

# 웹 상의 시각 정보 검색

한 태 화\*    민 경 필\*\*    전 준 철\*\*\*    최 용 길\*\*\*\*

## ◆ 목 차 ◆

- |               |                  |
|---------------|------------------|
| 1. 서 론        | 4. 3D 모델 검색 엔진   |
| 2. 시각 정보의 표현법 | 5. 결론 및 향후 연구 방향 |
| 3. 정보 검색 방법   |                  |

## 1. 서 론

지난 수년간 진행되어온 데이터 검색과 분석에 관한 연구 중에서 특히 멀티미디어 데이터에 관한 검색과 분석은 많은 진보를 해왔다. 어떤 질의에 부합하는 이미지를 검출하는 이미지검색은 웹상에서 미디어 정보의 빠른 증가로 인해 더욱 중요한 문제가 되고있다. 일반적으로 사용자들은 질의 이미지들을 주고 비슷한 이미지들의 집합을 제공할 수 있는 시스템을 사용하기를 원한다. 그렇기 때문에 특정 이미지에 대한 묘사를 어떻게 할지, 어떤 방법으로 다른 이미지와 비교하여 비슷한 이미지인지를 판단할지, 어떤 방법으로 이미지 데이터베이스의 색인을 구성할지, 어떻게 효과적으로 검색을 수행할지 등의 문제는 그러한 이미지 검색 시스템에서 가장 중요한 문제들이라 할 수 있다.

이미지는 형태(shape), 색상(color), 고유한 결(texture), 공간상의 관계(spatial relationship) 등 이미지의 콘텐츠라 하는 몇 가지 저수준의 특성으로 구별할 수 있다. 이런 특성을 이용하면 이미지를 묘사하고 모델링하는 것 뿐만 아니라 다른 이미지들 사이에서 비교할 수 있는 차이점을 만들 수도 있다. 현재까지의 이미지검색 시스템은 이런 특성을 기반으로 하여 구축되었으며 내용기반 이미지검색 (Content-based Image Retrieval ; CBIR)

시스템이라는 분야로 이미 인식되고있다.

이미지의 이런 특성은 이미지 자체가 가지고있는 고유한 인지적인 특징을 표시해주기 때문에 특정 이미지를 묘사하고 표현하는데 중요하게 사용되고있다. 현재 내용기반 검색시스템의 대부분이 이 특성을 이용하고 있다. 대표적인 IBM의 QBIC(Query By Image Content)시스템은 색상과 형태, 그리고 텍스처 유사도 측정을 이용하고있다[1,2].

## 2. 시각 정보의 표현법

80년대의 이미지를 이용한 정보검색의 시도 이래 90년대의 이미지검색기술은 방대한 이미지자체의 의미적 내용을 포함한 유사도를 이용한 검색방법을 연구하기 시작하였다. Katof[3]가 제안한 프로토타입은 이 문제를 더 일반화 시켰고, 유사도에 의한 검색시스템은 다음 세가지 요소에 근거하고있다.

- 이미지들로부터 특성을 추출하여 미리 계산된 데이터에 대해 효율적인 표현과 저장기준을 제공할 것
- 유사도 측정이 인지적인 의미를 충분히 가지고있으며, 예시이미지와 전체 데이터베이스를 비교할 때 효과적인 계산 가능여부
- 어떤 유사도를 선택하여 검색할 것인지, 검색된 이미지에 대한 순서적, 시각적으로 효과적인 표현과 타당한 피드백의 제공여부에 대한 사용자 인터페이스

\* 경기대학교 전자계산학과 박사과정  
\*\* 경기대학교 전자계산학과 박사과정  
\*\*\* 경기대학교 정보과학부 교수  
\*\*\*\* 경기대학교 경영학부 교수

## 2.1 유사도(similarity) 측면에서 본 이미지의 측정

여기서는 여러 가지 타입의 유사도 정의와 관련된 특성추출에 대해 간단히 설명한다. Gudivada는 [4]에서 검색을 위한 가능한 유사도의 종류를 형태(shape), 색상(color), 텍스처(texture), 공간상의 관계(spatial relationship) 등으로 구분하였다.

### • color similarity

이미지가 가진 여러 가지 특성 중에서 색상은 가장 직관적이면서 사용자가 이미지 간의 차이를 쉽게 분간할 수 있게 한다는 점에서 색상유사도 특성은 내용기반 이미지검색시스템을 구현하는데 매우 중요한 부분이다. 일반적으로 색상 히스토그램은 이미지의 회전, 물체 크기의 변화 및 이동등에 비교적 강한 특징을 가지고 있다. 이미지내의 전체적인 색상 분포의 속성을 알아 내기위해 HLS(Hue-Luminosity-Saturation)색상계를 사용하며 3개 모멘트를 사용하지만 빛에 독립적인 색상 분포 속성을 이용해야 할 경우는 색도(hue)와 채도(saturation) 공간의 분포만을 사용한다[5]. 색상유사도를 이용하여 검색할 때 한가지 어려운 점은 예를 들어, 사용자가 종종 붉은색의 물체를 가지고 똑같은 것을 찾으려 한다는 점이다. 이 문제는 색상유사도를 공간적요소로 제한하는 것으로 좀더 일반적으로 공간유사도와 색상유사도를 조합하여 텍스처유사도를 나타낼 수 있다. 이런 점은 프로토타입과 상업적인 시스템이 여러 가지 유사도 함수를 조합하기 위하여 사용자 인터페이스안에 복잡한 ad-hoc메커니즘을 갖는 이유이기도 하다.

