

## 지능형 내용기반검색을 이용한 시각정보 자동추출

송점동\*

### 요 약

지능형 기본구조에 의존하는 내용기반 정보검색 시스템을 발전시키기 위한 많은 노력들이 진행중이다. 실세계에 존재하는 상황을 구축하고 유지하며, 동적 지식 표현을 실용화하고, 문맥을 이용하고, 그리고 진보된 추론과 학습능력을 이용하여 실세계에 존재하는 상황을 구축하고 유지하는 능력을 지능이라고 간주하여 이 지능을 채택한 진보된 내용기반 정보검색시스템을 제안하였다. 지능의 이러한 요소들은 인간의 지각과 인지에 정합하는 의미적 수준에서 시각 정보를 검색하기 위한 효과적인 시스템을 만들어내기 위해 필요로하는 것들이다. 본 논문에서는 인간이 가지고 있는 지능과 인지 심리학, 인공지능, 기호학의 분야에 존재하는 지능시스템의 구축에 관련한 연구들을 조사해 본 후, 기존의 내용기반 정보검색 시스템의 능력을 개선하기 위해 본 논문에서 제안한 멀티미디어 지식 표현 모형인 MMNet을 알아보고, 이것을 이용하여 시각 정보를 자동적으로 추출하기 위한 지능형 내용기반검색시스템 모형을 만들어 보았다.

## 1. 서론

최근 데이터 저장장치의 고용량화와 프로세서의 고속화가 급속히 진전됨에 따라 멀티미디어 자료를 다루는 기술의 중요성이 증가되고 있다. 특히 정보의 멀티미디어화에 따라 컴퓨터에 저장되고 관리되는 이미지의 양이 계속 증가되는 추세이다. 이러한 이유로 효율적인 이미지 검색 시스템에 관한 연구가 필요하게 되었다.

내용기반 이미지 검색시스템에서 내용이라는 것은 패턴의 변화, 색상, 질감, 이미지에서의 개체모양, 이미지 안에 있는 개체간의 공간유사도 등을 말한다. 이러한 이미지의 내용은 이미지를 데이터베이스에 삽입할 때

자동 또는 반자동으로 속성추출을 하게된다. 내용기반 이미지 검색시스템의 사용자 질의어의 예로써, “빨강색이 20%, 파랑색이 30%를 차지하는 물고기 모양의 개체가 있는 이미지를 찾아라”, “제시한 예제와 유사한 특징을 가지는 이미지를 찾아라” 등을 사용하여 이미지 검색을 할 수 있어야 한다.

시각정보의 자동추출에 관한 최근 연구는 내용기반검색을 지원하고 한정되긴 하지만 자동화된 객체의 분류, 그리고 사용자로부터 관련 귀환에 의한 학습 메카니즘을 지원하는 멀티미디어 정보 검색 시스템의 제1세대 격인 시스템들을 연구해왔다. 이러한 시스템들의 개발은 색인과 검색을 위한 문자적인 중심어에 완전한 의존으로부터 진보되어 왔지만 시스템의 성능 향상을 위한 올바른 연구방향 제시에는 아직도 미흡한 실정이다.

\* 경문대학 컴퓨터정보과 조교수

본 논문에서 제시한 시스템은 기존시스템에 지능을 갖춘 시스템으로서 차세대 멀티미디어 정보검색 시스템으로서의 첫발을 내딛는 시스템이라 여겨지며, 신호처리와 패턴인식 및 컴퓨터비전 분야에서도 초기 공헌했던 인지심리학, 인공지능 및 기호학으로부터 연구를 이끌어 낼 것이다. 앞으로의 내용기반 정보검색 시스템은 보다 높은 의미적 단계에서 시각정보의 이해 효율을 높이기 위하여 사용자와 통신하는 능력이 갖추어져야 한다고 내다본다.

먼저 이미지 정보검색시스템을 개발하기 위한 기존연구를 2장에서 알아보고, 3장에서는 본 논문에서 제시한 지능형 내용기반 검색을 가능하게 하는 멀티미디어 지식 표현 모형인 MMNet의 기본 개념과 지능형 내용기반 이미지정보검색 시스템의 구조를 설명하며, 4장에서는 시스템을 이용한 실험과 그 결과를, 5장에서는 결론을 설명한다.

## II. 기존의 연구

이미지 검색시스템은 1990년대 초기부터 활발하게 연구가 진행되어왔고, 검색 효율을 높이기 위해서 많은 응용 분야에서 연구를 수행해 왔다[9, 26]. 이미지 검색시스템의 최종 사용자는 사람이기 때문에 검색 결과는 마치 사람이 수동적으로 데이터베이스 전체를 훑어보아 원하는 이미지를 찾아내는 것과 같은 시스템을 갖추어야 한다. 이것은 참 어려운 문제이다. 왜냐하면 이미지에 대한 사람의 의미적 생각을 자동화된 알고리즘으로 부호화 하는 일은 참으로 어렵기 때문이다.

