

비디오 자막 영역 원영상 복원 후 자막 교환 방법

전병태, 한규서, 배영래
한국전자통신연구원 컴퓨터·소프트웨어연구소
영상정보처리연구팀
E-mail : chunbt@etri.re.kr

A Method for Recovering Image Data for Caption Regions and Replacing Caption Text

Byung-Tae Chun, Kyuseo Han, Younglae Bae
Image Information Processing Team
Computer Software Technology Laboratory
ETRI(Electronics and Telecommunications Research Institute)

요 약

멀티미디어 정보 중 비디오 데이터는 정보량이 많기 때문에 자동화된 비디오 영상 처리 기술이 필요하다. 시청자의 이해와 시청의 편의성을 제공하기 위하여 대부분 비디오에 자막을 삽입하게 된다. 외국 방송물 및 영화에 삽입된 다른 언어 자막을 편집 과정에서 자막으로 교환 할 필요성이 종종 있게 된다.

기존의 방법들은 자막 부분을 충분히 포함하는 일정 영역에 특정 색상을 채운 후 새로운 자막을 삽입하게 된다. 기존 방법의 문제점은 많은 영역에 대해 비디오 영상 정보를 손실 시킴으로써 시청자에게 시청의 불편을 초래하고 자막 교환의 비 효율성과 부 자연스러움을 발생시킨다.

본 논문에서는 기존 방법의 문제점을 극복하기 위하여 자막 영역을 원영상으로 복원한 후 다른 자막으로 교환하는 방법을 제안하고자 한다. 원영상 복원을 위하여 비디오 정보와 *BMA(Block Matching Algorithm)*를 이용한 원영상 복원 방법을 제안하고, 복원된 영역에 다른 자막으로 교환함으로써 효과적인 자막 교환 방법을 제안하고자 한다.

실험 결과 원영상 복원을 이용한 자막 교환 방법은 기존 방법에 비해 자연스럽게 효과적인 교환 방법임을 볼 수 있었다.

1. 서 론

오늘날 멀티미디어 산업의 발달은 대용량 비디오 정보 처리 기술[1][2][3] 즉 비디오 영상 처리 기술, 저장 기술 및 검색 기술의 발전을 진일보 시켰다고 볼 수 있다. 비디오 매체는 방송, 교육, 가정, 영화 등과 같이 많은 분야에 사용되어 지고 있다.

비디오 자막은 사용자에게 비디오 내용에 관한 이해를 돕기 위하여 문자화된 형태로 비디오에 삽입된다. 삽입된 비디오 자막은 비디오에 관한 의미 정보를 함축적으로 나타내거나 비디오 내용에 대한 의미 전달 수단으로 사용되고 있다. 다른 언어 자막이 삽입된 외국 방송물 및 영화를 자막 교환 후 방송할 필요성이 종종 있게 된다.

기존의 자막 교환 방법은 자막 부분을 포함하는 일정 영역에 특정 색상을 채운 후 새로운 자막을 삽입하게 된다. 기존 방법의 문제점은 자막을 포함하는 일정 영역의 영상 정보를 손실 시킴으로써 시청자에게 시청의 불편과 자막 교환의 비 효율성 및 부 자연스러움을 초래한다.

본 논문에서는 기존 방법의 문제점을 극복하기 위하여 자막 영역을 원영상으로 복원한 후 다른 자막으로 교환하는 방법을 제안하고자 한다. 원영상 복원을 위해서는 비디오 정보와 *BMA*를 이용하여 자막 영역을 원영상으로 복원하는 방법을 본 논문에서 제안하고자 한다. 그리고 복원된 결과를 이용하여 자막 영역에 다른 자막을 삽입함으로써 효과적인 자막 교환 방법을 제안하고자 한다.

