

## 내용 기반 영상 검색 기술의 현황 및 전망

김낙우\* · 최종수\*\*

### 1. 서론

내용 기반 영상 색인 및 검색 기술은 최근 멀티미디어 정보의 폭발적 증가와 맞물려 가장 왕성한 연구가 진행되고 있는 분야이다. 야후나 라이코스, 구글과 같은 인터넷 검색 엔진은 사용자가 몇 개의 단어만 입력하면 방대한 양의 관련 문서들을 신속하고 정확하게 검색해 준다. 영상, 비디오, 음성 등을 포함한 디지털 문서가 점차로 늘어가고 있는 요즈음, 이러한 검색 엔진을 통해 새로운 미디어 영상물에 쉽게 접근할 수 있는가 하는 문제는 매우 중요하다. 그러나, 일반적인 텍스트 형태의 문서와는 달리 멀티미디어 영상으로부터 기존의 검색엔진을 통해 정보를 찾는 일은 쉽지 않다. 이에 따라 사용자에게 보다 효과적으로 멀티미디어 자료 검색 서비스를 제공하기 위하여 비디오 데이터베이스 관리 시스템(video database management system)이나 내용 기반 비디오 브라우저 등이 폭 넓게 연구되고 있다.

멀티미디어 데이터에서 그 내용을 대표할 수 있는 특징을 추출하고, 이를 기반으로 색인과 검색을 수행하는 것을 내용 기반 영상 검색 기법이라 한다[1]. 관리자에 의해 수동으로 삽입되는 주석 입력 방식이 아닌 영상 내 특징을 자동으로

추출하여 색인 과정에 사용함으로써 데이터베이스 구축에 필요한 시간 및 인력의 소모를 줄인다는 장점이 있다. 그러나, 멀티미디어 데이터가 나타내는 정확한 의미를 추출하기가 어렵다는 단점 또한 가지고 갖고 있다. 오늘날에는 2차원 영상의 검색을 지나 3차원 영상 검색에 이르기까지 그 연구 분야가 점차로 확대되고 있는 중이다.

2차원에서의 내용 기반 영상 검색 기술은 대상 데이터의 종류에 따라 정지 영상 검색과 동영상 검색으로 나뉜다.

정지 영상 검색 기술은 영상의 색상(color), 질감(texture), 형태(shape) 및 객체들의 공간적 위치(spatial location relationship)등이 주요 특징으로 사용되며, 동영상 검색 기술은 대상 비디오 데이터가 정지 영상이나 텍스트가 지닌 일반적인 속성 이외에 시간적 속성을 가지고 있고 비정형화된 구조를 하고 있기 때문에 비디오 분할(video segmentation), 특징 추출(feature extraction), 색인(indexing), 질의(query) 및 검색(retrieval) 등 비교적 복잡한 처리 기술을 필요로 한다[2,3].

3차원 물체의 검색에는 물체의 이동이나 회전 등에 대해 강건성을 지닐 수 있는 곡률, 체적, 포인트 거리 등이 특징 서술자(descriptor)로서 이용되며, 특히 특징 서술자를 선택할 때에 검색 시간에 대한 고려가 매우 중요하다.

\* 중앙대학교 첨단영상대학원 박사과정

\*\* 중앙대학교 첨단영상대학원 교수

※ 삼가 이 논문을故 박지환 교수님 영전에 올립니다.

현재의 '정보화 사회'는 다양한 종류의 멀티미디어 데이터에 대한 효과적이고 효율적인 데이터 관리 시스템 구축이 필수적이다. 우리 정부도 초고속 정보 통신망 구축 계획을 범 국가적인 차원에서 추진하여 서비스망을 구축, 성공적으로 시행하고 있으나, 여전히 실제 서비스 망에서의 멀티미디어 전반에 대한 유통, 저장, 처리 등이 취약한 것이 현실이다. 본 고에서는 현재 내용 기반 영상 검색 시스템의 현황을 분석하고 향후를 전망함으로써 멀티미디어 데이터의 효과적인 관리에 대한 필요성을 다시 한번 되새겨보고자 한다.