### • texture similarity

텍스처는 이미지의 특정영역에 대한 시각적 패턴으로, 대비(contrast), 거칠기(coarseness), 방향성(directionality), 규칙성(regularity), 주기성(periodicity), randomness등과 같은 특징을 가진다. 이미지의 텍스처는 신호를 시간영역에서 주파수영역으로 바꾸는 푸리에 변환과 동일한 방식의 웨이블릿 변환으로 표현한다. 이 방법은 건물이나 자동차 같은 물체를 가진 큰 사이즈의 이미지를 검색하는데 효과를 기대할 수 있다. 특히 Tamura의 텍스처 특

징 중 표면에 대한 측정치인 거칠기, 이미지의 gray level 분포인 대비, 이미지가 어떤 방향성분을 주고 가지는지의 방향성 성분을 이용하는 방법을 제시하였고 이 이론의 개선된 방법이 실제로 IBM사의 QBIC시스템에 응용되어 사용되고 있다[2].

### • shape similarity

물체의 형태는 그 경계면을 추적하여 얻어지는 모양으로 형태를 분석하는 두 가지 방법이 있다. 형태영역의 전체적인 형태특성에 근거한 것과 물체의 경계면에만 근거한 것들이 있다.

이미지들 안에 나타나는 기하학적 모양들 사이의 형태 유사도와 이미지로 묘사한 물체들 사이의 유사도로 구별한다. 전자의 경우 유사도는 정량적인 형태를 가진 정확한 이미지 데이터 베이스와 같이 물체의 특정한 종류에 대해서 적용하는 경우가 많으며 이를 이용하여 크기에 영향을 받지않는 곡률 등으로 형태 표현을 정의하고자 하였다. QBIC시스템은 원형성과 기형성, 주축의 방향성과 대수모멘트로 형태표현을 정의한다. 하지만 검색 시스템 사용자들은 실제로 방향각에 종속된 형태 유사도를 정의하기를 원하는데 사다리꼴의 검색 시 수평기저와 수직기저의 차이가 그 예라 할 수 있겠다.

Sclaroff와 Pentland[6]는 이차원 형태의 물리적 모델을 만들어서 변형 가능한 오브젝트의 프로토타입을 KLT(Karhunen-Loeve transform)를 이용하여 분석하였다. 이 방법은 변형 가능한 모드들의 리스트를 순서적으로 만들어 주는데 전체적인 변형과 연관된 저주파 비강체 모드와 부분적인 변형에 관계된 고주파 모드 모두를 가능하게 한다. 색상과 텍스처와 마찬가지로 형태 유사도를 이용한 방법 역시 이미지가 몇 개의 물체나 배경을 포함할 때 문제에 부딪히는데 이 문제를 해결하기위해서 형태들 사이의 사전 세그멘테이션과 공간관계를 만들어줄 필요가 있다.

### • spatial similarity

Gudivada와 Raghavan[7]는 이미지가 물체들로 분할되었고, 연관된 각 물체가 중심을 이루고 상징적 이름이 있는 상황에서는 공간적 유사도를 고려하였다. 이런 상황의 이미지표현을 상징적 이미지라 하고, 그

런 이미지에 대해서는 회전이나 크기변환 이동과 같은 이미지의 유사도 함수를 정의하기가 상대적으로 쉬울 수 있다.

이미지의 검색에 대한 많은 연구들이 대부분 유사도 정보를 이용하고있으며 이와 같은 특성들에 기반으로 많은 기술들이 개발되고있다. 그러나 절대적인 최적의 기술이 있을 수는 없으며 장단점을 포함한 가운데 서로 다른 상황에서 서로 다른 기술을 적용하여야 할 것이다. 그러므로 서로 다른 특성에 기반 한 기술이 효과적이고 정확한 검색에 유용할 것이며 실제로 대부분의 현재 내용기반 검색시스템은 이러한 유사도 측정제 하나 이상의 특성을 이용하고있다.

### 3. 정보 검색 방법

대량의 디지털 유통데이터에서 필요한 데이터만을 추출하는 데이터베이스의 응용은 웹의 범용화로 더욱 큰 영역을 확보하게 되었다. 숫자나 문자가 이루는 대부분의 정형데이터 형식에서 텍스트, 그래픽, 이미지, 오디오, 비디오 등과 같은 비정형 데이터를 포함한 멀티미디어데이터로의 확산이 그 대표적인 것으로 볼 수 있다. 웹이 수용할 수 있는 데이터 종류가 확대됨에 따라 웹을 통한 정보 검색 역시 단순한 정보검색차원을 넘어서 효율적인 저장방식과 효과적인 검색결과 제공을 통한 다양한 효과를 기대하게 되었다. 그러나 멀티미디어 데이터들은 종류와 특성이 데이터의 크기나 다양한 저장/출력방식, 시/공간 관계성, 패턴인식을 이용한 내용검색의 난해함이나 복잡한 구조 때문에 신속하고 효율적인 검색에 많은 어려움이 따른다. 이 장에서는 미디어별 검색 방법을 간단히 소개하고, 수많은 웹 상의 키워드 기반의 검색엔진이 보통 2D 이미지에 대한 검색을 제공하고 있다는 점에서 특히 이미지와 같은 시각정보에 대한 검색에 필요한 기술에 대해서술할 것이다. 그리고 현재 상용화 되고 있는 서비스와 내용기반 검색시스템에 대해 설명할 것이다.

#### 3.1 미디어별 검색방법

검색방법은 사용자가 검색조건과 대상 데이터 집합을 명시하여 그 조건에 따라 한꺼번에 검색이 가능한

질의 방법과 사용자가 한 단위의 데이터를 보거나 이와 관련된 다른 데이터를 찾기 위해 항해하면서 원하는 데이터를 찾는 브라우징 방법으로 나눈다.

