[UCI]I804:24011-200000234711

2005年 8月
博士學位論文

특징 기술자와 의미론적 키워드를 이용한 웹기반 이미지 검색 시스템 설계

朝 鮮 大 學 校 大 學 院

電 子 工 學 科

姜 成 官

2005年 8月 博士學位論文

특징 기술자와 의미론적 키워드를 이용한 웹기반 이미지 검색 시스템 설계

姜 成 官

특징 기술자와 의미론적 키워드를 이용한 웹기반 이미지 검색 시스템 설계

Design of the Web based Image Retrieval System Using Feature Descriptors and Semantic Keyword

2005年 8月 日

朝 鮮 大 學 校 大 學 院

電 子 工 學 科

姜 成 官

특징 기술자와 의미론적 키워드를 이용한 웹기반 이미지 검색 시스템 설계

指導敎授 朴 鍾 安

이 論文을 工學博士學位申請 論文으로 提出함.

2 0 05 年 4月 日

朝 鮮 大 學 校 大 學 院

電 子 工 學 科

姜 成 官

姜成官의 博士學位 論文을 認准함

委員長 光州科技阮 敎授 崔 泰 善 印

委 員 朝鮮大學校 敎授 韓 承 朝 印

委 員 순천大學校 敎授 文 庸 善 印

委 員 株)建和通信 代表 徐 載 宣 印

委 員 朝鮮大學校 敎授 朴 鍾 安 印

2 0 05 年 6 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 ····· 10
B. 멀티미디어 정보 추출 ····· 17
1. Color ····· 17
2. Texture ····· 27
3. Shape ····· 32
4. 기술자 설계 ····· 37

III. 검색 엔진 ····· 45
A. 이미지 정보의 검색 방법 ····· 45
1. 문자 기반 검색 ····· 45
2. 의미론적 기반 검색 ····· 46
3. 내용 기반 검색 시스템 ····· 46
B. 웹 기반 이미지 검색 시스템 ····· 54

IV. 시스템 설계 ····· 56
A. 기술자 설계 ····· 56
1. 색상 기반 검색 시스템 설계 ····· 56

2. 질감 기반 검색 시스템 설계 ····· 60
B. 검색 시스템의 설계 ····· 62
C. 데이터 구조 및 데이터베이스의 설계 ····· 63
1. 데이터베이스 설계 ····· 63
2. 데이터 구조 설계 ····· 64
D. 웹 기반 멀티미디어 데이터 관리 시스템의 설계 ····· 67
1. 시스템의 구성 ····· 67
2. 인터페이스 설계 ····· 68

***V.* 구현 및 분석** ····· 70
A. 구현 ····· 70
1. 시스템의 구성 ····· 70
2. 사용 이미지 ····· 70
3. 구현 시스템 ····· 74
B. 분석 ····· 77
1. 색상 기반 검색 ····· 77
2. 질감 기반 검색 ····· 87
3. 키워드 기반 검색 ····· 89

***VI.* 결론 및 향후 연구** ····· 94

참 고 문 헌 ····· 96

그 림 목 차

Fig. 2.1 Scope of MPEG-7 Standardization ····· 11
Fig. 2.2 Structure of Standard Elements and Application System of MEPG-7 ····· 13
Fig. 2.3 RGB Color Model ····· 18
Fig. 2.4 HSI Color Model ····· 20
Fig. 2.5 Color Histogram ····· 21
Fig. 2.6 Three Color Histogram whose Perceptual Similarities do not Correspond to their L_1 Distances ····· 25
Fig. 2.7 Cumulative Histograms of the Histograms Displayed in Fig. 2.6 ····· 26
Fig. 2.8 Classification of Retrieval Techniques based on Shape Information ····· 33
Fig. 2.9 The Phase of Zernike Polynomial ····· 35
Fig. 2.10 Structure of XML Document ····· 40
Fig. 2.11 Process of XML Document ····· 41
Fig. 3.1 Structure of the Content-based Retrieval System ····· 47
Fig. 3.2 QBIC System ····· 49
Fig. 3.3 VisualSEEK System ····· 49
Fig. 3.4 Chabot System ····· 50
Fig. 3.5 Photobook System ····· 51
Fig. 3.6 SIMPLIcity System ····· 52
Fig. 4.1 HSV Color Space ····· 57
Fig. 4.2 Procedure of Color Segmentation ····· 58
Fig. 4.3 Peak in Hue Histogram ····· 60
Fig. 4.4 2D DCT Cosine Basis Functions ····· 61
Fig. 4.5 Procedure of DCT Peak Histogram ····· 62
Fig. 4.6 Structure of Retrieval System ····· 63
Fig. 4.7 Structure of Database ····· 64
Fig. 4.8 Structure of Metadata ····· 65
Fig. 4.9 Structure of Total System ····· 67

Fig. 4.10 Diagram of Total System ····· 68
Fig. 4.11 Interface Design ····· 69
Fig. 5.1 Using Image ····· 73
Fig. 5.2 Start Screen ····· 74
Fig. 5.3 Screen for Information Extraction and Data Storage ····· 76
Fig. 5.4 Algorithm using Gray Color Histogram ····· 77
Fig. 5.5 A Peformed Result Graph using Gray Color Histogram ····· 78
Fig. 5.6 Algorithm using RGB Color Histogram ····· 80
Fig. 5.7 A Peformed Result Graph using RGB Color Histogram ····· 81
Fig. 5.8 Algorithm using HSV Color Histogram ····· 83
Fig. 5.9 A Peformed Result Graph using HSV Color Histogram ····· 84
Fig. 5.10 A Peformed Result Graph using HSV Color Segmentation ····· 87
Fig. 5.11 Algorithm using DCT Peak Histogram ····· 90
Fig. 5.12 Screen for Search Keyword Input ····· 92
Fig. 5.13 Proposed Key-Based Retrieval Technique ····· 93

표 목 차

Table 2.1 Example of Dominant Features and Descriptor of Audio Visual Data ····· 12
Table 2.2 The List of Zernike Moments ····· 36
Table 2.3 Comparison of HTML, SGML, XML ····· 38
Table 3.1 Comparison of Content-Based Image Retrieval System ····· 52
Table 5.1 Test Result using Gray Color Histogram ····· 78
Table 5.2 A Searched Images using Gray Color Histogram ····· 79
Table 5.3 Test Result using RGB Color Histogram ····· 81
Table 5.4 A Searched Images using RGB Color Histogram ····· 82
Table 5.5 Test Result using HSV Color Histogram ····· 84
Table 5.6 A Searched Images using HSV Color Histogram ····· 85
Table 5.7 Test Result using HSV Color Segmentation ····· 87
Table 5.8 A Searched Images using HSV Color Segmentation ····· 88
Table 5.9 Test Result using DCT Peak Histogram ····· 90
Table 5.10 A Searched Images using DCT Peak Histogram ····· 91

ABSTRACT

Design of the Web based Image Retrieval System Using Feature Descriptors and Semantic Keyword

Kang, Sung-Kwan
Advisor : Prof. Park, Jong-An, Ph. D
Department of Electronic Engineering
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 the way of solving the problem that has the incorrectness of the existing keyword-based multimedia retrieval system. Based on it, this paper try to construct the paper with the web-based multimedia retrieval system.

For the solution of the defect of the exiting simple keyword-based multimedia retrieval system, first of all, we applicate the semantic-based keyword method instead of it. Such makes a syntactic retrieval to be able through a input with the keyword that has the information matched a descriptor information instead of the simple keyword. The color keyword can accept only registered colors and gets rid of problem that the retrieval has exception though it is the same meaning of a keyword. Also, it is easy for you to draw a objective keyword and refer to syntactic retrieval when you make a input with segmenting texture, shape, object etc.

Secondly, the suggested system can refer to information including the descriptor information in case of necessity of a keyword-based. Descriptors can use color, texture, shape and add or remove various algorithm in case of need.

Thirdly, it provides a function of a integrative retrieval by using combination of a keyword and each descriptor information. As a rate set up by users, each method draws a degree of similarity and refers to the image of the highest similarity.

Lastly, all of the image information are stored in database and referred results are consisted of XML data to support the function such as retrieval efficiency, high- speed retrieval, various output methods.

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[4-7], 이러한 다양한 멀티미디어 데이터들을 효과적으로 검색하기 위해서는 각각의 미디어별 데이터 구조설계와 그에 적합한 검색 엔진의 개발이 필요하게 되었다.

하지만, 초기의 검색 시스템에서는 질의(Query) 자료로 문자 키워드를 제시하고 동일한 문자 색인(Index)을 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(Primitive) 특징으로서 이미지의 색상(Color), 형태(Shape) 및 질감(Texture) 정보 등이 이에 해당되고, 두 번째로 논리적인(Logical) 특징으로서 포함되어 있는 물체의 식별을 위한 논리적인 정보이다. 마지막으로 추상적인(Abstract) 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 등

이 제안한 히스토그램 인터섹션(Histogram Intersection) 방법이 대표적이며[8], 색상 비율의 히스토그램을 비교함으로서 물체를 인식하는 확장 Swain 방법[9], 그리고 색상 비교를 하기 위해 거리 방법(Distance Method)과 기준 색상 테이블 방법(Reference Color Table method)[10]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[9].

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[11]. 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(Chain Code)[15], 저나이크 모멘트(Zernike Moment)[12-14], 불변 모멘트 (Invariant Moment)[15, 16], 퓨리어 묘사(Fourier Descriptor)[15] 등을 특징 파라미터로 사용하는 기술들이 제안되었다.

질감 정보를 이용한 이미지 검색은 웨이블릿 변환(Wavelet Transform)과 같은 다해상도(Multi-resolution) 필터링 기술과 함께 발전되어 왔다. 질감은 많은 시각 요소들이 밀집해서 균일하게 정렬되어 있는 패턴이다[17]. 이를 분석하기 위해서 다 방향 필터뱅크를 이용하거나, Markov 모델을 사용한 모델 기반 분석[18]을 하는 연구가 많이 수행되었으며 최근에는 Gabor 웨이블릿 변환을 이용하여 이미지의 질감 정보를 주파수 대역별로 분해하여 분석하는 방법이 많이 사용되고 있다[19].

또한 두 가지 이상의 특징을 혼합하여 검색을 위한 정보로 이용하는 방법들도 제시되었는데[20-38], Jing Huang 등[37, 38]은 색상 상관도(Color Correlogram)를 도입하여 색상 정보 및 색상들의 공간적인 배치 정보도 동시에 고려하였으며, D.K. Park 등[23]은 색상, 그레이 정보 및 에지 정보를 동시에 고려하여 복합 히스토그램(Composite Histogram)을 생성하여 검색을 하였다. 그리고 Pass 등[22]은 색상 히스토그램의 각 색상 bin에 대해 정제화(Refinement)를 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 나와있다. 또한, 많은 실험적인 데모 시스템이 WWW 상에 존재하고 있는데, MIT 의 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(Query Specification), 메타데이터(Metadata) 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[39-43]. 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 가속화 시킬 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[44, 45].

이러한 배경에서 본 논문에서는 기존의 키워드 기반의 멀티미디어 검색 시스템이 갖는 부정확성이라는 문제를 해결하기 위한 방법을 제시하고 이를 기반으로 웹 기반 멀티미디어 검색 시스템을 구성하고자 한다.

첫째, 기존의 단순한 키워드 기반 방식의 단점을 해결하기 위해 먼저, 단순한 키워드 방식이 아닌 의미기반 키워드 방식을 사용하였다. 이는 단순한 키워드 대신 기술자 정보와 부합되는 정보를 키워드로 입력하게 함으로써 구문론적 검색이 이루어지도록 하였다. 색상 키워드는 규정된 색만을 입력가능 하도록 하여 같은 의미를 갖는 키워드일지라도 검색에서 제외되는 문제를 제거하였다. 또한, 질감, 형태, 객체 등으로 입력을 구분하여 줌으로써 입력 시 좀 더 객관적인 키워드 도출 및 구문 검색이 쉽도록 하였

다.

둘째, 제안된 시스템은 키워드 기반에 필요에 따라 기술자 정보를 포함한 검색이 가능하도록 구성하였다. 기술자는 색상, 질감, 형태를 사용할 수 있도록 하였고 사용자의 요구에 의해 다양한 알고리즘의 추가 또는 삭제가 가능하도록 하였다.

셋째, 제안된 시스템은 키워드 및 각 기술자에 대한 조합을 사용하여 통합 검색 기능을 제공한다. 사용자에 의해 설정된 비율에 의해 각 방법들은 통합 유사도를 도출하여 가장 유사도가 높은 이미지를 검색하게 된다.

마지막으로 모든 이미지 정보는 데이터베이스에 저장되고 검색된 결과는 검색의 효율성 , 고속 검색, 다양한 출력 방식의 제공 등의 기능을 지원하기 위해 XML 데이터로 구성되어 사용자에게 제공된다.

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 본 논문에서 사용한 키워드 및 색상, 질감 기술자를 사용자가 기술자를 선택하고 또한 선택한 기술자의 적용 비율을 설정 가능하도록 설계하였다. 이를 위하여 다양한 기술자를 결합하는 유사도 분석 수식을 적용하였다.

두 번째 단계로 XML를 이용하여 멀티미디어 데이터(MD) 구조를 설계한다. MD 구조에는 본 논문에서 사용한 키워드 및 색상, 질감 기술자에 이외의 MD 데이터 및 차후 데이터 및 알고리즘을 추가하기 쉬운 형태로 설계 하였다.

세 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.

네 번째 단계로 기존의 키워드 방식의 단점을 극복하기 위해 본 논문에서는 키워드를 분류하여 입력하는 방식을 채용하였고 색상과 질감 기술자를 추출하기 위하여 색상 기술자로써는 HSV Segmentation 알고리즘을 질감 기술자로써는 DCT Dominant

Histogram을 사용하였다. 또한, 추후 다양한 알고리즘을 추가 가능하도록 전체적인 시스템을 설계하였다.

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

여섯 번째 단계로 웹상에서 다양한 기술자 및 검색 알고리즘을 테스트하기 위한 툴을 제작하여 추후 개발된 다양한 알고리즘 검증에 사용하도록 하였다.

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 조작(interactive manipulation)을 쉽게 하고 압축된 데이터를 효과적으로 검색할 수 있도록 하는 도구도 필요하게 되었다. 이와 관련된 기술의 표준화가 ISO/IEC(International Organization for Standardization/International Electrotechnical Commission)를 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 기술개발담당 부그룹(Subgroup)과는 독립된 Requirement 부그룹을 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 통신분야의 응용에 대해서도 널리 적용할 수 있도록 하였다. 이러한 범용성을 추구하기 위하여 우선 다양한 응용의 기능과 특성에 대한 요구조건을 수립하는 것으로 시작되었다.

MPEG 뿐만 아니라 ITU-T SG15 ATM 비디오 전문그룹과 공동으로 수행되었고, IYU-R, SMPTE(Society of Motion Picture and Television Engineering) 그룹 등 외부 기관과의 협력을 통해 추진하기도 하였다. 이와 같은 요구조건은 기술담당 부그룹에서 관련된 부호화 요소기술로 변환하며 필요한 기술의 신규개발 및 기존 기술의 변경과 매개변수 선정 등에 이용되었다. 이와 같은 부호화 방식의 표준화가 1994년 11월 싱가포르 회의에서 시스템 부호화, 비디오 부호화, 오디오 부호화의 세부분으로 구분되어 각각 ISO/IEC 13818-1, 2, 3으로 제정되어 IS(International Standard)로 완성되었다. 이 가운데 시스템과 비디오 부호화의 표준은 ITU-T/SG15 WP1과 합동으로 표준화 작업이 이루어져 각각 ITU-T 권고 H.222와 H.262로서 1995년 7월에 승인되었다. 참고로 MPEG-2는 10Mbps까지의 전송률을 갖는 알고리즘으로 출발하였고 MPEG-3가 그 이상의 전송률을 갖는 것으로 정의되었으나 MPEG-2의 전송률에 있어서 10Mbps의 제한이 없어지고 향후 HDTV까지도 포함하는 것으로 변경되었다. 따라서 MPEG-3는 자연스럽게 MPEG-2로 흡수되게 되었다.

제정된 표준에는 64kbps～2Mbps 전송률의 화상전화 및 화상회의 등을 위한 ITU의 H.261, 1.5Mbps급 전송률의 디지털 저장매체를 위한 MPEG-1과 3～15Mbps급 전송률의 디지털TV 및 HDTV 수준의 응용을 위한 MPEG-2 그리고 가장 최근에 추진되어 온 64kbps 이하의 초 저속 전송률의 응용을 위한 단기 표준으로서 ITU의 H.263 등이 있다.

H.263에서는 H.261이나 MPEG-1, MPEG-2의 동화상 압축기법인 움직임 보상과 DCT(Discreet Cosine Transform)의 혼합 기법을 거의 그대로 사용하고 있기 때문에 기존의 H.261, MPEG-1, MPEG-2 시스템에서 고압축했을 때 발생하는 블록화 현상과 mosquito 현상이 심하게 나타나 재생 이미지의 화질이 급격히 열화되는 것으로 알려져 있다. 따라서 초저전송의 응용에 대한 수요가 확산되도록 하기 위해서는 보다 향상된 압축기법을 개발하여 재생이미지의 화질을 개선할 필요가 있는데, 이를 위한 표준화 작업은 1993년 7월경부터 시작하여 MPEG-4 규격을 제시하였다.

MPEG-4는 수요에 의해 개발되었다기보다는 표준화를 이끄는 기술자들의 기술적인 비전에 의하여 주도된 측면이 아주 크며 목표가 고정되어 있다기보다는 끊임없이 기술

의 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(Audio-visual) 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 DVD(Digital Versatile Disk)의 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 Web(WWW)으로 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(A picture is worth a thousand words)”든지, “보기(듣기) 전에는 믿기 어렵다(Seeing(Hearing) is believing)” 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(Metadata)”, 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.

