

SQL/MM Part5를 지원하는 쿼리변환기의 설계 및 구현

강기준*, 이부권*, 서영건**

요약

인터넷 기술의 발전으로 멀티미디어 데이터의 중요성과 활용도가 증가됨에 따라 데이터베이스에서 멀티미디어 데이터를 표현하고 처리하기 위한 연구가 필요하게 되었다. 관계형 데이터베이스에서는 멀티미디어 데이터를 저장하기 위한 단순히 저장구조만을 지원하지만 멀티미디어에 대한 데이터형, 표현, 질의에 대한 지원은 미흡하다. 이러한 문제를 해결하기 위해 ISO/IEC에서 멀티미디어 데이터를 위해 SQL 멀티미디어(SQL/MM)을 표준화하였다. 그러나 SQL/MM은 객체 관계형 데이터베이스에서는 지원하지만, RDBMS에서는 지원하지 않는다. 따라서 본 논문에서는 RDBMS 중 하나인 MS-SQL 2000에서 SQL/MM 지원하기 위해 쿼리변환기를 제안하고, 이를 이용한 이미지 검색 프로그램도 함께 보인다. 쿼리 변환기는 SQL/MM을 SQL로 변환하는 기능 외에 이미지 중복 체크기능을 추가로 제공한다. 그리고 쿼리 변환기를 이용한 이미지 처리 프로그램은 RDBMS 기반의 기존시스템과 쉽게 통합 운영이 가능하다.

A Design and Implementation of a Query Interpreter for SQL/MM Part5

Gi-Jun Kang*, Bu-Kwon Lee*, Yeong-Geon Seo**

Abstract

We need a research for representing and processing of multimedia data in database because of increasing the importance and utilization of the data owing to development of internet technology. RDBMS supports only the storing-structure to store multimedia, but the support for data type, representation and query of multimedia is insufficient. To cope with this problem, ISO/IEC standardized SQL multimedia(SQL/MM) for multimedia data. However, ORDBMS supports SQL/MM, but RDBMS does not support it. Therefore, this thesis proposes a query interpreter to support SQL/MM in MS-SQL 2000 as one of RDBMS and introduces a image retrieval application using it. The query interpreter supports the function to convert SQL/MM into SQL, and additionally the function of the image duplication check. The image processing application using a query interpreter can easily be integrated and operated with traditional RDBMS-based system.

Key words : SQL/MM Part5, 쿼리변환기, 확장저장프로시저

1. 서론

정보 전달 수단으로서 매우 효과적인 멀티미디어 정보는 컴퓨터 성능의 향상, 멀티미디어 기술의 발전, 정보통신 기술의 발전등 배경으로 상당히 발전 되었고, 앞으로 계속 발전될 것이다. 특히, 인터넷의 활용영역이 확대되면서 나날이 더 많은 멀티미디어 정보가 인터넷 환경에서 이용되고 있다. 이런 인터넷 기술의 발전은 예측할 수 없을 정도로 급격한 멀티미디어 정보의 생성과 분포를 가져 왔고 원하는 멀티미디어 정보를 빨리 처리하려는 이용자의 요구는 갈수록 많아지면서 다양해지고 있다[1].

이러한 이용자들의 요구를 충족시키기 위해서 멀티미디어 데이터의 데이터베이스화가 필수적이다. 그러나 멀티미디어 데이터는 개념적으로 복잡한 구조와 대용량이라는 특성을 가지고 있기 때문에 기존의 자료처리 기법으로 처리하기에는 어려운 점이 많다.

그리고 이러한 데이터베이스화를 위해서는 대용량의 자료를 저장하기 위한 저장구조, 멀티미디어 데이터형의 지원, 효율적인 질의처리 기능 등의 많은 요구사항들이 필요하다[2].

이를 위해 ISO/IEC에서 SQL 타입의 기능을 확장시킨 것으로서 영구적인 객체의 생성, 관리, 그리고 질의에 관한 기능을 포함하는 SQL3을 국제표준화 하였다.

특히 SQL3에 포함되어 있는 SQL/MM은 멀티미디어

※ 제일저자(First Author) : 강기준

접수일 : 2005년 4월 5일, 완료일 : 2005년 4월 22일

* 경상대학교 컴퓨터학과 박사과정

kijuny@empal.com

* 경상대학교 컴퓨터학과 교수

** 경상대학교 컴퓨터교육과 부교수

응용을 위한 SQL 클래스 라이브러리인 SQL 추상 데이터 타입(Abstract Data Type, ADT)을 제공한다[3].

