

□ 特 輯 □

멀티미디어 데이터 검색 기법

홍익대학교 전자계산학과 김 경 창*

● 목 차 ●

I . 서 론	IV . 매 치(Matching)
II . 관련 연구	4.1 매칭에서의 문제점
III. 멀티미디어 데이터 검색 시스템의 구조	4.2 도메인-종속 지식
3.1 정의와 배경	4.3 근사 매칭 알고리즘
3.2 구조	V . 요 약
3.3 멀티미디어 데이터에 대한 자연어 묘사	

요 약 : 멀티미디어 데이터베이스에서는 영상(image), 음향(sound) 등과 같은 멀티미디어 데이터의 복잡한 구조로 인해 데이터 조각이 기존의 데이터베이스에서 처럼 단순하지 않다. 데이터베이스로부터 멀티미디어 데이터를 검색하는데 있어서 중요한 것은 사용자의 질의와 멀티미디어 데이터의 내용(content)을 매칭하는 것이다. 일반적인 해결책은 키워드를 사용하거나 멀티미디어 데이터의 내용과 사용자 질의를 묘사하고 있는 텍스트 묘사(text description)를 사용하는 것이다. 이 때의 주된 문제점은 사용자나, 다른 시점의 같은 사용자가 동일한 개체(entity)를 다르게 서술하여 멀티미디어 데이터 내용의 묘사(description)와 사용자 질의의 묘사가 정확하게 매치되지 못하는 데 있다. 그러므로 멀티미디어 데이터를 검색하는 동안에, 저장된 멀티미디어 데이터와 사용자 질의의 묘사들간에 근사(approximate)매치가 통상 요구된다. 본 논문에서는 객체지향 및 자연어 인식 기법을 통합하여 근사매치에 지능적(intelligent)으로 접근하는 방법을 제안한다.

키워드 □ 객체지향 모델, 자연어 인식, 멀티미디어 데이터베이스 정보 검색, 클래스 계층(class hierarchy)

I. 서 론

멀티미디어 데이터베이스 관리 시스템은 기존의 데이터베이스의 기능을 수행할뿐 아니라, 영상, 음향 및 애니메이션과 같은 멀티미디어 데이터를 관리해주는 시스템이다. 오늘날 기술의 발달로 인하여 멀티미디어 데이터를 컴퓨터에 저장하는 것이 가능해 졌기 때문에 멀티미디어 시스템은 많은 관심을 모으고 있다. 멀티미디어는 컴퓨터와 사용자간에 통신의 폭을 넓힌다. 군사, 출판, 컴퓨터 보조 학습과 같은 많은 응용 분야에서, 멀티미디어 데이터를 필요로 하고 있다. 멀티미디어 데이터를 다루는데 필요한 하드웨어의 가격은 급속도로 떨어지고 있는데 비해 소프트웨어는 부족한 실정이다.

본 논문은 멀티미디어 데이터베이스 시스템에서 멀티미디어 데이터의 조각에 초점을 두고 있으며 특별히 근사 매칭 방법으로 멀티미디어 데이터를 검색하는 효과적인 방법을 상세하게 서술하겠다. 기존의 데이터베이스의 경우, 숫자나 영문 데이터의 조각은 데이터 내용을 이용함으로써 처리된다. 멀티미디어 데이터베이스의 문맥에서 접하게 되는 근본적인 문제는 내용을 어떻게 찾아 처리해야 하는가이다. 멀티미디어 데이터의 내용은 그 의미(semantics)가 본질적으로 다양하여 근사한 데이터를 편리하고 효과적으로 찾기가 어렵기 때문에 문제 해결이 쉽지 않다.

영상, 음향과 같은 멀티미디어 데이터의 내용은 대부분 구조를 갖추지 않은 복합 데이터이기 때문에 효과적인 검색 방법을 개발하는데 있어 데이터 내용을 직접 이

*중신회원

용한다는 것은 불가능하다. 대부분의 다른 시스템처럼 본 논문에서도 멀티미디어 데이터 내용의 묘사를 통해 탐색하는 방법을 따른다. 잘 알려진 키워드 방법은 부정확하고 사용자가 관심있는 데이터를 탐색하기가 어려우므로 적당치 않다. 따라서 본 논문에서는 더 좋은 대안으로 자연어 접근 방법을 제안한다.

자연어 접근 방법은 저장된 데이터의 내용을 묘사하기 위해 자연어 묘사(natural language description)를 결합(associating)하고 관련된 데이터의 검색을 위해 그 묘사를 이용한다. 즉 저장된 멀티미디어 데이터의 내용 묘사는 검색될 적당한(qualified) 멀티미디어 데이터를 설명하고 있는 사용자 질의 내의 묘사와 매치된다. 이 방법의 주된 문제점은 저장된 데이터의 내용 묘사가 사용자 질의에 포함된 묘사와 정확히 매치되지 않을 수도 있다는 것이다. 그 이유는 다른 사용자나 다른 시점의 한 사용자가 동의어를 사용하거나 개체의 내용을 일반화 내지 전문화 시킴으로써 같은 객체를 동일하게 설명하는 것이 어렵기 때문이다. 따라서 효과적인 검색과정을 위한 해결책은 정확한 매치가 불가능할 때마다 내용 묘사와 사용자가 준 묘사간에 부분적 매치나 근사 매치를 자동적으로 수행하는 것이다. 본 논문은 통합(integrating) 객체지향 및 자연어 인식 기법을 사용한 지능적(intelligent) 근사 매칭을 제안한다.

본 논문의 2절에는 관련 연구를, 3절에는 멀티미디어 데이터 검색 시스템의 근본적인 문제점을 다루고 있다. 4절에는 근사 매칭 알고리즘을 상세히 서술했고 5절에는 요약을 서술했다.

II. 관련 연구

지난 수 년동안, 학계나 산업계를 통해 멀티미디어 분야가 다양하게 연구되어 오고 있다. Toronto대학에서 개발한 MINOS 시스템[CHRIS86]은 텍스트, 영상, 음성뿐 아니라 속성(attribute)으로 구성된 고도의 구조화된 멀티미디어 객체를 다루고 있다. 고도의 브라우징(browsing)과 사용자 인터페이스는 동기화된(synchronized) 갱신뿐 아니라 스키마의 브라우징을 허용한다. MCC 데이터베이스 프로그램[WOEL86, 87] 또한 멀티미디어 응용에 필요한 데이터베이스 요구조건을 설립했다. 그들은 멀티미디어 데이터의 공유와 조작하기 위한 요구조건, 그리고 데이터 모델의 요구 조건을 확립했다. 하이퍼텍스트(hypertext) 또한 영상과 음향을 다루도록 확장되고 있다. 한 가지 주목할만한 성과는 Brown 대학에서 개발한 INTERMEDIA 시스템[YANK88]이다. [MASU

87]은 다른 연구들과 비교 분류하는 구조를 개발했다.

