

멀티미디어 정보검색의 응용

유성준, 김의선
서치캐스트 주식회사

1. 서론

멀티미디어 정보검색이란 오디오 또는 비디오 테이프나 필름 저장소 등의 아날로그 저장소와 디지털화된 멀티미디어 데이터 저장소와 같은 데이터베이스에서 멀티미디어 자료나 그와 관련된 정보를 찾는 것을 의미한다.

Offline 세계에서 멀티미디어 정보를 찾기 위한 방법으로서는 크게 두 가지가 있다. 자료가 있는 곳을 일상적인 말(텍스트)로 물어보거나, 찾고자 하는 멀티미디어 데이터와 비슷한 데이터로 관련 자료를 찾는 방법이다. 물론 이 두 가지 방법은 독립적이기보다는 상호 보완적으로 사용된다. 또한 컴퓨터에서 멀티미디어 정보를 찾을 때에도 같은 방법이 사용된다. 즉 텍스트 기반 정보검색 기술과 내용기반 멀티미디어 정보 검색(Content Based Multimedia Information Retrieval) 기술을 이용하는 것이다.

이 두 가지 기술의 기반이 되는 학문은 컴퓨터 및 전자공학의 다른 분야에 비해서 비교적 오랫동안 발전해 온 분야이지만 아직까지도 offline 상에서 이루어지는 인간의 인지능력에는 훨씬 못 미치는 수준에 머물고 있다. 물론 현재까지 이러한 기술이 사용자들의 정보검색에 대한 수요를 상당 부분 덜어주고 있지만 이 분야의 전문가들이 향후에도 많은 시간, 비용, 노력 등을 투자해서 완벽한 수준의 기술이 이루어져야 한다.

한때 일부에서는 멀티미디어 정보검색을 영상처리 및 음향 신호처리 기술에 기반한 내용기반 멀티미디어 정보검색이 전부라고 생각하는 경향이 있었다. 내용기반 멀티미디어 정보검색이라는 용어는 <Content Based Multimedia Information Retrieval>이라는 용어를 한국어로 그대로 번역하면서 생겨난 용어인데 자칫 텍스트 분야에서 사용하는 의미기반 정보검색과 혼동될 수 있다. 두 용어 모두 각 미디어(사진, 비디오, 텍스트)에 담겨 있는 내용과 사용자의 질의 내용이 일치, 또는 유사한 답을 검색해 주는 것을 뜻한다. 그러나 의미기반 정보

검색이란 텍스트에 기반을 둔 정보검색을 뜻하고 내용기반 정보검색이란 텍스트가 아닌 각 미디어의 특성에 기반을 둔 정보검색을 뜻한다.

2. 텍스트 기반 멀티미디어 정보검색

멀티미디어 정보검색은 기존의 일반 정보검색보다 객체의 속성에 대한 문제가 지원되어야 한다. 예를 들어 선인장이 들어있는 장면을 찾는 요구가 있다고 가정하자. 그리고 모든 선인장을 찾는 것이 아니라 빨간 선인장을 찾는다고 하자. 이런 경우 현재의 검색 기술은 '선인장'으로 찾은 검색결과 중 빨간 선인장에 해당하는 결과를 골라내는 정도에 그치고 있다. 만약 한 단계 더 나아가 '빨간'이라는 단어도 처리하는 검색엔진을 이용한다면 검색결과는 훨씬 정확해서 부적합한 검색결과의 양이 줄어들게 되므로 사용자는 원하는 정보를 얻기가 더욱 쉬워진다.

그러나 여기에서 다음과 같은 문제가 발생하면 안 된다. <빨간 선인장>에 관한 정보를 검색할 때, "빨간 화분에 담겨 있는 선인장을 보면서..."와 같은 문장이 든 문서가 "선인장이 빨갳다", "빨간 빛깔의 선인장", "선인장이 빨갳게 보인다" 등이 있는 문서보다 먼저 올라오는 경우이다. 이는 '빨간'이 가지는 형용사의 개념을 이해하지 못하고 독립된 명사로 처리를 했기 때문에 생기는 문제이다. 그러나 '빨간'은 활용이 가능한 형용사이므로 단어의 형태가 변한다고 해도 '빨갳다'라는 속성은 변함이 없으므로 '빨간 선인장'과 같은 의미로 처리해서 검색해야 한다. 뿐만 아니라 아름다운 선인장, 시든 선인장, 마른 선인장 등도 구분되어야 한다. 단지 '선인장'이라는 질의에 대해 빨간 선인장이라는 문장이 있는 문서와 같은 가중치의 검색결과로 제시되어서는 안 된다.