## 2. 내용 기반 영상 검색 기술의 현황

현재 내용 기반 영상 검색 시스템들의 대부분은 정지 영상을 대상으로 하는 시스템이며 인터넷을 통하여 시범적으로 서비스가 제공되고 있다. 대표적인 내용 검색 시스템의 개발 사례들은 아래 표 1과 같다.

이 외에 MIT대학에서 개발한 Photobook[10]과 같은 얼굴 영상 검색 시스템도 내용 기반 영상 검색의 특수한 응용 예라 할 수 있다.

검색을 위한 대표적인 시각특징정보로는 색상(color)과 질감(texture), 그리고 모양(shape)이 있다. 색상은 내용 기반 영상 검색 기법에 있어 가장 중요한 요소로서 모든 인지 색상은 RGB 색상의 조합으로 만들어 낼 수 있다. 일반적으로 색

상은 효율적인 특징 벡터로의 사용을 위해 양자화하여 사용한다. 질감은 영상 내의 시각 요소가 조밀하고 균등하게 나열 되어 있을 경우의 시각적 무늬를 말한다. 질감 특징은 크게 구조적인 분석과 통계적인 분석 기법에 의해 추출되는데, 통계적 분석 기법이 영상의 화소 단위 분석이라면 구조적인 분석은 좀 더 큰 매크로 영역에서의 질감 분석이라고 할 수 있다. 모양에 대한 특징 정보를 얻는 것은 다른 특징에 비해 매우 어렵다. 이미지에서 분할하고자 하는 대상 물체가 얼굴인지, 꽃인지, 비행기 인지 에 따라 각기 모양을 얻어내는 방법이 달라져야 하기 때문이다. 객체의 모양 특징은 영상에서의 에지 검출 후에 얻어지며, 해당 객체의 면적, 편심률, 원형성, 곡률 등의 특징 정보를 검색에 이용한다.

이러한 특징 정보를 이용한 정지 영상 검색 기법은 질의 영상에 따라 아래의 분류로 나뉜다.

- 아이콘에 의한 질의 - 미리 정의된 아이콘으로 관련 영상의 물체들을 표현한다. 이러한 아이콘들을 화이트보드 위에 배치함으로써 그 상관관계를 통해 영상 내 물체의 공간배열 등을 표현할 수 있다.

- 스케치에 의한 질의 - 사용자가 화이트보드 위에 검색하고자 하는 영상의 주요 특징을 영역별로 직접 스케치한다. 이러한 방법은 영상의 색상과 질감, 모양 정보 등의 검색이 가능하다.

표 1. 정지 영상 기반 영상 검색 시스템 개발 사례

개발 기관	시스템 명칭	웹 사이트 도메인 주소
IBM 社	QBIC [4]	<a href="http://wwwqbic.almaden.ibm.com">http://wwwqbic.almaden.ibm.com</a>
California 대학	Blobworld [5]	<a href="http://elib.cs.berkeley.edu/vision.html">http://elib.cs.berkeley.edu/vision.html</a>
Illinois 대학	MARS [6]	<a href="http://www-db.ics.uci.edu/pages/demos/index.shtml">http://www-db.ics.uci.edu/pages/demos/index.shtml</a>
Amsterdam 대학	ImageRETRO [7]	<a href="http://python.wins.uva.nl:6022/">http://python.wins.uva.nl:6022/</a>
Simon Fraser 대학	C-Bird [8]	<a href="http://jupiter.cs.sfu.ca/cbird/">http://jupiter.cs.sfu.ca/cbird/</a>
Columbia 대학	WebSEEK [9]	<a href="http://www.ctr.columbia.edu/webseek/">http://www.ctr.columbia.edu/webseek/</a>

■ 이미지에 의한 질의 - 샘플 이미지는 검색하고자 하는 영상의 기본형을 표현한다. 샘플 이미지의 시각 정보는 검색 영상의 시각 정보와 거의 유사할 것이며, 이는 전적으로 사용자의 선택에 달려있다.

