뉴스 비디오 인덱싱의 자동화

고경철, 신성윤, 이동섭, 이양원
군산대학교 컴퓨터 과학과
road@one.kunsan.ac.kr , syshin@cs.kunsan.ac.kr ,
leeds@ai.kunsan.ac.kr , ywrhee@ai.kunsan.ac.kr

Automation of News Video Indexing

Kyong-Cheol Ko, Seong-Yoon Shin, Dong-Seop Lee, Yang-Weon Lee
Dept. of Computer Science, Kunsan National Univ.
road@one.kunsan.ac.kr , syshin@cs.kunsan.ac.kr ,
leeds@ai.kunsan.ac.kr , ywrhee@ai.kunsan.ac.kr

요 약

뉴스 비디오 데이터베이스를 구축하기 위해서는 비디오 색인의 자동화에 관한 연구가 필연적이다. 그리하여, 본 논문에서는 비디오 색인화를 자동화 하는 방법으로, 이전에 제안한 키 프레임(key frame)의 추출 방법을 이용하여 추출된 키 프레임을 대상으로 색인화 하였다. 즉, 추출된 키 프레임 중에서 메타 색인에 해당하는 앵커 프레임(anchor frame)을 선택하는 방법을 제시하고, 선택된 앵커 프레임을 이용하여 뉴스 비디오를 색인화 하였다. 구현된 비디오 색인화는 사용자의 관점에 따라 여러 가지 방법으로 브라우징(browsing)되며, 사용자가 원하는 뉴스 비디오 씬(scene)들을 쉽게 선택하여 볼 수 있는 특징을 제공하였다.

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.

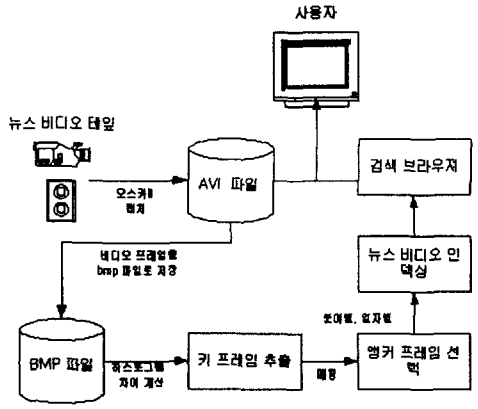
이러한 동영상 자료 중에서도 뉴스 비디오는 사회 전반의 중요한 정보를 함축적으로 짧은 시간대에 표현하는 것으로, 그 시대를 분석하는 귀중한 자료가 되고 구조적인 형태를 가지므로 내용 기반에 의한 비디오 검색의 주요 연구 대상이 되어 왔다.

본 논문에서는 비디오 색인화를 자동화 하는 방법으로, 이전에 제안한 키 프레임(key frame)의 추출 방법을 이용하여 추출된 키 프레임을 대상으로 색인화 하였다. 즉, 추출된 키 프레임 중에서 메타 색인에 해당하는 앵커 프레임(anchor frame)을 선택하는 방법을 제시하고, 선택된 앵커 프레임을 이용하여 뉴스 비디오를 색인화 하였다. 구현된 비디오 색인화는 사용자의 관점에 따라 여러 가지 방법으로 브라우징(browsing)되며, 사용자가 원하는 뉴스 비디오 씬(scene)들을 쉽게 선택하여 볼 수 있는 특징을 제공하였다.

본 논문의 구성은 다음과 같다. 2장에는 뉴스 비디오를 색인화하기 위한 시스템의 구성도에 대해 설명 하고, 3장에서는 [5]에서 제안한 키 프레임 추출 기법에 대해서 설명한다. 4장에서는 추출된 키 프레임을 이용하여 색인화 하는 과정에 대해서 설명하고, 5장에서는 현재의 뉴스 도메인을 실험 도메인으로 정하고, 제안한 방법을 적용한 후의 실험 결과에 대해서 분석하였다. 6장에서는 본 논문의 결론 및 향후 연구 방향에 대하여 설명하였다.

2. 제안된 모델

본 논문에서는 뉴스 비디오를 색인화하기 위하여 그림 1과 같은 모델을 제안하였다.



