[UCI]I804:24011-200000233851

2007年 2月

碩士學位論文

# HSV 칼라 색인을 이용한 CBIR 검색 시스템 설계

朝 鮮 大 學 校 大 學 院

情 報 通 信 工 學 科

白 宗 玄

# *HSV* 칼라 색인을 이용한 *CBIR* 검색 시스템 설계

Design of an Content based CBIR System Using HSV Color Index

*2007*年 *2*月 日

朝 鮮 大 學 校 大 學 院

情 報 通 信 工 學 科

白 宗 玄

# *HSV* 칼라 색인을 이용한 *CBIR* 검색 시스템 설계

指 導 敎 授　朴　鍾　安

이 論文을 工學碩士學位申請 論文으로 提出함.

*2007* 年　*2* 月　　日

朝 鮮 大 學 校　大 學 院

情 報 通 信 工 學 科

白　宗　玄

白 宗 玄의 碩士學位 論文을 認准함

委員長　朝鮮大學校　敎授　韓　承　朝　　印

委　員　朝鮮大學校　敎授　申　暎　淑　　印

委　員　朝鮮大學校　敎授　朴　鍾　安　　印

2007 年　2 月　　日

朝鮮大學校　大學院

# 목 차

그 림 목 차 ····· iii
표 목 차 ····· v
Abstract ····· vi

*I.* 서 론 ····· 1
A. 연구의 배경 및 목적 ····· 1
B. 연구의 내용 및 구성 ····· 5

*II.* 멀티미디어 관리 시스템
A. MPEG (Moving Picture Expert Group) ····· 6
1. MPEG의 표준화 동향 ····· 6
2. MPEG-7 표준 ····· 8
B. 멀티미디어 정보 추출 ····· 12
1. Color ····· 12
2. Texture ····· 22
3. Shape ····· 23

*III.* 검색 엔진 ····· 25
A. 이미지 정보의 검색 방법 ····· 25
1. 문자 기반 검색 ····· 25
2. 의미론적 기반 검색 ····· 26
3. 내용 기반 검색 시스템 ····· 26

*IV.* 시스템 설계 ····· 31
A. 기술자 설계 ····· 31
1. 내용 기반 검색 시스템 설계 ····· 31
2. 인터페이스 설계 ····· 35

*V.* **구현 및 분석** ····· 36
A. 구현 ····· 36
1. 시스템의 구성 ····· 36
2. 사용 이미지 ····· 36
3. 구현 시스템 ····· 38
B. 분석 ····· 41
1. 색상 분산차 코렐로그램을 이용한 방법 I ····· 41
2. 색상 분산차 코렐로그램을 이용한 방법 II ····· 43
3. 제안한 알고리즘 ····· 44

*VI.* **결론 및 향후 연구** ····· 46

**참 고 문 헌** ····· 47

# 그 림 목 차

Fig. 2.1 Scope of MPEG-7 Standardization ····· 9
Fig. 2.2 RGB Color Model ····· 13
Fig. 2.3 HSV Color Model ····· 15
Fig. 2.4 Color Histogram ····· 16
Fig. 2.5 Three Color Histogram whose Perceptual Similarities do not Correspond to their $L_1$ Distances ····· 20
Fig. 2.6 Cumulative Histograms of the Histograms Displayed in Fig. 2.6 ····· 20
Fig. 2.7 Classification of Retrieval Techniques based on Shape Information ····· 24
Fig. 3.1 Structure of the Content-based Retrieval System ····· 27
Fig. 4.1 HSV Color Space ····· 32
Fig. 4.2 HSV Color Indexing Flowchart ····· 33
Fig. 4.3 Color Space Partition for HSV Color Indexing ····· 34
Fig. 4.4 Suggested System Flowchart ····· 34
Fig. 4.5 Interface Design ····· 35
Fig. 5.1 Using Image ····· 38
Fig. 5.2 Start Screen ····· 39
Fig. 5.3 Screen for Information Extraction and Data Storage ····· 41
Fig. 5.4 Image Retrieval Algorithm using Color Variance Difference Correllogram ····· 42

# 표 목 차

Table 3.1 Comparison of Content-Based Image Retrieval System ····· 28
Table 5.1 Matching Rate of Test Image ····· 42
Table 5.2 Matching Rate of Test Image ····· 43
Table 5.3 Matching Rate of Test Image ····· 44

# *ABSTRACT*

## Design of an Content based CBIR System Using HSV Color Index

Beak, jong-hyun

Advisor : Prof. Park, Jong-An, Ph. D

Department of Information & Communications, Graduate School of Chosun University

A variety of web services are needed due to development of information and communication network and rapid increase of the Internet demand, but as data of enormous volume are needed for services on the multimedia data, which causes overload in network, it is hard to provide quick and various services.

Therefore, as a rapid development of multimedia technology and increase of population to use internet, the necessity of retrieval system for various images on the web is on the rise.

There are two kinds of image retrieval system on the web. One is a system that retrieval to present images from the image database contained in the server in itself through web interface. The other is a system that present the results of retrieval by the text-based retrieval method for the images on the web. The former is high in the confidence level, but is not satisfactory for the web user's needs that make request for various images on the web. The latter can retrieve the number of various images on the web, but has the problem of the fall in responsibility for retrieval.

To solve the problem, MPEG-7, which is an international standard of meta data representation in multimedia, was established in early 2002. accompanied by development of various types of services.

Also, an incorporating service is becoming a matter of common interest because a variety of users require One-Source Multi-Use which was designed to apply multimedia services at various platform.

In this background, this paper suggested a image retrieval system using HSV color index. It extract a representative value of each block which divide a fit image into a block of prescribed approximate figures. In the following, it assigned index code value which divided color space of HSV into 24 fragments through compared a representative value of HSV. Lastly, it output a high degree of matching image to calculate a distribution in the color space of each blocks. A index code of overall image is saved in the database and search for wanting image use this on the web.

Proposed system constructed web-based multimedia retrieval system with JSP to verify the suggested system and simulated it by using several images.

# I. 서 론

## A. 연구의 배경 및 목적

전 세계 인터넷 사용 인구는 매년 기하급수적으로 증가하여 10억 명이 넘어서고 있다. 세계 최고의 정보통신망을 보유하고 있는 국내에 경우 우리나라 인터넷 사용인구가 3천만 명을 넘어섰다. 이는 지난 94년 KT와 데이콤이 인터넷 서비스를 시작한 이래 10년만에 인터넷 이용자수가 200배 이상 증가했고, 인터넷은 이제 우리 생활에 없어서는 안될 중요한 생활필수품으로 등장한 것을 의미한다.

국내 인터넷 이용자의 평균 TV 시청시간은 16.7시간, 평균 신문구독시간은 4.8시간인 반면 인터넷 비이용자의 TV 시청시간은 23.2시간, 신문구독시간 5.3시간으로 나타나 인터넷이 정보매체를 대체하고 있는 것으로 조사됐다. 특히 12-19세 학생층 인터넷 이용자의 경우 인터넷을 통한 정보 입수비율이 84.3%로 가장 높은 반면 TV 81.4%, 신문 44.2%, 라디오 10.2%인 것으로 나타났다. 연령별로 볼 때 그동안 인터넷 이용률이 10대나 20대에 비해 상대적으로 낮았던 30대와 40대의 인터넷 이용률이 전년 대비 각각 7.9%, 7.5% 증가하였다[1].

또한, 중국이 인터넷 이용자 수가 9400만 명을 돌파하면서 이용자수 대비 미국에 이어 세계 2위의 인터넷 대국으로 올라섰다[2].

이러한 인터넷의 이용자 수의 증가는 반도체 기술의 발전으로 하드웨어의 성능이 100배 이상 향상 되었고, 광통신 기술의 발전으로 통신망의 속도가 1000배 이상 증가되어 양적/질적 정보를 제공할 수 있는 기반을 갖추었기 때문이다[3]. 그로인해 기존의 텍스트 기반의 웹 컨텐츠가 고용량의 멀티미디어 데이터가 포함되어 있는 웹 컨텐츠로 급격히 변하고 있다.

인터넷 이용자 수의 증가는 인터넷 서비스의 양적인 증가와 더불어 서비스 질적 상승을 요구하게 된다. 따라서, 정보의 바다인 인터넷에서 원하는 정보를 빠른 시간 안에 정확하게 얻기 위한 정보 검색 도구의 개발이 지속적으로 요구되고 있으며 많은 업체에서 이러한 정보 검색을 위한 시스템을 개발하여 운영하고 있다. 그러나 대부분의 정보 검색 시스템이 HTML 태그 정보를 이용한 텍스트 기반 검색에 치중하고 있어 문자로는 그 특성을 상세히 기술하기 어려운 이미지를 포함한 멀티미디어 정보 검색에는

문제점을 가지고 있다.

멀티미디어 컨텐츠들에는 이미지, 그래픽, 사운드, 동영상, 애니메이션 등과 같은 대용량의 다양한 멀티미디어 데이터로 이루어져 있으며, 이러한 다양한 멀티미디어 데이터들을 효과적으로 검색하기 위해서는 각각의 미디어별 데이터 구조설계와 그에 적합한 검색 엔진의 개발이 필요하게 되었다.

하지만, 초기의 검색 시스템에서는 질의 자료로 문자 키워드를 제시하고 동일한 문자 색인을 가진 이미지를 검색해내는 텍스트 기반 검색 시스템이 주류를 이루었으며, 이러한 검색 시스템은 몇 가지 문제점을 가지고 있다. 즉, 데이터베이스 상의 모든 이미지에 대해 각 이미지를 잘 표현할 수 있는 키워드를 설정해야 하는 수고가 요구되며 또한 키워드 설정 시 개인의 주관성이 포함되기 때문에 고유한 키워드를 설정할 수 없는 경우들이 존재한다. 또한 검색 시 사용자가 주어진 이미지에 대한 고유의 키워드를 기억하고 있어야 하는 단점도 지니고 있다. 따라서, 좀 더 객관적이고 자동화된 검색 시스템에 대한 요구가 나타났으며, 이에 대하여 문자 키워드 대신 검색하고자 하는 이미지와 유사한 이미지나 또는 이미지에 대한 정보(예를 들면 색상이나 형태 등)를 질의 자료로 제시하고 자동적으로 질의 자료와 유사한 이미지들을 순서대로 출력하는 시스템이 제시되었으며 이러한 시스템을 내용 기반 이미지 검색 시스템(Content-Based Image Retrieval : CBIR)이라고 한다.

내용 기반 이미지 검색 기술은 질의 자료로 이미지 자체 또는 이미지에 대한 관련된 정보를 제시하면 질의 자료를 분석하여 특징 정보를 추출한 후 이미지 데이터베이스 상에서 가장 유사한 특징을 가지는 이미지를 검색하는 기술이다. 내용 기반 이미지 검색 기술에 있어서 가장 중요한 요소는 이미지의 내용을 어떠한 특징으로 표현할 것인가 하는 문제이며, 이미지의 특징은 크게 3가지 레벨로 분류될 수 있다. 먼저 원시적인 특징으로서 이미지의 색상, 형태 및 질감 정보 등이 이에 해당되고, 두 번째로 논리적인 특징으로서 포함되어 있는 물체의 식별을 위한 논리적인 정보이다. 마지막으로 추상적인 정보이며 이미지의 느낌, 감정, 장면의 중요도와 같은 추상적인 표현을 말한다. 현재의 이미지 검색 기술은 대부분 이미지의 정보로서 색상, 형태 및 질감 등의 원시적인 특징을 사용하는 기술이 많이 연구되고 있다.

색상 정보를 사용하는 기술은 계산과 정보 표현에 있어서 간결하기 때문에 가장 많이 사용되는 방법으로서 색상 정보로서 주로 색상 히스토그램이 사용된다. Swain 등이 제안한 히스토그램 인터섹션 방법이 대표적이며[4], 색상 비율의 히스토그램을 비교

함으로서 물체를 인식하는 확장 Swain 방법[5], 그리고 색상 비교를 하기 위해 거리 방법과 기준 색상 테이블 방법[6]이 제안되었다. 색상 정보를 이용한 방법은 이미지 내에서 물체의 이동 및 회전, 왜곡 등의 변화에 강인하며 쉽게 적용될 수 있다는 장점을 가지지만 조명의 변화에 따라 성능이 크게 저하된다. 또한 내용은 전혀 다르지만 색상 분포가 비슷한 원하지 않는 이미지가 출력될 가능성이 있다[5].

형태 정보를 이용한 이미지 검색은 검색 기술 중 가장 적용하기 어려운 방법이다. 형태 정보를 이용하기 위해서는 이미지 내의 물체를 분리하는 것이 중요하며 따라서 물체의 윤곽이 비교적 뚜렷한 경우에 있어서 좋은 성능을 보인다[7]. 물체의 윤곽이 결정되면 물체의 모양은 면적, Eccentricity (Major 축과 Minor 축의 비율), 원형성(같은 면적을 가진 원과의 유사도), 형상 Signature (물체의 중심과 외곽과의 거리의 시퀀스), 곡률(Curvature; 윤곽이 회전하는 정도의 측정), 프랙탈 차원(자기 유사도의 정도) 등의 정보에 의해 표현될 수 있다. 형태 정보를 이용한 검색 기술로는 체인 코드[11], 저나이크 모멘트[8-10], 불변 모멘트[11, 12], 퓨리어 묘사[11] 등을 특징 파라미터로 사용하는 기술들이 제안되었다.

질감 정보를 이용한 이미지 검색은 웨이블릿 변환과 같은 다 해상도 필터링 기술과 함께 발전되어 왔다. 질감은 많은 시각 요소들이 밀집해서 균일하게 정렬되어 있는 패턴이다[13]. 이를 분석하기 위해서 다 방향 필터뱅크를 이용하거나, Markov 모델을 사용한 모델 기반 분석[114]을 하는 연구가 많이 수행되었으며 최근에는 Gabor 웨이블릿 변환을 이용하여 이미지의 질감 정보를 주파수 대역별로 분해하여 분석하는 방법이 많이 사용되고 있다[15].

또한 두 가지 이상의 특징을 혼합하여 검색을 위한 정보로 이용하는 방법들도 제시되었는데, Jing Huang 등은 색상 상관도를 도입하여 색상 정보 및 색상들의 공간적인 배치 정보도 동시에 고려하였으며, D.K. Park 등은 색상, 그레이 정보 및 에지 정보를 동시에 고려하여 복합 히스토그램을 생성하여 검색을 하였다. 그리고 Pass 등은 색상 히스토그램의 각 색상 bin에 대해 정제화를 통해 지역적인 정보를 포함시키는 방법을 제안하였다.

현재 내용기반 이미지 검색 기술을 구현하여 상용화된 제품으로 IBM의 QBIC, Virage의 VIR Image Engine, 그리고 Excalibur의 Image RetrievalWare 등의 검색 시스템이 있다. 또한, 많은 실험적인 데모 시스템이 WWW 상에 존재하고 있는데, MIT의 Photobook , Columbia 대학의 WebSEEk , 그리고 Carnegie-Mellon 대학의

Informedia 등이 대표적이다. 내용 기반 이미지 검색 시스템은 이제야 시장에 발을 들여놓기 시작한 상황이며, 주요 응용 분야로 범죄 방지(지문인식 및 얼굴인식 등), 지적 재산권 보호(상표 등록 등), 방송 매체 및 광고 (비디오 재료 관리 등), 그리고 Web 검색 등이 있다. 대표적인 Web 검색 엔진인 AltaVista 와 Yahoo의 검색 엔진들도 각각 Virage와 Excalibur의 검색 기능을 제공하고 있다.

내용기반 이미지 검색과 관련한 표준화 기술을 보면 크게 3가지 방향(이미지 압축, 질의 규정, 메타데이터 묘사)으로 나뉘어 질 수 있다. 현재 진행 중인 중요한 표준화로 멀티미디어에 대한 메타데이터 표현의 국제표준인 MPEG-7이 2002년 초에 제정되면서 다양한 형태의 서비스들이 개발되어 지고 있다[20-24]. 이 표준화에서는 내용 기반 이미지 검색의 핵심 기술인 멀티미디어 내용을 표현하기 위한 현재의 유일한 표준안으로 크기, 모양, 색상, 위치, 동작 등에 해당하는 객체의 저수준의 표현 방식과 객체의 종류, 사건, 장소, 장면 묘사 등의 고수준의 표현 방식을 모두 표현하게 됨으로써 현재의 내용 기반 이미지 검색 기법의 발전을 가속화시킬 것이다.

또한, 멀티미디어 서비스를 다양한 플랫폼에서 적용하고자 하는 플랫폼에 독립적인 One-Source Multi-Use 현상이 크게 대두되고 있어 각 플랫폼 간의 통합 서비스가 주요 관심의 대상이 되고 있다[25, 26].

