

샷의 타입을 이용한 뉴스 아카이브 시스템의 설계 및 구현

(Design and Implementation of a News Archive System using Shot Types)

한근주[†] 남종호^{**} 하명환^{***} 정병희^{***} 김경수^{****}
(Keunju Han) (Jongho Nang) (Myung-hwan Ha) (Byunghee Jung) (Kyeonsoo Kim)

요약 뉴스 아카이브 시스템을 구축하기 위하여서는 먼저 뉴스 비디오 스트림을 기사 단위로 인덱싱하고, 사용자가 기사 비디오를 모두 시청하지 않아도 그 내용을 이해할 수 있도록 하는 추상화 방법이 필요하다. 본 논문에서는 뉴스 비디오 스트림에 대하여 샷 타입을 이용하여 기사 단위로 인덱싱할 수 있는 새로운 기사 경계 검출 방법 및 기사 추상화 방법을 제안한다. 제안한 인덱싱 방법에서는 뉴스 비디오의 샷들을 앵커 샷, 인터뷰 샷, 연설 샷, 보도 샷, 그래픽 자료 샷 등으로 나눈다. 모든 기사는 앵커 샷으로 시작하고, 앵커 샷은 다른 샷에 비하여 길이가 길고 특별한 화면 구조를 가지고 있기 때문에 이를 이용하여 기사 단위의 인덱싱을 수행한다. 또한 각 기사에 대한 효과적인 추상화를 위하여 앵커 샷의 오른쪽 위에 있는 그래픽 데이터와 기사를 이루는 다른 샷들의 키 프레임들을 이용한 기사 포스터를 구성하는 방법을 제안하였다. 여러 종류의 뉴스 비디오 스트림에 대한 실험 결과에 의하면 본 논문에서 제안한 기사 경계 검출 알고리즘의 검출율 (recall) 및 정확도 (precision) 값이 각각 0.92 및 0.96 이상 됨을 알 수 있다. 또한 본 논문에서는 WWW상에서 수행되는 뉴스 아카이브 시스템의 프로토타입 시스템의 설계 및 구현에 대하여서도 설명한다.

Abstract In order to build a news archive system, the news video stream should be first segmented into several articles, and their contents are abstracted effectively. This abstraction helps the users to understand the contents of the article without playing the whole video stream. This paper proposes a new article boundary detection scheme for the news video streams together with a new news article abstraction scheme using the shot types of the news video data. The shots in the news video are classified into anchor person shots, interview shots, speech shots, reporting shots, graphic shots, and others. Since the news article starts with an anchor shot whose duration is relatively longer than other shots with a special screen structure, the article boundary is detected by the computing the length of the shot and checking the screen structure in the proposed scheme. For the effective abstraction of the article video, the graphic image located in the right-top of the anchor shot frames is primarily used in the proposed abstraction scheme since it is the abstraction of the article made by the producer of the news according to its contents so that it contains a lot of meaningful information. The key frames of the other shots except interview and report shots are also used to abstract the contents of the articles in the proposed scheme. Upon experimental results, the precision and recall values of the proposed article boundary detection scheme could be 92% and 96%, respectively. This paper also presents a design and implementation of a prototype news archive system on WWW that consists of an indexing tool, an authoring tool, a database for meta-data of the news, and a browsing tool.

[†] 비 회 원 : 한국 IBM 연구원

kjhan@kr.ibm.com

^{**} 종신회원 : 서강대학교 컴퓨터학과 교수

Jhnang@ccs.sogang.ac.kr

^{***} 비 회 원 : KBS 기술연구소 연구원

mhha@kbs.co.kr

bhjung@kbs.co.kr

^{****} 비 회 원 : KBS 기술연구소 차장

odyssey@kbs.co.kr

논문접수 : 2000년 1월 11일

실사완료 : 2001년 7월 12일

1. 서론

오늘날 저장 매체의 비용 절감, 네트워크의 발달, 압축 기술의 향상에 힘입어 멀티미디어의 사용이 급증하면서 텍스트, 오디오, 비디오와 같은 다양한 형식의 데이터가 이용되고 있고, 그 중에서 특히 전달력과 흥미 유발에 효과적인 비디오 데이터는 점점 더 중요한 역할을 하게 되었다. 이에 더하여 멀티미디어 정보 처리 기술 및 인터넷의 보편화에 따라 WWW상에서 디지털 비디오 라이브러리와 같은 첨단 비디오 서비스가 가능하게 되었다. 디지털 비디오 라이브러리는 사용자의 요구에 맞는 디지털 비디오 데이터를 빠르고 효과적으로 전달하기 위한 시스템으로 데이터의 저장 및 관리, 검색 기능을 갖추어야 한다. 이를 위해서는 우선 비디오 데이터의 효율적인 인덱싱과 브라우징 및 비디오 데이터의 내용에 관한 추상화가 필수적이다.

디지털 비디오 라이브러리 시스템 응용 분야의 일종인 뉴스 아카이브 시스템은 다양한 정보 중 우리가 매일 쉽게 접하고 자료적 가치가 높은 뉴스를 대상으로 방대한 양의 뉴스를 인덱싱하고 각 기사의 내용을 추상화하여 보관함으로써 사용자의 요구에 맞는 기사 검색 기능을 제공한다. 내용의 추상화를 위해서 저급 수준의 내용 정보(칼라, 움직임, 텍스트, 물체 등)와 고급 수준의 내용 정보(장소, 시간, 인물, 주제 등)를 자동 추출하게 된다. CMU의 뉴스 아카이브 시스템[1]에서는 고급 수준의 내용 정보를 추상화하는 방법으로 클로즈드 캡션(Closed caption)을 이용하거나 스크립트의 내용을 형태소 분석하여 키워드를 추출하는 방법을 사용하고 있지만, 국내 TV 방송용 뉴스들은 클로즈드 캡션을 포함하고 있지 않고 스크립트의 형태소 분석을 통하여 고급 수준의 내용 정보를 추출해내는 작업이 완전하지 못하므로 국내 뉴스에 맞는 비디오 데이터 분석 방법 및 추상화 방법이 필요하다.

본 논문에서는 우리나라 방송 현실에 맞는 뉴스 아카이브 시스템을 설계하고 구현함으로써 사용자가 방대한 뉴스 데이터를 사이에서 원하는 기사를 효율적으로 검색, 브라우징 할 수 있도록 한다. 이를 위하여 우선 TV 방송용 뉴스에 대하여 샷 단위로 인덱싱 하고 인덱싱된 샷을 대상으로 어떤 보도 유형인지를 판단한다. 인터뷰, 기자 보도, 회견 혹은 연설과 같은 특정 형태의 보도 유형을 갖는 샷인지의 판단은 샷의 길이와 샷내에서의 움직임에 이용한다. 그래픽 자료 화면으로 구성된 샷은 화면을 이루는 색상의 변화율과 구성 비율을 이용하여 검출하며 앵커 샷은 최소한의 간격을 유지하며 반복

되는 점을 이용하여 검출한다. 이렇게 검출된 샷의 보도 유형을 바탕으로 기사의 경계를 추출하고 내용을 추상화하여 함축적으로 표현한다. 기사의 추상화는 의미 전달에 효과적인 앵커 샷의 어깨 길이 그림과 그래픽 샷의 대표 화면과 상대적으로 샷의 길이가 긴 샷의 대표 화면으로 하나의 새로운 대표 화면을 이용하여 구성한다. 특히 앵커 샷의 어깨 길이 그림은 이미 방송사에서 시청자의 흥미와 이해를 돕기 위해 기사의 내용을 추상화하여 제작한 것이므로 기사의 내용 전달에 효과적이다. 이와 같은 추상화 방법은 기사의 전체 내용을 보지 않고 대강의 의미를 전달받을 수 있어 검색 결과를 확인하는데 유용하게 사용될 것이다.

본 논문은 다음과 같이 구성되어 있다. 제 2 장에서는 기존의 뉴스 아카이브 시스템에 대해 알아보고 시스템을 구성하는 각 요소의 구현 방법을 기존의 연구를 통해 살펴본다. 제 3 장에서는 본 시스템에서 필요한 기능들을 정의하고 전체 시스템의 설계할 제안하고, 제 4 장에서는 제 3 장에서 제안한 방법을 구현한 시스템의 특징과 각 구성 요소들의 역할을 설명하고 실험결과를 분석한다. 제 5 장에서 뉴스 아카이브 시스템에 관련한 향후 연구를 제안하고, 결론을 맺는다.