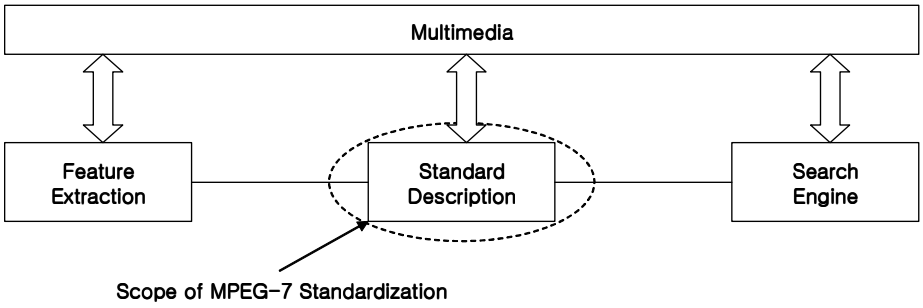


그림 2.1 MPEG-7 표준화의 범위

Fig. 2.1 Scope of MPEG-7 Standardization

이 중 특징 추출(Feature Extraction) 방식 및 검색 엔진(Search Engine)은 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)

MPEG-7에서 정의하는 용어는 다음과 같다. 데이터(Data)는 표현 형식에 상관없이 오디오 비쥬얼 정보를 의미하며, 특징(Feature)은 오디오 비쥬얼 정보의 성격이나 속성을 나타낸다. 기술자(Descriptor)는 특징을 어떤 값에 연결시켜주는 도구이며, 기술 구조(Description Scheme)는 데이터를 여러 개의 기술자로 나타내기 위한 그릇이다. 표 2.1에 이 용어들의 개념을 명확히 하기 위하여 대표적인 특징 형태, 특징 및 기술자의 예를 보였다.

표 2.1 오디오 비쥬얼 데이터의 대표적인 특징 및 기술자의 예

Table 2.1 Example of Dominant Features and Descriptor of Audio Visual Data

특징 형태 (Feature Type)	특징(Features)	기술자(Descriptors)
N-dimensional spatial-temporal structure	duration of a music segment	time code, etc.
	trajectory of objects	chain code, etc.
Statistical Information	color	color histogram, etc.
	audio frequency content	average of frequency components, etc.
Objective features	color of an object	color histogram, etc.
	shape of an object	a set of polygon vertices, a set of moments, etc.
Subject features	emotion(happiness, angry, sadness, etc.)	a set of eigenface parameters, text, etc.
	style	text, etc.
Concepts	event	text, etc.
	activity	text, etc.

한 종류의 기술자나 기술구조가 위에 열거한 모든 응용 분야를 다 만족시키는 것은 어려운 일이며, 효율적이지도 않다. 따라서, MPEG-7에서는 기술자나 기술구조의 제안 사항을 주요 응용 분야 별로 나누어 표준화되었다. 응용 분야와 관련하여 논의가 진행되고 있는 것 중의 하나는 Pull Model과 Push Model의 구분이다. 전자는 국지적으로나 네트워크상에 있는 데이터베이스로부터 사용자의 질의에 가장 가까운 데이터를 찾아주는 응용 분야이고, 후자는 방송에서와 같이 정보 제공자로부터 일방적으로 데이터가 공급될 때, 이로부터 원하는 정보를 필터링하여 주는 응용 분야이다. 두 응용 분야의 주요 특징상 차이점은 후자의 경우 정보 검색 시 실시간성이 강하게 요구된다는 것이다. 그림 2.2에는 MPEG-7의 표준화 요소와 응용 시스템 기능 구성 간의 관계를 보였다.

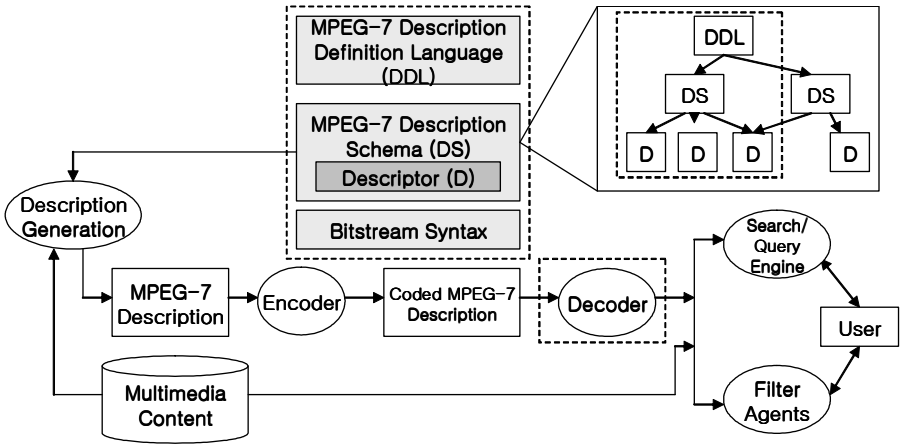


그림 2.2 MPEG-7의 표준화 요소와 응용 시스템 기능 구성

Fig. 2.2 Structure of Standard Elements and Application System of MEPG-7

MPEG-7에 관한 요구 사항은 기술정의 언어에 대한 요구 사항, 오디오 비쥬얼 데이터에 공통적으로 적용되는 요구 사항, 그리고 오디오 또는 비쥬얼 데이터별 특성에 연관된 요구 사항으로 나누어진다.

우선 기술정의 언어에 대한 요구 사항으로는

(1) MPEG-7 응용 분야를 지원하는 어떤 기술구조도 쉽게 생성되어야 한다

(2) 문법이 모호하지 않고 인터프리터 등에 의해 쉽게 파싱(parsing)될 수 있어야 한

다

(3) 오디오 비쥬얼 데이터를 포함한 복합적인 멀티미디어 정보에 대한 기술구조를 생성할 수 있어야 한다

(4) 기술구조 내에 그 코드를 포함시킬 수 있는 기능을 가져야 한다.

이 코드는 표준에서 정의되지 않은 새로운 기술자와 이 기술자에 대한 유사도 측정 기준 등을 사용자가 인식할 수 있도록 하는데 사용된다.

오디오 비쥬얼 데이터에 공통적으로 적용되는 기술자 및 기술구조에 대한 일반적인 주요 요구 사항으로는

(1) 데이터에 부가된 주석, 데이터의 시공간적인 구조, 주/객관적 특징, 또는 사건이나 활동 등의 개념까지도 포함하는 모든 특징의 표현을 지원하여야 한다

(2) 사용자의 요구에 적절히 대응할 수 있는 계층적인 내용 압축 및 표현 방식을 지원하여야 한다

(3) 비쥬얼 질의를 사용해서 오디오 정보를 찾고 오디오 질의를 사용하여 비디오 정보를 찾을 수 있도록 Cross-modality를 지원하여야 한다

(4) MPEG-7의 여러 가지 기술구조 상호간의 호환성을 가져야 한다

(5) 특징별 우선순위를 지정하여야 한다

(6) 기술자의 Scalability를 지원하여 한다

(7) 텍스트 기반 기술자의 경우 사용 언어의 지정, 모든 언어를 지원할 수 있는 글자 세트, 그리고 언어 사이의 변환을 위한 도구 등이 제공되어야 한다

기능적인 관점에서의 요구 사항으로는

(1) 내용 기반 검색 및 유사도 기반 검색 기능의 제공하여야 한다

(2) 데이터와 동기화 되어 제공되거나(Streamed) 또는 데이터와 별도로 제공되는 (Stored) 기술자의 지원하여야 한다

(3) 대화형 질의의 지원하여야 한다

(4) 소스 데이터 또는 관련 정보의 위치를 알려줄 수 있는 기능을 가지고 있어야 한다

(5) 정보의 개략적인 파악(preview)이 가능하도록 브라우징(Browsing) 기능 등을 지원하여야 한다

가 있다. 마지막으로 비쥬얼 또는 오디오 정보 각각의 특성에 따른 요구 사항으로는 지원되는 특징의 종류, 지원되는 데이터 포맷 및 클래스에 대한 요구 사항이 포함된다.

b. 연구 분야

MPEG-7은 멀티미디어 검색의 효율을 높일 수 있도록 다양한 멀티미디어 정보에 대한 특징들을 서술하고 이를 색인화하기 때문에 멀티미디어 콘텐츠에 대한 서술 인터페이스(Multimedia Content Description Interface)라고 한다. 이러한, MPEG-7 표준에 기반한 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[46-50].

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.

- 서술자(Descriptor)들의 조합 방법
- 특징 (Feature) 추출 방법
- 검색(Searching) 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.

이와같은 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, 질감을 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[51].

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

a. 색상모델

(1) *RGB*

RGB 색상 모델은 빛의 삼원색인 빨강(Red), 초록(Green), 파랑(Blue)를 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

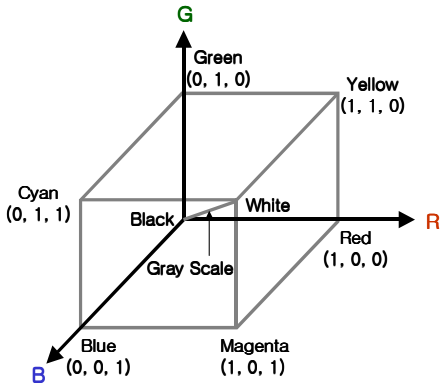


그림 2.3 RGB 색상 모델

Fig. 2.3 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 $0 \leq R, G, B \leq 1$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 밝기 (Brightness)에 관계하고 Cb와 Cr은 이미지의 색상 (Hue) 및 채도 (Saturation)에 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) *HSI*

HSI 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 HSI 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 명암도 (Gray Level)라고 한다.

HSI 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 HSI 모델은 만들 색상의 이름을 이용하여 색을 만드는 직관적인 모델이다.

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

.

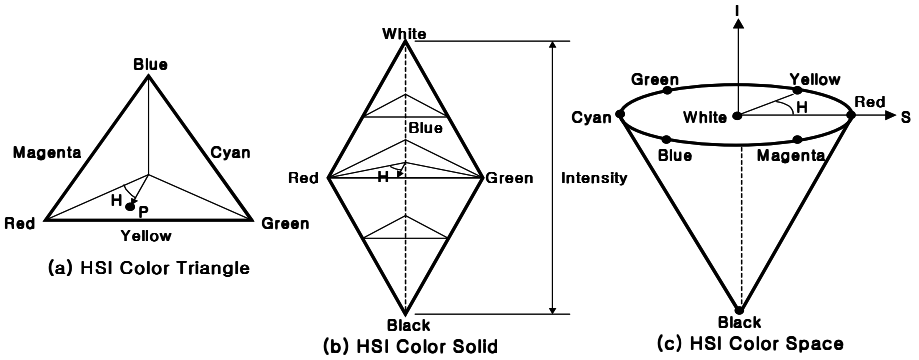


그림 2.4 HSI 컬러 모델

Fig. 2.4 HSI 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 HSI로의 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 히스토그램 (Color Histogram)에 의한 방법과 컬러로 유도된 감각 (Color-Induced Sensations)에 의한 방법이 있다.

컬러 히스토그램은 이미지의 하위 수준 색상 속성을 기술하는 가장 전통적인 방법이

다. 컬러 히스토그램은 이미지의 색상 내용에 대한 매우 효과적인 표현이며, 컬러 히스토그램의 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[52].

그림 2.5과 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.

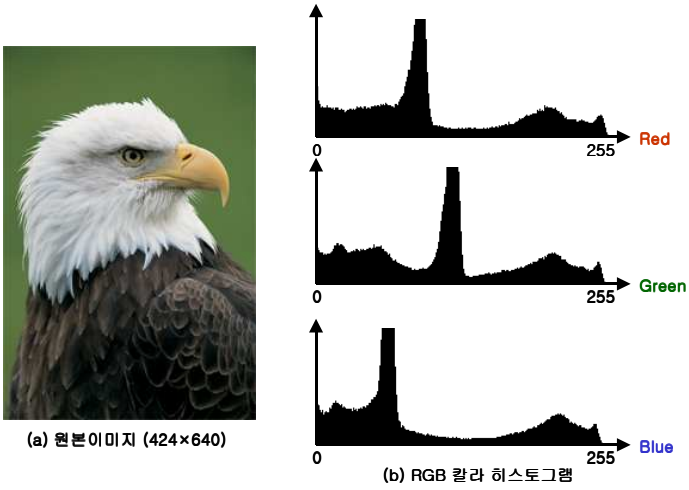


그림 2.5 칼라 히스토그램

Fig. 2.5 Color Histogram

컬러로 유도된 감각이란 색상으로 감각을 표현한 것을 말한다. 예를 들면, 빨간색이나 오렌지색은 따뜻함, 녹색이나 파란색은 차가움으로 표현하고, 비슷한 색상들이 인접해 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[53]. 이것은 하위 수준의 인지적 색상 특징을 관찰자의 직관적 관점이 잘 부합되도록 허용하는 이론으로 Itten[54]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[55].

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 Ballard[8]는 히스토그램 인터섹션(Histogram Intersection)이라는 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 Finlayson[9]은 일정한 색상 색인(Constant Color Indexing)이라는 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[56]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 방법(Distance Method)과 기준 색상표 방법(Reference Color Table Method)을 개발하였다[9]. 거리 방법은 특징으로서 세 가지의 색상 성분의 각각의 1차원 히스토그램의 평균값을 사용하였다. 두 특징 벡터간의 거리가 두 특징 사이의 유사도를 나타내기 위해 사

용된다. 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[57]. 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[8].

이 방법에서는 색상을 나타내는 색상 공간으로 식 (2-3)과 같은 RGB로부터 얻어지는 보색축(Opponent Color Axes)을 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 세기(intensity)축인 wb가 광원으로부터의 거리 및 밝기의 변화 등에 민감하기 때문에 영향을 줄이기 위해 구간을 더 크게 하였다. rg, by 는 16개의 구간으로, wb 는 8개의 구간으로 나누어 히스토그램은 총 2048 (16×16×8)개의 그룹(bin)을 가진다. 히스토그램이 구해지면 두 이미지의 색상 히스토그램은 히스토그램 인터섹션 (Histogram Intersection) 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 정규화시켜서 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 히스토그램(cumulative histogram)을 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.

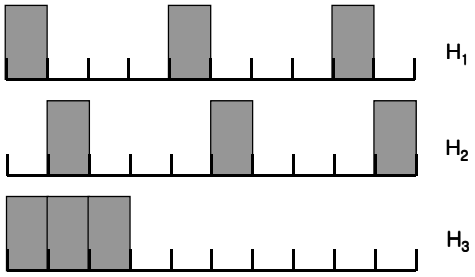


그림 2.6 L_1 거리에 의해 시각적 유사성을 갖지 않는 3개의 히스토그램

Fig. 2.6 Three Color Histogram whose Perceptual Similarities do not Correspond to their L_1 Distances

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

그림 2.6의 히스토그램을 보면 H_1과 H_2가 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.

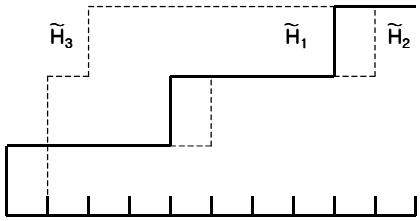


그림 2.7 그림 2.6에 주어진 히스토그램의 누적 히스토그램

Fig. 2.7 Cumulative Histograms of the Histograms Displayed in Fig. 2.6

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

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 분포(Color Distribution)의 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 모멘트(Central Moment)들에 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.

$$E_i = \frac{1}{N}\sum_{j=1}^{n} p_{ij}$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{n}(p_{ij} - E_i)^2\right)^{\frac{1}{2}} \qquad (2\text{-}8)$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{n}(p_{ij} - E_i)^3\right)^{\frac{1}{3}}$$

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, σ_i, ζ_i 그리고 s_i, t_i 라고 하면 두 이미지의 색상 유사도는 식 (2-9)와 같이 얻어진다.

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

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[58], 이미지의 지역적 색상 속성을 특성화하기 위해 컬러 히스토그램을 집합을 제안한 연구[59], VisualSEEK 시스템에서 컬러 이미지의 위치 영역에 사용한 색상 집합에 대한 연구[60] 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 영

역을 구분해내는데 유용하게 사용될 수 있다. 또한, 색상과 함께 질감은 특징을 식별하는데 강력한 기능을 지원한다[61].

질감은 실세계 이미지들의 내용을 묘사하기 위하여 사용되고, 인간 시각에 있어 중요한 요소이다. 색상 특징이 색상 공간 요소의 명암도 분포에 의해 결정되듯이 질감 특징은 연관된 화소들간의 상호 관계를 나타내는 이미지 명암도의 공간적인 분포에 의하여 결정된다[62]. 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)을 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[63].

질감을 인식하고 분별하는 방식은 크게 나누어 통계적인 방법, 구조적 분석 방법, 공간주파수 영역에서의 스펙트럼 분석 방법 등이 있다[64, 65].

통계적 방법으로는 자체 상관 함수, 주파수 공간의 역률 밀도 함수, 단위 면적당 윤곽점수, 수학적 형태학 등 다양하다[66, 67].

구조적 분석 방법은 질감 구조가 규칙적인 경우에 사용되는데 문법 구조에 의한 모델 (Grammatical Model)을 들 수 있다[68].

공간주파수는 화소의 변화율을 의미하는 것으로 이미지 자체를 공간주파수 영역으로 변환할 수 있다. 공간주파수 영역에서의 스펙트럼 분석을 통해 거침의 정도, 방향성 등을 얻을 수 있다[69].

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(Second-order Statistics) 자료를 비교하여 검색을 수행한다. 이러한 방법으로 1978년 Tamura 등은 대비(Contrast), Coarseness, 방향성(Directionality), 규칙성(Regularity), 선형성(Line-likeness) 및 거친 정도(Roughness) 등을 이용하여 질감 정보를 표현하였으며[70], 1996년 Liu와 Picard는 주기성(Periodicity), 방향성(Directionality) 그리고 Randomness 등을 이용하여 질감 정보를 계산하였다[71]. 이들과는 다른 방법으로 Manjunath와 Ma에 의한 Gabor 필터를 이용하는 방법[19], Kaplan 등의 프랙탈을 이용하는 방법[72] 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.

본 논문에서는 질감 특징을 표현하는 가장 대표적인 방법인 Tamura에 의해 제안된 질감을 표현하는 특징들에 대해 알아본다.

Tamura는 질감 이미지의 분석 및 분류를 위해 6가지의 질감 특징들을 정의하고 심리학적인 실험과 비교를 통하여 그 특징의 효과를 입증하였으며 이러한 특징 요소들은 IBM의 QBIC 등 다양한 시스템에서 유용하게 사용되고 있다. 제안된 질감 특징들은 다음과 같다[70].

a. *Coarseness - Coarse vs. Fine*

Coarseness는 가장 기본적인 질감 정보이며 반복되는 규칙적인 패턴의 미세한 정도를 나타낸다. 이미지 $f(x, y)$에 대해 Coarseness를 계산하는 과정은 다음과 같다.

