# SBS 디지털 비디오 검색 시스템

□ 박현주*/ *SBS 기술연구소 방식연구팀

## 1. 서 론

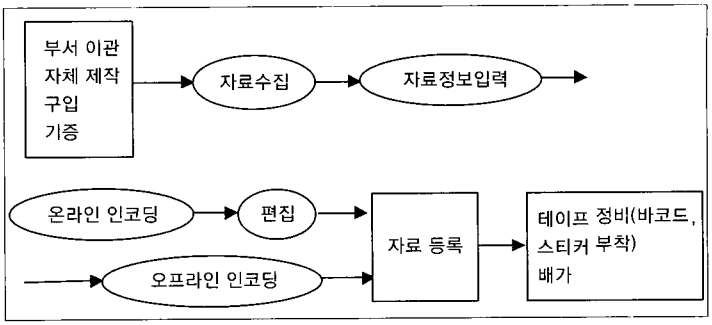
방송과 통신 및 컴퓨터의 융합이 그 어느 때 보다 가속화 되고 있다. 이러한 시대적 흐름 속에서 방송사는 기존에 제작되었던 방송 프로그램들 및 촬영 소재 들을 디지털 컨텐츠화함으로써 향후 다양한 분야에서 사용될 수 있는 미디어 자산으로의 가치를 극대화시킬 수 있다. 또한, 컨텐츠의 메타 데이터 정보를 데이터 베이스에 저장하고 이를 웹 서버와 연동시킴으로서 보다 효율적인 검색을 가능케 하며 자료의 이용도 신속 / 용이하게 할 수 있다.

SBS 기술연구소는 이러한 요구하에 디지털 비디오 검색 시스템을 개발하여 실제 현장에서 사용하고 있다. 본 서에서는 이 시스템의 개발 내용에 대한 소개를 하고저 한다.

## 2. 개발 배경

현재 SBS에서 테이프 형태로 보관하고 있는 영상 자료의 양은 약 42만개에 달한다. 여기에 음반과 효과 및 도서 자료까지 포함한다면 그 양은 실로 엄청나다. 이들 자료를 이용하기 위해서 사용자들은 그동안 웹상의 키워드 검색을 통해 원하는 후보 테이프들을 선정하고, 자료실에서 이들을 대여, 재생시켜 봄으로써 원하는 자료를 찾아내는 일들을 해야만 했다. 여기에서 생기는 문제점으로는 무엇보다 웹상에서의 검색과 실제 테이프를 재생시켜가며 검색을 하는 작업이 중복되고 많은 시간을 요한다는 것이다. 또한 반복적으로 사용되는 자료의 경우 테이프의 화질 열화도 피할 수 없는 일이었다. 자료 관리자의 측면에서 볼 때도 설치될 당시에는 기능이 훌륭했던 시스템이, 시대의 변화에 따라 네트워크 및 PC의 성능은 갈수록 향상되어 가는데, 이에

발맞추지 못해 느린 데이터 베이스 및 검색 엔진 그리고 자료 통계 및 테이프 관리에 있어서 여러가지 불편한 점들이 있어 업무상 비효율성을 가져올 수 밖에 없었다. 이러한 현실적인 상황과 영상 콘텐츠의 디지털화 및 이들의 통합적인 관리에의 필요성 그리고 인터넷과의 연동을 통한 새로운 멀티미디어 서비스의 창출등에의 요구로 새로운 디지털 비디오 검색 시스템을 개발하게 되었다.



〈그림 1〉 자료수집 및 등록 업무 흐름

## 3. 개발 인원/일정

▶ 개발 인원 : SBS 기술연구소 방식 연구팀 3인

▶ 개발 일정

| 기 간 | 내 용 |
|---|---|
| 2001. 6~2001. 7 | 프로젝트 검토 및 시스템 구성 세부 검토 |
| 2001. 8~2002. 1 | 프로토타입 시스템 구축 및 S/W 개발 |
| 2002. 2 | 프로토타입 시스템 개발 완료 |
| 2002. 3~2002. 4 | 시스템 통합 시험 및 문제점 보완 |
| 2002. 5. 2. | 서비스 개시 |

## 4. 개발 요구 사항

### 4.1 업무 흐름

영상 데이터의 디지털화에는 인코딩 과정이 추가됨으로써 자료 관리자들의 업무 증가를 가져올 수 있다는 점을 고려하여, 프로그램 자료들의 인코딩에는 on-line 입력을 통해 방송되고 있는 자료들을 자동으로 인코딩할 수 있게 하였다.