아이콘 질의 방식은 의미론적 내용 접근에 의한 검색에 적합하다[11]. 또한, 공간적인 상관관계를 기반으로 한 검색에 이용되기도 한다. 이러한 질의 방식은 물체들 사이의 공간 관계를 쉽게 특징지을 수 있다는 장점이 있으나, 공간 확장이나 물체의 시각 정보는 고려되지 않는다는 단점 또한 가지고 있다. 그림 1은 아이콘 질의 방식 검색 시스템의 예이다.

그림 2와 같은 스케치 질의 방식에서 효과적인 검색 이미지를 만들기 위해서는 사용자의 스케치 능력이 필수적이다. 하지만, 사용자는 일반적으로 검색하고자 하는 영상의 시각적 의미를 제대로 표현하지 못하기 때문에, 그 질의의 결과가 사용자가 원하는 결과를 도출하는 것은 대단히 어렵다. 보통 스케치 질의 방식은 모양의 유사성을 기반으로 검색할 때나 일부 영역에서의 강조 색상을 알고 있을 경우 사용된다.

이미지 질의 방식은 폭넓게 사용되고 있는 검색 방법이다(그림 3 참조). 다양한 이미지 목록으

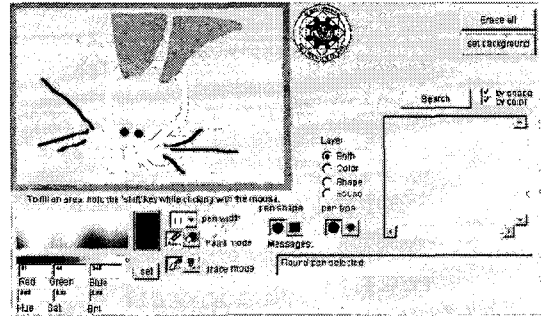


그림 2. 스케치 질의 방식의 예 - DrawSearch(13)

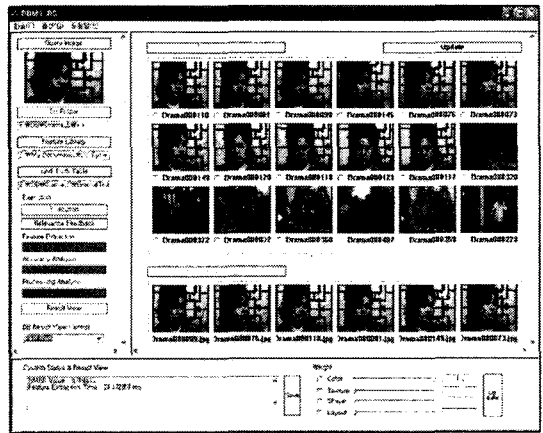


그림 3. 이미지 질의 방식의 예

로부터 질의 영상을 선택하여 데이터베이스의 영상과 특징을 비교한다. 이러한 방법은 스케치 질의 방법에 비해 사용자의 편의라는 측면에서 매우 유용하지만, 유사 영상을 검색하는 데에 있어 비슷한 배경의 영상이라 하더라도 주 물체의 축적, 회전, 특징 분포에 차이가 날 경우 제대로 찾지 못하는 등 좋은 검색 성능을 기대하기 어렵다는 단점이 있다. 질의 영상의 편집이나 relevance feedback 방식을 이용하면 성능은 더 나아진다. 이미지 질의 방식은 색상이나 질감 유사도에 의한 검색 방식에 사용되고 있다.