2. 자막 교환 시스템의 개요

자막 영역 원영상 복원 및 교환 방법의 시스템 개요는 그림 1에서 보여주고 있다. 먼저 효과적인 자막 교환을 위하여 자막 영역의 원영상 복원을 수행한다. 복원에 필요한 사전 정보(a priori-knowledge) 즉 자막 영역 정보, 장면 전환 정보, 복원의 방향성 정보를 추출하고, 추출된 비디오 사전 정보를 이용하여 BMA 방법[4]을 적용 자막 영역을 원영상으로 복원한다. 복원된 영상에 자막기를 이용하여 다른 자막으로 교환 함으로서 효과적이고 자연스러운 자막 교환을 수행할 수 있다.



그림 2. 추출된 비디오 자막 영역



그림 3. 문자의 구성 요소 정보

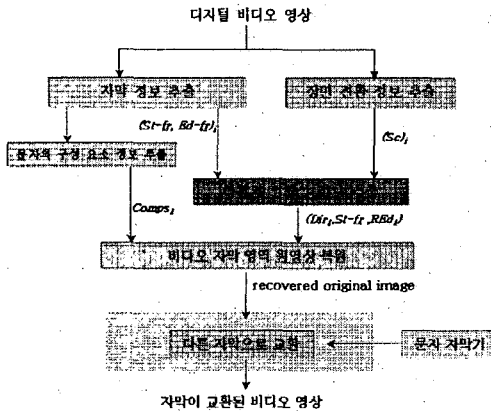


그림 1. 자막 영역 원영상 복원 및 교환 시스템

(2) 비디오 장면 전환 검출

장면 전환 정보는 복원의 방향성 결정에 중요한 정보가 된다. 효과적인 장면 전환 지점을 검출하기 위하여 본 논문에서는 모션 벡터를 이용한 장면 전환 검출 방법[7] 이용하였다. 모션 벡터(Mv)는 크기(M)와 방향성(D)으로 구성된다. 크기와 방향성 정보를 히스토그램으로 변환하고 각 프레임의 히스토그램 차로부터 얻어지는 교차율을 계산하여 장면 전환 지점을 검출한다. 이 방법을 히스토그램 비교법이라고 하며 장면 전환 지점의 교차율을 구하기 위해 식(2)을 이용한다.

$$\sum_j^g \frac{\min(H_i(j), H_{i+1}(j))}{\max(H_i(j), H_{i+1}(j))} \dots\dots(2)$$

여기서, H_i(j)는 j-th 번째 방향에 있는 i-th 프레임의 주파수

또한, 2 개이 연속된 프레임의 상관(correlation)도를 계산하여 장면 전환 지점을 추출한다. Pearson's 의 선형 상관계수는 식(3)을 이용하여 구한다. 상관계수의 범위는 -1 ≤ r ≤ 1이며, 1 또는 -1에 근접할수록 상호 프레임은 유사하며 r이 0에 가까울수록 프레임을 서로 다르다. 이와 같은 상관계수는 프레임에 상당히 큰 물체가 존재할 때 장면 전환 지점을 추출하는데 유용하다.

$$r = \frac{n \sum_{j=1}^n \bar{D}_i(j) \bar{D}_{i+1}(j) - \sum_{j=1}^n \bar{D}_i(j) \sum_{j=1}^n \bar{D}_{i+1}(j)}{\sqrt{\left[n \sum_{j=1}^n \bar{D}_i(j)^2 - \left(\sum_{j=1}^n \bar{D}_i(j) \right)^2 \right] \left[n \sum_{j=1}^n \bar{D}_{i+1}(j)^2 - \left(\sum_{j=1}^n \bar{D}_{i+1}(j) \right)^2 \right]}} \dots\dots(3)$$

점진적 장면전환 및 특수 효과에 의한 장면 전환 지점을 추출하기 위하여 식(2) 식(3)을 사용한다.

(3) 복원의 방향성 정보 추출

복원의 방향성 정보를 추출하는 이유는 자막의 시작되는 전 프레임이나 자막의 끝나는 다음 프레임은 자막이 삽입되지 않은 원 영상을 갖고있기 때문에 이 영상을 원 영상 복원에 중요한 정보로 이용하려는 것이다. 복원의 방향성은 추출된 장면 전환 정보에 따라 결정되며, 어느 방향으로 어디까지 복원을 진행할 것인가를 결정한다.