지능 정보검색(IR)은 인공지능과 정보검색 분야에서 중복하여 연구하고 있다. 몇년전 연구를 시작할 때 정보검색 연구자들은 일람표(tabular) 및 문자식(textual) 데이터에 직접 접근하는데 관심이 있었다. 그 당시 자연어 처리 연구자들은 그것이 지능적 정보검색에 기여할 수 있다는 견해를 가지고 있었다[SCHA81]. 특히 개념적 메모리를 조직하는 것에 대한 연구[DEJON79, KOLO80]는 지능적 사실 검색 시스템의 설계에 적당하다.

일찌기 정보검색 영역에서 AI기법에 대한 관심이 있었다[SPAR78, SMIT80]. IRUS 시스템[BATE83]에는 자연어를 통해 다른 종류의 데이터베이스를 처리하려는 현대적 시도가 잘 나타나 있다. RUBIC 시스템[TONG87]은 시스템의 인덱스에 텍스트내의 단어에 대한 위치 정보를 포함하고 있고 질의 처리 과정 동안에 위치 제어를 하도록 해주는 프로덕션 규칙을 기본으로 하는 시스템이다. BR 시스템[CROF87]은 검색 처리를 하는 모든 단계동안에 사용자를 보조해 주며, 스케줄러에 의해 관리되는 전문가 시스템들로 구성된다. 마지막으로, IOTA 시스템[CHIA87]은 키워드를 다양한 의미를 가진 명사 그룹으로 대체함으로써 정보검색 시스템의 성능을 질적으로 개선하고자 했다.

본 논문에서 제안한 방법은 위에 언급한 지능적 정보검색 시스템과는 약간 다르다. 이들 시스템의 대부분의 연구는 자연어 처리, 질의 처리 및 확장된 의미 모델을 기초로 하는 언역 기능에 주로 관심을 가지고 있다. 본 논문 또한 이런 특징을 갖고 있다. 그러나 시스템 개념과 사용자 개념간에 매칭기능에 있어서, 대부분의 시스템이 정확(exact) 매칭을 기본으로 하는데 비해 본 논문에서는 정확 매칭과 근사 매칭을 기본으로 하고 있다. 근사 매칭기능을 가진 다른 시스템의 경우에 사용된 매칭 함수는, 매칭 처리의 질을 향상시키기 위해 객체 지향 기법을 사용하고 있는 본 논문의 방법에 비해 초보적이고 피상적이다.

III. 멀티미디어 데이터 검색 시스템의 구조

이징은 멀티미디어 데이터베이스를 위한 지능적 검색 시스템의 여러 구성 요소로 이루어져 있다. 먼저, 멀티미디어 데이터의 지능적 검색과 관련된 여러 문제와 정의를 다루겠다.

3.1 정의와 배경

앞에서 언급한 바와 같이, 멀티미디어 데이터는 넓은

의미에서 알파벳, 숫자 데이터 뿐만 아니라 텍스트, 영상, 음성, 신호 등과 같은 형식화 되지 않은 데이터로 이루어진다. 멀티미디어 데이터베이스 관리 시스템(MDBMS)은 모든 멀티미디어 데이터를 관리하고, 질의어와 질의 처리를 제공할 뿐 아니라 동시성(concurrency), 일관성(consistency), 및 회복(recovery)을 다루는 기법을 제공해주는 시스템으로 정의된다.

데이터 모델과 구현 측면에서의 차이에도 불구하고, MDBMS에 관한 모든 연구는 추상 데이터 타입(abstract data type) 개념을 이용하여 멀티미디어 데이터를 조작하는 방향으로 이루어지고 있다. 이것은 일반적으로 적절한 접근 방법이다. 그러나, 어떤 연구도 멀티미디어 데이터의 검색과 내용 묘사의 문제점을 다루고 있지 않다.

멀티미디어 데이터를 다루는데 있어서 근본적인 어려움은 매우 다양한 의미를 가지고 있다는 것이다. 이런 어려움을 설명하기 위해 배의 영상을 보기로 하자. 그림이 주어졌을 때 그 그림속의 배가 어떤 종류인지를 어떻게 알 수 있는가. 다시 말하면, 배가 구축함, 순양함, 잠수함, 여객선 중 무엇인지 어떻게 알 것인가? 또 다른 예를 들면, 개와 고양이의 그림을 상상해 보자. 그들이 서로 쫓는지,늘고 있는지를 어떻게 알 것인가?

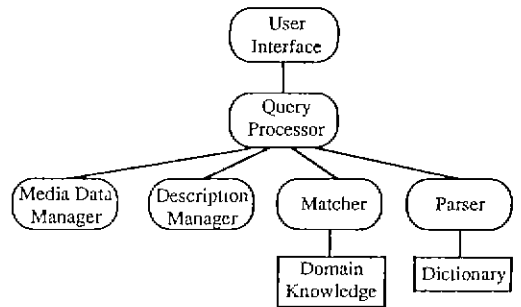
영상에 관한 질의에 대답하기 위해서는 인상을 살면서 겪게되는 풍부한 경험으로부터 좋은 답을 유도해야만 한다. 영상내의 다른 의미를 알기 위해서 영상의 내용을 분석하는 고도의 기술이 필요하게 된다. 그러나 오늘날의 기술은 이런 종류의 멀티미디어 질의에 답할 수 있는 기능을 가진 전문가 시스템이 기대할 만큼 발전되어 있지 않다. 그러나 그 대안으로 AI 및 정보검색 기술을 사용할 수 있다. 멀티미디어 데이터의 내용을 단어나 텍스트로 추상화할 수 있고 사용요구나 질의를 매치하기 위해 본래의 멀티미디어 데이터와 동등한 텍스트 묘사를 사용할 수 있다. 이것이 다른 응용을 위해, 멀티미디어 데이터를 다루는 MDBMS 설계에 사용되는 원리이다

멀티미디어 데이터는 크게 세 가지 부분으로 구성될 수 있다. 첫째, 가공되지 않은(raw) 데이터 부분은 본래의 멀티미디어 데이터를 주사(scanning) 또는 계수화(digitizing)함으로써 얻어진 영상, 음향, 신호 등의 비트열 표현이다. 둘째, 등록(registration)데이터 부분은 가공되지 않은(raw) 데이터에 관한 정보를 강화하고 중복을 없앤 것이다. 셋째, 묘사(description) 데이터 부분은 멀티미디어 데이터의 내용을 서술한다. 현재의 기술로는 묘사 데이터를 컴퓨터 만으로는 자동적으로 유도할 수 없으

므로 사용자가 자연어 형태로 멀티미디어 데이터에 대한 묘사 데이터를 부여한다.