- ***1* 단계** : 이미지내의 모든 픽셀들에 대해 2의 지수의 크기 즉, $1\times1, 2\times2, \ldots 32\times32$의 범위 내의 픽셀들에 대한 평균을 구한다. (x, y) 위치에서, 크기가 $2^k\times2^k$인 범위에 대한 평균은 식 (2-10)과 같다.

$$A_k(x,y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1}\sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i,j)/2^{2k} \tag{2-10}$$

- ***2* 단계** : 각 위치에 대해 수직, 수평 방향으로 평균값들의 차를 계산한다. 예를들면, 수평 방향으로의 차는 식 (2-11)과 같이 계산할 수 있다.

$$E_k(x,y) = A_k|(x+2^{k-1}, y) - (x-2^{k-1}, y)| \tag{2-11}$$

- ***3* 단계** : 각 위치에 대해 다음 식 (2-12)의 조건을 만족시키는 best size를 결정한다.

$$S_{best}(x,y) = 2^k \tag{2-12}$$

여기서, k는 수직 또는 수평 방향 중에서 E를 최대로 만드는 값이다. 즉,

$$E_k = E_{\max} = max\,(E_1, E_2, \ldots, E_L)$$

- ***4* 단계** : Coarseness F_{crs}를 얻기 위해 S_{best}의 평균을 구한다.

$$F_{crs} = \frac{1}{m \times n} \sum_{i}^{m} \sum_{j}^{n} S_{best}(i, j) \tag{2-13}$$

여기서, m, n은 Coarseness 계산 시 이미지의 유효 너비 및 높이이다.

b. 대비 - *High Contrast vs. Low Constrast*

이미지 내 패턴에서의 화소들의 대비 정도를 나타내며 다음과 같은 방법으로 얻어질 수 있다. 먼저 이미지내의 흰색, 검은색의 분포 정도를 고려하기 위해 식 (2-14)와 같이 kurt osis α_4를 계산한다.

$$\alpha_4 = \frac{\mu_4}{\sigma^2} \tag{2-14}$$

여기서 μ_4는 평균에 대한 4차 모멘트이고 σ^2는 분산이다. 여기에서 이미지내의 화소값들의 동적인 범위를 고려하여 대비는 식 (2-15)와 같이 정의된다.

$$F_{con} = \frac{\sigma}{(\alpha_4)^n} \tag{2-15}$$

여기서 n은 양수로서 실험결과 n=1/4 일 때 좋은 결과를 보였다.

c. 방향성

패턴들의 방향성 정도를 나타내며, 각 에지에서의 방향 히스토그램을 이용하여 얻어진다. 즉 수직 및 수평 방향으로 Sobel 연산자를 이용해 얻어진 값을 Δ_H, Δ_V 라고 하면

$$|\Delta G| = (\Delta_H + \Delta_V)/2$$
$$\theta = tan^{-1}(\Delta_V/\Delta_H) + \frac{\pi}{2} \tag{2-16}$$

이 얻어진다. $|\Delta G|$가 일정한 문턱값 이상인 픽셀에 대해 θ를 양자화하여 식 (2-17)과 같은 히스토그램을 만든다.

$$H_D(k) = N_\theta(k)/\sum_{i=0}^{n-1} N_\theta(i),\ i = 1, 2, 3, ..., n-1 \tag{2-17}$$

여기서 $N_\theta(k)$는 $|\Delta G| \geq t$, $(2k-1)\pi/2n \leq \theta < (2k+1)/2n$ 조건을 만족하는 포인트의 수치이다. 위의 $H_D(k)$의 피크들의 Sharpness를 고려하여 식 (2-18)과 같이 방향성을 계산한다.

$$F_{dir} = 1 - r \cdot n_p \sum_{p}^{n_p} \sum_{\phi \in w_p} (\phi - \phi_p)^2 \cdot H_D(\phi) \tag{2-18}$$

여기서 n_p는 피크의 수, ϕ_p는 H_D에서 p번째 피크의 위치, w_p는 p번째 피크의 범위, r은 ϕ의 양자화 레벨에 관한 정규화 계수, ϕ은 양자화된 방향값을 나타낸다.

d. 선형성

선으로 구성되어 있는 패턴의 정도를 말한다. 하나의 에지 방향과 그 주위의 에지들의 방향이 유사할 때 그러한 에지의 그룹을 선으로 간주한다. 선형성을 구하기 위해서 먼저 방향 Co-occurrence 행렬 $P_{Dd}(i, j)$을 구성해야 하는데, 이 행렬의 원소는 두 에지 쌍에 대해서 한 에지 픽셀이 방향 코드 i를 가지고 또 다른 에지 픽셀이 j 방향 코드를 가지며 서로 거리가 d 만큼 떨어져 있는 경우의 빈도수를 나타낸다. 이러한 $n \times n$ 크기의 Co-occurrence 행렬을 이용하여 선형성을 구하는 식은 식 (2-19)와 같다.

$$F_{lin} = \sum_{i}^{n}\sum_{j}^{n} P_{Dd}(i,j) cos\,[(i-j)\frac{2\pi}{n}] / \sum_{i}^{n}\sum_{j}^{n} P_{Dd}(i,j) \qquad (2\text{-}19)$$

e. 규칙성

이미지내의 질감 특성들의 규칙성 정도를 나타낸다. Tamura는 이미지 전체에 걸쳐서 질감 특징들이 변화한다면 불규칙적이라 가정하고 따라서 이미지를 분할하여 각 분할된 부분에 대한 질감 특성을 추출하여 특징들의 변화도를 고려하였다. 규칙성을 정의하는 식은 식 (2-20)과 같다.

$$F_{reg} = 1 - r(\sigma_{crs} + \sigma_{con} + \sigma_{dir} + \sigma_{lin}) \qquad (2\text{-}20)$$

여기서, r은 정규화 계수이며, 각 σ_{xxx}는 각 질감 특징 F_{xxx}의 표준편차이다.

f. 거친정도

시각에 대한 심리학적 실험에 바탕을 두어 질감 특징 중에서 가장 중요한 요소인 Coarseness 및 대비의 영향을 고려하기 위해 식 (2-21)과 같이 거친 정도를 정의하였다.

$$F_{rgh} = F_{crs} + F_{con} \qquad (2\text{-}21)$$

위와 같은 질감 특징들은 현재 QBIC 등의 시스템에서 사용되고 있으며 이미지의 질감 특성으로 위 특징들의 일부 또는 6가지 모두를 추출하여 이미지 간에 서로 비교함으로써 질감 정보 기반 이미지 검색이 가능하다.

3. *Shape*

형태란 이미지의 영역, 에지, 코너, 점 등의 요소들을 의미하는 것으로 모멘트, 에지 이미지 상호관계로서 그 특징을 나타낸다[73]. 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 것

으로서, 직선이 될 수도 있고 특별한 함수의 식으로 나타낼 수도 있다[74, 75].

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[76-78]. 따라서, 형태 정보를 표현하기 위해서는 일반적으로 객체의 크기 변화나 회전에 무관한 객관적인 속성 값을 정의하여야 한다[79]. 그리고 형태 정보에 의한 검색은 모양 데이터가 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 나뉘어질 수 있다[11]. 하나는 윤곽선 기반(Boundary-based) 기법이고 다른 하나는 영역 기반(Region-based) 기법이다. 윤곽선 기반 기법은 이미지 내의 물체의 외곽선에 기반하여 특징을 추출하는 방법으로서 체인 코드(chain code), 퓨리어 묘사(Fourier Descriptor), UNL 퓨리어 특징, 물체의 윤곽선의 반경, 특징점을 이용한 기법 등 다양한 기법들이 존재한다. 영역 기반 기법은 물체 영역의 기하학적인 특징을 추출하는 방법으로서 면적(Area), 신장율(Elongation), 오일러 수(Euler Number), 불변 모멘트(Invariant Moments), 저나이크 다항식을 이용한 저나이크 모멘트(Zernike Moments)를 비롯하여 많은 방법들이 있다.

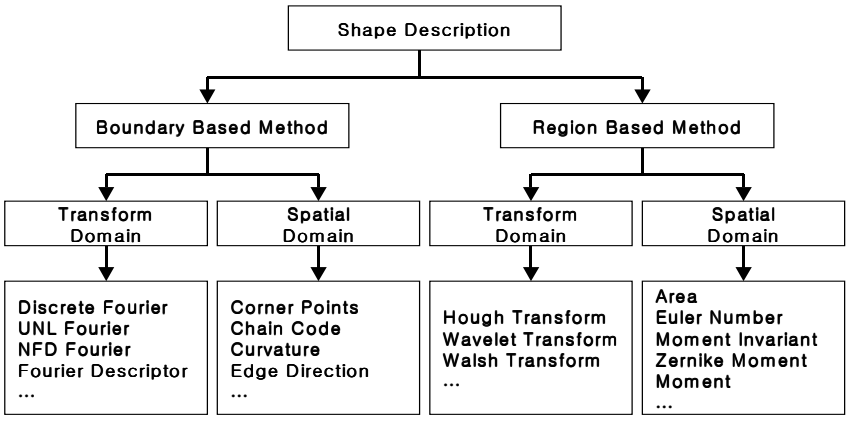


그림 2.8 형태 기반 검색 기법의 분류

Fig. 2.8 Classification of Retrieval Techniques based on Shape Information

본 논문에서는 대표적인 형태 기반 검색 기법인 Jain 등의 에지의 방향 히스토그램을 이용한 방법[16], 저나이크 모멘트를 이용한 방법[12] 및 지역적인 미분 불변치를 이용한 방법들에 대해 알아본다[80, 81].

a. 에지의 방향 히스토그램을 이용한 기법

Jain에 의해 제안된 방법으로 이미지 내의 물체의 에지의 방향 히스토그램을 형태 정보로 이용하였다. 즉, 에지의 방향 히스토그램을 구한 후 히스토그램 인터섹션 방법을 이용하여 각 이미지의 히스토그램을 비교하였다[16].

먼저 Canny 에지 연산자를 이용하여 에지의 위치를 찾은 후 에지 위치에서 그레디언트의 방향을 구하면 에지의 방향 정보를 얻어낼 수 있다. 즉, 픽셀 (m, n)에서 가로 수평 방향으로의 변화량을 $g_h(m,n)$이라 하고, 수직 뱡향으로의 변화량을 $g_v(m,n)$라고 하면 방향은 식 (2-22)을 이용하여 계산할 수 있다.

$$\theta = tan^{-1}\frac{g_v(m,n)}{g_h(m,n)} \tag{2-22}$$

이렇게 에지 위치에서 얻어진 방향들을 이용하여 히스토그램을 생성한 후 앞 절의 식 (2-5)의 히스토그램 인터섹션을 이용하여 유사도를 비교한다.

하지만 에지의 방향을 히스토그램화하는 것은 이미지의 이동에는 불변하지만 축척이나 회전에는 변화한다는 단점을 가지고 있다. 따라서 Jain은 이러한 다음과 같은 방법을 이용하였다. 축척에 대한 문제는 이미지내의 총 에지 수를 이용하여 히스토그램을 정규화함으로써 해결할 수 있다. 하지만 회전에 대한 변화는 축척에 대한 것보다 다소 어렵다. 물체가 회전하면 에지의 방향 히스토그램에서는 이동(Shift)이 일어나게 된다. 하지만 다음과 같은 문제가 발생한다. 즉, 에지 방향의 히스토그램의 그룹이 45° 의 각도로 나뉘어져 있을 경우, 30° , 40° 의 에지는 같은 그룹에 포함되게 된다. 그러나 10° 의 회전이 이루어질 경우 각기 서로 다른 그룹에 속하게 됨으로써 원래의 히스토그램과는 다른 모양의 히스토그램이 만들어지게 된다. 이러한 문제를 해결하기 위해서 히스토그램에 대해 평활화(Smoothing) 연산을 수행하였다. 하지만, 히스토그램을 매끄럽게 만들어도 급격한 회전에는 적용되지 못하는 단점을 가지고 있으므로, 이동 축척의 변화에는 어느 정도 좋은 결과가 나왔지만, 회전의 변화에는 좋지 못한 결과를 보였다.

b. 저나이크 모멘트를 이용한 기법

저나이크 모멘트는 단위원의 내부에서 완전한 직교집합을 형성하는 복소 저나이크 다항식으로부터 유도되며 저나이크 모멘트를 이용한 이미지의 인식에 대한 연구가 Khontanzad에 의해 제안되었다[12]. 저나이크 다항식의 집합은 식 (2-23)의 $V_{nm}(x,y)$에 의해서 나타내어진다.

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)e^{jm\theta} \tag{2-23}$$

여기서, n은 0과 양의 정수이며, m은 $|m| \le n, n-|m| = even$ 조건을 만족하는 정수, ρ, θ은 중심점까지의 거리와 시계 반대방향의 각도이며, $R_{nm}(\rho)$은 방사형 다항식이다.

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|}(-1)^s \frac{(n-s)!}{s!\,(\frac{n+|m|}{2}-s)!\,(\frac{n-|m|}{2}-s)!}\rho^{n-2s} \tag{2-24}$$

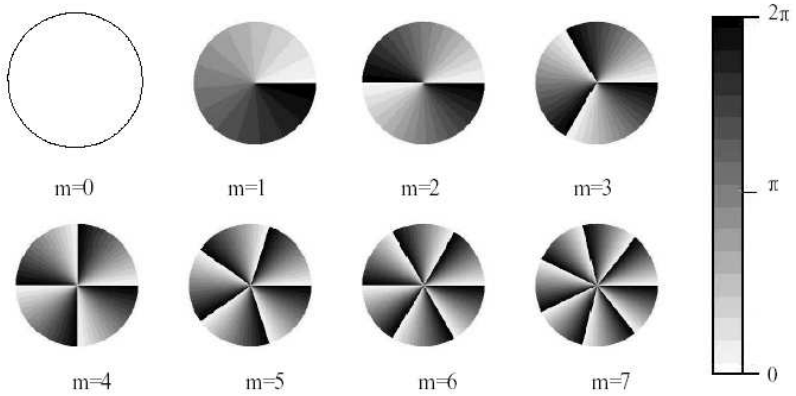


그림 2.9 저나이크 다항식의 위상

Fig. 2.9 The Phase of Zernike Polynomial

그림 2.9는 반복계수 m에 따른 저나이크 다항식의 위상을 나타내며 이로부터 저나

이크 모멘트가 물체의 대칭성 정보를 가지고 있음을 알 수 있다. 즉, m 차에 해당되는 모멘트의 크기가 가장 큰 경우, 원 내부의 물체는 m 차 대칭의 형태를 가진다.

디지털 이미지 $f(x,y)$에 대한 n 차, 반복계수 m의 저나이크 모멘트는 저나이크 다항식을 이용하여 식 (2-25)와 같이 정규 모멘트와 유사하게 정의된다.

$$A_{nm} = \frac{n+1}{\pi} \sum_{x} \sum_{y} f(x,y) V_{nm}^{*}(x,y) \tag{2-25}$$

여기서, $x^2 + y^2 \leq 1$이다.

차수에 따라 생성되는 저나이크 모멘트의 목록은 표 2.2과 같다. 각 이미지에 대해서 이렇게 추출된 모멘트들을 비교함으로서 이미지간의 형태의 유사성을 비교할 수 있다.

저나이크 모멘트의 경우 회전에 의한 모멘트의 크기가 불변이고, 그것들로부터 이미지의 재구성이 쉽다는 특징이 있다. 저나이크 모멘트를 적용하기 위해서는 반드시 크기와 이동에 대한 정규화 과정이 있어야 한다. 즉, 이미지 내의 물체가 영역 $x^2 + y^2 \leq 1$ 내부에 위치하도록 조정되어야 한다. 저나이크 모멘트는 완전한 직교집합으로, 저나이크 모멘트를 이용하여 원 이미지를 복원할 수 있으며 이미지의 기하학적 변형(이동, 회전, 축척)과 조명변화에 매우 강인하고 이미지 전체의 형태 정보를 효과적으로 표현할 수 있다는 장점이 있다. 그러나 이미지 열화에 매우 민감하고 이미지 내부의 세부 픽셀 값은 무시되는 약점이 있다.

표 2.2 저나이크 모멘트의 목록
Table 2.2 The List of Zernike Moments

차수 (n)	*Moments*	차수 (n)	*Moments*
0	A_{00}	7	A_{71}, A_{73}, A_{75}, A_{77}
1	A_{11}	8	A_{80}, A_{82}, A_{84}, A_{86}, A_{88}
2	A_{20}, A_{22}	9	A_{91}, A_{93}, A_{95}, A_{97}, A_{99}
3	A_{31}, A_{33}	10	$A_{10,0}$, $A_{10,2}$, $A_{10,4}$, $A_{10,6}$, $A_{10,8}$, $A_{10,10}$
4	A_{40}, A_{42}, A_{44}	11	$A_{11,1}$, $A_{11,3}$, $A_{11,5}$, $A_{11,7}$, $A_{11,9}$, $A_{11,11}$
5	A_{51}, A_{53}, A_{55}	12	$A_{12,0}$, $A_{12,2}$, $A_{12,4}$, $A_{12,6}$, $A_{12,8}$, $A_{12,10}$, $A_{12,12}$
6	A_{60}, A_{62}, A_{64}, A_{66}	13	$A_{13,1}$, $A_{13,3}$, $A_{13,5}$, $A_{13,7}$, $A_{13,9}$, $A_{13,11}$, $A_{13,13}$

c. 지역적인 미분 불변치를 이용한 기법

이미지에서 많은 정보를 포함하고 있는 부분들 즉, 변화가 큰 부분들을 따로 추출하여 그 부분들에 대한 지역적인 특징을 검색에 이용한다. 이를 위해 이미지에 있어서 특징점으로서 관심점을 정의하고 그 특징점에 대해 미분 불변치를 구함으로서 지역적인 특징을 구하고 각 이미지간의 유사한 지역적인 특징을 가지는 특징점들을 비교함으로써 유사성을 측정할 수 있다.

4. 기술자 설계

a. *XML (eXtensible Markup Language)* 표준

웹의 대중화에 가장 큰 공헌을 한 것 중의 하나가 HTML(Hyper Text Markup Language)이다. 인터넷 붐을 몰고 온 HTML언어는 문서의 외양만을 제어하는 데에는 유용하게 사용될 수 있으며, 쉽게 홈페이지를 만들고 볼 수 있다는 편리함으로 인해 웹의 발전에 크게 기여하였으며, 현재까지도 웹 문서를 저작하는 언어로 보편적으로 사용되고 있다.

그러나 제한된 태그로만 분류되어 지정되지 않은 태그의 사용이 불가능하다거나 단일 링크형으로서 문서들 사이의 다양한 관계를 지원하지 못하여 검색하기가 쉽지 않거나, 구조 기반 설계라기보다 표현 양식 기반 설계이기 때문에 표현 형식의 변환이 어렵다는 등의 한계를 드러내고 있다.