3. 비디오 자막 영역 원영상 복원 방법

3.1 비디오 정보 추출

(1) 비디오 자막 및 구성 요소 정보 추출

자막 추출 방법은 지형학적 특징점(topographical features)을 이용한 자막 영역 추출 방법[5][6]을 사용하였다. 이 방법은 추출된 문자의 지형학적 특징점을 점-선-면 확장법을 이용하여 문자 영역을 추출함으로써 기존 문자 영역 추출의 문제점인 휴우리스틱 변수의 사용을 최소화하고 임계치 값을 일반화함으로써 일반화된 문자 영역 추출 방법이다. 지형학적 특징점(TF_{ij})은 식(1)을 이용하여 구한다.

$$TF_{ij} = \begin{cases} 1 & : \text{if } ((P_y (= A_1) > A_2 + \alpha_{-th}) \& \& ((A_2 > A_3 + \alpha_{-th}) \& \dots \\ & \dots\dots(A_{n-1} > A_n + \alpha_{-th})) \\ 0 & : \text{otherwise} \end{cases} \dots\dots(1)$$

추출된 지형학적 특징점은 근접성을 고려하여 선으로 연결하고, 연결된 선은 다시 면 영역을 구한다. 여러 단계에서 구해진 면 영역은 중첩성을 조사하여 최종 자막 후보 영역을 추출한다. 추출된 자막 후보 영역은 검증을 통하여 자막 영역을 추출한다. 추출된 자막 영역은 그림 2에서 보여주고 있으며, 각 문자의 구성 요소 정보는 그림 3에서 보여주고 있다.

추출된 장면 전환 정보와 자막 정보를 이용하여 복원의 방향성과 복원의 중점을 그림 4와 같이 3가지로 분류할 수 있다.

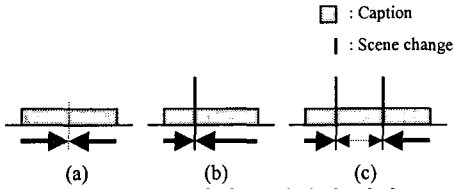


그림 4. 복원의 방향성 결정

3.2 BMA 을 이용한 자막 영역 원영상 복원

문자 영역을 복원하기 위해서 복원 하고자 하는 자막 영역이 원 영상의 어느 부분과 밀접한 관계가 있는지를 찾아내야 한다. 원 영상 영역을 찾기 위하여 블럭 정합 방법을 사용한다. 블럭 정합 방법은 동영상 압축 및 전송에 많이 사용되고 있는 방법으로 현재 프레임과 참조 프레임에서 최대 유사도를 찾기 위한 방법으로 사용된다. 블럭 정합의 기준은 최소 절대차(MAD : Minimum Absolute Difference)이며 식(4)와 같다.

$$MAD(d_1, d_2) = \sum_{(x,y) \in R} |F(x, y, t) - F(x+d_1, y+d_2, t-1)| \quad ..(4)$$

여기서 R 은 문자 구성요소 단위의 문자의 경계 화소, |d1|, |d2| 값은 탐색 범위 보다 작은 값이고 F(x,y,t) 는 시간 t 프레임에 있는 좌표(x,y)의 화소를 말한다.

블럭 정합 방법에 의한 자막 영역의 원 영상 복원 방법은 비교될 자막 영역은 문자 단위가 아니라 문자의 구성 요소 단위이다. 또한 문자의 구성 요소 영역 전체에 대하여 블럭 정합을 수행하는 것이 아니라 문자의 윤곽선을 포함하는 외곽선 화소에 대하여 블럭 매칭을 수행한다. 최적의 위치가 찾지면 그 영역의 원영상 정보를 이용하여 원영상으로 복원한다.

4. 원영상 복원 후 자막 교환

복원된 비디오 영상에 교환하고자 하는 다른 언어를 문자 자막기를 이용하여 그림 5와 같이 문자를 삽입한다.