이러한 문제를 해결하기 위해 검색 시스템은 인간이 이미지를 범주화하는 것처럼 속성의 일부를 사용해서 인간과 같은 지능을 갖춘 시스템이 되기를 요구한다. 사람은 이미지 안에 존재하는 객체를 기반으로 이미지를 묘사하는 경향이 있다. 그래서 이미지 안에서 찾아진 객체들은 이미지 내용을 인간이 지각하도록 결과를 만들어 낸다. 그런데, 일반적인 이미지 영역에서 객체 인식은 매우 어려운 문제이고 쉬운 해결 방법이 존재하지 않는다. 이러한 객체 인식 문제를 피하기 위해 연구자들은 이미지 내용에 관련된 저수준 특징들을 찾아내기 위한 연구를 계속해 왔다. 이미지는 이러한 저수준 특징들, 또는 속성들로서 표현된다. 이것은 이미지가 그 의미적 내용과 관련된 저수준 특징들과 조화를 이룬다고 가정한다. 예를 들면, 기계 부품들은 그들의 모양을 근거로 구분될 수 있으며 판매하기 위한 제품들은 그들의 색상과 질감에 의해 식별될 수 있다.

이차원 형상은 다른 객체들로부터 어떤 객체를 구분하기 위한 하나의 중요한 특징이 될 수 있다. 패턴인식에서 취해왔던 괄목할 만한 연구성과는 객체 서로간의 형상을 일치시키는 것이다. Mehre 등[6]은 상표이미지 데이터베이스 상에 내용기반 검색을 위한 다양한 형태 측정에 관한 비교 연구를 했다. 형상을 표현하는데 사용된 특징들은 문자열을 부호화하고 Fourier 표현 계수와 같이, 그리고 그들을 다각형 근사[23]와 불변모멘트[7]과 같은 이미지에서 영역을 나타내는 등의 객체의 경계를 표현하는 방법으로 그들을 분류하는데 사용되었다. 그러나 이러한 많은 연구들은 객체가 형상 특징들이 계산되기 전에 배경으로부터 구분될 수 있다라는 가정하

에서 수행되었다. 이것은 객체가 평탄한 배경에 반대해서 묘사되는 곳이 데이터베이스에 대한 문제가 아니라 이것은 일반적인 이미지 데이터베이스에 대한 심각한 문제인 것이다. 일반적으로 이미지에서의 객체의 식별은 객체의 3차원 형상뿐만 아니라 객체와 카메라에 조명이 어떻게 비추는지에 따른다.

동물의 모피, 섭취하는 식물들과 같은 객체들의 주제들은 질감 패턴에 관련된 특징을 구분하기 위해 아주 중요하다. Ma와 Manjunath[29]는 이미지 정보검색을 위해 질감기반 패턴을 사용해왔다. Liu와 Picard[11]는 인간이 가지고있는 질감을 인식하는 성질인 주기성, 방향성, 어림짐작의 상호 직교좌표적 하부분야로 동종 난수분야에 근거한 이미지 모형을 제안하였다. 이러한 질감 특징은 지각적으로 유사한 자연적 질감들을 검색하는데 아주 효과적인 것으로 나타났다.

데이터베이스의 자료가 색상으로 구축된 저수준 특징에 색상을 일반적으로 많이 사용하였다. 색상은 상품, 깃발, 우표, 새, 물고기와 같이 아주 특별한 색상을 가지는 객체를 색인화하기가 아주 쉽다. Swan과 Balard[19]는 색상이미지를 색인하기 위해 색상 히스토그램의 사용을 제안하였고 유사 이미지를 찾기 위한 효율적 히스토그램 교차기술을 구현하였다. 히스토그램 교차 기술과 더불어 색상히스토그램 정규화는 유사 이미지의 검색 속도가 높고 일반적으로 전환, 회전과 크기 변화가 없다는 장점으로 색상이미지를 색인화하는데 아주 인기가 있는 방법이었다[15]. 그러나 색상 히스토그램은 색상을 구성하는 픽셀값의 공간적 구성에서는 정보를 통합하지 않는다. 이미지는 다른 구성에 있는 유사한 색상을 포함하는 이미지에는 많은 잘못된

정합이 일어날 수 있기 때문이다. 몇몇 연구자들은 검색 효율을 높이기 위해서 표현에 이러한 정보를 포함하려고 시도하였다. Zabih 등[13]은 이미지에서의 색상 분배에 더하여 색상들 쌍의 공간적 상호협력에 관한 정보를 포함하는 색상 상호협동을 제안하였다. Matas 등[14]은 다중 색상 객체를 표현하는데 사용할 수 있는 색상 인접 그래프를 이용하였다. 그러나 정합단계는 커다란 이미지 데이터베이스에서 사용한다면 너무 많은 계산을 거쳐야하는 문제가 있다. Das 등[18]은 이미지 질의에 의해 요구된 색상들의 공간적 성질들을 강화하기 위해 걸러내는 단계에서 사용된 단순 공간 인접 그래프구조를 제안하였다. 색상기반 이미지정보검색이 지니고 있는 주된 문제는 특징으로서의 색상이 일반적 이미지 데이터베이스에 들어있는 이미지 내용과 상호작용이 잘 이루어지지 않는다는 것이다. 예를 들어 붉은공을 찾는 질의어가 붉은색 자동차, 붉은색 꽃, 붉은색 옷을 입은 사람 또는 불자동차를 검색하는 것과 같다. 더욱이 색상만을 사용하는 것은 원숭이나 호랑이 등의 이미지들과 같이 단지 몇 개 안되는 색상만으로 이미지 데이터베이스를 식별하기에는 불충분하다.