최근 3D 인터페이스 기반의 정지 영상 검색 방법도 등장하고 있다. 기존의 검색 방법에서는 같은 장소라도 보는 시점을 달리하여 찍은 영상의

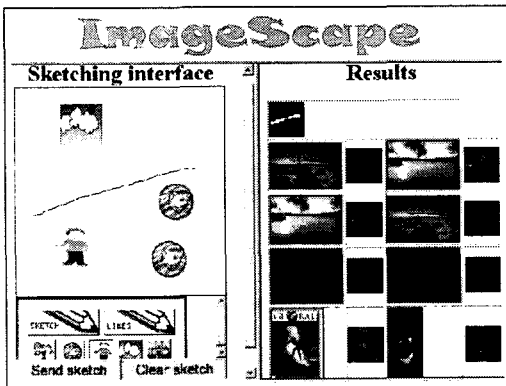


그림 1. 아이콘 질의 방식의 예 - ImageScape(12)

경우 색상 특징이 크게 달라질 수 있다는 문제로 인해 서로 간의 유사 영상으로 검색될 확률은 적었다. 그러나, 3D 인터페이스를 가지고 영상의 질의를 주었을 경우 보는 시점을 사용자가 정해줄 수 있게 되기 때문에 사용자가 원하는 영상을 제대로 검색할 수 있다는 장점을 가지게 된다. 이는 기존의 방법에서 영상의 시점 변환을 지원하지 않았기 때문이다. 그림 4는 3D 인터페이스 기반의 검색 시스템을 보여주고 있다.

동영상을 기반으로 한 영상 검색 시스템은 아직까지는 본격적인 서비스가 제공되지 않고 있으나 급격히 증가하고 있는 비디오 데이터에 대해 효과적이면서 동시에 효율적인 검색을 수행할 수 있는 동영상 검색 시스템 개발은 필수적이다.

동영상 검색 시스템의 개발은 다음 기술의 기반 하에 이루어져야 한다.

가) 비디오 인덱싱을 위한 알고리즘 개발

- 비디오 입력 처리 기능
  - 비디오 분할(video segmentation) 및 특징 분석(feature analysis)
- 비디오의 내용 특징 추출 기능
  - 대표 프레임 추출(key frame extraction)을 통한 시각 정보의 획득
  - 시간 정보를 통한 물체의 움직임 궤적 추적

- 추출된 내용을 인덱싱 기능
    - 추출된 각 내용 특징들을 표현할 수 있는 인덱싱 적용
    - 비디오의 전체적인 인덱싱 특징 벡터 도출
  - 인덱싱된 특징을 데이터 베이스화
- 나) 비디오 검색을 위한 알고리즘 개발
- 질의 비디오에 대한 특징 추출 및 인덱싱 알고리즘 수행
  - 질의 비디오와 데이터 베이스와의 비교/분석 기능
    - 질의 비디오의 인덱싱 결과를 분석
    - 질의 비디오의 인덱싱 특징 벡터와 데이터베이스 내의 인덱싱 특징 벡터들과의 유사도 비교
  - 질의 비디오와 가장 유사한 후보 비디오 순으로 출력 기능
    - 검색 결과를 유사도 순으로 배열
    - 후보 영상/비디오 순으로 사용자에게 출력

3. 차세대 내용 기반 영상 검색 기술

최근 내용 기반 영상 검색 기술은 새로운 환경의 변화를 맞고 있다. CAD시스템과 같이 3D 데이터를 다루는 작업군이 늘어나면서 3D 물체를 가진 3차원 영상들이 인터넷 상에서 점점 증가하고

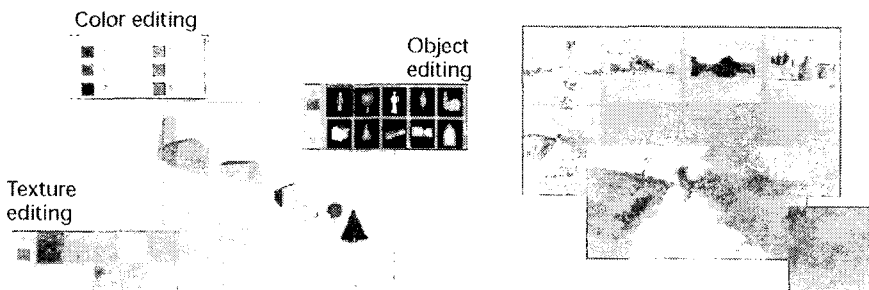


그림 4. 3D 인터페이스 기반 영상 검색 시스템의 예 - Query by Photographs(14)

있다. 기존의 2차원 이미지를 다루던 검색 기법은 3차원 영상의 검색에 직접적으로 활용되기 어렵기 때문에 최근 이러한 3차원 물체를 위한 영상 검색 시스템이 등장하고 있다.