이러한 색상, 질감, 형태 정보 중에서 가장 많은 정보를 포함하고 있는 색상 정보가 이미지 검색의 가장 핵심적인 기술이라 할 것이다. 특히, 색상 공간 중에서도 HSV는 RGB가 색상만을 표현하는 것과는 달리 색상, 채도, 명도를 모두 표현함으로써 이미지의 밝기가 변경되거나 크기, 회전 등의 변화가 있더라고 RGB에서 발생하는 오류를 해결할 수 있는 장점을 가지고 있다.

이러한 배경에서, 본 논문에서는 HSV 칼라 인덱스를 사용한 이미지 검색 시스템을 제안한다. 해당 이미지를 일정한 개수의 블록으로 분할하고 각 블록의 대표값을 추출한다. 다음단계로 HSV 대표값을 비교하여 HSV 색상공간을 24개로 분할한 색인 코드 값을 할당한다. 마지막으로 각 블록들의 색상공간에서의 분포도를 계산하여 매칭도가 높은 이미지를 출력하게 된다. 전체 이미지의 색인코드는 데이터베이스에 저장되고 웹상에서 이를 이용하여 원하는 이미지를 검색하게 된다.

제안된 시스템을 검증하기 위하여 웹 기반 멀티미디어 검색 시스템을 JSP로 구축하였고 몇 개의 이미지를 사용하여 이를 검증하였다.

## B. 연구의 내용 및 구성

본 논문에서 구현하고자 하는 웹 기반 멀티미디어 관리 시스템을 다음과 같은 단계를 거쳐 구현하였다.

첫 번째로 다양한 MD 데이터를 저장하기 위해 My-SQL 데이터베이스를 사용하였고, 해당 데이터베이스에 접근하기 위해 JSP-Bean 기반으로 웹에서 제어할 수 있도록 하였고, 데이터베이스 드라이버로 Type 4 방식의 JDBC 드라이버를 사용하였다.

두 번째 단계로 기존의 색상 검색 방식의 단점을 극복하기 위해 본 논문에서는 색상 기술자로써는 HSV 색상을 인덱싱하는 알고리즘을 사용하였다.

세 번째 단계로 본 논문에서 제안한 알고리즘을 테스트하기 위해 테스트용 멀티미디어 데이터베이스 및 이미지 디렉토리를 구축하였다.

마지막으로 전체적인 시스템이 웹상에 운영되는 시스템을 JSP-Bean을 이용하여 구축하였다.

이와 같은 단계를 거쳐서 구현한 시스템은 다음과 같은 요소들로 구성된다.

- 데이터 구조에 맞도록 해당 이미지의 기술자 정보를 추출하고 등록하는 기능
- MD 데이터베이스로부터 선택된 검색 도구에 적합한 검색을 하는 검색 기능
- 기본적인 멀티미디어 데이터 처리를 할 수 있는 기능

이를 위해서 II장에서는 MPEG, 기술자들에 대한 기존의 연구, 데이터 구조 설계에 대한 내용과 기존 연구에 대해서 살펴보고, III장에서는 검색 엔진에 대한 연구에 대해 살펴보고, IV장에서 제안한 시스템을 설계하고, V장에서 이를 시험 이미지를 통해 시뮬레이션을 하여 이를 분석하고, VI장에서 결론을 맺는다.

# II. 멀티미디어 관리 시스템

## A. *MPEG-7 (Moving Picture Expert Group)*

### 1. *MPEG* 표준화 동향

오디오와 동영상 서비스를 실현하기 위해서는 방대한 양의 데이터를 효과적으로 저장하거나 전송하기 위한 압축기술의 개발과 표준화가 필수적이다. 1980년대 초반부터 시작된 디지털 이미지기술의 발전이 정지이미지에서 시작되어 동영상으로 옮겨졌고, 결국 멀티미디어 서비스를 위한 발판을 구축하게 되었다. 그러나 멀티미디어 서비스를 효율적으로 제공하기 위해서는 오디오와 비디오 같은 데이터를 압축하는 기술뿐만 아니라 데이터를 효과적으로 연관(link)시키고, 사용자에 의한 대화형 조작을 쉽게 하고 압축된 데이터를 효과적으로 검색할 수 있도록 하는 도구도 필요하게 되었다. 이와 관련된 기술의 표준화가 ISO/IEC를 중심으로 진행되어 왔는데 그 주요활동은 다음과 같다.

MPEG(Moving Picture Experts Group) 표준화는 저장 미디어에 동영상을 압축하여 저장하기 위해서 시작되었다. 1988년 1월 ISO/IEC JTC1/SC2(Joint Technical Committee 1/Sub Committee 2, 현재는 SC29)에 동영상의 부호화를 연구할 전문가 그룹 MPEG을 설립하였다. 이는 이미 진행되었던 통신분야에서 ITU(International Telecommunication Union)의 H.120, H.261 등과 같은 부호화 표준과는 다른 기능과 특성을 요구하는 상황에 대응하기 위한 활동을 시작하며 표준화를 지향하였다. 이때의 표준화 대상은 약 1.5Mbps의 전송속도를 가지는 CD-ROM이었고 ISO/IEC 11172로 완성되어 1992년에 승인되었는데 이것이 잘 알려진 MPEG-1이다.

보다 고품질의 디지털 방송을 위한 수 Mbps~수십 Mbps급의 이미지부호화 표준화 작업은 1990년 7월부터 MPEG-2를 목표로 시작되었는데, MPEG에서는 기술개발담당부 그룹과는 독립된 Requirement 부 그룹을 설치하여 범용 표준 실현을 위한 요구조건을 취합하였다. 동시에 부호화 작업에 있어서 다양한 응용들이 공통적으로 사용할 수 있도록 하기 위한 요구 조건들을 만들어 표준화 작업을 지원하였다. MPEG-2의 표준화 과정은 MPEG-1을 기반으로 하여 매우 광범위한 분야를 지원할 수 있도록 하여 디지털 저장매체 뿐만 아니라 방송이나 통신 분야의 응용에 대해서도 널리 적용할 수

있도록 하였다. 이러한 범용성을 추구하기 위하여 우선 다양한 응용의 기능과 특성에 대한 요구조건을 수립하는 것으로 시작되었다.

MPEG-4는 수요에 의해 개발되었다기보다는 표준화를 이끄는 기술자들의 기술적인 비전에 의하여 주도된 측면이 아주 크며 목표가 고정되어 있다기보다는 끊임없이 기술의 발전과 더불어 수정되고 보완되었다고 볼 수 있다. MPEG-4는 원래 1998년 11월에 IS를 완성하도록 되어 있었으나 여러 차례 목표를 수정해 온 결과 현재까지 고려한 모든 기능들에 대해 1998년까지 표준화하기 어렵게 되었다. 그러나 마이크로소프트사가 MPEG-4의 일부 기능만을 가지고 상용화에 돌입하겠다고 선언하는 등 외부로부터의 압력으로 인하여 표준화의 시점을 늦추고 있을 수만은 없게 되어 이에 대한 대책으로 1997년 7월의 40차 회의에서는 1998년 11월까지 MPEG-4 version 1 DIS(Draft International Standard)를 완성하고, 1999년 11월까지 MPEG-4 version 2 DIS를 완성하기로 하는 2단계 전략을 수립하였다.

MPEG-7은 영화, 케이블TV, 홈비디오 및 음반 산업의 발전에 따른 멀티미디어 콘텐츠 제작의 급증과 함께, 컴퓨터의 활용을 한 단계 높여 사용자가 멀티미디어정보를 효율적으로 제작, 편집. 검색 등의 처리를 효율적으로 도와주는 도구를 제공하는데 목적을 두고 있다. 이러한 도구에 있어 핵심이 되는 기술은, 멀티미디어 정보를 효과적으로 기술하는 메타데이타 기술로서, MPEG-7은 정지이미지, 동영상, 음악 등 멀티미디어 정보에 대한 메타데이타의 표준을 다루고 있다. 향후 전자도서관, VOD, 멀티미디어 검색 소프트웨어 등의 시장을 창출할 것으로 예상된다.

MPEG-21은 MPEG 표준화위원회가 2000년 5월 표준화 작업을 시작하였으며, 영화, 음악, 소프트웨어, 전자서적 등의 멀티미디어 컨텐츠가 제작되면서부터 최종소비자에게 이르기까지의 과정을 컨텐츠 제작자, 배급업자, 최종 사용자가 쉽게 접근할 수 있는 틀을 제공하는 표준으로서, 멀티미디어 컨텐츠의 전자상거래가 대표적인 응용이다. MPEG-21은 이러한 멀티미디어 틀에 대한 각 사용자들의 요구사항을 분석하고, 기존의 다양한 네트워크, 단말기, 전송규약, 디지털저작권 보호 등의 표준과 기술을 검토하여, 이를 보완하여 통합하는 표준을 제정하는 것을 목적으로 하고 있다. 지금까지 언급한 내용들을 요약하면 다음과 같다.

- MPEG-1 : VCD, MP3 표준
- MPEG-2 : Digital TV, DVD 표준
- MPEG-4 : 유무선 멀티미디어 표준
- MPEG-7 : Descriptor, 오디오, 비디오 컨텐츠의 검색 표준
- MPEG-21 : Multimedia Framework의 표준 또는 유비쿼터스 멀티미디어 표준

## 2. *MPEG-7* 표준

현재의 세계는 급격한 멀티미디어 정보의 증가와 활용을 경험하고 있다. 이는 다음과 같은 여러 가지 기술 발전에 기인한다. 첫째로 디지털 오디오 비쥬얼 데이터 처리 및 압축 기술의 발전과 이와 관련된 국제 표준화의 성공적인 도출이다. MPEG-2 표준은 통신, 방송, 저장 매체에 폭 넓게 적용할 것을 목표로 제정된 디지털 오디오 비디오 데이터의 압축 및 전송에 관한 국제 표준으로서 고선명 TV를 포함하는 디지털 TV 방송, 차세대 가전기기의 핵심 장치가 될 DVD의 표준으로 채택되었다. 디지털 방송환경에서는 기존의 아날로그 방송 환경에서보다 훨씬 작은 전송 대역폭에 더 많은 양질의 프로그램을 공급할 수 있게 되었으며, 이미 미국의 DirecTV 등에서 150채널 이상의 방송 채널을 갖고 서비스를 제공하고 있다. 둘째로 고성능 개인용 컴퓨터, 대용량 저장 장치의 보편화 및 World Wide Web으로 대변되는 컴퓨터 네트워크의 발전에 따라 디지털로 표현된 멀티미디어 정보의 생성, 전송, 가공이 매우 용이해졌다. WWW은 E-mail, Newsgroup등을 통한 정보 교환, HomePage를 통한 기업의 홍보, 상품의 광고 및 거래, Web 방송 등 이미 그 활용도는 일상의 주요 활동에 없어서는 안될 수단으로 자리잡고 있다. Yahoo, AltaVista 등의 Web 검색 엔진이 가장 사람들이 많이 이용하고 있는 Site가 되고 있는 점은 효율적인 정보 검색을 위한 기술과 도구의 필요성이 절실함을 말해주고 있다. 그러나 엄청난 속도로 증가하는 멀티미디어 정보 중에서 사용자가 필요로 하는 내용의 정보를 찾기 위해서는 기존의 키워드 기반의 검색은 한계에 도달한 상황이기 때문에 사용자가 원하는 정보를 내용에 기반하여 검색할 수 있는 방법이 요구되고 있다.

이러한 최근의 기술 발전 추세 및 시장 요구를 바탕으로 하여, 국제 표준화 기구인 ISO와 IEC의 연합기술위원회 산하의 MPEG(공식명칭: ISO/IEC JTC1 SC29/WG11)에서는 MPEG-7: Multimedia Content Description Interface 라는 이름으로 멀티미디어 데이터의 내용기반 검색을 위한 내용 표현 방식에 관한 국제 표준이 제시 되었다.

### a. 표준화 동향

"한 장의 그림은 천 마디의 말로도 표현이 어렵다" 든지, "보기(듣기) 전에는 믿기 어렵다" 등의 말은 멀티미디어 정보의 속성을 단적으로 드러내고 있다. 일반적으로 비디오 Archive로부터 "Terminator II 영화에서의 오토바이가 질주하는 장면"을 찾으려고 하거나, 또는 "오늘 내가 즐겨 시청한 것과 유사한 종류의 프로그램"을 방송하는 TV 채널을 찾으려고 하는 경우, 이를 키워드 기반의 표현 및 검색 기술로 구현하기란 대단히 어렵다. 이러한 문제를 해결하기 위해 내용 기반 멀티미디어 정보 검색을 효율적으로 지원하기 위한 기술을 개발하고 이를 국제 표준화하고자 하는 것이 MPEG-7이다. 즉, 기존에 표준화되었거나 지금 표준화가 진행되고 있는 MPEG-1/2/4 등은 오디오 비쥬얼 데이터의 데이터 압축을 그 목표로 하였으나, MPEG-7은 데이터 그 자체가 아닌 데이터의 내용에 대한 표현 방법을 다루는 것이다. 이를 다른 말로 "메타데이터", 또는 "Bits about bits"라고 표현하기도 한다.

MPEG-7에서는 주로 오디오 비쥬얼 정보(정지화상, 픽쳐, 그래픽, 3D 모델, 오디오, 스피치, 비디오)의 표현을 그 대상으로 하고 있으나, HTML, SGML, 또는 RDF 등이 목표로 하고 있는 텍스트 문서의 표현에 대한 것은 표준화 범위에 포함하지 않는다. 필요한 경우 이러한 문서 포맷에 관한 표준화가 적용될 수 있도록 하는 기능은 갖게 될 것이다. 그림 2.1은 MPEG-7과 관련된 정보 처리 과정과, 이 중 MPEG-7이 표준화 하고자 하는 범위를 개략적으로 보여준다.

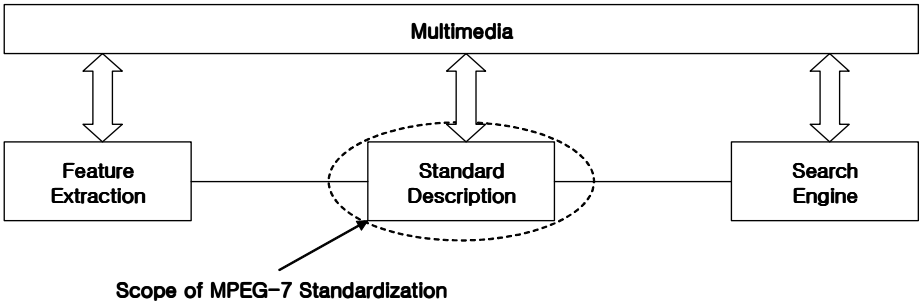



그림 2.1 MPEG-7 표준화의 범위

Fig. 2.1 Scope of MPEG-7 Standardization

이 중 특징 추출 방식 및 검색 엔진은 표준화의 대상으로 하지 않는다. 왜냐 하면,

이들은 여러 멀티미디어 정보 처리 시스템간의 호환성을 획득하는 것과는 상관이 없으며, 업체간의 경쟁이 가능한 부분으로 남겨 놓는 것이 향후 계속적인 기술 발전을 유도하기 위하여 바람직하기 때문이다. MPEG-7은 멀티미디어 정보의 특징 및 내용 표현 방식에 초점을 맞추어 다음의 사항을 표준화하고자 한다.

- 기술자(Descriptors : D)와 기술 구조(Description Scheme : DS)
- 기술 구조를 표현하기 위한 기술정의언어(Description Definition Language : DDL)
- 색인/저장/전송을 효율적으로 하기 위하여 사용될 코딩된 기술(Coded Description)

### b. 연구 분야

MPEG-7은 멀티미디어 검색의 효율을 높일 수 있도록 다양한 멀티미디어 정보에 대한 특징들을 서술하고 이를 색인화 하기 때문에 멀티미디어 콘텐츠에 대한 서술 인터페이스라고 한다. 이러한, MPEG-7 표준에 기반 한 서술자를 검색에 이용할 경우 텍스트 주석 기반과 내용 기반의 이미지 검색에 사용되는 미디어의 본질적인 특징을 효과적으로 표현할 수 있어, 하나의 MPEG-7 문서를 이용하여 다양한 질의 인터페이스를 제공하는 검색 및 브라우징 시스템을 구성할 수 있다[26-31].

따라서, MPEG-7은 멀티미디어 정보를 쉽고 정확하게 검색할 수 있도록 하는 다양한 정보들을 메타데이터 형태로 붙여 사용하게 된다.

일반적으로 MPEG-7 표준에서는 세부적인 사항을 규정하고 있는 것이 아니라 단지, 서술자에 사용되는 기본 요소만을 정의하고 있다. 따라서, 그 이외에 사용하는 알고리즘은 연구자에 따라서 다양한 적용이 가능하다고 할 수 있다. 이러한 연구 분야로써는 크게 다음과 같이 3가지 영역으로 구분할 수 있다.

