BC 필터기법을 적용하였으며 자막프레임 검출을 위하여 뉴스에서 자주 등장하는 자막영역을 대상으로 IT 기법을 적용하여 자막프레임을 검출하였다.

본 연구에서는 자막이 없는 경우 중요한 장면일지라도 검출을 하지 않기 때문에 자막프레임과 더불어 영상에 중요한 장면이 되는 프레임을 검출하는 방법에 대한 연구가 필요하다. 또한 최근 3D로 작업이 된 자막의 경우와 흐르는 자막의 경우 등 특수한 작업이 이루어진 영상의 경우에는 검출을 못하는 문제점이 발생하였다. 향후 연구과제로는 이러한 자막 프레임의 검출 연구와 검출된 데이터를 효과적으로 색인화 하는 연구가 필요하다.

참고문헌

[1] B. Yeo B Liu, "Raped Secne Analysis on Compressed Videos", IEEE TCSVT II, vol. 5, No. 6, pp. 533-544, 1995.

[2] C. Berrut and M. Mechkour, "Representation of Images for Multimedia: A Preliminary study," Proc. IFIP WG 2.6 2nd Working Conf. on Visual Database System, IFIP, Elsevier Science Publishers, Amsterdam, pp. 317-326, 1991.

[3] H.Zhang, A. Kankanhalli and S. W. Smoliar, "Aotomatic Partitioning od Full-motion Video", Multimedia System, Vol. 1, No. 1, pp. 10 - 28, 1993.

[4] John S. Boreczky and Lawrence A. Rowe, "Comparison of Video Shot Boundary Detction Techniques, "Storage and Retrieval Image and Video Datavase VI, Proc. of IS&T/SPIE 1996 Symp. on Elec. Imaging: Science and Techonlogy, February 1996.

[5] 권오덕, 최창규, 장용석, 김승호, "가변길이 프레임 생략 방법을 이용한 장면 전환 검출", 한국 정보 처리 학회 가을 학술발표집 vol. 28, No. 2, pp. 268-270.