개별적인 특징만을 사용하는 것 보다 여러 개의 특징들을 결합해서 정보검색의 효율을 높이는 연구가 진행되어왔다[8, 24]. Jain과 Vailaya[3]는 상표이미지의 데이터베이스를 색인하는데 특징들로서 색상 히스토그램과 형상을 사용하였다. 형상은 이미지에 표현된 다른 방향을 고려하는 히스토그램으로서 묘사된다. Belongie 등[25]는 색상과 질감 특징을 응집색상과 질감 영역으로 분리하는데 사용하였고 내용기반검색에서 이미지를 blobs

라는 용어로 표현하였다.

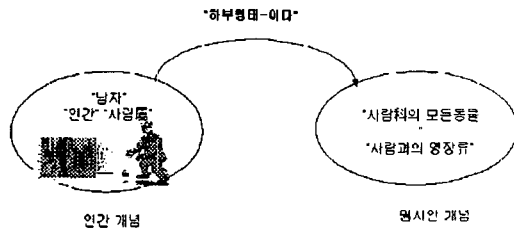


그림 1 인간과 원시인 개념의 다중 매체 표현과 '\*허무형태-이다\*' 관계

색상, 질감과 형상과 같은 특징들을 결합하여 이미지를 검색하는 응용시스템으로서 QBIC[28]은 예제이미지, 스케치 또는 선택된 색상과 질감 패턴의 예를 근거로 질의를 한다. 사용자는 최종 결과에 있는 각각의 특징에 붙는 중요한 특징들을 선택할 수 있다. Virage[16]는 플러그인으로 사용된 특징의 특징 영역과 같이 일반적인 색상, 형상과 질감을 사용하도록 허용하는 또다른 일반적인 이미지정보검색 시스템이다. Photobook[4]은 문자 주석에 추가한 특징으로서 형상, 질감과 고유이미지를 사용한 시스템이다. 이 시스템은 이미지의 특정 클래스에 관해 작동하도록 훈련된다. 다중 특징과 다중 질의 모드를 사용한 기존시스템의 또다른 예로서 Candid[22]과 Chabot[27]이 있다. 일반적 이미지 검색에서 돌출된 문제는 웹으로부터 관련 이미지를 검색하는 것이다. Smith와 Chang[17]은 VisualSEEK라 불리는 웹에서의 이미지 검색시스템을 구현하였다. 이 시스템은 이미지를 표현하는 이미지에서의 공간적으로 국부화시킨 색상 영역을 사용하였다. Sclaroff 등[17]은 웹으로부터 이미지를 모으고 색상, 질감, 위치와 다른 특정화된 특징들을 사용

해서 그들을 색인하는 ImageRover 시스템을 개발하였다.

위에서 논의한 것을 근거로 알아본 바는, 일반적인 이미지 검색 시스템은 거의 대부분은 특정화된 특징들과 마찬가지로 저수준 특징들을 사용했었다는 것을 알 수 있다. 그러나 사람들은 자기의 질의에 관련된 특징이나 특징들의 결합이 선택되기를 기대한다. 원하는 특징을 찾는 것은 아주 어려운 문제이고 특징을 사용하기 위해 특징들에 대한 지식과 경험이 요구된다. 다중 특징들을 사용함으로써 일어나는 보다 심각한 문제는 특징들을 어떻게 결합하는가를 아는 것이다. Nastar 등[8]이 만든 Surfimage 시스템은 정규화된 선형 결합과 특징들의 결합에 근거한 이미지의 계층을 계산하기 위해 투표 방법을 사용한다.

이미지 정보검색기술의 또 하나의 약점은 만들어진 시스템을 평가하는 방법이다. 대부분의 연구자들은 그들 자신의 개별적 데이터베이스에 관한 그들의 기술을 평가해왔다. 그와같은 방법은 항상 그 평가결과는, 특히 일반적인 이미지 수집에 집중하는 기술에서는 무엇이 그 평가의 표준인지가 분명하지 않다. 그런 데이터베이스에서의 예로서 유사성은 어떤 경우에는 정하기가 어렵다. 얼굴 검색 응용시스템인 FERET 데이터베이스[21]는 표준 검사 데이터베이스를 제공한다.