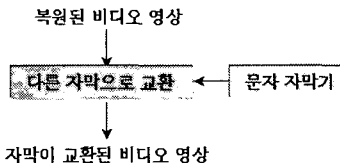


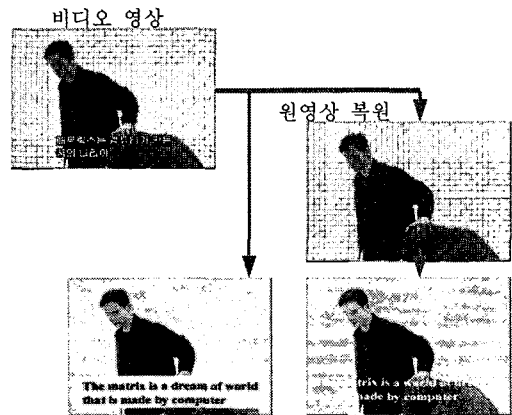
그림 5. 문자 자막기를 이용한 자막 교환

5. 실험 결과 및 결론

실험 환경은 Windows NT 상에서 750 MHz 속도를 갖는 PC로 구현하였다. Visual C++6.0 언어를 사용하였으며 비 압축 비디오 영상을 실험 영상으로 사용하였다. 비디오 영상은 RT5 Encoding 보드를 사용하여

file.mpg 파일을 획득한 후, 소프트웨어적으로 Decoding 하여 실험에 사용하였다. 그림 6은 한글 자막을 영어 자막으로 교환하는 과정을 보여주고 있다. 그림 6(a)는 기존의 자막 교환 방법을 보여주고 있으며 그림 6(b)는 제안된 자막 교환 방법을 보여 주고 있다.

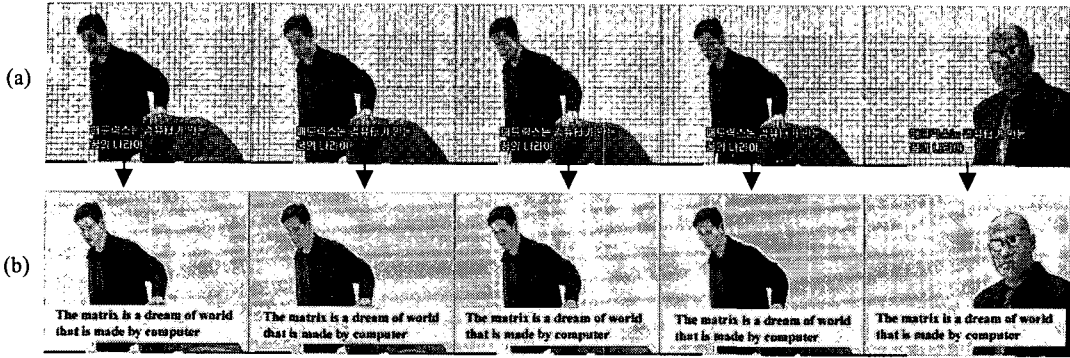
실험 결과 원영상 복원 방법을 이용한 자막 교환 방법은 기존 방법에 비해 자연스럽게 효과적인 교환 방법임을 볼 수 있었다. 그림 7은 기존 방법, 그림 8은 제안된 방법의 수행 결과를 나타내고 있다. 기존 방법과 제안된 방법을 비교 결과는 그림 9에서 보여 주고 있다.



