

# 대규모 이미지 데이터베이스에서 고차원 색인 구조를 이용한 효율적인 내용 기반 검색 시스템

## (An Efficient Content-based Retrieval System using High-Dimensional Index Structure in Large Image Databases)

이 등 호 \* 박 주 홍 \* 정 진 완 \*\* 김 형 주 \*\*\*  
 (Dong-Ho Lee) (Ju-Hong Park) (Chin-Wan Chung) (Hyoung-Joo Kim)

**요약** 이미지나 비디오, 오디오와 같은 멀티미디어 데이터들은 기존의 단순한 텍스트 기반의 데이터에 비하여 대용량적인 특성과 비정형적인 특성을 가지고 있어서 검색시 많은 어려움이 따른다. 본 논문에서는 대규모의 이미지 데이터베이스에서 효율적이고 신속하게 사용자가 원하는 이미지를 검색할 수 있는 내용 기반 검색 시스템을 제시한다. 이를 위해서 본 논문에서는 최근 여러 장점으로 인하여 신호 분석이나 이미지 압축 분야에 많이 사용되는 웨이브릿 변환을 이용하여 이미지 데이터로부터 내용 기반 검색에 사용되는 특징 벡터를 효율적으로 추출하는 기법과 유사성 측정 방법을 제안한다. 그리고, 이러한 특징 추출 방법과 유사성 측정 방법을 이용하여 내용 기반 질의 및 검색을 수행할 경우, 검색 조건을 만족하는 객체인데 실수로 검색해내지 못하는 경우인 false dismissals이 발생하지 않음을 보인다. 또한, 대규모 이미지 데이터베이스에서 신속한 내용 기반 검색을 지원하기 위하여 고차원 데이터에 대한 효율적인 색인을 제공하는 X-tree를 이용한 이미지 색인 방법을 보이며, 이것이 기존의 순차 검색이나 R\*-tree를 이용한 색인 방법보다 신속하게 이미지 데이터들을 검색할 수 있다는 것을 다양한 실험을 통해 보인다. 마지막으로, QBIC에서 제안한 검색 적합성 측정 방법을 이용하여 본 논문에서 제안하는 내용 기반 이미지 검색 시스템의 검색 적합성을 보인다.

**Abstract** There are many technically difficult issues in searching multimedia data such as image, video, and audio because they are more massive and more complex than simple text-based data. In this paper, we propose a content-based image retrieval method which can search large image databases efficiently by image content. For efficient feature extraction, we present a new method of extracting feature vector from image data using wavelet transform which is widely used in digital signal analysis and image compression. Then, we prove that there are no false dismissals in the case of searching image database using our proposed feature extraction and similarity measurement. We also show an image indexing method using X-tree that provides the efficient indexing for high-dimensional data to support fast retrieval in large image databases, and show that it supports faster retrieval than other method such as sequential scan or R\*-tree through various experiments. Finally, we show the retrieval effectiveness of our content-based image retrieval system using the measurement proposed by QBIC system.

### 1. 서론

#### 1.1 연구의 배경 및 목적

이미지 데이터는 기존의 텍스트 기반의 데이터에 비하여 대용량적인 특성과 비정형적인 특성을 가지고 있어서 신속하고 효율적인 검색에 많은 어려움이 따른다. 최근, 이러한 이미지 데이터를 신속하고 효율적으로 검색하기 위하여 다양한 내용 기반 이미지 검색 기법(content-based image retrieval)들이 연구되고 있는데, 이는 이미지 데이터로부터 내용으로 표현되는 특징 데이터를 추출하여 이것을 기반으로 원하는 이미지를 검색하는 방법이다.

\* 본 논문은 한국과학재단의 특정 연구 과제 No 95-0100-23-04-3, "다중매체 요구형 시스템 기술에 관한 연구"와 과학기술처의 국가 지리 정보 시스템 기술 개발 사업(1998)에서 지원 받았음

\* 비회원 서울대학교 컴퓨터공학과  
 dhlee@oopsla.snu.ac.kr  
 jhpark@oopsla.snu.ac.kr

\*\* 종신회원 : 한국과학기술원 전산학과 교수  
 chungcw@cs.kaist.ac.kr

\*\*\* 종신회원 : 서울대학교 컴퓨터공학과 교수  
 hjk@oopsla.snu.ac.kr

논문접수 : 1998년 6월 1일  
 심사완료 : 1998년 11월 17일

이미지 데이터로부터 추출할 수 있는 특징 데이터들은 이미지의 모양(shape), 색상(color), 질감(texture) 등과 같은 기본적인 특징(primitive feature)과 이미지의 다양한 의미 정보(semantic information)를 텍스트 형태로 표현하는 논리적인 특징(logical feature)으로 나눌 수 있다[1].

기존의 이미지 검색 시스템들은 대부분 논리적인 특징에 기반한 검색 방법만을 제공하였는데, 이는 논리적인 특징이 텍스트 형태로 표현되기 때문에, 기존의 상업용 데이터베이스 시스템의 질의 언어(SQL)와 쉽게 연동하여 사용할 수 있으며, 사용자에게 익숙한 의미 정보에 의한 검색이 가능하기 때문이다. 하지만, 논리적인 특징에 기반한 검색 방법은 사람이 일일이 의미 정보를 기술해야 된다는 단점과 의미 정보를 작성하는 사람과 실제 검색하는 사람 사이의 관점의 불일치로 인하여 검색이 비효율적일 수 있다는 단점이 존재한다[2,4,5]. 이러한 단점을 극복하기 위하여 최근 이미지 처리(image processing)나 컴퓨터 비전(computer vision) 등과 같은 분야의 연구 결과를 이용하여 이미지 데이터로부터 기본적인 특징을 자동으로 추출하여 이를 기반으로 검색을 수행하는 방법을 많이 연구하고 있다[2,3,4,5]. 이러한 검색 방법은 이미지 데이터의 특징이 행렬이나 벡터의 형태로 표현되기 때문에 완전 일치 검색(exact match)이 아닌 유사 검색(approximate match, or similarity retrieval)을 수행하여 사용자가 원하는 이미지와 유사한 이미지들을 검색한다. 즉, 질의 이미지가 주어졌을 때, 데이터베이스에 저장된 모든 이미지들 중에서 질의 이미지와 유사한 모양이나 색상, 질감 등을 가지는 이미지들을 검색할 수 있다. 따라서, 이러한 내용 기반 검색 기법은 시계열 데이터베이스(time sequence database), 과학 데이터베이스(scientific database), 이미지/비디오/오디오 데이터베이스(image/video/audio database) 및 DNA 유전자 데이터베이스(DNA/genome database)와 같은 다양한 응용에서 효율적이고 중요한 기술로 사용될 수 있다[6].

예를 들어, 주식 시장 관리를 위한 데이터베이스의 경우에 전형적으로 사용되는 질의로는 “주식 가격이 특정 회사 주식 가격과 유사하게 변동하는 회사들을 모두 검색하라” 등이 있을 수 있으며, 이미지 데이터베이스의 경우에는 “특정 이미지와 유사한 색상 분포를 가지고 있는 이미지를 모두 검색하라”, 의학 이미지 데이터베이스인 경우에는 “암세포의 질감과 유사한 것을 포함하는 X-ray 이미지를 모두 검색하라” 등과 같은 질의가 있을 수 있다.

내용 기반 검색 기법이 위와 같은 다양한 응용들에서 효과적이고 핵심적인 기술로 사용될 수 있으나, 아직까지 다음과 같은 두 가지 미해결 문제(open issues)로 인하여 실제로 개발하는데 많은 어려움이 따른다[7].

#### • 특징 추출의 완전성(Completeness of feature extraction) 문제

특징 데이터를 추출하는 과정의 완전성에 관한 문제이다. 즉, 이미지 데이터로부터 특징 추출 기법에 의해 추출된 특징 벡터(feature vector)를 이용하여 내용 기반 검색을 수행할 경우, 검색 조건을 만족하는 객체(qualified object)인데 실수로 검색해내지 못하는 경우인 false dismissals이 발생하지 않는다는 것을 보장해야 한다.

#### • 차원의 저주(Dimensionality “curse”) 문제

대부분의 다차원 색인 방법의 성능은 색인에 사용되는 데이터의 차원이 증가함에 따라 지수적(exponential)으로 감소한다. 그래서, 결국은 순차 검색을 하는 것보다 못한 성능을 나타내게 된다. 이러한 것을 일반적으로 차원의 저주 문제라 한다.

본 논문에서는 이미지의 전체적인 모양을 나타내는 특징 데이터를 자동으로 추출하여 사용자가 원하는 이미지와 유사한 모양을 가지고 있는 이미지들을 신속하고 효율적으로 검색할 수 있는 내용 기반 이미지 검색 기법을 제안한다. 그리고, 특징 추출의 완전성 문제를 해결하기 위하여 본 논문에서 제안하는 특징 추출 및 유사성 측정 방법을 이용하여 내용 기반 검색을 수행할 경우, 항상 false dismissals이 발생하지 않음을 증명한다. 또한, 차원의 저주 문제를 해결하기 위하여 최근 많은 연구가 수행되고 있는 고차원 데이터 색인 방법들 중에서 X-tree를 이용한 이미지 색인 방법을 사용한다. 본 논문에서 연구된 대규모 이미지 데이터베이스를 위한 효율적인 내용 기반 검색 기법은 디지털 화랑, 위성 이미지 데이터베이스, 의학 이미지 데이터베이스, 기타 상업적인 멀티미디어 정보 검색 시스템과 같은 대규모 멀티미디어 데이터베이스 시스템에서 효율적이고 핵심적인 기술로 사용될 수 있다.