결국 '빨간 선인장'이 갖는 속성을 만족시키는 것이 가장 정확한 결과로 우선적으로 제시되어야 한다. 즉,

단어가 갖는 각각의 개념을 정확하게 이해해야 한다는 뜻이다. 하지만 현재 키워드형 검색 엔진으로는 이러한 것을 처리하기가 어렵고 자연어처리를 하는 검색기술이 개발되어야 한다. 이러한 기술을 기반으로 등장한 것이 진정한 의미의 자연어처리형 검색엔진이다.

3. 내용기반 멀티미디어 정보 검색

약 6~7년 정도 전부터 영상정보검색에 대한 논의가 영상처리 전문가들에 의해 본격적으로 시작되었다. 영상처리기술을 이용한 내용기반 영상정보검색이라는 기술이 학문적으로 이슈가 되면서 많은 사람들이 영상정보검색에 관심을 갖기 시작했다. 마침 최근에 멀티미디어 데이터가 서서히 보급되기 시작하고 있으며 아울러 이에 따른 통신 등의 기술이 멀티미디어의 보급을 지원할 수 있도록 환경이 형성되어가고 있다.

초창기에는 마치 내용기반 멀티미디어 정보검색 기술로 모든 멀티미디어 자료를 완벽하게 찾을 수 있을 것이라고 기대를 했다. 그러나 이 논의가 본격화된 지 수년이 지난 오늘날 기술적으로 어렵다는 것이 증명되기에 이르렀다. 물론 최근 특정 도메인에서는 특정 알고리즘을 사용하여 어느 정도 수준의 정확한 결과를 얻을 수 있는 정도는 가능하다. 하지만 초창기에 기대를 하던 정도에는 못 미치는 수준이다.

원래 내용기반 멀티미디어 정보검색은 사진이나 비디오 속에 있는 내용을 컴퓨터가 이해하고 이를 기억하고 있다가 그와 내용이 같은 질의가 들어왔을 때 해당 결과를 찾아 주는 것을 의미한다. 즉 어떤 사진 속에 탱크라는 물체가 있다면 그와 다른 기종, 다른 각도에서 찍은 탱크 사진이 들어와도 탱크를 찾아 주어야 하는 것이다. 그러나 실제로 이러한 과정은 이미지 이해라는 분야의 기술적 한계 때문에 현재는 완벽하지 못한, 나름대로의 결과를 제시하고 있는 실정이다. 이에 따라 현재 개발된 기술은 이미지의 특징(feature) - 색상, 질감, 모양 등의 정보 등을 이용하여 이미지를 검색하는 데에 사용하고 있어서 처음에 의도했던 내용기반 정보검색의 의미는 퇴색하였다고도 볼 수 있다.

멀티미디어 자료에 대한 표준을 주로 다루었던 MPEG-1, 2 그리고 4에 이어서 MPEG-7에서는 멀티미디어 정보검색을 위한 표준을 다루고 있으며 이를 바탕으로 현재는 멀티미디어 프레임워크에 대한 표준인 MPEG-21에 대한 논의가 시작되고 있는 상황이다. MPEG-7에서는 멀티미디어 정보검색을 위하여 멀티미디어 데이터 특징의 종류를 정하고 그것들을 어떻게 표현할 것이며 그 표현방법의 표준을 정하고 있다. 오는 2001년 9월에 최종 표준이 완성될 것이며 현재 우리

나라에서는 산학연이 협조하여 매우 활발하게 이 분야에 대한 표준화를 추진하고 있다.

내용기반 영상정보검색 초창기에는 색상, 질감, 모양 등을 어떻게 잘 추출해내고 잘 검색하는가에 연구의 초점이 모아졌었다. 그러나 최근에는 연구분야가 보다 다양하고 세분화되고 있어 영상 정보의 저장 및 검색이라고 하기보다는 관리라는 보다 폭 넓은 차원의 이슈가 연구되고 있다. 예를 들어 문자 다중 방송 등에서 제공되는 문자 정보를 어떻게 바로 색인 정보로 이용할 수 있을 것인가, storage hierarchy를 포함한 저장 구조의 문제, 유사도에 대한 성능 측정, 검색 결과의 효율적인 브라우징, 사용자 인터페이스, 메타검색엔진, 지적소유권 및 보안에 관한 문제 등을 들 수 있다. 또한 사용자 피드백을 받아서 그것을 검색 엔진에 적용하는 'Relevance Feedback' 이슈도 텍스트 검색 분야에서만큼 중요한 문제이다.