(그림 1) 개략적인 시스템 구성도

입력되는 뉴스 비디오는 우선적으로 각 장면이 분할되어야 한다. 각 장면의 분할은 키 프레임을 추출하고, 키 프레임에서 사전지식을 이용하여 앵커 프레임을 추출하는 두 단계로 구성된다. 선택된 앵커 프레임과 키 프레임을 이용하여 뉴스 비디오의 색인화 작업을 한다. 그리고 사용자는 검색 브라우저를 통해 원하는 뉴스 스토리를 볼수 있다.

3. 비디오 키프레임 추출

[5]에서는 기존의 히스토그램 차이값을 계산하는 알고리즘 중에서 Color Histogram과 χ^2 Histogram를 이용하여 프레임간의 차이값을 계산하는 식을 제안한다. 두 알고리즘의 장점을 복합하여 변형한 (식 1)은 Color Histogram의 각각 RGB를 따로 계산하여 차이값을 좀더 세분화 하였으며, χ^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.

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

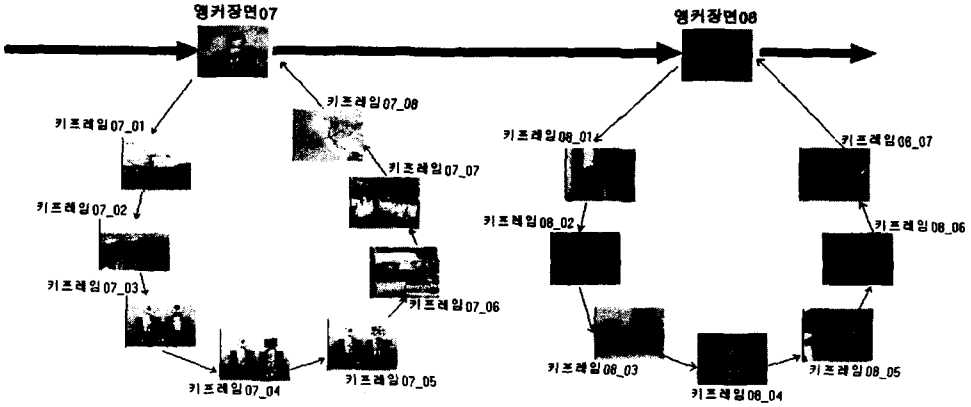
(식 1)에서 곱해진 웨이트 값은 영상을 RGB 칼라에서 명암도 등급으로 변환하기 위한 것으로 NTSC 표준에서 비롯된 것이다. RGB 칼라 공간은 서로 가산될 수 있는 삼원색으로 구성되어 컬러 분광 정도에 따라 색이 결정되는 것이다. 특히 초록은 명암도 등급 중에서 상당히 큰 요소로, 간혹 초록 요소만을 명암도 단계 데이터로 사용하기도 한다.

4. 비디오 색인화

뉴스 비디오에서 분할된 첫 프레임들 중에는 앵커 장면이 존재한다. 앵커 장면을 검출하는 이유는 뉴스가 구조적인 특징으로 이루어져 있어 사건별로 색인화 할 때 앵커 장면이 그 사건의 시작 장면이 되기 때문이다. 그래서 앵커 장면만으로도 색인화가 가능하다. 그리고, 앵커 장면은 앵커와 뉴스를 가장 함축적으로 표현하는 뉴스 아이콘이 존재하기 때문이다. 추출하는 방법으로는 키 프레임들을 추출하는 도중 앵커 프레임이 키 프레임으로 추출되면 그 이후에 추출되는 키 프레임들의 대표 프레임으로 선택하여 저장한다.

키 프레임들 중에서 앵커 장면을 추출하게 되면, 다음 앵커 장면이 나타나기 전까지의 모든 키 프레임들을 이전의 앵커 프레임에 종속시킨다. 그래서 하나의 뉴스 스토리는 앵커 장면이 나온 후에 다음 앵커 장면이 나올 때까지이며, 키 프레임들도 마찬가지로 다음 키 프레임이 나오기 전까지의 이웃 프레임들을 묶어 하나의 씬(scene)으로 구성한다.

아래의 (그림 2)에서 나타나듯이 키 프레임들은 추출된 순서에 의해 자동으로 일련의 번호로 저장되게 되며 이러한 일련의 번호는 물리적인 비디오 프레임 번호와 연결되어 원하는 뉴스의 어떠한 씬이라도 직접 찾아갈수 있다.