### 1.2 논문의 구성

본 논문의 구성은 다음과 같다. 먼저, 2절에서는 그동안 수행되어 왔던 내용 기반 검색 및 다차원 색인 기법에 관한 관련 연구를 소개하고, 3절에서는 본 논문에서 제안하는 특징 추출 기법과 유사성 측정 기법, 그리고 신속한 검색을 지원하기 위한 내용 기반 이미지 색인 방법에 대하여 설명한다. 4절에서는 3절에서 제안된 여러 기법들을 기반으로 구현된 시스템의 전체적인 구

조를 설명하고, 실제 구현된 시스템상에서 내용 기반 이미지 질의를 위한 사용자 인터페이스 및 질의 검색 결과를 보인다. 5절에서는 본 논문에서 제시한 여러 기법들에 의해 구현된 내용 기반 이미지 검색 시스템의 검색 효율성(efficiency)과 적합성(effectiveness)을 다양한 실험을 통해 보인다. 마지막으로 6절에서는 요약 및 앞으로의 연구 방향에 대해서 설명한다.

## 2. 관련 연구

최근 신속하고 효율적인 내용 기반 검색에 관한 연구가 활발히 진행되고 있는데, 이 절에서는 이전에 수행되었던 내용 기반 검색 기법과 다차원 색인 기법에 대하여 살펴본다.

### 2.1 내용 기반 검색에 관한 연구

QVE[8]는 Hirata와 Kato가 만든 내용 기반 이미지 검색 시스템으로, 처음으로 시각적 질의를 이용한 내용 기반 검색 방법을 제안하였지만, 이미지로부터 추출한 외곽선 데이터를 그대로 특징 데이터로 사용하기 때문에 실제 질의 처리 및 검색 과정에서 유사성을 검사하기 위해서 데이터베이스에 저장된 각각의 이미지에 대해 이동(translation)이나 스케일링(scaling), 회전(rotating)과 같은 기하학적인 변환을 일일이 고려해야 한다는 단점을 가진다.

QBIC[3,5]은 IBM Almaden 연구소에서 개발된 내용 기반 이미지 검색 시스템으로, 이미지의 색상이나 질감, 모양 등과 같은 기본적인 특징에 기반한 다양한 시각적 질의를 제공한다. 또한, 데이터베이스에 저장된 이미지에 대하여 사람이 의미 정보를 부여할 수 있도록 허용함으로써 제한적이거나 의미 정보에 의한 검색도 가능하다. 그러나, 기본적인 특징과 의미 정보를 표현하는 논리적인 특징이 체계적으로 통합되어 있지 않으며, 다양하고 복잡한 질의 인터페이스가 체계적으로 통합되어 있지 않아서 오히려 사용자 측면에서는 이용하기가 어려울 수 있다는 단점이 존재한다. 또한, 특징 데이터를 완전 자동으로 추출하지 못하고 부분적으로 사람의 수작업을 이용해야 한다는 단점이 존재한다.

VisualSEEK[9]는 WWW를 이용한 내용 기반 이미지 검색 시스템으로 Columbia 대학에서 개발되었다. 이 시스템의 가장 큰 특징은 시각적 특징(visual feature)에 의한 내용 기반 질의(content-based querying)와 공간 정보에 의한 공간 질의(spatial querying)가 동시에 가능하다는 것이다. 즉, 먼저 이미지에서 색상 영역(color region)에 관한 정보를 추출한 후에, 이러한 정보를 이용하여 각 색상 영역간의 공간적인 관계를 추론

함으로써 공간 관계에 대한 특징을 추출하여 질의에 사용한다. 그러나, 이러한 추론 과정은 시간이 많이 소요되는 작업으로 비효율적일 수 있으며, 색상 정보에 의존한 검색만 가능하다는 단점을 가진다.

Jacobs[2] 등이 제안한 내용 기반 이미지 검색 시스템은 특징 데이터의 추출 과정에서 처음으로 웨이브릿 변환을 사용하였다. 그들의 연구에서는 칼라 이미지(color image)에 대하여 직접 웨이브릿 변환을 적용하여 얻은 계수(coefficients)들을 특징 벡터로 사용하는데, 칼라 이미지의 각 채널("R"ed, "G"reen, "B"lue)마다 따로 특징 벡터를 구성하며, 각 특징 벡터의 차원이 너무 커서 기존의 색인 기술을 이용하는데 적합하지 않다. 실제로, 그들의 연구에서는 특징 벡터의 추출 및 유사성 측정과 관련된 컴퓨터 비전 및 그래픽 기술에 초점을 두고 있으며, 효율적이고 신속한 검색을 중요시하는 데이터베이스 측면의 기술은 고려하지 않고 있다. 따라서, 그들의 연구에서는 효율적이고 신속한 검색을 지원하기 위한 색인 기술에 대한 언급이 없다.

본 연구에서는 [2] 연구의 단점을 극복하기 위하여 특징 벡터의 차원을 효율적으로 줄이면서 동시에 이미지의 전체적인 모양에 의한 효율적인 내용 기반 검색이 가능한 방법을 제안하고, 이러한 검색 과정에서 항상 false dismissals이 발생하지 않음을 보인다. 또한, 특징 벡터 추출 과정을 통하여 추출한 특징 벡터에 적합한 색인 구조를 이용하여 색인을 구성하고 검색함으로써 대규모 이미지 데이터베이스에서도 효율적이고 신속한 내용 기반 검색을 지원하는 방법을 제시한다.

### 2.2 다차원 색인(Multi-dimensional indexing) 기술에 관한 연구

최근 다양한 멀티미디어 데이터 처리에 대한 요구가 증가함에 따라 데이터베이스 측면에서는 유사 검색에 대한 필요가 증가하게 되었다.

Shasha와 Wang[10]은 삼각 부등식(triangular inequality)과 불필요한 검색을 사전에 제거하는 색인 기법을 제안하였다. 그러나, 이 색인 기법의 공간 사용은 데이터베이스에 저장되는 객체의 수에 2의 지수승태(quadratic)로 증가하기 때문에 대용량 데이터베이스에는 적절하지 못하다는 단점이 있다.

Aurenhammer[11]는 최근점 질의(nearest neighborhood query)에 사용되는 Voronoi 다이어그램을 제안하였다. 이 색인 방법은 2차원 공간에서는 잘 동작하나, 3차원 이상으로 확장되기 위해서는 상당히 복잡한 변형 과정이 필요하며, 더 높은 차원에서는 제대로 동작하지 않는다는 단점을 가지고 있다.

이외에 K-D-B tree[12], R\*-tree[13]와 나머지 R-tree 계열[14,15], Linear quadtrees[16], Grid-files[17] 등이 있는데, 이것들은 색인에 사용되는 데이터의 차원이 저차원이나 중차원일 경우에는 잘 동작하지만, 고차원 데이터들에 대해서는 제대로 동작하지 못한다. 즉, 데이터의 차원이 증가함에 따라 검색에 들어가는 비용이 지수적으로 증가하기 때문에, 결국에는 순차 검색을 하는 것보다 못한 결과를 초래한다.

이러한 기존의 다차원 색인 방법의 단점을 극복하기 위하여 최근에는 고차원 데이터 집합을 위한 여러가지 고차원 색인(high-dimensional index) 구조들이 제안되고 있다.

TV-tree[18]는 고차원 데이터에서 색인에 사용되는 차원의 수를 줄이고 동적 차원(active dimension)의 이동(shifting or telescoping)을 통하여 고차원 데이터를 효율적으로 색인할 수 있다. 즉, 차원의 감소는 고차원 데이터를 각 차원의 중요도에 따라 순서화한 후에 그 중에서 몇 개의 중요한 차원만 색인에 사용하는 방법이고, 차원의 이동은 만약 서브트리(sub-tree)상에서 색인에 사용되는 동적 차원의 값들이 서로 동일할 경우에, 그 차원을 무시하고 다음에 있는 중요도가 떨어지는 차원들의 값을 색인에 이용하는 방법이다. 이렇게 함으로써 R\*-tree의 성능을 향상시켜 고차원 데이터를 효율적으로 색인할 수 있으나 특징 데이터들의 차원이 중요도에 따라 순서화되어야 하며, 동적 차원의 이동이 가능한 특징 데이터들이어야 한다는 단점이 존재한다.

X-tree[19]는 TV-tree와 마찬가지로 R\*-tree의 한 변형으로 겹침이 없는 분리(overlap-free split) 방법과 슈퍼노드(supernode) 개념을 이용하여 R\*-tree의 성능을 향상시킨 것이다. 겹침이 없는 분리 방법은 K-D-B 트리처럼 탐색 공간을 서로 분리된 다른 공간으로 분리하여 검색을 효율적으로 할 수 있도록 한다. 또한, 슈퍼노드는 노드들 사이의 겹침을 극복하기 위하여 사용되는 확장 가능한 노드이다.

이외에도 고차원 데이터의 효율적인 색인을 위한 구조들로 VAMSplit R-tree[20]나 SS-tree[21], SRtree[22] 등이 있으며, 모두 트리 기반의 색인 방법들이다.

