

# 임베디드 시스템을 이용한 XML 데이터베이스를 기반으로 이미지 검색 시스템의 설계

김 경 수\*

요 약

본 연구에서는 검색의 정확도가 떨어지는 문제들을 해결하기 위해, meaning-based 검색과 content-based 검색을 결합하여 검색 시스템을 설계하였다. 설계된 검색 시스템은 임베디드 시스템에 적용할 수 있도록 했으며, 휴대용 PDA 및 스마트 폰과 같은 휴대용 장치로 이미지 데이터를 검색하고 관리가 가능하다. 아울러 검색 시스템은 멀티미디어 데이터를 효율적으로 활용할 수 있다.

## Design of Image Retrieval System Based on XML Database Using Embedded System

Kyung-Soo Kim\*

### ABSTRACT

This study to solve these problems a search system has been designed by combining the two methods. Also the search and manage image data by handheld devices such as portable PDA or smart phone, a system has been designed application to an embedded system. Once this is used, multimedia data can be efficiently searched and utilized by handheld devices.

Key words : MPEG-7, Audio/Image Retrieval System, Multimedia Data

---

접수일 : 2009년 6월 1일; 채택일 : 2009년 6월 18일

\* 백석문화대학 컴퓨터정보학부

## 1. 서 론

최근 디지털 카메라, 캠코더 등 컴퓨터에 연결 할 수 있는 멀티미디어 기기들이 일반 국민들에게로부터 큰 호응 얻고 있다. 또한, 인터넷의 개방형 네트워크 시스템으로, 사용자가 클릭 버튼으로 멀티미디어 데이터를 쉽게 접근하고 공유할 수 있다[7]. 멀티미디어 데이터는 자기 자신 그 자체로, 특이하고, 높은 용량을 포함하고, 부정형(불규칙)적인 자연에 관련된 정보를 수반하지만, 텍스트 데이터와 달리, 효율적인 방법으로 관리하기는 어렵다. 그런 관점에서, 멀티미디어 데이터는 효과적으로 데이터를 관리하기 위해, 텍스트-based 쪽으로 더 많이 skewed된 현재 가능한 검색 보다 검색 방법과 도구에 다른 접근법이 필요하다[5].

이 논문은 이미지 검색 시스템이 설명되어있다. 이미지의 예 그리고 시맨틱-based의 조사와 키워드의 조사를 통한 콘텐츠 베이스의 통합을 통해 그리고 XQuery 검색 기능과 강력한 XML 메커니즘을 제공하여, 현재 사용할 수 있는 검색의 한계를 극복하기 위해 노력했다. 스키마 형태를 통해 제공되는 MPEG-7와 같이, 보다 효율적인 운영을 위해, 임베디드 XML 데이터베이스인 Berkeley DB를 XML이 사용될 것이다. 이 논문은 임베디드 시스템과 MPEG-7을 제 2장에서 소개할 것이며, 제 3장에서는 검색 시스템의 기능 및 구조에 대해 설명하고, 마지막 제 4장에서는 결론을 맺을 것이다.

## 2. 관련 연구 및 기술

멀티미디어 데이터는 자기 자신 그 자체로, 매우 독특하고 중요하고, 높은 용량을 포함하고, 부정형(불규칙)적인 자연에 관련된 정보를 수반함으로써, text-based data에 비교해서 데이터를 관리하기는 어렵다. 다음으로, 더욱 텍스트 쪽을 향하여 맞춘 엔진 검색과 닮지 않게, 멀티미디어 데이터에

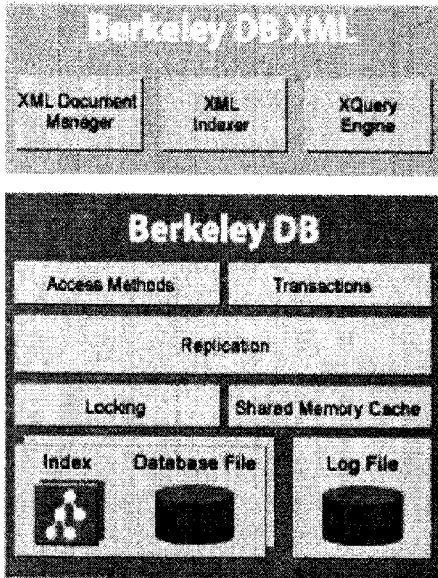
해당하는 접근 방식과 검색 도구를 고안하는데 필요로 한다[5]. 현재 사용 가능한 이미지 검색 시스템에는 그들이 지원할 수 없는 복잡한 쿼리 구조뿐만 아니라, 연구 결과가 정확하지 않을 수 있는 공통적인 문제를 안고 있다[4]. 따라서 XQuery가 위의 문제를 해결할 수 있는 쿼리 구조를 지원할 수 있다[2]. 따라서 높은 함수 기능을 가지고 있는 작은 크기의 장치의 사용 증가와 함께, 임베디드 시스템의 장점을 활용할 수 있는 시스템에 대한 필요성도 극적으로 증가했다.