- 서술자들의 조합 방법
- 특징 추출 방법
- 검색 방법

첫 번째 연구 분야로는 MPEG-7에서 정의하고 있는 다양한 서술자들 중에서 어떠한 서술자를 사용할 것인가, 서술자 정보를 어떻게 추출할 것인가, 각 서술자들을 어떻게

사용할 것인가 등에 대한 것이다, 두 번째 연구 분야로는 서술자에 해당하는 어떠한 특징을 사용할 것인가, 해당 특징들은 어떠한 알고리즘을 사용하여 추출할 것인가에 대한 것이다. 세 번째 연구 분야로는 해당 데이터를 어떠한 방법을 사용하여 검색할 것인가에 대한 것이다.

이와 같은 멀티미디어를 검색하기 위한 MPEG-7의 절차는 다음과 같다.

- Descriptor의 선정 (D)
- Description Scheme 선정 (DS)
- Description Definition Language 생성 (DDL)
- Multimedia Data Description 코딩 (DC)

### c. *Descriptor*의 기본 요소

MPEG-7 표준에서 규정하고 있는 세부 서술자들에는 다음과 같은 것이 있다. 이러한 서술자들을 조합하여 해당 멀티미디어 데이터를 표현할 수 있다.

#### (1) 기본구조

Grid Layout과 Histogram이 있다.

#### (2) 기본특징

##### *(a)* 색상 기술자

RGB, YUV, HSV, HMMD, Mono Chrome, RGB에 대한 선형 변환을 표현하는 Color Space, 이미지에 큰 영향을 미치는 색상을 표현하는 Dominant Color, 이미지를 구성하는 색들의 분포도를 표현하는 Color Histogram, 선형, 비선형 양자화, Look-up Tables 지원하는 Color Quantization, Average, Median, Intersection Histogram를 지원하는 GoF/GoP Color Histogram, 이미지를 구성하는 색상의 구조를 표현하는 Color-Structure, 색의 공간적 분포도를 표현하는 Color Layout, 색상 히스토그램의 유사도 분석하는 Harr Transformed Binary Histogram의 8가지가 있다.

##### *(b)* 질감 기술자

4개의 방향성 edge와 1개의 비방향성 edge의 분포를 표현하는 Edge Histogram, 질

감을 이용한 규칙과 조잡도의 판별하는 Texture Browsing, 양적인 질감표현인 Homogeneous Texture의 3가지가 있다.

*(c)* 형태 기술자

Visual Object를 감싸는 최소의 사각형인 Object Bounding Box, Object의 영역을 표현하는 Region-based Shape, Object의 윤곽을 표현하는 Contour-based Shape의 3가지가 있다.

*(d)* 움직임 기술자

카메라의 동작에 의한 움직임 표현인 Camera Motion, 이동 궤적의 표현인 Object Motion Trajectory, 시공간적 움직임 특성인 Parametric Object Motion, 움직임의 정도 표현인 Motion Activity의 4가지가 있다.

# B. 멀티미디어 데이터 정보 추출

## 1. *Color*

색상이란 이미지를 볼 때 즉각적으로 인지되는 시각적 특징이며 밝기, 색도, 채도의 세 가지 속성을 가지는 객체를 식별하게 하는 하나의 특징이다[32].

이러한 색상의 모델에는 RGB, CMYK, YIQ, HSV, YCbCr, HSI 등이 있다.

### a. 색상모델

#### (1) *RGB*

RGB 색상 모델은 빛의 삼원색인 빨강, 초록, 파랑을 기본색으로 하여 3차원 좌표공간으로 되어있으며 빛의 가산 혼합 원리에 의해 색을 만든다. 가산 혼합 원리란 색상을 혼합하여 더욱 밝은 색을 얻을 수 있다는 색체 원리로 삼원색을 모두 합하면 흰색, 전혀 빛이 없는 곳은 검정의 색상으로 나타난다. 명암도는 검정색에서 흰색으로 이어지는 선을 따라 표현된다. 빛을 이용한 출력 장치가 색상을 표현 할 때 이 모델을 사용한다. 즉, RGB 색상 모델은 이미지에서 다루기는 어려우나 대부분의 CRT 모니터와 컴퓨터 그래픽에서 널리 사용되고 있다.

그림 2.3은 RGB 색상 모델의 좌표계를 표현하고 있다

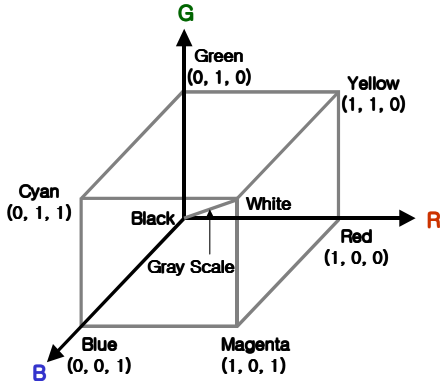



그림 2.2 RGB 색상 모델

Fig. 2.2 RGB Color Model

컬러 채널당 8비트를 가지는 24비트 컬러 그래픽스 시스템에서 빨강색은 (255, 0, 0)이고, 컬러 입방체 상에서는 (1, 0, 0)이다. 컬러 입방체 상에서의 빨강, 초록, 파랑 값의 범위는 $0 \leqq R, G, B \leqq 1$이다. RGB 색상 모델은 간단하지만, 빨강, 초록, 파랑의 컬러요소들의 상호관계가 너무 크다는 단점을 가지고 있다.

### (2) $YCbCr$

YCbCr은 색상 정보로부터 광도를 분리하는 또 하나의 색상공간이다. 광도는 Y (Luminance Signal)로 기호화되고 푸른 정도는 Cb (Chrominance Blue), 붉은 정도는 Cr (Chrominance Red)로 기호화 된다.

Y는 이미지의 밝기에 관계하고 Cb와 Cr은 이미지의 색상 및 채도에 관계한다.

사람의 눈은 밝기에 민감하고 상대적으로 색상이나 채도에 덜 민감하기 때문에 Cb와 Cr은 최소화 데이터로 압축하여 표현할 수 있다.

Y는 0~1의 값을 가지며 Cb와 C은 -0.5~0.5의 값을 가진다.

식 (2-1)은 RGB에서 YCbCr로의 변환 공식을 나타낸다. 이것은 CCIR (International Radio Consultive Committee)의 권고안 601-1이며, JPEG 압축에서 사용되는 전형적인 방법이다.

$$
\begin{aligned}
Y &= 0.29900R + 0.58700G + 0.11400B \\
Cb &= -0.16874R - 0.33126G + 0.50000B \\
Cr &= 0.50000R - 0.41869G - 0.08131B
\end{aligned}
\tag{2-1}
$$

YCbCr 색상 모델에서의 색상 분포 특성은 Y 값의 평균과 분산 값 Cb와 Cr의 값에 0.5를 감한 경우의 평균과 분산 값을 일렬로 정렬한 컬러그램을 사용한다.

### (3) *HSV*

HSV 색상 모델은 컬러를 색상, 채도, 명도로 정의하는 모델이다.

사람의 컬러 지각 능력은 세 가지 cone들의 반응들에 기초하고 있다. 이것 때문에 컬러 공간은 이런 삼중 자극값을 기반으로 한다. 따라서, 이와 유사한 컬러 공간은 HSV 컬러 공간이다.

보통 컬러라는 용어는 색상을 가리키며, 빨강색, 초록색과 같이 색상을 식별하는 역할을 한다. 따라서 색상은 다양한 파장이 복합됨으로서 받아들여지는 컬러에 대한 느낌이라 할 수 있다.

여기서, 검정색, 흰색, 회색은 색깔을 가진다고 하지만 색상은 없다.

채도란 흰색으로 희석되지 않은 색깔의 정도를 말한다. 순수한 색상에 첨가되는 자연 색깔의 양이 증가할수록 채도는 감소한다. 채도는 흔히 색깔이 얼마나 순수한가를 의미하기도 한다. 채도가 약한 색깔은 색이 바래거나 희미해져 보이며, 채도가 강한 색깔은 뚜렷하고 활기가 넘쳐 보인다. 붉은 색은 채도가 가장 높은 색인 반면, 분홍색은 채도가 떨어지는 색이다. 순수한 색깔은 100% 채도를 가지며 흰색이 전혀 섞이지 않은 것을 말한다. 흰색과 순수한 색깔을 혼합하여 0%에서 100%로 채도를 만들 수 있다.

명도란 빛이 물체에 반사되어 느껴지는 강도를 말하는 것이다. 즉, 흰색에서 검정색까지 이르는 모든 범위를 의미하며, 이러한 범위를 명암도라고 한다.

HSV 컬러 공간을 사용할 때, 어떤 컬러를 만들어 내기 위해서 몇 퍼센트의 파랑색이나 녹색이 필요한지 알 필요가 없다. 진한 빨강색을 분홍색으로 바꾸기 위해 단순히 채도를 조절하고, 어두운 것을 밝게 하려면 명도를 변경한다. RGB 모델은 색상 명을 모르는 상태에서 R, G, B 값을 혼합하여 새로운 색상을 만드는 수학적인 모델이다. 반면 HSV 모델은 만들 색상의 이름을 이용하여 색을 만드는 직관적인 모델이다.

HSV 색상 모델은 그림 2.4와 같이 원통 모양의 좌표계로 모형화 된다. 하나의 원뿔 밑에 또 하나의 원뿔이 뒤집혀져 있는 이중 원뿔 모형을 사용한다

.

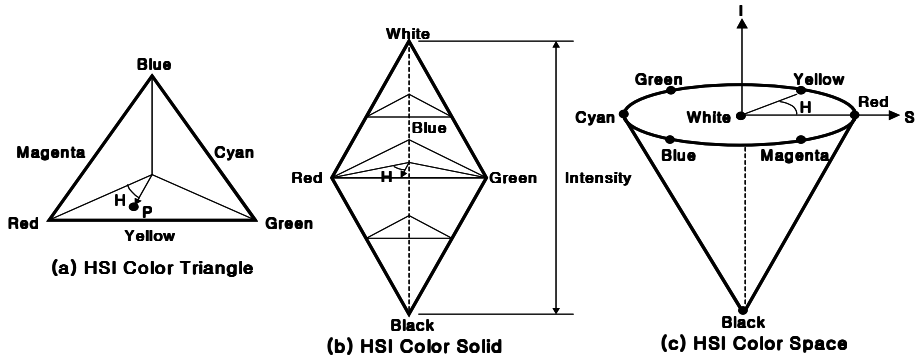



그림 2.3 HSV 컬러 모델

Fig. 2.3 HSV Color Model

색상은 0°에서 360°의 범위를 가진 각도로 표현된다. 채도는 색상의 순수도를 나타내는 것으로 0에서 1까지의 범위를 가지는 반지름에 해당한다. 명도는 평균 흑백 계조를 표현하는 것으로 Z축에 해당하는데 0일때는 검정색을, 1일때는 흰색을 나타낸다.

식 (2-2)는 RGB에서 HSV로의 변환 공식을 나타낸다.

$$
\begin{aligned}
I &= \frac{1}{3}(R+G+B) \\
S &= 1-\frac{3}{R+G+B}[min(R,G,B)] \\
H &= cos^{-1}\left(\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right)
\end{aligned}
\tag{2-2}
$$

### b. 색상정보를 이용한 연구

이러한 색상 특징을 나타내는 대표적인 방법으로는 컬러 히스토그램에 의한 방법과 컬러로 유도된 감각에 의한 방법이 있다.

컬러 히스토그램은 이미지의 하위 수준 색상 속성을 기술하는 가장 전통적인 방법이다. 컬러 히스토그램은 이미지의 색상 내용에 대한 매우 효과적인 표현이며, 컬러 히스

토그램의 계산 방법은 각 색상 값에 얼마나 많은 화소가 속하는지 계소하는 통계적 방법을 사용한다.

이미지간의 유사성은 각 화소가 갖는 색상 값의 차이에 의해 결정된다. 색상을 이용한 대부분의 연구들은 색상 값을 얻기 위하여 컬러 히스토그램 방법을 이용한다[33].

그림 2.4과 같이 R, G, B의 세 가지 색상 채널별 히스토그램을 정렬한 컬러 히스토그램은 주어진 컬러 이미지는 초기에 특별한 변환을 가하지 않으면 R, G, B 값을 보유하고 있기 때문에, 이미지를 구성하는 각 픽셀의 RGB 색상 공간에서의 분포 특성은 검색과정에서 유용하게 사용될 수 있다.

컬러 히스토그램으로서 표현되는 색상 분포는 이미지를 나타내는 객체의 각 구성 요소에 대한 내용보다는 전체적인 이미지의 내용을 나타내는데 적합하다. 또한 컬러 히스토그램 방법은 객체의 회전이나 이동 등과 같은 기하학적 변형에도 비교적 영향을 받지 않으며 구현 알고리즘도 간단하다.

그러나, 빛의 밝기와 이미지의 배경과 객체의 크기에 민감하고, 전혀 다른 종류의 이미지도 같은 색상 분포를 가질 수 있다는 단점이 있다. 또한 색상은 화소들 사이의 위치 영역에 대한 정보가 부족하기 때문에 객체의 대략적인 구분 정보만을 가진다.

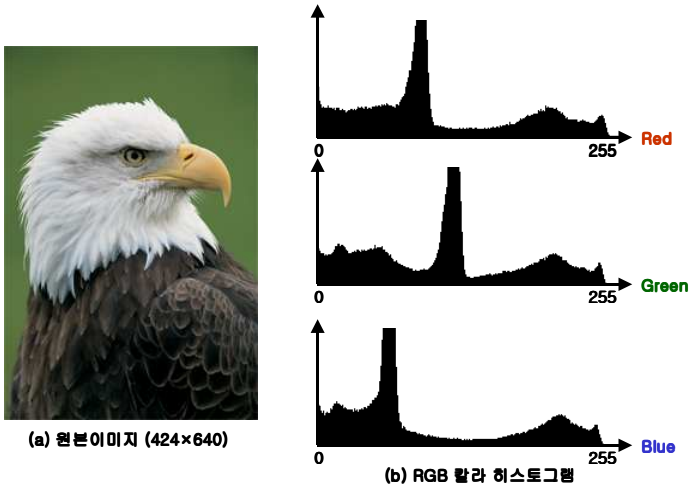



그림 2.4 칼라 히스토그램

Fig. 2.4 Color Histogram

컬러로 유도된 감각이란 색상으로 감각을 표현한 것을 말한다. 예를 들면, 빨간색이나 오렌지색은 따뜻함, 녹색이나 파란색은 차가움으로 표현하고, 비슷한 색상들이 인접

해 있으면 고요함과 평온함을 느끼고, 단독으로 우위에 있는 색상의 존재와 대조의 부재는 근심을 결정하게 하며, 어두운 노란색과 자주색은 이런 감정을 강화시키며, 조화된 일치는 기쁨과 안정의 감각을 결정한다는 등이다[34]. 이것은 하위 수준의 인지적 색상 특징을 관찰자의 직관적 관점이 잘 부합되도록 허용하는 이론으로 Itten[35]이 표현상의 수준에서 색상 조합의 효과를 알맞게 하기 위해 180가지의 기본색상으로 배열된 Itten-Runge 색상구조 정규 규칙들 (색상언어)을 정의했다. Itten의 이론에 따라 색상 조합의 규칙은 감정적 수준에서 이미지 내용의 표현을 정의하는 데 사용되고 이러한 표현은 개개인의 행동과 문화에 의존한다[36].

색상 정보 기반 인덱싱 기법은 이미지 데이터베이스의 검색에 있어서 아주 중요한 방법 이다. 이미지의 다른 정보들과 비교할 때 색상 특징은 이미지에 대한 인덱싱을 구축하는데 가장 중요한 저수준의 정보를 가지고 있다. 색상 정보는 잡음이나 다른 배경의 복잡도에 덜 민감하며 이미지 내의 물체의 이동, 회전 등의 변화에 영향을 적게 받는 통계정보를 계산하는데 사용될 수 있다. Swain과 Ballard는 히스토그램 인터섹션이라는 색상 검색 방법을 제안하였다. 이 방법에서는 각 이미지에 대해 색상 히스토그램을 구성한 후 각 히스토그램간의 공통부분을 취하는 방식으로 거리를 구하여 비교하였다. 이 방법은 계산과 정보 표현에 있어서 간결하기 때문에 많은 상황에서 유용하게 사용되고 있지만, 조명이 변화하는 경우에서는 성능이 저하되는 단점을 가지고 있다. 이러한 문제를 해결하기 위해서 Funt와 Finlayson은 일정한 색상 색인이라는 알고리즘을 개발하여 확장된 Swain 방법을 발표하였다. 이 방법은 색상 비율의 히스토그램을 비교함으로써 검색을 수행한다. Funt 등의 논문에서는 제안한 알고리즘이 일정한 조명에서는 Swain의 방법보다 조금 성능이 떨어지지만 조명이 변화하는 경우에서는 성능이 크게 향상됨을 보였다. 또한 Stricker 등[36]도 히스토그램 인터섹션 방법의 단점을 보강하여 누적 히스토그램을 이용하는 방법 및 각 색상 채널의 모멘트를 이용하여 색상 정보를 표현하고 유사도를 비교하는 방법을 제시하였다. 반면에 Mehtre 등은 색상 비교를 위해 거리 방법과 기준 색상표 방법을 개발하였다. 거리 방법은 특징으로서 세 가지의 색상 성분의 각각의 1차원 히스토그램의 평균값을 사용하였다. 두 특징 벡터간의 거리가 두 특징 사이의 유사도를 나타내기 위해 사용된다. 기준 색상표 방법에서는 한 세트의 기준 색상들이 먼저 색상표로 정의된다. 이미지내의 각 색상 픽셀들은 색상표 상의 가장 가까운 색상으로 정의된다. 새롭게 할당된 색상을 가진 픽셀들의 히스토그램이 이미지의 색상 특징으로 사용된다.