### 3. 내용 기반 이미지 검색 시스템

이 절에서는 본 논문에서 제안하는 내용 기반 이미지 검색 시스템에서 사용하는 특징 추출 기법과 유사성 측

정 기법, 그리고 대규모 이미지 데이터베이스에서의 신속한 검색을 지원하기 위한 내용 기반 이미지 색인 방법에 대하여 설명한다.

#### 3.1 특징 추출 과정(Feature Extraction)

이미지로부터 검색에 사용되는 특징 벡터를 추출하기 위하여 먼저 RGB의 세 채널을 갖는 완전 칼라 이미지를 한 채널을 갖는 흑백 이미지로 변환한 후에 외곽선 데이터를 추출하고 이 외곽선 데이터에 대하여 웨이브릿 변환을 적용한다. 이러한 결과로 생성되는 웨이브릿 계수들 중에서 가장 큰 값을 갖는 몇 개의 계수들만을 취하여 최종적으로 검색에 사용하는 특징 벡터를 추출한다.

웨이브릿 변환은 기본 함수(basic function)로 사인(sine), 코사인(cosine) 함수 뿐만 아니라 다양한 웨이브릿 모함수(wavelet mother function)를 사용할 수 있어서 신호 데이터(signal data)에 불연속성이 있거나 특별히 날카롭게 도출된 부분이 있을 경우, 이러한 지역적인 특징을 분석하는데 매우 유용하며, 신호의 전체적인 특징을 추출하는데도 매우 효과적이라는 장점을 가지고 있다[23,24,25]. 또한, 웨이브릿 변환은 이미지에 대한 다해상도 분석(multiresolution analysis)이 가능하다는 장점을 가지고 있다[23,24,25]. 이것은 이미지에 웨이브릿 변환을 적용하여 추출한 특징 벡터는 이미지의 해상도에 대하여 독립적인 정보를 제공한다. 즉, 해상도가 서로 다른 두 이미지의 전체적인 모양이 유사하다면, 두 이미지는 해상도에 관계없이 유사한 특징 벡터를 가진다. 이러한 이유로 질의 이미지와 데이터베이스에 저장된 이미지간의 해상도(resolution)가 서로 다르더라도 검색이 가능하게 된다.

실제로, 본 연구에서 SOP(SNU OODBMS Platform)<sup>2)</sup> 이미지 데이터베이스에 저장된 이미지들의 해상도는 128×128이고, 질의 이미지는 512×512로 서로 다르지만, 유사한 이미지에 대해서는 거의 유사한 특징 벡터들이 얻어짐을 확인할 수 있었다. 이러한 장점으로 인하여 웨이브릿 변환은 이미지 압축이나 특징 추출 및 시계열 데이터에서의 잡음(noise) 제거와 같은 여러 분야에서 최근 많이 사용되고 있다.

본 논문에서는 이러한 유용한 특성을 가지고 있는 웨이브릿 변환을 이용하여 이미지로부터 특징 벡터를 추출한다. 알고리즘 1은 완전 칼라 이미지로부터 특징 벡터를 추출하는 과정이다.

1) 고차원 데이터들 중에서 실제 색인에 사용되는 차원을 말한다

2) 서울대학교 컴퓨터 공학과 객체지향 시스템 연구실에서 개발한 ODIMG-98 표준을 따르는 객체지향 데이터베이스 시스템이다.

**알고리즘 1: 특징 추출 과정**

- 입력 : 완전 칼라 이미지(full color image)
  - 출력 :  $f$ -차원의 특징 벡터
1. 완전 칼라 이미지를 흑백 이미지( grayscale image)로 변환한다.
  2. 흑백 이미지로부터 외곽선을 추출한다.
  3. 2의 과정에서 추출된 외곽선 데이터에 대하여 웨이브릿 변환을 적용하여 특징 벡터를 추출한다.
    - 3.1 외곽선 데이터에 대하여 웨이브릿 변환을 적용한다.
    - 3.2 웨이브릿 변환을 적용한 결과로 나오는 웨이브릿 계수(Wavelet coefficient)를 정규화(normalization)한다.
    - 3.3 정규화한 웨이브릿 계수를 정렬하여 가장 큰 계수값(coefficient value)을 갖는  $f$ 개의 계수만을 취하고 나머지 계수들을 무시하는 절단(truncation) 작업을 수행한다.
    - 3.4 3.3의 절단 과정을 거쳐 생성된  $f$ 개의 계수값을 이용하여  $f$ -차원의 특징 벡터를 구성한다.
  4. 3.4의 과정을 통해 생성된 특징 벡터를 최종적으로 검색에 사용하는 특징 벡터로 한다.

[2]의 연구에서는 완전 칼라 이미지에 대하여 직접 웨이브릿 변환을 적용하여 칼라 이미지의 각 채널("R"ed, "B"lue, "G"reen)마다 따로 특징 벡터를 구성하기 때문에, 실제 한 이미지에 대한 특징 벡터의 차원이 너무 커서 색인을 구성하여 검색하기에 적합하지 않으며, 순차 검색을 할 경우에도 상당히 비효율적일 수 있다<sup>3)</sup>. 따라서, 대규모 이미지 데이터베이스를 검색하는 경우에는 적합하지 못하다는 단점이 존재한다. 반면, 본 논문에서는 알고리즘 1에서의 같은 과정을 통하여 특징 벡터의 차원을 효율적으로 줄임으로써 기존의 다차원 색인 구조를 이용할 수 있으며, 따라서 대규모의 이미지 데이터베이스에 대해서도 효율적이고 신속한 검색을 보장한다는 장점을 가지고 있다.

또한, [2]의 연구에서는 칼라 이미지에 대하여 직접 웨이브릿 변환을 적용하여 특징 벡터를 추출했을뿐 그 특징 벡터가 이미지의 어떤 속성을 표현하는지는 명확히 구별할 수가 없다. 따라서, 그들의 연구에서는 질의 이미지의 전체적인 모양이 질의 대상이 되는 데이터베이스에 저장된 이미지와 유사하더라도 색상의 구성이 다르게 되면 사용자가 원하는 이미지들을 제대로 찾지

못하는 결과를 초래한다. 본 논문에서 제안하는 방법은 이미지의 외곽선 데이터에 대하여 특징 벡터를 추출하므로, 특징 벡터는 이미지의 전체적인 모양에 대한 특징을 표현한다. 따라서, 사용자는 자신의 질의 이미지와 유사한 모양을 가지고 있는 이미지들을 일관성 있게 검색할 수 있다

**3.2 유사성 측정 과정(Similarity Measurement)**

1절에서도 언급했듯이 이미지 데이터에서 추출한 특징 벡터들 간의 유사성을 측정하여 내용 기반 검색을 수행할 경우, 항상 false dismissals이 발생하지 않음을 보이는 것은 중요한 문제이다. 즉, 내용 기반 멀티미디어 정보 검색 시스템의 경우에는 제안하는 특징 추출 기법이나 유사성 측정 기법을 사용하여 질의 검색을 수행할 경우, 항상 검색 조건을 만족하는 모든 적합한 객체(qualified object)들이 질의에 대한 결과로 검색됨을 보장해야한다[5].

이 절에서는 3.1절에서 제안한 특징 추출 기법을 사용하여 추출된 특징 벡터를 이용하여 이미지 객체간의 유사성을 측정하는 기법을 설명하며, 이러한 방법을 이용하여 내용 기반 검색을 수행할 경우 항상 false dismissals이 발생하지 않음을 증명한다.

본 논문에서는 이미지들 간의 유사성(비유사성)을 측정하기 위하여 유클리드 거리 함수(Euclidean distance function)를 이용한다. 유클리드 거리 함수의 특징은 정규 직교 변환(orthonormal transform)하에서 그 거리가 항상 보존된다는 것이다<sup>4)</sup>. 웨이브릿 변환은 정규 직교 변환이며, 따라서 이러한 변환을 적용하여 생성된 특징 벡터는 항상 유클리드 공간상에서 그 거리가 보존된다. 이러한 성질을 이용하여 본 논문에서 제안하는 특징 추출 과정과 유사성 측정 과정에서 항상 false dismissals이 발생하지 않음을 보일 수 있다.

[정의 1] 두 개의 이미지 객체  $O_1, O_2$  간의 유클리드 거리는 다음과 같이 표현된다.

$$D(O_1, O_2)$$

이미지 객체로부터 추출한  $f$ -차원의 특징 벡터는 이미지 객체를  $f$ -차원의 특징 공간(feature space) 상의 한 점으로 표현할 수 있다 따라서 다음의 정의가 가능하다.

[정의 2]  $F_s(O)$ 는 객체  $O$ 를  $f$ -차원의 특징 공간상의 한 점으로 사상(mapping)시키는 함수이다.

3) 실제로 [2] 연구에서는 각 채널당 40~60개의 웨이브릿 계수 로 특징 벡터를 구성한다

4) 유클리드 공간상에서 정규 직교 변환이 가지는 성질에 대해서는 [5] 연구를 참고하십시오.

위의 [정의 1], [정의 2]를 이용하여, 본 논문에서 제안한 특징 추출 및 유사성 측정 기법을 이용하여 내용 기반 검색을 수행할 경우 false dismissals이 발생하지 않음을 보이는데 필요한 다음의 정리를 유도할 수 있다.