2. 연구배경

사람들에게 키워드 기반의 질의가 월드 와이드 웹(World Wide Web)의 사용 경험이 증가함에 따라 친숙한 부분이 되고 있다. 그러나 방송용 뉴스와 같은 데이터는 그 양이 급속도로 증가하고 있어 단순한 글자의 나열로 원하는 데이터를 찾는 것이 제한적임을 느끼게 된다. 멀티미디어로 표현되는 정보의 증가는 이러한 데이터를 자동적으로 처리하여 사용자의 요구에 적합한 데이터를 공급할 수 있는 시스템을 요구한다. 특히 비디오와 오디오 데이터에 대한 요구가 많아짐에 따라 이들을 대상으로 의미 있는 단위로의 분할 작업을 실시하고, 분할된 데이터에 대한 내용을 추출하여 구조적인 데이터를 구성하여 사용자의 요구에 따라 정보를 제공하게 되는데 각각의 단계를 자동적으로 처리하기 위한 연구가 계속되고 있다. 이 중 뉴스는 그 구조가 비교적 정형화되어 있어 분할 작업과 정보 추출의 자동화가 다른 영역에 비해 성과가 좋고 뉴스 데이터는 오디오, 비디오 정보 외에도 자막과 스크립트 등 이용 가능한 정보가 다양한 방법으로 제공되고 있어 연구가 활발한 분야이다. 이들에 대하여 좀 더 자세히 설명하면 다음과 같다.

2.1 기존의 뉴스 아카이브 시스템

기존에 연구된 뉴스 아카이브 시스템은 음성 인식, 이

미지 처리, 기계 학습(Machine Learning), 자연어 처리 등 다양한 분야의 기술을 혼합하여 이전의 텍스트 기반의 키워드 검색 수준에서 오디오와 비디오 데이터로부터 추출한 데이터를 바탕으로 내용 기반의 검색을 지원하는 수준으로 발전되어 가고 있다.

CMU의 Informedia 디지털 비디오 라이브러리 시스템[1]은 방송용 뉴스에 대한 음성인식 기술과 이미지 및 자연어 처리 기술을 혼합하여 좀 더 진보된 검색을 가능하게 하고 뉴스를 기사단위로 자동 분할하여 원하는 장면으로 빠르게 접근 가능하게 하는 것을 목표로 개발되고 있다. <그림 1>은 Informedia 디지털 비디오 라이브러리 시스템에서 어떻게 메타 데이터를 생성하는지 그 과정을 설명하고 있다. 우선 비디오와 오디오 자료를 분리한 후 비디오 정보를 이용하여 샷 경계를 검출하고 얼굴 인식 과정과 자막 분석을 통해 얻어진 자료와 오디오 정보를 이용하여 음성 인식 과정을 통해 추출한 자료를 함께 이용하여 기사의 경계를 추출하고 앞서 추출된 자료를 이용하여 기사 단위로 메타 데이터를 생성한다. 이 메타 데이터를 바탕으로 사용자가 원하는 뉴스 기사를 검색하여 기사 단위로 결과로 보여준다. 그러나 기사 단위의 분할과 메타 데이터 생성에 가장 큰 영향력을 미치는 자료는 클로즈드 캡션으로부터 획득한 정보인데 클로즈드 캡션 내에 포함된 내용이 실제로 방송된 이야기 내용과 반드시 일치하지도 않고 작성 시 입력 오류를 포함하기도 해서 완벽하다고 할 수 없기 때문에 이 연구에서도 인정된 바와 같이 클로즈드 캡션 정보를 사용하지 않고 수행할 수 있는 방법이 제시되어야 한다.

MITRE에서 연구 개발 중인 NOD시스템[2]은 뉴스만의 고유한 형식과 한 기사가 몇 분을 넘지 않는다거나 이 시간대에는 광고가 나온다거나 하는 시간에 따른 내용 정보에 더하여 '여러분, 안녕하십니까?' 혹은 'KBS 뉴스 아무개였습니다'와 같이 자주 나오는 구에 대한 템플릿을 이용하여 뉴스 기사를 인덱싱 한다. 이외에 화자 전환 검출과 기사 텍스트를 이용한 키워드(인물, 장소 정보 등) 추출을 통해 인덱싱된 데이터에 내용 정보를 추가하게 된다. MITRE의 기사 경계 추출 방법은 도메인 특성을 분석하여 미리 도메인에 맞는 모델을 구성해야 하는데 이러한 모델링 과정 시간이 많이 소요되고 다른 도메인에는 적용하지 못하며 사용된 언어에 의존적이라는 단점이 있지만 앞서 설명한 Informedia에 비해 좀 더 정확한 결과를 추출할 수 있다는 장점을 지녔다.

2.2 뉴스 아카이브 시스템 구성요소에 대한 기존연구

2.2.1 샷 경계 검출

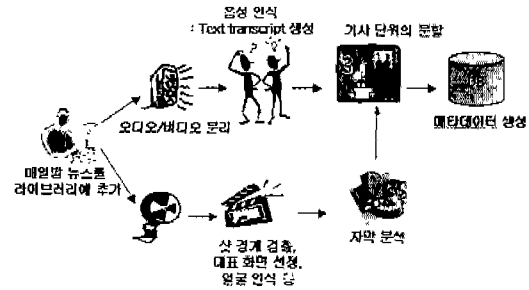


그림 1 Informedia 디지털 비디오 라이브러리 시스템

주어진 비디오 시퀀스의 내용을 분석하기 위해서 제일 먼저 비디오 시퀀스를 샷 단위로 분할한다. MPEG 형식의 압축된 비디오 시퀀스에 대한 기존의 샷 경계 검출 방법은 어떤 정보를 이용하는가에 따라 두 종류로 나누어진다. 첫번째 방법은 DCT상수 정보를 이용하는 방법[3,4,5,6,7,8]으로 비교 방식에 따라 DCT 상수의 위치를 고려하여 1:1로 비교하는 방법[4,5,6,7]과 위치에 상관없이 히스토그램을 작성하여 비교하는 방법[3,8]이 있다. 두 번째 방법은 움직임 벡터 정보를 이용하는 방법[6,7,8]으로 비디오에 대한 압축 방법 중 MPEG은 B 프레임과 P프레임을 부호화할 때 앵커 프레임과 유사한 부분이 있으면 이를 참조한다는 인코딩 정보를 이용하는 방법이다.

2.2.2 신 경계 검출

샷의 변화는 카메라의 움직임이라는 물리적인 원인에 의한 것이므로 프레임과 프레임 사이에 어떤 값들을 비교함으로써 그 경계를 찾을 수 있지만, 신은 배경 장소, 등장 인물, 앞 뒤 내용과의 관계에 의해 정의되는 의미적인 단위이므로 단순히 연속된 두 샷을 비교해서 찾을 수 없다. 신 경계에 대한 기존의 연구는 특정 비디오 분야를 대상으로 하는 도메인 종속적인 방법[6]과 영화와 같은 일반적인 비디오를 대상으로 하는 도메인 독립적인 방법[9,10,11,12]으로 구분된다. 우선 도메인 독립적인 방법으로, Yeung의 방법[10,11]을 설명하겠다.

Yeung의 방법에서는 우선 모든 샷에 대하여 프레임들을 시공간적으로 서브샘플링 한다. 시공간적으로 서브샘플링 된 샷들에 대하여 신의 경계를 찾기 위한 첫 단계로 유사한 샷들을 하나로 묶는 클러스터링 작업을 수행한다. 생성된 클러스터들에 대해 시간적인 전후 관계를 화살표로 연결한 STG(Scene Transition Graph)를 작성하여 클러스터간의 시간의 흐름에 따른 참조 관계를 바탕으로 STG내의 cut edge를 찾아서 이를 장면의

경계로 한다.