이를 지원하기 위해 기존의 RDBMS를 기반으로 객체 지향 모델을 지원하도록 확장하는 방식으로 SQL3에 기반을 둔 Informix의 “Universal Server”, Oracle의 “Oracle8”, IBM의 “DB2”, Universal DB(UDB)’ 등의 제품들이 출시되고 있다[4].

그러나 대부분의 현존시스템에서 사용되는 RDBMS는 SQL3을 지원하지 않으며, 단순히 대용량의 자료를 저장하기 위한 저장구조만 지원한다.

그러다보니 RDBMS에서 멀티미디어 데이터 처리는 단순히 멀티미디어 데이터를 저장해서 각 응용에 의존적인 DB 스키마, 질의어, 프로그래밍 등의 작업들을 각각 처리해야 하는 문제점이 발생한다[5]-[7].

따라서 본 논문에서는 RDBMS에서 SQL/MM Part5를 지원할 수 있도록 SQL 확장을 통한 쿼리변환기를 설계 및 구현하였다. 그리고 쿼리변환기를 이용한 이미지 검색 프로그램도 함께 제안한다.

본 논문에서는 멀티미디어 구성요소 중에 이미지를 고려하였다. 쿼리변환기는 응용 레벨에서 SQL/MM Part5를 SQL로 변환해주며, SQL/MM Part5를 사용함으로써 표준화된 방법으로 이미지 처리를 할 수 있다.

2. 관련 연구

2.1 SQL 멀티미디어(SQL/MM)

SQL3에 포함되어 있는 SQL/MM는 멀티미디어 응용을 위한 SQL 클래스 라이브러리인 SQL 추상 데이터 타입을 제공하는 ISO/IEC의 새로운 국제 표준이다. 이러한 SQL/MM은 과학과 산업에 자주 이용되는 이미지, 사운드, 애니메이션, 음악, 비디오 등의 멀티미디어 개체를 다루기 위한 방법을 제공하는 클래스 라이브러리를 표준화하였다.

SQL/MM은 여러 분야로 나눠 표준화되었다.

분야 1은 프레임워크로서, 다른 분야에서 사용되는 공통개념들의 정의를 제공한다. 분야 2는 보통 문자 데이터로서, full-text 데이터들을 위한 정의를 제공한다. 분야 3은 공간 데이터로서, 기하, 위치, 위상들의 공간 데이터들을 위한 정의를 제공한다. 분야 5는 정지 영상 데이터로서, 그래픽, 그림 사진을 포함한 이미지 데이터들을 위한 정의를 제공한다. 분야 6은 데이터 마이닝으로서, 대용량의 데이터로부터 이전에는 몰랐지만, 유용하게 활용될 수 있는 지식을 효과적으로 찾아내는 지식 탐사의 한 연구 분야이다. 이것은 현대의 데이터 관리에 대한 중요한 관점을 나타낸다[3].

본 논문의 관심부분인 분야 5는 정지영상을 위한 구조화된 사용자 정의 타입과 내용 기반 이미지 검색을 위해 이미지의 특징들에 대한 타입도 제공한다.

그림 1은 SQL/MM Part5를 위한 타입들을 나타내고 있다.

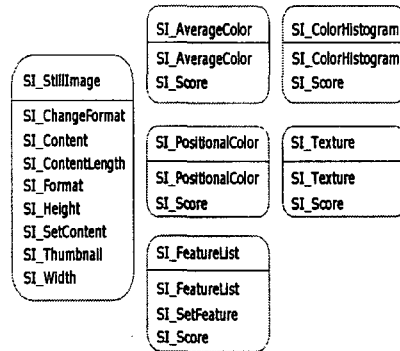


그림 1. SQL/MM Part5를 위한 타입들

SI_StillImage 타입은 JPEG과 같은 이미지 데이터를 검색하는 방식이나 기법을 표현하고 저장하기 위해 정의되었고, 이미지 데이터 자체를 나타내는 SI_Content, 이미지 데이터의 크기를 나타내는 SI_ContentLength, 이미지 데이터에 대한 참조값을 나타내는 SI_Reference, 저장 형식을 나타내는 SI_Format, 이미지의 폭을 나타내는 SI_Width, 이미지의 높이를 나타내는 SI_Height 등의 속성이 있다.

이미지 데이터의 특징을 위한 타입들은 평균 컬러 특성을 나타내는 SI_AverageColor, 컬러 히스토그램 특성을 나타내는 SI_ColorHistogram, 위치 컬러특성을 나타내는 SI_PositionalColor, 텍스처 특징을 나타내는 SI_Texture와 특징들을 조합해서 나타내는 SI_FeatureList로 구성된다.