그러나 색상 히스토그램을 사용하는 방법들은 다음과 같은 단점들을 가지고 있다[38]. 첫째, 히스토그램은 차원을 줄이기 위해 양자화를 필요로 한다. 전형적인 24 비트 컬러 이미지는 224 개의 bin을 발생시키며, 이것은 최소한 2MB의 저장 공간을 필요로 한다. 따라서 bin의 수를 줄이기 위해 양자화가 필요하며, 이것에 의해 색상 정보의 손실이 일어나며 또한 어느 정도의 양자화 과정이 최적인지, 어느 정도까지 양자화를 수행해야 하는지에 대한 규칙이 없다. 둘째, 히스토그램이 만들어지는 색상공간이 검색결과에 상당한 영향을 줄 수 있으며 또한 양자화의 방법 및 양자화 정도를 결정한다. 예를 들면, 균일한 양자화는 RGB 공간에서는 유용하게 사용될 수 있는데, 그것은 RGB 공간에서 각 색상 성분의 분포가 균일하기 때문이다. 그러나 Mensell이나 LUV 공간에서는 균일한 양자화로는 불충분하다. 셋째로 색상 히스토그램을 사용하면 색상의 제외를 고려하기가 어렵다. 양자화의 정도에 따라 bin들이 나타내는 색상들이 변화하기 때문에 제외하고자 하는 색상을 나타내는 bin을 찾아내기가 어려우며 또한 제외하고자 하는 색상과 다른 색상들이 제외되는 경우가 발생할 수 있다. 마지막으로 히스토그램은 이미지의 전체적인 색상 정보만을 추출한다. 즉, 어떠한 공간적인 정보도 이용되지 않는다는 단점을 가진다. 본 논문에서는 색상 정보를 이용한 이미지 검색 기법 중에서 대표적인 Swain의 히스토그램 인터섹션 방법 및 Stricker 등의 누적 히스토그램 방법, 색상 채널의 모멘트를 이용하는 방법을 자세하게 분석한다.

#### (1) 히스토그램 인터섹션

Swain이 제안한 방법으로 색상 이미지의 경우 밝기 이미지에 비해 정보량이 많음을 이용하여 이미지의 통계학적 특성을 이용한 3차원 히스토그램을 생성한 후에 히스토그램 공간상에서 모델과 입력이미지의 정합을 시도하는 알고리즘이다. 이 알고리즘은 간단하고 적용하기 쉬우면서도 물체의 회전, 이동 및 어느 정도의 크기 변화, 그리고 이미지 획득 시 각도의 변화 등에 강인하게 동작한다는 장점을 가진다.

이 방법에서는 색상을 나타내는 색상 공간으로 식 (2-3)과 같은 RGB로부터 얻어지는 보색축을 이용하였다.

$$\begin{aligned} rg &= R - G \\ by &= 2B - R - G \\ wb &= R + G + B \end{aligned} \tag{2-3}$$

이미지의 색상 히스토그램을 구하기 위해서 위 세 축의 구간을 나누어야 하는데, 여기서는 세기축인 $wb$가 광원으로부터의 거리 및 밝기의 변화 등에 민감하기 때문에 영향을 줄이기 위해 구간을 더 크게 하였다. $rg$, $by$ 는 16개의 구간으로, wb는 8개의 구간으로 나누어 히스토그램은 총 2048 (16×16×8)개의 그룹을 가진다. 히스토그램이 구해지면 두 이미지의 색상 히스토그램은 히스토그램 인터섹션 방법에 의해 비교된다. n 개의 그룹을 가진 두 이미지 I , M 이 주어졌을 때 히스토그램 인터섹션은 식 (2-4)와 같이 정의된다.

$$\sum_{j=1}^{n} min\,(I_j, M_j) \tag{2-4}$$

위 식의 결과는 두 이미지 내에서 같은 색상 값을 가지고 있는 화소의 수를 나타낸다. 이 결과는 이미지의 크기에 따라 값이 변하므로 식 (2-5)과 같이 정규화 시켜서 사용한다.

$$H(I, M) = \frac{\sum_{j=1}^{n} min\,(I_j, M_j)}{\sum_{j=1}^{n} M_j} \tag{2-5}$$

만일 두 이미지가 동일하다면 위의 결과는 1이 되고 두 이미지의 색상이 전혀 다른 이 방법은 구현이 간단하고 비교적 좋은 성능을 보이지만, 조명의 변화에 민감하고 각기 다른 그룹들간의 시각적 유사성을 전혀 고려하지 않고 각각의 그룹을 독립적으로 간주하고 비교를 했다는 점에서 문제점을 가지고 있다. 또한 어떠한 공간적인 정보도 이용하지 않고 이미지 전체에 대한 색상 특징을 구하기 때문에 잘못된 결과가 나타날 수 있으며, 특히 데이터베이스의 크기가 커질수록 그 문제가 심각해진다.

#### (2) 누적 히스토그램을 이용한 기법

Stricker 등[56]은 Swain의 히스토그램 인터섹션 방법을 개선시킨 누적 히스토그램을 이용한 방법을 제시하였다. 그림 2.6에서와 같이 픽셀의 수가 N 개인 3개의 히스토그램이 있다고 할 때 보통 히스토그램 사이의 유사성 정도를 측정하는데 많이 사용되는 $L_1$ 거리를 사용하여 거리를 계산하면 식 (2-6)와 같다.

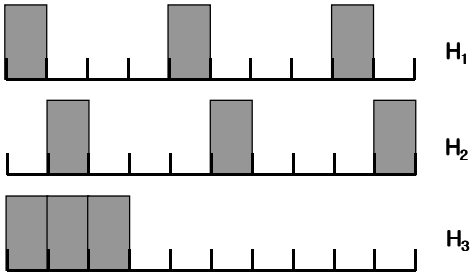



그림 2.5 $L_1$ 거리에 의해 시각적 유사성을 갖지 않는 3개의 히스토그램

Fig. 2.5 Three Color Histogram whose Perceptual Similarities do not Correspond to their $L_1$ Distances

$$L_1 = d_{L_1}(H, I) = \sum_{l=1}^{n} |h_{cl} - i_{cl}|$$
$$d_{L_1}(H_1, H_2) = 2N > d_{L_1}(H_1, H_3) = d_{L_1}(H_2, H_3) = 1.33N \tag{2-6}$$

그림 2.6의 히스토그램을 보면 $H_1$과 $H_2$가 더 유사하지만 실제 계산 결과는 다르게 나타남을 알 수 있다. 따라서 Stricker는 색상 히스토그램 H(M)에 대해 누적 히스토그램인 $\hat{H}(M) = (\widetilde{h_{c1}}, \widetilde{h_{c2}}, \ldots \widetilde{h_{cn}})$을 정의하였다. 여기서, $\widetilde{h_{cj}} = \sum_{c_i \le c_j} h_{ci}$이다. 그림 2.6의 히스토그램들에 대해 누적 히스토그램을 구하면 그림 2.7와 같다. 다시 누적 히스토그램에 대해 $L_1$ 거리를 이용하여 유사도 정도를 측정하면 식 (2-7)과 같다.

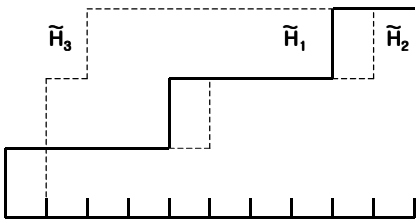



그림 2.6 그림 2.6에 주어진 히스토그램의 누적 히스토그램

Fig. 2.6 Cumulative Histograms of the Histograms Displayed in Fig. 2.6

$$d_{L_1}(H_2, H_3) = 4N > d_{L_1}(H_1, H_3) = 3N > d_{L_1}(H_1, H_2) = N \tag{2-7}$$

위 결과를 보면 기존의 히스토그램을 이용한 방법보다 더 좋은 결과를 보임을 알 수 있다. 즉, 누적 히스토그램을 사용하게 되면 히스토그램을 구성할 때 잡음에 의해 유사한 색상이 다른 그룹에 속하게 되더라도 유사도 측정 시 어느 정도 반영이 됨을 알 수 있다.

### (3) 색상 모멘트를 이용한 기법

색상 히스토그램이나 누적 히스토그램은 검색에 적용 시 색인 과정에서 양자화가 필요하며 최적의 양자화 과정을 찾아내는 것은 아주 어려운 일이다. 따라서 인덱스 생성 과정에서 파라미터에 의존하지 않는 방법이 요구되며 이 중의 하나가 색상 분포의 특징을 이용하는 것이다. Stricker 등은 색상 검색을 위한 색상 분포의 특징으로서 각 색상 채널의 모멘트를 이용하는 방법을 제안하였다[56]. 색상 모멘트를 이용한 방법을 보면 이 방법은 확률 분포가 그것의 중심 모멘트들에 의해 유일하게 표현될 수 있다는 사실을 바탕으로 하고 있다. 이미지의 색상 분포 즉 색상 히스토그램은 확률 분포로 해석될 수 있으며 따라서 색상 분포는 모멘트에 의해 표현될 수 있다. 여기서는 색상 정보를 위한 인덱스로서 각 색상 채널의 1, 2, 3차 중심 모멘트를 이용하였다. j 번째 이미지 픽셀에서 i번째 색상 채널의 값이 $p_{ij}$ 라면, 이 채널과 관계된 인덱스는 식 (2-8)과 같이 계산된다.

$$
\begin{aligned}
E_i &= \frac{1}{N}\sum_{j=1}^{n} p_{ij} \\
\sigma_i &= \left(\frac{1}{N}\sum_{j=1}^{n}(p_{ij}-E_i)^2\right)^{\frac{1}{2}} \\
s_i &= \left(\frac{1}{N}\sum_{j=1}^{n}(p_{ij}-E_i)^3\right)^{\frac{1}{3}}
\end{aligned}
\tag{2-8}
$$

여기서, N 은 이미지의 픽셀 총 수이다.

r 개의 색상 채널을 가지는 두 이미지 I, H 의 각각의 색상 정보 인덱스를 $E_i$, $F_i$, $\sigma_i$, $\zeta_i$ 그리고 $s_i$, $t_i$ 라고 하면 두 이미지의 색상 유사도는 식 (2-9)와 같이 얻어진다.

$$d_{mom}(H, I) = \sum_{i=1}^{r}(w_{i1}|E_i - F_i| + w_{i2}|\sigma_i - \zeta_i| + w_{i3}|s_i - t_i|) \quad (2\text{-}9)$$

여기서 $w_{kl} \geq 0\,(1 \leq l, k \leq 3)$는 사용자가 정의한 계수이다. 계수를 조정함으로써 적용 환경에 맞게 유사도 함수를 정할 수 있다.

이 방법은 이미지의 색상 정보를 적은 수의 인덱스를 이용하여 나타낼 수 있는 장점을 가지며 성능에 있어서도 히스토그램을 이용한 방법 이상의 결과를 보여주었다. 그러나 각 색상 채널사이의 상관관계에 대한 어떠한 정보도 이용하지 않으므로 그 성능에 있어서 한계가 있다.

색상 정보를 이용한 대표적인 연구로는 색상 모델의 선택이 어떻게 히스토그램에 매칭되는가를 분석한 연구[39], 이미지의 지역적 색상 속성을 특성화하기 위해 컬러 히스토그램을 집합을 제안한 연구[40], VisualSEEK 시스템에서 컬러 이미지의 위치 영역에 사용한 색상 집합에 대한 연구[41] 등이 있다.

## 2. *Texture*

질감은 특정 패턴이 유사적 규칙성을 보이는 지역적인 영역의 특징을 말한다. 이러한 질감 특징의 유사성에 기초하여 검색하는 기법은 다양한 상황에 적용될 수는 없지만 유사한 색상을 가진 이미지들(예를 들면 하늘과 바다, 또는 나뭇잎과 잔디 등)의 영역을 구분해내는데 유용하게 사용될 수 있다. 또한, 색상과 함께 질감은 특징을 식별하는데 강력한 기능을 지원한다[42].

질감은 실세계 이미지들의 내용을 묘사하기 위하여 사용되고, 인간 시각에 있어 중요한 요소이다. 색상 특징이 색상 공간 요소의 명암도 분포에 의해 결정되듯이 질감 특징은 연관된 화소들 간의 상호 관계를 나타내는 이미지 명암도의 공간적인 분포에 의하여 결정된다[43]. 질감 정보를 사용한 검색 방법은 주로 이미지 전체가 동일한 질감 특징을 갖는 이미지나 전체가 몇 개의 두드러진 질감 특징을 갖는 이미지 (직물 구조, 벽지의 무늬, 생체 조직, 세포의 모습 등)를 찍은 사진 이미지 등을 대상으로 연구되고 있다. 따라서, 다양한 객체를 담은 실제 이미지에서 질감을 특징으로 추출하는 것은 아직까지 한계가 있고, 사용자가 시각적인 예를 제시하는 경우에는 적용하기 어렵다[44].

질감을 인식하고 분별하는 방식은 크게 나누어 통계적인 방법, 구조적 분석 방법, 공간주파수 영역에서의 스펙트럼 분석 방법 등이 있다[45, 46].

통계적 방법으로는 자체 상관 함수, 주파수 공간의 역률 밀도 함수, 단위 면적당 윤곽점수, 수학적 형태학 등 다양하다[47, 48].

구조적 분석 방법은 질감 구조가 규칙적인 경우에 사용되는데 문법 구조에 의한 모델을 들 수 있다[49].

공간주파수는 화소의 변화율을 의미하는 것으로 이미지 자체를 공간주파수 영역으로 변환할 수 있다. 공간주파수 영역에서의 스펙트럼 분석을 통해 거침의 정도, 방향성 등을 얻을 수 있다[50].

이와 같은 질감에 대한 연구 중 다양한 기법들이 질감 정보의 유사성을 계산하기 위해서 제안되었는데 그 중 많은 방법들이 질의 자료와 저장된 이미지들로부터 계산된 2차 통계 자료를 비교하여 검색을 수행한다. 이러한 방법으로 1978년 Tamura 등은 대비, Coarseness, 방향성, 규칙성, 선형성 및 거친 정도 등을 이용하여 질감 정보를 표현하였으며[51], 1996년 Liu와 Picard는 주기성, 방향성 그리고 Randomness 등을 이용하여 질감 정보를 계산하였다[52]. 이들과는 다른 방법으로 Manjunath와 Ma에 의한 Gabor 필터를 이용하는 방법, Kaplan 등의 프랙탈을 이용하는 방법[53] 등이 있다. 질감 정보를 이용한 검색 시 질의는 팔레트로부터 원하는 텍스쳐를 선택하거나 또는 샘플 이미지를 제시함으로서 검색이 가능하다. 개발된 시스템으로는 1998년 Ma와 Manjunath에 발표된 "Texture Thesaurus"가 있다.

## 3. *Shape*

형태란 이미지의 영역, 에지, 코너, 점 등의 요소들을 의미하는 것으로 모멘트, 에지 이미지 상호관계로서 그 특징을 나타낸다[54]. 영역이란 전처리 과정에서 얻은 윤곽점, 질감, 운동, 색상 등에 대해 동일한 특징을 가지는 이미지의 일부분을 일컫는다. 에지는 윤곽점을 구하는 과정에서 나타나거나 여러 가지 구획 방법에 의하여 얻어지는 것으로서, 직선이 될 수도 있고 특별한 함수의 식으로 나타낼 수도 있다[55, 56].

또한, 형태 정보는 객체를 이루는 윤곽선이 객체가 놓여 있는 위치나 크기 등에 영향을 받지 않는다는 장점이 있지만 객체의 형태 유사성에 대한 주관적인 해석이 가능하므로 객체의 형태를 정확히 정의한다는 것은 어렵다. 그리고 객체의 윤곽선은 객체의 크기 변환이나 회전에 따라 민감하기 때문에 형태의 추출 및 모델링이 어렵다[57-59]. 따라서, 형태 정보를 표현하기 위해서는 일반적으로 객체의 크기 변화나 회전에 무관한 객관적인 속성 값을 정의하여야 한다[60]. 그리고 형태 정보에 의한 검색은 모양 데

이터가 미리 등록되어야 하므로 응용 분야가 국한되어 있다.