도메인 종속적인 방법으로 뉴스 영역을 대상으로 한 Zhang의 방법[6]을 살펴보면 뉴스 비디오는 앵커가 말을 하는 샷(Anchor person shot)과 그 뉴스의 내용을 보여 주는 샷(Reporting shot)이 반복되어 나타나는 특성을 이용하여 주어진 샷이 앵커 샷인지를 판단하여 앵커 샷이면 새로운 신¹⁾의 시작으로 판단한다. 앵커 샷은 <그림 2>와 같이 4가지의 유형으로 모델링할 수 있는데 이 때, 앵커 샷의 모델링을 위한 모델 이미지를 선택하기 위해서 앵커 샷은 샷 내에서 움직임이 거의 없으므로 샷의 각 프레임 색깔에 대한 평균과 분산을 구하여 그 값이 주어진 임계값보다 작으면 4가지 모델 중 하나가 될 수 있는지 검사한다. 앵커 샷에 대한 모델 이미지가 모두 선택되면 새로운 샷이 나타날 때마다 모델링된 앵커 샷과의 유사도를 검사해서 유사한 모델이 있으면 앵커 샷으로 하고, 그렇지 않으면 보도 샷으로 한다. 이 방법은 미리 앵커 샷의 모델을 구축하기 때문에 특정 뉴스 형태에서는 정확한 검출율을 나타내지만 모든 형태의 뉴스에 적용될 수 없고 모델링된 앵커 샷과 유사한 구도의 샷을 앵커 샷으로 오인하는 단점이 있다.



<모델 1> <모델 2> <모델 3> <모델 4>

AB : 앵커, C : 배경, D : 어깨 길이

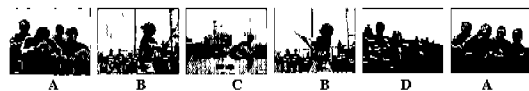
그림 2 뉴스 비디오에서 앵커 프레임의 4가지 모델

2.2.3 내용의 추상화 방법

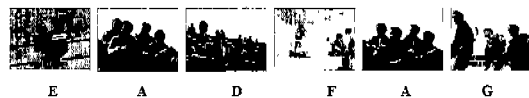
디지털 비디오에서 임의 위치의 내용을 보기 위해서는 비디오 데이터에 대한 분석과 인덱싱 외에도 브라우징 기능이 구비되어야 하며, 사용자에게 전체 혹은 일부의 내용을 잘 축약하여 효율적으로 전달하는 기능은 중요하다. 가령 MPEG-1 비디오로 압축된 30분 짜리 데이터(대략 300Mbyte)의 경우 고속의 인터넷상에서도 전송률이 초당10Kbyte를 넘지 못한다는 점을 생각해보면 전체 비디오의 내용이 다 필요하지 않을 경우 대강의 의미 전달이 될 수 있도록 재구성하여 전송하는 것이 효과적임을 알 수 있다. 따라서 사용자가 내용을 파악할 수 있도록 내용의 추상화 방법에 대한 연구를 도메인 종속적인 방법과 도메인 독립적인 방법으로 구분하여 설명한다.

1) 뉴스에서의 신은 기사(Article)를 의미한다.

도메인 독립적인 방법으로 가장 보편적인 방법이 각 샷의 대표 화면 혹은 일정 간격으로 선택된 화면을 시간 순으로 나열하여 만드는 스토리보드(Story board)[5, 13], 정지된 배경 혹은 전경과 함께 카메라의 패닝과 킬링이 있는 경우 파악이 용이한 이미지 아이콘[10,14] 등이 있다. 도메인 종속적인 방법으로인 Video Poster[15]는 스토리가 있는 영화, 시트콤과 같은 비디오 데이터와 뉴스와 같이 스토리를 갖지 않는 비디오 데이터에 대해 각기 다른 방법을 제시한다. <그림 3>의 (a)와 같이 유사한 샷들에 대하여 동일한 라벨을 부여하여 <그림 3>의 (b)와 같이 한 신 안에서 많이 반복되는 샷의 비중을 높여 그 비중에 따라 이미지 사이즈를 결정하여 하나의 이미지로 합성하여 적은 양으로 내용을 파악할 수 있도록 표현하는 방법이다. 일반적으로 많이 반복되는 샷의 경우 의미적으로 중요한 샷이라는 이유에서인데 뉴스의 경우에는 반복적으로 같은 샷이 나오는 경우가 드물기 때문에 우선 순위의 결정을 반복성에 두지 않고 샷의 길이로 한다. 실제 보도 기사의 구성을 살펴보면, 앵커의 리드 인 부분과 인터뷰, 기자의 보도 장면이 다른 샷에 비해 길이가 길다. 그러나 이런 부분들은 실제 기사의 내용 전달 측면에서는 중요도가 높지 않으므로 이러한 샷의 대표 화면이 포스터에서 큰 비중을 차지하게 되는 것은 문제가 있다. 따라서 단순히 샷의 길이나 반복성을 기준으로 추상화하는 방법이 아닌 뉴스의 특성에 적합한 추상화 방법이 제안되어야 한다.



(a) 하나의 신에 속하는 샷들에 대한 라벨 부여



(b) (a)에 대한 Video Poster

그림 3 Video Poster 제작 과정

3. 시스템설계

웹상에서 서비스되는 뉴스 아카이브 시스템 가져야 하는 기능을 사용자 요구 분석을 통해 분석하여 전체 시스템의 구조를 설계하고 각 요소의 세부 기능을 설명한다. 특히 기사의 추상화 방법과 추상화를 위한 기사의 보도 유형 판별 방법을 제안한다.

3.1 요구 분석

현재 국내 방송 3사와 같이 제목, 스크립트에 대한 제한된 검색 기능만을 제공하는 뉴스 아카이브 시스템[16, 17]의 확장을 위해서 우선 방송용 뉴스의 형식적 특성을 살펴보고 이를 바탕으로 데이터베이스의 설계를 위한 사용자의 질의를 분석한다.

3.1.1 방송용 뉴스의 형식

방송용 뉴스는 공통된 특별한 형식이나 정해진 규정을 갖고 있지 않지만 일반적으로 많이 적용되고 있는 특성은 다음과 같다. 방송 기사는 주제별로 정치/경제/사회/문화/과학/스포츠 등의 리포트로 나뉘어 지며 잘 짜여진 리포트는 도입부(Lead, Top), 본문(Narration), 인터뷰(Interview), 기자의 현장 목소리(On Mike Comment), 그리고 클로징(Closing Comment)으로 구성된다[18]. TV 방송용 뉴스는 앵커가 보도될 뉴스에 대하여 리드하는 멘트(Lead In)로 하나의 기사에 대한 보도가 시작되는데 그 시간은 10초에서 15초 사이가 된다. 이때 화면 우측 상단에 보도 내용에 대한 그림을 보여줌으로써 시청자의 시선을 끌게 되는데 이 역할을 하는 그림을 어깨 결이라고 한다. 어깨 결이의 위치는 지정된 것은 아니지만 오른쪽 상단인 구도가 가장 시선을 끌기 좋은 구도이다[20].

보도 기자의 오프닝 멘트나 중간 혹은 마지막 멘트는 15초 안팎이 좋으며 30초를 넘기면 좋지 않다. 인터뷰 컷은 대개 20초 안팎이 좋은데 어떤 때는 불과 5초 안팎의 짧은 컷이 나가는 경우도 있다. 기사를 구성하는 하나의 장면 길이는 3초가 적당하다. 우리 인간의 눈은 1/16초까지 지각할 수 있지만 최소한 2-3초는 가야 제대로 알아 볼 수 있고 3초를 넘기면 싫증을 낸다는 조사 결과에 근거한 것이다. 다만 인터뷰 컷트는 말하는 한 장면이 15초에서 20초 계속 될 수밖에 없다. 한 뉴스의 보도 길이에 대해서 KBS와 MBC 9시 뉴스를 대상으로 분석한 결과 우리 나라의 리포트 뉴스는 짧게는 50초 안팎에서 길게는 기획 뉴스의 경우 3분 안팎으로 제작되지만 일반적으로 1분 10초를 기준으로 해서 조금 짧거나 조금 길게 제작된다[21]. 이 중에 앵커의 리드 인 부분이 차지하는 시간은 평균 15초 안팎이며 한 뉴스는 리포트 뉴

스와 단신을 포함하여 평균 34건의 기사를 포함한다.