SQL/MM Part5 표준에 정의된 각 클래스를 구현하기 위해서 각 클래스들의 관계를 고려해야 한다.

그림 2는 이미지 처리를 위해 정의된 클래스들간의 관계를 나타내고 있다.

SQL/MM Part5의 구현을 위해 전역 메소드로서 SQL-invoked regular function을 정의하고 있다. 표 1은 메소드에 대응하는 SQL-invoked regular function을 나타내고 있다[8].

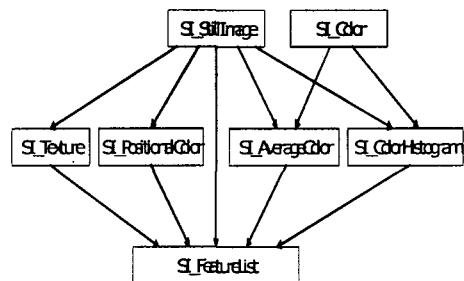


그림 2. 클래스들간의 관계

표 1. 메소드에 대응하는 SQL-invoked 함수

Type Name	Method Name	Parameter Types(if any)	Function Name
SL_StillImage	SL_StillImage	BLOB	SL_mkStillImage1
	SL_StillImage	BLOB, Varchar	SL_mkStillImage2
	SL_StillImage	DATALINK	SL_mkStillImage3
	SL_StillImage	DATALINK, Varchar	SL_mkStillImage4
	SL_setContent	BLOB	SL_chgContent
	SL_changeFormat	Varchar	SL_convertFormat
	SL_Scale	Integer, Integer	SL_scaleImage
	SL_Scale	Double Precision	SL_zoomImage
	SL_Resize	Integer, Integer	SL_resizeImage
	SL_Rotate	Double Precision	SL_rotateImage
	SL_Thumbnail		SL_getThmbnI
	SL_Thumbnail	Integer, Integer	SL_getSizedThmbnI
	SL_content		SL_getContent
	SL_contentLength		SL_getContentLngh
	SL_format		SL_getFormat
	SL_height		SL_getHeight
SL_width		SL_getWidth	
SL_AverageColor	SL_Score	SL_StillImage	SL_ScoreByAvgClr
SL_ColorHistogram	SL_Score	SL_StillImage	SL_ScoreByClrHstgr
SL_PositionalColor	SL_Score	SL_StillImage	SL_ScoreByPstnlClr
SL_Texture	SL_Score	SL_StillImage	SL_ScoreByTexture
SL_FeatureList	SL_Score	SL_StillImage	SL_ScoreByFtrList

2.2 확장 저장 프로시저

일반적으로 RDBMS의 응용분야를 확대하기 위해서, 다양한 객체 타입의 지원이 필수적이다. 다양한 형태의 객체를 데이터베이스에서 표현하기 위한 방법은 ADT 개념을 정의한 후 확장된 SQL 구문을 사용하여 ADT 객체와 연산자를 선언하고, 기존 프로그래밍 언어를 사용하여 함수와 프로시저의 형태로 선언자를 구현한 후, 데이터베이스 시스템에 통합시킨다.

MS-SQL Server에서는 확장 저장 프로시저를 통해 SQL확장을 할 수 있다. 확장 저장 프로시저는 SQL Server 주소 공간에서 실행되며, SQL Server 개방형 데이터 서비스 API를 사용하여 만들 수 있다.

확장 저장 프로시저 사용 방법은 먼저 C나 C++ 프로그래밍 언어로 동적 연결 라이브러리(DLL) 형태로 외부 루틴을 만든다. 그런 다음 고정 서버 역할의 구성원이 SQL Server에 확장 저장 프로시저를 등록하고 다른 사용자에게 이 프로시저를 실행할 수 있는 권한을 부여한다. MASTER 데이터베이스에만 확장 저장 프로시저를 추가할 수 있다.

확장 저장 프로시저는 일반 저장 프로시저와 비슷하며 일반 저장 프로시저와 같은 방식으로 실행된다. 매개 변수를 확장 저장 프로시저에 보낼 수 있고, 결과 및 결과 상태를 반환할 수 있다[9].