3.2 구조

본 절에서는 멀티미디어 데이터 검색을 다루는 MDBMS의 여러 구성요소를 다룬다. 구성요소는 그림 1과 같다.



(그림 1) 멀티미디어 DBMS 구조

그림 1에서 나타난 바와 같이, 구성요소는 질의처리기(Query Processor), 묘사관리기(description manager), 사용자 인터페이스, 파서, 매치(matcher)로 이루어져 있다. 질의처리기는 사용자로부터 질의를 받아 다른 요소를 호출함으로써 질의를 처리한다. 멀티미디어 데이터에 관한 새로운 묘사가 들어오면, 질의처리기는 파서(parser)를 호출한다. 따라서 일차 프레디케트(predicate) 계산을 하기위해 사전(dictionary)을 사용한 후 질의 처리기에 다시 반환한다. 질의처리기는 멀티미디어 데이터와 그것의 묘사를 연결해주는 묘사관리기에게 프레디케트를 넘겨준다.

질의 처리기가 텍스트 묘사를 가진 질의를 받으면, 그와 동등한 프레디케트를 알기위해 파서를 호출한다 그후 프레디케트는 매치에게 넘어간다. 매치는 질의 프레디케트와 저장된 멀티미디어 데이터의 프레디케트를 비교함으로써 적절한(자격을 갖춘) 멀티미디어 데이터와 질의를 매치시킨다. 매치는 묘사관리기를 호출하고 도메인 지식을 사용하면서 이 작업을 한다. 또한 정확한 매치가 불가능하면 매치는 자동적으로 근사 매치와 연결한다. 질의가 끝났을 때, 질의 처리기는 적절한 멀티미디어 데이터로 연결을 한다.

3.3 멀티미디어 데이터에 대한 자연어 묘사

위에서 언급했듯이, 본 논문은 저장된 멀티미디어 데

이타의 내용에 대한 자연어 묘사와 사용자가 관심있는 멀티미디어 데이터의 질의묘사를 매칭함으로써 검색하는 방법을 제안하고 있다. 키워드는 불연속적이고 멀티미디어 데이터의 내용을 적절히 포착하기 위한 합성 연결 기법(complex linking mechanism)이 부족하기 때문에 키워드 탐색을 사용하지 않는다. 또한 키워드만으로 정확한 의미를 전달하는 것이 항상 가능하지 않을때도 있다.

오늘날의 AI 기술로는 자연어 처리가 제한없이 이루어지기가 어렵다. 그러나 멀티미디어는 특별한 응용의 경우에만 묘사 영역을 제한한다. 그러므로 자연어 묘사 대신, 본 논문에서는 멀티미디어 데이터를 묘사하기 위해 캡션(caption) 기능을 사용한다. 캡션은 자연어의 일부를 가지고 기록묘사(writing description)를 하는 것이 자연스럽지만 특별하고 틀에 맞추어진 방법이다. 캡션에 대한 상세한 내용과 제한점은 본 논문에서는 다루지 않으며 [HOLT90, ROWE91]에 나와있다.

자연어 묘사는 모든 사람에게 친근하다는 장점이 있다. 그러나 그것은 묘사인식의 문제를 해결하지 못하며, 특정 응용 분야에 초점을 맞춘 매칭처리가 필요하다. 이 문제를 처리하기 위해 사용자가 응용의 범위를 정의하는 사전(dictionary)을 제공하여, 어휘(vocabulary)와 의미 및 시스템의 지식(knowledge)을 제한한다. 따라서 텍스트 묘사를 프레디키트들로 변환한다. 자연어 묘사의 부정확성과 모호성은 프레디키트로 변환됨으로써 상당히 줄어든다. 이들 프레디키트는 성질(property)이나 관계(relationship)와 같은 멀티미디어 데이터를 가지고 실세계 개체에 대한 사실을 나타내고 있다. 대부분의 파서방법에서 처럼, 묘사의 형식적 표기로서 1차 프레디키트 계산(first order predicate calculus)을 선택 사용한다. 사전을 사용하여 묘사를 프레디키트로 변환하는 것이 파서의 일이다. 또한, 본 논문의 파서는 사용자 질의를 자동적으로 주어, 동사, 목적어로 분해하는 기법을 제공한다. 분해된 요소는 주어, 명사, 목적어 범주로 분석하는 도메인 의존지식(domain dependent knowledge)과 매치시키기 위해 사용된다. 파서와 프레디키트에 대한 상세한 내용은 본 논문의 범위 밖이며[LUM89, HOLT90, DULL90]에 나와있다. 자연어 묘사 및 파서를 이용하여 동등한 프레디키트들로 변환한 예가 다음과 같이 주어졌다.

묘사: "A car with red body"

프레디키트: car(x), component(x,y), body(y), color(y, red)

올바른 프레디키트 집합을 선택하는 것은, 전문가 시

스템에 대한 지식을 획득하는 것과 비교해 볼 때 매우 어려운 작업이다. 본 논문의 경우에는, 사전(dictionary)에 특정 단어와 연관된 분야에서 사용하는 모든 단어와, 단어를 묘사상에 나타낼 때 사용할 프레디키트들을 나열하였고, 이를 파서가 인식할 수 있다고 가정하는 것임으로도 충분하다. 그러므로, 묘사상에 사용될 수 있는 모든 프레디키트 집합이 사전에 정의되어 있어야 한다.

IV. 매칭(Matching)

이 장에서는 질의에 명시된 묘사와 내용을 자연어 묘사로 표현된 멀티미디어 데이터와 매치시키는 새로운 방법을 제안한다. 매칭 처리에서 중요한 것은 객체지향 클래스 계층 구조를 사용하여 도메인 지식(domain knowledge)을 표현하는 것이다. 먼저, 앞에서 설명을 회피한 매칭 알고리즘의 문제점을 살펴보고자 한다. 이 문제점들이 본 논문에서 소개할 근사 매칭에 대한 새로운 접근법을 제시한 동기가 되었다.

