

인터넷을 이용한 영상자료 검색시스템 개발

Development of Image Database Using Internet

이세경 · 이인성
서울시립대학교 조경학과

Lee, sekjung · Lee, Insung
Dept. of Landscape Architecture, The University of Seoul

I. 서론

대용량 DB와 디지털입력기술의 발달은 과거에는 생각하지 못하였던 영상자료의 디지털 데이터베이스화를 가능하게 하고 있다. 영상 데이터베이스는 자료의 훼손을 방지할 수 있을 뿐 아니라 효과적인 검색과 배포를 가능하게 해준다. 최근 인터넷의 발전은 이러한 이미지 데이터베이스를 네트워크를 통해 수많은 사람들이 동시에 접속하여 사용할 수 있게 하여 그 효용가치와 활용잠재력을 더욱 높이고 있다. 영상자료를 많이 활용해야하는 조경학 분야에서는 이러한 영상 데이터베이스의 구축과 활용은 매우 중요하다.

현재 웹(Web)상에 활용 가능한 영상데이터베이스는 매우 다양한데, 그 중 AMORE(www.ccril.com/amore/)는 NEC USA에서 개발된 영상 검색 엔진으로 30,000개의 이미지와 300개의 해당 관련 이미지 웹사이트로 연결을 시켜준다. 이미지들은 웹페이지에서 유사한 이미지나 유사한 문맥 또는 키워드로 찾을 수 있도록 구성되어있으며 효율적인 검색을 위해 6개의 분류항목이 제공된다. 콜롬비아 대학의 WebSEEK(disney.ctr.columbia.edu/webseek/)은 650,000여개의 이미지와 10,000여개의 비디오가 수록되어 있다. 검색은 16개의 카테고리를 이용한 주제별 검색과 해당 관련 단어를 이용한 검색이 가능하다. Yahoo의 Image Surfer(ipix.yahoo.com/)는 분류항목을 이용한 주제별 검색이 가능하며 주제를 이용하여 해당되는 자료를 찾을 수 있도록 구성되어 있다. 검색결과는 작은 이미지(thumbnail)로 보여지며 이를 클릭하면 해당관련사이트로 접속이 되어 해상도가 높은 이미지를 얻을 수 있다.

본 연구에서는 서울학연구소 「서울성장 50년사 영상자료 탐사」의 일환으로 서울시 공보와 사진실에서 보관하고 있는 1957-1994년 사이의 서울시의 주요 행사, 사건 및 경관변화를 수록한 수십만장의 사진 자료 중 가치가 높다고 판단되는 사진을 선별하여 이를 디지털 데이터베이스로 구축하여 CD-ROM과 인터넷 검색시스템으로 개발하였다.

II. 연구방법

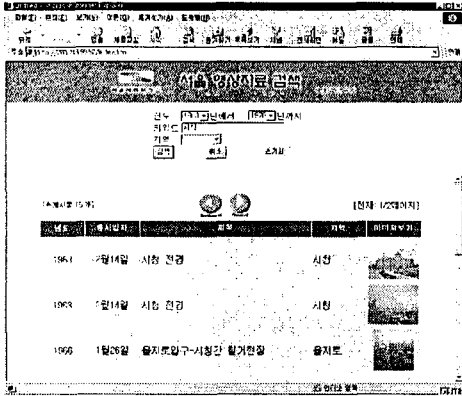
서울시의 사진자료 중 서울시의 성장과정과 생활사, 주요 행사 및 역사적 사건, 등 서울의 특징적인 이미지를 보여 주는 사진자료 2,617장을 선별하고 고품질 스캐너와 Photoshop을 이용하여 음화를 직접 읽어들이어 디지털 데이터화하였다. 사진 필름들은 4×5", 120mm, 135mm, Slide 등 다양한 크기를 가지고 있었으나, 입력시에는 원판의 크기와 관계 없이 입력된 이미지의 크기가 비교적 일정하게(장축이 대략 1,000 pixel 정도로) 유지되도록 하였다.

입력된 파일은 저장공간의 효율성이 높고 압축에 따른 손상의 정도가 비교적 적은 JPEG 포맷으로 저장되었으며, 파일 이름은 촬영날짜를 기준으로 분류입력되었다. 사진자료 목록은 '연도', '날짜', '제목', '지역', '주제어', '파일이름' 등 여섯 항목으로 작성되었다. 촬영일자는 서울시 목록을 그대로 수용하였으나, 제목은 사진내용과 현저히 다른 경우가 많이 발견되어 관련전문가의 자문과 확인 작업을 거쳐 수정, 입력하였다. 목록자료는 초기에는 Excel을 이용하여 작성되었으며, MS Access DBMS 파일로 전환시킨 후 외부 접속을 위하여 ODBC로 연결해 주었다.