다음은 확장 저장 프로시저를 등록하고 사용하는 명령어이다.

```

USE master
-- ESP.DLL에 구현된 xp_ex1이라는
-- 확장 저장 프로시저 등록
EXEC sp_addextendedproc
        'xp_ex1', 'C:\WESP3.DLL'
-- xp_ex1 실행
EXEC xp_ex1
    
```

3. 시스템 설계 및 구현

3.1 구현 환경

본 시스템의 구현환경은 IBM 호환 PC, 운영체제는 Windows 2003 Server 상에서 개발하였으며, 구현언어는 Visual C++ 6.0, Visual Basic 6.0, DBMS는 MS-SQL Server 2000을 이용하였다.

3.2 시스템 구성

전체 시스템은 기존 프로그램과 이미지처리 프로그램으로 나뉜다. 기존 프로그램은 RDBMS를 사용하여 현재 운영되고 있는 프로그램이다.

이미지 처리 프로그램은 이미지 검색이나 이미지 조작 등의 이미지 처리를 수행하는 응용프로그램이다. 전체 시스템은 ADB(Application Database), SL_IPV (Images and Parameter Values), SL_IDD(Images Data Definition), SL_TEMP, MASTER 데이터베이스, 쿼리 변환기로 구성된다.

ADB는 기존 프로그램에서 사용하는 데이터베이스이다. SL_IPV는 이미지와 이미지 속성 값들을 저장하고 있는 데이터베이스이다. SL_IDD은 이미지 처리를 위한 스키마를 저장하고 있는 데이터베이스이다. SL_TMEP는 이미지 중복 체크를 위해 이미지를 임시로 저장하고 있는 데이터베이스이다. MASTER는 이미지 처리를 위해 미리 구현된 확장 저장프로시저를 저장하고 있는 데이터베이스이다. 중복체크는 동일한 이미지의 중복 입력을 방지하기 위해 모듈이다. 이는 사용자가 미리 설정한 적절한 임계값을 가지고 처리한다.

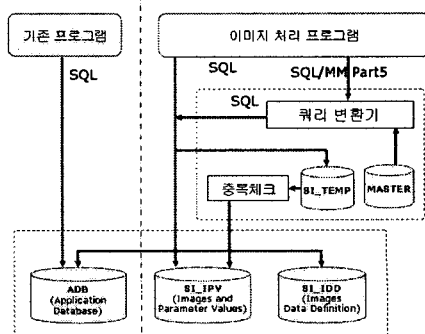


그림 3. 전체 시스템 구성도

쿼리 변환기는 SI_IPV, SI_IDD, MASTER 데이터베이스를 이용하여 SQL/MM Part5를 다수의 SQL과 확장 저장프로시저들로 변환해 준다. 결론적으로 이미지 처리 프로그램에서 SQL/MM Part5를 입력받아서 쿼리변환기가 RDBMS에서 처리할 수 있는 SQL과 확장 저장프로시저로 변환을 해서 RDBMS에서 이미지 처리를 한다.

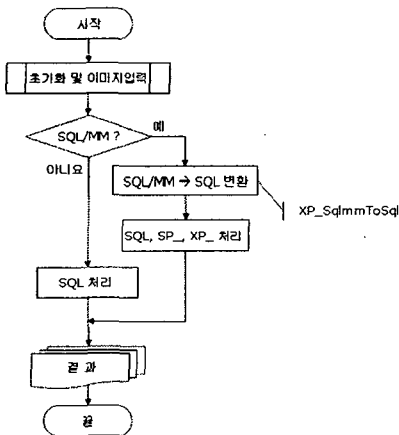
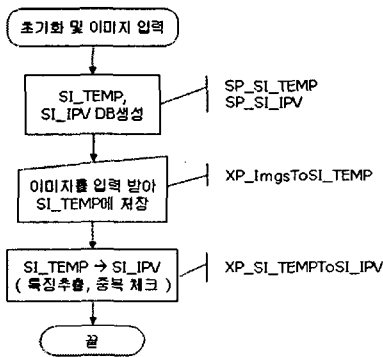


그림 4. 쿼리변환기 작업 흐름도

3.3 쿼리변환기

쿼리변환기는 RDBMS에서 SQL/MM Part5를 지원하기 위해서 SQL/MM Part5를 SQL로 변환하는 작업을 한다.

그림 4는 쿼리변환기 작업 흐름도를 나타내고 있다.

쿼리 변환과정은 "SI_"로 시작하는 토큰이 질의문에 포함되면 SQL/MM Part5로 간주하여 파싱 단계를 거쳐 다수의 SQL, 사용자 정의 저장프로시저, 확장 저장프로시저로 변환 작업을 수행한다.