또다른 몇몇 시스템이 일반적인 데이터베이스에서 이미지 검색 문제를 풀기위해 시도해왔던 동안 사용자가 정말 필요로 하는 질문에 대해 아직까지도 답을 주지 못하고 있다. 대부분의 공통적인 질의 형식은 예제 이미지를 제공하는 것이지만, 이것은 사용자의 의도를 통찰하기에는 부족하다. 예를 들면,

사용자는 해가 떠있는 대낮에 어떤 건물 앞에 정차하고있는 차가 그려진 그림을 보고, 어떤 사람은 동일한 건물의 다른 그림을, 유사한 차들이 있는, 다른 햇빛을 받는 광경을 또는 건물 앞에 정차한 차를 가진 건물의 그림들 중에 어떤 한 그림을 원하는 것을 의미한다. 또 다른 시스템으로서 Surfimage[8]는 일찍이 문자기반 정보검색 시스템에서 검색 결과를 정제하기 위한 메카니즘으로 많이 사용해왔던 관련 귀환(relevance feedback)을 사용하였다.

앞서 알아본 여러 내용기반검색 시스템은 이미지 검색시스템은 저수준 시각적 특징들을 사용하여 이미지 정보를 색인화하기 위한 가능성을 연구해왔다. 이 시스템들은

- (1) 시각적 자료로부터 바로 특징들을 자동적으로 추출하는 것
- (2) 빠른 접근을 위해 추출된 기술어들을 색인화하는 것
- (3) 시각 자료의 검색을 위한 기술어들을 질의하고 정합하는 것으로 작동을 한다. 이러한 기존 능력을 뛰어넘는, 사용자가 찾고자하는 예제를 통해 질의와 학습을 정제하는 연관 귀환을 지원하는 연구가 시도되었다[25, 28].

### III. 멀티미디어 지식표현 시스템(MMNet)

MMNet은 기호와 이미지, 동영상, 음성, 그래픽과 문자와 같은 음성-시각 정보와 추출된 특징을 기술한 내용을 통해 실세계에

존재하는 지식을 표현하는 의미망이다. MMNet에서 실세계에 존재하는 개체들은 개념과 개념 사이의 관계로 표현되고, 멀티미디어 도구들에 의해 그 개체가 정해지고 만들어지게 된다. 예를 들면, 인간(Human)이란 개념은 "인간(human)", "사람(man)"과 "사람屬(homo)"과 같은 단어, 사람의 이미지, 그리고 사람이 말한 음성을 기록하여 나타낼 수 있고, "하부형태이다"라는 관계로 "유인원" 개념과 관련을 갖는다. 유인원 개념은 "유인원"이란 단어와 "사람과에 속하는 원시인"으로 정의할 수 있다. 그림 1에 앞선 예를 그림으로 나타냈다.

#### 3.1 개념

개념은 MMNet에서 지식의 기본 단위이다. 개념은 실세계에 있어서 객체들을 의미적으로 생각하는 것으로 간주한다. 객체는 무생물 객체, 생물 개체, 사건들과 특성과 같은 실세계에 존재하는 요소들이다. 개념의 예를 보면, 그림 1에서와 같이 어떤 살아있는 개체인 인간과 무생물 객체인 차를 들 수 있다. 결혼은 이벤트 및 젊음과 소유의 개념이다. 개념은 자동차와 같은 실세계에서의 객체 클래스로, 마이클 잭슨과 같은 유일하고 식별될 수 있는 객체로, 그리고 자유와 같은 실세계에서 물리적으로 존재하지 않는 추상적 객체로 간주할 수 있다. 언어와 개념 사이에 존재하는 단어들의 차이점을 찾아내는 것은 중요한 일이다. 앞선 예에서 개념은 통상적으로 한 단어로 이름을 부여한다. 그러나 인간들은 개념을 가리키는 한 단어 이상으로 어떤 동일한 개념을 유일하게 명시하거나, 그 개념을 유일하게 명시하고자 하는

단어들을 가지지 못하는 경우가 다반사이다. 첫 번째 경우는 직물의 예로서 아주 특별하지만 그 것이 무엇인지 알지 못하는 천 조각의 질감을 예로 들 수 있다. 두 번째 경우는 동의어, 즉 동일하거나 “인간(human)”과 “사람(man)”과 같은 언어에 있어서의 거의 동일한 의미 또는 감각을 가지는 단어에 일치하는 경우이다. 마지막의 경우로서 다의(polysemy), 즉 스튜디오(studio)와 같은 단어는 “작업실 또는 예술가의 아뜨리에”, “라디오 방송 또는 텔레비전 프로그램을 방영하기 위해 특별히 설비한 방이나 방들의 집합”, 그리고 “안방, 주방, 목욕탕을 구성하는 아파트”와 같은 언어에 있어서 한 단어가 가지고 있는 다중 의미를 나타내는 경우이다.

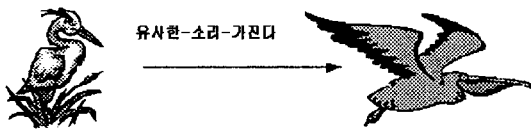


그림 2 소리 관계의 예

### 3.2 개념의 매체 표현

개념은 문자, 이미지, 동영상과 같은 멀티미디어 콘텐츠에 의해 표현할 수 있는 실세계에 존재하는 객체로 간주한다. 예로서 인간(Human)이란 개념은 “인간(human)”, “사람(man)”, 그리고 “사람屬(homo)”이란 단어와 인간을 나타내는 이미지, 그리고 인간의 음성 기록으로 나타낼 수 있다.