3.1.2 질의 분석

뉴스 아카이브 시스템을 통해 검색자가 원하는 기사의 검색을 지원하기 위해서는 질의 처리를 효과적으로 할 수 있고 질의의 대상이 되는 메타 데이터의 저장에 유용한 데이터베이스가 구축되어야 한다. 이러한 데이터베이스의 설계를 위하여 가능한 질의문을 예상하고 분석한다. 기본적으로 뉴스는 보도 날짜와 보도 내용의 범주에 따라 구분할 수 있다. 이외에 인터넷 신문에서 독자의 편이를 위해 읽고 있는 기사와 관련된 기사의 목록을 함께 제공하는 것처럼 동일한 이슈 혹은 사건에 대한 기사를 보기 위해서는 기사가 최근의 어떤 사건 혹은 이슈와 관련된 경우 명시해 주는 것이 이러한 서비스를 위해 필요하다. 이처럼 주요 사건을 하나의 요소로 간주하는 것은 옷 로비 사건과 관련된 기사를 보고 싶다고 같은 질의를 가능하게 한다. 뉴스의 문장은 육하원칙(5W1H)라 해서 누가(Who), 언제(when), 어디서(where), 무엇을(What), 왜(Why), 어떻게(How) 했다는 것이 나타나야 한다는 [22]의 연구에 기초해 볼 때 육하원칙을 대상으로 하는 탈주범 관련된 기사를 보고 싶다고 같은 질의가 가능하다. 이를 위해서는 육하원칙에 해당하는 요소를 추출하여 메타 데이터를 생성해야 한다. 육하원칙에 해당하는 내용은 주관적 관점을 최대한 배제하고 보도된 기사에 따라 객관적으로 추출해 낼 수 있기 때문에 좋은 추상화 대상이 될 수 있다. 위의 분석을 바탕으로 처리 가능한 질의의 예를 정리하면 다음과 같다.

- * 2월 20일 경제면 기사 목록을 보고 싶다.
- * 지난 30일 동안 발생한 정치면 이슈 목록을 보고 싶다.
- * 1999년 1년간 발생한 살인 사건에 대한 기사를 보고 싶다.
- * 국회에서 국회 의원들의 봄 싸움 장면이 담긴 기사를 찾고 싶다.
- * 설날, 추석과 같은 명절의 교통상황에 대한 기사들만 보고 싶다.
- * '카메라 출동'에서 보도된 자료 중 불법 시공에 관련된 기사들을 보고 싶다.
- * 판문점을 배경으로 한 자료 화면이 실린 기사를 보고 싶다.
- * '구제역 과문'과 같은 가축의 전염병에 관련된 이전 기사를 보고 싶다.
- * 지난 대선의 각 지방 투표율에 대해 알고 싶다.

3.2 시스템 설계

뉴스 아카이브 시스템은 사용자가 원하는 뉴스 기사를 찾아서 제공하는 서비스를 목적으로 한다. 이를 위해서는 우선 뉴스 데이터를 확보해야 하고 확보된 뉴스 데이터를 서비스 방법에 맞게 가공해야 한다. 가공된 뉴

스 데이터 중에 어느 것이 사용자가 원하는 것인지 판단하기 위한 처리 작업을 거친 후 작업 결과를 다시 사용자에게 전달하게 된다. 이러한 일련의 과정을 바탕으로 <그림 4>의 뉴스 아카이브 서비스를 위한 전체 시스템을 구성한다. 뉴스 아카이브 시스템을 메타 데이터의 생성을 위해 뉴스 데이터에 대한 인덱싱 작업과 내용 정보의 추상화 작업을 수행하는 <그림 4>의 (a)와 질의문을 처리하여 기사의 추상화 된 결과와 기사 관련 데이터를 사용자에게 전해주는 <그림 4>의 (b), 그리고 <그림 4>의 (c)의 사용자의 질의와 결과를 위한 인터페이스 부분으로 나누어 설명한다.

3.2.1 데이터 인덱싱

공중파를 통해 방송되는 뉴스를 MPEG 형식의 데이터로 압축한 MPEG 스트림에 대하여 자동 샷 경계 검출 방법[23]을 이용하여 구현한 프로그램을 통해 샷에 대한 인덱싱을 실시하고 그 결과를 바탕으로 기사 경계 검출 작업을 수행한다. 새로운 기사는 <그림 6>에 설명된 알고리즘에 따라 검출된 앵커 샷을 바탕으로 기사의 경계를 판단한다. 뉴스의 기사 경계 추출은 뉴스의 특성을 활용하기 때문에 높은 정확도를 나타내지만 샷 혹은 기사 경계 검출의 결과가 100% 정확하지 않기 때문에 보완 작업을 위한 인덱싱 도구를 사용한다. 수정 작업을 통해 확정된 샷과 기사의 인덱싱 결과는 데이터베이스에 저장되고 인덱싱 결과는 전체 기사 중에서 사용자가 원하는 기사만을 전달하고 해당 부분의 동영상만을 재생할 수 있도록 한다.

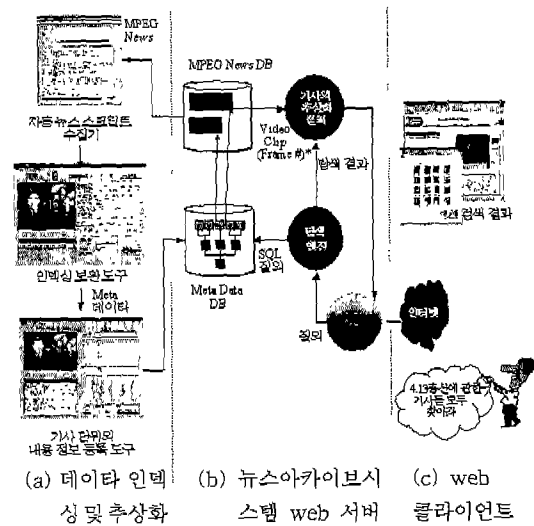


그림 4 뉴스 아카이브 시스템 구성도

3.2.2 저작 도구 및 뉴스 스크립트 수집기

인덱싱된 데이터에 대하여 3.1.2에서 설명한 질의를 처리하기 위한 메타 데이터를 추출하여야 한다. 그 중 기사 내용의 분야(section)와 어떤 주요 사건에 연관된 기사인가에 대한 정보의 추출을 위해서 기존의 시스템에서는 스크립트의 분석을 통해 동일한 어휘의 반복 회수를 이용하였는데 형태소 분석에 걸리는 시간과 정확성을 생각해 볼 때 아직 기대만큼의 성과를 내고 있지 못하다. 또한 [22]에서 설명한 바와 같이 뉴스의 문장이 포함하는 육하원칙의 기준에 따른 내용을 추상화하기 위한 메타 데이터의 자동 추출은 어렵다. 이런 측면에서 어느 정도 사람의 보완 작업이 필요하게 되고 이 작업을 도울 수 있는 도구가 필요하다. 본 시스템에서는 육하원칙에 해당하는 정보를 저작 도구를 이용하여 추상화 한다. 이 때 어휘 사전을 사용토록 하여 주관적 요소는 최대한 배제하고 저작 될 수 있도록 한다. 저작 도구를 통해 기사의 내용을 육하원칙에 기반한 객관적인 사실로 추상화하기 위해서 보도를 위해 제작된 스크립트(대본)에 기술된 것을 참고하게 된다. 스크립트는 각 방송사에서 웹 환경의 뉴스 아카이브 시스템에서 제공하는 내용을 자동 뉴스 스크립트 수집기를 통해 읽어와 데이터베이스에 등록된다.

3.2.3 데이터베이스 설계

내용 기반의 검색을 위해 질의의 대상이 되는 메타 데이터를 정의하고 전체 데이터베이스를 설계하였다. 데이터베이스는 시스템 안에 저장되는 멀티미디어 데이터에 대한 메타 데이터와 하드에 물리적으로 저장되어 있는 디지털 미디어 파일과의 관계를 포함하고 있다. 또한, 멀티미디어 데이터의 내용 설명과 멀티미디어 데이터간의 계층적 구조를 설명할 수 있어야 한다. <그림 5>의 (a)는 TV 방송용 뉴스 데이터에 대한 계층적 구조를 설명한 것이다. 이러한 관계 구조와 기사의 주제별 분류와 주요 사건별로 관련 기사의 검색을 지원하기 위한 테이블을 추가하여 <그림 5>의 (b)의 관계형 데이터베이스를 설계한다.

‘뉴스’ 테이블은 ‘KBS 9시 뉴스’와 같은 하나의 프로그램 단위의 뉴스 방송에 대한 정보를 포함한다. 일회의 뉴스 프로그램을 통해 보도되는 25여건의 기사는 ‘기사’ 테이블로 구성하여 자동 스크립트 수집기를 통해 읽어온 스크립트 정보와 동영상 파일에 대한 위치 정보, 기사 제목, 분류(Section) 정보와 함께 저장한다. 저작 도구를 통해 등록되는 기사의 내용 정보는 ‘키워드’ 테이블의 내용을 바탕으로 ‘속성’ 테이블에 등록된다. ‘키워드 구분’ 테이블에는 육하 원칙의 ‘누가’, ‘언제’, ‘어디’

서, '어떻게', '무엇을'에 해당하는 구분이며 필요에 따라 추가가 가능하다.