4. MPEG-7

지금의 MPEG 표준화 중에서 MPEG-1은 CD-ROM 등의 저장(1.5MBps)을 위한 표준화이고 MPEG-2(4~80MBps)는 디지털 TV 등의 고선명 데이터를 위한 표준화이며, MPEG-4는 저전송선로인 64KBps수준의 전송 속도내에서도 동영상 데이터의 송수신이 가능한 표준을 정한 것이며 마지막으로 MPEG-7은 데이터 그 자체가 아닌 데이터의 내용에 대한 특징표현 방법을 다루는 것이며 이를 'Meta Data' 또는 'Bits about bits' 라고 표현하기도 한다.

이는 Still Image, Video, Audio, Graphic, 3D Model, Speech 등의 Audio-Visual 정보를 대상으로 하고 있으나, HTML, SGML, RDF 등의 문서의 표현에 대한 것, Feature Extraction 방식 및 Search Engine 등은 표준화의 대상으로 하지 않는다.

다음은 MPEG7의 표준화 사항들이다.

첫째, Descriptor 와 Description Scheme

여기서 Descriptor는 서술자로서 데이터의 특징을 나타내는 도구이며 Description Scheme은 데이터를 여러 개의 서술자로 나타내는 구조를 나타낸다.

둘째는 Description Definition Language 이고 이는 서술자 정의 언어로서 XML과 같은 언어로 표현되고 있다.

셋째는 색인, 저장 및 전송을 위한 Coded Description 등이다.

MPEG에서 취하고 있는 표준화 과정을 살펴보면, 성능이 우수한 제안서들을 바탕으로 통합 시험 모델(Experimentation Model : XM)을 만든다. XM은

MPEG-2에서는 TM(Test Model), MPEG-4에서는 VM(Verification Model)에 해당된다. 이러한 XM을 통해 세부 항목별 공동 실험(Core Experiment)을 통해 성능과 기능 보완. 표준안으로서의 통합 기능 검증 을 거쳐 WD(Working Draft), CD(Committee Draft) 등을 만들어 간다. 이러한 경쟁단계와 협력 단계 (Collaboration Phase)를 거쳐 CD에서 실질적인 기술 적 사항들이 모두 확정되고, 그 이후에는 편집 상의 보완 및 국가별 투표를 거쳐(DIS) 국제 표준(IS)으로 확정된다.

5. 멀티미디어 정보검색의 응용분야

키워드 기반의 검색에서 사용자가 원하는 단어를 입력하여 웹 문서를 찾듯이 원하는 Color, Shape, Edge 등의 특징만을 입력하는 것으로 원하는 자료를 찾을 수 있다. 이는 다양한 분야에서 활용할 수 있는 기술이

며 예를 들어 특정 배우의 이름을 입력하면 그 배우가 등장하였던 사진,영화나 드라마에서 등장하였던 장면 들을 쉽게 찾아 볼 수 있다.또다른 예로 특정 가수의 한 구절의 가사만으로 그 가수에 대한 정보와 음반 및 뮤직 비디오등의 자료를 얻을 수 있는 것이다.이러한 기술을 일반 사용자들이 웹상에서 이미지와 비디오등의 멀티미디어 검색이 가능하게 개발한 일부 사이트가 실험적으로 운영되고 있으며 이를 상용화하는 단계까 지 개발한 상태이다.

하지만 현재 MPEG 7에서는 여러가지의 응용분야 를 한 종류의 Descriptors 와 Description Scheme 으로 만족시킬 수 는 없기에 응용분야에 맞추어 표준화할 계획이다.

이를 응용분야와 관련하여 진행되고 있는 모델이 Pull Model과 Push Model로 구분할 수있다.

우선 Pull Model이라 함은 local 이나 네트워크에 위치 한 DB로부터 원하는 자료를 제공하는 분야이고 Push Model은 컨텐츠 제공자로부터 자료가 제공될 때



그림 1. 색상 기반 이미지 검색

이를 filtering하여 원하는 자료만을 제공하는 분야이다. 하지만 Push Model같은 경우 real time으로 자료를 제공하기에는 시간적 제약이 뒤따른다.

구체적으로 Pull Model에 해당하는 분야를 언급하게 되면 다음과 같다.