4.1 매칭에서의 문제점

현행 시스템[LUM89, HOLT90]에서는 파싱을 하면, 데이터베이스에 저장된 멀티미디어 데이터 인스턴스에 대한 프레디키트의 집합이 결과로 나온다. 질의는 자연어로 입력되어 파스(parse)된다. 질의 프레디키트의 인수는 변수가 될 수 있다. 질의 프레디키트와 연관된 멀티미디어 데이터의 내용 묘사(content description) 프레디키트가 있을 경우에는 그 멀티미디어 데이터가 질의의 결과로 선택된다. 사용자 질의에 표현된 묘사와 저장된 멀티미디어 데이터의 내용 묘사는 정확하지 일치하지 않을 수 있다. 때때로 도메인에 적합한 몇몇 규칙을 두어서, 동일한 사물이 서로 다른 프레디키트로 표현되는 상황을 지정한다. 이 매칭 기술은 같은 의미를 가지나 다르게 표현되는 자연어 구문을 파악할 수는 있지만, 프레디키트 간의 의미론적 관계는 파악할 수 없다.

다음과 같은 영상 데이터가 있다고 가정하자; "A car with red body". 이 예에서 생성되는 프레디키트는 "car(x), component(x,y), body(y), color(y,red)"이다. 인수에 대해 "a red car"라는 질의를 고려해 보자. 이 질의는 "car(x), color(x,red)"로 해석된다. 이때 시스템은 car의 body color가 car의 color와 동일하다는 것을 파악하지 못하므로 매치를 시키지 못한다. 이 문제를 해결하기 위해 프레디키트간의 의미론적 관계를 표현하기 위한 규칙을 제시하였다. 위의 예는 다음과 같은 규칙이 적

용될 수 있다.

IF(car(X), component(X,Y), body(Y), color(Y,Z)) then color(X,Z)

위의 규칙을 사용하면 color(X,red)가 위의 예로부터 추론되고 질의와 묘사간의 매칭이 일어난다. 아직 해결되지 않은 중요한 문제는 프레디카트의 어느 리터럴(literal)을 매치시킬 것인가와 어느 정도 일반화(generalize)하여 매치를 시킬 것인가이다. 이것은 앞서도 언급했던 근사 매칭에도 해당되는 것이다. 이에 대한 해답은 도메인-종속(domain-dependent) 지식을 사용하는 것이다. 저장된 멀티미디어 데이터의 내용 묘사에 대해 사용자 질의가 정확하게 매치되기를 원할 때는 현행 매칭기술 [LUM89, HOLT90]이 정확하다. 그러나 이 방법은 사용자 질의에 대해 정확한 매치가 일어나지 않을 때, 사용자는 아무런 응답을 받지 못한다는 문제가 생긴다. 이런 경우 지능 검색 시스템(intelligent retrieval system)은 근사 매치를 시도한다. 이 시스템은 멀티미디어 데이터의 내용묘사가 사용자 질의에 주어진 묘사를 일반화시킨 것을 어느 정도 만족하면 그것을 질의 결과로 취급한다. 여기서의 목적은 효과적으로 근사 매치를 수행하는 것이다. 앞서도 언급했듯이 본 논문에서 제안한 근사 매치 알고리즘은 이러한 목적을 위해 도메인-종속 지식을 사용한다.

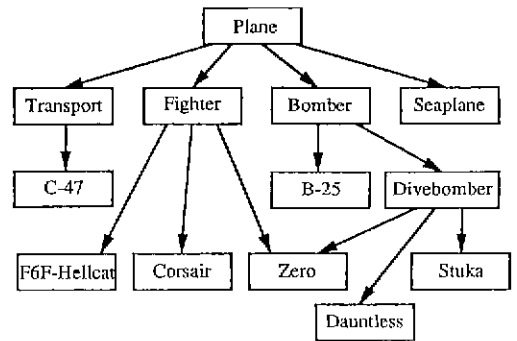
4.2 도메인-종속 지식

앞에서 이미 멀티미디어가 응용되는 분야에서 멀티미디어 데이터의 묘사 범위가 한정된다는 것을 언급하면서, 멀티미디어 데이터를 묘사하는데 캡션(caption)을 사용하는 것이 바람직하다는 것을 보여 주었다. 이것은 캡션의 도메인이 각 멀티미디어 응용에만 제한된다는 것을 의미한다. 도메인-종속지식은 캡션의 도메인에서 중요한 개념이다. 본 논문에서는 도메인-종속 지식에 명사와 동사 개념만을 포함시킨다.

도메인-종속 지식을 표현하기 위해 객체지향 데이터 모델[BANE87, KIM89, ZDON90]을 사용했다. 객체지향 모델은 고도로 구조화되어 있고, 복잡객체를 지원하며 모든 mini-world 객체를 나타낼 수 있다. 이 데이터 모델은 CAD/DAM, VLSI, 사무 자동화, 소프트웨어 공학 및 인공지능 분야 등에서 광범위하게 사용되어 왔다. 도메인-종속 지식을 표현하기 위해 객체지향 데이터 모델을 사용해야만 하는 당위성은 다음과 같다.

먼저, 이 데이터 모델은 캡션의 내용을 개념적으로

일반화시킬 수 있는 일반화(generalization)와 특수화(specialization) 추상화를 지원한다. 두번째로, 연구가 [WOEL87, HOLT90]들이 이미 멀티미디어 데이터베이스 응용분야에서 객체 지향 모델을 사용하는 것이 적당하고 앞으로도 전망이 있다는 것을 확인했다. 본 논문에서는 제 2차 세계 대전 중 태평양양전에 참가한 미국 군대를 도메인으로 제안했다. 그러나 이 근사 매칭 알고리즘은 다른 멀티미디어 응용에도 모두 적용될 수 있다.