사용자 정의 저장프로시저는 MS-SQL 상에서 사용자에 의해 정의된 저장프로시저로서 주로 테이블을 생성하는 작업을 수행하며, 프로시저명은 "SP_"로 시작한다. 확장 저장프로시저는 표 1에서 정의된 "SI_"로 시작하는 이미지 처리함수에 대응하는 작업을 하며, 프로시저명은 "XP_"로 시작한다.

표 2는 쿼리변환을 위한 몇 가지 변환 예를 보여주고 있다.

4. 실험 및 평가

4.1 실행화면

그림 5는 본 논문에서 제안하는 쿼리변환기를 이용한 이미지 검색 프로그램의 실행화면을 나타내고 있다.

이미지 검색 프로그램은 질의부분과 출력부분으로 구성되어 있다. 질의부분에서는 텍스트 기반으로 SQL이나 SQL/MM Part5를 입력을 한다. 그리고 출력부분에서는 필요에 따라 쿼리변환 작업을 거친 후에 질의문에 대한 결과를 문자와 이미지로 출력하는 곳이다. 출력부분의 첫 번째 이미지는 질의 이미지를 나타낸다.

표 2. 쿼리변환을 위한 변환 예

SQL/MM	SQL, 저장프로시저
SI_StillImage("img.jpg")	XP_SI_StillImage("img.jpg")
SI_ConvertFormat("img.jpg", "img.bmp")	XP_SI_ConvertFormat("img.jpg", "img.bmp")
SI_GetThmbnl("img.jpg")	XP_SI_StillImage "img.jpg" XP_SI_GetThmbnl img
SELECT * FROM RegLogoes WHERE 1.2 < SI_findTexture(NewLogo).SI_Score(Logo)	XP_SI_StillImage "NewLogo" XP_SI_FindTexture RegLogoes, NewLogo SELECT * FROM RegLoges where 1.2 < RetScore ORDER BY RetScore DESC

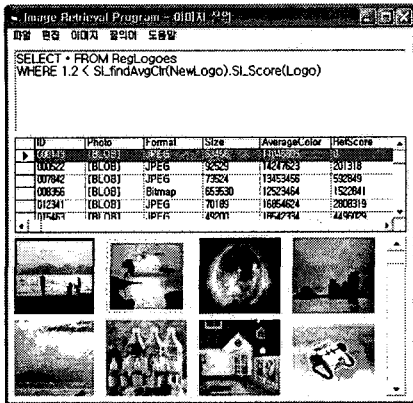


그림 5. 이미지 검색 프로그램 실행화면

4.2 평가

아직까지는 SQL/MM Part5를 완벽하게 지원하는 ORDBMS는 없다. 그렇지만 많은 ORDBMS는 비슷한 방법으로 정지영상을 처리할 수 있는 기능을 제공한다. 하지만 MS-SQL과 같은 RDBMS에서는 SQL/MM Part5를 지원하지 않는다[6].

본 논문에서는 RDBMS에서 SQL/MM Part5를 지원 하도록 하는데 초점을 맞추었다. 기능비교는 SQL/MM Part5를 지원하는 RDBMS는 없기 때문에 RDBMS간의 비교는 어렵다. 따라서 본 시스템과 SQL/MM Part5를 지원하는 ORDBMS의 기능비교를 하였다. 표 3은 이미지 처리를 지원하는 DBMS의 기능 비교를 나타내고 있다[10, 11, 12, 13].

대부분의 ORDBMS에서는 정지영상 조작기능과 내용기반 이미지검색을 위해 표준에서 정의된 average color, color histogram, positional color, texture 특징을 제공한다.

특히 Excalibur Data-Blade와 Oracle interMedia은

표준에서 정의되어 있지 않는 shape 특징을 제공하며, Oracle interMedia는 얼굴 인식도 제공한다.

내용기반 이미지 검색은 우선 특징(들)을 적절하게 이용하여 유사도 검사를 한 후 사용자에게 의해 결정된 적절한 임계값을 이용하여 이미지 검색을 한다.

5. 결론 및 향후 과제

RDBMS의 응용분야를 확대하기 위해서는 시스템의 장점을 유지하는 동시에 새로운 응용분야를 동적으로 적응력이 있는 확장형 데이터베이스의 개발이 효과적이다. 본 논문에서는 이미지 처리를 위한 질의어 표준인 SQL/MM Part5를 RDBMS에 지원하기 위해 MS-SQL의 확장 저장 프로시저를 사용하여 쿼리변환기를 구현 하였다.