[정리 1]

$$D(\vec{Q}, \vec{T}) \leq \epsilon \Rightarrow D(F_e(\vec{Q}), F_e(\vec{T})) \leq \epsilon \quad (1)$$

- $\epsilon$  : 사용자가 정의한 유사성 허용 오차(tolerance)
- $Q, T$  : 질의 이미지(query image) 객체와 질의의 대상이 되는 데이터베이스에 저장된 대상 이미지(target image) 객체
- $\vec{Q}, \vec{T}$  :  $Q, T$ 에 대하여 웨이브릿 변환을 적용하여 얻은 계수들로 구성된 벡터

[증명] 특징 추출시 웨이브릿 변환을 이용하는데, 웨이브릿 변환은 정규 직교 변환으로 거리가 보존되는 변환(distance-preserving transform)이다. 만약 웨이브릿 변환을 적용하여 얻은 웨이브릿 계수들을 정렬하여 그 중에서 가장 큰 값을 갖는  $f(\leq n)$ 개의 계수만을 취하여 특징 벡터로 사용한다면 다음과 같은 식이 가능하다

$$D^2(F_e(\vec{Q}), F_e(\vec{T})) = \sum_{i=0}^{f-1} |Q_i - T_i|^2 \leq \sum_{i=0}^{f-1} |Q_i - T_i|^2 = D^2(\vec{Q}, \vec{T}) \quad (2)$$

따라서, 다음과 같은 식이 성립된다.

$$D(F_e(\vec{Q}), F_e(\vec{T})) \leq D(\vec{Q}, \vec{T}) \leq \epsilon \quad (3)$$

식 (3)에 의해 [정리 1]이 항상 만족됨을 알 수 있다. [Q.E.D.]

따라서, 본 논문에서 제안한 특징 추출 과정을 통하여 추출된 특징 벡터들 간의 차이는 실제 이미지들 간의 차이(거리)보다 항상 작기 때문에 false dismissals이 발생하지 않는다. 다시 말해, 항상 검색 조건을 만족하는 적합한 객체들을 모두 검색함을 보장할 수 있다. 그러나, 이것은 검색 조건을 만족하지 못하는 객체인데 실수로 검색해내는 경우(false alarm)가 발생하지 않음을 보장하지는 못한다. 실제, 질의 결과로 검색된 적절하지 못한 객체들은 후처리 과정(post-processing)을 통하여 제거할 수 있다.

3.3 내용 기반 이미지 색인(Content-based Image Indexing)

대규모 이미지 데이터베이스 상에서 효율적이고 신속

한 검색을 지원하기 위해서는 이미지 색인 기법이 필수적이다. 이 절에서는 대규모 이미지 데이터베이스를 대상으로 하는 내용 기반 검색 시스템에서 신속한 검색을 지원하기 위한 효율적인 이미지 색인 기법에 대하여 설명한다.

기존의 대부분의 내용 기반 검색 시스템들은 공간 데이터베이스(spatial database) 상에서 많이 사용되는 R\*-tree를 이용한 색인 방법을 제공하고 있다. 그러나, R\*-tree는 색인에 사용되는 데이터 차원이 저차원이거나 중차원인 경우에는 효율적인 색인을 제공하나, 10차원을 넘는 고차원 데이터인 경우에는 효율적인 색인을 제공하지 못한다[18,19]. 따라서, 저차원이나 중차원 데이터를 주로 사용하는 공간 데이터베이스에서는 R\*-tree를 효율적으로 사용할 수 있으나, 고차원 데이터를 주로 사용하는 내용 기반 검색 시스템에서는 비효율적이다[18,19].

최근 이러한 제약을 극복하기 위하여 고차원 데이터를 위한 색인 구조들이 제안되고 있다[18,19,20,21,22]. 이러한 고차원 데이터 색인 구조들은 각기 다른 접근 방법과 특징을 가지고 있기 때문에 특정 응용 환경에 적용할 때는 그 특성을 잘 파악하여 응용에 적합한 색인 구조를 선택하는 것이 바람직하다. 특정 고차원 데이터 색인 구조를 본 연구에서와 같은 내용 기반 검색 시스템에 응용하기 위해서는 다음과 같은 사항들을 고려해야 한다.

- **특징 벡터의 성질** - 특징 추출 과정을 통해 생성된 특징 벡터를 사용하여 내용 기반 검색을 수행할 경우, 만약 특징 벡터의 각 차원의 중요도를 고려하여 가중치를 두어 검색한다면 그러한 조건을 고려한 색인 구조를 이용하는 것이 효율적이다. 즉, 특정 색인 기술을 이용할 경우 항상 사용하는 특징 벡터의 성질이 잘 맞는지 여부를 판단하는 것이 중요하다.
- **유사성 측정 함수의 성질** - 내용 기반 검색에서 두 이미지간의 유사성을 측정하기 위해 사용하는 함수의 성질을 잘 알아야 한다. 유클리드 거리 함수는 많은 경우에 유용하게 사용될 수 있으나 모든 경우에 일반적으로 적용될 수는 없다. 만약 색상 히스토그램처럼 특징 벡터들 사이에 간섭(cross-talk)이 존재한다면 이러한 것을 반영할 수 있는 2의 지수승배 형식을 가지는 거리 함수(quadratic distance bounding)를 이용해야 한다[5]. 만약 이러한 거리 함수를 사용한다면 어떠한 다차원 색인 구조도 그대로 사용할 수 없다[5]. 따라서, 내용 기반 검색 시스템에 색인 기술을 적용할 때는 사용하는 유사성 측정 함수의 성질을 잘

5) 식 (2)에 대한 상세한 증명은 [5]를 참고하십시오.

파악해야 한다.

- **사용하는 질의의 종류** - 내용 기반 이미지 검색에서 사용하는 질의의 종류를 잘 파악하여 이에 적합한 색인 구조를 이용하는 것이 바람직하다. 어떤 색인 구조들은 특정 질의에 대한 성능을 향상시키는데 초점을 두어 개발하였기 때문에 다른 종류의 질의에 대해서는 비효율적일 수 있다[21,22,26,27].

본 연구에서는 최근 발표된 중요한 고차원 데이터 색인 구조인 TV-tree[18], X-tree[19], VAMSplit R-tree[20], SS-tree[21], SR-tree[22]를 대상으로 위의 세 가지 사항을 고려한 결과, 본 연구의 응용 환경에 가장 적합한 색인 구조로 X-tree를 선정하였다. 그 이유는 다음과 같다.

TV-tree는 색인에 사용되는 데이터들이 차원의 중요도에 따라 순서화되어 있어야 한다. 그러나, 본 연구에서 사용하는 특징 벡터는 이미지의 외곽선 데이터에 대하여 웨이브릿 변환을 적용한 결과로 생성되는 웨이브릿 계수들 중에서 그 값이 가장 큰 몇 개의 계수들만을 추출하여 구성되는데, 각 차원의 중요도는 고려하지 않는다. 즉, 특징 벡터의 앞쪽에 있는 차원들이 뒤쪽에 있는 차원들보다 중요하지 않으며, 특징 벡터의 각 차원의 중요도는 모두 동일하다. 따라서, 본 응용 환경에서는 TV-tree를 이용하기에는 적합하지 않다. 또한, VAMSplit R-tree는 색인에 사용되는 고차원 데이터들에 대한 구성이나 분포에 대한 정보를 사전에 알고 있을 때 가능한 색인 방법이다. 따라서, 동적인 갱신을 제공하지 못한다[20, 22]. 그러나, 본 연구에서 사용하는 특징 벡터들은 동적으로 생성되는 정보들로 사전에 그 구성이나 분포를 알 수가 없으며 특징 벡터의 삽입이나 삭제가 자유롭게 일어날 수 있다. 그리고, SS-tree나 SR-tree는 주로 고차원 데이터에 대한 최근접 질의(nearest neighborhood query)의 성능을 향상시키는데 초점을 둔 색인 구조이나, 본 연구에서 사용하는 내용 기반 질의는 범위에 의한 질의(range query) 방법이다. X-tree는 관련 연구에서 기술한 여러 특징들을 가지고 있으며 색인에 사용되는 특징 데이터에 대한 제한이 없다. 또한, 특정 질의에 대한 성능을 향상시키기 위하여 설계된 색인 구조가 아니며 동적 갱신이 가능한 색인 구조이다[19].

본 논문에서는 위와 같은 이유들로 인하여 여러 고차원 색인 구조들 중에서 X-tree 색인 구조를 선정하였으며, 이를 내용 기반 이미지 색인 방법으로 사용한다.

특징 추출 과정을 통하여 추출된 특징 벡터들은 다차원 특징 공간상에서 하나의 점으로 표현할 수 있다[6,7].

따라서, 다음과 같은 알고리즘 2를 이용하여 많은 양의 이미지 객체들로부터 추출한 특징 벡터들로 색인을 구성하고, 알고리즘 3에 의해 다차원 공간상에서 질의 이미지 객체에 해당하는 점으로부터 일정한 유사성 허용 오차 안에 있는 모든 점들을 검색함으로써 질의 이미지와 유사한 이미지들을 신속하게 검색할 수 있다. 이 방법은 멀티미디어 객체를 색인하는 일반적인 방법으로 제안된 GEMINI[6]에 기반을 두고 있다.