##### 3.1.1 이미지 데이터 검색

이미지검색은 초기에 이미지 내용을 속성집합으로 모델화하는 방향과 이미지변역 연구가들에 의한 통합 시스템개발 방향이 모색되었지만 내용기반 검색에서는 특성추출을 어느 정도 자동화할지 등의 문제에 관련하여 시스템 개발에 역점을 두고있다. 전통적인 종래의 검색방법은 사용자가 제시한 적절한 키워드를 같이 입력하였으나 사용자가 시각적인 예를 보여주기 위한 색상, 텍스처, 스케치, 모양, 공간제약, 브라우징 방식 등의 질의 클래스를 분류하여 이를 바탕으로 내용을 기반으로 하는 이미지 검색 시스템개발이 모색되고있다. 사진과 같은 이미지는 대부분 탐색 시 키워드 매칭을 이용하여 요청한 이미지에 대한 분류번호로 원하는 사진을 찾는다. 그러나 철자의 오류 등으로 매칭이 이루어 지지 않을 수 있기 때문에 정확성을 기하기 위하여 이미지의 특성을 특징 데이터로 가지고 있다가 질의 시에 개념 질의를 이용하여 빠른 시간 내에 찾을 수 있도록 한다. 특히 기존의 사람 얼굴 사진과 같은 검색에서는 얼굴 인식성을 향상시키기 위하여 얼굴 모델이 요구되는데, 얼굴을 모델하기에는 너무 많은 형태가 존재하기 때문에 실제로 검색하는데 어려움이 있다. 최근에는 캡션과 같은 부제목이 달려있는 경우 실제 사진 속의 내용을 검색하기위해 부제목 정보를 이용한다. 비디오 이미지의 경우에는 먼저 데이터 베이스에 이미지 및 비디오를 저장하기 위하여 이미지에서 객체를 구분하고 비디오를 샷 단위로 나누어 색, 짜임새, 형태, 위치 등의 특성을 계산한 후 질의 단계에서 질의에 사용된 데이터 이미지와 샷을 검색하여 같은 이미지를 찾기 위하여 계산 후 저장된 이미지를 검색하도록 한다.

특히 특징기반 3D 물체의 정보검색 경우는 시스템이 3D 물체의 데이터를 입력 받아 기하모델을 구성한 후 모델로부터 특징들을 추출하여 데이터 베이스에 저장하는 인식 서브시스템과 사용자 질의를 이용하여 데이터베이스에 저장된 3D 물체의 정보를 검색해내는 탐색 서브시스템으로 구성한다. 이때 중요한 것은 사

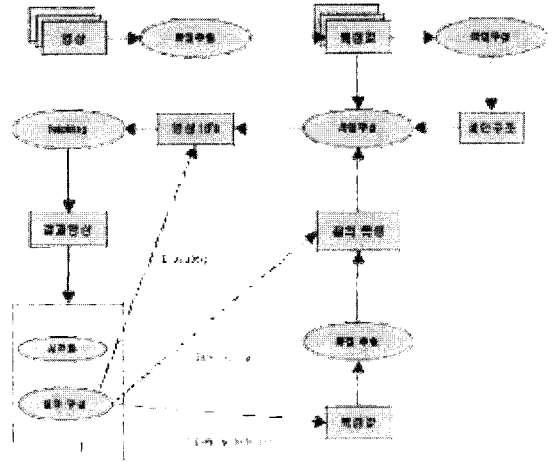
사용자의 질의와 데이터를 매칭 시키는 방법이다. 이 방법은 묘사기반 매칭으로서 키워드를 사용하는 방법과 텍스트 묘사를 이용하는 방법을 사용한다. 또한 물체의 기하정보나 위상정보를 이용하여 정확한 매칭을 지원하는 방법과 물체의 데이터로부터 비문자 형태의 특징을 추출하여 매칭에 이용하는 방법이 있다.

### 3.1.2 음성 및 문서 데이터 검색 기법

음성정보 처리기술을 이용한 검색으로, 통계적 패턴 매칭의 음성인식기술은 스펙트럼 분석 단계에서 입력된 음성특징을 가진 스펙트럼특징을 추출하며, 단어 단위 매칭에서 사전에 따라 부단어 모델을 연결한 단어모델을 작성한 후 이를 입력 특징벡터와 매칭하여 가장 유사한 단어를 찾는다. 그리고 문장단위 매칭에서 언어 모델을 이용하여 가장 높은 확률로 발생할 단어열을 구한다. 구문론적인 규칙과 의미론적인 규칙은 사람이 직접 작성하거나 통계적 언어 모델에 따라 구성되고, 탐색과 인식 단어열의 결정은 모든 단어열을 고려한 후 가장 가능성이 높은 단어열로 선택 된다 문서 데이터 검색 기법에서는 데이터베이스를 구성할 때 문서를 하나씩 받아들여 텍스트를 단어로 분해하고 색인이 불필요한 불용어 목록과 비교하여 그 안에 없는 경우 어간이 구분된다. 문서와 데이터베이스 전체에 걸쳐 검색에 사용되는 단어의 빈도에 따라 단어들의 수가 계산되어 문서, 문서 내 항목, 총 단어 수와 같은 관련정보와 단어들 이 데이터베이스로 입력되어 데이터베이스는 키워드와 문서식별자의 쌍으로 구성되는 역파일 구조가 된다. 하나의 정보검색시스템에서 각 문서들은 유일한 식별자를 가지며, 항목연산이 가능하다면 각 항목은 유일한 항목명을 가져야 한다[8].

### 3.2 내용기반 이미지 검색시스템

대부분의 이미지검색 시스템의 프레임워크를 개념적으로 표현해보면 그림 1과 같이 일반화하여 표현할 수 있다[9]. 이런 시스템들은 사용자가 질의를 어떻게 구성하는지, 적절한 검색결과에 대한 피드백을 줄 수 있는지 준다면 어떻게 주는지, 어떤 특성을 검색에 이용하는지, 질의 이미지와 데이터베이스이미지에서 특성이 어떻게 매칭 되는지, 사용자에게 검색 결과를 어떻게



(그림 1) 이미지검색 시스템의 프레임워크

보여주는지 등을 검색시스템에서 해결해야 할 과제로 본다. 사용자 인터페이스는 주로 질의를 구성하는 부분과 결과를 나타내는 부분으로 구성되며 이미 언급한 것처럼 하나씩 데이터베이스에서 브라우징 되거나, 키워드나 이미지 특징 등으로 검색되는 등 여러 방법을 통해 검색된다. 또 다른 방법으로는 이미지를 주거나 특징을 스케치하여 검색하는 방법도 있다.