2차원 영상의 비교를 위한 주요 특징이 색상이었다면, 3차원 영상에서의 주요 특징은 모양이다. 3차원 영상에서의 모양 특징을 올바르게 얻기 위하여 우리는 다음의 세가지 요소를 반드시 고려해야 한다.

- 3차원 물체에 대한 독립적인 표현, 이동이나 회전에 대한 강건성
- 잡음에 대한 강건성
- 빠른 계산 속도 및 쉬운 인덱스 기법의 사용

첫째, 3차원 물체는 다면체의 메쉬나 체적 데이터, 매개 변수 방정식 등과 같은 많은 표현 방법을 가지고 있기 때문에, 3차원 영상에서의 특징 추출 방법은 다양한 3차원 데이터 표현에 대해 독립적이어야 한다. 또한, 3차원 물체는 다양한 자세나 축척에서 저장될 수 있기 때문에, 특징 추출 방법은 3차원 물체의 이동이나 회전, 축척에 대해 불변해야 한다. 둘째, 3차원 물체는 3D 그래픽 프로그램에서 뿐만 아니라 3D 입력 장치를 통해서도 만들어질 수 있다. 이는 3차원 물체가 어느 정도의 에러를 포함할 수도 있다는 것을 의미한다. 때문에 특징 추출 방법은 에러에 대한 강건성을 확보해야만 한다. 마지막으로 특징 정보를 인덱스하기 쉬어야 하며 유사도 계산에 있어서 빠른 처리가 가능해야 한다.

이러한 요구를 수용하기 위해 많은 연구가 활발히 진행 중에 있으며, 주된 연구 경향은 다음과 같다.

- 2D view 기반의 방법  
3차원 물체는 각각 다른 각도에서의 시각 시점

으로부터 얻어진 2차원 영상들의 연속으로 표현될 수 있다. 이렇게 얻어진 2차원 영상들의 각각을 서로 비교함으로써 유사도를 얻는 방법을 2D view 기반의 방법이라 한다[15,16]. 예를 들면 다음과 같다. 먼저 주어진 3D 물체를 10개 시점에서의 2D 영상으로 분할하고, 이를 이진화하여 모양 정보를 취득한다. 취득된 10개 시점 모두에서 각각 영상 간 조합을 맞추며 매칭을 하고, 유사도 함수를 통해 최대의 유사도 값을 가지는 조합을 찾아내 그 결과를 사용자에게 보여줌으로써 검색을 마친다 [15].

- 히스토그램 기반의 방법

3차원 물체는 일반적으로 면이나 꼭지점, 수직 벡터 등 몇 개의 성분으로 분해될 수 있다. 이러한 벡터 성분으로부터 우리는 3차원 모델의 속성 히스토그램을 만들 수 있고, 이 히스토그램을 이용하여 각 3차원 물체를 비교한다. 3차원 모델을 표현하는 대표적인 방법으로는 가우시안 곡률, 수직 벡터의 변위, 중점 등을 들 수가 있다[17,18]. 전처리 과정으로 PCA(principle component analysis)를 이용하기도 한다.

- 토폴로지(topology) 기반의 방법

3차원 물체들 간의 구조를 그래프로 묘사하여 이를 이용하는 방법이다. 물체 표면의 모든 점에 대해 거리 계산하여 그래프를 생성한다. 아직까지 이 방법은 잡음에 민감하고 계산 시간이 오래 걸린다는 문제가 남아있다[19].

그림 5는 3차원 영상 검색 시스템의 예를 보여주고 있다.