(그림 2) "Plane"의 일반화 계층 구조

그림 2는 도메인에서 명사개념인 "Plane"의 일반화 계층 구조를 보여준 예로 세계 2차 대전중 태평양양전에 참가한 비행기에 대한 도메인-종속 지식이다. 논문은 객체, 클래스, 클래스 계층 구조 또는 속(lattice)에 의한 계승(inheritance) 메소드(method)와 같은 객체지향 개념에 익숙하다는 가정하에 쓰여졌다. 그림 2는 한 클래스로부터 그 서브클래스로 화살표 방향이 그려져 있다. 그림 2에서 Plane 클래스는 Transport, Fighter, Bomber 및 Seaplane 클래스로 특수화되어 있다. Transport 클래스는 C-47 클래스로 특수화되고 Fighter 클래스는 F6 F-Hellcat, Corsair 및 Zero 클래스로 특수화된다 또한, Bomber 클래스는 B-25와 Zero, Dauntless 및 Stuka 클래스로 특수화되는 Divebomber로 특수화된다. "Plane"의 일반화 계층은 Zero 클래스기 Fighter 클래스 및 Divebomber리는 두 개의 슈퍼클래스를 가지므로 클래스 속(lattice)이다. 또한, 슈퍼클래스의 특성은 슈퍼클래스/서브클래스 계층구조를 따라 그들의 모든 서브클래스로 계승된다. 그러나 역은 성립되지 않는다.

그림 2는 도메인에서 명사(즉 Plane)에 해당되는 도메인-종속 지식의 한 예이다. 필요에 따라, 도메인에 모든 명사 및 동사 개념에 대한 도메인-종속 지식을 가질 수 있다. 명사 및 동사 개념 중 몇 개는 동일한 클래스 또는

일반화 계층에 속할 수 있다. 그러므로 일반화 계층 구조를 명사나 동사 개념 각각에 대해 생성할 필요는 없다.

4.3 근사 매칭 알고리즘

이 절에서는 근사 매칭 알고리즘에 대해 설명한다. 명확성을 위해 예를 들면서 근사 매칭 알고리즘을 설명하고자 한다. 특별히 언급하지 않는 한 그림 2에 주어진 일반화 계층 구조 예를 가지고 설명하는 것이다. 우선 하고자 하는 일이 어떤 것인지 알아본다. 먼저, 멀티미디어 데이터베이스에 저장된 비행기의 영상이 있고 그 영상은 "Transport plane"으로 묘사되어 있다고 가정한다. 이제 사용자가 "C-47"인 모든 비행기를 요청하는 질의를 했다고 가정하자. 정확한 매칭은 이루어지지 않더라도, 도메인-종속 지식에 의해 모든 C-47은 Transport plane이므로, Transport plane과 연관된 모든 영상을 검색해야 한다. 사용자가 모든 Fighter plane을 요청할 때 단순히 Transport plane과 연관된 모든 영상 데이터를 검색할 수는 없는데, 이는 사용자가 원하는 것이 아니기 때문이다.

여기서 제시하는 알고리즘은 설계하는 사람에 따라 일반화 계층 구조의 정의가 달라지는 정도를 최소화하기 위한 것이다. 계층 구조 설계자는 다른 사람들과 도메인-종속 지식을 보는 관점이 다를 수 있다. 이러한 현상으로 인해 근사 매칭을 하는 동안 일반화 계층 구조의 어떤 가지(branch)가 다른 특징 가지에 비해 길게 지우치게 될 수 있다. 효과적인 근사 매칭 알고리즘이라면, 위에서 언급한 모든 문제를 해결하고 일반적인 해결책을 제시해야 한다. 본 논문에서는 이러한 문제를 일반화 계층 구조의 각 노드(클래스)에 가중치를 지정하는 휴리스틱(heuristic)을 사용하여 해결했다. 이제, 주어진 사용자 묘사에 대해서 저장된 멀티미디어 데이터를 검색하기 위한 공경하고도 정확한 가중치 부여 스킴을 세우고자 한다.

4.3.1 일반화 계층 구조상에서의 가중치 부여 스킴

이 절에서는 근사 매칭 알고리즘에 사용될 일반화 계층의 가중치 부여 전략을 설명한다. 일반화 계층상에서의 가중치 부여 전략은 객체 지향 데이터 모델에서 지원되는 클래스 계층(속)의 의미론이나 IS-A 계층 개념을 사용한다. 가중치 0을 가지고 있는 클래스 C가 있다고 할 때, 명사나 동사 개념에 대한 일반화(클래스) 계층 구조에 다음과 같은 두 개의 일반화 휴리스틱을 적용할 수 있다.

휴리스틱 1 : C의 모든 직접(간접) 서브클래스는 양수 가중치를 가진다.

휴리스틱 2 : C의 모든 직접(간접) 슈퍼클래스는 음수 가중치를 가진다.

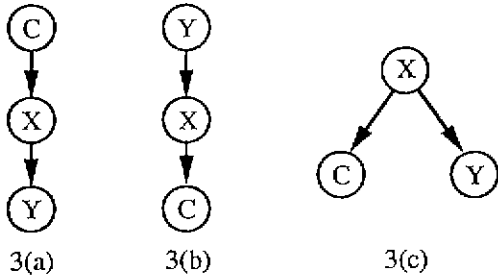
여기서 C는 질의 묘사에서 지정한 클래스이다. 음수 (-) 가중치를 일반화 노드에 지정하는 것은 자명한 것이다. 특수화 노드에 양수(+)를 지정하는 것은 특수화가 자신의 부가적인 정보뿐 아니라 모든 부모(parent)노드의 특성을 계승받는다는 사실에 기초한 것이다. 그러므로 특수화 계층 구조를 따라 연결된 경로상에 있는 노드에 양수 또는 그 이상의 가중치가 지정되어야 한다는 것을 알 수 있다.

휴리스틱이 주어지면, 명사 또는 동사 개념으로 사용자 질의 묘사에 지정된 클래스 C에 대한 상대적인 양수 가중치를 가지는 클래스 계층에 있는 모든 클래스가 근사 매칭시에 선택된다. 이것은 양수 가중치를 가지는 모든 클래스가 클래스 C의 서브클래스(특수화된 클래스)이기 때문이다. 각 클래스는 C클래스의 특수화된 버전이므로 C클래스의 특성을 포함하며, 따라서 클래스 C라고 할 수 있다.

반면, 클래스에 상대적인 음수 가중치를 가지는 클래스 계층에 있는 모든 클래스는 가중치에 따라 신중하게 선택되어야 한다. 이것은 음수 가중치를 갖는 모든 클래스는 클래스 계층상에 지정된 클래스 C의 슈퍼클래스이기 때문이다. 각 클래스는 클래스 C의 일반화된 버전이므로 클래스 C의 모든 특성을 포함하지는 않으며, 따라서 클래스 C가 아니라고 할 수 있다. 그중 어느 클래스를 선택할 것인가는 사용자로부터 어느 정도까지 일반화시킬 것인가에 대해 얻은 정보에 달려있다.