키워드 기반의 웹 상의 검색엔진 중 HotBot(<http://hotbot.lycos.com/>)과 NBCi(<http://www.nci.com/>)에서 이미지를 그 매체로 검색한다. 좀 더 일반적으로는 야후의 이미지 서퍼(<http://isurf.yahoo.com/>)나 라이코스의 멀티미디어 검색기(<http://multimedia.lycos.com/>)가 있지만 이 시스템은 키워드 기반의 검색방법을 이용한다. 웹상에서 알파벳순으로 특정 이미지를 모으는 시스템으로는 ImageFinder(<http://sunsite.berkeley.edu/ImageFinder/>)가 웹 상의 위치한 사진들을 모아서 모집된 이미지의 리스트를 제공한다. 알타비스타의 Photofinder는 내용기반 이미지 검색엔진으로 특정하게 수집된 이미지 리스트와 웹을 모두 사용하여 검색하는 엔진이다. 다음은 현재 상용화 및 생산 되어지고 있거나, 연구되는 시스템들이다.

#### • AltaVista Photofinder

(<http://image.altavista.com/cgi-bin/avncgi>)

DEC에서 개발한 것으로 현재 알타비스타에 의해 운영되며, 주된 색상과, 모양, 텍스트와 같은 특징기반 유사도를 이용하며, 사용자가 먼저 키워드를 입력하는

질의방식을 적용하고, 인덱싱 방법은 VIR이라는 기존의 이미지 엔진상에 시스템이 존재하며, 결과출력은 순서 없이 간단한 그림 형태로 제공된다.

- Berkely Digital Library Project  
(<http://elib.es.berkely.edu/photos/all.shtml>)

버클리 대학에서 개발하고, 수집된 데이터, 키워드, 위치, 작가 등의 질의를 사용할 수 있다. 질의방식은 내용기반 검색을 위하여, 사진에 사용된 13개의 색상과 색상정도를 세단계로 구분할 수 있다. 이미지의 특징 값들을 텍스트 스트링으로 저장하여 매칭 시키며, 모든 특징 값들을 관계형 데이터베이스에 저장하여 색인한다. 결과물은 색인번호나 작가, 수집키 등 일정 순서 없이 나타낸다.

- C-bird(Content-Based Image Retrieval from Digital libraries)  
(<http://jupiter.cs.sfu.ca/cbird/>)

네 개의 벡터(color, MFC, MFO, chromaticity vector)로 구성된 특징 기술자와 색상 레이아웃과 에지 레이아웃을 이용한 레이아웃 기술자를 각 이미지에 대하여 계산하여야 하며, 임의위치에서 사용자가 질의하게 되면 각 질의에 대한 연속적인 격자모양 이미지를 데이터베이스에서부터 나타낸다. 채도 벡터사이의 거리를 가지고 매칭하며, 사용자가 행과 열로 나타난 디스플레이 이 이미지에서 선택할 수 있도록 결과를 나타내준다.

- Chabot(<http://elib.cs.berkely.edu/photos/>)

버클리 공대에서 개발한 POSTGRES를 기반으로 한 특징 기반 검색이 추가 되는 시스템이다. 초창기에 개발된 것으로 California수자원 관리 조직에서 소장하고 있는 이미지를 저장하고 검색하는 것을 위주로 시스템을 개발하였으며, 사용자는 검색분류 리스트를 제공받아서 질의 할 수 있으며 이미지의 매칭을 위해서는 색상 기준을 이용한다.

- PicToSeek([http://zomax.wins.uva.nl:5345/ret\\_user](http://zomax.wins.uva.nl:5345/ret_user))

네덜란드 암스테르담 대학에서 개발한 시스템으로 고정적인 색상과 모양이, 카메라 뷰포인트, 오브젝트의 기하적 특징과 조명에 독립적인 질의가 내용기반 질의 안에 특징값으로 이용된다. 질의 이미지는 웹크롤러나

사용자 URL을 통해서 모아진 데이터베이스에서 선택될 수 있고 질의 히스토그램과 데이터베이스 이미지 히스토그램 사이의 유사도 특징은 정규화 교차로 매칭한다.

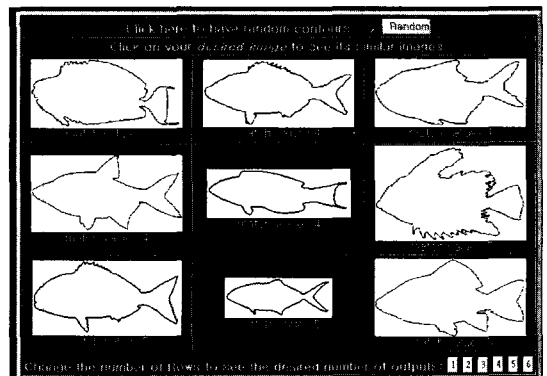
- QBIC(<http://www.qbic.almaden.ibm.com/cgi-bin/stamps-demo>)

IBM에서 개발한 정지 이미지 및 동이미지 검색 엔진으로써 특징 기반 검색 기능이 추가 되고, 주석 기반 검색기능이 추가된 시스템이다. 점차 기능이 향상되고 있으며 장면 전환검색 기능과, 스토리 보딩에 의한 비디오 요약 출력 기능도 첨가 되었다. 질의 형태는 예제 이미지와, 사용자가 구성한 스케치물이나 선택한 색상, 텍스처 패턴 등을 기반으로 하며, 질의 오브젝트와 데이터베이스 오브젝트간의 거리는 유클리드 거리를 사용한다. 그리고 검색결과로 주어진 것을 새로운 검색의 시작으로 사용할 수 있는 피드백을 제공한다.