매체 표현은 전체 이미지와 같은 전체 멀티미디어 도구를 필요로 하지 않는다. 그들은 멀티미디어 도구의 구체일 수 있고 매체

로부터 추출된 특징들과 관련을 가질 수 있다. 인간이란 개념의 매체 표현은 인간 객체와 그 이미지 영역에 대한 이미지 형태의 윤곽 특징의 값을 포함한다. 개념은 멀티미디어 도구를 갖지 않는 특징 값들에 의해 증명될 수 있다. 예로서 인간 개념은 모양 윤곽 특징의 값에 의해 나타낼 수 있다. 이미지 형태의 윤곽 특징의 값은 인간을 묘사하는 이미지 영역의 집합에 대한 모양 윤곽 특징 값들의 평균이다. 그러나 이미지 자료는 매체표현에서 모양 윤곽 특징의 값만으로는 포함되지 않는다. 비록 이미지의 매체 표현이 어떤 멀티미디어 도구의 구체일 수 있고 멀티미디어 도구로부터 추출된 특징들이라는 사실을 예시하는데 사용되었다고 하더라도 결국 그 매체는 문자, 음성, 동영상 이외에는 되지 못할 것이다. 개념의 매체 표현을 좀더 살펴보자. 문자 매체로서, “차(car)”라는 단어의 정의는 “자동차(automobile)”이고, 이미지 매체로서, 자동차의 사진으로부터 추출된 모양 특징을 갖는 이미지로 나타내고, 음성 매체로서, 달리는 자동차의 소리 기록과 같이 멀티미디어로 차의 개념의 매체 표현을 나타낼 수 있다.

### 3.3 개념 사이의 관계

개념은 의미적이고 지각적인 관계에 의해 다른 개념과 관련을 갖는다. 의미적 관계는 그들의 의미에 근거한 관련을 갖는다. 렉시칼 시스템인 WordNet[12]에서의 모든 의미적 관계는 표 1에서 나타낸 정의와 예제에서와 같다. 반의어는 개념사이에 대칭 관계이다. 표1의 나머지 관계는 이진표현과 타동사 표현이다.

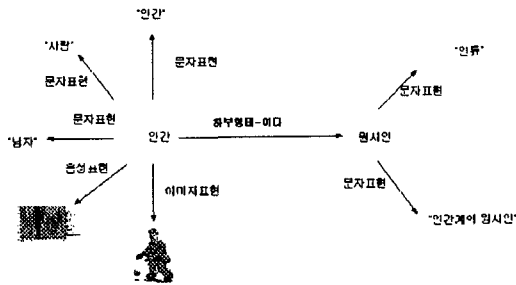


그림 4 MMNet의 예제에 일치하는 의미 네트워크

<표 1> 의미적 관계의 예

관계	정의	예제
반의	반대이다	흰색은 검정색에 반대이다
상하부	상부형태이다 하부형태이다	유인원은 인간의 상부 형태이다 인간은 유인원의 하부 형태이다.
부분	부분, 소속, 또는 대체 된다 부분, 소속, 또는 대체를 갖는다	배는 함대소속이다. 미티니는 진을 대체한다
수반	수반하기 또는 필요성이나 중요성에 의해 포함되거나 야기 된다.	이혼은 결혼을 수반한다.
방법	방법이 된다	속삭이는 것은 말의 방법이다.

개념은 감각에 의해 감지된 객체로 간주한다. 그러므로 개념은 감각을 사용함으로써 얻어진 인상에 근거한 지각적 관계를 가질 수 있다. 지각적인 관계의 예는 시각적 관계(예, 하늘 “유사한-색상-가진다” 바다)와 소리관계(예, 황새 “유사한-소리-가진다” 비둘기)이다.

### 3.4 의미망

인공지능 분야에서의 첫 번째 목적은 지능 개체들을 이해하고 지능 시스템을 구축하는

것이다. 인공지능이 컴퓨터에 제공한 중대한 공헌은 의미망과 같은 컴퓨터기반 지식 표현을 개발했다는 것이다. 의미망[20]은 객체, 개념 또는 상황 표현을 노드로 하고 아크는 노드들 사이의 관계를 나타내는데 사용한다.

MMNet의 의미망 구성을 살펴보면, 개념과 매체 표현은 노드로, 아크는 개념과 매체 표현의 연결을, 그 연결에는 표현의 형식을 표시했다. 의미적 관계는 개념노드 사이에 아크로 나타냈고, 매체 표현의 지각적 관계는 노드 대신에 아크로 나타냈다. 그림 3에 의미망의 예를 보였다.

### 3.5 지능형 내용기반 검색시스템

MMNet은 멀티미디어 정보를 포함함으로써 기존의 지식 표현 시스템을 확장하였다. 지식의 개념적이고 지각적인 관계를 통합함으로써 MMNet은 의미적이면서 특징 단계에서의 멀티미디어 콘텐츠를 다루는 폭넓은 응용시스템에 적용할 수 있다.