- 비디오영화, 라디오 데이터베이스 저장과 검색
- 전문적인 이미지, 비디오 제작과 전송서비스
- 상업적인 음악 시장(karaoke와 음반 판매)
- 효과적인 음향 도서관
- 역사적인 연설등의 오디오 데이터 베이스
- 언론 분야(이름, 음성 또는 얼굴 정보를 사용하여 어떤 정치가의 연설 장면을 찾을 경우)
- 특정 데이터의 저장과 검색
- 교육에서의 멀티미디어 자료 이용
- 역사 박물관, 미술품 전시 등의 문화적인 서비스
- 게임 검색, 가라오케에서의 노래 검색 등 오락 분야
- 사람의 특징 인식 등 범죄 수사
- 지도 제작, 생태 조사, 자연 자원 관리 등의 원격 탐사
- 원격 구매
- 건축, 부동산 내부 장식 관련 정보 검색

다음은 Push Model로서 다음과 같은 응용분야가 해당이 된다.

- 사용자의 원하는 자료에 대해 전문적인 검색서비스
- 개인의 특성화된 TV 방송 서비스
- 인텔리전트한 멀티미디어 제공서비스
- 사용자의 개별화로 검색, 필터링, 브라우징 서비스
- 여행 정보 안내, 지리 정보 서비스

5.1 이미지 내용 기반검색의 활용

이미지에 대한 내용기반 검색 기법들은 이미지의 color, shape, edge 등의 특징을 가지고서 원하는 이미지 영상을 찾을 수 있다. 이를 위해 우선 상당한 양의 이미지를 데이터 베이스화하여 각각의 이미지가 가지고 있는 공통된 특징들을 자동으로 추출하여 인덱싱을 거쳐 데이터 베이스화하여 저장된다. 이런 상태에서 사용자는 자연어, 또는 color, shape, sketch 하여 query(질의)를 주면 자연어의 경우 description의 text 정보에서 찾아 해당되는 이미지를 보여지게 된다. color의 경우 query image의 color histogram 값과 인덱싱된 특징 추출 DB와 비교하여 가장 가까운 값을 가진 이미지를 순서대로 열거하여 보여지게 된다.

이러한 대략적인 이미지 내용기반 검색은 정확성면에서 WD에 채택된 기술을 적용하여 만족할 만한 답을 출력하는 기술을 보여주고 있다

5.2 비디오 내용 기반검색의 활용

현재 비디오 데이터와 함은 압축된 상태의 mpeg 파일과 스트리밍 상태의 asf, rm 형태의 동영상일 것이다. 이러한 동영상을 검색하려면 우선 동영상의 장면 전환 점을 찾아 내어 이를 사용자에게 보여주며 사용자가 임의로 찾고자하는 프레임을 클릭하여 전체 동영상에서 원하는 영상의 길이만을 클리핑하여 편집하는 기능을 제공할 수 있는 기술 수준에 와있다. 단순한 검색정보만을 제공할 때에는 대부분 동영상의 장면 전환 프레임을 나타내어 웹 브라우징을 사용하는 방법이 많이 사용된다.

그리고 장면 전환 점을 찾는 방법중에는 압축된 상태에서 장면 전환점을 찾는 방법과 압축 비디오를 display 하면서 uncompressed 된 상태에서 장면 전환 점을 찾는 방법이 있다. 장면전환점을 찾는 방법은 color, edge, luminance 등을 사용하여 찾을 수 있다. 또한 이러한 방법을 사용하여 이미지 검색에서와 마찬가지로 비디오 검색에서도 사용자의 임의적인 특징 성분의 query로 원하는 동영상속에서의 프레임을 찾을 수 있는 것이다.

그리고 비디오 내용기반 검색을 위하여서는 시간의 흐름 구조에 따라 구성이 되어있는 동영상을 논리적이고 물리적인 방법으로 재구조화 되어야만 한다. 또한 의미적인 구조를 위하여 사용자가 임의대로 동영상의 구조를 편집하여 story board라는 구조를 생성시킬 수 있다.

동영상내 객체의 움직임은 찾아내는 연구로 VideoQ 라는 시스템에서는 사용자의 움직이는 스케치(animated sketch) 질의를 하게 해준다. 이 시스템은 웹기반 비디오 검색이 가능한 시스템으로 객체와 배경이 대조가 잘되는 야구, 축구, 골프, 스키 등의 조건에 좋은 검색결과가 나온다.