- SQUID(Shape Queries Using Image Database)

(<http://www.ec.surrey.ac.uk/Research/VSSP/imagedb/squid.html>)

영국의 Surrey 대학에서 개발한 시스템으로, 이미지에서 경계선 윤곽을 추출하여 전체 형태를 표현하여 이 형태정보를 가지고 검색하는 방법이다. 사용자가 디스플레이 된 오브젝트의 경계선 중에 선택하여 질의를 주고, 질의한 전체 파라미터와 가장 근접한 파라미터를 가지는 모델이 있는 이미지를 선택하여 CSS 매칭 알고리즘을 적용한다. 결과는 내림차순 유사도 순서로 디스플레이 된다. 실제로 이 시스템은 1,100개의 단일 배경 해상 생물 이미지를 데이터베이스로 구성하여 응용하였다.



(그림 2) SQUID 시스템

- WebSEEk(<http://www.ctr.columbia.edu/WebSEEk/>)

이미지의 카탈로그와 웹 상에서 모든 비디오를 통해서 텍스트와 색상 기반의 질의를 한다. 사용자는 먼저 카탈로그를 선택하여 주제를 넣어 질의다. 질의 색상 히스토그램과 목표 히스토그램간의 거리로 매칭하며 모든 데이터베이스상의 이미지를 계산하여 개별적으로 색인한다. 검색 결과는 내림차순 유사도 순서로 페이지별로 디스플레이 된다.

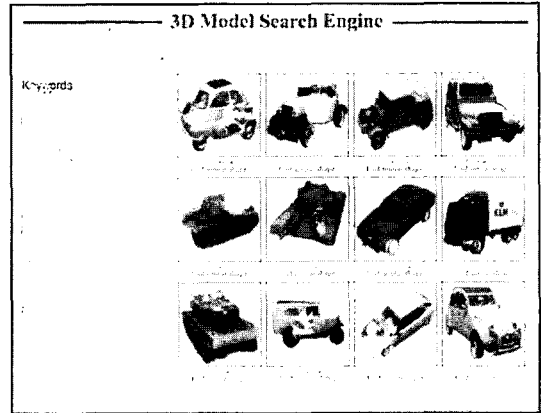
- WISE(Wavelet Image Search Engine)

(<http://bergman.stanford.edu/~zwang/project/imsearch/WBIS.html>)

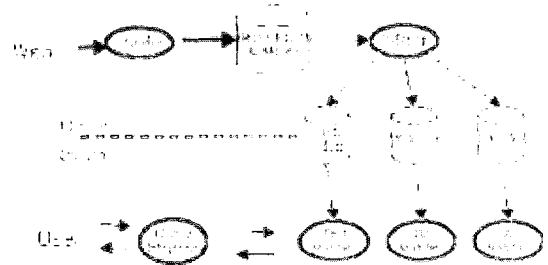
WBIS(Wavelet-Based Image Indexing and Searching)으로 불리기도 하는 이 시스템은 스탠포드 대학에서 개발한 것으로 Daubechies 웨이블릿 변환을 이용하여 색상 레이아웃 정보를 인코딩하여 질의를 수행한다. 데이터베이스 이미지나 직접 손으로 그린 스케치를 질의 이미지로 사용하여 예제 이미지에 의한 질의를 수행하며, 매칭 단계는 질의이미지에서 표준분포를 계산하여 매칭하는 단계와 이단계를 거쳐 웨이블릿 변환 계수들 사이의 가중치 유클리디언 거리를 이용하여 질의와 비교하여 매칭하는 두 단계로 구성된다. 이 시스템은 스탠포드 대학의 도서관과 연구 프로젝트에서 이용되고있다.

#### 4. 3D 모델 검색 엔진

웹상에 이용할 수 있는 3D 모델이 급증하면서 이를 찾아서 이용하고자 하는 검색엔진을 필요로 하지만 전통적인 문자 기반 검색 기술은 3D 데이터를 다루는 데는 효과적이지 않다. 3D 모델 검색을 위해서는 주로 형태 기반 검색 방법(shape-based search method)을 이용하며, 쉽게 이용할 수 있는 질의 방법과 임의의 다각형모델에 적용할 수 있는 매칭 알고리즘의 연구에 노력하여야 한다. 본 장에서는 3D과 2D스케치, 3D 모델 텍스트 검색어를 기반으로 질의를 지원하는 웹 기반 검색 엔진 시스템에 대하여 설명할 것이다. 형태기반 질의를 위해 모델의 수정이나 방향 정렬 없이 유사도 특징을 계산하기위해 구면조화 함수(spherical harmonics)를 이용하는 새로운 매칭 알고리즘을 이용



(그림 3) 3차원 모델 검색 엔진



(그림 4) 시스템 개략도

한다. 이 알고리즘은 다른 형태 매칭 알고리즘보다 향상된 수행능력과 20,000개의 저장소에서 일초 내에 결과를 보여줄 만큼이나 빠른 속도를 제공한다[10].

- 시스템 개략도(system overview)

프린스턴 대학이 개발한 이 시스템은 그림 4와 같이 crawling, indexing, querying, matching의 4단계로 구성되어 수행된다. 수집과 인덱싱의 과정은 오프라인으로 수행되고, 질의와 매칭은 각 사용자의 질의에 대해 수행된다. 각 단계의 주된 수행기능은 다음과 같다.

##### Crawling 단계 :

웹을 뒤져서 3D모델의 데이터 베이스를 형성한다. 웹상의 3D 데이터는 여전히 그 비율이 낮고 전체 3D 데이터 중에 상세한 데이터를 가진 모델 역시 얼마 되지 않으므로 3D모델의 측정을 페이지랭크 하는 방법의 크롤러를 개발하여 17,834개의 VRML모델을 웹에서 수집하였다. 또한 기존의 3D모델 업체들에 의해

더 추가된 상업적 모델을 포함할 수도 있다.