## 4. 결 론

멀티미디어 정보는 현재 양적, 질적으로 엄청난 증가세를 보이고 있으며, ADSL과 같은 초고

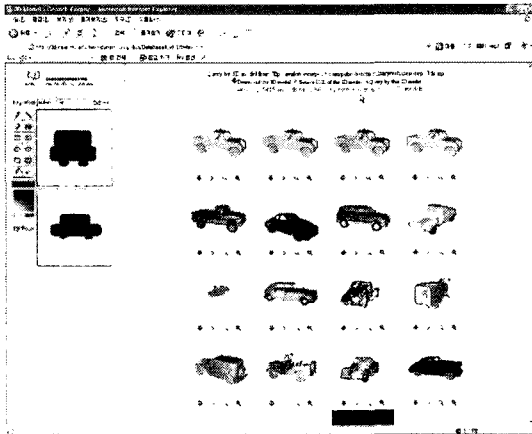


그림 5. 3차원 영상 검색 시스템의 예 3D model retrieval [15]

속 인터넷 환경의 보급과 더불어 고용량, 고화질의 데이터가 인터넷 콘텐츠의 대다수를 차지하고 있다. 국내의 소프트웨어 진흥원이나 한국전산원에서는 수백억의 예산을 편성하여 영상의 디지털화를 수행, 데이터 베이스화하였으며, 국외의 코비스와 같은 경우, 6천 5백만장에 이르는 영상 정보를 웹 사이트를 통해 판매하고 있는 등, 인터넷 환경을 통해 접근할 수 있는 디지털 시각 정보의 양은 기하급수적인 증가세를 나타내고 있다. 그러나, 요즈음의 인터넷 검색 환경은 여전히 텍스트 기반의 검색 및 브라우징 환경으로서 효과적으로 멀티미디어 데이터를 검색할 수 있는 시스템의 개발은 당면한 문제라 할 수 있다.

본 고에서는 질의 영상의 2D, 3D 환경을 분류 기준으로 하여 각각의 내용 기반 영상 검색 기술의 특징을 살펴보았다. 2차원 기반의 검색 기법은 영상에서의 색상, 질감, 모양 등의 시각 특징 정보와 움직임과 같은 시간 특징 정보를 바탕으로 이루어진다. 전자의 경우 정지 영상 검색에 사용되며, 시간상의 특징을 이용할 경우 동영상 검색도 가능해진다. 최근 3D 그래픽 산업의 발전으로 인한 3D 데이터의 증가로 3차원 기반의 검색 기법이

연구되고 있다. 3차원 데이터의 검색은 모양 특징을 위주로 하고 있으며, 크게 2D view 기반의 방법, 히스토그램 기반의 방법, 그리고 토폴로지 기반의 방법 등으로 나눌 수 있다.

내용 기반 검색 기법에서의 여러 가지의 분류 방법 중 단 하나의 최적 기법을 찾아내기란 대단히 어려운 일이다. 텍스트 기반의 검색 기법과는 달리 사용자마다의 시각적 주관에 강하게 개입이 되기 때문이며, 같은 질의를 냈을 경우라도 때에 따라서 원하는 결과 영상이 다를 수 있기 때문이다. 사용자의 주관 뿐 아니라, 현 기술 단계에서의 내용 기반 검색 시스템이 사용자가 원하는 수준과는 다소 거리가 있음도 사실이다. 이는 사용자가 단순히 보이는 것만의 유사상으로 질의에 대한 결과 영상을 찾지 않기 때문이다. 사용자가 원하는 의미 기반의 영상 검색 시스템의 개발의 위해서는 보다 체계적으로 집중적인 연구가 필요하며, 이러한 연구 분야에 대한 조기 접근을 통해 효율적인 멀티미디어 데이터베이스의 구축 및 제반 관련 사업의 발전 기반을 쌓을 수 있을 것이다.

참 고 문 헌

- [1] 이미숙, 황분우, 이성환, 내용 기반 영상 및 비디오 검색 기술의 연구 현황, *정보과학회지*, 제 15권, 제 9호, pp. 10-19, 1997.
- [2] A.K. Elmagarmid, et al., Video Database Systems: Issues, Products, and Applications, *Kluwer Academic Publishers*, 1997.
- [3] B. Prabhakaran, Multimedia Database Management Systems, *Kluwer Academic Publishers*, 1997.
- [4] W. Niblack, et al., The qbic project: Querying images by content using color, texture, and shape, *In Proc. of the SPIE Conf. on Storage and Retrieval for Image and Video Databases*, pp. 173-187, 1993.