이러한 문제점들을 해결하기 위해서 CSS(Cascading Style Sheets), HTML DOM (Document Object Model), DHTML(Dynamic HTML) 등 많은 대안들이 등장하였으나 근본적인 문제점을 해결하지 못하고 있다. HTML이 대중적으로 보급되기 이전에, 절차적 markup이 아니라 기술적인 markup을 이용함으로써 하드웨어 및 소프트웨어의 종류에 관계없이 이 기종간의 정보 교환이나 검색, 처리를 지원하기 위해 1986년에 문서의 내용이나 구조를 정의할 수 있는 메타언어로써 SGML(Standard Generalized Markup Language)이 국제표준기구(ISO : International Standardization Organization)에서 국제표준(ISO 8879:1986)으로 제정되었다.

SGML은 유연성이 많고 시스템이나 플랫폼에 독립적으로 운용되는 등 많은 장점을 가지고 있지만 사용이 어렵고 DTD 생성이나 이해가 쉽지 않은 등 시스템을 개발하는데 많은 어려움이 있다. 그러나, SGML은 지나치게 복잡하여 관련 소프트웨어의 개발이 쉽지 않으며, 인터넷상에서의 서비스를 위한 목적으로 만들어진 것이 아니기 때문

에 웹상에서 일반적인 기능을 수용하지 못하며, 관련 툴의 가격이 비싸다는 등의 이유로 대중화되지 못했었다.

이와 같은 한계를 뛰어넘으며 등장한 것이 XML이다. XML은 1996년 W3C(World Wide Web Consortium)의 후원으로 형성된 XML Working Group에 의해 개발된 것으로, HTML과 SGML이 갖는 단점은 보완하고 장점은 결합하고 있다.

먼저 SGML은 매우 복잡하기 때문에 전체를 지원하는 소프트웨어의 개발이 용이하지 않은데, 복잡하고 어려운 과정을 대폭 완화시켜 간단하고 매우 융통성이 있는 텍스트 포맷으로 완성시켰다. 또한 구조화된 문서를 정의하고 자유롭게 태그를 정의할 수 있는 등 SGML의 장점은 그대로 취합하고 있다.

HTML은 제한된 태그로만 분류되어 지정되지 않은 태그의 사용이 불가능하던 단점을 보완하여 XML 문서를 만드는 개발자가 임의대로 항목을 정할 수 있도록 하였으며, 인터넷상에서 손쉽게 하이퍼미디어 문서를 제공할 수 있는 HTML의 장점은 그대로 가져가고 있다.

1998년 2월 10일에는W3C 권고에 따르는 XML 1.0[82] 사양이 발표되었다.

표 2.3 HTML, SGML, XML의 비교
Table 2.3 Comparison of HTML, SGML, XML

항목	*HTML*	*SGML*	*XML*
태그	사용자 정의 불가능, 제한적	사용자 정의 가능, 무제한적	사용자 정의 가능, SGML 보다는 제한적
문서의 재사용	불가능	가능	가능
응용분야	홈페이지 제작용, 단순한 컨텐츠 표현	방대한 내용과 구조를 요하는 기술적 문서/전자사전 등	SGML과 동일, 웹상의 거래문서
문서작성	쉽고 간결함, 논리 구조 작성 어려움	복잡하고 어려움, 논리 구조 작성 어려움	SGML을 단순화시켜 편리하게 작성가능
문서검색	효과적 검색 어려움	자료의 표현 구조와 자료의 내용 분리 / 정확한 검색 가능	SGML과 동일
링크	HTML <a> 단순 링크	HyTime	XLL (Xlink, Xpointer)
출력형식 언어	CSS	DSSSL	DSSSL, XSL (XSLT)

XML은 구조화된 문서를 정의하여 태그를 자유롭게 정의할 수 있는 SGML의 장점과 인터넷 상에서 손쉽게 하이퍼미디어 문서를 제공할 수 있는 HTML의 장점을 모두 가질 수 있도록 제안된 웹 표준포맷으로 다음과 같은 설계 목표를 가지고 설계되었다.

- 인터넷을 사용할 수 있어야 한다
- 응용프로그램의 폭 넓은 다양성을 지원하여야 한다
- SGML과 호완성이 있어야 한다
- XML 문서를 처리하는 프로그램을 작성하기 쉬워야 한다
- 문서는 인간이 읽을 수 있어야 하며 합리적이고 명확해야 한다
- 설계는 빠르게 준비되어야 한다
- 설계는 형식에 맞아야 하고 간결해야 한다
- 문서는 만들기 쉬어야 한다

이러한 XML은 다음과 같은 것들을 표현할 수 있다

- 일반문서
- 구조화된 정보 : 규정된 문서, 구조화된 문서 등
- 데이터와 메소드를 가지고 있는 객체들 : Java Applet, ActiveX control 등
- 데이터베이스 관련 : Query 문장 등의 집합
- 웹 관련 메타 데이터
- 화면 표현 방법

XML은 데이터를 설명하는 표준으로서 다음과 같은 이점들을 가져다 준다

- 웹 어플리케이션, 인터넷소프트웨어 개발자에게 XML은 클라이언들과 서버사이의 데이터 전송과 어플리케이션들 사이의 데이터 공유나 디스크에 저장되는 데이터를 위한 강력하고 유연한 포맷을 제공한다
- XML은 self-describing 스키마를 가지고 있어서 구조화된 데이터를 조작하는 애플리케이션 사이의 상호운용을 제공할 수 있다. 예로, Push Publishing에 있어서는 컨텐츠들의 계층적 구조를 쉽게 표현할 수 있는 포맷을 제공한다. 웹 사이트

에 대한 메타데이터를 제공하기 위한 수단으로 손쉽게 사용할 수 있다.

b. *XML* 문서

XML 문서는 문서규칙(DI), 구조, 출력형태로 구성된다. XML 문서의 DTD (Document Type Definition)는 DI (Document Instance)의 형태의 명세이고, 문서의 규칙을 알아낼 수 있다. 즉, DI의 작성 규칙을 정의하는 부분이다. 이 DTD는 그림 2.10과 같이 외부 DTD 서브셋과 내부 DTD 서브셋의 두 부분으로 분류할 수 있는데, 외부 DTD 서브셋은 DTD가 DI와 분리되어 있는 형태로서 일반적으로 타인에 의해서 작성된 DTD이다. 내부 DTD 서브셋은 DI를 포함하고 있는 형태이다. XML 문서의 DI는 XML 문서의 실체로서 사용자가 작성하는 본문에 해당하고 DTD에서 정의한 규칙에 따라서 작성되어진다.

출력형태는 문서의 출력 형식을 지정하는 부분이다. 즉, 문서가 어떻게 보여질 것인지를 정의하는 부분이다. XML을 위해서 제안된 XSL (eXtensible Style Language)는 DSSSL (Document Style Semantics and Specification)과 CSS (Cascading Style Sheet)의 장점을 합친 것이라고 할 수 있다. CSS는 단순한 구조의 XML 문서를 출력하는데 사용되고, XSL은 좀 더 복잡한 구조의 문서와 더 강력한 형식의 출력이 가능하다.

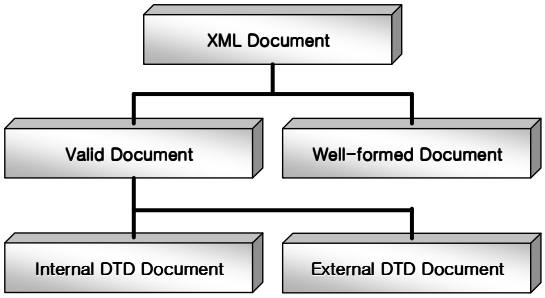


그림 2.10 XML 문서의 구조

Fig. 2.10 Structure of XML Document

모든 XML문서는 논리적 구조와 물리적 구조로 구성된다. 물리적 구조는 개체(Entity)라는 요소들로 이루어지고, 논리적 구조는 문서의 전체적인 구조를 표현하는 부분으로 요소, 선언, 속성 등의 구성 요소를 이용하여 표현한다. 이 두 구조는 올바르

게 혼합되어 XML 문서의 구조를 나타내는 DTD를 형성한다.

그림 2.11은 XML 문서가 처리되는 과정을 표현한 것이다. 일반 텍스트 문서가 의미 없는 내용으로 구성된 것에 반해 위에서 설명한 것처럼 XML 문서는 내용과 구조를 포함하고 있다. 이러한 XML 문서는 포맷팅 출력 규격에 따라 다양한 형태로 출력을 가질 수 있다.

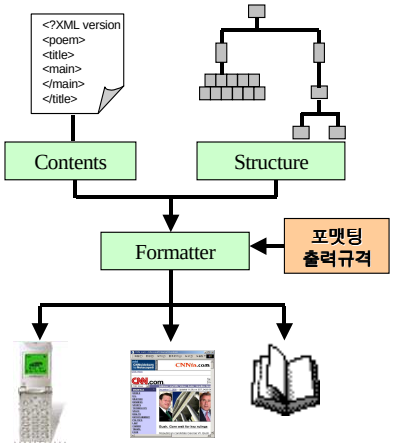


그림 2.11 XML 문서의 처리

Fig. 2.11 Process of XML Document

(1) 문서 규칙 (Document Instance, DI)

문서의 실제적인 부분인 DI는 태그와 문자열로 구성되어 있는데 DI가 어떻게 작성되었는가에 따라서 XML 문서는 두 가지 형태로 분류된다. 이 중 유효한 문서 (Valid Document)는 DI가 정의된 DTD에 맞게 작성된 경우의 문서를 의미한다. 즉, DTD에서 정의된 문서의 구조에 맞게 DI를 작성한 경우를 의미한다. 이는 XML 문서 작성에 관한 전문가나 능숙한 사용자 또는, XML 문서의 작성을 지원하는 편집기로 작성될 수 있는 형태이다. 따라서, 문서의 전체 구성은 알고 있으나, DTD를 분석 및 설계할 수 없는 사용자들에게는 적합하지 못한 문서의 형태이나 검색의 대상이 되는 문서나 중요도가 높은 문서, 같은 DTD를 여러 사람이 사용하는 경우에 적합하다. 또 다른 형태의 문서인 문법에 맞는 문서 (Well-formed Document) XML 문서를 작성하려고 할 때 DTD에 대한 사전 지식이 없는 사용자가 DTD 없이 또는 DTD를 무시하고 사용할

수 있는 문서의 형태이다. 이 형태의 문서는 DI 자체의 문서 오류만을 검증하게 된다.

(a) Well-formed XML 문서

문법상에는 오류가 없지만 문서와 관련된 DTD에 대해서는 전체 논리 구조가 유효하지 않은 문서를 의미한다. 이 개념이 유효한 이유는 DTD를 읽는 데에는 많은 시간과 대역폭이 필요하기 때문이다. 문서가 문법적으로 유효하며 손상되지 않았다는 것을 미리 보장한다면 많은 시간을 절약할 수 있게 된다. Well-formed 문서는 다음의 조건을 갖추어야 한다.

- 전체적으로 문서의 양식과 맞아야 한다.
  (1) 하나 또는 그 이상의 요소들을 포함해야 한다.
  (2) 최상위에는 root 또는 document element라는 하나의 요소만 표시해야 한다.
- 스펙에 서술된 모든 규약을 만족해야 한다.
- 문서 내에서 직간접적으로 참조되는 개체 파싱이 잘 구성되어야 한다.

(b) Valid XML 문서

Valid 문서는 반드시 DTD를 구성해야 할 뿐만 아니라 각각의 요소들은 DTD에 정의된 규칙을 따라야 하는 등 모든 규칙을 반드시 준수해야 한다. 다음으로 어떤 문서 구조로 작성할 것인지 DTD를 정의하고 나면, 구체적인 문서의 내용을 쓰게 되는데, 여기서부터는 HTML과 동일하게 작성하는데 태그를 정의할 수 있다는 것이 다른 점이다. 이 때 영문뿐만 아니라, 한글로도 태그를 사용할 수가 있다.

(2) 문서 구조 *(Document Type Definition, DTD)*

DTD는 Document Type Definition의 약자로 XML Document Type을 기술하는 표준화된 방법으로 XML 문서에서 사용되는 마크업을 선언한 곳이 DTD라고 할 수 있으며 문서에 대한 논리, 물리 구조를 정의한다. 즉, DTD는 XML 문서 안에 결합되거나 외부에 독립적으로 존재할 수 있는 XML 문서이다. DTD에는 XML 문서가 어떻게 구조화되고 어떤 요소를 포함해야하며 어떤 종류의 데이터가 포함되어야함을 정해주고 기본 값이 무엇인지 등을 규정하는 규칙을 정의하며 다음과 같은 사항을 정의하게 된다.

- 내용 모델과 더불어 문서에서 허용되는 요소 형태를 정의한다.
- 각 요소에 할당되어 있는 속성을 정의한다.
- 문서에 허용되는 개체를 정의한다.
- 외부 개체와 함께 사용되는 표기법을 정의한다.

위의 개념에서 보듯이 모든 애플리케이션이 공유하는 XML 데이터의 설명을 XML Vocabulary라 하고 이것을 공유하는 XML 도큐먼트들의 모임을 Document Type이라고 한다. 즉 DTD란 이러한 동일 XML Vocabulary를 공유하는 XML 도큐먼트들의 형태를 정의한 것이라고 할 수 있다. 이러한 DTD의 기능은 다음과 같다.

- XML Vocabulary에 대한 견고한 정의를 가능케 한다.
- XML Document의 구조에 대한 기술을 공유할 수 있다.
- Instance의 Validation을 검증할 수 있다.
- Document Instance마다 하나만의 DTD를 가진다.

프로그래밍 언어와 비교하면 DTD는 언어의 정의와 같다. 따라서 문서의 구조를 정의하고, DTD의 정의에 따라서 DI가 작성되어야 유효한 문서가 된다. DTD의 구성은 요소(Element), 속성(ATTLIST), 개체(Entity)가 있는데, 요소는 가장 상위 요소로부터 여러 개의 요소에 대한 속성들을 정의하는 것이며, 개체는 보통 프로그래밍 언어에서 사용되는 매트로 기능과 흡사하다. 정의와 참조되는 형태에 따라 일반개체 (Genneral Entity)와 매개변수 개체 (Parameter Entity)로 나누어지는데, 일반개체는 DTD에서 정의되고 DI에서 참조되며, 매개변수 개체는 DTD에서 정의되고 DTD에서 참조된다. 그 이외에 선택된 부분 (Marked Section)과 처리 명령 (Processing Instruction)이 있는데, 선택된 부분은 문서의 일부 선택된 부분을 특별히 처리할 때 사용되고, 처리명령은 문서에서 외부 프로그램을 사용하여 처리할 수 있도록 한다. 이와 같은 구성을 갖는 DTD를 파싱하기 위해서는 DTD의 문법을 검사하고, 문서의 구조적인 정보를 이용하여 DI 문서 규칙을 생성해야 한다. 이 과정에서 생성되는 문서의 구조를 BNF (Backus-Naur Form) 형태의 문법으로 생성해서 파서 생성기의 입력으로 사용할 경우에는 문법에 대한 오류와 모호성이 발생할 수 있기 때문에 오류와 모호성을 배제하기

위해서 트리 형태의 DI 문서 규칙을 검사해야 한다.

(3) 출력형태 (eXtensible Style Language, XSL)

XML 문서에 대한 사양이 발표된 후에 XML 문서를 브라우징하기 위해서 문서에 대한 포맷을 정의하는 스타일시트로 XSL이 제안되었다. XML 문서 이외에 SGML과 HTML를 위한 스타일시트가 표준으로 정의되어 있는데, XSL은 SGML을 위한 스타일시트 표준인 DSSSL을 기반으로 하고 있고, HTML을 위한 스타일시트 표준인 CSS로 변환이 가능하며 더 강력한 기능을 갖는다. 이러한 XSL은 다음과 같은 설계 원칙으로 설계되었다.

- XSL은 인터넷상에서 직접 사용할 수 있도록 설계하였다
- XSL은 XML 문법을 기반으로 한다
- XSL은 모든 포맷팅 작업을 수행하기 위한 선언적인 언어이다
- XSL은 DSSSL의 부분집합이다
- CSS를 XSL로 변환이 가능하도록 설계되었다
- XSL에서 사용하는 선택적 특성의 수를 최소화시켰다
- XSL은 사람이 판독할 수 있는 형태이다
- XSL 설계는 신속하게 이루어져야 한다
- XSL은 이용자들이 쉽게 만들 수 있어야 한다

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[85].

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[86].

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[87, 88]. 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[89].

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[90].

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[91]과 구글 검색 엔진[92] 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[93, 94]. 예를 들면, 원시적 의미론에 근거한 검색 시스템에는 사람의 얼굴 식별을 위한 눈, 입, 코, 이마가 있는 얼굴의 요소를 특징화하여 모델화하거나 전역적인 얼굴형을 데이터화한 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[95].

3. 내용 기반 검색 시스템

a. 내용 기반 이미지 검색

내용 기반 이미지 검색 시스템(Content-based Retrieval)은 1992년 Kato[96]에 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[97]. 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[98, 99].

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.

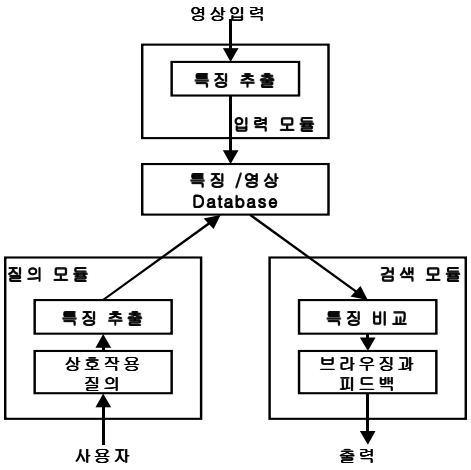


그림 3.1. 내용 기반 검색 시스템의 구조

Fig. 3.1 Structure of the Content-based Retrieval System

이미지의 입력 과정에서는 입력된 이미지를 가장 잘 표현할 수 있는 특징(Feature)을 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 이

미지의 특징과 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 구문적(semantic)이며, 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.