이러한 물체의 형태 정보는 시각적으로 이미지의 매우 다양한 정보를 가지고 있으며 형태 정보 기반 검색 기법은 구현하기 어렵지만 형태 정보가 물체를 가장 잘 표현하므로 내용 기반 검색 시스템에서 가장 중요한 부분이다. 형태 표현 기법은 그림 2.8과 같이 크게 2가지 부류로 나뉘어 질 수 있다. 하나는 윤곽선 기반 기법이고 다른 하나는 영역 기반 기법이다. 윤곽선 기반 기법은 이미지 내의 물체의 외곽선에 기반하여 특징을 추출하는 방법으로서 체인 코드, 퓨리어 묘사, UNL 퓨리어 특징, 물체의 윤곽선의 반경, 특징점을 이용한 기법 등 다양한 기법들이 존재한다. 영역 기반 기법은 물체 영역의 기하학적인 특징을 추출하는 방법으로서 면적, 신장율, 오일러 수, 불변 모멘트, 저나이크 다항식을 이용한 저나이크 모멘트를 비롯하여 많은 방법들이 있다.

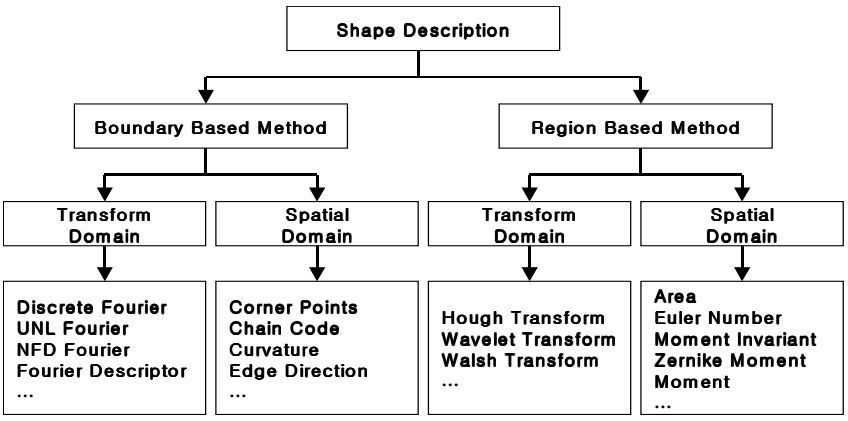



그림 2.7 형태 기반 검색 기법의 분류

Fig. 2.7 Classification of Retrieval Techniques based on Shape Information

# III. 검색 엔진

## A. 이미지 정보의 검색 방법

이미지 정보 검색의 목적은 데이터베이스로부터 질의와 관련있는 이미지들을 찾기 위한 것으로 전통적 정보 검색 방법을 시각 매체를 포함하는 것으로 확장한 것이다.

멀티미디어 데이터 중에서도 이미지는 시각적이며 공간적인 정보로 멀티미디어에서 가장 많이 사용되고 있으며, 저장과 출력 방식의 다양성, 대용량, 공간 관계 표현 및 비형적이라는 특징을 가지고 있다[61].

이미지 정보 검색은 기존의 전통적인 검색 방법인 문자 정보 검색 방법을 이용하여 시각적 정보를 포함하도록 확장한 것이다. 이미지 정보를 효율적으로 질의 및 검색하기 위한 방법은 문자 기반 검색, 의미론적 기반 검색, 내용 기반 검색의 세 가지로 분류된다[62].

### 1. 문자 기반 검색

문자 기반 검색 (Text-based Retrieval) 방법은 검색 대당이 되는 모든 이미지 데이터들에 대해서 사람이 수작업으로 의미 정보를 부여하고, 그 의미 정보를 이용하여 원하는 이미지를 검색하는 방법이다[64, 65]. 1세대 이미지 검색 시스템들은 문자 속성을 통해 이미지를 검색하였다. 예를 들면, "숲에서 빨간 우선을 들고 서있는 소녀"와 같은 형식의 검색어를 사용하여 검색하는 것이다. 문자 기반 검색은 이미지의 특징을 키워드로 묘사하는 것이다. 키워드는 이미지나 동이미지 안에 들어 있는 중요한 개체나 내용에 종속적인 용어 (숲, 우산, 소녀 등)를 사용하는 것이 검색하기에 적합하다. 이때 검색은 SQL이나 전체 문자검색과 같은 일반적인 질의 언어를 사용하여 이루어진다[66].

이 방법은 키워드의 자동적인 생성이나 다양한 종류의 이미지들을 구분하기 위한 특징 추출이 어렵고, 사람이 색인 작업을 하기 때문에 개인의 주관성이 포함되어 색인 정보의 일관성과 객관성이 보장되지 않는 단점이 있다.

그러나, 웹 이미지 검색에서의 문자 기반 검색은 웹 문서의 태그 정보를 분석하여 이미지에 대한 키워드로 삼아 검색 정보로 활용한다. 이 방법은 키워드의 자동 생성이 가능한 장점은 있지만, 키워드와 이미지 내용과의 일치성은 보장할 수 없다는 단점이

있다.

기존의 문자 기반 웹 이미지 검색 시스템으로는 엠파스 검색 엔진[68]과 구글 검색 엔진[69] 등이 있다. 엠파스 이미지 검색 엔진은 검색 결과를 이미지 URL 형태로 보여주고, 구글 이미지 검색 엔진은 검색 결과를 이미지와 이미지 URL 형태로 보여준다.

## 2. 의미론적 기반 검색

의미론적 기반 검색 (Semantic-based Retrieval)은 의미를 갖는 키워드 조합을 사용하여 검색하는 방법이다. 이 방법은 추상화된 특징들을 사용하는데, 이러한 특징들을 적용 분야와 상당히 일정한 관계를 갖고 있으므로 분야가 제한된 환경에서만 사용할 수 있다. 이미지의 의미체계는 전통적 관점의 의미론과 추론에 근거한 해석으로 나눌 수 있다. 전통적 관점은 객체, 역할, 행동, 사건 등과 같은 원시적 의미론의 식별로 구성되고, 추론에 근거한 해석은 전역적 묘사와 탬플릿 매칭에 의해 이루어진다. 예를 들면, 원시적 의미론에 근거한 검색 시스템에는 사람의 얼굴 식별을 위한 눈, 입, 코, 이마가 있는 얼굴의 요소를 특징화하여 모델화하거나 전역적인 얼굴형을 데이터화 한 안면 데이터베이스가 있고 해석을 포함하는 검색 시스템을 안면 표현을 비교하기 위해 추론을 사용하는 안면 데이터베이스 시스템이나 특정 병리학에 대해 탐색하는 의학적 데이터베이스, 혹은 지리적 지도에서 탐색과 네비게이션에 대해 용이하게 하는 공간적 추론을 가진 지리정보 시스템 그리고 장면 이해를 위한 야외 장면 동이미지 데이터베이스 등이 있다.

## 3. 내용 기반 검색 시스템

### a. 내용 기반 이미지 검색

내용 기반 이미지 검색 시스템(Content-based Retrieval)은 1992년 Kato에 의해 제시되었으며 색상 및 형태 특징을 이용하여 데이터베이스 상에서 자동적으로 이미지를 검색할 수 있는 시스템을 제안하였다.

즉, 기존의 문자 기반 이미지 검색 방법들의 단점을 극복하기 위하여 이미지 데이터에서 색상, 질감, 형태 등의 표현 요소들을 자동으로 추출하여, 이를 검색에 이용하는 방법이다[93]. 이 방법은 이미지를 표현하는 문자 정보에 의존하는 것이 아니라 이미지에서 직접 추출한 시각적인 특성을 이용하기 때문에 키워드와 관계없이 이미지의 자체

정보인 시각적인 예로서 사용자가 원하는 이미지를 질의 및 검색할 수 있다. 즉, 색상, 질감, 형태 등의 내용 표현 요소들을 통하여 얻어진 특징 값을 사용하여 질이 이미지와 검색 대상 이미지 사이의 유사도를 계산하고 유사도 값에 의한 검색을 수행한다. 그러므로 다양한 이미지에 대하여 효율적으로 적응할 수 있으며, 다양한 질의 방법을 통한 검색을 제공해주는 장점이 있다. 이 경우, 효율적인 이미지 검색을 위해서는 무엇보다 각 이미지를 대표하는 내용 표현 요소들의 효과적인 특징 추출이 중요하다.

내용 기반 이미지 검색은 질의 자료로 이미지 자체 또는 이미지에 대한 관련된 정보를 제시하면 질의 자료를 분석하여 특징 정보를 추출한 후 이미지 데이터베이스 상에서 가장 유사한 특징을 가지는 이미지를 제시하는 기술이다. 그림 3.1은 내용 기반 이미지 검색 시스템의 일반적인 구조를 나타낸 것이다. 그림에서 보는 바와 같이 이미지 검색 시스템은 크게 이미지의 입력 과정, 질의 과정 그리고 검색 과정으로 구성되어 있다.

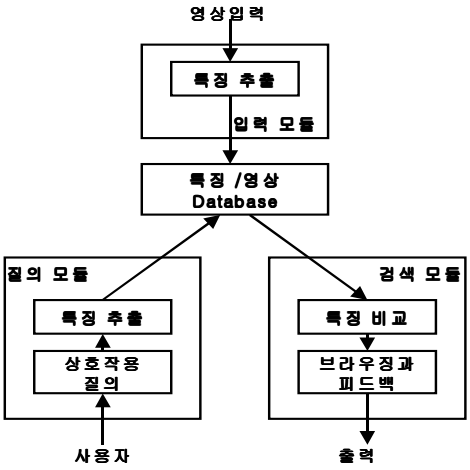



그림 3.1. 내용 기반 검색 시스템의 구조

Fig. 3.1 Structure of the Content-based Retrieval System

이미지의 입력 과정에서는 입력된 이미지를 가장 잘 표현할 수 있는 특징을 추출해서 색인화하여 이미지 데이터베이스를 구축하는 단계이고, 이미지의 질의 과정은 사용자가 찾고자 하는 이미지를 질의 자료로 제시하는 특징을 추출하는 과정이다. 이 단계에서는 질의 이미지에 대해 데이터베이스의 구축 시 특징을 추출할 때와 동일한 방법으로 같은 종류의 특징을 추출해야 한다. 마지막으로 검색 과정은 질의 이미지의 특징

과 데이터베이스 상의 이미지들의 특징들을 비교하여 가장 유사한 특징을 가지는 이미지들을 결과로 출력하는 단계이다.

검색 시스템에 있어서 가장 중요한 요소는 이미지를 어떠한 특징을 가지고 표현할 것인가와 추출된 특징들을 어떠한 방법으로 비교할 것인가 하는 문제이다. 검색에서 사용되는 특징은 원시적이거나 또는 구문적이며, 특징 추출은 전적으로 자동적으로 수행되어져야 한다. 키워드가 이미지의 내용을 나타낸다고 하더라도 수동적으로 할당된 키워드에 의한 검색은 내용 기반 이미지 검색이라 하지 않는다.

기존의 내용 기반 이미지 검색 시스템으로는 QBIC, VisualSEEK, Chabot, Photobook, SIMPLIcity 등이 있다. 표 3.1은 내용기반 이미지 검색시스템을 비교한 것이다.

표 3.1 내용기반 이미지 검색 시스템의 비교
Table 3.1 Comparison of Content-Based Image Retrieval System

| 시스템 | 검색방법 | 특징정보 추출방식 | 단점 |
|---|---|---|---|
| QBIC | 사용자 스케치 검색, 샘플 이미지에 의한 유사성 검색, 색상/모양/질감 기반 검색 | 점(point)으로 특징 벡터 표현, R*-트리를 이용한 인덱싱 | 통합되지 못한 검색 인터페이스, 의미 기반 정보 검색 불가능, 지식기반 검색 불가능 |
| VisualSEEK | 평균 색상 및 공간 검색, 시각적 특징에 의한 검색 | 영역적 평균 색상 히스토그램 추출과 자동화된 영역 추출 | 색상정보 검색의 큰 비중, 비효율적인 공간 관계 연산 시간 |
| Chabot | 텍스트 및 색상 기반 검색 | 평균 색상 히스토그램과 수동적인 키워드 입력 | 복잡한 질의 인터페이스, 내용정보간의 관계성 부족 |
| Photobook | 객체 기반 검색, 색상 및 질감, 형태에 기반한 검색 | 객체(윤곽선) 정보 추출, 색상/질감/형태 특징 추출 | 통합되지 않은 검색 인터페이스, 의미 기반 검색 불가능, 객체간의 관계성 부족 |
| SIMPLIcity | 영역 매칭 검색 | 영역 세그멘테이션 분할, 의미적 분류, 영역 특징 추출 | 지식기반 검색 불가능, 영역적 연산 시간 소요 |

### b. 질의 이미지의 특징

내용 기반 이미지 검색 시스템에서는 이미지의 내용을 바탕으로 특징을 추출하여 검색을 수행하기 때문에 이미지의 어떠한 특징을 어떠한 방식으로 추출할 것이지는 매우 중요한 문제이다. 일반적으로 데이터베이스 상에서 이미지를 검색한다는 것은 어떠한

색채의 비율을 포함하는 이미지, 특별한 유형의 물체나 장면을 포함하는 이미지를 찾는다거나, 특별한 분위기를 나타내는 이미지, 또는 특별한 질감이나 패턴을 포함하는 이미지를 찾아내는 것을 말한다. 이러한 목적의 이미지 검색을 위해서 다음과 같이 이미지의 특징들이 3가지의 레벨로 분류될 수 있다. 레벨 1은 색상, 질감, 형태 및 이미지 내 구성 요소들의 공간적인 배치 등과 같은 원시적인 특징들에 의한 검색을 말한다. 이러한 수준의 검색은 객관적이고 이미지 자체로부터 바로 추출할 수 있는 특징들을 이용하기 때문에 추가로 외적인 정보를 필요로 하지 않는다. 이러한 검색 방식은 상표 등록 시 검색이나, 의류 등의 색상 매칭 등에서 적절하게 사용될 수 있다. 레벨 2는 이미지 내에 존재하는 물체의 인식에 대한 논리적인 추론을 바탕으로 한 논리적인 특징을 이용한 검색을 말한다. 예를 들면, 주어진 것과 같은 종류의 물체의 검색, 즉, 2층 버스의 이미지를 검색할 때 2층 버스와 같은 질의 내용이 이에 해당된다. 이러한 검색을 수행하기 위해서는 물체에 대한 추가적인 정보가 미리 구축이 되어야 한다. 레벨 3은 이미지내의 물체나 장면에 대한 의미와 목적과 같은 고수준의 추론을 포함하여 추상적인(Abstract) 정보를 바탕으로 검색을 수행한다. 예로는, 감정적인 또는 종교적인 의미를 가진 이미지의 검색(고생을 나타내는 이미지 등)과 같은 것이 있다. 검색이 성공적으로 되기 위해서는 이미지의 내용과 그것이 묘사하는 추상적인 개념을 연관시키는 추론이 적절하게 수행되어야 한다. 이러한 검색은 예술 작품 데이터베이스나 신문 자료 등에서의 검색에서 사용될 수 있다.

현재 대부분의 검색 시스템은 레벨 1의 질의 내용을 바탕으로 검색을 하고 있다. 이미지를 표현하는 특징들은 이미지에서 특징을 추출하는 방법에 따라 크게 전역적인 특징과 지역적인 특징으로 나눌 수 있다. 전역적인 특징은 이미지 전체를 동시에 고려하여 추출된 특징을 말하며 색상 정보로 사용되는 색상 히스토그램, 물체의 형태 정보로 사용되는 물체의 기울기 정보, 저나이크 모멘트 등이 대표적이다. 지역적인 특징은 이미지의 각 일부분에서 추출된 특징들로서 주로 텍스쳐 정보 추출 시 이용되며 특징점들에서 얻어진 주파수 성분값, 대비, 방향성, Coarseness, Roughness 등이 이에 해당된다. 두 이미지간에 특징을 비교할 때 같은 지역에 해당하는 특징사이의 거리를 비교함으로써 이미지의 지역적인 특성을 고려할 수 있다.