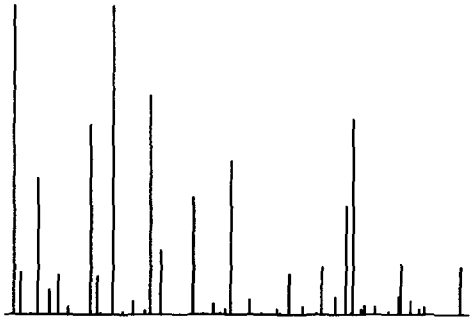
(그림 2) 뉴스 스토리의 색인화 구조

5. 실 험

본 논문의 실험에서 하드웨어 측면에서는 펜티엄 100에 주 기억용량이 32메가인 개인용 컴퓨터와 비디오 플레이어를 기본 실험 기기로 하였다. 구현되는 소프트웨어 측면의 환경은 윈도우 98 환경에서 마이크로소프트사에서 개발한 비쥬얼 C++ 5.0의 객체언어를 이용하였다. 실험 데이터가 포함된 비디오 테이프를 기본 입력 매체로 활용하였다. 비디오 테이프는 한국방송공사에서 녹음하여 판매한 이틀치(1998년 11월 22일, 12월 11일, KBS 저녁 9시 뉴스)를 대상으로 하였다. 비디오 테이프를 비디어 플레이어를 사용하여 구동시킨 후에, 비디오 데이터들은 오스카II라는 비디오 캡쳐 보드를 이용하여 AVI 압축 형태의 자료를 획득하였다. 획득되는 AVI 데이터는 초당 다섯 개의 프레임 비율로 캡쳐 하였고, 한 프레임은 400 × 300의 영상 크기로 변환한 자료를 실험에 사용하였다.

5.1 키 프레임 추출 예

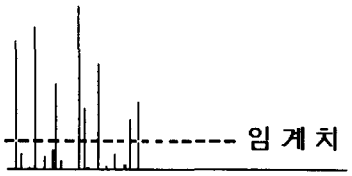
(그림 3)은 [5]에서 제안한 방법을 적용하여 얻어진 히스토그램 차이값을 나타낸다. 그림에서 값이 높게 나타난 곳은 장면의 변화가 심한 것을 의미하며 키 프레임으로 분할될 수 있다. 반면 값의 차가 거의 없는 것은 장면 간의 변화가 거의 없는 것이므로 분할되지 않는다. 이 방법은 프레임간의 작은 차이에는 민감하지 않지만 프레임간의 차이가 조금 크면 히스토그램 값이 높게 나타난다는 장점을 가지고 있다. 반면 프레임간의 칼라수에 의한 차이값의 계산으로 인해 장면의 변화가 일어나도 색의 수가 비슷하면 히스토그램 차이값이 높게 나타나지 않아 키 프레임으로 선택되어야 할 프레임이 선택되지 않는 오류를 범할 수도 있다.



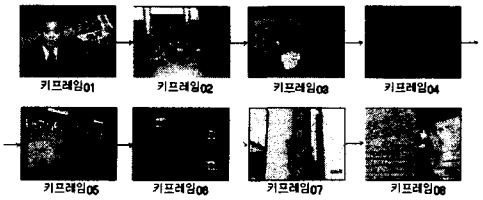
(그림 3) (식 1) 적용한 히스토그램 결과

(그림 4)는 얻어진 히스토그램 차이값에 임의의 임계치를 적용하는 그림이다. 임계치를 결정하는 방법은 사용자가 임의의 값을 전체 비디오에 적용하는 방법과 프로그램 자체적으로 일정한 간격의 범위내에 있는 차이값들을 비디오의 특성에 따라, 서로 감안하여 임계치를 적용하는 방법, 두가지가 대표적이다. 전자의 경우는 정적 임계치라고 부르며, 후자의 경우는 동적 임계치라 한다. 본 논문에서는 기존의 방법중 동적 임계치 방법을 사용하여 비디

오 마다 다른 임계치를 결정하여 사용하였다.



(그림 4) 히스토그램 차이값에 임계치 적용



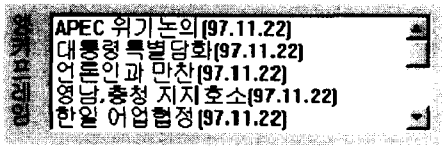
(그림 5) 추출된 키 프레임의 예

(그림 5)는 히스토그램 차이값에 임계치를 적용하여 선택된 키프레임의 예를 나타내고 있다. 이렇게 선택된 키 프레임들은 순서에 의한 일련의 번호로 저장되며 키 프레임들 사이의 프레임들의 대표 프레임으로 사용된다. 사용자가 특정 키 프레임을 브라우저에서 선택하면 그 키 프레임이 대표하는 특정 부분의 프레임들이 보여지게 되는 것이다.