먼저, QBIC (Query by Image Content) 시스템은 IBM의 Niblack 등이 만든 시스템으로 정지 이미지를 위한 내용기반 검색과 동이미지를 위한 내용 기반 검색 및 시각화를 위한 목적으로 만들어진 완벽한 플레임 워크이며 거리 측정에 있어 가중치를 적용한 유클리디안 거리 (Weighted Euclidean Distance)를 사용하여 컬러 히스토그램의 유사도를 측정하여 검색하는 시스템으로 칼라, 형태, 질감 등의 내용을 기반으로 하는 이미지 검색 및 제한된 범위 내의 비디오 검색이 가능한 시스템이다. QBIC 시스템은 그림 3.2과 같이 사용자에게 이미지의 칼라나 형태, 질감 등과 같은 다양한 속성을 기반으로 하는 시각적 질의를 제공한다. 시각적 질의란 사용자가 이미지 검색을 위하여 실제 이미지나 스케치 이미지 등의 시각적 요소를 이용하여 질의하는 것을 의미한다.

형태를 이용하는 경우에는 유클리디안 거리를 유사도 계산으로 사용하며 사용자가 찾고자 하는 이미지를 스케치하면 그것과 유사한 이미지를 찾아주는 기능도 제공된다. 그리고 사용자가 이미지에 대하여 키워드도 주석을 붙일 수 있도록 허용하여 텍스트를 기반으로 하는 검색 기능도 제공한다.

QBIC 시스템은 상표 검색이나 우표 검색과 같은 특정 응용에만 적합한 시스템으로 질의 인터페이스가 복잡하고 체계적이지 않아서 오히려 전문가가 아닌 일반 사용자 측면에서는 이용하기 어렵다. 그리고 키워드를 입력하려면 부분적인 수작업 단계를 거쳐야 한다. QBIC 시스템은 현재 웹상에서 서비스를 제공하고 있으면 지속적으로 발전하고 있다.

RGB 컬러 공간은 256 색상의 집합으로 축소하고, 계산을 용이하게 하기 위해 히스토그램을 정규화 한다. 정지 이미지에 대해서는 의미론적 내용에 의한 질의, 광범위한 색상 유사성, 색상 영역, 질감 유사성 형태와 공간 관계 유사성에 의한 질의가 있다.

동이미지에 대해서는 자동분할 및 단축키 프레임 추출이 있는데 정적인 예들을 사용하는 내용에 의한 질의, 스크립트와 캡션들을 사용하는 내용에 의한 질의 등이 있다[100].

그림 3.2 QBIC 시스템

Fig. 3.2 QBIC System

컬럼비아 대학 센터의 통신 연구소에서 개발된 Smith와 Chang의 VisualSEEK 시스템은 컬러 이미지 내의 지역 영역에 대한 컬러 집합을 사용하였다.

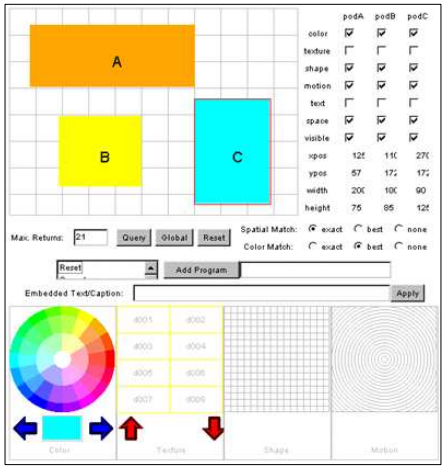


그림 3.3 VisualSEEK 시스템

Fig. 3.3 VisualSEEK System

히스토그램과는 다른, 색상의 총계를 타나내는 컬러 집합은 이진 벡터로 색상의 선택에 일치된다. 이는 이미지의 두드러진 영역들이 단진 몇 가지의 도미넌트 색상을 갖는 것을 가정한다. 이 시스템은 정지 이미지와 동이미지에 ego 색상 및 직감 특징에 대한 검색을 지원하며 통합된 특징 맵 내에서 색상과 질감의 조합이 가능하다.

Chabot 시스템은 그림 3.4와 같이 버클리 공대에서 개발한 관계형 데이터베이스 시스템인 POSTGRESS를 기반으로 개발된 내용기반 이미지 검색 시스템으로 텍스트와 칼라 특성을 사용한다. 이 시스템은 캘리포니아 수자원 관리 부서에서 소장하고 있는 이미지를 저장하고 검색하기 위해 개발되었다[101].

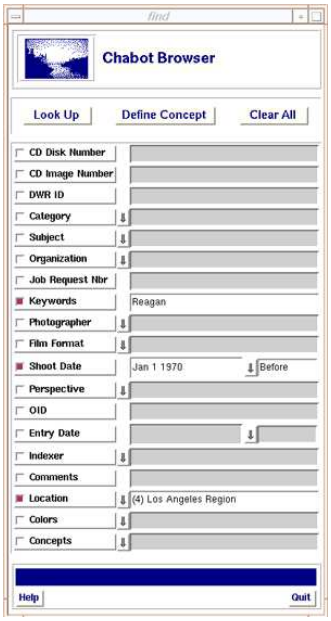


그림 3.4 Chabot 시스템

Fig. 3.4 Chabot System

이 시스템에서 하나의 이미지는 고해상도부터 저해상도까지 다섯 단계의 이미지 형태로 서비스되므로 방대한 저장 용량을 필요로 한다. 현재는 Cypress 시스템이라고 부르고 있다. 그림 3.4와 같이 질의 인터페이스에서는 일반적인 서지 정보 외에 지역, 칼라, 개념 들을 이용하여 검색할 수 있다. 이 시스템의 특징은 이미지에 대한 히스토그램을 구한 후, 칼라에 대해 적절한 주석을 기술하여 이미지를 검색한다. 주석을 이용하여 기존의 관계형 데이터베이스 시스템에 있는 질의어와 연결하여 검색할 수 있는 장

점도 있다. 그러나 내용기반 검색의 측면에서 보면 칼라 특징 정보만을 이용하는 것이 단점이다.

Photobook 시스템은 그림 3.5와 같이 MIT 공대에서 개발한 내용기반 이미지 검색 시스템으로 형태, 질감 특징 등을 기반으로 하고 있다. 이 시스템은 이미지의 통계적인 성질에 기반을 둔 KL(Karhunen-Loeve) 변환을 사용하여 이미지를 몇 개의 주성분 값으로 표현한다. 이미지를 공간으로 변환하기 위하여 이미지의 벡터로부터 구한 공분한 행렬의 고유 벡터를 사용한다. 이 방법은 이미지 식별에 필요한 성분만을 추출하여 압출할 수 있고 다시 원래의 이미지로 복원할 수 있는 기능을 가지고 있다. 이 시스템을 얼굴 인식 분야 등의 특정 분야에 사용되며, 기존이 얼굴 윤곽선 등의 특징점들을 통한 매칭 방법과는 달리 얼굴의 표정 등 얼굴에 약간의 변형을 주어도 같은 얼굴을 찾아낼 수 있다.

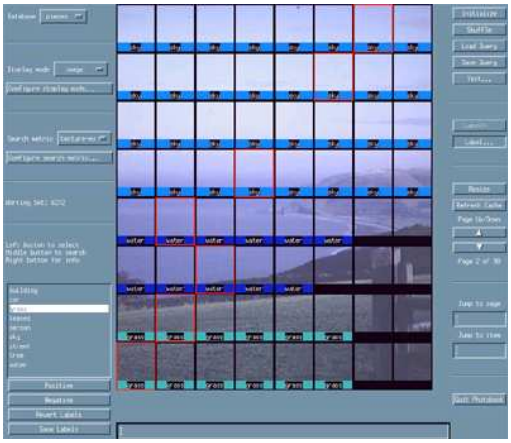


그림 3.5 Photobook 시스템

Fig. 3.5 Photobook System

SIMPLIcity (Semantics-sensitive Integrated Matching for Picture LIbraries) 시스템은 스탠포드 대학에서 개발한 내용기반 이미지 검색 시스템으로 의미 분류와 웨이블릿 변환을 이용한 IRM(Integrated Region Matching) 등의 기법을 사용한다. 이 시스템은 기존 웨이블릿 기반의 시스템보다 정확도, 검색 속도 면에서 향상되었다. 그러나 이미지의 의미 분류 시에 하나의 이미지가 동시에 여러 개의 의미로 분류되는 현상이 발행하는 단점이 있다. 그리고 그림 3.6와 같이 질의 인터페이스 기능도 보다 다양한

사용자의 요구에 부응하기 위하여 보완이 필요하다.

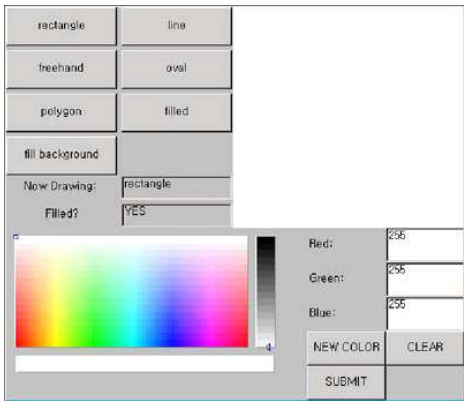


그림 3.6 SIMPLIcity 시스템

Fig. 3.6 SIMPLIcity System

표 3.1은 위의 시스템 및 기타 시스템을 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(Primitive) 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(Logical) 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(Contrast), 방향성(Directionality), Coarseness, Roughness 등이 이에 해당된다. 두 이미지간에 특징을 비교할 때 같은 지역에 해당하

는 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(Geographic Information System : GIS), 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.

B. 웹 기반 이미지 검색 시스템

기존의 웹 이미지 검색 시스템은 두 종류로 나눌 수 있다.

먼저 문자 기반 검색을 사용하는 엠파스 검색 엔진은 웹상의 이미지를 대상으로 문자 기반 검색을 하며 검색 결과로 이미지의 URL을 출력하는 시스템이다. 검색된 이미지의 URL 중에는 유효한 것 외에 그렇지 못한 것, 예를 들면 폐쇄된 사이트의 이미지 URL이 포함된 경우도 있다. 또한, 검색된 이미지 URL은 다르지만 내용은 동일한 이미지인 경우도 있다. 이러한 모든 것이 이미지의 URL 형태로만 검색결과가 출력되기 때문에 사용자는 찾고자 하는 이미지에 대해 막연히 제시된 URL을 보고 추출하여 선택할 수밖에 없으며 여러 번의 시행착오를 거쳐야만 원하는 이미지를 얻을 수 있게 되어 있다. 또한, 엠파스 이미지 검색 엔진은 문자 기반 검색 방법을 사용하고 있으며 검

색 결과의 신뢰도가 낮다는 단점이 있다.

다음은 서버의 이미지 데이터베이스를 이용한 검색을 웹 인터페이스를 통해 지원해 주는 시스템으로 Viper가 있다. Viper 시스템은 특징 그룹을 색상 레이아웃, 전역적 색상, 질감 레이아웃, 전역적 질감으로 나누어 사용자가 선택한 특징에 의한 내용 기반 검색을 지원하고 있다. 그러나 이 시스템은 웹상의 이미지를 대상으로 하는 것이 아니라 자체적으로 구축한 이미지 데이터베이스를 이용한 검색이기 때문에 사용자의 요구를 만족시키지 못한다.

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 칼라와 질감 기술자의 조합만을 사용하고자 한다. 본 논문에서 테스트하고자하는 이미지 데이터가 일반적인 웹상에 존재하는 이미지이므로 형태나 움직임 기술자들은 제외하였다.

1. 색상 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.

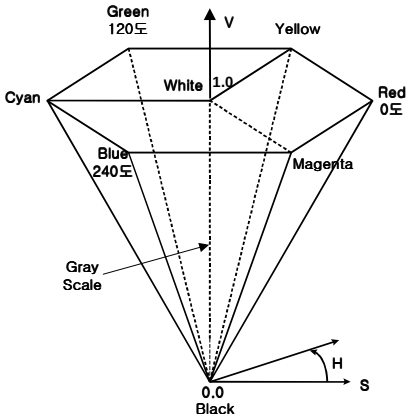


그림 4.1 HSV 칼라 공간

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. 대표 색상 검색

유사한 색상을 포함하는 이미지 내의 영역을 분리해내기 위해 회귀적인 HSV 색상 공간에서의 분할 방법을 적용한다. HSV 공간에서의 분할은 이미 그 성능이 검증되었으며 또한 빠르고 자동적으로 수행될 수 있기 때문에 RGB 공간대신 이용되었다.

이 방법은 클러스터링 기법이나 영역 성장(Region Growing) 기법에서와는 달리 초기 시드(Seed) 픽셀이나 색상 수의 선택과 같은 변수들에 독립적이므로 사람의 개입 없이 과정이 수행될 수 있다.

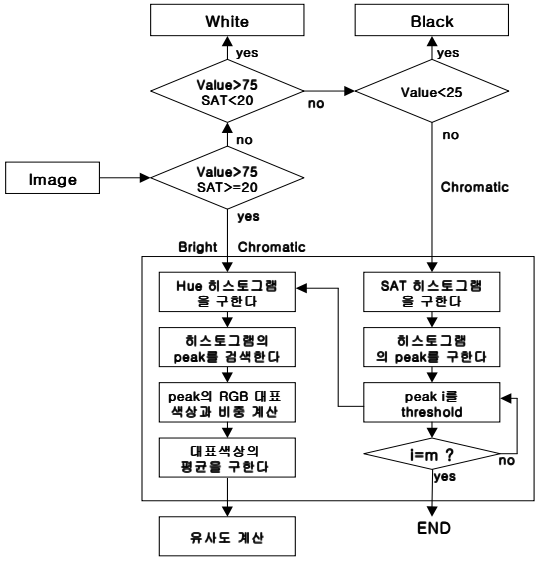


그림 4.2 대표 색상 검색 과정

Fig. 4.2 Procedure of Representation Color Retrieval

HSV 공간에서는 유사한 Hue 방향에 놓이면 유사한 색상으로 분류된다. 이러한 성질을 이용하여 이 방법에서는 그림 4.2의 알고리즘에서와 같이 Saturation과 Value 정보를 고려하면서, 색상 정보의 대부분을 가지고 있는 Hue의 히스토그램을 Thresholding 한다. 첫 번째 단계는 모든 Bright Chromatic 픽셀에 대한 Hue 히스토그램을 생성한다. Bright Chromatic 픽셀은 Value>75%, Saturation>=20%를 가지는 픽셀로 정의하였다. 일단 이러한 조건을 만족하는 픽셀들이 확인되면 Hue 히스토그램이 생성되고 Hue 히스토그램은 m개의 Bright Color들로 Thresholding 된다. 여기서

m개의 색상은 Hue 히스토그램이 보여주는 피크의 수에 의해 결정되는 이미지에 종속적인 수이다.

남아있는 이미지 픽셀에 대해 Saturation과 Value가 Achromatic 영역을 결정하기 위해 사용된다. 특히 실험적으로 Value<25%를 만족하는 색상은 검정색으로, Saturation<20%, Value>75%의 색상은 흰색으로 분류되는 것으로 알려져 있다.

마지막으로 남아있는 모든 픽셀들은 chromatic 영역에 해당된다. 이 영역의 색상에는 넓은 범위의 Saturation 값이 존재할 수 있다. 이것을 고려하기 위해서 모든 Chromatic 픽셀들에 대해서 Saturation 히스토그램을 생성하였다. 히스토그램이 p개의 피크를 가진다고 가정하면, p개의 피크를 Thresholding 하여 각각의 피크에 포함된 픽셀들에 대해서 다시 Hue 히스토그램을 생성한다. 다시 각각의 Hue 히스토그램의 피크에서 대표 색상을 추출한다. 이 과정은 p개의 Saturation 피크 각각에 대해서 반복된다. 이 색상분할 과정을 거치면 각 이미지에서 대표 색상을 얻을 수 있으며 각 영역에서 RGB 성분의 평균값을 계산하여 그 영역의 대표색으로 선정한다. 최종적으로 얻어진 색상 정보는 총 색상 수 및 각 영역의 대표색 그리고 이미지에 있어서 각 색상이 차지하는 비율 등이 얻어지며 이 정보가 색상정보 기반 검색에서 사용되어진다.

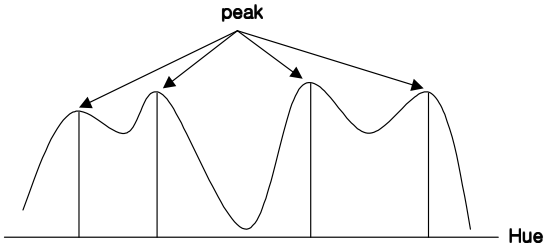


그림 4.3 Hue 히스토그램에서의 peak

Fig. 4.3 Peak in Hue Histogram

c. 유사도 계산

위의 색상 분할 과정을 거치게 되면 각 이미지에 대한 대표색들이 추출된다. 이러한 색상 정보를 이용하여 이미지 검색을 수행하기 위해 색상 정보를 비교하는 과정이 요구되며 이 과정이 검색 성능을 좌우하는 중요한 부분이다. 본 논문에서는 대표 색상들의 거리 비교하기 위하여 다음과 같은 방법을 사용하였다.

먼저 이미지 I_n과 I_m에 대해 색상 분할을 통하여 얻어진 대표 색상의 수를 각각 N_n,

N_m 이라 하면, 각 이미지에서 추출된 대표색과 그 색상의 비율은 식 (4-2)와 같이 표현할 수 있다.

$$I_n\text{의 대표색 : } C(I_n) = \{C_i\} = \{R_i, G_i, B_i, r_i\}$$
$$I_m\text{의 대표색 : } C(I_m) = \{C_i\} = \{R_i, G_i, B_i, r_i\} \qquad (4\text{-}2)$$

여기서, R_i는 i 번째 색상의 Red 성분, G_i는 i 번째 색상의 Green 성분, B_i는 i 번째 색상의 Blue 성분이고 r_i는 이미지 내에서 i 번째 색상이 차지하는 비율을 나타낸다.

위에서 구해진 대표 색상과 비율을 이용하여 두 이미지의 유사도를 식 (4-3)을 이용하여 구하였다.

$$Similarity = \frac{\sum_{i}^{m}(\sum_{j}^{n} avg(I_n(R_{ij}, G_{ij}, B_{ij}))) \times I_n(r_i)}{\sum_{i}^{m}(\sum_{j}^{n} avg(I_m(R_{ij}, G_{ij}, B_{ij}))) \times I_m(r_i)} \qquad (4\text{-}3)$$

2. 질감 기반 검색 시스템 설계

a. *DCT Peak* 히스토그램

일반적으로 DCT 변환은 MPEG 압축 표준에서 핵심이 되는 기술이다. DCT 변환의 장점은 변환 전에는 화면에 불규칙하게 퍼져 있던 화소 값이 변환 후에는 저주파 항으로 집중되는 경향이 있다는 것이다[102]. 따라서 고주파 항을 버리는 조작 (Removing Redundant Data)을 통해 정보손실이 거의 없이도 정보압축이 가능하다. 왜냐하면 이미지를 구성하는 대부분이 저주파 성분으로 구성되어 있고 고주파 성분은 적게 분포하고 있기 때문이다.