초창기의 영상정보검색엔진은 주로 박물관, 백과사전 등의 사진을 검색하는 데에 사용이

되다가 야후나 라이코스 등에서 범용 검색엔진으로 채택하여 운용 중이다. 즉, 초기에는 주로 일반 사용자들을 위한 검색에 주로 사용이 되었는데 최근 비디오 색인 및 검색엔진을 중심으로 인터넷 방송 제작에서 사용되거나 일반 방송국 내에서의 콘텐츠 제작에 사용되고 있다. 한 단계 더 나아가 영상검색 기술은 인터넷 서비스의 수익 모델을 증진시키는 데에도 사용이 될 수 있다. 뿐만 아니라 디지털 방송 계획 등으로 방송 자료를 디지털화해야 할 필요성이 생기면서 영상정보의 저장 및 검색에 대해서 눈을 돌리기 시작했다.

그러나 앞서 전술한 내용기반 멀티미디어 정보 검색 기술이 ABC, CNN, CBS, BBC, C-SPAN, PBS 등 수많은 외국방송사 및 콘텐츠 업체의 예를 들어 볼 때

방송사의 인터넷 사이트나 인터넷 방송사이트에서의 영상정보 검색에 사용된다기 보다는 방송의 제작에 널리 쓰이고 있다. 즉 기존의 방송제작과정에서는 그림 2와 같이 입력, 편집, 포스팅 과정에서 파일을 관리하는 수준이었지만 콘텐츠의 양이 늘어나서 관리할 대상이 많아짐에 따라 그림3과 같은 식으로 SQL 및 키워드 기반의 비디오 자료의 저장, 검색이 필요하게 된다. 국내의 일부 방송국에서 이같은 기능의 비디오 관리 시스템을 운영하고 있다. 그러나 향후 콘텐츠가 디지털화됨으로써 그림4와 같이 보다 고급의 비디오 관리 시스템이 필요하리라 예측된다.