지금까지 제시한 가중치 부여 스킴은 모호하고 명확히 정의된 것이 아니다. 다만 명확한 것은, C를 포함하면서 양수 가중치가 지정된 동일한 클래스 계층 구조상의 모든 클래스는 항상 선택된다는 것이다. 반면, C를 포함하면서 음수 가중치가 지정된 동일한 클래스 계층 구조에 있는 클래스는 사용자에게 의해 주어진 어떤 임계값(threshold value)을 초과해야만 선택된다는 것이다. 이제 클래스 계층 구조에서 서로 다른 클래스들의 가중치 지정에 대해 설명하고자 한다.

클래스 계층 구조에서 클래스에 가중치를 저장할 때 세 가지 상황이 발생한다. 이러한 상황이 그림 3에 나와 있다. 사용자의 질의에 명시된 클래스를 C라고 가정한다. 화살표 방향이 한 클래스에서 그 서브클래스로 표시되어 있다고 앞에서 이미 가정했다. 예를 들어, 그림 3(a)에서는 클래스 C가 클래스 X의 슈퍼클래스 X의 슈퍼클



(그림 3)

래스이며 클래스 X는 클래스 C의 서브클래스이다. 그림 3(a)에 나와 있듯이 첫번째 상황은 클래스 C의 서브클래스인 클래스(X 또는 Y)에 가중치를 지정하는 것이다. 두번째 상황은 그림 3(b)에 나타나 있는 클래스 C의 슈퍼클래스(X 또는 Y)에 가중치를 지정하는 것이다. 세번째는 그림 3(c)에 나와 있듯이 클래스 C의 슈퍼클래스(X)의 서브클래스(Y)에 가중치를 지정하는 것이다.

가중치 부여 스킴 원리는 아주 간단하다. 양수 가중치를 갖는 모든 클래스와 음수 가중치 값을 가지지만 임계값을 초과하는 몇몇 클래스를 근사 매칭시에 선택한다. 먼저, 사용자 질의에 명시된 클래스 C의 가중치를 0으로 지정한다. 앞에서 언급했듯이 클래스 C는 자연어 질의 묘사에 표현된 명사 또는 동사 개념이 될 수 있다. 클래스 C는 C를 포함하는 클래스 계층구조상의 모든 다른 클래스에 대한 참조 지점이 된다. 클래스 C의 서브클래스에 대해 휴리스틱 1을 적용하여, 양수 가중치를 지정할 수 있는데, 이것은 이 클래스들이 클래스 C의 특수화된 버전이기 때문이다. 클래스 C의 특수화된 버전은 C 자체보다 더 자세하고 정확한 정보를 가지므로 0대신 양수 가중치를 가진다. 필요에 따라, C의 모든 서브클래스에 동일한 양수 가중치를 지정한다.

C의 슈퍼클래스 또는 C의 슈퍼클래스의 서브클래스인 모든 클래스에 대해서는 휴리스틱 2에 따라 음수 가중치를 지정하는데 이것은 이 클래스들이 클래스 C의 일반화된 버전이기 때문이다. 클래스 C의 일반화된 버전은 C 자체보다 더 적은 정보나 일반화된 정보를 가지므로 음수 가중치를 가지는 것이다. 서로 다른 일반화된 버전은 서로 다른 음수 가중치를 가진다. 그러나 음수 가중치를 지정할 때 모델의 정의에 따른 영향을 최소화해야 한다. 클래스 계층 구조에서 클래스 C와 밀리 떨어진 클래스일수록 그 클래스에는 더 큰 음수 값이 지정된다.

대부분의 시스템에서는, 클래스 계층 구조에서 한 클래스(또는 노드)의 가중치가 사용자 질의에 명시된 C클래스 레벨에 대한 상대적인 클래스의 레벨 깊이에 선

형적으로 반비례하여 지정된다. 이것은 정확한 접근 방법이 아니라고 생각되는데, 그 이유는 동일한 클래스 계층 구조에서 특정 클래스의 참조 클래스(이 예에서는 클래스 C)에 대한 상대 거리는 절대적인 것이 아니라 도메인 지식에 대한 특정 설계자의 관점에 따라 달라 지므로 어느 정도 인위적이라 할 수 있기 때문이다. 이 접근 방법의 주된 문제점은 어떤 긴 가치에 속하는 클래스는 참조 클래스로부터의 상대적인 거리로 인해 더 큰 음수 값을 지정받을 수 있으므로 불공평하게 자격을 부여받지 못하게 된다는 것이다. 본 논문에서의 가중치 부여 스킴은 다른 짧은 가치에 비해 클래스 계층 구조 상의 긴 가치에 있는 노드에 생기는 치우침을 최소화하는 것이 목적이다. 클래스 C가 그림 3에서처럼 사용자 질의에 명시된 클래스라고 하면, 위의 세 가지 상황에서, 클래스들에 대한 가중치 지정 공식은 다음과 같이 지정할 수 있다.

- (1) 사용자 질의에 명시된 클래스(그림 3(a)의 클래스 C)
가중치 = 0
- (2) 상황 1에서 클래스 C의 서브클래스(그림 3(a)의 클래스 X 또는 Y)
가중치 = α , α 는 정수 상수
- (3) 상황 2에서 C의 슈퍼클래스(그림 3(b)의 클래스 X 또는 Y)

$$\text{가중치} = -\left[\alpha \times \sum_{i=1}^n \left(\frac{1}{\beta}\right)^i \right]$$

α, β 는 정수로된 상수이며 n은 클래스 C의 상대적인 슈퍼클래스의 레벨 수

- (4) 상황 3에서 클래스 C의 슈퍼클래스의 서브클래스 (그림 3(c)에서 클래스 Y)

$$\text{가중치} = -\left[\alpha \times \sum_{i=1}^h \left(\frac{1}{\beta}\right)^i \right] - \left[\gamma \times \sum_{j=1}^l \left(\frac{1}{\nu}\right)^{(i+1-j)} \right]$$

$\alpha, \beta, \nu, \gamma$ 는 정수로된 상수, h는 클래스 C에 상대적인 슈퍼클래스 X의 레벨 수이며 l은 슈퍼클래스 X에 상대적인 서브클래스의 레벨 수

이 스킴에서는 양수 가중치가 지정된 클래스가 항상 선택된다. 음수 가중치를 가지는 클래스는 사용자가 설정한 임계값을 초과하지만 않으면 선택될 수 있다. 서로 다른 클래스에 대한 가중치 지정을 이해하기 위해 그림 2의 클래스 계층 구조를 사용하여 예를 들어 보겠다. 사용자 질의가 주어졌는데 사용자 묘사에 매치하는 영상 데이터의 내용 묘사가 데이터베이스에 없으면, 매치는

자동적으로 군사 매칭을 수행한다. 가중치 지정 공식과 주어진 사용자 질의 묘사를 사용하여 클래스 계층 구조에 있는 클래스에 대한 가중치를 다음과 같이 나타낼 수 있다. 인수인 α, β, ν 및 γ 에는 40, 2, 48 및 2를 각각 주었다.