### 2.1 Berkeley DB

XML 버클리 DB를 XML은 특별한 목적의 데이터 관리지, 네이티브 XML 데이터를 관리하기 위해 버클리 DB의 상단에 내장되어 있다. 이것은 빠르고, 신뢰성, 비용을 절약하는 스토리지가 가능하고 Native XML 데이터와 Semi-구조화된 데이터의 검색을 제공한다. 또한, 버클리 DB XML는 의심할 나위 없이 네이티브 스토리지와 XML 문서의 표준화를 위해 설계되었고, 번역 또는 다른 데이터 구조로 변환하지 않고도 데이터를 저장할 수 있는 임베디드 데이터베이스이다[1].

버클리 DB를 XML은 오픈 소스이다, 윈도우 NT, Unix, Linux 그리고 다른 주요 플랫폼에서 사용할 수 있는 database 엔진이 내장되어 있으며, 작으면서 빠르고, 이해하고 사용하기 쉽고, 신뢰할 수 있다. 이 시스템은 Key/data pair의 검색과 높은 저장 기능을 필요로 하는 응용 프로그램의 수요를 충족 할 수 있는 프로그램에 직접 연결 하도록 허용하는 라이브러리를 제공한다.

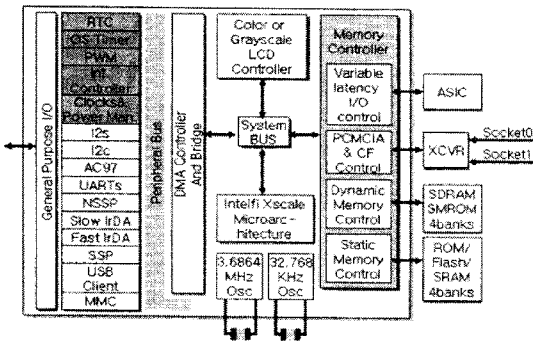
현재, 이걸 C++, Perl, Tcl, Java 그리고 다른 프로그램 언어를 지원할 수 있도록 인터페이스를 제공한다[1]. 더구나, Java, C, C++, Tel, Python 그리고 PHP의 API들에게만 제공하는것이 아니라 유닉스, 리눅스, 윈도우즈 그리고 다른 플랫폼, 인터넷 서버, 데스크탑, 노트북, 그리고 다른 장치에서도 운영 될 수 있다.



(그림 1) Berkeley DB XML

## 2.2 XScale 프로세서

XScale 프로세서는 32-bit 프로세서, ARM5TE core based, 그리고 마이크로프로세서의 인텔 StrongARM line의 후속이다. 인텔 pXA 255은 ALU, 어드레스 버스, 데이터 버스, 그리고 레지스트리 기본 기능들을 하나의 칩에 추가해, 산술 작업, 출력장치, 메모리 제어 및 기타 주변 장치에 대한 레지스트리를 위한 임베디드 장치이다.



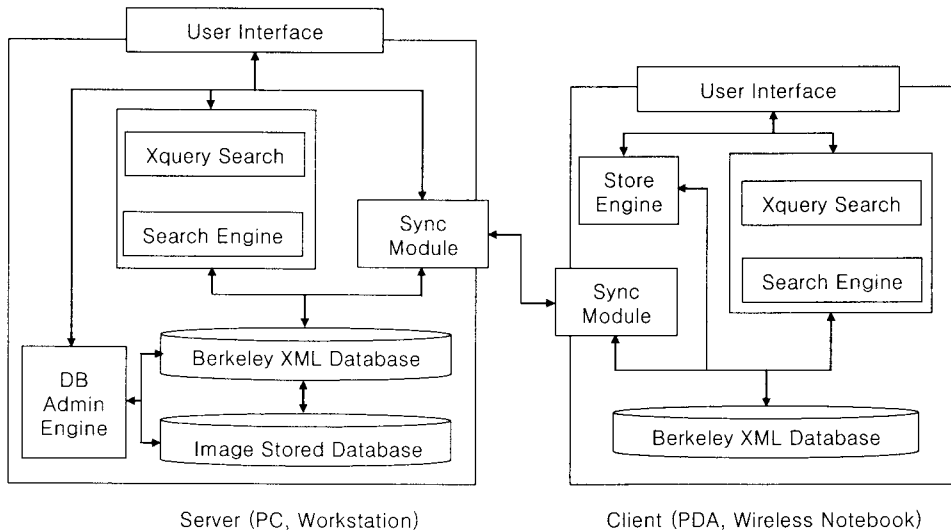
(그림 2) PXA255 CPU 구조

위의 기능과 유사한 마이크로 컨트롤러, 컨트롤을 위해 사용되며, Rom과 다양한 주변 장치의 제어 컨트롤을 위해 필요한 입력 및 출력 장치 임베디드이다. XScale 프로세서는 높은 함수 32-비트 마이크로 프로세서, 0.35um, 3-계층 메탈 CMOS 와 2.6밀리온 트랜지스터를 가지고 있다. 이 프로세서는 다양한 속도로 데이터를 전송할 수 있으며 적은 전력을 소모하고 다양한 캐시 모드와 다양한 주변 기기와의 인터페이스를 지원한다.