전통적인 내용기반 검색시스템은 저수준 시각 특징을 사용함으로써 이미지 정보를 색인하고 검색한다. 이 시스템들은 자동적으로 시각 자료로부터 저수준 특징들을 추출하고 데이터베이스 내부에 보다 빠른 접근을 위해서 추출한 기술자를 색인하고, 검색한다.

확장된 내용기반 검색시스템에서는 색상과 질감 특징을 사용한다. 색상 특징은 색상 히스토그램과 색상 결합성을 가진다. 질감 특징은 Wavelet 질감과 Tamura 질감이 있다. 내용기반 검색엔진에 요청하는 질의들은 데이터베이스에 있는 이미지의 이름으로서 시각 질의일 수 있고, 이미지 데이터베이스에 있는 저수준 특징들의 일부에 대한 값일 수

도 있다.

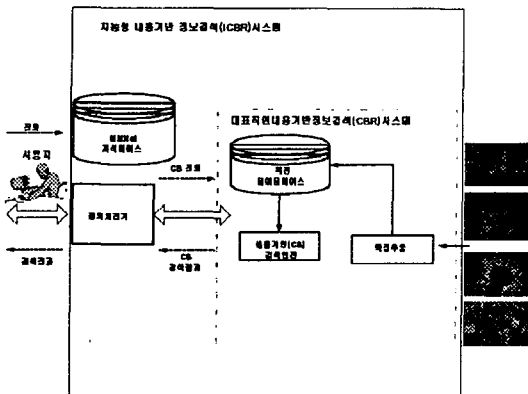


그림 4 자형형 내용기반 정보검색 시스템

MMNet에서의 지식 베이스 구축은 분류, 기계학습과 인공지능 도구를 사용해서 수동적 또는 자동적으로 이루어진다. 확장된 내용기반검색 시스템을 위한 지식베이스는 데이터베이스에 있는 이미지의 문자 주석을 사용해서 반자동으로 만들어지며, 지식베이스의 단어 입력은 자동화 렉시칼 시스템인 WordNet으로 이루어진다.

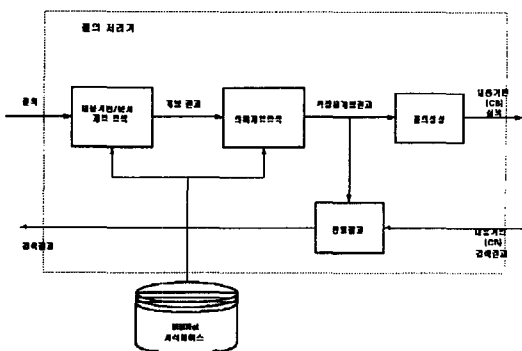


그림 5 이미 분류기

MMNet 시스템에 지식을 구축하기 위한 첫 번째 단계는 개념과 개념의 문자 표현을 WordNet과 사람의 도움을 사용하여 만들어 내는 것이다. 먼저, 문자 주석은 문자열 순으로 정렬시킨다. 전치사와 분사 그리고 중복된 단어들과 같은 불필요한 단어들을 제거한 후, WordNet을 사용하여 각각의 단어에 대해 동의어와 의미를 만들어낸다. 감독자는 문자 주석의 문맥에서의 각각의 단어에 대한 올바른 의미와 이미지 자료를 선택한다. 예를 들면 “스튜디오에 있는 가수”라는 주석에 대해 “스튜디오”에 걸 맞는 의미인 “라디오 방송 또는 텔레비전 프로그램을 방영하기 위해 특별히 설비한 방이나 방들의 집합”을 선택한다. 감독자는 역시 각각의 단어(명사, 동사 등)에 대한 문법적 범위를 지정한다. 개념은 각각의 단어/의미 쌍을 만들어 내며, 문자 표현으로서 WordNet에 의해 제공된 의미와 유의성을 할당한다.

다음 단계는 개념 사이에 존재하는 관계를 생성하는 것이다. WordNet은 모든 반의어와 동의어를 자동적으로 만들어 내고, WordNet의 출력을 자동적으로 파싱하고, 개념사이의 관계를 획득하며, MMNet 지식베이스에 그 결과를 저장한다.

마지막으로 개념의 시각적 표현들은 색상과 질감 특징 추출 도구를 사용해서 자동적으로 만들어낸다. 개념과 관련된 모든 이미지에 대해 색상과 질감 특징들을 추출해 내고 특징 중심을 계산한다. 개념 각각에 대한 시각적 표현은 관련된 특징 값을 지닌 개념에 대한 이미지 리스트이고 특징 중심이다.