- (1) 주어진 질의 묘사: "A transport plane sank in the Pacific"
가중치: transport=0, C-47=40, Fighter=-44, Corsair=-56
- (2) 주어진 질의 묘사: "A F6F-Hellcat sank in the Pacific"
가중치: F6F-Hellcat, Plane=-30, Seaplane=-54, Stuka=-72, C-47=-56
- (3) 주어진 질의 묘사: "A bomber sank in the Pacific"
가중치: Bomber=0, Stuka=40, Plane=-20, Seaplane=-44, C-47=-56

위의 예에서는, 양수 가중치가 지정된 모든 클래스가 부분 매칭동안에 선택된다. (1)의 예에서 클래스 C-47은 양수 가중치를 갖는다. 이것은 "A C-47 sank in the Pacific"인 내용 묘사를 가지고 있는 저장 영상이 40이라는 가중치를 가지며 군사 매칭동안 질의 결과로 선택됨을 의미한다.

이 예에서 보여주듯이 사용자 질의 묘사에 지정된 클래스의 서브클래스인 클래스 계층 구조상의 모든 클래스에는 양수 가중치가 지정됨을 알 수 있다. 다른 모든 클래스는 음수 가중치가 지정된다. 이러한 클래스의 가중치는 사용자 질의에 명시된 참조 클래스의 레벨에 대한 상대적인 클래스 레벨 깊이에 반비례하나, 엄격히 말해서 선형적으로 반비례하지는 않는다.

일반적인 선형 방법에서는 클래스의 가중치가 사용자 질의에 명시된 클래스 C의 레벨에 대한 상대적인 클래스의 레벨 깊이에 반비례한다. 만약 -10과 같은 음수 상수 가중치를 클래스 C로부터 떨어진 각 레벨에 지정하면 클래스 C로부터 5 레벨 떨어진 클래스는 -50이라는 가중치를 가지며, 클래스 C에서 2 레벨 떨어진 클래스는 -20을 가진다. 예를 들어 사용자 묘사가 "A transport sank in the Pacific"이라면 클래스 Transport의 가중치는 0이며, 클래스 Seaplane의 가중치는 -20이고, 클래스 Stuka의 가중치는 -40이다. 우리가 제시한 공식을 사용하면 같은 사용자 질의에 대해서 클래스 Transport, Seaplane 및 Stuka에 각각 0, -44 및 -62가 지정된다. 클래스 Stuka의 가중치에 본 논문에서 제시한 방법 대신 선형 방법을 사용하면 클래스 Seaplane의

가중치에 비해 상대적으로 더 치우치게 된다.

Stuka가 Transport에 가까운 것에 비해 클래스 Seaplane이 어느 정도 의미론적으로 Transport에 가까운지 결정하기가 어려운데 이것은 Seaplane과 Stuka가 둘다 Plane 유형이기 때문이다. 본 논문에서 제안한 방법 대신 선형 방법을 사용할 경우, 클래스 Stuka가 클래스 Seaplane에 비해서 군사 매칭시에 선택될 기회가 적어지는 방향으로 임계값이 선택된다. 또 다른 어려운 일은 임계값뿐 아니라 지정 공식에 적용되는 상수값을 설정하는 것이다. 사용자는 군사 매칭시에 사용할 멀티미디어 데이터의 갯수에 따라 상수 및 임계값을 정확히 선택해야 한다. 그러므로 시스템이 매칭 프로세스를 수행하는 동안에는 사용자 인터페이스를 통해 사용자와 상호 대화를 해야 한다.

4.3.2 가중치 알고리즘의 응용

가중치 알고리즘을 응용할 때는 사용자 질의에 주어진 묘사뿐만 아니라 저장된 각 멀티미디어 데이터와 함께 저장된 자연어 묘사를 이해하기 위해 파서가 필요하다. 이미 언급했듯이 묘사가 파스되어 시스템에 1차 프레디키트로 저장된다. 그런 다음 질의는 다음과 같이 치러진다. 질의를 사용자로부터 받으면, 파서는 자연어 묘사를 더 작은 구성요소 그룹인 명사, 동사 및 목적어 명사구로 분할한다. 이러한 프레디키트들이 저장된 멀티미디어 데이터의 묘사에 해당되는 프레디키트와 정확히 매치하면, 선택된 멀티미디어 데이터가 검색된다. 그렇지 않으며 멀티미디어 데이터를 찾기 위해 matcher가 객체 클래스의 명사 및 동사 일반화 계층 구조를 찾아서 제안된 가중치 지정 스킴에 따라 사용자로부터 받은 해당 가중치 인자를 지정하면서 클래스에 가중치를 지정하도록 했다. 그런 다음 사용자가 설정한 임계값을 초과하는 가중치와 관련된 멀티미디어 데이터를 검색한다. 자연어 질의 묘사의 분할을 위에서 언급한 체 가지 그룹보다 더 작은 구성요소로 나눌 수 있다. 예를 들어 복합 명사구는 몇 개의 더 작은 명사 그룹으로 분리되고 가중치 알고리즘이 결합 가중치를 얻기 위해 이 그룹에 적용될 수 있다. 예를 들어, "The man with a mustache"는 두 개의 클래스, "man"과 "mustache"로 나누어진다. 분할을 더 세분화할수록 더 크고 복잡한 처리가 요구된다.

V. 요 약

데이터베이스에 저장된 음향 또는 영상과 같은 멀티

미디어 데이터를 검색할 때의 주된 문제점은, 그 데이터들이 본질적으로 풍부한 의미를 가지며, 데이터베이스에 사용되는 기존 탐색 방법이나 정보 검색 시스템은 작동을 하지 못하거나 거의 소용이 없다는 것이다. 지능 정보 검색 시스템에 대한 대부분의 연구는, 자연어 처리와 확장 의미론 모델에 기초한 문서 내용 또는 사용자 질의의 연역 추론에 관한 것이다. 그러나, 그들 중 대부분은 정확한 매칭 또는 단순 선형 방법을 사용한 원시 근사 매칭을 다루고 있다.