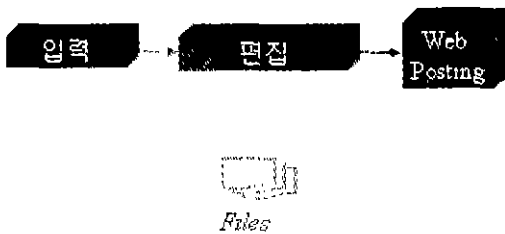


그림 2. 초창기의 비디오 웹 포스팅 과정

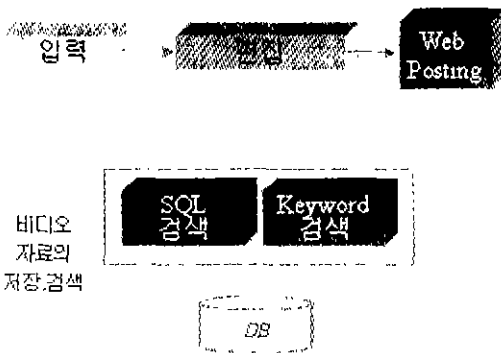


그림 3. 초기 비디오 관리 시스템을 이용한 비디오 웹 포스팅 과정

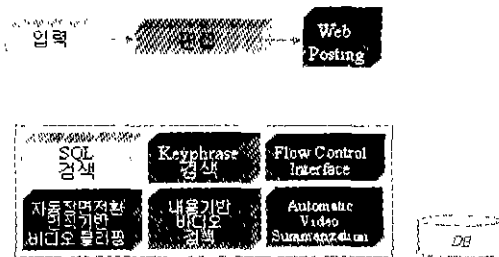


그림 4. 비디오 관리 시스템을 이용한 비디오 웹 포스팅 과정

실제로 선진국의 일부 방송사에서는 이같은 개념의 비디오 관리 시스템을 도입하여 사용하고 있다. 이 비디오 관리 시스템을 이용하면 영상 자료의 비디오 클리핑을 위하여 장면 전환 자동 인식 엔진을 이용하여 장면이 바뀌는 곳의 첫 번째 프레임들을 찾아낼 수 있다. 여기에서 해당 부분만 잘라낸다면 굳이 비선형 편집기를 사용하지 않고도 해당 범위의 비디오 데이터를 전체에서 잘라낼 수 있어서 바로 웹으로 전환할 수 있다. 만일 다른 비디오 클립과 이어 붙이는 작업을 하고 했을 때에도 굳이 비선형 편집기와 같은 시스템을 열지 않고서도 간단한 편집기를 이용하여 원하는 비디오 클립을 만들 수도 있을 것이다. 이렇게 만든 자료를 보관했다가 추후 이를 이용하기 위해서는 검색을 할 필요성이 있는데 원활한 검색 및 관리를 위하여서는 메타 데이터를 함께 입력하여야 한다. 그러나 주석정보의 입력은 주저하다시피 시간과 노력을 요하는 작업으로서 이를 해결하기 위해서는 문자 다중 방송에서 제공되는 문자 정보의 자동 입력, 시나리오, 대본 등의 자동 입력 등이 요구되며 시스템 구축시 이러한 장비와 기능도 함께 구축할 필요가 있다. 이때 텍스트 검색의 결과는 문서의 일부가 제시된다. 그러나 비디오 검색에서는 현재 이주 간결한 정도만 제공이 되는데 만약 스토리보드나 preview 기능까지 제공한다면 검색의 효율성을 훨씬 높일 수 있을 것이다.

인터넷 서비스 측면에서는 비디오 데이터의 내용을 모두 제공함으로써 네트워크의 구축 없이 사용할 수 있다. 즉 요약 화면이나 스토리보드만 제공함으로써 디스크 및 네트워크에 대한 부하도 줄이면서 사용자 하여금 보다 빠르고 정확하게 찾고자 하는 비디오 데이터에 접근할 수 있도록 도와줄 수 있다. 나아가 요약 비디오는 인터넷 상에서 preview를 하고 비디오 대역에서 비디오를 빌려 볼 수 있도록 하는 사업 모델에도 사용될 수도 있을 것이다.

5.3 오디오 내용 기반검색의 활용

오디오에 대한 내용기반 검색방법은 브라우징과 인덱스 등의 방법이 있다. 인덱스 검색 방법은 오디오 데이터를 특징벡터로 인덱싱하여 사용자가 오디오 데이터를 질의하면 원하는 정보를 보여주는 방법과 질의된 오디오 데이터를 음성인식하여 키워드로 전환하여 결과를 음성이나 텍스트로 나타낸다. 이러한 오디오 내용기반의 응용에는 노래방과 karaoke 등에서 특정 부분의 멜로디만으로도 원하는 곡과 가수를 찾게 해준다.

5.4 비디오, 문자, 오디오정보를 사용한 내용기반 검색

동영상에서 문자와 오디오 정보를 이용하여 효과적인 멀티미디어 검색시스템을 구성할 수 있다. 실질적으로 현재의 멀티미디어 정보들은 어느 하나의 데이터만이 아니라 동영상과 그속의 캡션(closed caption, open caption), 오디오 정보 등의 복잡한 형태의 데이터가 대부분이다. 그러므로 이들 문자정보나 오디오를 이용한 내용기반 검색은 현실적으로 커다란 필요성을 충족시켜 줄 기술인 것이다. 동영상 속의 문자정보를 이용한 방법은 이들 문자정보를 인식하여 인덱싱하여 키워드나 자연어 질의를 통해 검색결과를 보내줄 수 있다. 위의 기술을 가지고 방송 뉴스, 교육용 자료 등을 일정 단어들과 해당하는 영상을 검색결과로 보여줄 수 있는 것이다.

6. 결어

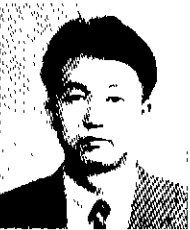
이상에서 멀티미디어 정보검색의 구성요소, 기술개발의 흐름, 향후 발전 방향 및 활용에 대하여 살펴보았다. 초창기에는 검색기술로 시작하였는데 최근에는 검색기술을 바탕으로 비디오 제작 및 응용 시스템 개발 플랫폼으로 발전해 나가는 추세에 있다. 결국 방송자료의 저장 및 효율적인 검색은 콘텐츠를 보유하고 있는 방송사 등의 장기적인 해결 과제이다. 이러한 기능은 방송제작의 여타 환경과 효율적으로 연동되어야 하며 최근 이러한 요구사항을 만족시키는 비디오 저장, 관리, 검색시스템이 개발 및 출현하고 있다.

필자소개



유성준

- 미국 시리큐스대 전산학 (박사)
- ETRI 지식정보 검색팀장, 책임 연구원
- 현 서치캐스트(주) 기술 연구소장



김익선

- 송실대학교 전자공학과
- 현 서치캐스트(주) 멀티미디어 기반연구팀장