3.3 기사의 추상화

웹 브라우저에 바탕을 둔 인터페이스를 이용하여 뉴스 아카이브 시스템 사용자가 질의를 던지게 되면 그 질의에 해당하는 검색 결과는 때에 따라 수 백건에 이른다. 이런 데이터 가운데 어느 것이 사용자가 원하는 데이터인지 확인하기 위해 기존의 시스템에서 제공되는 바와 같이 단순히 기사 제목이나 스크립트의 내용을 읽어보거나 동영상상을 재생하는 것은 많은 시간이 소요된다. 따라서 검색 결과의 내용을 쉽고 빠르게 파악할 수 있도록 기사의 내용을 추상화하는 것은 뉴스 아카이브 시스템이 갖추어야 할 필수적 요소이다.

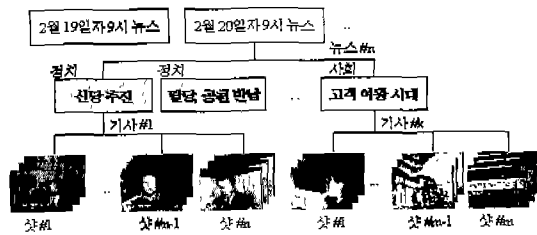
본 논문에서는 뉴스의 특성에 맞추어 기사 내용을 한 눈에 파악할 수 있는 추상화 방법을 제안한다. 기사 분석을 통해 보도 형태를 판별하여 내용 전달에 효과적인 보도 형태를 갖는 샷의 대표 화면들과 방송을 위해 기사의 내용을 잘 축약하여 표현한 앵커 샷의 어깨 결이를 이용하여 기사의 내용을 추상화한다.

간에 파악할 수 있도록 인덱싱 된 샷 혹은 신의 대표 화면들을 이용하여 추상화하는 방법이 많이 사용되었다. 한 신의 내용을 전달하기 위해 신을 구성하는 모든 샷의 대표 화면을 이용하여 추상화 할 경우 이해도를 높일 수 있는 반면 너무 많은 대표 화면의 나열로 인하여 시간이나 데이터 양의 측면에서 효과적이지 못하다. 이러한 이유로 적은 양의 데이터로 비디오의 내용을 효과적으로 추상화 할 것인가에 대한 연구가 활발해 졌다.

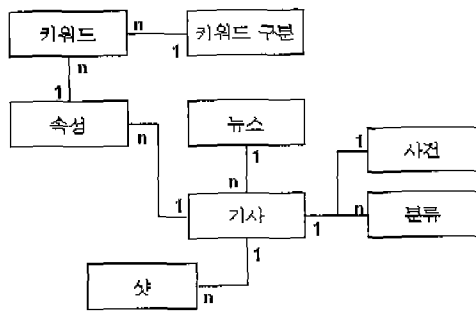
영화와 같이 줄거리를 갖는 비디오에 대해서는 한 신 안에서 자주 반복되는 샷이 다른 샷에 비해 그 신 안에서 비중 있는 샷이란 점을 이용하여 추상화한다[15]. 즉 신의 내용을 표현하기 위해서 신을 구성하는 샷 중에서 반복성이 높은 샷의 대표 화면을 이용하여 추상화하는 것이다. 그러나 뉴스의 경우에는 한 기사 안에 유사한 샷이 반복되는 경우가 거의 없어 이러한 특성을 반영할 수 없다. 뉴스의 기사를 구성하는 샷은 길이가 길수록 의미 전달에 있어 효과적인 샷이므로 이러한 특성을 이용하여 영화와 같은 비디오와는 틀리게 샷의 반복성보다는 샷의 길이가 긴 샷의 대표 화면으로 기사를 추상화하게 된다.

하나의 기사 안에서 샷의 길이가 긴 샷들의 보도 유형은 앵커가 리드인 멘트를 하는 부분, 인터뷰, 회견, 기자의 보도 부분이다. 대부분의 앵커 보도, 인터뷰, 기자 보도, 회견 혹은 연설을 보도하는 샷은 화자의 얼굴을 담고 있어 음성 정보를 듣지 않는 한 의미 전달 효과는 전혀 없다고 할 수 있다. 따라서 무조건 샷의 길이가 긴 샷의 대표 화면만을 이용하여 기사를 추상화하는 것은 효과적이지 못하다. 본 논문에서는 이러한 점을 해결하기 위해 보도 유형을 자동으로 판별하여 앵커 리드인 멘트 부분, 인터뷰, 회견 및 연설, 기자 보도 부분을 제외한 샷들로 기사를 추상화하는 방법을 제안한다.

하나의 기사는 앵커 리드인 멘트 부분, 인터뷰, 기자 보도, 회견 혹은 연설, 그래픽 자료화면, 일반 자료화면 등의 보도 형태를 갖는다[21]. 뉴스의 기사는 보통 인터뷰와 같은 보도 형태가 연속된 2개 이상의 샷에 걸쳐 나타나지 않기 때문에 보도 형태를 판별하는 것이 샷 단위로 가능하다. 우선 상영 시간(Duration)이 길고 화면내의 움직임이 적은 샷을 선택한다. 앞서 설명한 바와 같이 뉴스의 샷은 3초를 넘지 않지만 인터뷰한 사람의 의견, 기자의 보도 멘트를 전달받기 위해서는 일반 보도 자료 화면과는 달리 상영 시간이 길어지게 된다. 하나의 인터뷰, 앵커 보도, 회견, 기자 보도 화면은 하나의 샷으로 구성되므로 이러한 특성을 이용해 일반 보도 자료와 그 외의 것을 구분할 수 있고 앵커 보도 화면은



(a) 뉴스데이터의 계층적 구조



(b) 테이블 정의와 테이블간의 관계

그림 5 테이블 정의와 테이블간의 관계

3.3.1 샷 단위의 보도 형태 결정

길이가 긴 비디오 데이터의 내용을 관찰자가 빠른 시

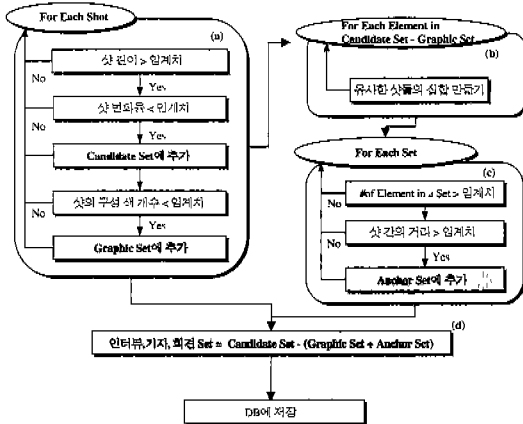


그림 6 샷의 보도 유형 판별

반복적으로 나타나므로 이를 이용하여 앵커 보도 화면을 구별한다. 그래픽 자료 화면은 일반 화면과는 달리 인위적으로 그려진 그림이기 때문에 화면에 사용된 색상이 단순하다는 점을 이용하여 구별한다. <그림 6>은 샷 중에서 앵커 리드인, 인터뷰, 기자 보도, 회견 혹은 연설, 그래픽 자료 화면을 구별하는 알고리즘이다. 여기서 각 보도 형태의 특징에 따라 샷이 어떤 보도 형태를 갖는지 파악하는 것은 하나의 기사에 대한 추상화 작업의 선행 작업임과 동시에 앵커 보도 부분이 새로운 기사가 시작되는 샷이므로 기사 단위로 분할하는 작업이기도 하다. <그림 6>의 (a) 단계에서 샷의 길이가 임계치보다 길고 샷내의 프레임간의 변화가 적은 샷을 추출하여 Candidate Set을 구성한다. 일반적으로 그래픽, 앵커, 인터뷰, 회견, 기자 보도 샷이 다른 샷에 비하여 그 길이가 길고 샷내의 움직임이 적기 때문에 위의 방법으로 추출된 Candidate Set은 그래픽, 앵커, 인터뷰, 회견, 기자 보도 샷 중 하나일 가능성이 높다. Candidate Set 가운데서 샷을 구성하는 색의 개수가 임계치 보다 적은 경우를 그래픽 샷으로 정의하며 (b) 단계에서 Candidate Set의 샷들간의 유사도를 비교하여 유사한 샷들의 집합을 만든다. (c) 단계에서 (b)에서 구성한 집합의 원소의 개수가 많은 것(즉 자주 반복적으로 나타나는 샷의 집합)을 대상으로 같은 집합에 속한 샷간의 거리를 측정한다. 샷간의 거리가 일정 거리를 넘지 않을 경우 이를 앵커 샷으로 판단할 수 있다. <그림 6>의 (a)에서 구한 샷의 집합에서 (b)와 (c)의 과정을 통해 구한 그래픽 보도 타입을 갖는 샷과 앵커 샷을 제외한 나머지 샷은 인터뷰 혹은 회견, 기자의 보도 타입의 샷이라고 판단할 수 있게 된다.