본 논문에서는 MDBMS가 데이터를 검색할 때 발생하는 근본적인 문제점들을 살펴보고, 멀티미디어 데이터베이스에 대한 검색 시스템의 구조를 대략 살펴 보았다. 본 논문은 객체 지향 데이터 모델을 이용하여 표현된 도메인 지식을 사용한 근사 매칭 알고리즘과, 데이터베이스에 저장된 서로 다른 멀티미디어 데이터에 동적으로 가중치를 부여하는 가중치 부여 스킴을 제시 하여서, 주어진 사용자 질의 묘사에 부분적으로 매치되는 멀티미디어 데이터를 선택할 수 있도록 하였는데 주된 의미가 있다. 본 논문에서 제시한 매칭 스킴은 또한, 근사 매치가 올바르고 정확하게 사용자에게 정보를 제공하도록 사용자와 대화를 하도록 되어 있다.

파서는 사용자 질의를 주어 명사, 동사 및 목적어 명사 구성요소로 자동적으로 분할하는 메카니즘을 제공한다. 데이터를 검색하는 동안에, 명사 및 동사 범주를 다루는 도메인-중속 지식의 일반화 계층 구조에 매치시키기 위해 분할된 구성요소를 사용하므로 이것은 필수적인 것이라고 할 수 있다. 근사 매칭을 할 때 캡션을 완전히 이해하고 처리하기 위해서는 자동적으로 형용사와 다른 캡션 구성요소를 구별해 낼 수 있도록 파서를 개선시키는 연구가 필요하다.

본 논문의 접근 방법은 간단하다. 이러한 단순성은 객체지향 데이터 모델의 일반화 및 특수화 추상의 의미론에 기인한 것이다. 또한, 본 논문의 접근 방법은 IR 또는 AI분야의 응용에 실질적으로 적용될 수 있는 일반적인 것이라고 생각한다. 본 논문은 저자의 관심 연구 분야의 한 영역일 뿐이며 현재 다른 멀티미디어 데이터 탐색 방법을 연구중이다.

참 고 문 헌

[BANE87] Banerjee, J. *et al.*, "Data Model Issues for Object-Oriented Application," ACM TOOLS, January 1987.
 [BATE83] Bates, M. and B. Bobrow, "Information retrieval using a transportable natural language inter-

face," Proc. of the 6th. ACM SIGIR conf. on R & D in information Retrieval, Bethesda, MD. 1983.
 [CHIA87] Chiaramella, Y. and B. Defude, "A prototype of an intelligent system for information retrieval: IOTA," Information Processing and management, Vol. 23, No 4, pp. 285~303, 1987.
 [CHRIS86] Christodoulakis, S. *et.*, "Multimedia Document Presentation, Information Extraction, and Document formation in MINOS: A MODEL and a System," ACM TOOLS, Vol. 4, Oct, pp. 345~383.
 [CROFT87] Croft, W. and R. Thompson, "I3R: A New Approach to the Design of Document Retrieval Systems," Journal of the American Society for Information Science, 38, pp. 389~404. 1987.
 [DEJON79] Dejon, G. F., "Prediction and substantiation: A new approach to natural language processing," Cognitive Science 3,3 (July 1979), pp. 251~273.
 [DULL90] Dulle, J. "The Scope of Descriptive Captions for Use in a Multimedia Database System," M. S. Thesis, Computer Science Department, Naval Postgraduate School, Monterey, CA, June 1990.
 [HOLT90] Holtkamp, B. *et. al.*, "Demon- A media object model incorporating natural language descriptions for retrieval support," Tech. Report NPS52-90-019, Computer Science Department, Naval Postgraduate School, Feb. 1990.
 [KIM89] Kim, Kyung-Chang *et al.*, "Cyclic Query Processing in Object-Oriented Databases." Proc. 7th. Int'l Conf. on Data Engineering, L. A., CA, Feb. 1989
 [LUM89] Lum, V. and K Meyer-Wegener, "A Multimedia Database Management System Supporting Content Search in Media Data," Tech. Report NPS 52-89-020, Computer Science Department, Naval Postgraduate School, Monterey, CA, March 1989.
 [MASU87] Masunaga, Y. "Multimedia Databases: A Formal Framework." Proc IEEE CS Office Automation Symp. IEEE CS Press. order no 770, Washington 1987. pp. 36~45.
 [ROWE91] Rowe, N. and E. Guglielmo, "Exploiting Captions for Access to Multimedia Databases," To appear in IEEE Computer, 1991.
 [SCHA81] Schank, R. *et. al.* "Conceptual information retrieval," Information Retrieval Research. Oddy, Robertson, van Rijsbergen, Williams, Eds., London, 1981.
 [SMIT80] Smith, L. C. "Artificial intelligence: what can it offer to information retrieval," Annual Re-

vue on Information Science Technology, 15 : 67~115, 1980.

[SPAR78] Spark, K. Jones, "Artificial intelligence: what can it offer to information retrieval," Proc. of the Informatics 3. Aslib, ed., London, 1978.

[TONG87] Tong, R. *et. al.*, "Conceptual Information Retrieval using RUBRIC," Proc. ACM SIGIR Conference. pp. 247~253, 1987.

[WOEL87] Woelk, D. and W. Kim, "Multimedia Information Management in an Object-Oriented Databases System," Proc. of VLDB, Brighton, England, Sept. 1987.

[YANK88] Yankelovich, N. *et. al.*, "Intermedia: The Concept and the Construction of a Seamless Information Environment," IEEE Computer, Vol. 21. No. 1, pp. 81~96, Jan. 1988.

[ZDON90] Zdonik, S. and D. Maier, "Readings in Object-Oriented Databases Systems," Morgan Kaufmann Publishers, Inc. 1990.

김 경 창



- 1978 홍익대학교 전자계산학과 학사
- 1980 한국 과학기술원 전신과 석사
- 1990 미국 텍사스 대학(오스틴) 건설학 박사
- 1980 ~ 1983 경제기획원 조사통계국 사무원
- 1990 ~ 1991 미국 해군 대학원 객원 교수
- 1991 ~ 현재 홍익대학교 전자계산학과 조교수

관심 분야: 객체지향 데이터베이스 및 프로그래밍 언어, 다중매체 시스템, 병렬처리 시스템, 소프트웨어 공학