**알고리즘 2: 내용 기반 이미지 색인의 구성**

- 입력 : 데이터베이스에 적재할 이미지들
- 출력 : X-tree 색인 파일

1. 데이터베이스에 적재할 모든 이미지들에 대하여 다음과 같은 과정을 반복 수행한다.

- 1.1 이미지로부터 알고리즘 1을 이용하여  $f$ -차원의 기본 특징 벡터를 추출한다.
- 1.2 1.1의 과정에서 추출된 특징 벡터를 X-tree에 삽입함으로써 색인을 구성한다.

다음의 알고리즘 3은 알고리즘 2에 의해 생성된 색인을 이용하여 질의 이미지와 유사한 이미지들을 검색하는 과정이다.

**알고리즘 3: 내용 기반 이미지 색인의 검색**

- 입력 : 질의 이미지, X-tree 색인 파일
- 출력 : 질의 이미지로부터  $\epsilon$ 의 유사성 허용 오차 안에 있는 모든 이미지들의 포인터<sup>6)</sup>

1. 질의 이미지 객체(Q)로부터 알고리즘 1을 이용하여 검색에 사용하는 특징 벡터를 추출한다.
2. 1의 과정에서 추출된 질의 이미지의 특징 벡터를 다차원 특징 공간상의 한 점  $F_q(\vec{Q})$ 로 사상한다.
3. 2의 과정을 통하여 얻어진 점과  $\epsilon$ 을 이용하여 범위에 의한 질의를 수행한다. 즉, 알고리즘 2에 의해 만들어진 색인을 이용하여 특징 공간상에서  $F_q(\vec{Q})$ 로부터  $\epsilon$ 의 유사성 허용 오차 안에 있는 모든 점들을 검색한다
4. 3의 과정을 통해 검색된 점들과 질의 이미지 객체에 해당하는 점간의 실제 거리를 계산하는 후처리 과정(post-processing)을 통하여 거짓으로 검출된 객체들을 제거한다.
5. 4의 과정을 통해 얻어진 점들에 해당하는 포인터를 반환한다.

알고리즘 3의 과정 4에서 후처리 과정이 필요한 이유는 그림 1로 설명될 수 있다. 그림 1은  $f$ -차원의 특징 공간

6) 실제로는 객체지향 데이터베이스 시스템(SOP)에서 객체를 구별하기 위해 사용하는 객체 식별자(OID)이다.

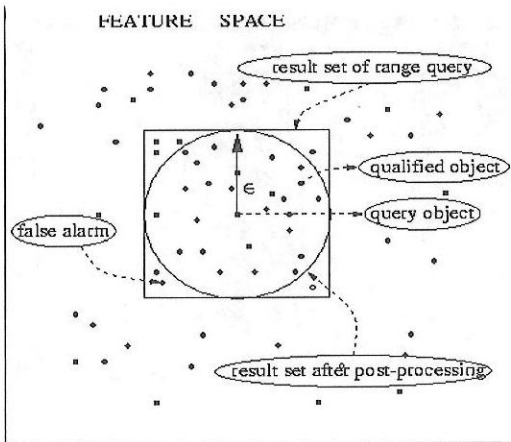


그림 1 f-차원의 특징 공간상에서의 내용기반 질의 및 후처리 과정

상에서 X-tree를 이용한 내용 기반 이미지 질의 처리 및 후처리 과정을 보여주는 개념적 그림이다. 알고리즘 1에 의해 이미지들로부터 추출된 특징 벡터들은 그림 1에서와 같이 특징 공간상의 점들로 표현될 수 있다. 또한, 알고리즘 3의 1~2과정을 통하여 질의 이미지는 그림 1에서 사각형 중앙에 있는 점으로 표현되고, 과정 3의 범위에 의한 질의를 거쳐 사각형 안에 있는 점들을 검색해 낸다. 이처럼 사각형 안에 있는 점들을 검색하는 것은 X-tree도 R\*-tree와 마찬가지로 삽입이나 영역 검색이 사각형을 기반으로 이루어지기 때문이다. 따라서, 3의 과정으로 검색을 끝내게 되면 그림 1에서와 같이 거짓으로 검출된 객체들이 결과로 남아 있게 된다. 그래서, 4의 과정에서 이렇게 거짓으로 검출된 객체들을 제거하는 후처리 작업을 수행하며, 그 결과 그림 1에서와 같이 원안에 있는 적합한 객체들만을 질의의 결과로 얻어낸다. 마지막으로, 5의 과정에서는 SOP 이미지 데이터베이스에 저장된 유사한 이미지들의 포인터들을 반환한다.

알고리즘 2,3은 본 연구에서와 같은 응용 환경하에서 기존에 존재하는 고차원 데이터를 위한 색인 구조인 X-tree를 이용하여 색인을 구성하고, 내용에 기반한 질의 및 검색을 수행하는 총체적인 방법이다.

#### 4. 내용 기반 검색 시스템의 구조 및 질의 검색 결과

이 절에서는 3절에서 제안한 특징 벡터의 추출 기법과 유사성 측정 기법, 그리고 X-tree를 이용한 내용 기

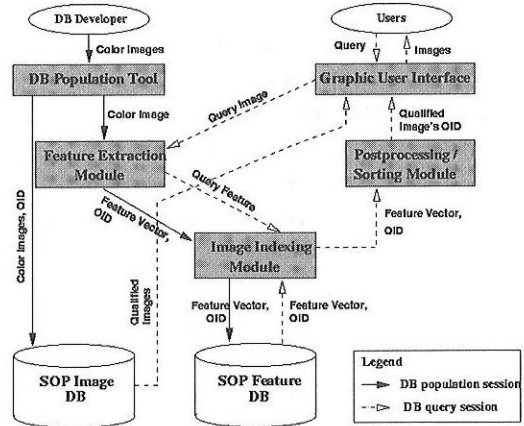


그림 2 내용 기반 이미지 검색 시스템의 전체적인 구조

반 이미지 색인 기법을 이용하여 구현된 내용 기반 이미지 검색 시스템의 전체적인 구조와 실제 구현된 시스템 상에서의 질의 및 검색 결과를 설명한다.

#### 4.1 내용 기반 이미지 검색 시스템의 구조

그림 2는 본 연구에서 제안된 내용 기반 이미지 검색 시스템의 전체적인 구조를 보여주고 있다.

이미지 데이터베이스 적재 단계(DB population session)에서는 특징 추출 모듈을 이용하여, 많은 양의 이미지들로부터 특징 벡터를 추출하고, 이 특징 벡터들과 해당 OID들을 이용하여 색인을 구성한다. 그리고 본래의 이미지 데이터와 특징 벡터 및 식별자는 SOP내의 이미지 데이터베이스와 특징 데이터베이스에 각각 저장한다.

데이터베이스 질의 단계(DB query session)에서는 시스템에서 제공하는 사용자 인터페이스를 이용하여 질의 이미지를 구성한 후, 이 질의 이미지에 대하여 마찬가지로 특징 추출 과정을 통하여 특징 벡터를 추출하고, 이미지 데이터베이스 적재시 구축된 색인구조를 이용하여 일정한 유사성 허용 오차를 부과하여 질의를 수행한다. 후처리 및 정렬 모듈에서는 이러한 질의 수행 결과로 얻어지는 특징 벡터들과 해당 OID들 중에서 거짓으로 검출된 객체들을 제거하는 후처리 과정을 수행하고, 식 (4)를 이용하여 이미지간의 유사성을 사용자가 인지하기 쉬운 백분율로 변경하여 유사한 순서대로 정렬한다.

$$\text{Similarity}(Q, T) = 100 \times \left( \frac{\epsilon - D(F_e(\vec{Q}), F_e(\vec{T}))}{\epsilon} \right) (\%) \quad (4)$$

사용자 인터페이스에서는 후처리 및 정렬 모듈을 거



처 얻어진 적합한 이미지 객체들의 *OID*를 이용하여 SOP 이미지 데이터베이스에 저장되어 있는 해당 이미지 객체들을 가져와서 사용자에게 보여준다.

### 4.2 질의 및 검색 결과

본 연구에서 제안하는 내용 기반 이미지 검색 시스템에서는 두 가지 시각적 질의(visual query)를 제공한다. 하나는 사용자가 직접 그래픽 도구를 사용하여 그린 그림과 비슷한 이미지를 검색하는 사용자 그림에 의한 질의이며, 다른 한가지는 기존에 존재하는 스틸 이미지(still image)와 같은 예제 이미지를 이용하여 비슷한 이미지를 검색하는 예제 이미지에 의한 질의이다.

그림 3, 4는 이 두 가지 질의 방법 중 사용자 그림에 의한 질의와 그 검색 결과를 보여주고 있다. 그림 3은 구현된 시스템에서 제공하는 그래픽 도구를 이용하여 사용자가 검색하고자 하는 이미지와 유사한 이미지를 그린 화면을 보여준다. 그림 3에서 사용자가 찾고자 하는 그림을 그린 후, 좌측 하단에 있는 검색 버튼("Search")을 클릭하면 검색이 이루어지게 되며, 그림 4는 이러한 질의에 대한 검색 결과이다. 그림 4에서 위로 갈수록, 그리고 왼쪽에 있을수록 질의 이미지와의 유사성이 높은 그림이다. 즉, 왼쪽 가장 상단에 있는 것이 가장 유사한 그림이고, 그 오른쪽에 있는 것이 두 번째로 유사한 그림이 된다. 그림 4를 보면 그림 3의 질의 이미지와 유사한 사람의 모양을 나타내는 이미지들이 검색됨을 볼 수 있다. 또 다른 사용자 질의 방법인 예제 이미지에 의한 질의는 사용자 그림에 의한 질의 방법과 동일한 방법으로 사용 가능하며, 단지 질의 이미지를 사람이 직접 그리는 것이 아니라 시스템에서 제공하는 적절한 예제 이미지들 중에서 하나를 선택하여 질의를 한다는 것만이 다르다.