## 3. 이미지 검색 시스템 구조

이미지 검색 시스템에서 사용자가 저장된 메타 데이터와 이미지에 대한 모듈을 통해 Berkely DB 검색 기능을 실행을 이 서류에서 설명했다. 검색은 키워드와 이미지의 예를 통합하는 keyword, Xquery, similar image search via example를 제공한다. 본질적으로, Xquery는 XML를 최고로 활용하는 강력한 쿼리 방법이다. 사용자가 데이터베이스에서 관련 데이터에 대한 쿼리를 추구하는 과정을 통해 모색해야 결과가 반환된다. 키워드 검색 기반의 의미론(semantics-based keyword 검색)은 contents based 검색 안에서 불가능하다. 더하자면, vector based에 기반을 둔 example image queries를 통한 검색 이미지 때문에, 비록semantic based 검색을 할 수 없는 것이 단점이지만, content-based 검색은 가능하다. 그렇게 볼 때, 위의 한계를 극복하기 위해, 이 서류는 키워드 검색과 example 이미지 검색을 통합해 모색했다.

게다가, MPEG-7, 멀티미디어 데이터에서의 메타데이터 표준을 적용한 XML 스키마가 Berkeley DB XML을 기반으로 XML 스키마의 효율적인 관리를 위해 사용된다. (그림 3)은 임베디드 시스템에 있는 XML을 DB-기반 이미지 검색 시스템 구조입니다.



(그림 3)

#### 4. 결 론

이 논문에서는 현재 사용 가능한 이미지 검색 도구를 극복하기 위해 노력하여 검색 시스템을 제안하다.

의미-기반 검색은 내용-기반 이미지를 검색할 수 없는 높은 수준의 메타 데이터를 사용하고, 반대로 내용 기반 검색은 의미 기반 검색을 검색할 수 없는 낮은 수준의 메타 데이터를 사용한다. 두 가지 검색 방식의 통합하여 그 부정확한 점들이 반환되는 성과를 가져오게 한다.

더욱이 Xquery와 강력한 XML 메커니즘을, 그리고 XML 스키마의 효율적인 관리를 위해 이미지 검색을 추가했다. 임베디드 시스템과 정보를 제공하는 서버 시스템이 Berkeley DB XML의 위에 있는 네이티브 XML 데이터베이스에 놓여진다.

결과적으로, XML 문서의 분석을 통해 검색 속도 저하를 극복했다. 더 나아가서 Berkeley DB 기능의 활용, 업무, 복원, 그리고 DBMS와 다른 다양한 접근 교환 수단에 의해 이 시스템의 장점을 보낼 것이다.

#### 참 고 문 헌

- [1] Berkely DB XML, <http://www.sleepycat.com/products/xml.shtml>.
- [2] Jan Axelson, Lakeview Research, Embedded Ethernet and Internet Complete : Designing and Programming Small Devices for Networking, Paperback, 2003.
- [3] Open Source Native XML Database, <http://exist.sourceforge.net/>.
- [4] Virginia E.Ogle and Michael Stonebraker, "Chabot : Retrieval from a Relational Database of Images", IEEE Computer, Vol. 28, No. 9, pp. 40-46, September 1995.
- [5] 황명균, 공현장, 김판구 "Internet Information Retrieval : Semantic Image Retrieval Using RDF Metadata Based on the Representation of Spatial Relationship", Journal of Korea Information Processing Society, Vol. 11, No. 5, pp. 573-580.
- [6] 이재호, 김형준, 김혜울, "Study of Image information Processing : MPEG-7 based Video/

Image Retrieval System(VIRS)”, Journal of Korea Information Processing Society, Vol. 10, No. 5, pp. 543-552.

- [7] 신영숙, 김기형, 송해상, 안광선 “Embedded Linux using Intel PXA 255, scitech media, 2004.
- [8] David E. Simon, An Embedded Software Primer, Addison-Wesley, 1999.



### 김경수

2001년 순천향대학교 전산학과  
(공학박사)

2005년~2006년 VCU DBLab  
Visiting Scholar

1998년~현재 백석문화대학  
컴퓨터정보학부 조교수