**Indexing 단계 :**

인덱싱 단계에서는 텍스트와 형태질의 기준으로 3D모델을 검색하기 위해 색인을 계산한다. 특히 3D모델의 변형에 강하고, 방향에 고정적인 효과를 가진 형태 묘사자를 이용하는데, 계산성을 고려하기 위하여 묘사정도, 간결성, 효과성을 고려한 구면조화함수를 기반으로 만들었다.

**Querying 단계 :**

이 시스템은 사용자가 직접 3D모델을 검색하게 하고, 텍스트 키워드와 2D 케치, 3D스케치, 모델 매칭 그리고 반복적인 상세 찾기 등을 기반으로 하는 질의 방식을 지원한다. 텍스트나 형태만을 가지고 질의한 결과보다 좋은 결과를 볼 수 있는, 두 질의형태를 합성한 방법을 이용한다.

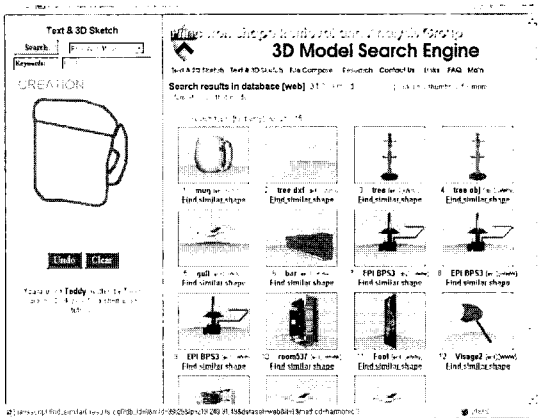
**Matching 단계 :**

각 사용자의 질의에 대해 웹 서버는 질의에 가장 잘 매칭하는 색인을 계산하여 16개의 3D모델을 반환한다. 일초 내에 저장소에서 결과물을 나타내며, 색인된 모델의 개수로 표현된다.

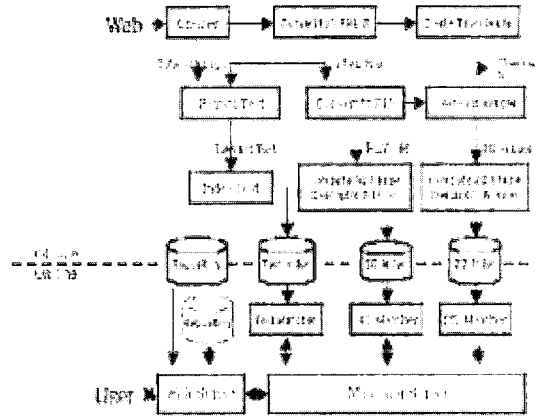
이 시스템에서 가장 중요한 것은 어떻게 형태 기반의 질의 인터페이스를 제공하느냐 와 쉽고 효율적으로 3D모델을 대용량 저장소에서 검색하는 것을 가능하게 하는 매칭 방법이다. 3D모델을 검색엔진의 질의로 주어 비슷한 것을 검색하도록 하는 형태기반질의의 중요한 점은 모델의 형태를 계산 가능한 표현, 즉 형태 묘사자로 표현하는 것이다. 형태 묘사자는 일반적으로 빠른 계산, 간결한 저장, 쉬운 색인, 유사 변환상의 고정성, 잡음이나 적은 특성에 덜 민감하고, 3D 물체의 표현, 배열, 종류에 독립적이고, 임의의 위상변화에 강하며, 다양한 크기에서도 형태의 차이를 구별할 수 있는 특성들을 가지고 있어야 한다. 그러나 이러한 특성을 모두 가진 형태 묘사자는 없기 때문에 좀더 일반화된 실린더 등과 같은 방법의 고수준의 형태표현은 물체의 내·외부모델의 일치성을 요구하고, 확장 가우시안 이미지나 모멘트, 웨이블릿과 같은 형

태 표현은 차이를 확연히 구별할 수 있는 규범적인 좌표계에 사전등록 하도록 요구하는 방법을 이용한다. 구면조화함수기반의 형태 묘사자를 이용할 경우 중요한 것은 방향 정보를 버리기 위한 구면조화함수를 이용하기 위해 3D모델을 동심구(concentric spheres)로 정의된 함수들의 집합으로 분해하는 것이다. 이렇게 되면 방향 고정과 묘사성을 가진 형태묘사자가 도출되는데 원래 형태가 이 표현으로 재구성될 수 없는 반면, 두 묘사자의 비교가 L2거리의 하한경계를 제공한다. 이 접근법은 3D모델을 규범 좌표계에 등록하지 않고도 색인 할 수 있는 장점이 있다. 폴라코너로 표현된 3D모델의 구면 조화함수 형태 묘사자는 다음 과정을 통해 구할 수 있다. 우선 다각형 표면을  $2R \times 2R \times 2R$ 의 복셀 격자에 주사하여, 다각형 표면이 복셀 하나의 폭 안에 들어있으면 1로, 그렇지 않으면 0으로 값을 정한다. 이때 R값은 32까지의 적절한 값으로 정하고, 정규화 과정이 필요하다. 다음단계에서 복셀 격자를 실수 값 함수로, R보다 작거나 같은 길이로 간주하여, 구좌표내에 함수로 표현한다.  $f(r, \theta, \phi) = \text{fr}(\theta, \phi) = \text{Voxel}(r \sin(\theta) \cos(\phi) + R, r \cos(\theta) + R, r \sin(\theta) \sin(\phi) + R)$  에서 구면함수집합  $\{f_0, f_1, \dots, f_R\}$ 을 구할 수 있다. 이때 구한 구면조화함수를 이용하여 서로 다른 주기들의 합  $\bar{m}$ 으로 표현할 수 있으며, 이때  $\|f\|$  값은 함수 fr이 회전한다 하더라도 변하지않는다. 이렇게 다른 반지름들을 가진 서로 다른 기호들을 조합하여, 3D 모델의 반지름 r0와 m0 주기에 대해 (r0, m0)색인에 대한 값을 가진 이차원 회전고정 구면 조화 묘사자를 구할 수 있다[10]. 두 구면 조화 묘사자를 비교는 간단하게 유클리디언 거리를 이용한다.