### 4.2 시스템 요구 사항

▶ 현재 구축되어 있는 사내 LAN을 최대한 활용한다.

▶ 기존에 입력된 메타 데이터는 새로운 시스템에서 가공하여 재사용 한다.

▶ 네트워크 상에서 동영상의 검색이 가능해야 한다.

▶ 이미 등록된 영상 자료는 자동으로 동영상 인코딩이 가능해야 한다.

▶ 사용자는 H/W나 S/W의 설치 없이 간단하게 검색이 가능해야 한다.

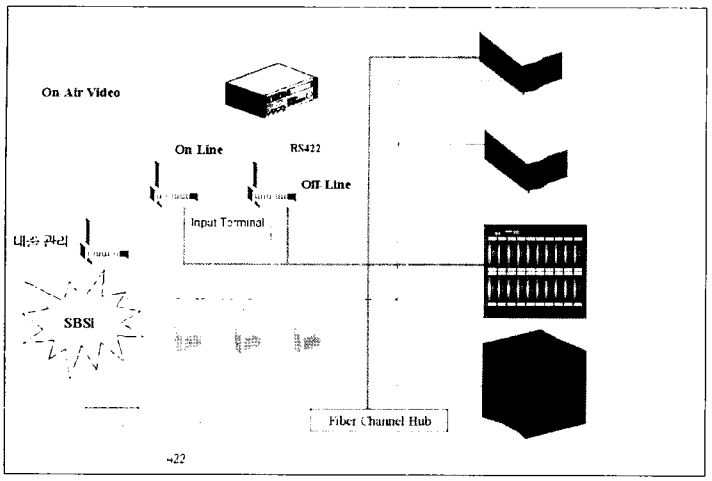
▶ 동영상의 인코딩 포맷은 어느 곳에서도 재사용이 가능한 표준 포맷을 사용한다.

▶ 사용자 관리에 있어 영상 자료 사용의 권한에 대한 보안 등급을 설정할 수 있어야 한다.

▶ 아카이브 시스템은 향후의 확장성을 고려하여 설계한다.

## 5. 시스템 구성

디지털 비디오 검색 시스템은 2단계의 개발 과정을 거쳐서 개발할 것을 고려하였으며, 이번 개발은 그 1차 단계로 설정되었다. 그것은 무엇보다 고비




〈그림 2〉 SBS 디지털 비디오 검색 시스템 구성

용이 소요되는 시스템 구축에 있어서의 시행착오를 줄이면서 충분한 사전 검토 작업을 하기 위함이었다.

### 5.1 스트림 서버

웹상에서 동영상을 검색하기 위해서는 필수적으로 저해상도의 동영상 압축이 필요하다. 현재 웹상에서 가장 많이 사용되고 있는 동영상은 MPEG-4 포맷이다. MPEG-4는 스트리밍이 가능한 압축 포맷으로 MS의 wmv파일 포맷으로의 변환이 가능하다. 현재 220kbps의 압축률로 동영상 인코딩을 하고 있다. 동영상을 요구한 클라이언트 측에 동영상 서비스를 수행하는 스트림 서버는 현재 MS의 Windows 2000 Server에서 제공하고 있는 Media Service S/W를 사용한다. 스트림 서버의 사양은 다음과 같다.

- INTEL PENTIUM 4 CPU
- 500MB RAM
- OS : Windows 2000 Server

### 5.2 DB 서버

DB 서버 및 DBMS 소프트웨어는 범용성이 있고 안정성이 검증된 시스템으로 선택하였다. DB서버의 사양은 다음과 같다.