### c. 내용 기반 이미지 검색 시스템의 응용 분야

내용 기반 검색 시스템의 응용 분야는 상당히 다양하다. 각 분야별로 활용 가능한

응용을 살펴보면 먼저 범죄예방 분야로 범죄자의 얼굴, 지문, 발자국 등의 인식에 사용할 수 있으며, 군사 분야에서는 레이더 사진으로부터 적의 무기 확인, 위성사진으로부터 목표물 확인 등에 적용 가능하다. 또한 지적 재산권 분야로는 상표 이미지 등록 및 불법 복제 보호 등에 응용할 수 있고, 의학 분야에서는 X선 사진, 초음파 사진에 시각적 정보를 이용하여 병의 진단 및 환자의 상태를 파악할 수 있다. 패션 디자인, 그래픽 디자인, 산업 디자인 분야에서 고객에서 정보를 제공하거나 재창작 과정에서 요구되는 2, 3차원의 스케치나 모델링에 이용되며, 문화적 유산을 보관하는 박물관이나 미술관, 신문, 잡지 등의 기사나 광고에 적합한 그림을 찾기 위해 사용될 수 있고, 지리 정보 시스템 교육 등 그 활용 분야가 매우 다양하다. 최근에는 인터넷의 활성화로 다양한 분야의 전문가들에게 원하는 이미지를 제공하기 위한 검색 엔진이 서비스되고 있으며, 대표적인 것으로 Excalibur 검색 시스템에 기반을 둔 Yahoo의 Image Sufer와 Virage의 검색 시스템을 이용한 AltaVista의 AV Photo Finder가 있다.

# IV. 시스템 설계

## A. 기술자 설계

본 논문에서는 제안한 전체 시스템은 사용자가 원하는 기술자들을 모두 테스트되도록 구현되었지만 실제 구현한 시스템에서는 MPEG-7에서 기술하고 있는 기술자들 중에서 칼라 기술자의 조합만을 사용하고자 한다. 본 논문에서 테스트하고자하는 이미지 데이터가 일반적인 웹상에 존재하는 이미지이므로 형태나 움직임 기술자들은 제외하였다.

### 1. 내용 기반 검색 시스템 설계

본 절에서는 색상 분할 기법을 이용하여 이미지의 색상 정보를 추출한 후 각각의 이미지에 대해 얻어진 색상들을 비교하는 방법을 사용하며, 이를 이용한 색상 기반 검색 기법을 제안하고자 한다. 최근에 색상 기반 검색 기법이 기존의 히스토그램을 사용하는 방법에서 색상 분할을 이용하는 방법으로 많이 전환되었다. 색상 이미지의 분할은 현재 많은 연구가 수행되고 있는 분야이며 특히 인간의 시각 처리의 기본 요소라는 점에서 관심을 끌고 있다. 인간의 시각 처리의 초기 단계에서는 저수준 처리로서 이미지에서의 작은 영역내의 세밀한 부분을 평균하여 이미지의 대표적인 색상들을 인지한다. 이렇게 추출된 색상 정보들이 시각 처리의 다음 단계에서 뇌에 의해 물체로 분류 또는 인식된다. 본 연구에서도 이러한 개념을 이용하여 이미지를 HSV 공간에서 색상 분할하여 대표 색상 정보를 추출한 후 추출된 색상들을 비교하는 방법을 제안하여 색상 기반 이미지 검색을 수행할 수 있는 기법을 제시하고자 한다.

#### a. *HSV* 색상 변환

색은 명도, 채도, 색상이라는 3가지 특성으로 정의된다. 이러한 특성을 이용하여 인간이 느끼는 색상들을 표현하는 것이 HSV (Hue, Saturation, Value or Lightness) 이다. Hue는 빨강, 노랑, 파랑 등 인간이 감지하는 색감을 나타내며, 이것은 빛의 파장에 의해 결정된다. Saturation은 채도를 나타내며, 단일 색상의 빛에 흰색의 빛을 가함으로써 생기는 변화를 나타낸다. 즉, 같은 빨강의 색상을 지니고 있다하더라도 거의 무채

색에 가까운 빨강으로부터 아주 선명한 빨강색까지 존재하게 되는데 이는 Saturation이 다르기 때문이다. 이것은 한 명도에서 어떤 영역의 색채의 풍부함의 정도를 나타내는 단위이다. 그리고 Value는 명도를 나타내며, 색깔의 밝기, 즉 그 색이 반사되거나 투과되어 우리 눈에 도달하는 양의 차이를 기술하는 양이다. 예를 들어, 무채색은 RGB 삼원색이 같은 비율로 섞여 있는 경우에 해당하며 이렇게 섞여 있는 빛이 우리 눈에 들어오는 빛의 양에 의해서 흰색부터 점점 회색이 되다가 빛이 들어오지 않게 되면 결국 검정색으로 된다.

이러한 HSV 색상 공간은 그림 4.1과 같으며 Hue는 0° ~360° 의 범위를 가지고 Saturation과 Value는 0과 1사이의 값을 갖는다.

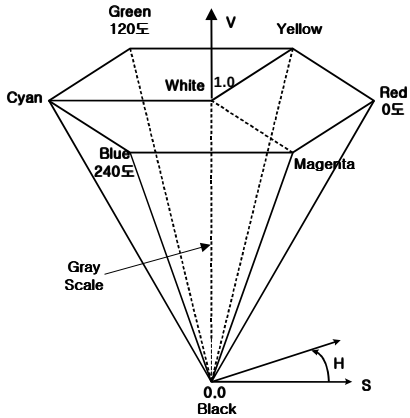



그림 4.1 HSV 칼라 공간

Fig. 4.1 HSV Color Space

색상을 표현하는 가장 일반적인 RGB 공간에서 HSV 공간으로의 변환은 식 (4-1)에 의해서 이루어진다.

$$H = \begin{cases} H_1 = cos^{-1}\left( \dfrac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right) & , if\ B \le G \\ 360^\circ - H_1 & , otherwise \end{cases}$$
$$S = \frac{max(R, G, B) - min(R, G, B)}{max(R, G, B)} \qquad (4\text{-}1)$$
$$V = \frac{max(R, G, B)}{255}$$

여기서 R, G, B는 각각 0~255 사이의 값을 가지는 빨강, 초록, 파란색 성분이다.

### b. *HSV* 칼라 색인

RGB 칼라를 HSV 칼라로 변환하면 이를 데이터베이스에 저장하기 위해서는 해당 데이터의 크기를 줄여야 한다. 본 논문에서는 HSV 칼라를 색인화 하고 이를 코드화 함으로써 데이터의 크기를 줄이고자 하였다. 그림 4.2는 전체 알고리즘을 표현한 것이다.

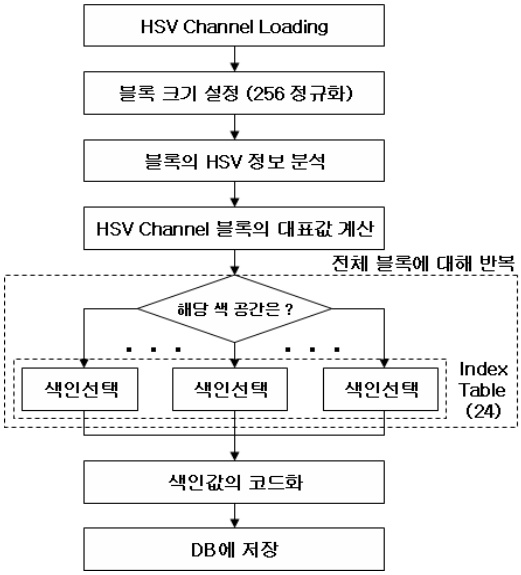



그림 4.2. HSV 칼라 색인화 흐름도

Fig. 4.2. HSV Color Indexing Flowchart

먼저 칼라 값을 정규비트 크기로 변환하기 위해 이미지 크기에 따른 검색 블록 크기를 결정한다 (이미지 당 256개의 데이터). HSV 칼라를 각 블록에서 기준이 되는 대표값으로 평균을 계산한다. 이 값을 비교하여 설정된 색인 값으로 변환한다.

색인 값은 그림 4.3과 같이 HSV 색상 공간을 24개로 분할하고 각각에 색인 값이 주어졌다. 이러한 수행을 통해서 해당 이미지에 분포해 있는 색상에 대한 분포 특성을 얻을 수 있게 된다. 이러한 동작을 전체 블록에 대해서 수행을 하고 수행이 끝나면 색인 값들을 코드화 하고 이를 데이터베이스에 저장 한다

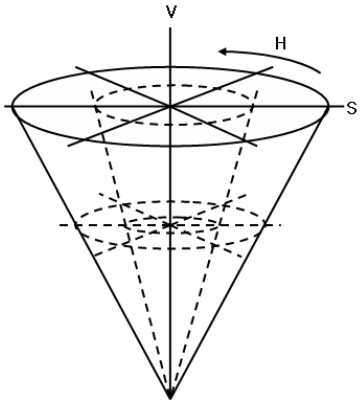



그림 4.3. HSV 칼라 색인화를 위한 칼라 공간 분할

Fig. 4.3. Color Space Partition for HSV Color Indexing

### c. 유사도 계산

그림 4.4는 본 논문에서 구현한 웹 기반 이미지 검색 시스템의 흐름도이다.

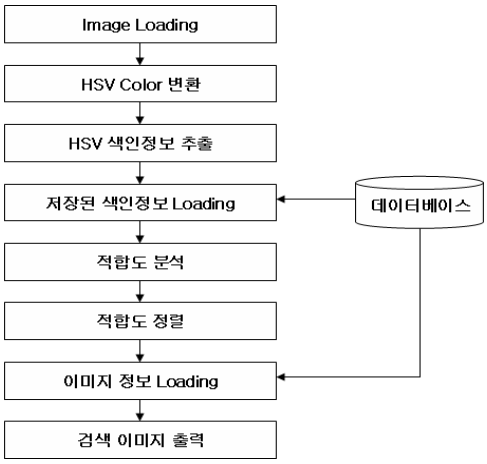



그림 4.4. 제안된 시스템의 흐름도

Fig. 4.4. Suggested System Flowchart

검색하고자 하는 기준 이미지를 로딩하고 검색을 하게 되면 기준 이미지에 대한 HSV 색인 정보를 추출하고 추출된 정보를 이용하여 데이터베이스에 저장된 비교이미지들과 비교를 하여 각각에 대한 적합도를 계산한다. 적합도는 수식 (4-2)에 의해서

계산된다.

$$A(i) = \sum_{i=0}^{256} Cp(I_b, I_d) \tag{4-2}$$

여기서 $I_b$는 기준이미지 코드이며 $I_d$는 비교이미지 코드이다. Cp() 함수는 각 비트의 코드를 비교하여 동일하면 0을 다르면 1을 반환한다.

적합도 계산이 완료되면 적합도를 크기순으로 정렬하고 해당 정보를 갖는 이미지의 위치를 데이터베이스로부터 검색하여 검색화면에 출력하게 된다.

## 2. 인터페이스 설계

그림 4.5는 제안된 시스템을 웹상에서 구현하기 위해 설계된 메인 화면의 인터페이스 설계이다. 화면은 크게 검색하고자 하는 비교 이미지를 출력하는 부분과 검색 방법 및 검색 키워드를 입력하는 부분, 검색된 이미지를 출력하는 부분을 구성되어 있다. 또한 관리자가 멀티미디어 데이터를 관리하도록 다양한 정보추출 알고리즘의 비교 및 검증, DB 설계, 테이블 생성, DB 제어가 가능하도록 하단에 추가 버튼을 구현할 수 있도록 하였다.

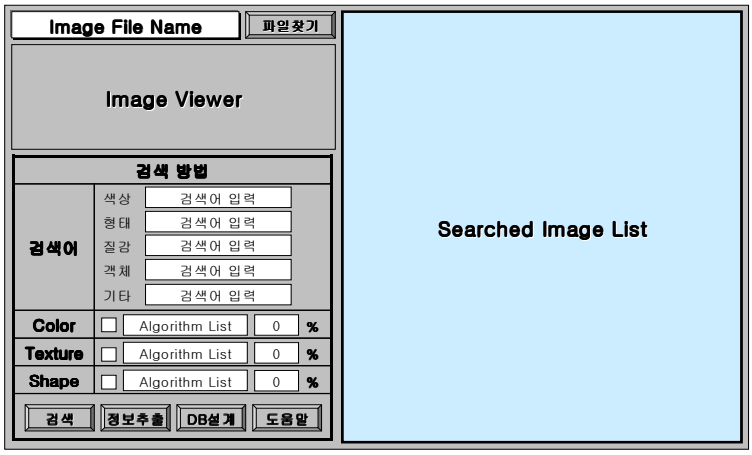



그림 4.5 인터페이스 설계

Fig. 4.5 Interface Design

# V. 구현 및 분석

## A. 구현

### 1. 시스템 구성

제안된 멀티미디어 검색 시스템은 윈도우 기반에서 제작되었지만 다른 OS에서도 수정없이 동작되도록 설계되었다. 웹 서버로는 플랫폼에 독립적인 Apache를 사용하였고, 데이터베이스로는 MySQL를 사용하였다. JSP에서 데이터베이스에 접근하기 위한 드라이버는 Type 4 형태인 JDBC 드라이버를 사용하였고, JSP 컨테이너로 JSP/Bean를 비교적 잘 지원하는 Resin를 사용하였다.

데이터 설계를 위해 XML를 사용하였으며 데이터 구조 설계는 DTD를 이를 검색하기 위해 DOM을 사용하였다. 현재 구현된 시스템은 JPEG와 GIF만 지원하고 있지만 차후 시스템에서는 모든 형태의 이미지를 지원할 수 있도록 모듈을 추가하기 편하게 설계되었다.

| 구분 | 종류 | 비고 |
|---|---|---|
| OS | Window XP or 2000 | |
| Web Server | Apache Web Sever 2.0.5 | |
| Database | MySQL, JDBC Driver 4 Type | |
| JSP Engine | Resin 3.0.5 | |
| XML | DTD, DOM, XQL | |
| Language | HTML, JavaScript, JSP/Bean, Applet | |
| Image Format | JPEG, GIF | |

### 2. 사용 이미지

그림 5.1은 테스트에서 사용된 이미지는 크기가 각기 다른 임의의 이미지 16개를 사용하였다. (1)-(2), (8)-(9)-(10), (11)-(12), (13)-(14), (15)-(16)이 각각 유사 이미지이다, (4) 이미지는 (3) 이미지를 확대한 것이고, (5) 이미지는 (3) 이미지를 축소한 것이다. (6) 이미지는 (3) 이미지의 밝기를 밝게 변화시킨 이미지이고 (7) 이미지는 (3) 이

미지를 180도 회전시킨 이미지이다. 전체 이미지는 미리 제안된 시스템에서 구현한 정보 추출 도구를 사용하여 각 이미지에 대한 기본 정보와 기술자 정보가 데이터베이스에 저장되었다.

(1)

(2)

(3)

(4)

(5)

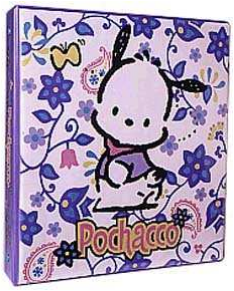
(6)

(7)

(8)

(9)

(10)

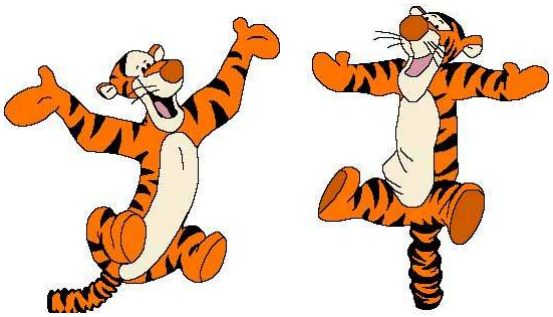


(11)

(12)

(13)

(14)



(15) (16)

그림 5.1 사용 이미지

Fig. 5.1 Using Image

## 3. 구현 시스템

### a. 시작 화면

그림 5.2는 설계한 인터페이스 화면을 실제로 웹상에서 구현한 화면이다. 상단의 “찾아보기” 버튼을 눌러서 검색하고자 하는 이미지를 선택하면 바로 아래에 있는 출력화면에 이미지가 출력된다. 초기화면에는 설정된 초기이미지가 출력되어 있다. 출력화면 아래쪽에는 검색 환경을 설정하는 부분이다. 키워드는 각각 색상, 질감, 형태, 객체, 기타로 구분하여 입력하도록 설계되었고, 색상은 사용자의 주관적인 색상에 대한 생각을 배제하기 위하여 기준 색을 설정하여 선택하도록 하였다.

사용자는 칼라, 질감, 형태 기술자를 사용할 것인지를 선택할 수 있고 각각의 기술자를 추출하는 알고리즘을 선택할 수 있으며 최종적으로 각 기술자들의 적용율을 설정할 수 있다.

다양한 알고리즘을 테스트하고 데이터베이스 관리를 하기 위한 기능을 사용하기 위한 버튼이 화면 하단에 설계되었다.

우측 화면은 검색된 결과를 출력하는 화면이다.