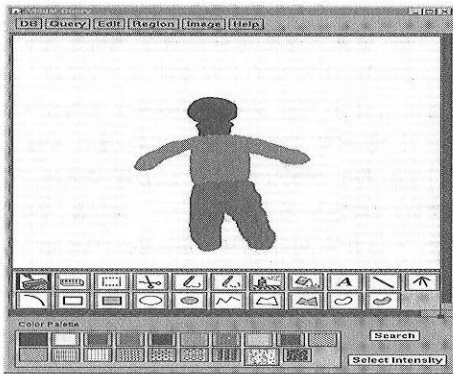


그림 3 사용자 그림에 의한 질의

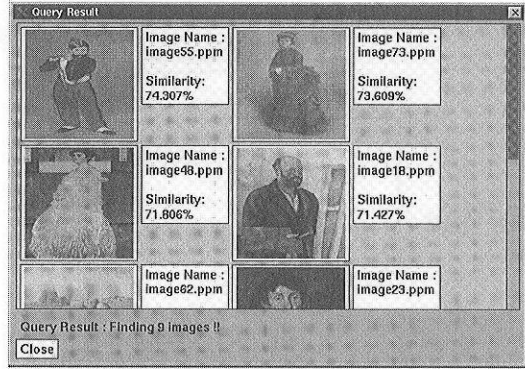


그림 4 그림 3의 검색 결과

이러한 두 가지 사용자 질의 방법을 제공하는 것은 사용자가 이 두 가지 질의 방법을 상호 보완적으로 사용함으로써 자신이 원하는 이미지들을 좀 더 쉽게 찾을 수 있도록 도와주기 위해서이다. 즉, 사용자가 대규모 이미지 데이터베이스에서 어떤 이미지를 찾고자 할 때, 먼저 시스템에서 제공하는 적절한 예제 이미지들 중에서 자신이 찾고자 하는 이미지와 유사한 예제 이미지가 있으면 이를 선택하여 질의를 수행하면 되고, 만약 유사한 예제 이미지가 없으면 직접 그려서 원하는 이미지들을 검색하면 된다.

그림 4에서 백분율로 표현되는 이미지간의 유사성은 상대적인 값으로 유사성 허용 오차의 값에 따라 변경될 수 있는 값들이다.

## 5. 실험 및 결과 분석

본 논문에서 제안된 내용 기반 검색 시스템은 Unix 상에서 C/C++ 언어와 X/Motif 언어를 이용하여 구현되었으며, 구현된 시스템의 성능을 측정하기 위하여 크게 두 가지 종류의 실험을 수행하였다. 첫 번째 실험은 검색의 효율성을 측정하기 위하여 대규모의 이미지 데이터베이스를 대상으로 검색에 소요되는 시간을 측정하였으며, 두 번째 실험은 검색의 적합성을 측정하기 위하여 5개의 대표적인 질의 이미지를 선정하여 각각의 질의 이미지에 대하여 얼마나 적합한 이미지들이 검색되는지를 측정하였다.

### 5.1 검색의 효율성(Retrieval Efficiency)

본 실험에서는 검색의 효율성을 측정하기 위하여 128 × 128의 해상도를 가지는 이미지 1,000여개에서 추출한 특징 벡터와 13,000여개의 임의로 생성된 특징벡터를 이용하여 실험하였으며, 64M 바이트의 주기억 장치와 4G

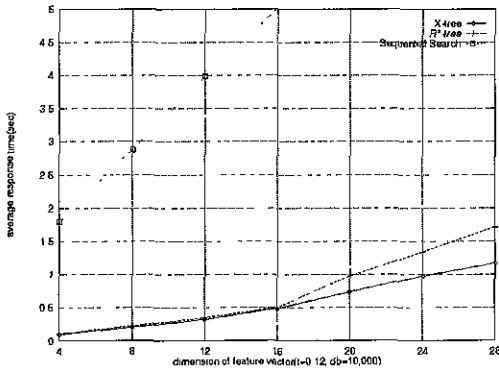


그림 5 특징 벡터의 차원에 따른 검색 시간: 데이터베이스의 크기가 10,000이고 유사성 허용 오차가 0.12일 경우

바이트의 보조 기억장치를 가지는 Sun Sparc 10 상에서 10개의 질의 이미지들에 대한 평균 응답 시간을 측정하였다<sup>7)</sup>. 질의 이미지의 전처리에 소요되는 시간은 모든 검색 방법에 동일하게 적용되므로 평균 응답 시간에서 제외시켰으며, R\*-tree나 X-tree를 이용한 색인 경우에는 후처리 과정에 소요되는 시간을 평균 응답 시간에 포함시켰다. 그리고, 페이지 크기는 모두 4K 바이트로 고정시켜 실험하였다.

그림 5는 데이터베이스에 저장된 이미지 객체의 수(특징 벡터의 수와 같음)가 10,000개일 경우, 특징 벡터의 차원에 따른 응답 시간의 변화를 보여주고 있다. R\*-tree를 이용한 색인 방법과 X-tree를 이용한 색인 방법의 경우, 모두 순차 검색보다 훨씬 빠른 응답 시간을 보임을 알 수 있다. 하지만, 특징 벡터의 차원이 16 차원을 넘으면, R\*-tree를 이용한 색인 방법은 일반적으로 X-tree를 이용한 색인 기법에 비해 응답 시간이 떨어짐을 알 수 있다.

최근 [27] 연구에서는 고차원 데이터에 대하여 순차 검색이 R\*-tree나 X-tree를 이용한 색인 방법에 비하여 보다 효율적이라는 연구 결과를 발표했다. 그들의 실험에서는 고차원 데이터들이 균등하게 분포되어 있는 환경(uniform distribution)에서 실험하였으며, 순차 검색의 경우에는 주기억 장치의 공간이 허용되는대로 가능한 많은 양의 데이터를 한번에 읽어들이어서 실험하였다[26]. 그러나, 고차원 데이터 색인 구조의 성능은 데

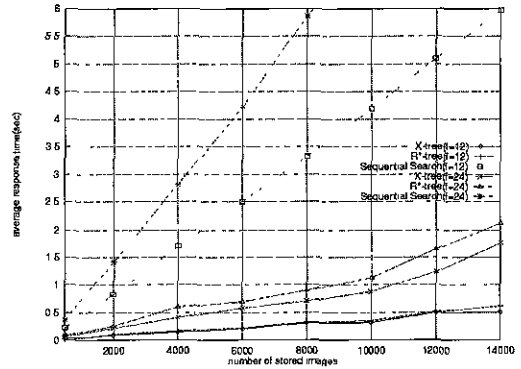


그림 6 데이터베이스의 크기에 따른 검색 시간: 특징 벡터의 크기가 12, 24이고, 유사성 허용 오차가 0.12일 경우

이터들의 분포에 많은 영향을 받는다. 즉, 데이터들이 클러스터링되어 있다면, 고차원 색인 구조는 좋은 성능을 보이나, 클러스터링이 잘 되어 있지 않다면 성능이 떨어진다[22,26].

본 실험에서 사용된 데이터들은 정규화 과정을 거치기 때문에 고차원 단위 공간상에서 하나의 점으로 표현된다. 이러한 전체 단위 공간상에서 질의 영역이 차지하는 비율을 측정함으로써 클러스터링 정도를 추정할 수 있다.

그림 5에서 데이터베이스의 크기가 10,000개이고 유사성 허용 오차( $\epsilon$ )가 0.12일 경우, 데이터의 차원이 20일 때 전체 단위 공간상에서 질의 영역이 차지하는 비율은  $(2 \times 0.12)^{20} = 10^{-13}$  으로<sup>8)</sup> 상당히 적은데도 불구하고, 10개 이상의 질의 결과를 얻을 수 있었다<sup>9)</sup>. 이는 전체 데이터베이스 크기의  $10^{-3}\%$ 로, 전체 단위 공간상에서 질의 영역이 차지하는 비율과 비교하면 상당히 큰 비율이다. 따라서, 본 실험 환경에서 사용한 데이터들은 상당히 클러스터링이 잘 되어 있음을 알 수 있다. 또한, 본 실험에서는 모든 검색 방법에 대하여 동일한 크기의 버퍼를 사용하였다. 위와 같은 이유로 본 실험 환경에서는 X-tree나 R\*-tree를 이용한 색인 방법이 순차 검색보다 좋은 성능을 보인다.

그림 6은 데이터베이스의 크기에 따른 응답 시간의 변화를 보여주고 있다. 먼저, 순차 검색인 경우, 데이터베이스의 크기에 비례하여 응답 시간이 선형적으로 증가

7) 평균 응답 시간은 UNIX에서 제공하는 times 시스템 콜을 이용하여 측정하였으며, 단일 사용자 모드에서 실험을 수행하였습니다

8)  $\epsilon$ 은 질의 영역의 반지름이고 차원이 20이므로,  $(2 \times 0.12)^{20}$ 은 질의 영역의 크기이다.