- INTEL PENTIUM 4 CPU
- 500MB RAM
- OS : RED HAT LINUX
- DBMS : Oracle 8i

데이터 베이스 서버에는 텍스트 검색을 위한 고성능의 검색 엔진을 탑재하였다. 검색 엔진은 문서의 규모가 수십만에 달하는 점을 감안하여 현재의 상용 RDBMS의 쿼리문으로는 검색 요구를 실시간으로 처리하는 것이 어렵다는 판단하에 설치하였다. 검색 요구에 대한 질적인 판단은 응답 시간(response time)과 산출량(throughput)면에서 살펴볼 수 있다. 응답시간은 사용자가 쿼리문을 입력하여 결과를 받기까지의 시간을 말한다. 수십만건의 문서와 인덱스를 하나의 DB에 저장하여 검색할 경우 쿼리문 처리에 수십초 혹은 수분의 시간이 경과하므로 검색 서비스의 제공에 장애가 생긴다. 산출량은 하나의 서버가 단위 시간당 얼마나 많은 쿼리문을 처리할 수 있느냐를 말하는 것으로서, 검색 포탈과 같은 대용량 검색 시스템의 경우 피크 타임 때에는 초당 수십 혹은 많게는 수백개의 질의가 들어오므로 초당 하나씩 처리한다 해도 항상 수백대의 서버가 대기하고 있어야 한다는 얘기가 된다. 상용 RDBMS를 사용하는 경우 대용량 환경에서 발생하는 문제인 응답시간과 산출량 저하의 문제를

해결하기 위한 경제성 있는 방법을 찾기는 당분간 어렵지 않을까 생각된다. 이러한 점을 고려하여 쿼리 기반의 검색이 아닌 스크립트 인터페이스를 갖춘 상용 검색 엔진을 도입하게 되었다.

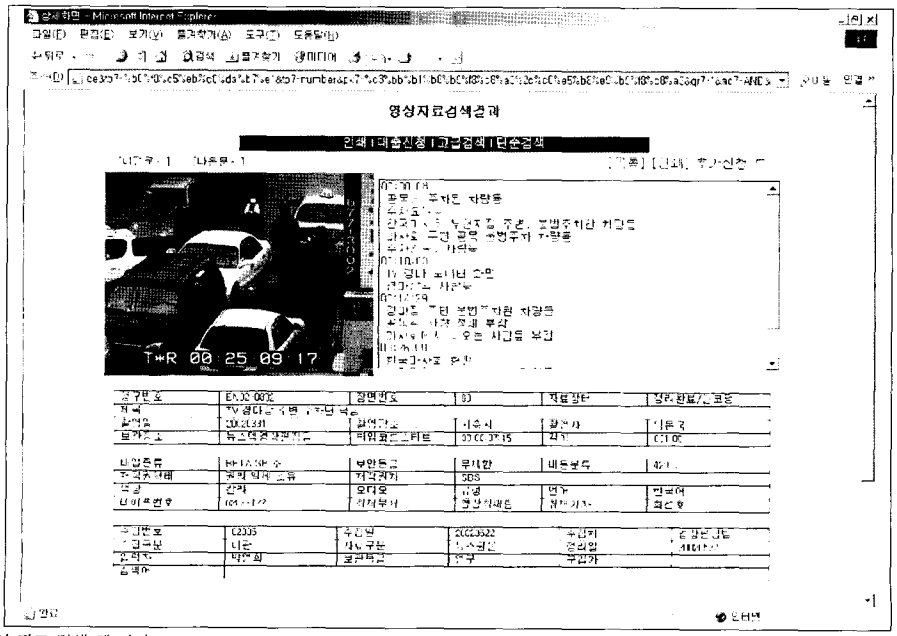
### 5.3 웹서버

웹서버를 통해 사용자들은 찾고자 하는 자료에 대한 텍스트 기반의 검색과 동영상 검색을 할 수 있다. 또한, 원하는 자료를 찾았을 때 기존에 신청서 양식을 손으로 써서 대출 신청을 하였던 것에서 탈피하여 웹상에서의 대출 신청이 가능하게끔 하였다. 웹서버의 사양은 다음과 같다.

- ▶ INTEL PENTIUM 4 CPU
- ▶ 500MB RAM
- ▶ OS : RED HAT LINUX
- ▶ Web Service S/W : Apache

### 5.4 SAN (Storage Area Network) 및 FC Storage

향후 MPEG-2 인코딩 및 스트리밍 서비스 등을 위해서는 데이터의 고속 전송을 가능케하는 네트워크가 필수적이다. 이를 위해 Fiber Channel 네트워크를 구성하였으며, SAN 제어 S/W로 IBM의 SAnergy를 사용하였다. SAN은 Fiber Channel Hub를 중심으로 각각의 Host들이 F/C Host Adaptor를 통해 고속의 데이터 전송을 하고 있으며, MDC(Meta Data Center)기능을 하는 서버쪽으로는 LAN을 통해 SAN 메타 데이터를 전송한다. Storage는 F/C 접속이 가능한 제품을 선정하여 설치하였다. 초기 용량은 MPEG-4 200kbps를 기준으로 하여 5000시간 저장이 가능하도록, 500GB로 시작하여 향후 요구에 따라 저장 용량을 점차 확장해 나갈 예정이다.