### 3.6 질의 처리기

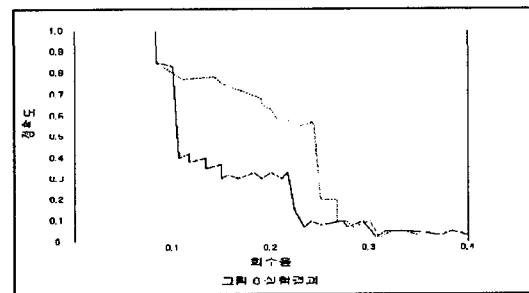
확장된 내용기반 검색시스템은 개념의 매체 표현을 가진 MMNet 지식베이스와 MMNet 지식베이스가 정제하고, 확장하고 사용자 질의로 번역하는데 사용하는 질의 처리기를 포함하고 있다. 비록 내용기반 검색 엔진이 단지 색상과 질감 특징에 기반한 내용기반 검색을 제공한다고 하더라도, 검색 엔진은 MMNet이 그들을 시각 자료로 변환하는데 사용할 수 있기 때문에 문자 질의를 받아들일 수 있는데 이것은 MMNet에 의해 검색 엔진에 그 기능성을 추가한다.

사용자가 문자 또는 시각적으로 질의를 제출할 때 질의 처리기는 중요도를 지닌 질의에 관련된 개념을 식별한다. 이것은 지식 베이스에 있는 개념의 매체 표현에 그 질의를 정합하는 것으로 처리된다. 해당 개념에 최고로 부여된 점수와 의미적으로 그 개념에 유사한 질의를 고려한 후 질의는 관련된 개념의 매체 표현은 내용기반 검색 엔진에 다중 질의들을 보냄으로서 이미지 데이터베이스에 그 값이 정합된다. 마지막으로 그 결과는 사용자에게 그들을 보이도록 유일한 리스트에 병합하고 수집된다.

## IV. 실험

이 실험의 목적은 이미지 검색 응용에서 MMNet을 평가하기 위한 것이다. 실험은 MMNet을 사용하는 확장된 내용기반 검색 엔진과 MMNet을 사용하지 않는 검색엔진과의 수행성능을 비교하였다. 양 엔진은 동일

한 내용기반 검색엔진과 이미지 데이터베이스를 사용하였다. MMNet 지식베이스와 질의 프로세서는 단지 전자에만 사용하였다. 이미지 검색 효율은 정확도와 회수율을 측정하였다. 회수율과 정확도는 검색 엔진의 효율성을 평가하는데 사용하는 표준적인 측정 방법이다. 회수율은 검색된 관련이미지의 비율로 정했다. 정확도는 관련된 이미지를 회수한 비율로 정했다. 실험에 사용된 이미지는 표준[23]에 제안된 색상 묘사 기술을 비교하고 평가하는 MPEG-7에서 사용된 5000개의 이미지들이다. 이미지의 영역은 스포츠, 뉴스, 가정생활에서 찍은 사진들이다. MMNet 지식 베이스를 구축하기 위해 사용된 문자 표현은, 50개는 일반화 특별화(hypernymy/hyponymy) 관계에 의해 다른 개념과 관련된 개념에, 34개는 구성, 합성 또는 대치(meronymy/hyponymy) 관계에 관련된 개념들이다.



실험을 통해본 결과, 확장된 내용기반 검색 엔진에서의 문자 질의에 대한 검색 효율은 이미지 질의 보다 아주 낮다. 그 결과는 확장된 내용기반 검색 시스템에 의해 이미지 질의에 대한 검색효율의 여분 획득을 보인다. 만일 두 개의 이미지가 회수되었다면, 양

시스템에 대한 회수율과 정확도의 값은 각각 0.5와 0.8이다.

## V. 결론

본 논문에서 의미적이고 지각적인 정보를 표현하기 위해서 멀티미디어 콘텐츠를 사용하는 지식 표현 골격을 갖춘 MMNet을 제안하였다. MMNet의 주된 성분들은 개념적 개체들과 의미적이고 개념사이에 존재하는 지각적인 관계를 포함한다. 지식의 개념적이고 지각적인 표현을 통합한 MMNet은 의미적이고 지각적인 단계에서 멀티미디어 콘텐츠를 다루는 응용의 광범위한 응용에 지대한 영향을 준다. 초기 실험 결과는 내용기반 검색시스템에서 MMNet을 사용해서 검색효율을 높인 것을 알 수 있었다.

## 참고문헌

- [1] 김철원, 최기호, "칼라 지정을 이용한 내용기반 화상검색 시스템 구현", 한국정보처리학회 논문지, 제4권, pp. 933-943, 1997.
- [2] 류근호, 김지호, "정보검색", 시그마프레스, 1995.
- [3] A. K. Jain and A. Vailaya. Image retrieval using color and shape. *Pattern Recognition*, 29: 1233-1244, 1996.
- [4] A. Petland, R. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image database. *International Journal of Computer Vision*, 18(3): 233-254, 1996.
- [5] B. Eisen, H. G. Tillman, and C. Draxler, "Consistency of judgments in manual labeling of phonetic segments: The distinction between clear and unclear cases," *Proceeding of the ICSLP(Barrff)*, pp. 871-874, 1992.
- [6] B. Mehre, M. Kankanhalli, and W. Lee. Shape measures for content-based image retrieval : A comparison. *Information Processing and Management*, 33, 1997.
- [7] C. C. Chen. Improved moment invariants for shape discrimination. *Pattern Recognition*, 26(5):683-686, 1993.
- [8] C. Nastar, M. Mischke, C. Meilhac, and N. Boujemaa. Surfimage: A flexible content-based image retrieval system. *Proc. ACM Multimedia*, Sept 1998.
- [9] D. Forsyth, J. Malik, and R. Wilensky. Searching for digital pictures. *Scientific American*, 276(6):72-77, Jun 1997.
- [10] D. Zier, J.-R. Ohm, "Common Datasets and Queries in MPEG-7 Color Core Experiments", ISO/IEC JTCl/SC29/WG11/MPEG99/M5060, Melbourne, Australia, Oct. 1999.
- [11] F. Liu and R. W. Picard. Periodicity, directionality and randomness : Wold