쿼리 변환기를 사용한 이미지 처리의 장점은 SQL/MM Part5의 표준화된 이미지 질의어 지원, 확장 저장 프로시저로 구현함으로써 응용프로그램과 데이터 베이스 간에 독립성 제공, RDBMS를 사용하는 현존 시스템을 그대로 유지하면서 이미지 처리를 쉽게 통합 운영할 수 있다. 뿐만아니라 확장 저장프로시저의 추가 구현으로 쉽게 쿼리변환기의 기능 확장이 가능하다.

향후 과제로는 효과적인 이미지 처리 알고리즘과 다양한 색인기법 등에 대한 연구가 필요하다. 그리고 JPEG2000과 같은 새로운 이미지 형식에 대한 이미지 타입 지원과 shape feature와 face recognition 등의 새로운 이미지 특징을 위한 타입 추가가 향후 과제로 남아있다.

(표 3) 이미지 처리를 지원하는 DBMS의 기능 비교

(Table 3) Comparison for function of RDBMS supporting image processing

	SQL/MM 표준	DB2 Image Extender	Excalibur Data-Blade	Image Foundation DataBlade	Oracle interMedia	MS-SQL 쿼리변환기
scaling	√	√	√	√	√	√
rotation		√		√		√
color reduction		√	√	√	√	√
format conversation	√	√	√	√	√	√
cropping				√	√	
sharpening/blurring				√		
average color feature	√	√				√
color histogram feature	√	√	√		√	√
positional color feature	√	√	√		√	√
texture feature	√	√	√		√	√
shape feature			√		√	
face recognition					√	

참고 문헌

- [1] 최윤철, 고견, 임순범 공저, "멀티미디어 배움터", p25-p34
- [2] 이종화, 박유현, 하창석, 김경석, "멀티미디어 데이터를 지원하기 위한 SQL 확장", 멀티미디어학회 논문지, 제 2권, 제 2호, 1999. 6.
- [3] Jim Melton and Andrew Eisenberg, "SQL Multimedia and Application Packages (SQL/MM)", ACM SIGMOD Record, Volume 30, Issue 4, December 2001
- [4] D. Barry, "ODBMS Feature Coverage", Object Magazine, July, 1997.
- [5] Cindy X. Chen, Jiejun Kong, Carlo Zaniolo, Design and implementation of a temporal extension of SQL, IEEE ICDE, 2003
- [6] MyrianR.B.Araujo, Caetano Traina Jr., Agra Traina, Jpsiane M. Bueno, Humberto L. Razente, Extending relational databases to support content-based retrieval of medical images, IEEE CBMS, 2002
- [7] Knut Stolze, "Still Image Extensions in Database Systems System", Datanbank-Spektrum, 2002/2
- [8] <ISO/IEC JTC 1/SC 32/WG VIE-009, December 10, 2001> < 2003-08-15 >
- [9] MS-SQL Server 2000 온라인 설명서, 확장 저장 프로시저
- [10] IBM Corporation : DB2 Universal Database Image, Audio, and Video Extenders Administration and Programming, Version 7, 2000
- [11] Informix Corporation : Excalibur Image DataBlade Module, Version 1.2-User's Guide, March 1999
- [12] Informix Corporation : Informix Image Foundation DataBlade Module, Version 2.00-User's Guide, December 2000
- [13] Oracle Corporation : Oracle8i interMedia Audio, Image, and Video-User's Guide and Reference. 2000

강기준



2000년 경상대학교 컴퓨터과학과 학사
2002년 경상대학교 컴퓨터교육과 석사

2003년-현재 : 경상대학교 컴퓨터과학과 박사과정
관심분야: MPEG, JPEG2000, 영상처리

이부권



1972년 경상대학교 농경제학과 학사
1978년 미시건주립대학 대학원 시스템 공학전공 석사
2001년 경남대학교 대학원 컴퓨터공학과 박사

1980년-현재 경상대학교 컴퓨터과학과 교수
2001년-현재 경상대학교 컴퓨터정보통신연구소원
관심분야: 시뮬레이션, 멀티미디어

서영건



1987년 경상대학교 전산통계학과 학사
1989년 숭실대학교 대학원 전자계산학과 석사
1997년 숭실대학교 대학원 전자계산학과 박사

1989년-1992년 삼보컴퓨터 근무
1997년-현재 경상대학교 컴퓨터교육과 부교수
2001년-현재 경상대학교 컴퓨터정보통신연구소원
관심분야 : 멀티미디어통신, 영상인식