〈그림 3〉 영상 자료 검색 웹 서비스

# 6. S/W 구성

## 6.1 개발 환경

이번 프로젝트의 S/W 개발을 위한 TOOL 및 기반 S/W들은 다음과 같다.

- ▶ Windows 2000 Professional, Windows 98
- ▶ Visual C++ 6.0 SP4
- ▶ MS ADO(ActiveX Data Objects)
- ▶ MS Windows Media Encoder SDK 7.0
- ▶ Oracle Client 8i

## 6.2 DB 관련 S/W

- DB 테이블 설계

DB 테이블의 설계는 다음과 같은 기준을 갖고 하였다.

- ▶ 자료의 중복성 배제
- ▶ 무결성(Data Integrity) 보장
- ▶ 저장 용량의 최소화
- ▶ 처리 효율의 극대화
- ▶ 이해하기 쉬운 구조
- ▶ 다수의 쿼리문 처리를 위한 Stored Procedure 사용

테이블 디자인시에는 무엇보다 중복된 자료를 배제하는 것을 주요 원칙으로 하였다. 방송 자료의 경우 하나의 테이프 안에는 다수의 장면들이 존재하는데 만일 이들을 하나의 테이블로 처리할 경우 장면 정보마다 중복되는 테이프 관련 정보들이 있게 된다. 이러한 중복을 막기 위해 테이프 테이블을 기준으로 장면 테이블을 따로 설계하여 1:n의 구조를 갖게 하였다.

- 기존 DB의 Migration

기존에 사용하고 있던 DBMS는 IBM의 DB2였다. 테이블의 필드가 필요에 따라 임의적으로 중복되게 사용되고 있었고, 예전에 사용하던 테이블이 따로 존재하고 있는 등 데이터의 일관성이 많이 결여되어 있었다. 따라서 예전의 데이터를 새로운 DB에 Migration하는 작업이 쉽지만은 않았으며, 많은 시행 착오를 거쳤다.

## 6.3 OFF LINE 영상 입력 단말 S/W

자체 제작 하였거나 구입 또는 기증 등을 통해 수집된 비디오 테이프는 자료 관리 부서로 이관되어 메타 데이터를 DB에 입력하고, 저해상도의 MPEG-4 영상으로 인코딩하여 Storage에 저장한다. 또한 폐기되거나 손망실로 유실된 테이프의 정보는 DB에서 제

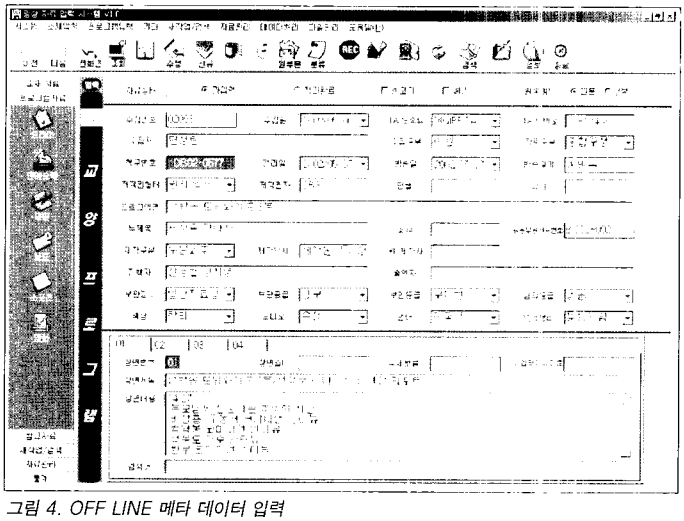
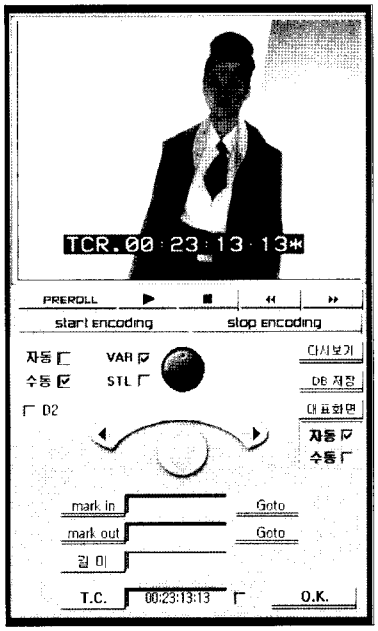