- features for image modeling and retrieval. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 18(7):722-733, Jul 1997.
- [12] G. A. Miller, "WordNet: A Lexical Database for English", *Communication of the ACM*, Vol. 38, No. 11, pp. 39-41, Nov. 1995.
- [13] J. Huang, S. Ravi Kumar, M. Mitra, W. J. Zhu, and R. Zabih. Image indexing using color correlograms. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 762-768, Jun 1997.
- [14] J. Matas, R. Marik, and J. Kittler. On representation and matching of multi-coloured objects. *Fifth International Conference on Computer Vision (ICCV)*, pages 726-732, Jun 1995.
- [15] J. P. Eakins, K. Shields, and J. Boardman. ARTISAN - a shape retrieval system based on boundary family indexing. *SPIE Proc. on Storage and Retrieval for Image and Video Databases*, 2670:17-28, Feb 1996.
- [16] J. R. Bach, C. Fuller, A. Gupta, and et al. The Virage image search engine: An open framework for image management. *Proceedings of SPIE*, 2670: 76-87, 1996.
- [17] J. R. Smith and S. F. Chang. Visually searching the web for content. *IEEE Multimedia*, 4(3): 12-20, Sep 1997.
- [18] M. Das, E. M. Riseman, and B.I A. Draper. FOCUS: Searching for multi-colored objects in a diverse image database. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 756-761, Jun 1997.
- [19] M. J. Swain and D. H. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11-32, 1991.
- [20] M. R. Quillian, "Semantic Memory", *Semantic Information Processing*, M. Minsky (ed), MIT Press, Cambridge, MA, 1968.
- [21] P. J. Phillips, H. J. Moon, P. Rauss, and S. A. Rizvi. The FERET evaluation methodology for face recognition algorithms. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 137-143, Jun 1997.
- [22] P. Kelly, M. Cannon, and D. Hush. Query by image example : the CANDID approach. *SPIE Proc. on Storage and Retrieval for Image and Video Databases*, pages 238-248, Feb 1995.
- [23] R. Mehrotra and J. E. Gary. Similar-shape retrieval in shape data management. *IEEE Computer*, 28(9):57-62, Sep 1995.
- [24] R. Picard, T. Minka, and M. Szummer. Modeling subjectivity in image libraries. *International Conference on Image Processing*, Sept

- 1996.
- [25] S. Belongie, C. Carson, H. Greenspan, and J. Malik. Color and texture-based segmentation using EM and its application to content-based image retrieval. Sixth International Conference on Computer Vision (ICCV), pages 675-682, Jan 1998.
- [26] V. N. Gudivada and V. V. Raghavan. Content-based image retrieval systems. *IEEE Computer Magazine*, 28(9):61-70, Sep 1995.
- [27] V. Ogle and M. Stonebraker. Chabot : Retrieval from a relational database of images. *IEEE Computer*, 28(9):40-48, Sep 1995.
- [28] W. Niblack, R. Barber, and et al. The QBIC project: Querying images by content using color, texture and shape. *SPIE Proc. on Storage and Retrieval for Image and Video Databases*, 1908: 173-187, Feb 1993.
- [29] W. Y. Ma and B. S. Manjunath. Texture-based pattern retrieval from image databases. *Multimedia Tools and Applications*, 2(1):35-51, Jan 1996.

## Automatic Selection of Visual Information using Intelligent Content-Based Retrieval

Jeom-Dong, Song\*

### Abstract

In this paper, we examine work in the evolution of content-based retrieval systems that rely on an intelligent infrastructure. Here, we refer to intelligence as the capabilities of the systems to build and maintain situational or world models, utilize dynamic knowledge representations, exploit context, and leverage advanced reasoning and learning capabilities. We argue that these elements are essential to producing effective systems for retrieving visual information at semantic levels matching those of human perception and cognition. In this paper, we review relevant research on the understanding of human intelligence and construction of intelligent systems in the fields of cognitive psychology, artificial intelligence, semiotics. We also discuss how some of the principal ideas from these fields lead to new opportunities and capabilities for content-based retrieval systems. Finally, we describe some of our efforts in these directions. In particular, we present MediaNet, a multimedia knowledge presentation framework that facilitates and enables intelligent content-based retrieval.

---

\* Dept. of Computing & Information Processing, Pyongtaek Institute of Technology