3.3.2 기사의 추상화

본 논문에서 제안하는 각 기사에 대한 추상화 방법은 기사를 구성하는 중요한 몇 개의 샷의 대표 화면과 앵커 리드인에 나오는 어깨 결이를 이용한다. 어깨 결이로 이용되는 그림은 보도하는 방송사 측에서 이미 기사의 내용을 잘 축약하여 표현한 것일 뿐만 아니라 보도하고자 하는 기사의 제목을 5글자 미만의 글로 표현하고 있어 기사 내용의 전달력이 그 어떤 화면보다 높다. 앵커 리드인의 형태를 갖는 샷의 대표 프레임에서 어깨 결이 부분을 추출하여 기사의 대표 화면 제작시 가장 시선을 끌기 좋은 곳에 위치시킨다.

기사의 내용 전달에 효과적인 몇 개의 샷의 선정은 샷의 길이와 보도 유형에 따라 결정된다. 보도 유형 중 그래픽 자료 화면은 보통 기사의 내용을 일목요연하게 정리하여 설명할 때 많이 사용되므로 기사의 내용 전달력이 높기 때문에 기사가 그래픽 자료 화면을 포함한다면 그 샷의 우선권을 어깨 결이 그림 다음으로 한다. 반면 주로 인물 위주로 구성된 인터뷰, 회견, 기자 보도 화면은 화면 속의 인물이 말하는 내용을 전달하기 위한 것이기 때문에 오디오적 정보는 의미가 있지만 비디오적 정보는 기사 내용 전달에 큰 도움을 주지 못한다. 게다가 정치면을 제외하고는 대부분 일반인들에게 인지도가 낮은 사람들이므로 인물 자체에 대한 정보로서의 의미도 떨어진다. 샷의 보도 형태를 구별하는 단계에서 인터뷰, 회견, 기자 보도 타입의 샷으로 구별하는 것은 어렵지만 이 셋 중 하나의 타입이라는 것은 판단할 수 있으므로 대표 화면 구성시 인터뷰, 회견, 기자 보도 화면은 추상화 대상에서 제외한다. 일반 보도 화면에 대해서는 샷의 길이가 긴 것부터 높은 우선 순위를 부여한다. 즉 우선 순위는 어깨 결이 그림, 그래픽 화면, 일반 보도 화면 중 길이가 긴 샷의 대표 화면 순으로 부여된다.

추상화 화면 구성은 그래픽 샷의 개수에 따라 결정되는데 각 타입의 종류는 [15]의 내용을 바탕으로 800x600 resolution의 사용자가 웹 브라우저를 통해 결과를 확인할 경우 전달 측면에서 효과적인 구성을 갖도록 재구성한 것이다. 효과적인 구성이란 352x240 MPEG 스트림에서 대표 화면을 JPEG로 전환시 화면의 본래 크기 비율을 유지하면서 그림의 내용을 인지하기 용이한 정도의 크기적 조건을 갖고 한 화면 내에 몇 개의 결과씩을 보여주는 것이 효율적인가 등의 측면을 고려한 것이다.

각 샷에 부여된 우선 순위를 바탕으로 기사에 대하여 <그림 7>과 같이 추상화하여 표현한다. 사람의 시선이 좌우, 상하의 순서로 머문다는 점을 반영하여 시선을 끌기 좋은 곳부터 우선 순위가 높은 그림을 배치하고 기

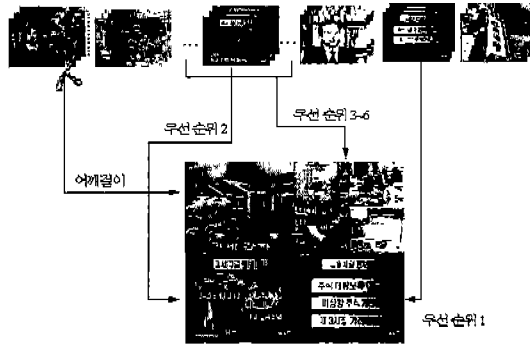


그림 7 기사의 추상화 방법

사를 구성하는 샷의 형태에 따라 그 크기와 그림의 개수를 조절한다. 뉴스 아카이브 시스템 사용자의 질의에 따라 어떤 형태로 기사를 추상화 할 것인가에 대한 결정이 달라지는 것이 아니므로 미리 기사의 추상화에 대한 정보를 데이터베이스에 저장하여 질의 결과로 사용한다.

4. 뉴스 아카이브 시스템의 구현

4.1 시스템 구현

본 연구에서는 앞 장에서 설명한 기사 단위의 인덱싱 방법 및 추상화 방법을 바탕으로 WWW상에서 수행되는 뉴스 아카이브 프로토타입 시스템을 구축하였다. 실제 구현에 대하여 설명하면 다음과 같다.

* 데이터베이스 구현

3 장에서 설계한 데이터베이스를 뉴스 아카이브 시스템의 관련 모든 프로그램에서 사용가능하고 효과적으로 데이터를 관리하게 되는 것을 보이기 위해 MS-Windows98에서 MS-Access97을 이용하여 구현하였다. 테이블과 테이블간의 관계는 <그림 5>에서 설명한 바와 같이 정의하였다.

* 인덱싱 도구

샷의 자동 경계 검출은 [23]에서 제안한 방법으로 구현한 프로그램을 이용하였고, 그 결과를 바탕으로 앵커 샷 검출을 통해 기사 경계를 자동으로 추출한다. 검출 결과를 확인하고 수정하기 위한 인덱싱 도구는 MS-Windows98에서 MS-Visual Basic 6.0을 사용하여 구현하였고 <그림 8>에 나타낸 것과 같이 저작자가 내장되어 있는 MPEG 상영기를 이용하여 대상 MPEG 데이터를 상영하면서 인덱싱할 수 있도록 도와주는 기능을 한다. 이 도구는 사람의 개입을 최소화, 간편화하기 위한 것으로서 재생 화면을 보면서 저작자가 그 샷과 기

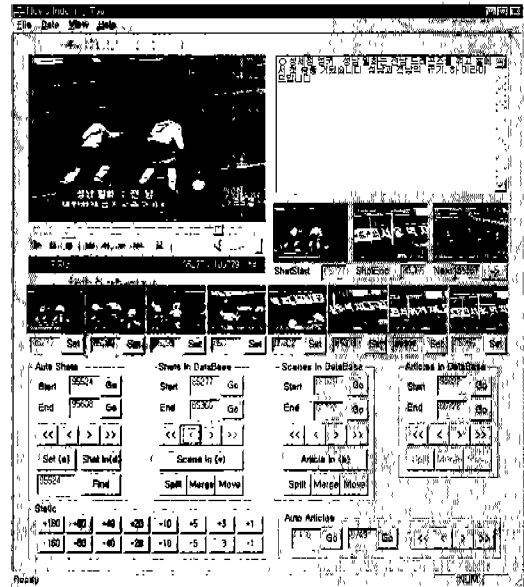


그림 8 인덱싱 도구

사의 대표 화면을 선택할 수 있는 기능과 구간 재생 등의 편의 기능을 제공한다.

* 저작 도구

MS-Windows98상에서 MS-Visual Basic 6.0을 이용하여 구현한 저작 도구를 통해 각 기사가 갖는 내용 정보를 추상화한다. 저작자의 주관을 최대한 배제하고 객관적인 정보로 추상화 할 수 있도록 제안한 육하원

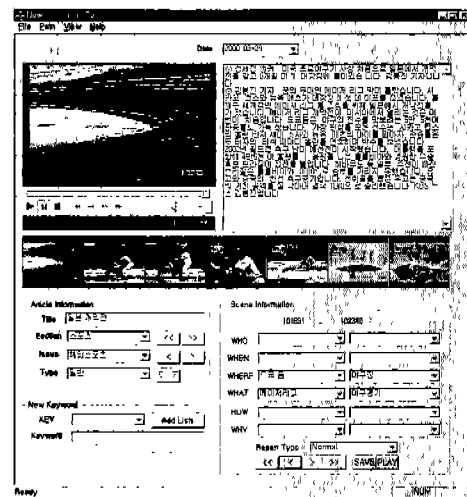


그림 9 저작 도구

칙에 따른 저작 방법을 반영하였고 기사 제목과 스크립트는 자동 스크립트 수집기를 통해 가져온 데이터로 자동 추가된다. MPEG 스트림의 재생은 Active-X Control을 통해 이루어진다. <그림 9>는 본 연구에서 설계/구현한 저작 도구의 수행 화면 예이다.