5.2 브라우징 방법

(그림 6)은 비디오를 검색할 수 있는 검색창에 대한 그림이다. 앵커프레임을 선택하면 해당 뉴스 비디오가 화면에 출력되고, 키 프레임의 선택으로 의미있는 장면의 변화 부분을 찾아갈 수 있다.

(그림 7)은 뉴스 스토리 이름 리스트의 예를 그림으로 나타낸 것이다. 원하는 리스트를 선택하면 해당 부분으로 이동한 후, 원하는 프레임 번호로 이동한다.



(그림 7) 뉴스 스토리 예

그리고 (그림 8)에서 나타나는 키 프레임들은 뉴스 스토리를 구성하고 있는 작은 씬을 의미하며 선택하면 해당 부분의 비디오가 보여지게 된다.

(그림 8) 키 프레임 선택 예

(그림 6). 인덱스된 비디오 검색의 전체 화면

(그림 9) 비디오 제어 버튼

(그림 9)는 사용자가 선택한 비디오를 제어할 수 있는 버튼의 예를 나타낸 것이다. 그림에서 **(처음)**은 선택된 뉴스 스토리의 맨 앞, 즉 앵커 장면으로 이동하는 버튼이고, **(10F 앞)**과 **(5F 앞)**은 비디오가 재생되는 현재 부분에서 10 프레임 혹은 5프레임 앞으로 이동하는 버튼이고 **(Play)**는 정지되어 있는 비디오를 재생시키는 버튼이다. 그리고 **(5F 뒤)**와 **(10F 뒤)**버튼은 5 프레임 혹은 10 프레임 다음으로 이동하는 버튼이다. **(끝)**은 선택된 뉴스 스토리의 끝 프레임, 즉 가장 마지막 프레임으로 이동하는 버튼이고, **(정지)**는 재생되고 있는 비디오를 일시 정시킬 때 사용하는 버튼이다.

6. 결론 및 향후 연구 방향

본 논문에서는 뉴스 비디오의 효율적인 검색을 위한 색인의 생성 목적을 위해 기존의 장면 분할 기법 중 [5]에서 제안한, Color Histogram에서 각각의 RGB를 따로 계산하여 차이값을 세분화하는 장점과, χ^2 Histogram에서 차이값을 강조하는 장점을 이용하여 NTSC 표준에 따른 가중치를 적용한 장면 분할 방법을 이용하여 키 프레임 추출하고, 자동 추출된 키 프레임을 이용하여 색인화하여 사용자가 원하는 장면을 직접 볼 수 있도록 도와주는 검색 브라우저를 개발하였다.

향후 연구 과제로는, 추출된 키 프레임 중에서 앵커 프레임을 자동선택하고 선택된 앵커 프레임을 이용하여 색인화 하는 연구와 그리고, 키 프레임 추출시 임계치의 선택이 결과에 중요한 영향을 미치므로 알맞은 임계치 자동 선택에 대한 기법이 추후 계속 연구되어져야 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 27-07, 1998.

[6] Y. Gong et al. "An image database system with content capturing and fast image indexing abilities,"Proceedings of the International Conference on Multimedia Computing and Systems, Boston, MA, pp.121-130, May 1994.

[7] S. Devadiga, D. Kosiba, U. Gargi, S. swald, P.Kasturi, "A semiautomatic video database system," SPIE conference on Storage & Retrieval for Image & Video Databases, vol.2420, pp.262-267,1995.

[8] U. Gargi, S. Oswald, D. Kosiba, S. Devadiga, and R. Kasturi "Evaluation of video seqpuence indexing and hierarchical video indexing," SPIE conference on Storage & Retrieval for Image & Video Databaes, vol.2420, pp.144-151,1995.

[9] A. Akutsu(ed.), "Video indexing using motion vector," Visual Communications and Image Processing '92, Boston, MA, pp.1522-1530, November 1992.

[10] W. Wolf, "Multimedia Information on the Internet,"97International Conference Multimedia database on internet, pp.3-23, 1997.