형태유사도 질의를 이용한 검색은 대표적인 3D모델이 이미 존재할 경우에 가능하지만 그렇지 않을 경우에 사용자는 텍스트질의나 최소한의 검색의 시발점이 되는 스케치정보를 이용할 수 있어야 한다. 웹의 3D 기하모델 검색에 익숙하지 않는 사용자의 경우 그 상세정도나 숙련된 스케치를 하기란 어렵기 때문에 가능한 쉽고 간단하게 3D모델을 검색할 수 있도록 인터페이스를 제공할 수 있는 것이 바로 사용자로 하여금 기본적인 스케치가 가능하도록 하는 것이다. 교육을 받지 않고도, 형태질의를 3D스케치 틀을 이용해서 줄 수 있는 첫번째 방법은 이 시스템이 제공하는 Teddy나



(그림 5) 2D 스케치를 이용한 3D 모델 검색 결과



(그림 6) 3D모델 검색 엔진의 데이터 흐름

Sketch를 이용하는 방법이다[10]. 이 툴들은 3D모델을 간단하게 생성할 수 있도록 도와주는 질의 인터페이스로 위의 매칭 알고리즘을 통하여 검색결과를 출력한다. 이 방법의 단점은 비록 간단하게 구성되었다 하더라도 3D모델의 특성 때문에 어느 정도의 3D물체 묘사교육이 필요하다는 것이다. 두 번째 방법은 2D형태의 그림을 그린다음 3D 물체의 2D 투영을 이용하는 방법인데, 이는 대부분의 사용자가 이미 이런 2D스케치에 익숙하기 때문에 위의 방법보다 배우기 쉽다는 장점을 가지고 있지만, 2D이미지의 경우 3D모델보다 형태정보가 적다는 점을 감안해야 한다는 단점도 있다. 그리고 2D 스케치를 2D물체와 매치하기 위한 알고리즘을 개발해야 하는 것도 주된 문제점이 될 수 있다. 사용자가 형태를 그리는 방법, 뷰포인트를 어떻게 선택해서 그릴지, 사용자의 입력제한을 어디까지 할 것인지, 수작업 스케치를 인식하기에 충분한 알고리즘은 어떤 것인지 등이 문제점이다.

대부분의 사용자가 스케치할 때 물체의 외형을 경계선형태로 표현하고 일반적으로는 측면, 전면, 윗면의 세가지 상태로 구분하며, 또 이 모양들은 대부분 물체의 중심을 기준으로 그리는 경향이 있기 때문에 직교 좌표계에 대입하여 물체를 표현할 수 있는 편리함도 있다. 스케치한 그림과 3D모델의 투영된 실루엣을 비교하여 매칭하는 데에는 물체의 변형이나 기하학적 부정확성 때문에 일반적인 다각형 모델의 경계선 매칭 알고리즘을 이용해서는 안 된다. 대신, Chamfer

matching[11] 과 Hausdorff matching[12]를 이용하여 색인할 수 있는 거리를 계산하는데 사용한다. 3D 모델이 임의의 방향을 가지고 있는 경우에는 모델을 이미지 평면의 반사와 회전에 견고한 이미지 매칭 방법이어야 한다. 이 매칭 과정은 (1)경계선의 거리 변환을 계산하고 (2)다른 반지름들과의 차이를 제한하는 원형 함수들을 모아 (3)각 원형 함수를 삼각 함수의 합으로 확장하여 (4)회전이 어떤 주기의 폭 안에서는 변하지 않는다는 사실을 이용하여 각 원형 함수의 기호를 구성 삼각함수의 크기의 리스트로 정하고 (5)경계선의 2D기호를 구하기 위해 서로 다른 기호를 조합한다. 이런 일련의 방법들은 Zahn과 Roskies의 경계곡선에 대한 회전고정기호를 제공하는 퓨리에 묘사자를 이용한 방법과 비슷하다. 3D모델을 검색하는 텍스트질의 묘사에 사용된 문자를 키워드로 매칭한다. 3D모델에 대한 문서를 구성하여 모델의 파일명과 연결된 웹 페이지, 모델파일내의 아스키라벨 등의 텍스트가 사용된다. 그리고 사용자 정의 키워드나 다른 문서와 매치하기위해 텍스트문서의 분류에 많이 적용되는 분류 방법인 TF-IDF 방법을 이용한다.

텍스트와 형태 질의의 방법을 함께 이용하는 다중 질의(multimodal query)방식은 텍스트와 2D/ 3D스케치를 하나의 질의로 넣는 방식이다. 먼저 텍스트질의 결과에서 비슷한 형태를 링크하여 텍스트와 3D 형태를 조합하여 두 번째 질의를 하는 방법이다. 이런 웹상의 질의를 이용한 3D모델의 검색 엔진의 데이터 흐름은 그림 6과 같다.