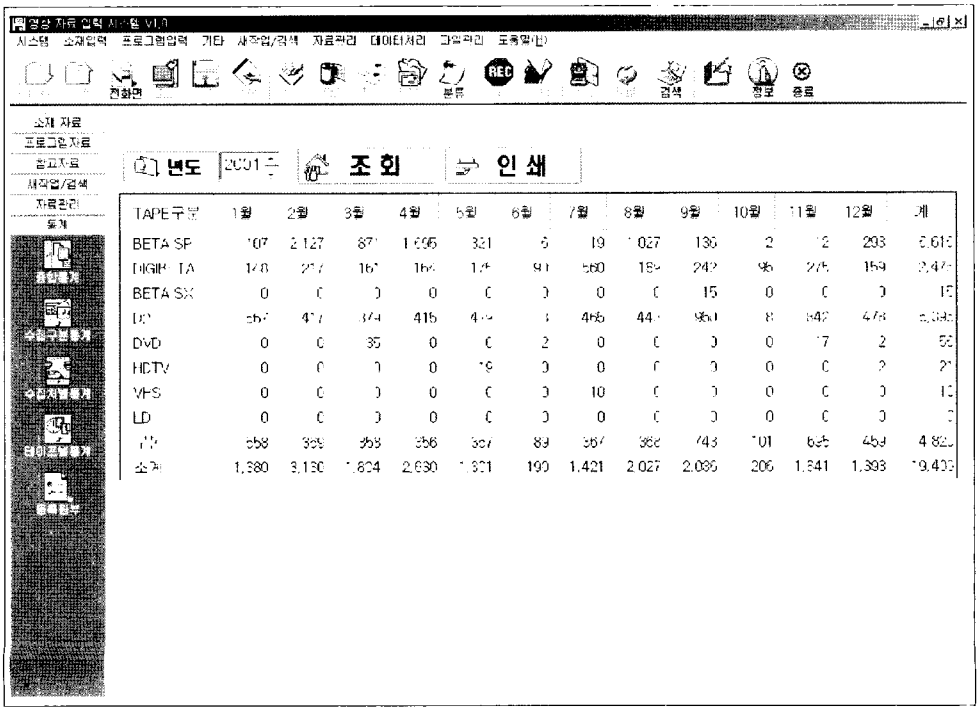
그림 4. OFF LINE 메타 데이터 입력



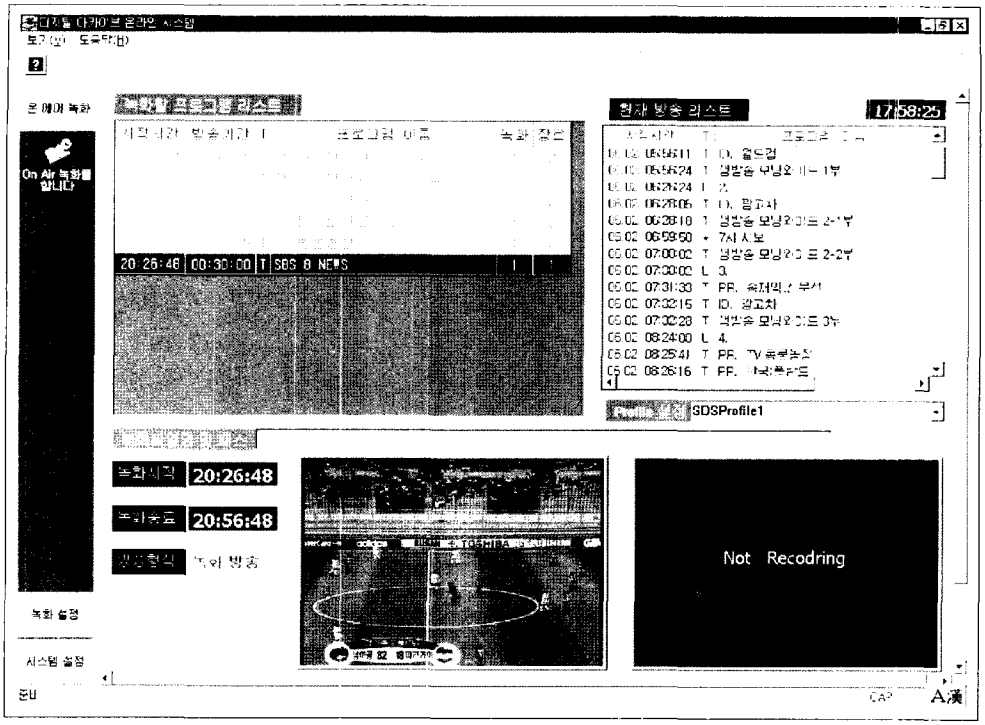

〈그림 5〉 OFF LINE 영상 인코딩

거되어 따로 관리 되는데 이에 대한 지원도 가능해야 하며, 자료의 정리시 일련번호나 보관장소 등의 변경 등도 신속하게 가능하도록 지원되어야 한다. OFF LINE 입력 S/W에서 개발한 모듈들은 다음과 같다.

- ▶ Oracle DB Interface 모듈 : 각종의 쿼리문의 실행이 가능하다.
- ▶ MPEG-4 인코딩 및 캡쳐 보드 제어 모듈
- ▶ VTR 제어 모듈 : SONY 프로토콜의 지원으로 자동 편집이 가능하다.
- ▶ USER INTERFACE 모듈