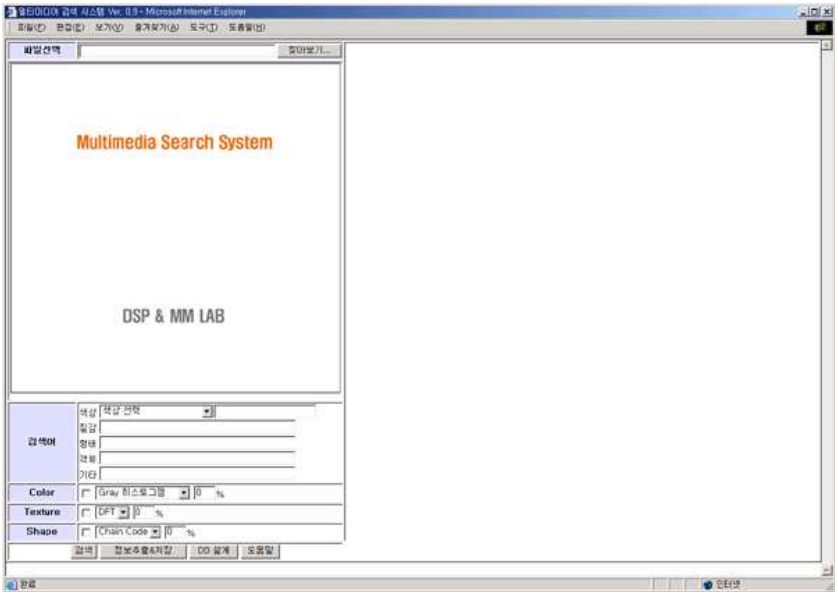



그림 5.2 시작 화면

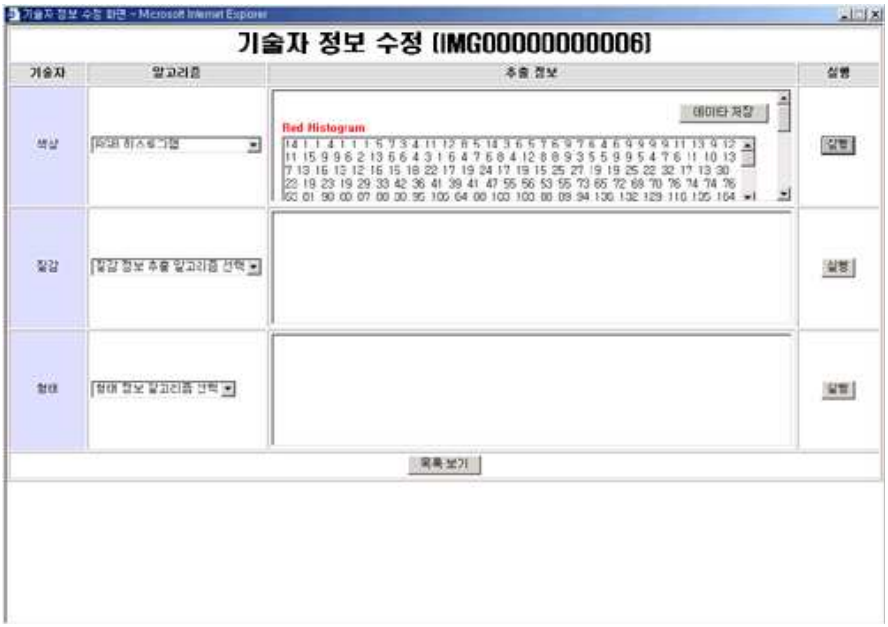
Fig. 5.2 Start Screen

### b. 정보 추출 및 저장 화면

그림 5.3는 설계한 인터페이스 화면을 실제로 웹상에서 구현한 화면이다. (a)는 해당 이미지에 대한 기본적인 정보를 입력하는 화면이다. 이미지의 코드, 이름, 위치, 형태, 크기, 입력한 사람이 자동적으로 입력되고 키워드와 기타 정보를 관리자가 입력할 수 있다. (b)는 색상, 질감, 형태에 대한 다양한 방법에 대한 값을 자동 생성하여 이를 데이터베이스에 저장할 수 있다. (c)는 데이터베이스에 등록된 이미지 목록을 보는 화면으로 자세히 보기, 수정, 삭제 기능을 지원한다.

기본정보 입력
파일선택
입력 정보
이미지 코드
File 이름
File 형태
File 위치
등록자 ID
admin
등록 일자
파일 크기
400
600
색상
질감
형태
기타
기타정보
목록보기
다시 작성

(a)


기술자 정보 수정 (IMG000000000006)
기술자
알고리즘
추출 정보
실행
색상
데이터 저장
Red Histogram
질감
형태
목록보기

(b)

(c)

그림 5.3 정보 추출 및 저장 화면 (a)기본 정보 추출 및 입력 화면, (b)기술자 정보 추출 및 입력화면, (c)이미지 목록 보기 화면

Fig. 5.3 Screen for Information Extraction and Data Storage (a)Screen for Basic Information Extract and Data Input, (b)Screen for Descriptor Information Extract and Data Input, (c)Screen for Image List Viewer

# B. 분석

## 1. 색상 분산차 코렐로그램을 이용한 방법 *I*

본 논문에서 제안한 알고리즘을 검증하기 위하여 먼저 색상 분산차 코렐로그램을 이용한 이미지 검색 알고리즘의 결과와 비교하였다. 기존의 칼라 히스토그램이 칼라 영상의 단순한 칼라 빈도수를 나타냄으로써 영상은 다르지만 비슷한 칼라 빈도수를 가지면 같은 이미지로 인식하는 점과 영상내의 밝기나 칼라 변화에 민감하여 데이터베이스가 커질수록 오류가 심각해지는 문제를 가지고 있으며, 칼라 코렐로그램 알고리즘도 특징 값 추출 및 거리에 의한 결과 값이 다르게 나올 수 있다는 문제점을 내포하고 있다. 이러한 알고리즘들의 문제점을 해결하기 위한 것이 색상 분산차 코렐로그램이다.

그림 5-4는 알고리즘을 나타내고 있다.

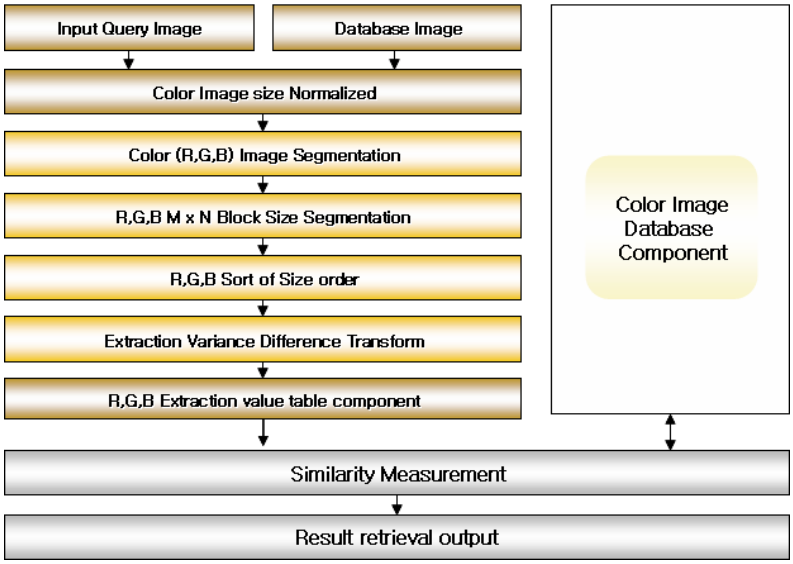



그림 5.4 색상 분산차 코렐로그램을 이용한 검색 알고리즘

Fig. 5.4 Image Retrieval Algorithm using Color Variance Difference Correllogram

표 5-1. 테스트 이미지의 매칭도

Table 5-1. Matching Rate of Test Image

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | *13* | *14* | *15* | *16* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1* | 100 | 64 | 28.6 | 27 | 25.9 | 28.8 | 28.6 | 31.1 | 37.7 | 34.2 | 36 | 38.7 | 36.6 | 34.8 | 40.3 | 39.5 |
| *2* | | 100 | 28.6 | 28 | 28.4 | 29.2 | 28.8 | 32.7 | 39.5 | 35.4 | 37.2 | 39.9 | 38.5 | 37 | 40.9 | 41.2 |
| *3* | | | 100 | 23.7 | 51.4 | 57.8 | 71.6 | 25.5 | 30.2 | 32.7 | 35.4 | 32.3 | 31.1 | 35 | 35.4 | 34.4 |
| *4* | | | | 100 | 23 | 24.9 | 23 | 19.6 | 23.5 | 22.4 | 24.9 | 29.8 | 23.7 | 24.7 | 24.1 | 24.7 |
| *5* | | | | | 100 | 46.9 | 46.9 | 27.8 | 31.9 | 35.6 | 35.4 | 33.5 | 32.5 | 34.2 | 34.2 | 35.4 |
| *6* | | | | | | 100 | 60.1 | 28.6 | 32.1 | 30.9 | 34 | 33.3 | 33.1 | 32.5 | 36 | 37.4 |
| *7* | | | | | | | 100 | 26.1 | 30.7 | 31.1 | 34.2 | 32.3 | 31.5 | 33.1 | 33.5 | 33.3 |
| *8* | | | | | | | | 100 | 58 | 48.1 | 31.9 | 34.8 | 40.1 | 37.9 | 53.1 | 54.1 |
| *9* | | | | | | | | | 100 | 57.8 | 44.9 | 49.6 | 53.7 | 52.9 | 63.4 | 64.8 |
| *10* | | | | | | | | | | 100 | 47.3 | 48.8 | 53.9 | 51.2 | 54.3 | 54.3 |
| *11* | | | | | | | | | | | 100 | 53.3 | 50.6 | 49.2 | 45.7 | 45.7 |
| *12* | | | | | | | | | | | | 100 | 56.6 | 55.1 | 50.4 | 49.8 |
| *13* | | | | | | | | | | | | | 100 | 67.3 | 52.5 | 54.5 |
| *14* | | | | | | | | | | | | | | 100 | 49.4 | 48.6 |
| *15* | | | | | | | | | | | | | | | 100 | 86.2 |
| *16* | | | | | | | | | | | | | | | | 100 |

표 5-1은 색상 분산차 코렐로그램 알고리즘을 이용한 이미지 검색 결과이다. 표에서 반전된 부분이 실제 유사 이미지 목록으로 대부분 유사 이미지를 찾는 것을 볼 수 있다. 하지만, 유사 이미지와 다른 이미지 간의 구분을 명확하게 하지 못하고 영상의 변화가 심한 경우에 오류가 발생하는 것을 알 수 있다. 이러한 문제가 발생하는 것은 단순히 R, G, B를 기준이 없이 색상차만을 이용하기 때문에 발생하는 문제이다.

## 2. 색상 분산차 코렐로그램을 이용한 방법 *II*

이전의 알고리즘의 문제점을 해결하기 위해서 제안된 것이 단순히 R, G, B의 크기 순이 아니라 G를 기준으로 R, B 색상의 색상 분산차를 이용하여 알고리즘을 개선한 것이다.

표 5-2. 테스트 이미지의 매칭도

Table 5-2. Matching Rate of Test Image

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | *13* | *14* | *15* | *16* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1* | 100 | 74.1 | 44.2 | 42.4 | 43.6 | 45.9 | 43.8 | 46.7 | 50.8 | 47.3 | 50.2 | 51 | 50 | 47.9 | 53.9 | 53.1 |
| *2* | | 100 | 43.8 | 42.2 | 43.6 | 45.1 | 43.2 | 46.5 | 51.6 | 47.3 | 48.8 | 50.2 | 50.4 | 49.8 | 53.3 | 53.1 |
| *3* | | | 100 | 39.1 | 61.5 | 68.1 | 81.7 | 41.4 | 43.6 | 45.9 | 48.8 | 46.1 | 44.4 | 45.9 | 47.5 | 47.9 |
| *4* | | | | 100 | 39.1 | 40.5 | 38.9 | 36.8 | 39.3 | 38.3 | 40.3 | 44.4 | 39.7 | 40.1 | 39.7 | 40.5 |
| *5* | | | | | 100 | 60.5 | 58.8 | 44 | 46.3 | 48.4 | 50 | 48.4 | 45.1 | 47.3 | 47.5 | 49.2 |
| *6* | | | | | | 100 | 71.4 | 42.2 | 46.1 | 44 | 49 | 47.9 | 46.7 | 46.5 | 49.6 | 50 |
| *7* | | | | | | | 100 | 41.4 | 44.4 | 45.3 | 49 | 46.7 | 43.4 | 45.3 | 46.9 | 46.9 |
| *8* | | | | | | | | 100 | 65.8 | 55.3 | 44.7 | 46.7 | 50.4 | 50.6 | 59.5 | 60.7 |
| *9* | | | | | | | | | 100 | 61.5 | 53.7 | 57.2 | 59.9 | 60.9 | 67.1 | 66.9 |
| *10* | | | | | | | | | | 100 | 54.3 | 55.1 | 60.5 | 59.3 | 59.3 | 60.3 |
| *11* | | | | | | | | | | | 100 | 60.3 | 59.3 | 57.4 | 54.3 | 53.5 |
| *12* | | | | | | | | | | | | 100 | 63.4 | 62.8 | 58.8 | 58 |
| *13* | | | | | | | | | | | | | 100 | 69.8 | 61.1 | 61.5 |
| *14* | | | | | | | | | | | | | | 100 | 58.4 | 57.4 |
| *15* | | | | | | | | | | | | | | | 100 | 88.3 |
| *16* | | | | | | | | | | | | | | | | 100 |

표 5-2는 개선된 알고리즘을 이용한 이미지 검색 결과이다. 이전 알고리즘에 비해

유사 이미지를 더 명확하게 검색하는 것을 알 수 있다. 하지만, 영상이 변화가 심한 경우에는 이전 알고리즘보다는 정확하게 검색할 수 있지만 여전히 오류를 내포하고 있음을 알 수 있다.

## 3. 제안한 알고리즘

이전에서 실험된 RGB 색상 기반의 검색 알고리즘의 단점을 해결하기 위해 HSV 색상을 검색 기술자로 사용하였다. HSV 색상은 RGB가 색상만을 표현하기 때문에 명암이나 채도 등의 변화가 있는 이미지 검색에서는 적절하기 못하는 방법이다. 하지만, HSV는 색상뿐만 아니라 채도와 명도를 모두 표현하고 있기 때문에 RGB에서 발생하는 문제를 해결할 수 있다. 본 논문에서는 이러한 장점을 갖는 HSV 색상을 기반으로 하는 알고리즘을 제안하였고, HSV가 갖는 색상 정보 데이터의 크기를 줄이기 위해서 HSV 색상 공간을 24개로 분할하여 각 공간의 색상분포도를 정보로 하여 데이터베이스에 저장하고 이를 웹에서 검색함으로써 검색속도를 높일 수 있었다. 본 논문에서 제안한 알고리즘을 검증하기 위해 웹상에서 이를 구현하였다.

표 5-3. 테스트 이미지의 매칭도

Table 5-3. Matching Rate of Test Image

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | *13* | *14* | *15* | *16* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *1* | 100 | 89.6 | 44.4 | 43.2 | 48 | 10 | 43.6 | 31.6 | 22 | 24.8 | 14.8 | 10.8 | 26.4 | 16.4 | 21.6 | 25.2 |
| *2* | | 100 | 47.6 | 46.4 | 51.2 | 14 | 46.8 | 36.4 | 26.8 | 28.8 | 14.8 | 14 | 34 | 18.4 | 24.8 | 30 |
| *3* | | | 100 | 96.4 | 75.2 | 66.4 | 94.8 | 63.2 | 56 | 36 | 50 | 66.4 | 38.4 | 60.4 | 59.2 | 63.2 |
| *4* | | | | 100 | 76 | 67.6 | 96 | 61.2 | 54.4 | 34 | 46.8 | 67.6 | 36.8 | 63.2 | 56 | 60 |
| *5* | | | | | 100 | 62.8 | 80 | 58.8 | 48 | 35.6 | 34.8 | 48.8 | 40.4 | 57.2 | 44.8 | 48.8 |
| *6* | | | | | | 100 | 67.2 | 52.8 | 54 | 33.2 | 52.4 | 74.4 | 34.8 | 71.6 | 62.4 | 63.2 |
| *7* | | | | | | | 100 | 63.2 | 56.4 | 36 | 46.4 | 65.2 | 38 | 65.2 | 55.6 | 48.8 |
| *8* | | | | | | | | 100 | 88.4 | 72.4 | 48 | 52.8 | 39.6 | 52.8 | 58.8 | 67.6 |
| *9* | | | | | | | | | 100 | 74.8 | 52.4 | 54 | 36 | 54 | 62 | 67.6 |
| *10* | | | | | | | | | | 100 | 46 | 33.2 | 32 | 33.2 | 44.4 | 49.6 |
| *11* | | | | | | | | | | | 100 | 78 | 23.6 | 41.2 | 86.4 | 78 |
| *12* | | | | | | | | | | | | 100 | 34.8 | 63.2 | 88 | 80.8 |
| *13* | | | | | | | | | | | | | 100 | 63.2 | 32.8 | 35.2 |
| *14* | | | | | | | | | | | | | | 100 | 51.2 | 52 |
| *15* | | | | | | | | | | | | | | | 100 | |
| *16* | | | | | | | | | | | | | | | | 100 |

표 5-3은 본 논문에서 제안한 시스템에 테스트 이미지 16개를 적용한 매칭도이다. 숫자가 적을수록 매칭도가 높다. 실험 결과에서 보이는 것처럼 이전의 알고리즘들에 비해 유사 이미지를 명확하게 검색하는 것을 볼 수 있으며 이전의 RGB 기반의 알고리즘의 방법의 결과에 비해 이미지의 크기가 변하거나 회전하거나 밝기가 변하여도 해당 이미지를 검색함을 알 수 있다. 하지만, 다른 이미지 일지라도 동일한 색상 분포를 갖는 이미지에서는 오류가 발생하는 것을 알 수 있다. 즉, 3번과 4번 이미지를 검색할 때 다른 이미지인 12번 이미지가 비슷한 이미지로 검색되고 있으며, 6번 이미지를 검색할 때 12번과 14번과 같은 다른 이미지가 검색되고 있으며, 11번 이미지에서는 15번, 16번 이미지가 잘못 검색되는 것을 알 수 있었다. 이러한 문제가 발생하는 것은 비슷한 색상 분포를 갖기 때문으로 이를 해결할 수 있는 연구가 필요할 것이다. 이러한 문제를 해결하기 위해서는 본 논문에서 제안한 알고리즘에서 임의로 S와 V 공간의 임계치를 0.5로 선택하였기 때문에 이 임계치를 계산하는 알고리즘을 추가해야 할 것이다.