이러한 DCT 변환은 실수부인 코사인 함수만 남는 특성을 제외하면 Fourier 변환과 유사한 형태의 결과를 얻는다.

2D DCT의 경우는 1D DCT를 확장한 것이므로 Basis Cosine Waveform이 64개가 필요하다. 다만 수직과 수평축으로 분리되는 것이 다르다. Basis f_n은 맨 좌측상단에

위치하게 된다. 2차원 이미지 공간상의 영역은 1차원 주파수 영역으로 변환하여 계산 처리 되는데 이때 지그재그 패턴을 이용하여 1차원의 배열 형태로 변환 작업한다.

그림 4.4는 64개의 Cosine Basis Function 집합을 보여준다. 이는 1D 8-point Basis Function의 집합에서 출발한 수평성분을 같은 Basis Function을 갖는 수직 성분의 집합을 곱함으로써 생성된 것이다.

수평 Basis Function의 DC 성분은 왼쪽에 위치하며, 수직의 DC성분은 맨 위쪽에 위치하므로 전체 DC 성분은 맨 왼쪽의 위에 위치하게 되는데 흰 부분에 +진폭을 검정부분에 -진폭을 주게 될 경우 DC 성분은 회색이 된다. 즉 평균값이 회색이 되는 것이다.

64개의 계수값이 적절히 스케일되면 어떠한 64개의 샘플값들이라도 8×8 block의 샘플로서 표현될 수 있다. 저주파에서 점차적으로 고주파 항으로 basis 함수를 나타내기 위하여 지그재그 패턴을 사용한다. DC 성분이 있는 맨 좌측상단에서 시작하여 지그재그로 내려가다 보면 맨 좌측하단의 위치가 가장 높은 주파수를 나타내게 된다.

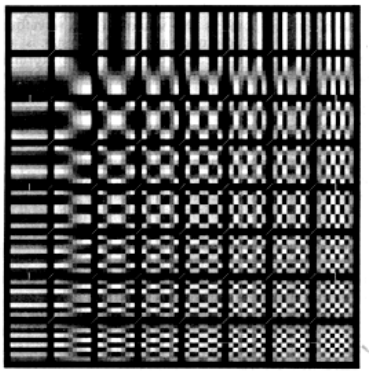


그림 4.4 2차원 DCT 코사인 기초 함수

Fig. 4.4 2D DCT Cosine Basis Functions

식(4-4)는 2차원 DCT 변환을 나타낸 것이다.

$$F(u,v) = \frac{1}{4} C(u) C(v) \sum_{i=0}^{7} \sum_{j=0}^{7} f(i,j) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16} \qquad (4\text{-}4)$$

본 논문에서는 DCT 변환을 실행하고 DCT 결과에서 저주파 성분을 64X64 블록의 정보만을 추출하여 히스토그램을 구하고 이 중에서 Peak를 검사하여 일부의 값들만을

사용함으로써 데이터량을 감소시켜 실행 시간을 줄일 가능성을 높였다. 그림 4.5는 본 논문에서 사용한 알고리즘을 보여주고 있다. 차후 형태 기반 분석에서는 고주파 성분 부분을 사용함으로써 한번 DCT 변환을 통해서 이미지의 질감특성과 형태특성을 모두 분석할 수도 있을 것이다.

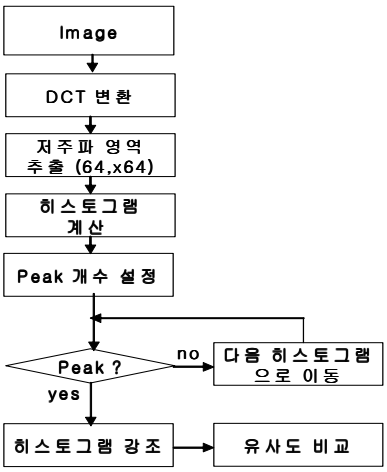


그림 4.5 DCT Peak 히스토그램의 과정

Fig. 4.5 Procedure of DCT Peak Histogram

B. 검색 시스템의 설계

그림 4.6은 전체 검색 시스템의 구조를 표현하고 있다. 위절에서 설명한 색상 기술자와 질감 기술자를 메타데이터로 변환하여 서버 측 데이터베이스에 저장하게 된다. 또한, 기술자들과 더불어 기존의 키워드 기반 방식을 사용하기 위해여 해당 키워드도 동시에 입력하여 저장하게 된다.

전체 검색은 3계층으로 이루어지게 된다. 1계층에서는 기존의 키워드 기반 방식을 사용하여 키워드와 간단한 이미지 데이터에 대한 정보 (크기, 종류 등)를 사용하여 검색하게 된다. 2계층에서는 1계층의 결과를 바탕으로 세부 검색을 실행하는 단계이다. 즉, 기존의 키워드 방식의 검색 방식에 의한 오류를 줄이기 위해 색상 기술자나 질감 기술자를 사용하여 검색을 실행한다. 3계층에서는 더욱 정확한 정보를 얻기 위해서 다양한 정보와 기술자들의 조합을 통해서 이미지를 검색하게 된다.

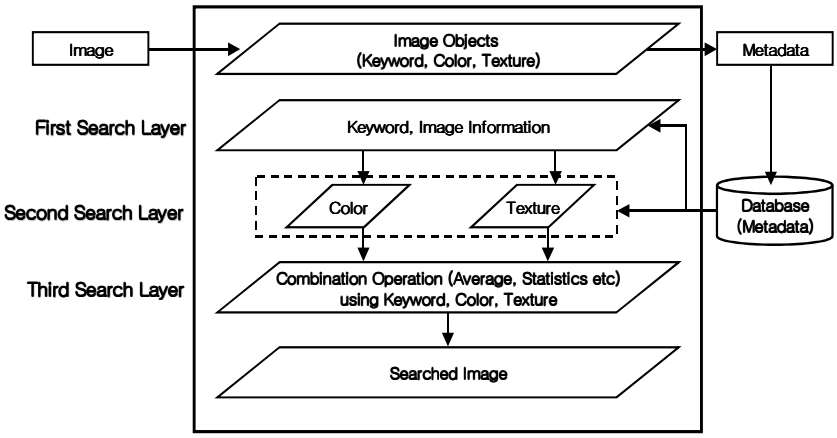


그림 4.6 검색 시스템의 구조

Fig. 4.6 Structure of Retrieval System

본 논문에서는 키워드와 색상기술자와 질감기술자에 대한 비율을 사용자가 조절가능하도록 설계하였고 사용자의 설정이 없는 경우에는 색상 기술자에 대한 유사도와 질감 기술자에 대한 유사도 조합을 식(4-5)를 이용하여 연산되도록 하였다. 각 계층에 따른 검색 레이어 설정은 사용자의 옵션 설정에 의해 결정할 수 있다.

$$S = \frac{w_c \times S_c(Q, D) + w_t \times S_t(Q, D)}{w_c + w_t} \tag{4-5}$$

여기에서 S는 조합 유사도, w_c는 색상 기술자에 대한 가중치, S_c는 색상 기술자에 대한 유사도, w_t는 질감 기술자에 대한 가중치, S_t는 질감 기술자에 대한 유사도, Q는 질의 이미지이며, D는 데이터베이스에 저장된 이미지를 타나낸다. 초기값으로 w_c와 w_t는 0.5로 설정되어 있다.

C. 데이터 구조 및 데이터베이스의 설계

1. 데이터베이스 설계

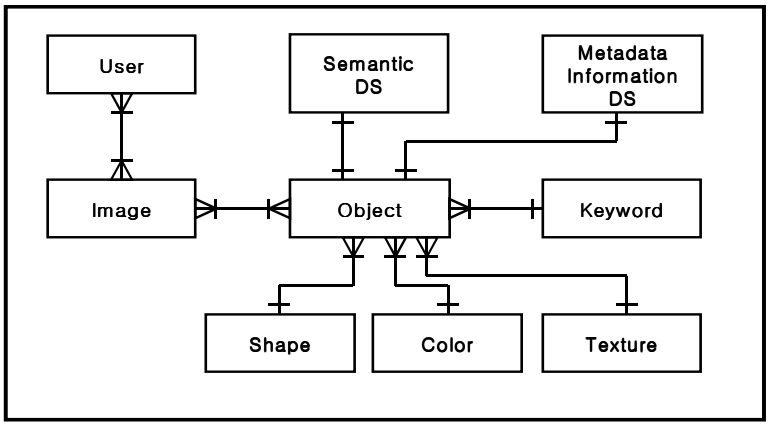


그림 4.7 데이타베이스의 구조

Fig. 4.7 Structure of Database

그림 4.7은 본 논문에서 제안된 이미지 데이터베이스의 E-R Diagram을 표현한 것이다. 현재 시스템에서는 User, Keyword, Texture, Color 테이블을 사용할 것이다. Energy와 Shape 테이블은 각각 이미지의 에너지 분포와 형태 기반 검색에서 사용할 데이터가 저장되는 곳이고, Semantic DS는 의미론적 기반 검색 시스템에서 사용할 데이터의 구조가 저장되는 곳이다. 또한, Metadata Information DS는 이미지 전체 정보에 대한 데이터 구조가 저장되는 곳이다.

2. 데이터 구조 설계

그림 4.8은 Metadata의 구조를 타나낸 것이다. 이미지 디렉토리는 여러 개의 이미지들로 구성되고 각 이미지는 아이디, 객체, 카테고리, 설명, 등록자로 구성하였고, 사용자 정보는 아이디, 패스워드, 등급, 가입날짜로 구성되었다. 이미지 객체로는 객체 아이디, 기본정보, 의미론적 정보, 특징정보로 구분되어 설계되었다. 본 논문에서는 이 중에서 기본정보와 특징정보만을 구현하여 테스트하였다. 특징정보는 칼라, 질감, 형태, 에너지, 키워드의 형태로 설계되었지만 본 논문에서는 칼라, 질감, 키워드만을 사용하였

다.

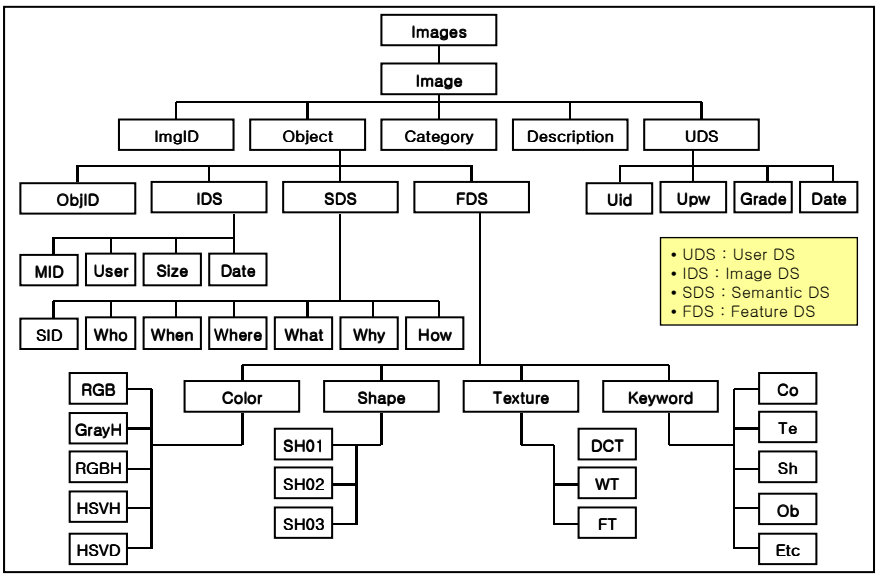


그림 4.8 Metadata의 구조

Fig. 4.8 Structure of Metadata

위에서 설계한 Metadata 구조를 기반으로 XML DTD를 설계하였다.

```
<?xml version="1.0" encoding="UTF-8" ?>

<!ELEMENT Images (Image*)>
<!ELEMENT Image (Imgid, Object*, Category, Description, UDS*)>

<!ELEMENT ImgID (#PCDATA)>
<!ELEMENT Object (ObjID, IDS*, SDS*, FDS*)>
<!ELEMENT Category (#PCDATA)>
<!ELEMENT Description (#PCDATA)>
<!ELEMENT UDS (UID, Upw, Grade, IDate)>

<!ELEMENT Object (ObjID, IDS*, SDS*, FDS*)>
<!ELEMENT ObjID (#PCDATA)>

<!ELEMENT IDS (MID, User, Size, Date)>
<!ELEMENT MID (#PCDATA)>
<!ELEMENT User (#PCDATA)>
```

```
<!ELEMENT Size (#PCDATA)>
<!ELEMENT Date (#PCDATA)>

<!ELEMENT SDS (SID, Who, When, Where, What, Why, How)>
<!ELEMENT SID (#PCDATA)>
<!ELEMENT Who (#PCDATA)>
<!ELEMENT When (#PCDATA)>
<!ELEMENT Where (#PCDATA)>
<!ELEMENT What (#PCDATA)>
<!ELEMENT Why (#PCDATA)>
<!ELEMENT How (#PCDATA)>

<!ELEMENT FDS (Color*, Shape*, Texture*, Keyword*)>
<!ELEMENT Color (RGB, GrayH, RGBH, HSVH, HSVD)>
<!ELEMENT Shape (SH01, SH02, SH03)>
<!ELEMENT Texture (DFT, DCT, WT)>
<!ELEMENT Keyword (Co, Te, Sh, Ob, Etc)>

<!ELEMENT RGB (#PCDATA)>
<!ELEMENT GrayH (#PCDATA)>
<!ELEMENT RGBH (#PCDATA)>
<!ELEMENT HSVH (#PCDATA)>
<!ELEMENT HSVD (#PCDATA)>

<!ELEMENT SH01 (#PCDATA)>
<!ELEMENT SH02 (#PCDATA)>
<!ELEMENT SH03 (#PCDATA)>

<!ELEMENT DFT (#PCDATA)>
<!ELEMENT DCT (#PCDATA)>
<!ELEMENT WT (#PCDATA)>

<!ELEMENT Co (#PCDATA)>
<!ELEMENT Te (#PCDATA)>
<!ELEMENT Sh (#PCDATA)>
<!ELEMENT Ob (#PCDATA)>
<!ELEMENT Etc (#PCDATA)>

<!ELEMENT UDS (UID, Upw, Grade, UDate)>

<!ELEMENT UDI (#PCDATA)>
```

```
<!ELEMENT Upw (#PCDATA)>
<!ELEMENT Grade (#PCDATA)>
<!ELEMENT UDate (#PCDATA)>
```

D. 웹 기반 멀티미디어 데이터 관리 시스템의 설계

1. 시스템의 구성

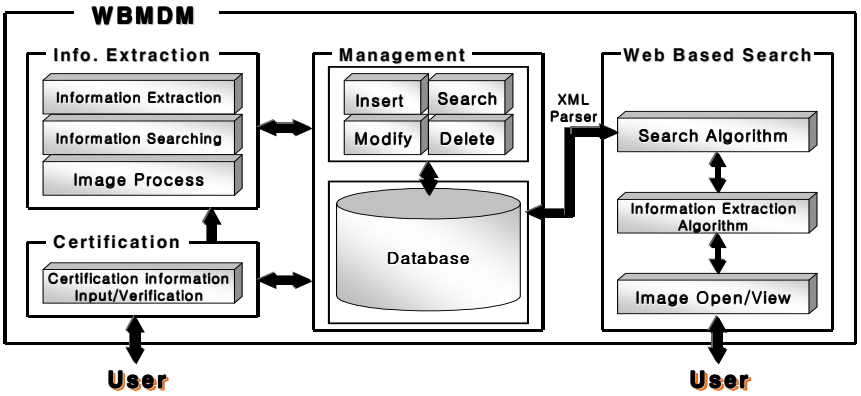


그림 4.9 전체 시스템의 구조

Fig. 4.9 Structure of Total System

그림 4.9는 본 논문에서 제안한 검색 시스템을 포함하는 전체 멀티미디어 데이터 관리 시스템의 구성도이다. 전체 시스템은 이미지 정보 추출부, 사용자 인증부, 이미지 정보 관리부, 웹 기반 이미지 검색부로 구성되어 있다. 이미지 정보 추출부는 해당 이미지에서 필요한 정보나 기술자들을 추출하여 DB에 저장하는 기능과 각종 이미지 처리 알고리즘을 구현하고 테스트하는 기능을 갖는다. 사용자 인증부는 구현된 시스템을 사용하기 위한 관리자 및 사용자 권한을 설정하고 검사하는 부분이다. 이미지 정보 관리부는 이미지 정보 추출부 및 사용자 인증부에서 처리된 데이터를 추가, 삭제, 변경, 검색하는 기능을 갖는다. 이미지 검색부는 기존의 웹 기반 이미지 검색 엔진과 같이 원하는 이미지를 검색하는 기능을 갖는다. 검색결과는 XML 파서에 의해 XML Document로 변환되어 XSL로 설정된 형태로 웹상에서 출력하게 된다.

그림 4.10은 웹상에서 이미지를 검색하는 과정을 나타내고 있다. 사용자가 원하는 이미지를 불러와서 해당 알고리즘에의 의해 분석한 데이터나 키워드를 입력하여 서버로 검색 키워드로써 전송한다. 서버는 전송된 키워드를 이용하여 데이터베이스에 저장된 이미지 정보들과 비교하여 유사도 높은 순으로 이미지 위치를 전송하면 클라이언트는 전송된 위치에 맞는 이미지 디렉토리로부터 이미지를 불러와서 웹상에 출력하게 된다. 관리자는 이미지 관리 툴을 이용하여 지속적으로 데이터베이스에 멀티미디어 정보를 추가/수정/삭제/검색 등을 할 수 있다. 전체시스템은 JSP/Bean으로 구축되었다.