- ▶ 메타 데이터 입력 모듈 : 자료와 관련된 정보들을 입력하고 변경할 수 있다.
- ▶ 자료 일괄 수정 모듈 : 폐기 처리 및 일괄 수정이 가능하다.
- ▶ 통계 처리 모듈 : 자료의 이용 통계를 리포트한다.




〈그림 6〉 OFF LINE 통계 조회




〈그림 7〉 ON LINE 영상 입력 단말

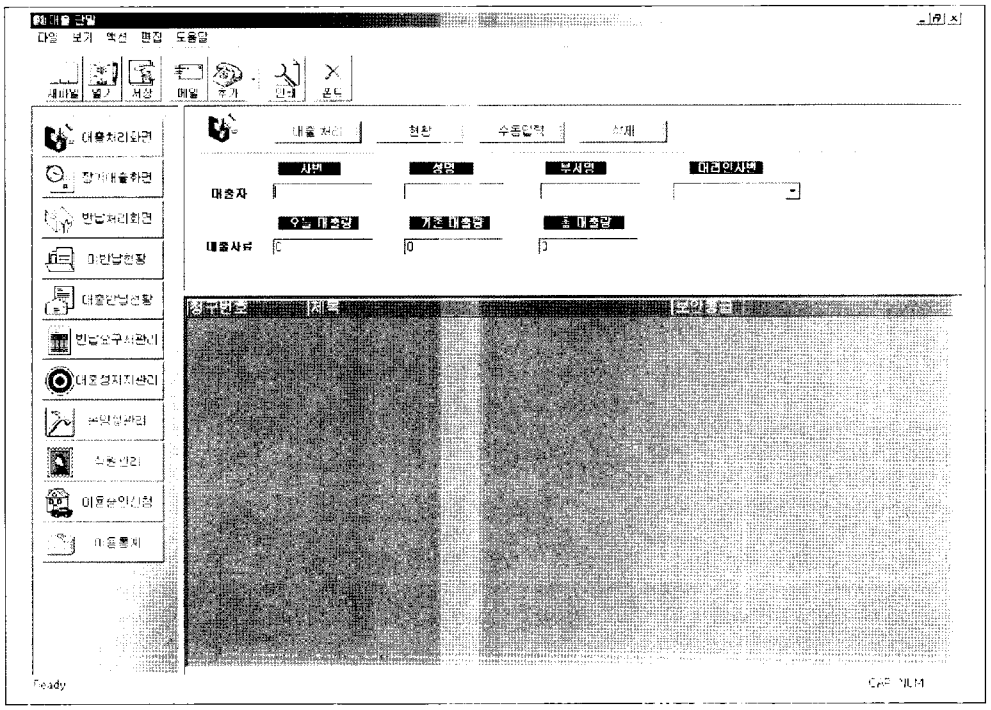
## 6.4 ON LINE 영상 입력 단말 S/W

생방송인 프로그램들을 중심으로 방송되고 있는 소재들을 바로 MPEG-4로 인코딩하여 보다 빠른 시간 내에 디지털 컨텐츠를 생산할 수 있게끔 하였다. ON LINE S/W는 주조정실의 자동운행 시스템의 데이터 베이스에 연결하여 편성표를 읽어 와서 프로그램별로 MPEG-4인코딩을 할 수 있게 한다. 개발 모듈은 다음과 같다.