## 5. 결론 및 향후 연구 방향

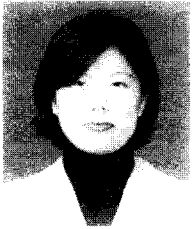
시각 데이터가 멀티미디어 컴퓨팅의 중심이 되어 가고있고 많은 연구들이 진행되고 있으며, 컴퓨팅과 커뮤니케이션 시장에서 시각데이터의 이용영역이 확대되고 있는 것이 현실이다. 2D/3D 시각정보 검색기술은 웹상에서 이용되는 데이터의 종류가 다양해 질 수록 더 복잡해 지고 다양한 기술을 필요로 한다. 본 논문에서는 시각정보를 검색하기 위하여 여러 가지 유사도를 이용하여 표현하는 방법과, 검색을 위하여 이미지나 기존의 문서 데이터들이 검색되는 방법, 특히 현재 개발되었거나 사용중인 내용기반 이미지 검색 시스템들의 특징들에 대해 간단히 소개하였다. 그리고 최근 이런 검색 시스템이 2D 이미지의 검색에 한정되지 않고 웹상에서 3D모델을 검색할 수 있는 형태기반 모델 검색 시스템개발로 이어지면서 2D에 비해 기술적으로 고려해야 할 점들에 대해 소개하였다.

지금도 여러 분야에서 많은 연구들이 시도되고 있지만, 시각데이터의 의미적인 정보 (semantic information)를 추출하고 운용하는 방법 등이 여전히 남아있는 문제이며, 3D모델의 검색에 있어서는 좀더 나은 2D 이미지 매칭 방법과, 새로운 질의 인터페이스 구축에 대한 문제나 새로운 매칭과 색인 알고리즘들의 필요성, 그리고 의학이나 분자 물리학과 같은 새로운 응용분야로의 확대 등에 대한 향후 연구가 필요할 것이다.

## 참 고 문 헌

- [1] <http://wwwqbic.almaden.ibm.com/stage/detail.html>, 2001.
- [2] J.Smith, and S.Chang. Automated image retrieval using color and texture. Technical report CU/CTR 408-95-14, Columbia University, July 1995.
- [3] T.Kato, Database Architecture for Content-Based Image Retrieval, Proc.of SPIE Conf. on Image Storage and Retrieval Systems, Vol.1662, San Jose, February 1992, pp.112-123.
- [4] V.N.Gudivaea, On spatial similarity measures for multimedia applications, porc. SPIE Conf. Storage and Retrieval for Image and video Databases III, Vol.2420, San Jose, CA, USA, February 1994, pp.363-380.
- [5] K.Hirata and T.Kato, Query by Visual Example:Content-Based Image Retrieval, Proc.E.D.B.T.'92 Conf. On Advances in Database Technology, In Pirott and Delobel and Gottlob editors, Springer-Verlag, Lecture Notes in Computer Science, Vol.580, 1994, pp.56-71.
- [6] I.K.Sethi and N.Patel, A Statistic Approach to Scene Change Detection, Proc. SPIE Storage and Retrieval for Image and video Database III, Vol.2420, San Jose, CA, USA, February 1995, pp.329-338.
- [7] V.N.Gudivada and Vijay V.Raghavan, Design and Evaluation of Algorithms for Image Retrieval by Spatial Similarity, ACM Transactions on Information Systems, 1995, April, Vol.13, No. 2, pp.115-144.
- [8] P. Aigrain and H. Zhang and D. Petkovic. Content-Based representation and retrieval of visual media: A state-of-the-art review. Multimedia tools and applications, volume 3, pages 179-202. Kluwer Academic Publishers, 1996.
- [9] Remco C. Veltkamp and Mirela Tansase. Content-Based Image Retrieval Systems: A Sruvey. Department of Computing Science, Utrecht University, Technical Report, 2000.
- [10] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A Search Engine for 3D Models. Submitted for publication, 2002.
- [11] H.Barrow, J.Tenenbaum, R.Bolles, and H.Wolf. Parametric correspondence and chamfer matching: two new techniques for image matching. In proceedings of the International Joint Conference on Artificial Intelligence, pp.659-663, 1977.
- [12] D.Huttenlocher, D.Klanderma, and Rucklidge. Comparing images using the Hausdorff distance. IEEE Transaction in Pattern Analysis and Machine Intelligence, 15(9): 850-863, 1993.
- [13] <http://shape.cs.princeton.edu/search.html>.

○ 저 자 소개 ○



한 태 화

1995년 경기대학교 전자계산학과(학사)

1998년 경기대학교 전자계산학과(석사)

1999년~현재 : 경기대학교 전자계산학과 박사과정

관심분야 : multiresolution modeling, 3D data compression, 시각 정보 검색



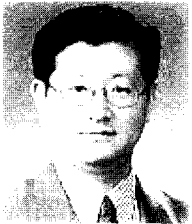
민 경 필

1996년 경기대학교 수학과(학사)

1998년 경기대학교 전자계산학과(석사)

2000년~현재 : 경기대학교 전자계산학과 박사과정

관심분야 : face detection and recognition, 3D face modeling



전 준 철

1984년 중앙대학교 전자계산학과(학사)

1986년 중앙대학교 전자계산학과(석사)

1987년~1988년 : 삼성반도체 통신연구소 연구원

1992년 The Univ. of Connecticut, 컴퓨터공학과(석사)

1995년 The Univ. of Connecticut, 컴퓨터 공학과(박사)

2001년~2002년 미시건 주립대학 패턴인식 및 영상처리연구실 객원교수

1995년~현재 : 경기대학교 정보과학부 부교수

관심분야 : 컴퓨터 그래픽스(3D face modeling & feature detection), 영상처리(2D, 3D face detection), 의료영상처리(MRI, CT 영상 처리)



최 용 길

1976년 고려대학교 무역학과, 심리학과(학사)

1980년 고려대학교 마케팅(석사과정)

1985년 미시간주립대 국제 경제학(석사)

1988년 조지아대학교 경영교육학과(석사)

1992년 조지아대학교 경영교육학과(박사)

1992년~1995년 조지아대학교 경영대학 기업경영 연구소 연구원

2001년~2002년 Texas A&M 주립대학 국제경영학과 객원교수

1997년~현재 : 경기대학교 경영학부 교수

관심분야 : 전자상거래, 사무자동화, Office System Management, 컴퓨터보안