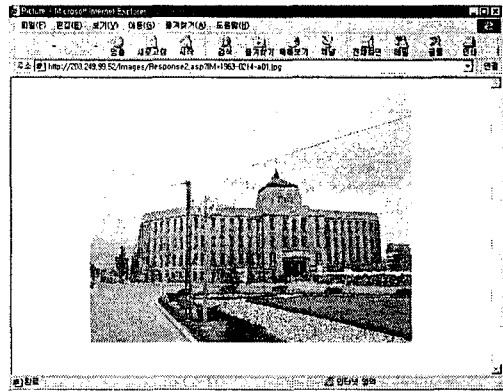
영상자료의 검색시스템은 인터넷과 CD-ROM의 두가지 매체를 대상으로 개발되었다. 인터넷 검색시스템은 MS의 웹서버 개발도구인 ASP(Active Server Page)와 Visual InterDev를 이용하여 개발되었으며, 서울시립대학교의 GIS/CAD 연구실에서 시험운용을 하여 문제점을 보완하였다. 인터넷 검색시스템은 서울시와 사진자료의 배포에 관한 저작권 문제가 해결되지 않아 아직 외부로 공개되지는 않고 있다. CD-ROM 검색시스템은 Visual Basic을 이용하여 개발되었다.

III. 영상자료 검색 시스템

인터넷과 CD-ROM 검색시스템은 동일한 인터페이스로 구성되어 있다. 시작화면은 데이터베이스에 대한 간단한 설명과 검색방법에 대한 안내문을 보여준다. 영상자료를 검색하는 방법에는 단일검색 및 다중검색이 있는데, 단일검색은 연도, 검색어, 지역 중 하나를 기준으로 검색하는 방법이며, 다중검색은 연도, 검색어, 지역 중 둘 이상의 기준으로 검색하는 방법으로 AND검색이 적용된다. <그림 1>에서와 같이 연도를 '1963'년에서 '1970'년으로, 검색어를 '시청'으로 입력하고 '검색'을 누르면 '1963'년에서 '1970'년 사이의 '시청'에 관련된 사진들을 작은 이미지로 보여준다. 검색된 목록에서 사진을 보고 싶을 때는 해당되는 작은 이미지를 눌러 <그림 2>와 같이 고해상도의 이미지를 볼 수 있다.



<그림 1> 인터넷 검색시스템 결과 화면



<그림 2> 검색된 영상자료의 전체화면

IV. 결론 및 향후 연구방향

디지털 데이터베이스화는 자료의 영구보존과 편리한 활용에도 중요한 목적이 있지만 이를 통해 정보가 효과적으로 공유될 수 있다는 점에 더욱 큰 의미를 부여할 수 있다. 서울시 성장 50년사 영상자료는 우리나라 공공기관이 소장한 영상자료 중에서는 최초로 디지털 데이터화된 사례로 생각된다. 공공자료의 데이터베이스 작업은 국가 정보인프라의 중요한 바탕이며, 앞으로 여러 기관이 소장한 방대한 자료들이 이러한 방식으로 데이터베이스화되고 필요로 하는 학계와 일반인에게 공개되어 활용된다면 국가 생산력과 연구역량을 높이는 귀중한 밑거름이 될 수 있을 것이다.

특히 조경학과 같이 영상자료가 교육과 연구에 필수적으로 활용되는 분야에서 이러한 영상 데이터베이스를 구축, 활용하는 것은 중요한 과제로 생각된다. 만약 통합적인 영상 데이터베이스가 구축된다면 현재의 슬라이드를 사용하는 영상교육의 내용적 충실성과 준비의 효율성을 한층 개선하여 그 차원을 한단계 높일 수 있을 것이다.

본 연구는 영상데이터베이스 구축작업보다는 시범시스템 제작을 통해 영상데이터베이스의 가능성과 효과를 확인하고 이에 따른 문제들을 파악하여 앞으로의 본격적인 구축 작업의 방향을 설정하고 이에 필요한 경험과 교훈을 얻는 데 보다 근본적인 목적이 있었다. 제작과정에서 발견된 문제와 보완되어야 할 사항들을 정리해 보면 다음과 같다.

- 다양한 문자 및 영상 데이터베이스와의 접합: 영상 데이터베이스에서 영상자료뿐 아니라 영상에 대한 설명과 관련자료를 동시에 제공할 수 있으면 보다 충실한 정보의 전달이 가능해진다. 현재 여러 기관에서 신문기사의 목록화 작업이 이루어지고 있는데 영상 데이터베이스의 검색 결과에 관련 기사목록이 즉각 참조 가능하도록 부가되고 신문기사의 검색시스템에 관련 영상이 참조가능하도록 부가된다면 양쪽 데이터베이스의 효율을 동시에 높일 수 있을 것이다.

- 검색기능과 사용자 인터페이스의 보완: 현재의 검색시스템에서는 단일 검색과 다중 검색 등 간단한 검색방법만을 제공하고 있는데, 앞으로 저장 자료의 수가 많아지면 간단한 검색 방법으로는 충분하지 않을 것이며 보다 다양한 검색기능이 제공되어야 할 것이다. 아울러 본 제작작업에서는 CD-ROM 검색시스템과 인터넷 검색시스템이 각각 다른 개발도구로 구축되었으나 장애는 개발도구가 통일되고 일관성 있는 사용자 인터페이스 체계를 만들어 나가야 할 것이다. 현재의 관련 기술개발의 추세와 인터넷의 확산 속도로 볼 때 양쪽 시스템 모두 인터넷 브라우저를 통해 접근할 수 있도록 통일되는 것이 바람직한 것으로 판단된다.