9) 실제로는 12개의 결과를 얻었음.

함을 알 수 있다. 특징 벡터의 차원이 12차원일 경우에는 데이터베이스의 크기에 상관없이 R\*-tree를 이용한 색인 방법과 X-tree를 이용한 색인 기법은 거의 비슷한 응답 시간을 보이나, 차원이 24차원일 경우에는 데이터베이스의 크기가 증가함에 따라, X-tree를 이용한 색인 기법의 응답 시간이 R\*-tree를 이용한 색인 방법보다 빠름을 알 수 있다.

그림 7은 데이터베이스의 크기가 10,000이고 특징 벡터의 차원이 20차원일 때, 유사성 허용 오차의 증가에 따른 응답 시간의 변화를 보여주고 있다. 순차 검색일 경우에는 유사성 허용 오차의 변화에 관계없이 일정하

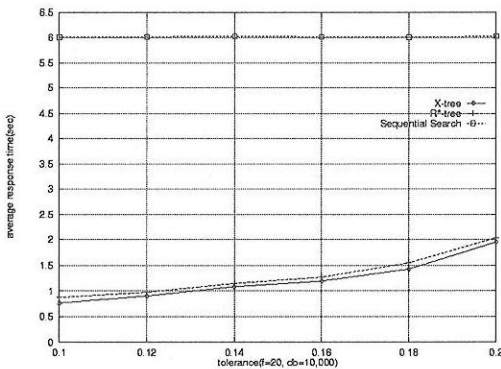


그림 7 유사성 허용 오차에 따른 검색 시간: 특징 벡터의 차원이 20이고, 데이터베이스의 크기가 10,000일 경우

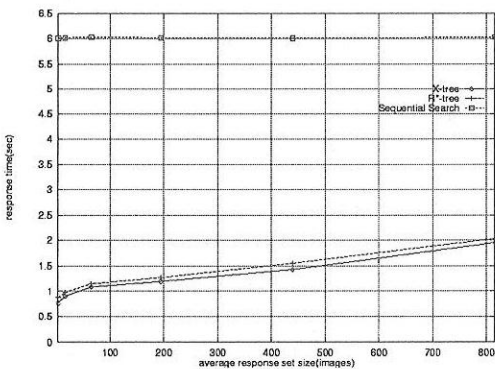


그림 8 검색 결과의 크기에 따른 검색 시간: 특징 벡터의 차원이 20이고, 데이터베이스의 크기가 10,000일 경우

게 약 6초대에서 검색됨을 알 수 있다. 이것은 순차 검색의 경우에는 유사성 허용 오차에 상관없이 데이터베이스에 저장되어 있는 모든 이미지들의 특징 벡터를 한 번씩 읽어야 하기 때문이다. R\*-tree를 이용한 색인 방법과 X-tree를 이용한 색인 기법의 경우에는 일반적으로 R\*-tree를 이용한 방법의 응답 시간이 떨어짐을 알 수 있다. 또한, 유사성 허용 오차가 증가함에 따라 두 색인 방법 모두 응답시간이 조금씩 증가하였다. 이것은 유사성 허용 오차가 증가함에 따라 질의 결과의 집합이 커지고, 그에 따라 후처리 과정에 소요되는 시간이 조금씩 증가하기 때문이다.

그림 8은 데이터베이스의 크기가 10,000이고 특징 벡터의 차원이 20차원일 경우, 질의의 결과로 얻어지는 이미지의 개수에 따른 응답 시간의 변화를 보여주고 있다. 순차 검색의 경우에는 그림 7에서와 같은 이유로 거의 동일한 응답 시간을 보임을 알 수 있다. R\*-tree를 이용한 색인 방법과 X-tree를 이용한 색인 기법의 경우에는 X-tree를 이용한 색인 기법이 R\*-tree를 이용한 방법보다 조금 빠른 응답 시간을 보임을 알 수 있다.

위의 다양한 실험을 통하여, 우리는 본 논문에서 제안한 X-tree를 이용한 내용 기반 이미지 색인 기법이 순차 검색보다 훨씬 뛰어나며, 또한 데이터베이스 크기가 증가하고 특징 벡터의 차원이 증가할수록 기존의 R\*-tree를 이용한 색인 방법보다 더 신속한 검색을 지원함을 알 수 있었다.

## 5.2 검색의 적합성(Retrieval Effectiveness)

이 절에서는 본 논문에서 제안된 여러 기법에 따라 구현된 내용 기반 이미지 검색 시스템의 검색 적합성<sup>10)</sup>에 대한 실험 결과를 설명한다.

완전 일치 검색을 기반으로 하는 정보 검색(IR: Information Retrieval) 분야에서는 재현율(recall)과 정확률(precision)을 이용하여 검색의 적합성을 측정하고 있다. 어떤 질의에 대하여,  $T$ 를 질의에 적합한 모든 객체의 개수라 하고,  $R$ 은 검색된 모든 적합한 객체의 개수,  $T$ 은 검색된 모든 객체의 개수라고 하면, 정확률은  $R_p/T_p$ 로 정의되며, 재현율은  $R_r/T_r$ 로 정의된다[5, 29].

그러나, 본 논문에서와 같이 완전 일치 검색이 아닌 유사 검색(similarity retrieval)을 수행하는 시스템에서는 이러한 정확률이나 재현율을 그대로 이용하는 것보다는 표준 정확률과 재현율(normalized precision and recall)을 이용하여 평가하는 것이 바람직하다[5,28,29].

10) 정보 검색(IR) 분야에서 사용하는 용어들에 대한 한글 표기는 [28]을 참고하십시오.



(a) 9 (b) 8 (c) 6 (d) 12 (e) 4  
 그림 9 질의 이미지 및 그와 유사한 적절한 이미지들의 개수

이것들은 적합한 이미지 객체들이 실제 검색된 결과의 순서 상에서 어떤 위치에 나타나는지를 반영하는 것이다. 예를 들어, 표준 재현율은 T개의 적절한 객체가 처음부터 T번째 위치까지 순서대로 나타나는 이상적인 검색 결과와 실제 시스템에서의 검색 결과의 차이를 측정하는 것이다[28,29].

QBIC[5]의 연구에서는 시스템의 검색 적합성을 측정하기 위하여 이러한 표준 정확률과 재현률을 변형한 방법을 제시하였다. 다음의 AVRR과 IAVRR은 QBIC에서 제안한 검색 적합성 측정 방법이며, 본 논문에서는 이 방법을 이용하여 검색 적합성을 측정하였다.

- AVRR(average rank of all relevant items) : 모든 적절한 이미지 객체들의 순위(rank)들에 대한 평균값 (첫 번째 위치, 즉 질의 이미지와 가장 유사한 이미지의 순위는 0임).
- IAVRR(ideal AVRR) : 이상적인 평균 AVRR 값(모든 적절한 이미지 객체들이 상위에서 차례로 검색된 경우 최대가 된다.) :  $IAVRR = (0+1+\dots+(T-1))/T$

여기서, IAVRR에 대한 AVRR의 비율이 검색의 적합성 정도를 보여준다[5]. 즉, AVRR/IAVRR의 값이 1에 근접하면 근접할수록 실제 검색 결과가 가장 이상적인 결과에 근접하게 됨을 의미하게 된다.

그림 9는 본 실험을 위해 사용한 5개의 예제 질의 이미지를 보여주고 있으며, 아래의 숫자는 데이터베이스에 저장된 이미지들 중에서 각 질의 이미지와 유사한 적절한 이미지들의 개수이다. 실제 실험에서 사용된 데이터베이스의 크기는 100이고, 5개의 질의 이미지에 대하여 적합한 이미지들의 총 개수는 39개이다. 각 이미지로부터 20차원의 특징 벡터를 추출하여 실험하였으며, 적절한 이미지들이 모두 검색될 수 있도록 충분한 유사성 허용 오차(0.1~0.3)를 고려하였다.

표 1은 그림 9에서 제시한 5개의 질의 이미지에 대한 IAVRR과 AVRR에 대한 실험값과 5번의 실험 결과에 대한 평균값을 보여주고 있다. 평균 AVRR 값이 6.96이라는 것은 질의 이미지와 유사한 적절한 이미지가 평균

표 1 그림 9의 질의 이미지에 대한 적합성 실험 결과

질의 이미지	(a)	(b)	(c)	(d)	(e)	평균
IAVRR	4	3.5	2.5	5.5	1.5	3.4
AVRR	8.2	7.6	5.4	10.1	3.5	6.96

7번째 순위(rank = 7)에 나타나게 된다는 것을 의미한다. 이것은 만약 어떤 질의 이미지에 대한 결과로 20여개의 이미지들을 보여지게 된다면, 대부분의 적절한 이미지들이 모두 보여지게 됨을 의미한다. 본 실험에서 IAVRR에 대한 AVRR의 비율(AVRR/IAVRR)은 2.05이며, 이것은 QBIC 시스템의 검색 적합성과 거의 유사한 수준이다[1].

### 6. 요약 및 향후연구 방향

본 논문에서는 대규모 이미지 데이터베이스에서 효율적으로 내용 기반 질의 및 검색을 수행할 수 있는 기법을 제안하였다. 본 논문의 중요한 기여를 정리하면 다음과 같다.