[5] C. Carson, et al., Blobworld: A system for region-based image indexing and retrieval, *In 3th Int. Conf. on Visual Info. Systems*, LNCS 1614, pp. 509-516, 1999.

[6] M. Ortega, et al., Supporting similarity queries in MARS, *in Proc. of the 5th ACM Int. Multimedia Conf.*, pp. 403-413, 1997.

[7] J. Vendrig, et al., Filter image browsing: Exploiting interaction in image retrieval, *In 3th Int. Conf. on Visual Info. Systems*, LNCS 1614, pp. 147-154, 1999.

[8] Z.N. Li, et al., C-bird: Content-based image retrieval from image repositories using chromaticity and recognition kernel, Technical Report CMPT98-03, Simon Fraser University, 1998.

[9] J.R. Smith, Integrated spatial and feature image systems: Retrieval, compression and analysis, PhD thesis, Columbia Univ., 1997.

[10] A. Pentland, et al., Photobook: Content-based manipulation of image databases, *Int. Jour. of Computer Vision*, Vol. 18, Num. 3, pp. 233-254, 1996.

[11] J. Assfalg, et al., Three-dimensional interfaces for querying by example in content-based image retrieval, *IEEE Tran. on Visualization and Computer Graphics*, Vol 8, No. 4, 2002.

[12] J.M. Buijs, et al., Visual learning of simple semantics in imagescape, *In 3th Int. Conf. on Visual Info. Systems*, LNCS 1614, pp. 131-138, 1999.

[13] E.D. Sciascio, et al., Content-based image retrieval over the web using query by sketch and relevance feedback, *In 3th Int. Conf. on Visual Info. Systems*, LNCS 1614, pp. 123-130, 1999.

[14] J. Assfalg, et al., Querying by photographs: a VR metaphor for image retrieval, *IEEE Multimedia*, Vol. 7, pp. 52-59, 2000.

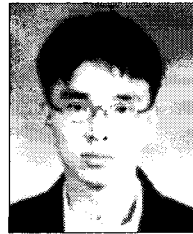
[15] D. Chen, et al., On visual similarity based 3D model retrieval, *Eurographics 2003*, Vol. 22, Num. 3, pp. 223-232, 2003.

[16] P. Min, et al., A 2D sketch interface for a 3D model search engine, *Siggraph 2002*, p. 138, 2002.

[17] J. Tangelder, et al., Polyhedral model retrieval using weighted point sets, *Int. Jour. of Image and Graphics*, Vol. 3, No. 1, pp. 209-229, 2003.

[18] T. Funkhouser, et al., A search engine for 3D models, *ACM Tran. on Graphics*, Vol. 22, pp. 83-105, 2003.

[19] M. Hilaga, et al., Topology matching for fully automatic similarity estimation of 3D shape, *SIGGRAPH 2001*, pp. 203-212, 2001.



김 나 우

- 1997년 2월 중앙대학교 제어계측공학과 졸업(공학사)
- 2002년 2월 중앙대학교 첨단영상대학원 영상공학과 졸업(공학석사)
- 2002년 3월~현재 중앙대학교 첨단영상대학원 영상공학과 박사과정 재학중
- 관심분야: 영상처리, 영상정보기술



최 종 수

- 1975년 2월 인하대학교 전기공학과 졸업(공학사)
- 1977년 2월 서울대학교 전자공학과 졸업(공학석사)
- 1981년 2월 일본 Keio University 전기공학과 졸업(공학박사)
- 1981년 9월~1999년 8월 중앙대학교 전자공학과 교수
- 1999년 9월~현재 중앙대학교 첨단영상대학원 교수
- 관심분야: 컴퓨터비전, 영상정보기술