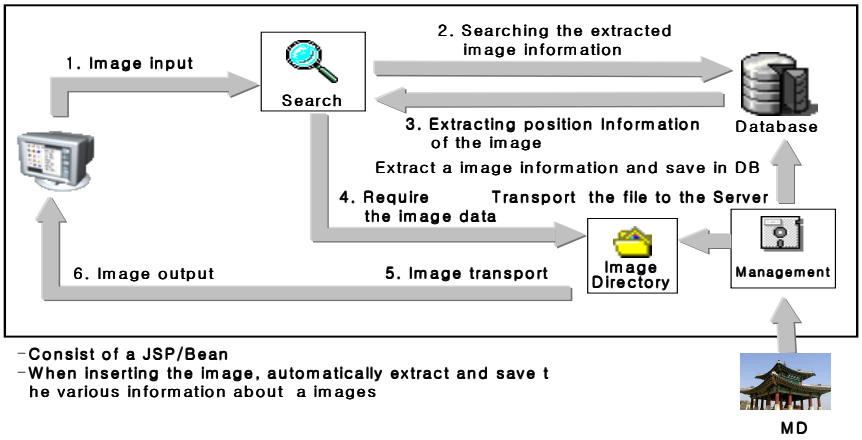


그림 4.10 전체 시스템의 흐름도

Fig. 4.10 Diagram of Total System

2. 인터페이스 설계

그림 4.11은 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.

검색된 이미지 목록을 보여주는 부분은 데이터베이스로부터 검색된 이미지 정보를 XML 데이터로 변환하여 해당 XSL에 따라 다양한 형태로 출력 가능하도록 하고 다양한 세부 정보를 보여줄 수 있도록 설계 하였다.

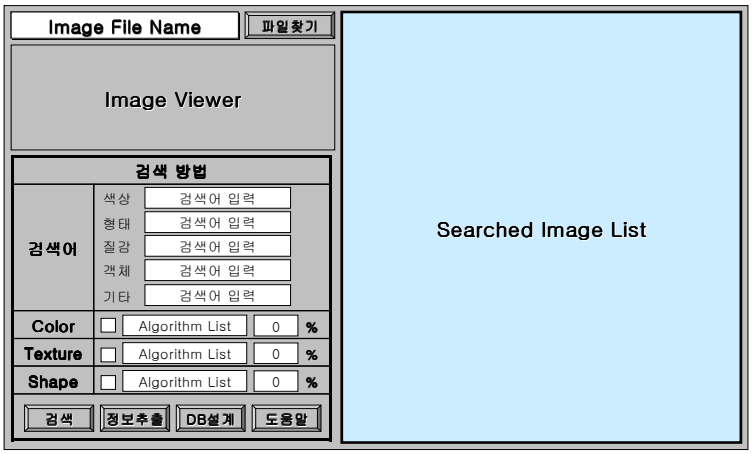


그림 4.11 인터페이스 설계

Fig. 4.11 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	

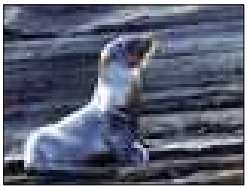
2. 사용 이미지

그림 5.1은 테스트에서 사용된 이미지는 크기가 각기 다른 임의의 이미지 20개를 사용하였다. 구현된 시스템을 검증하기 위하여 테스트 이미지 중에서 (f), (j), (o), (s) 이미지를 기준 이미지로 하여 검색을 실행하였다. 전체 이미지는 미리 제안된 시스템에서 구현한 정보 추출 도구를 사용하여 각 이미지에 대한 기본 정보와 기술자 정보가

데이터베이스에 저장되었다.

(a) animal00005 (400×600)

(b) animal00005 (400×600)



(c) animal00010(400×300)

(d) animal00027(400×267)

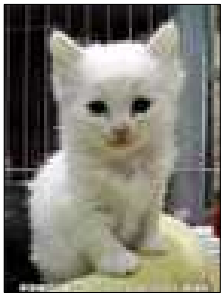
(e) animal00031(400×300)

(f) animal00076(640×469)

(g) animal00078(640×465)

(h) animal00082(640×435)

(i) animal00085(640×479)



(j) animal00093(360×480)

(k) animal00094(360×480)

(l) animal00099(316×480)

(m) animal00112(800×600)

(n) animal00113(800×600)

(o) animal00122(800×600)

(p) animal00123(800×600)

(q) animal00126(800×600)

(r) animal00130(800×600)

(s) animal00134(800×600)

(t) animal00136(800×600)

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.

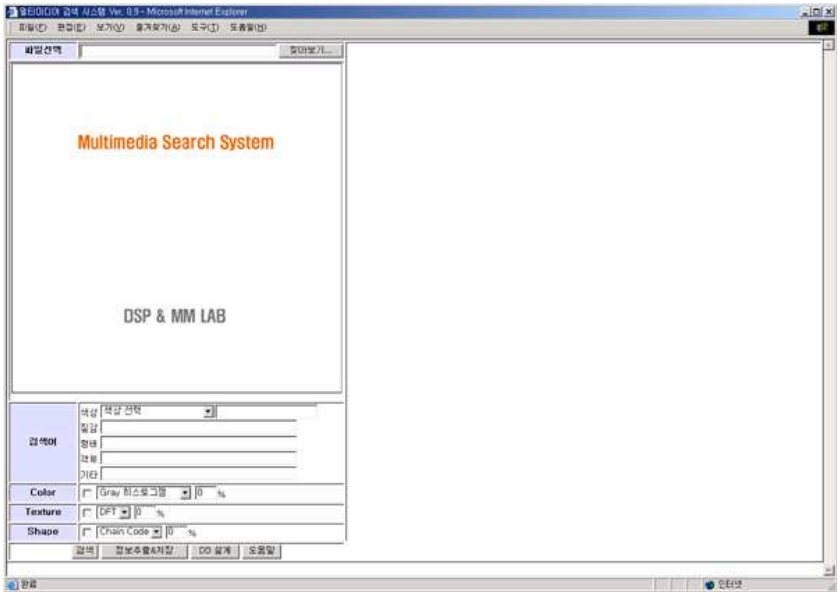


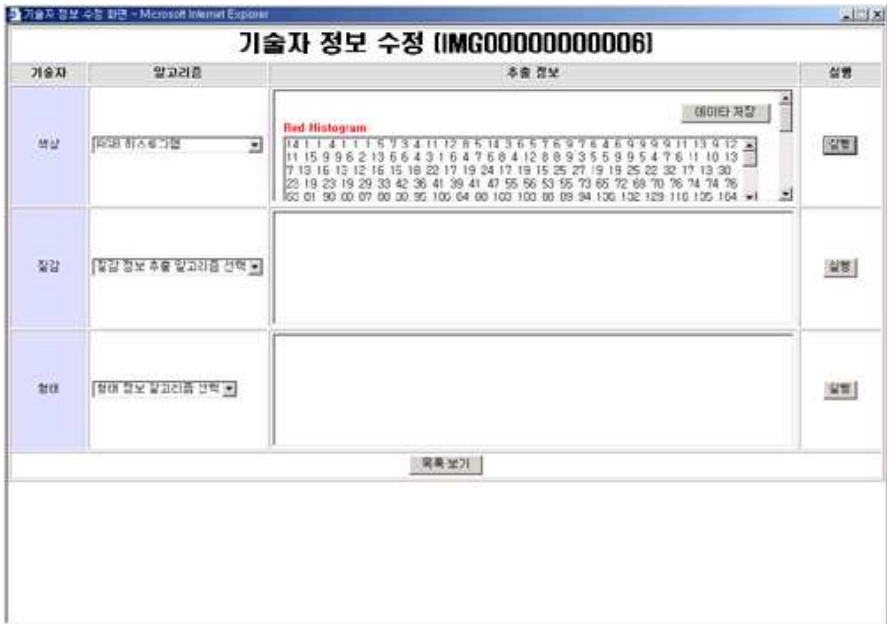
그림 5.2 시작 화면

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.

(a)



(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 기반 검색

본 논문에서 기술자 정보로써 색상과 질감 기술자를 사용하였다. 색상 기술자를 검증하기 위하여 Gray 칼라 히스토그램을 사용한 방법, RGB 칼라 히스토그램을 사용한 방법, HSV 칼라 히스토그램을 사용한 방법, 제안된 알고리즘을 사용한 방법을 비교하였다.

a. *Gray Color Histogram*

그림 5.4의 흐름도는 본 논문에서 제안한 알고리즘과 비교하기 위한 Gray Color 히스토그램 알고리즘을 보여주고 있다. 이미지가 로드되면 RGB 칼라 맵을 추출하고 이를 기반으로 Gray 칼라 맵으로 변환한다. 변환된 칼라 맵을 이용하여 Gray 히스토그램을 구하고 비교 이미지들의 Gray 칼라 히스토그램과 비교하여 유사성을 계산한다. 비교하고자하는 이미지들의 히스토그램 값은 미리 계산되어 데이터베이스에 저장되어 있기 때문에 따로 계산하지 않고 데이터베이스에 저장된 값을 사용하여 계산된다.

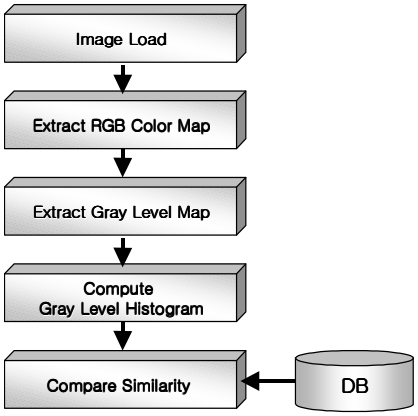


그림 5.4 Gray Color Histogram을 이용한 알고리즘

Fig. 5.4 Algorithm using Gray Color Histogram

표 5.1은 위의 알고리즘을 이용하여 이미지를 검색한 결과이다. 각 값들은 유사도를 표시하고 있다. 최대 유사도는 100%이고 최소 유사도는 0%이다. 굵은 글씨로 이태릭

체로 보여지는 값들이 유사도가 가장 높은 4개의 이미지를 표시한 것이다.

표 5.1 Gray Color Histogram을 이용한 검색 결과 (%)
Table 5.1 Searching Result using Gray Color Histogram (%)

이미지	*(f)*	*(j)*	*(o)*	*(s)*	이미지	*(f)*	*(j)*	*(o)*	*(s)*
(a)	61	66	43	68	(k)	70	75	57	66
(b)	55	54	50	60	(l)	56	63	44	61
(c)	57	59	50	65	(m)	67	66	58	69
(d)	87	*82*	*69*	*77*	(n)	83	77	*70*	73
(e)	53	59	43	61	(o)	65	61	*100*	56
(f)	*100*	75	65	73	(p)	75	70	58	75
(g)	*91*	*79*	68	76	(q)	52	47	45	48
(h)	*91*	75	65	73	(r)	46	52	44	55
(i)	*91*	*80*	*69*	*77*	(s)	73	69	56	*100*
(j)	75	*100*	61	69	(t)	73	71	56	*81*

그림 5.5는 표 5.1을 그래프로 표현한 것으로 검색 결과를 시각적으로 보여주고 있다.

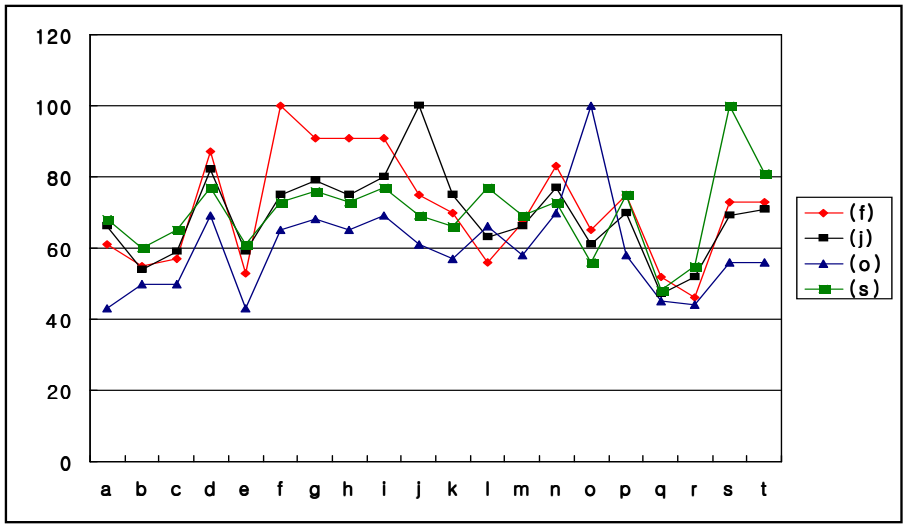


그림 5.5 Gray Color Histogram을 이용한 수행결과 그래프
Fig. 5.5 A Peformed Result Graph using Gray Color Histogram

표 5.2 Gray Color Histogram을 이용한 검색된 이미지들
Table 5.2 A Searched Images using Gray Color Histogram

기준이미지	검색이미지 (유사도순)

표 5.2는 위의 알고리즘을 사용하여 실제 검색된 이미지를 나열할 것이다. 결과에서 (f)이미지에 대한 검색은 비교적 정확하게 추출하였지만 (j), (o) 이미지에 대한 검색결

과는 정확하지 않음을 알 수 있다. 마지막으로 (s) 이미지에 대한 검색결과는 유사한 이미지인 (t) 이미지는 검색을 하였지만 다음 유사도를 갖는 이미지에서 정확하지 않은 검색을 하였다.

b. *RGB Color Histogram*

그림 5.6의 흐름도는 본 논문에서 제안한 알고리즘과 비교하기 위한 RGB Color 히스토그램 알고리즘을 보여주고 있다. 이미지가 로드되면 RGB 칼라 맵을 추출하고 이를 이용하여 RGB 히스토그램을 구하고 비교 이미지들의 RGB 칼라 히스토그램과 비교하여 유사성을 계산한다. 비교하고자하는 이미지들의 히스토그램 값은 미리 계산되어 데이터베이스에 저장되어 있기 때문에 따로 계산하지 않고 데이터베이스에 저장된 값을 사용하여 계산된다.

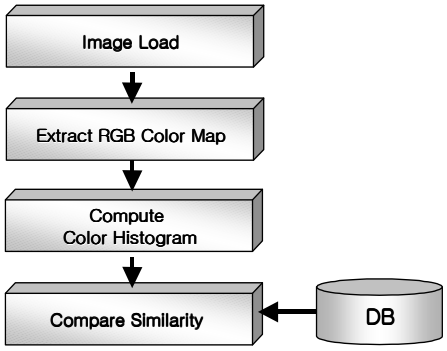


그림 5.6 RGB Color Histogram을 이용한 알고리즘

Fig. 5.6 Algorithm using RGB Color Histogram

표 5.3는 위의 알고리즘을 이용하여 이미지를 검색한 결과이다. 각 값들은 유사도를 표시하고 있다. 최대 유사도는 100%이고 최소 유사도는 0%이다. 굵은 글씨로 이태릭체로 보여지는 값들이 유사도가 가장 높은 4개의 이미지를 표시한 것이다.

표 5.3 RGB Color Histogram을 이용한 테스트 결과 (%)

Table 5.3 Test Result using RGB Color Histogram (%)

이미지	*(f)*	*(j)*	*(o)*	*(s)*	이미지	*(f)*	*(j)*	*(o)*	*(s)*
(a)	62	69	47	67	(k)	69	73	54	65
(b)	55	53	47	55	(l)	57	63	48	60
(c)	56	55	47	62	(m)	65	61	56	67
(d)	86	***80***	***67***	***74***	(n)	82	74	65	69
(e)	53	58	43	61	(o)	65	57	***100***	55
(f)	***100***	73	65	70	(p)	71	69	55	72
(g)	***90***	***76***	***67***	73	(q)	52	46	43	48
(h)	***90***	73	65	70	(r)	47	52	43	55
(i)	***90***	***77***	***68***	***74***	(s)	70	66	55	***100***
(j)	73	***100***	57	66	(t)	71	67	56	***80***

그림 5.7는 표 5.3을 그래프로 표현한 것으로 검색 결과를 시각적으로 보여주고 있다.

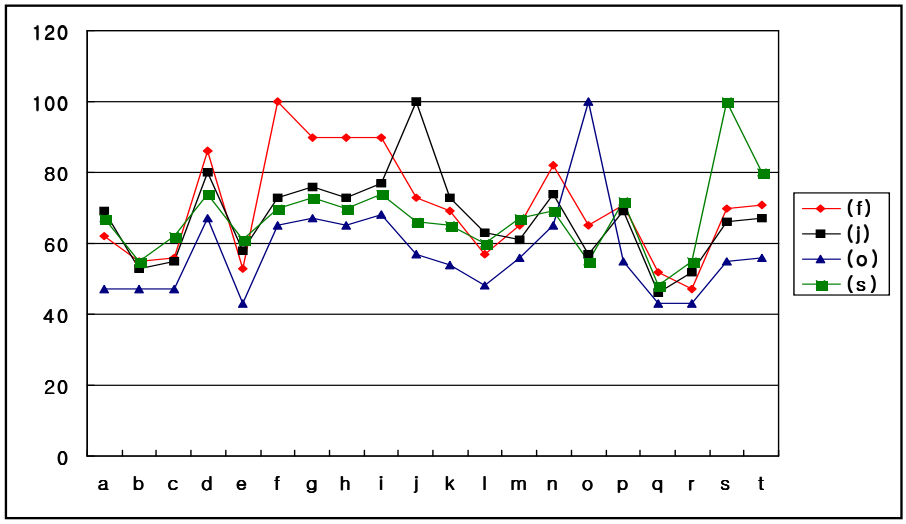


그림 5.7 RGB Color Histogram을 이용한 수행결과 그래프

Fig. 5.7 A Peformed Result Graph using RGB Color Histogram

표 5.4 RGB Color Histogram을 이용한 검색된 이미지들
Table 5.4 A Searched Images using RGB Color Histogram

기준이미지	검색이미지 (유사도순)

표 5.4는 위의 알고리즘을 사용하여 실제 검색된 이미지를 나열할 것이다. 결과에서 (f)이미지에 대한 검색은 비교적 정확하게 추출하였지만 (j), (o) 이미지에 대한 검색결

과는 정확하지 않음을 알 수 있다. 마지막으로 (s) 이미지에 대한 검색결과는 유사한 이미지인 (t) 이미지는 검색을 하였지만 다음 유사도를 갖는 이미지에서 정확하지 않은 검색을 하였다.

c. *HSV Color Histogram*

그림 5.8의 흐름도는 본 논문에서 제안한 알고리즘과 비교하기 위한 HSV Color 히스토그램 알고리즘을 보여주고 있다. 이미지가 로드되면 RGB 칼라 맵을 추출하고 이를 기반으로 HSV 칼라 맵으로 변환한다. 변환된 칼라 맵을 이용하여 HSV 히스토그램을 구하고 비교 이미지들의 HSV 칼라 히스토그램과 비교하여 유사성을 계산한다. 비교하고자하는 이미지들의 히스토그램 값은 미리 계산되어 데이터베이스에 저장되어 있기 때문에 따로 계산하지 않고 데이터베이스에 저장된 값을 사용하여 계산된다.