4.2 실험 및 분석

본 논문에서 구현한 기사 경계 검출 알고리즘의 테스트에서 샷의 길이와 유사도, 샷간의 간격에 대한 임계값은 실험에 의해 결정하였다. 실험에 사용된 데이터는 모두 352x240 크기의 MPEG-1 시스템 스트림으로 KBS 9시 뉴스를 대상으로 하였다.

다음 <표 1>은 실험 데이터의 특성을 설명한 것이며, <표 2>는 실험 데이터에 대해 가장 좋은 실험 결과를 나타내는 임계값을 적용하여 샷의 보도 타입을 판별하는 알고리즘을 수행한 결과이다. 이 결과는 검출 결과를 이용하여 계산한 리콜(recall)과 프리시전(precision)의 값도 포함하고 있는데 리콜과 프리시전을 구하는 식은 다음 <식 1>과 같다. <식 1>을 보면 리콜은 찾지 못한 결과를 측정하는 기준이고, 프리시전은 잘못 찾은 결과를 측정하는 기준임을 알 수 있다.

	Relevant	Not Relevant
Retrieved	A Correctly retrieved	B Incorrectly retrived
Not Retrieved	C . Missed	D Correctly reject

$$Recall = \frac{\text{Relevant retrieved}}{\text{All relevant}} = \frac{A}{A + C}$$

$$Precision = \frac{\text{Relevant retrieved}}{\text{All retrieved}} = \frac{A}{A + B} \quad (1)$$

표 1 실험에 사용한 뉴스 비디오 데이터

비디오종류	뉴스1	뉴스2	뉴스3	뉴스4	뉴스5
프레임개수 (시간 m:s)	77965 (43:21)	30059 (16:42)	32513 (18:04)	42858 (23:08)	38099 (17:52)
샷 개수	634	255	276	270	268
기사 개수	25	10	11	15	11
인터뷰/기자/ 회견 샷 개수	71	21	27	32	21
앵커 샷 개수	25	10	11	15	11
그래픽 샷 개수	8	9	7	9	3

이 실험 결과에서 샷의 타입을 결정하지 못하는 경우는 샷의 길이가 임계치보다 작은 경우인데 3장에서 설명한 바와 같이 인터뷰나 회견 등과 같이 화자의 음성

표 2 샷의 보도 타입 판별기 수행 결과

비디오종류 실험결과	뉴스1	뉴스2	뉴스3	뉴스4	뉴스5
찾은 개수	67	23	26	30	21
잘못 찾은 개수	8	4	2	4	4
찾지 못한 개수	12	2	3	6	4
Recall	0.83	0.91	0.89	0.81	0.81
Precision	0.88	0.83	0.92	0.87	0.81
찾은 개수	26	10	10	16	14
잘못 찾은 개수	2	0	0	1	4
찾지 못한 개수	1	0	1	0	1
Recall	0.96	1.0	0.91	1.0	0.92
Precision	0.92	1.0	1.0	0.94	0.71
찾은 개수	8	9	6	12	4
잘못 찾은 개수	2	2	1	3	1
찾지 못한 개수	2	2	2	1	0
Recall	0.75	0.78	0.71	0.9	1.0
Precision	0.75	0.78	0.83	0.75	0.75

을 이용해 내용을 전달하는 샷의 경우는 그 길이가 20 초 안팎인 것이 일반적이나 불과 5초 안팎의 짧은 컷이 나가는 경우도 있기 때문이다.

그래픽 샷은 설명을 위해 사용되는 화면이기 때문에 샷의 길이가 일반 자료 화면보다 길고 애니메이션 효과 등을 사용하여 색의 변화가 다른 샷에 비해 크다는 것을 타입 결정의 기준으로 삼았는데 이러한 조건을 만족하지 못하는 <그림 10>의 (a)와 같은 그래픽 샷일 경우 찾지 못하였다. 앵커 샷의 경우에는 단신을 보도할 경우 앵커의 얼굴이 잠시 비춰지고 바로 기사 화면으로 넘어가기 때문에 앵커 샷으로 판별하지 못한다. 인터뷰나 회견, 기자 보도 샷은 일반적으로 화자의 얼굴을 중심으로 잡는 화면 구도를 갖고 카메라의 위치나 각도의 변화가 거의 없기 때문에 화면의 변화율이 작은 샷을 대상으로 판별하게 되는데 때로는 <그림 10>의 (b)와 같이 전체 모습에서 특정 인물로 클로즈업 되는 구도를 갖는 경우가 있어서 잘못 판단하는 경우가 발생한다. 인터뷰 샷을 앵커 샷으로 잘못 판단하는 경우가 있는데 이런 경우는 인터뷰 화면의 구조가 앵커 샷의 구조와 유사하면서 동시에 앵커와 유사한 색의 양복을 입은 경우이다. 샷의 유사도는 [19]에서 제안한 chi-square 방법을 사용하여 측정하였고 앵커 샷은 유사한 샷이 반복된다는 성질을 이용하여 검출하였는데 <그림 10>의 (c)처럼 '국무 회의'에서 대통령의 발언 장면은

한 기사 안에서 반복되어 나와서 앵커 샷의 set으로 잘못 판단 할 수 있기 때문에 이를 위해서 앵커 샷간의 최소 거리를 유지하는가를 함께 고려하였다.

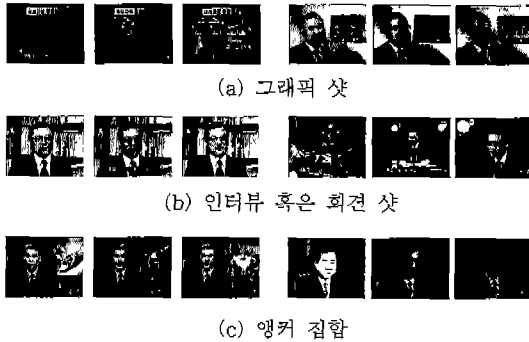


그림 10 샷 타입별 일반적 예와 오류를 발생하는 예

본 논문에서 제안한 추상화 방법에서는 판별된 보도 타입을 바탕으로 기사를 추상화하여 표현한다. 우선 앵커 샷의 어깨 절이 부분만을 추출하여 좌측 상단에 위치시킨 후 그래픽 화면이 있을 경우 그 다음으로 우선 순위를 갖는 곳에 위치시킨다. 남은 부분은 인터뷰, 기



(a) 기사를 구성하는 모든 샷의 대표화면



(b) 기사의 추상화 결과

그림 11 기사의 추상화 예

자 보도, 회견 혹은 연설 등의 샷을 제외한 샷 가운데 샷의 길이가 가장 긴 것부터 채워나간다. 이렇게 만들어진 기사의 추상화 예는 <그림 11>과 같다.

5. 결론 및 향후 연구 방향

디지털 라이브러리의 일종인 뉴스 아카이브 시스템과 같은 디지털 비디오 데이터를 이용한 응용 프로그램은 비디오 데이터의 내용 정보를 추상화하여 사용자에게 의미를 쉽고 빠르게 전달할 수 있도록 하기 위해 먼저 비디오 데이터에 대한 인덱싱과 내용의 추상화 기술이 필요하다. 본 논문에서는 뉴스 아카이브의 시스템에 대한 요구사항을 분석하여 시스템을 설계하고 구현하기 위하여 MPEG 뉴스 비디오 스트림을 샷 단위로 인덱싱하고 이를 대상으로 샷이 갖는 보도 타입을 판별하여 이를 이용한 기사 경계 검출 알고리즘과 기사 내용의 추상화 방법을 제안하였다. 제안한 알고리즘을 바탕으로 샷 타입 판별 도구를 구현하였고 이를 이용하여 기사 단위의 인덱싱을 실시하였다.