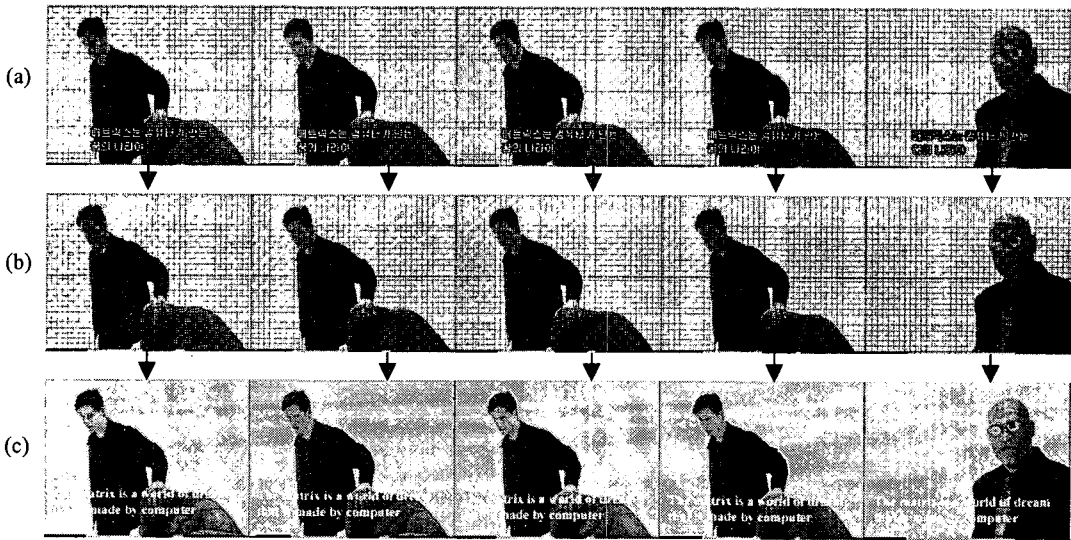
(a) 기존 방법 (b) 제안된 방법
그림 6. 기존 및 제안된 자막 교환 방법

참고문헌

- [1] Smoliar, S and Zhang, H., "Content-based video indexing and retrieval," IEEE Multimedia, summer, pp.63-72, 1994.
- [2] ITU-T Recommendation H.263, Video Coding for Low Bit Rate Communication, Mar. 1996.
- [3] H.Reeve and J.S.Lim, "Reduction of blocking effects in image coding," Optical Engineering, Vol.23, No.1, pp.34-37, Jan., 1984
- [4] Joan L. Mitchell, William B. Pennebaker, Chad E. Fogg, and Didier J. LeGall, MPEG video compression standard, Chapman & Hall, 1996.
- [5] 전병태, 배영래, 김태운, "일반화된 문자 및 비디오 자막 영역 추출 방법, 정보과학회논문지: 소프트웨어 및 응용, Vol. 27, No.6, pp.632-641, 6. 2000.
- [6] Byung Tae Chun, Younglae Bae, Tai-Yun Kim, "Text extraction in videos using topographical features of characters," FUZZ-IEEE'99, Vol.II, pp.1126-1130, Korea, 8. 1999
- [7] Ok-Bae Chang, Myung-Sup Yang, Jae-Hyun Lee, "Segmentation of gradual scene transitions using motion vector", IS&T/SPIE, Vol. 3422, pp.187-198, July. 1998



(a) 디지털 비디오 영상
 (b) 기존의 자막 교환 방법
 그림 7. 기존의 자막 교환 방법 및 결과



(a) 디지털 비디오 영상
 (b) 자막 영역 원영상 복원
 (c) 제안된 자막 교환 방법
 그림 8. 제안된 자막 교환 방법 및 그 결과

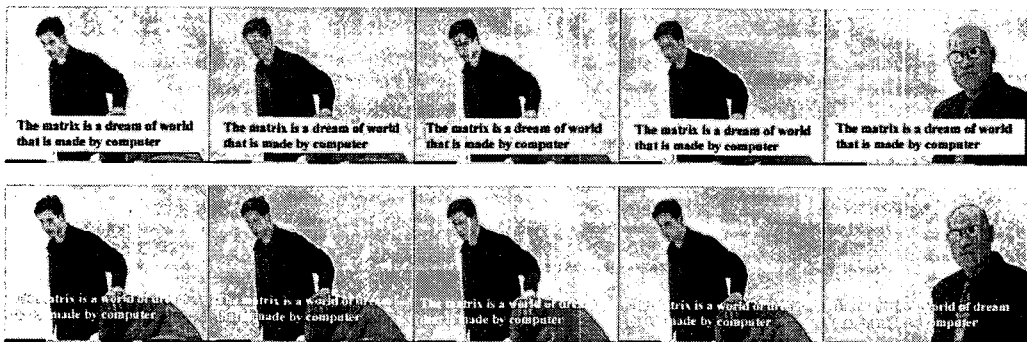


그림 9. 기존 방법과 제안된 방법의 비교