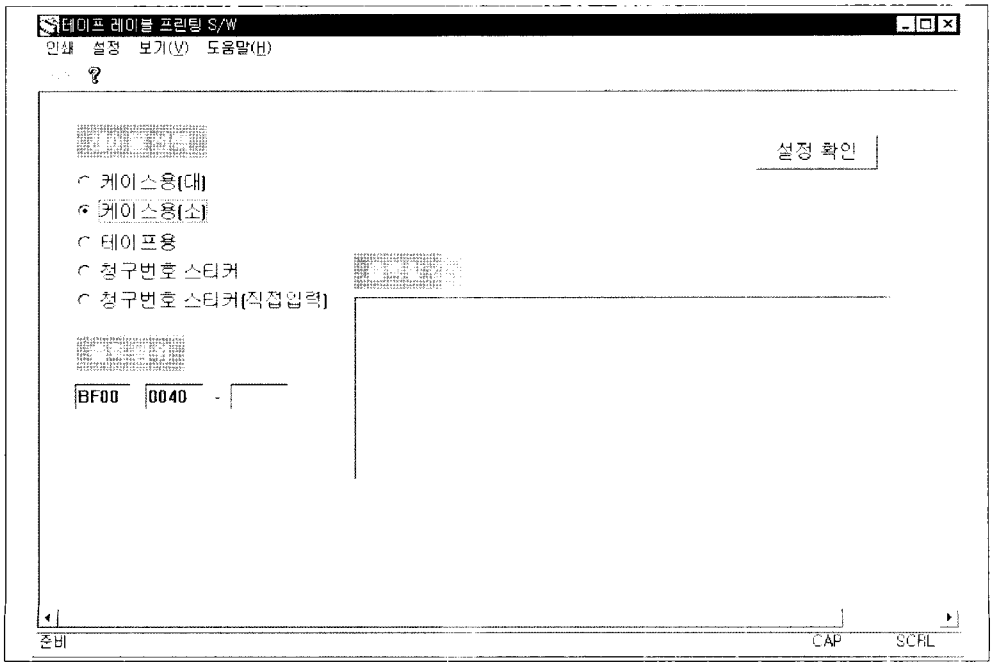
- ▶ Oracle DB Interface 모듈 : 각종의 쿼리문의 실행이 가능하다.
- ▶ SQL 2000 DB Interface 모듈 : 주조의 자동 운행 DB에 접속 편성표를 읽어온다.
- ▶ MPEG-4 인코딩 및 캡쳐 보드 제어 모듈
- ▶ USER INTERFACE 모듈
- ▶ 메타 데이터 입력 모듈 : 자료와 관련된 정보들을 입력하고 변경할 수 있다.

## 6.5 대출 관리 단말 S/W

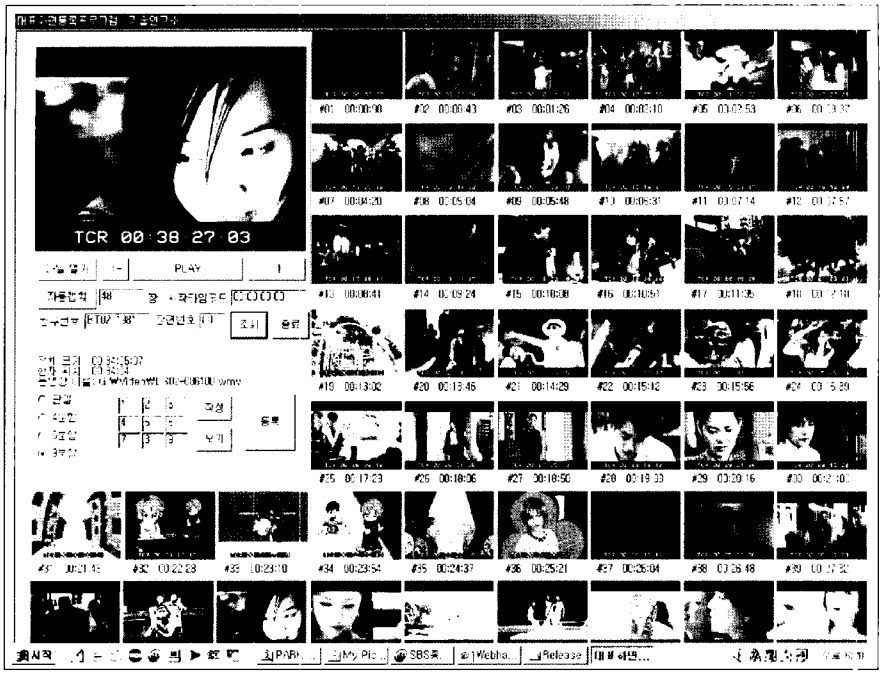
자료를 원하는 사용자들을 관리하고 대출 관련 데이터 처리를 위해 개발되었다. 신청자들의 사원증에 있는 바코드를 읽어들여 이들의 신원을 확인하고 대출 승인을 해 주며, 테이프 이용 상황에 대한 통계 처리를 리포트 해 줌으로써 대출 현황에 대한 통계를 알 수 있게 해준다. 또한, 대출 관리에서는 미반납자에 대한 관리도 중요한데, 미반납자에 대한 이메일 고지 등을 가능케 한다. 개발 모듈은 다음과 같다.