첫째, 내용 기반 검색을 위한 효율적인 특징 벡터의 추출 과정과 유사성 측정 방법을 제시하였으며, 이러한 기법들을 사용하여 내용 기반 검색을 수행할 경우, 항상 false dismissals이 발생하지 않음을 보였다.

둘째, 여러 고차원 색인 구조들 중에서 본 연구에서와 같은 내용 기반 이미지 검색 환경에 적합한 색인 구조를 선택하는 기준을 제시하였고, 이러한 기준에 기반하여 X-tree를 이용한 효율적인 이미지 색인 방법을 사용하였으며, 이것이 기존의 순차 검색이나 R\*-tree를 이용한 검색 방법보다 효율적임을 다양한 실험을 통하여 보였다.

셋째, QBIC 시스템에서 제안한 검색 적합성을 측정하는 방법을 사용하여 본 논문에서 제안하는 시스템이 어느정도의 만족스러운 검색 결과를 보이는지를 살펴보았다.

본 논문에서 제시한 내용 기반 검색 기법은 앞으로 고차원 특징 데이터를 다루는 이와 유사한 내용 기반 멀티미디어 정보 검색 시스템에서 중요하고 핵심적인 기술로 사용될 수 있을 것이다.

표 2는 관련 연구에서 기술하였던 내용 기반 검색 시스템들에서 제공하는 질의의 종류 및 사용하는 색인 방

11) QBIC 시스템에서 색상에 의한 질의의 경우에는 AVRR/IAVRR = 1.93이고, 모양에 의한 질의의 경우에는 2.05이다 [5].

법을 비교한 것이다. 본 논문에서 제안된 시스템은 아직 다양한 종류의 질의를 제공하지 못하고 있다. 따라서, 향후 연구 과제로는 이미지의 다양한 속성을 표현하는 특징 데이터를 효율적으로 추출하는 방법과, 이들을 서로 통합하여 복합적인 질의를 수행할 수 있는 효율적인 방법에 대해서 연구할 예정이며, 이를 확장하여 대규모 비디오 데이터베이스에서 신속하고 효율적인 검색을 지원하는 내용 기반 비디오 검색 연구를 진행할 예정이다.

표 2 내용 기반 검색 시스템들의 비교

시스템	질의의 종류				색인 방법
	모양	색상	질감	공간정보	
QBIC	○	○	○	○	R*-tree
VisualSeek	×	○	×	○	R-tree, Quad-Tree
QVE	○	×	×	×	없음
Jacobs[6]	△	△	×	△	없음
본 시스템	○	×	×	×	X-tree

Legend :

- 가능
- × 불가능
- △ 구분 모호

감사의 글

본 연구를 위하여 X-tree 소스 코드를 제공해 주신 AT&T 연구소에 계신 Stefan Berchtold 박사에게 감사사를 드립니다.

참 고 문 헌

[1] V. N. Gudivada and V. V. Raghavan, "Content-Based Image Retrieval Systems," *IEEE Computer*, 28(9), 1995.

[2] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast Multiresolution Image Query," *Proc. ACM SIGGRAPH, New York*, 1995.

[3] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC Project: Querying Image by Content using Color, Texture, and Shape," *Proc. SPIE Storage and Retrieval for Image and Video Databases*, pages 173-187, February 1993.

[4] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Database," *International Journal of Computer Vision*, fall 1995.

[5] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equiz, "Efficient and Effective Querying by Image Content," *Journal of Intelligent Information System(JIIS)*, 3(3):231-262, July 1994.

[6] C. Faloutsos, "Fast Searching by Content in Multimedia Databases," *Data Engineering*, 18(4), 1995.

[7] A. Swami, R. Agrawal, and C. Faloutsos, "Efficient Similarity Search in Sequence Databases," *FODO: International Conference on Foundations of Data Organization and Algorithms, Chicago, Illinois, USA*, 1993.

[8] K. Hirata and T. Kato, "Query by Visual Example-Content Based Image Retrieval," *Advances in Database Technology(EDBT '92)*, pages 56-71, 1992.

[9] J. R. Smith and S-F. Chang, "VisualSBEK: A Fully Automated Content-based Image Query System," *ACM Multimedia 96, Boston, MA*, 1996.

[10] D. Shasha and T-L Wang, "New techniques for best-match retrieval," *ACM TOIS*, 8(2):140-158, April 1990.

[11] F. Aurenhammer, "Voronoi Diagrams - A Survey of a Fundamental Geometric Data Structure," *ACM Computing Surveys*, 23(3):345-405, September 1991.

[12] J. T. Robinson. "The K-D-B-tree: A Search Structure for Large Multidimensional Dynamic Indexes," *Proc. ACM SIGMOD, Ann Arbor, USA*, pages 10-18, April 1981.

[13] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R\*-tree: An Efficient and Robust Access Method for Points and Rectangles," *Proc. ACM SIGMOD*, pages 322-331, May 1990.

[14] A. Guttman, "R-trees: a dynamic index structure for spatial searching," *Proc. ACM SIGMOD*, pages 47-57, June 1984.

[15] T. Sellis, N. Roussopoulos, and C. Faloutsos, "The R+-trees: a Dynamic Index for Multi-Dimensional Objects," *Proc. of 13th VLDB Conference, Brighton, England*, pages 507-518, September 1987

[16] H. Samet, *The Design and Analysis of Spatial Data Structures, Addison-Wesley*, 1989.

[17] J. Nievergelt, H. Hinterberger, and K. C. Sevcik, "The Grid File: An Adaptable, Symmetric Multitree File Structure," *ACM TODS*, 9(1):38-71, March 1984.

[18] K.-I. Lin, H. V. Jagadish, and C. Faloutsos, "The TV-tree: An Index Structure for High-Dimensional Data," *VLDB Journal*, 3(4):517-542, 1994.

[19] S. Berchtold, D. A. Keim, and H.-P. Kriegel, "The X-tree: An Indexing Structure for High-Dimensional Data," *Proc. the 22nd VLDB Conference, Bombay, India*, pages 28-39, September 1996.

[20] D. A. White and R. Jain, "Similarity Indexing: Algorithms and Performance," *Proc. SPIE Vol.2670, San Diego, USA*, pages 62-73, 1996.

[21] D. A. White and R. Jain, "Similarity Indexing with the SS-tree," *Proc. of Data Engineering*, pages 516-523, 1996.

[22] N. Katayama and S. Satoh, "The SR-tree: An Index

Structure for High-Dimensional Nearest Neighbor Queries," *Proc. ACM SIGMOD, Tucson, Arizona*, May 1997.

- [23] Y.T. Chan, WAVELET BASICS, *Kluwer Academic Publishers*, 1995.
- [24] A. Graps, "An Introduction to Wavelets," *IEEE Computational Science and Engineering*, 2(2), 1995.
- [25] D. H. Lee, Y. J. Song, and H.-J Kim, "SCARLET: Design and Implementation of Content-based Image Retrieval System using Wavelet Transform," *Journal of KISS*, 3(4), 1997.
- [26] Private Communication with Stefan Berchtold who is the author of [25]. Apr 22 - Apr 27 1998.
- [27] S. Berchtold, C. Bohm, and H.-P. Kriegel, "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality," *Proc. ACM SIGMOD, Seattle, Washington*, 1998.
- [28] 정영미, 정보검색론, 정음사, 1988
- [29] G. Salton and M. J. McGill, Introduction to Modern Information Retrieval, *McGraw-Hill*, 1983.



김형주

1982년 2월 서울대학교 컴퓨터공학과 졸업. 1985년 8월 Univ. of Texas at Austin, 전자계산학 석사. 1988년 5월 Univ. of Texas at Austin, 전자계산학 박사. 1988년 5월 ~ 9월 Univ. of Texas at Austin. Post-Doc. 1988년 9월 ~ 1990년 12월 Georgia Institute of Technology, 부교수. 1991년 1월 ~ 현재 서울대학교 컴퓨터공학과 부교수. 관심분야는 객체지향 시스템, 사용자 인터페이스, 데이터베이스



이동호

1995년 2월 홍익대학교 컴퓨터공학과(학사). 1997년 2월 서울대학교 컴퓨터공학과(석사). 1997년 3월 ~ 현재 서울대학교 컴퓨터공학과 박사과정 재학중. 관심분야는 객체지향 시스템, 멀티미디어 정보 검색, 데이터베이스



박주홍

1993년 2월 서울대학교 계산통계학과(학사). 1995년 2월 서울대학교 컴퓨터공학과(석사). 1995년 3월 ~ 현재 서울대학교 컴퓨터공학과 박사과정 재학중. 관심분야는 객체지향 시스템, 데이터베이스, 질의어 처리



정진완

1973년 서울대학교 전기공학과(학사). 1983년 University of Michigan 컴퓨터공학과(박사). 1983년 ~ 1987년 미국 GM 연구소 선임 연구원. 1987년 ~ 1993년 미국 GM 연구소 책임 연구원. 1993년 ~ 1996년 한국과학기술원 정보 및 통신공학과 부교수. 1996년 ~ 현재 한국과학기술원 전산학과 부교수. 관심분야는 GIS, 객체지향 데이터베이스, 멀티미디어 데이터베이스, 분산 데이터베이스, CIM.