# VI. 결론 및 향후 연구과제

이미지 검색에서 사용하는 기술자로 색상, 형태, 질감 등이 사용된다. 그 중에서 색상은 가장 많은 정보를 포함하고 있는 기술자이다. 색상 기술자 중에서 가장 기본이 되는 색상 공간이 RGB이지만 RGB가 색상만을 표현함으로 해서 밝기 변화 등의 이미지가 변형되는 상태에서는 오류를 발생시키는 단점을 가지고 있다. 하지만, HSV 색상 공간은 색상, 명도, 채도를 표현하는 색상공간으로 RGB가 가지는 단점을 해결할 수 있는 인간의 시각에 가장 가까운 모델이라 할 수 있다. 하지만 RGB에 비해 계산시간이 오래 걸리는 단점과 정보 데이터의 크기가 크다는 단점을 가지고 있다.

이와 같은 문제를 해결하기 위해 본 논문에서는 기존의 칼라 기반 이미지 검색 시스템에서 발생하는 검색 오류를 HSV 칼라 색인을 이용한 알고리즘을 제안하여 개선하였다.

이와 같이 개선된 알고리즘과 구축된 이미지 데이터베이스를 이용함으로써 이미지의 크기 변화, 회전, 밝기 변화에도 강인한 검색을 할 수 있음을 실험을 통해 검증하였다. 또한, 칼라 정보를 색인화 하여 데이터베이스에 저장함으로써 데이터의 크기를 줄여 웹 콘텐츠에서도 이미지 검색을 가능하도록 하였다.

향후 연구과제로는 본 논문에서 제안한 알고리즘이 비슷한 색상 공간 분포를 갖는 이미지에서 오류를 발생한다는 단점을 가지고 있으므로 이를 해결함으로써 정확한 검색이 가능한 알고리즘에 대한 연구가 진행되어야 할 것이다. 또한, 내용기반의 다른 기술자들인 질감 기술자와 형태 기술자 등과 연계한 통합 검색 알고리즘에 대한 연구도 진행되어야 할 것이다. 마지막으로 기존의 키워드 기반의 검색엔진과 특징 기술자 기반의 검색엔진을 통합하여 최종적인 검색 엔진을 개발해야 할 것이다.

# 참고 문헌


[1] "2004년 상반기 정보화실태조사", 정보통신부, 한국인터넷진흥원, 2004. 6

[2] "제15차 중국 인터넷 발전현황 통계보고", 중국인터넷정보센터(CNNIC), 2005. 1

[3] 장시웅, 정기동, "분산 UNIX 환경에서 Shared-Concurrent File System의 성능 및 구현", 한국정보처리학회논문지 제3권 제3호, pp.617~630, 1996. 5

[4] M. J. Swain and D. H. Ballard, "Color Indexing", International Journal of Computer Vision Vol.7 No.1, pp.11~32, 1991.

[5] B.V. Funtand G.D. Finlayson , "Color constant color indexing ", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 17, pp.522~529, 1995.

[6] B.M. Mehtre, M.S . Kankanhalli, A .D. Narsimhalu, and G.C. Man, "Color matching for image retrieval", Pattern Recognition Letter, Vol.16, pp.325~331, 1995,

[7] M. Safar, C. Shahabi and X. Sun, "Image retrieval by shape: a comparative study", ICME 2000, Vol.1, pp.141~144, 2000.

[8] A. Khotanzad and Y.H. Hong, "Invariants Image Recognition by Zernike Moments", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.12, No.5, pp.489~497, 1990.

[9] H.K. Kim, J.D. Kim, D.G. Sim and D.I. Oh, "A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval", ICME 2000, Vol.1, pp.307~310, 2000.

[10] Y.S. Kim and W.Y. Kim, "Content-based trademark retrieval system using a visually salient feature", Image & Vision computing, Vol.16, No.12~13, pp.931~939, 1998.

[11] A.K. Jain, "Fundamental of Digital Image Processing", Prentice Hall International, 1989.

[12] A.K. Jain and A. Vailaya, "Image Retrieval using Color and Shape", Pattern Recognition, Vol.29, No.8, pp.1233~1244, 1996.

[13] S. Belongie, C. Carson, H. Greenspan and J. Malik, "Color and Texture-based

Image Segmentation Using EM and Its Application to Content-based Image Retrieval", Computer Science Division , University of California at Berkeley, ICCV'98, 1998.

[14] R. Chellappa and S. Chatterjee, "Classification of Textures Using Gaussian Markov Random Fields", IEEE Trans. on Acoustics, Speech, and Signal Processing, vol.ASSP-33, no.4, pp.959, 1985.

[15] B.S. Manjunath and W.Y. Ma, "Texture Features for Browsing and Retrieval of Image Data", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.18, No.8, pp.837~842, 1996.

[16] T. Gevers and Arnold W.M, "PicToSeek : Combining Color and Shape Invariant Features for Image Retrieval", IEEE Trans. on Image Processing, Vol.9, No.1, pp.102~119, 2000.

[17] R. Milanese, "A Rotation, Translation, and Scale-Invariant Approach to Content-based Image Retrieval", Journal of Visual Communication and Image Representation, Vol.10, pp.186~196, 1999.

[18] G. Pass and R. Zabih, "Comparing images using joint histogram", Multimedia Systems, Vol.7, pp.234~240, 1999.

[19] D.K. Park, W.S. Jeon, C.S . Won, S.J. Park and S.J. Yoo, "A Composite Histogram for Image Retrieval", ICME, Vol.1, pp.355~358, 2000.

[20] ISO/IEC JTC 1/SC 29/WG 11: "Information Technology-Multimedia Content Description Interface-Part 5: Multimedia Description Schemes", MPEG document N4005, Singapore, March, 2001.

[21] 권순범, "웹과 무선이동통신 시스템과의 통합을 위한 기술소개와 인터넷 비즈니스모델의 확장사례 연구", 추계지능정보시스템 학회, 1999. 11.

[22] 정갑영 외, "유무선 통신의 균형발전을 위한 정책방향", 한국통신경영연구소, pp.31~56, 2000. 12.

[23] 강호경, 정용주, 유기원, 노용만, 김문철, 김진웅, "MPEG-7 질감 기술자", 방송공학회논문지, Vol.5, No.1, pp.10~22, 2000.

[24] 박동권, 전윤석, 박수준, 원치선, "MPEG-7을 위한 에지 히스토그램 서술자", 방송공학회논문지, Vol.5, No.1, pp.31~40, 2000.

[25] Bober, M., "The Contour Shape Descriptor for MPEG-7 and Its Applications", IEEE, pp.286~288, 2000.

[26] 이동호, 유광석, 김희율, "컬러와 모양 정보를 이용한 캐릭터 이미지 검색", 방송공학회 논문지, Vol.5, No.1, pp.50~60, 2000.

[27] 이상미, "MPEG-7 기술을 이용한 내용 기반 정보검색의 활용", 통신소프트웨어 학술대회, Vol.1 No.1, pp.253~256, 1997.

[28] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann, pp.81~113, 2001.

[29] Jain A. K, "Fundamentals of Digital Image Processing", Englewood Cliffs. NJ: Prentice-Hall, 1989.

[30] Otha Y, T. Kanade and T. Sakai, "Colour Information for Region Segmentation", Computer Graphics and Image Processing, pp.222~241, 1980.

[31] Itten. J, "색체의 예술", 1976.

[32] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann Publishers, pp.97~99, 2001.

[33] M.A. Stricker and M. Orengo, "Similarity of Color Images". Storage and Retrieval for Image and Video Databases (SPIE), pp.381~392, 1995.

[34] D. Androutsos, K.N. Plataniotis, and A.N. Venetsanopoulos, "A Novel Vector-Based Approach to Color Image Retrieval Using a Vector and Image Angular-Based Distance Measure", Computer Vision Understanding, Vol.75, no.1/ 2, pp.46~58, 1999.

[35] Gevers T. and A. Smeulers, "A Comparative Study of Several Colour Models for Colour Image Invariant Retrieval", Proc. of First Int. Workshop on Image Database and Multimedia Search, Amsterdam, pp.17~27, 1996. 8

[36] Colombo C, A. Rizzi and I. Genoversi, "Histogram families for colour-based retrieval in image dagabase", Proc. ICIAP'97, Florence, pp.204~211, 1997. 12

[37] Smith J. R. and S. F. Chang, "VisualSEEK: A fully automated content-based image query system", ACM multimedia'96, Boston, MA. 1996. 11

[38] Alberto Del. Bimbo, "Visual Information Retrieval", Morgan Kaufmann, pp.117~131, 2001.

[39] Rafael C, Gonzalez, Peter Neo, Rajiv Mebrotra, "A Pictorial Index Mechanism for Model-based Matching", Data and knowledge engineering. Vol.8, pp.309~327, 1992

[40] 김진아, "내용기반 영상 데이터베이스 검색을 위한 질감 특징 추출", 창원대학교석사학위 논문, 1997.

[41] Crois G. R and A. K. Jain, "Markov random field texture models", IEEE Trans. Pattern Analysis and Machine Intelligence, 5CD:25~39, 1983.

[42] Bajscy R., "Computer identification of visual surfaces", Computer Graphics and Image Processing, pp.118~130, 1973.

[43] Laws K., "Rapid texture identification", SPIE Vol.238, Image Processing for Missile Guidance, pp.376~380, 1980.

[44] Kbotanzad A. and R. L. Ksyshap, "Feature selection for texture recognition based on image synthesis", IEEE Trans. on system, Man and Cybernetics, 17(6):1087~1095

[45] Picard R., "A Society of Models for Video and Image Libraries", Technical Report no. 360, MIT Media Lab Perceptual computing Section, 1995

[46] 이문호, "내용기반 영상 검색 시스템의 구현", 경상대학교 대학원 석사학위 논문, 2000

[47] H. Tamura, S. Mori and T. Yamawaki, "Textural Features Corresponding to Visual Perception", IEEE Trans. on System, Man, And Cybernetics, Vol.SMC-8, No.6, pp.460~473, 1978.

[48] F. Liuand, R.W. Picard, "Periodicity, directionality, and randomness : Wold features for image modeling and retrieval", IEEE Trans. on PAMI, Vol.18, No.7, pp.722~733, 1996.

[49] L.M. Kaplan et al, "Fast texture database retrieval using extended fractal features", in Storage and Retrieval for Image an d Video Databases VI, Proc SPIE 3312, pp.162~173, 1998.

[50] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann, pp.134~156, 2001.

[51] S.K. Kang, Y.C. Jung, J.A. Park, "Image Corner Detection Using Hough Transform", Springer(LNCS3523), pp.279~286, 2005.

[52] M.H. Jang, H.J. Park, K.S. Seo, J.A. Park, "Histogram as a Measure of Shape Matching", SPIE, pp.431~438, 2004.

[53] Randy crane, "A simplified approach to image processing", Prentice-Hall, pp.18~29, 1997.

[54] Blum H. and R. Nagel, "Shape description using weighted symmetric axis feature", Pattern Recognition, pp.167~180, 1978.

[55] Shapiro L. G., "A structural model of shape", IEEE Trans, on Pattern Analysis and Machine Intelligence, pp.10~20, 1979.

[56] William I. Grosky, Peter Neo, Rajiv Mebrotra, "A Pictorial Index Mechanism for Model-based Matching", Data and knowledge engineering. Vol.9, pp.309~327, 1992.

[57] S. T. Campbell and S. M. Chang, "The Role of Database System in the Management of Multimedia Information", Proc. of International Workshop on Multi-Media Database Management System, pp.4~11, 1995. 8.

[58] Rajiv Mebrotra and James E. Gary, "Feature-Based Retrieval of Similar Shape", IEEE 9th International Conference on Data Engineering, pp.108~115, 1993.

[59] Chock M., A. Cardenas and A. Klinger, "Database structure and manipulation capabilities of a picture database management system", IEEE Trans. on Pattern Analysis and Machine Intelligence, pp.484~492, 1984.

[60] Joseph T. and A. Cardenas, "PICQUERY: A High level query language for pictorial database management", IEEE Trans. on Software Engineering, pp.630~638, 1998.

[61] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann Publisher, pp.1~19, 2001.

[62] 김진아, 이승훈, 우용대, 정성환, "내용기반 영상 검색을 위한 특징 추출 및 영상 데이터베이스 검색 시스템 구현", 한국정보처리학회논문지 제5권 제 8호, 1998.

[63] http://search.empas.com/search/image_search.html

[64] http://images.google.com/

[65] Chang C. W., K. F. Lim and S. Y. Lee, "The characteristics of digital video

and considerations in designing video database". Proc. 4th Int. ACM Cont. on Information and Knowledge Management, Baltimore, MD, pp.370~377, 1995.

[66] Minka T. and R. Picard, "Interactive learning with a society of models", Pattern Recognition, 30(4), pp.565~582, 1997.

[67] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann Publisher, pp.26~29, 2001.

[68] K. Hirata and T. Kato, "Query by visual example-content based image retrieval", Advances in Database Technology (EDBT'92), pp.56~71, 1992.

[69] Bather, W. Equitz, "Query By Content for Large On-Line Image Collection", IEEE, 1995.

(별 지)

# 저작물 이용 허락서

| 학 과 | 정보통신공학과 | 학 번 | 20027401 | 과 정 | 석사 |
|---|---|---|---|---|---|
| 성 명 | 한글: 백 종 현 한문 : 白 宗 玄 영문 : Beak Jong Hyun | | | | |
| 주 소 | 광주광역시 동구 서석동 21-11 | | | | |
| 연락처 | E-MAIL : muruuu@hanmail.net | | | | |
| 논문제목 | 한글 : HSV 칼라 색인을 이용한 CBIR 검색 시스템 설계<br>영문 : Design of an Content based CBIR System Using HSV Color Index | | | | |

본인이 저작한 위의 저작물에 대하여 다음과 같은 조건아래 조선대학교가 저작물을 이용할 수 있도록 허락하고 동의합니다.

- 다 음 -

1. 저작물의 DB구축 및 인터넷을 포함한 정보통신망에의 공개를 위한 저작물의 복제, 기억장치에의 저장, 전송 등을 허락함
2. 위의 목적을 위하여 필요한 범위 내에서의 편집형식상의 변경을 허락함. 다만, 저작물의 내용변경은 금지함.
3. 배포전송된 저작물의 영리적 목적을 위한 복제, 저장, 전송 등은 금지함.
4. 저작물에 대한 이용기간은 5년으로 하고, 기간종료 3개월 이내에 별도의 의사표시가 없을 경우에는 저작물의 이용기간을 계속 연장함.
5. 해당 저작물의 저작권을 타인에게 양도하거나 또는 출판을 허락을 하였을 경우에는 1개월 이내에 대학에 이를 통보함.
6. 조선대학교는 저작물의 이용허락 이후 해당 저작물로 인하여 발생하는 타인에 의한 권리 침해에 대하여 일체의 법적 책임을 지지 않음
7. 소속대학의 협정기관에 저작물의 제공 및 인터넷 등 정보통신망을 이용한 저작물의 전송출력을 허락함.

200 년 월 일

저작자: 백 종 현 (서명 또는 인)

## 조선대학교 총장 귀하

2007年 2月 碩士學位論文

# HSV칼라색인을이용한CBIR검색시스템설계

白 宗 玄