〈그림 8〉 대출 관리 단말 S/W




〈그림 9〉 테이프 라벨 프린트 S/W




〈그림 10〉 대표화면 등록 S/W

- ▶ Oracle DB Interface 모듈 : 각종의 쿼리문의 실행이 가능하다.
- ▶ USER INTERFACE 모듈
- ▶ 통계 처리 모듈 : 자료의 이용 통계를 리포트한다.
- ▶ 대출 및 반납 관리 모듈
- ▶ 예약 대출 및 미반납 현황 처리 모듈
- ▶ 대출 정지자 관리 및 손망실 처리 모듈
- ▶ 바코드 Reader 인터페이스 모듈

### 6.6 기타 S/W

- 테이프 라벨 프린트 S/W

테이프에 붙일 라벨을 프린트 한다. 사용하고 있는 테이프 사이즈에 따라 필요한 라벨을 적절한 사이즈로 프린트 할 수 있다.

- 대표화면 등록 S/W

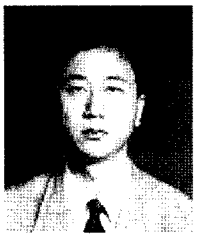
인코딩 된 MPEG-4파일의 대표화면을 등록하는 기능을 한다. 일정 시간 단위로 MPEG-4영상을 추출하여 이중 대표 화면을 선택할 수 있게 한다.

## 7. 맺음말

디지털 비디오 검색 시스템의 개발을 통한 기반 기술의 축적과 자체 개발을 통한 비용 절감 및 경제적인 시스템의 구축은 이번 프로젝트의 소중한 자산이 되었다. 무엇보다 어려웠던 점은 기존의 시스템이 존재한다는 것을 항상 고려해야 한다는 점이었다. 차라리 처음부터 새로이 시작되는 시스템이라면 오히려 어려움이 덜 했을 것이다. 디지털 비디오 검색 시스템은 향후 MPEG-2 동영상 인코딩 시스템의 도입과 테이프 라이브러리의 도입 및 고속의 네트워크 구성을 통한 직접적인 고화질 영상의

전송이 주요한 과제가 될 것이다. 뉴스룸 시스템과의 연계 또한 앞으로의 중요한 이슈가 될 것이다. 이 모든 것들이 갖추어 진다면 방송사의 제작 스텝들은 테이프를 대여하러 수고스럽게 갈 필요 없이 자신의 자리에서 동영상을 검색하고 원하는 영상을 불러와 자유롭게 편집하는 일이 가능해 질 것이다. 디지털 컨텐트 자산의 확보는 효율적인 영상 자료 관리와 함께 증대될 다양한 디지털 컨텐츠 공급의 핵심 인프라가 되어 방송사의 경쟁력 확보와 수익 창출에 지대한 영향을 끼칠 것이다.

## 필 자 소 개



**박 헌 주**

- 1995년 : 서울대학교 전기공학과 졸업
- 1995년 : SBS 입사, TV 기술 1부 근무
- 1996년 : 기술연구소 근무
- 현재 : SBS 기술연구소 방식연구팀 선임연구원
- 주관심분야 : DTV, 비디오 아카이브