본 논문에서 제안한 샷의 타입 판별은 샷의 길이와 샷 내의 변화율, 샷을 구성하는 색상의 변화율 등을 이용하여 앵커, 인터뷰, 기자, 회견 혹은 연설, 그래픽 보도 유형을 갖는 샷을 구별해 낸다. 특히 앵커 샷의 판별은 시간축을 기반으로 움직임이 거의 없는 샷 가운데 최소 간격을 유지하며 반복적으로 나타나는 샷의 특성을 이용하였다. TV 방송용 뉴스의 특징을 분석한 결과 앵커의 리드 인 멘트로 기사가 시작된다는 점을 이용하여 앵커 샷을 기준으로 기사를 인덱싱하며, 기사의 내용을 함축적으로 표현하기 위해서 의미 전달에 효과적인 앵커 샷의 어깨 길이와 그래픽 샷을 이용하여 적은 양의 그림으로 내용 전달의 효과를 높였다.

움직임이 거의 없는 샷을 찾고 샷의 유사성을 판단하기 위해 색상에 대한 히스토그램을 사용하였는데 MPEG 스트림에서 I 프레임의 DC값만을 복호화 하였기 때문에 오버헤드가 적은 연산과정을 수행하게 된다. 기사의 내용 추상화에 있어서도 샷 타입의 판별 결과를 이용하여 모든 샷들의 대표 화면으로 내용을 전달하던 기존의 방법이 지닌 단점을 보완하였다. 기존의 방법은 추상화 된 데이터 자체가 너무 많은 그림으로 이루어져 있어 그림을 전송 받고, 전송 받은 그림을 이해하는데 많은 시간이 필요하지만 샷의 타입을 판별하여 기사의 내용을 잘 전달할 수 있는 샷들로 새롭게 구성한 대표 화면은 시간적 측면이나 전송 속도, 전달력 등에서 좋은 결과를 나타낸다.

앞으로의 연구에서는 어떠한 형식의 뉴스 비디오 데

이타에 대해서도 좋은 검출율을 나타낼 수 있도록 도메인 독립적인 방법이 제안되어야 할 것이다. 또한 자막을 포함하고 있는 샷이 단순한 화면보다 내용 전달에 효과적이므로 기사의 추상화를 위해서 자막을 포함한 샷을 추출하게 된다면 좀 더 효과적인 추상화가 가능할 것이다. 이 외에 좀 더 자동화된 뉴스 아카이브 시스템의 구축을 위해서 수동 작업을 통해 내용 정보를 추상화 하던 단계에서 스크립트를 자동으로 형태소 분석하여 주요 단어를 추출하고 자막의 내용을 자동 분석하여 기사의 내용을 추상화 하고 화자 전환 인식을 통해 앵커 샷, 인터뷰나 회견 등의 샷을 구별해 내는 연구가 진행 될 것이다.

참 고 문 헌

[1] Christel, M., Kanade, T., Mauldin, M., Reddy, R., Stevens, S., Wactlar, H., Techniques for the Creation and Exploration of Digital Video Libraries, *Multimedia Tools and Applications*, Vol. 2, Borko Furht, editor. Boston, MA: Kluwer Academic Publishers, 1996

[2] A.G.Hauptmann and M.J.Witbrock, Story Segmentation and Detection of Commercials in Broadcast News Video, *Proc. of ADL-98 Advances in Digital Libraries Conference*, 1998.

[3] B.L.Yeo and B.Liu, Rapid Scene Analysis on Compressed Video, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.5, No.6, pp.533-544, 1995.

[4] F.Arman, A.Hsu and M.Chiu, Feature Management for Large Video Database, *Proc. of Storage and Retrieval for Still Image and Video Databases*, Vol. SPIE 1908, pp.2-12, 1993.

[5] F.Arman, A.Hsu and M.Chiu, Image Processing on Compressed Data for Large Video Databases, *Proc. of First ACM International Conference on Multimedia*, pp.267-272, 1993.

[6] H.J.Zhang, Y.Gong and S.W.Smoliar, Automatic Parsing of News Video, *Proc. of IEEE Conference on Multimedia Computing Systems*, pp. 45-54, 1994.

[7] H.J.Zhang, C.Y.Low and S.W.Smoliar, Video Parsing and Browsing using Compressed Data, *Multimedia Tools and Applications*, Vol. 1, pp. 89-111, 1995.

[8] J.Meng, Y.Juan and S.F.Chang, Scene Change Detection in MPEG Compressed Video Sequences, *Digital Video Compression : Algorithms and Technologies*, Vol. SPIE 2419, pp.14-25, 1995.

[9] J.R.Kender and B.L.Yeo, Video Scene Segmentation

via Continuous Video Coherence, *Proc. of Conference on Computer Vision and Pattern Recognition*, pp.367-373, 1998.

[10] M.Yeung, B.L.Yeo and B.Liu, Segmentation of Video by Clustering and Graph Analysis, *Computer Vision and Image Understanding*, Vol. 71, No. 1, pp.94-109, 1998.

[11] M.Yeung and B.Yeo, Video Visualization for Compact Presentation and Fast Browsing of Pictorial Content, *IEEE Transactions on Circuits and systems for Video Technology*, Vol. 7, pp. 771-785, 1997.

[12] Y.Tonomura, A.Akutsu, K.Otsuji, and T.Sadakata, VideoMAP and VideoSpaceIcon: Tools for Annotating Video Content, *Proc. of INTERCHP93, Human Factors in Computing Systems Conference*, pp.131-136, 1993.

[13] Yasuo Arika, Telop and Flip Frame Detection and Character Extraction from TV News Articles, *Proc. of International Conference for Information and Documentation*, pp.198-202, 1994.

[14] L.Packer and T.Griffiths, *High Performance, TV News Reporting*, 1984.

[15] S.W.Smolia and H.J.Zhang, Content-based Video Indexing and Retrieval, *Proc. of International Conference on Multimedia Computing and Systems*, pp.62-72, 1994.

[16] <http://knews.kbs.co.kr>, KBS 뉴스 아카이브 시스템.

[17] <http://news.sbs.co.kr>, SBS 뉴스 아카이브 시스템.

[18] 윤재홍, *TV 뉴스 취재에서 보도까지*, 커뮤니케이션북스, pp.217-221, 1998.

[19] I.K. Sethi and N. Patel, A Statistical Approach to Scene Change Detection, *Proc. of SPIE 2420*, pp. 329-339, 1995.

[20] 한국언론연구원, *TV 뉴스 보도* pp.52-84, 1990.

[21] 한국방송학회, *보도 프로그램 제작의 실제, 방송 보도론*, 나남, pp.127-199, 1993.

[22] 이계경, 김학희, *방송 뉴스 취재와 보도*, 나무와 숲, 1999.

[23] J.H. Nang, S.W. Hong and Y.I. Ihn, "An Efficient Video Segmentation Scheme for MPEG Video Stream using Macroblock Information," *Proc. of ACM Multimedia Conference'99*, pp.23-26, 1999.



한 근 주
1996년 2월 서강대학교 수학과 졸업(학사). 2000년 8월 서강대학교 대학원 컴퓨터학과 졸업(석사). 2000년 8월 ~ 현재 한국 IBM 근무. 관심분야는 멀티미디어 시스템, E-Business, 인터넷 프로그래밍 등



남 중 호

1986년 서강대 전자계산학과 졸업. 1988년 한국과학기술원 석사. 1992년 한국과학기술원 박사. 1992년 ~ 1993년 Fujitsu연구소 연구원. 1993년 ~ 현재 서강대학교 컴퓨터학과 교수. 현재 서강대학교 컴퓨터학과 부교수



하 명 환

1993년 2월 경북대학교 공과대학 전자공학과 졸업(학사). 1995년 2월 한국과학기술원 전기 및 전자공학과 졸업(석사). 1995년 2월 ~ 현재 KBS 기술연구소 연구원. 관심분야는 멀티미디어 방송 제작 시스템, 비디오 인덱싱, 영상 처리



정 병 회

1994년 2월 이화여자대학교 전자계산학과 졸업(학사). 1996년 2월 한국과학기술원 전산학과 졸업(석사). 1996년 1월 ~ 현재 KBS 기술연구소 연구원. 2000년 9월 ~ 현재 한국과학기술원 전산학과 박사과정 재학중. 관심분야는 멀티미디어 방송 시스템, 미디어 아카이브, 초고속 네트워크 시스템



김 경 수

1983년 2월 서울대학교 공과대학 제어계측공학과 졸업(학사). 1985년 2월 서울대학원 제어계측공학과 졸업(석사). 1985년 3월 ~ 현재 KBS 기술연구소 차장. 관심분야는 멀티미디어 방송 제작 시스템, 스트리밍 미디어 및 서비스, 콘텐츠 보호 및 관리 기술