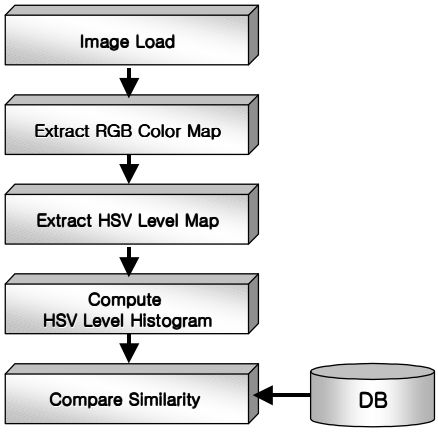


그림 5.8 HSV Color Histogram을 이용한 알고리즘

Fig. 5.8 Algorithm using HSV Color Histogram

표 5.5는 위의 알고리즘을 이용하여 이미지를 검색한 결과이다. 각 값들은 유사도를 표시하고 있다. 최대 유사도는 100%이고 최소 유사도는 0%이다. 굵은 글씨로 이태릭체로 보여지는 값들이 유사도가 가장 높은 4개의 이미지를 표시한 것이다.

표 5.5 HSV Color Histogram을 이용한 테스트 결과 (%)
Table 5.5 Test Result using HSV Color Histogram (%)

이미지	*(f)*	*(j)*	*(o)*	*(s)*	이미지	*(f)*	*(j)*	*(o)*	*(s)*
(a)	39	50	49	44	(k)	38	***67***	39	41
(b)	40	49	49	44	(l)	39	48	39	37
(c)	38	43	51	39	(m)	39	47	***55***	***50***
(d)	39	***52***	53	37	(n)	38	40	***55***	44
(e)	41	***56***	***55***	38	(o)	37	40	***100***	44
(f)	***100***	36	37	39	(p)	36	42	***59***	43
(g)	***53***	37	38	39	(q)	38	38	49	***51***
(h)	***49***	35	37	38	(r)	38	41	50	49
(i)	***57***	35	42	43	(s)	40	43	44	***100***
(j)	37	***100***	39	43	(t)	38	41	42	***64***

그림 5.9는 표 5.5을 그래프로 표현한 것으로 검색 결과를 시각적으로 보여주고 있다.

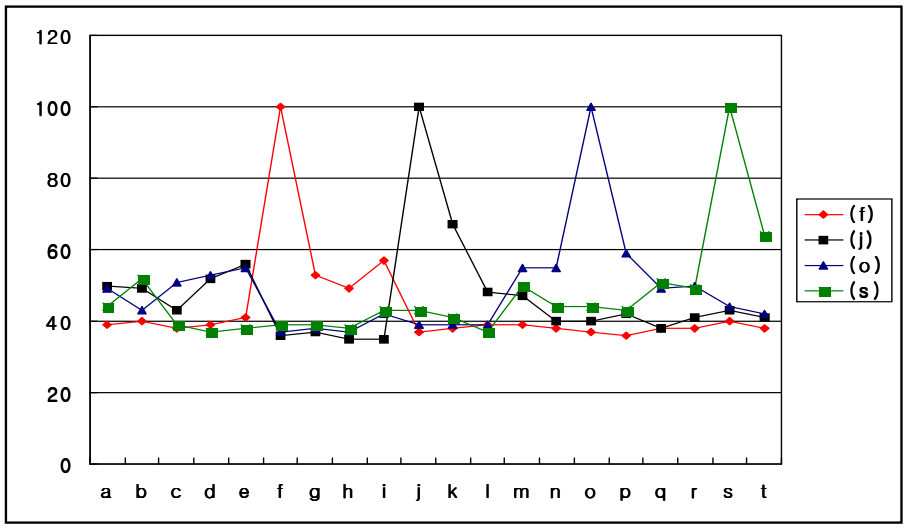


그림 5.9 HSV Color Histogram을 이용한 수행결과 그래프
Fig. 5.9 A Peformed Result Graph using HSV Color Histogram

표 5.6 HSV Color Histogram을 이용한 검색된 이미지들
Table 5.6 A Searched Images using HSV Color Histogram

기준이미지	검색이미지 (유사도순)

표 5.6은 위의 알고리즘을 사용하여 실제 검색된 이미지를 나열할 것이다.

결과에서 (f)이미지에 대한 검색은 비교적 정확하게 추출하였다. 이전의 a, b절에서 사용한 알고리즘의 결과와 비교할 때 이전의 알고리즘에서는 유사도가 동일하게 계산되었지만 비슷한 이미지이지만 (f), (g), (h) 이미지에 비해 (i) 이미지는 크기 및 이미지의 변화율이 큰 것을 육안으로도 알 수 있다. 이전의 알고리즘에서는 이러한 차이를 분석할 수 없었지만 이 절에서 사용한 알고리즘은 이러한 차이를 구분하는 것을 알 수 있다.

(j), (o) 이미지에 대한 검색결과도 이전의 알고리즘에 비해 비교적 정확한 결과를 얻을 수 있을 것을 알 수 있다. 하지만 (j) 이미지와 유사한 (l) 이미지를 검색하지 못하는 것을 알 수 있다. 또한, (o) 이미지에 대해서도 유사한 이미지를 검색하지만 다음 유사도에서 검색된 이미지가 오차가 많은 검색을 하고 있음을 알 수 있다.

(s) 이미지에 대한 검색결과는 유사한 이미지인 (t) 이미지는 검색을 하였고 이전에 검색에 비해 색깔에 대한 유사도 검색 성능이 늘었지만 역시 마지막 검색 결과에서 오차가 발생하는 것을 알 수 있다.

d. 사용된 알고리즘 : *HSV Color* 분할

위에서 사용한 칼라 기술자를 추출하기 위한 알고리즘들의 문제점을 극복하기 위해 본 논문에서는 HSV 히스토그램에서 두드러지는 색을 추출하고 그러한 색에 가중치를 부여하는 방법을 사용하였다. 일반적으로 HSV가 주위의 밝기 변화에도 강인한 성질을 가지고 있다. 따라서, 동일한 색의 분포를 가지고 있는 객체를 분할하는 곳에 자주 사용되고 있다. 하지만 위의 검색 결과에서 보듯이 무작위적인 이미지를 검색할 시에는 오차가 발생하는 것을 알 수 있었다. 이러한 문제를 해결하기 위해 본 논문에서는 S와 V의 값에 따라 H 값에 가중치를 주는 방법을 사용하였다.

표 5.7는 그림 4.2와 같이 본 논문에서 제안한 알고리즘을 이용하여 이미지를 검색한 결과이다. 각 값들은 유사도를 표시하고 있다. 최대 유사도는 100%이고 최소 유사도는 0%이다. 굵은 글씨로 이태릭체로 보여지는 값들이 유사도가 가장 높은 4개의 이미지를 표시한 것이다.

표 5.7 HSV Color 분할을 이용한 테스트 결과 (%)
Table 5.7 Test Result using HSV Color Segmentation (%)

이미지	*(f)*	*(j)*	*(o)*	*(s)*	이미지	*(f)*	*(j)*	*(o)*	*(s)*
(a)	13	29	22	25	(k)	17	***80***	19	26
(b)	17	32	23	24	(l)	20	***42***	16	17
(c)	17	19	26	15	(m)	13	24	25	32
(d)	11	33	20	11	(n)	12	16	22	21
(e)	17	***40***	19	10	(o)	13	16	***100***	20
(f)	***100***	17	13	17	(p)	10	21	***53***	22
(g)	***52***	19	12	18	(q)	13	18	***28***	***41***
(h)	***45***	18	14	18	(r)	17	23	***26***	***43***
(i)	***50***	13	14	19	(s)	17	28	20	***100***
(j)	17	***100***	16	28	(t)	16	26	23	***80***

그림 5.10는 표 5.7을 그래프로 표현한 것으로 검색 결과를 시각적으로 보여주고 있다.

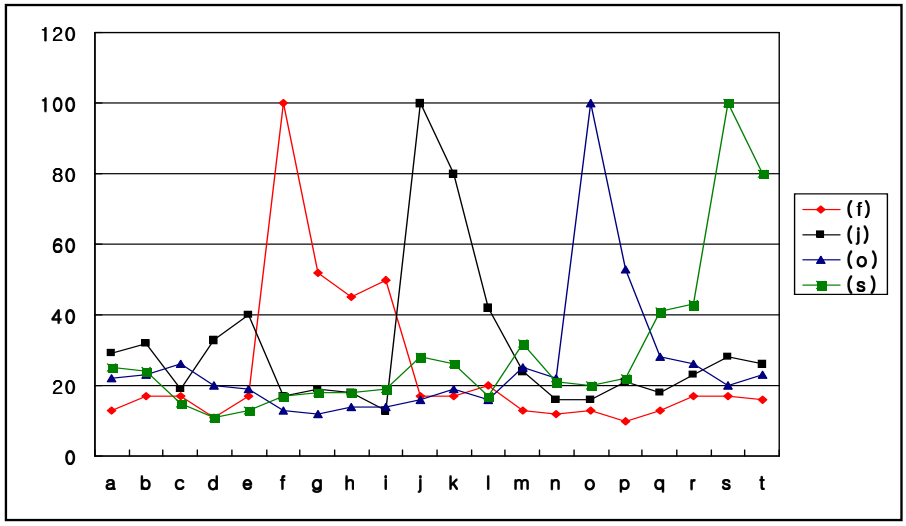


그림 5.10 HSV Color Segmentation을 이용한 수행결과 그래프
Fig. 5.10 A Peformed Result Graph using HSV Color Segmentation

표 5.8 HSV Color Segmentation을 이용한 검색된 이미지들
Table 5.8 A Searched Images using HSV Color Segmentation

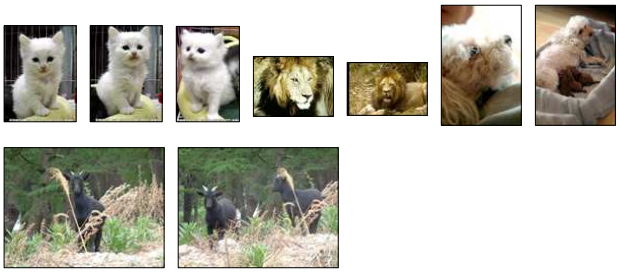
기준이미지	검색이미지 (유사도순)
	

표 5.8은 위의 알고리즘을 사용하여 실제 검색된 이미지를 나열할 것이다.

결과에서 (f)이미지에 대한 검색은 비교적 정확하게 추출하였다. 이전의 이전에서 사용한 알고리즘의 결과와 비교할 때 이전의 알고리즘에서는 유사도가 동일하게 계산되었지만 비슷한 이미지이지만 (f), (g), (h) 이미지에 비해 (i) 이미지는 크기 및 이미지의 변화율이 큰 것을 육안으로도 알 수 있다. 이전의 알고리즘에서는 이러한 차이를

분석할 수 없었지만 이 절에서 사용한 알고리즘은 이러한 차이를 구분하는 것을 알 수 있다. 또한 c절에서 사용한 알고리즘과 비교하면 c절에서는 (i)의 이미지의 유사도가 높다고 나오지만 본 논문에서 제안한 알고리즘에서는 (g)의 이미지의 유사도가 높게 나오는 것을 알 수 있다. 그것은 육안으로 보이듯이 전체적인 객체들 중에서 붉은 색 계열의 공과 흰 고양이와 노란 인형을 볼 때 (g)의 이미지가 더 유사함을 알 수 있을 것이다.

(j) 이미지에 대한 검색결과도 c절의 알고리즘에서 대해서는 검색이 되지 않았던 (l) 이미지가 검색되어지는 것을 볼 수 있다.

(o) 이미지에 대해서도 c절에서는 유사한 이미지를 검색하지만 다음 유사도에서 검색된 이미지가 오차가 많은 검색을 하지만 본 논문에서 제안한 알고리즘은 녹색계열의 이미지가 주를 이루는 (q)와 (r) 이미지가 검색되는 것을 알 수 있다.

(s) 이미지에 대해서는 c절에서는 유사한 이미지인 (t) 이미지는 검색을 하였지만 유사가 차이가 나는 검색 결과에서 오차가 발생하였지만 녹색계열과 갈색계열이 주를 이루는 이미지인 (r)과 (q) 이미지를 비교적 정확히 검색하는 것을 알 수 있다.

2. 질감 기반 검색

표 5.9는 그림 4.5와 같이 본 논문에서 제안한 알고리즘을 이용하여 이미지를 검색한 결과이다. 본 논문에서는 질감 기술자로써 DCT 알고리즘을 사용하였다. 따로, 객체를 분할 하지 않고 무작위 이미지를 사용하였기 때문에 질감의 에너지 요소를 분석하여 이를 비교함으로써 이미지를 검색하고 하였다. 단, 기존의 DCT 변환 결과 중에서 저주파 성분 중 64×64 부분의 값에서 Peak 히스토그램 구하여 이를 이용한 알고리즘을 사용하였다.

각 값들은 유사도를 표시하고 있다. 최대 유사도는 100%이고 최소 유사도는 0%이다. 굵은 글씨로 이태릭체로 보여지는 값들이 유사도가 가장 높은 4개의 이미지를 표시한 것이다.

표 5.9 DCT Peak Histogram을 이용한 테스트 결과 (%)
Table 5.9 Test Result using DCT Peak Histogram (%)

이미지	*(f)*	*(j)*	*(o)*	*(s)*	이미지	*(f)*	*(j)*	*(o)*	*(s)*
(a)	41	***65***	54	47	(k)	39	***69***	41	43
(b)	41	***59***	53	45	(l)	42	59	42	40
(c)	39	52	51	42	(m)	45	54	***59***	54
(d)	41	58	53	40	(n)	40	46	***59***	55
(e)	43	58	57	39	(o)	38	***60***	***100***	54
(f)	***100***	37	37	43	(p)	38	57	***63***	53
(g)	***55***	37	38	43	(q)	41	52	54	***58***
(h)	***52***	36	37	42	(r)	39	55	56	***58***
(i)	***62***	36	41	46	(s)	43	52	48	***100***
(j)	38	***100***	40	44	(t)	42	50	47	***67***

그림 5.11는 표 5.9를 그래프로 표현한 것으로 검색 결과를 시각적으로 보여주고 있다.

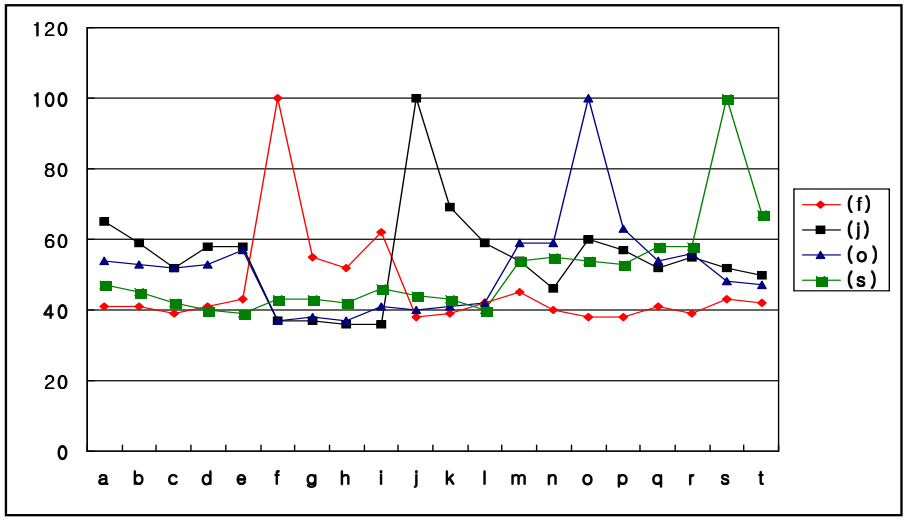


그림 5.11 DCT Peak 히스토그램을 이용한 알고리즘의 수행결과 그래프

Fig. 5.11 A Peformed Result Graph using DCT Peak Histogram

표 5.10 DCT Peak 히스토그램을 이용한 검색된 이미지들
Table 5.10 A Searched Images using DCT Peak Histogram

기준이미지	검색이미지 (유사도순)

표 5.10은 위의 알고리즘을 사용하여 실제 검색된 이미지를 나열할 것이다.
결과에서 (f)이미지에 대한 검색은 비교적 정확하게 추출하였다. 하지만 일부 오차를

포함하고 있는 것을 알 수 있다.

(j), (o) 이미지에 대한 검색은 유사도가 멀어지면 오차를 많이 발생하였다.

(s) 이미지에 대한 검색결과는 유사한 이미지인 (t) 이미지는 검색을 하였고 풀 질감을 갖는 (q)와 (r) 이미지가 검색되었다.

3. 키워드 기반 검색

본 논문에서는 기존의 단순한 키워드 입력방식이 가지는 비정확성 문제를 해결하기 위하여 키워드를 세분화하여 입력하는 방식을 사용하였다. 색상, 질감, 형태, 객체 등의 키워드를 분리하여 입력함으로써 입력자의 주관적인 단어 선택이나 의미 선택에 의한 오류발생 문제를 감소시킬 수 있으며 각 키워드를 조합함으로써 구문론적 검색 방식을 적용할 수도 있게 된다.

예를들어, “하얀색 개”나 “백색 개”는 같은 의미이지만 검색하는 사람이 입력하는 모든 경우의 수를 등록하는 것은 어려움이 있다. 따라서, 기준 색상을 설정함으로써 동일한 색상을 다르게 표현함으로써 발생하는 문제점을 해결하였다. 또한, 질감, 형태에 대한 이미지 정보를 따로 입력함으로써 기존에 객체만을 입력하는 이용하는 것보다 더욱 많은 검색 키워드를 제공하게 하였다.

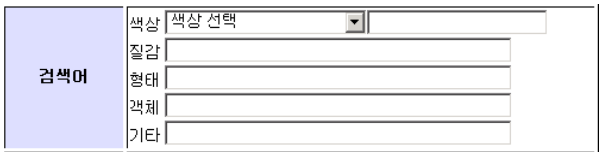


그림 5.12 검색 키워드 입력 화면

Fig. 5.12 Screen for Search Keyword Input

사용자가 키워드를 분리하여 입력하면 검색 엔진은 입력된 항목의 키워드를 서버로 전송하고 데이터베이스의 동일 항목의 키워드와 비교한다. 기존에 한 항목에 여러 가지의 키워드들이 저장되어 있어서 단순히 텍스트만을 비교하는 것으로 결과를 얻는 것에 반해 제안된 시스템은 키워드를 의미론적으로 분류함으로써 각 항목을 조합한 구문론적 키워드 검색이 가능하다.

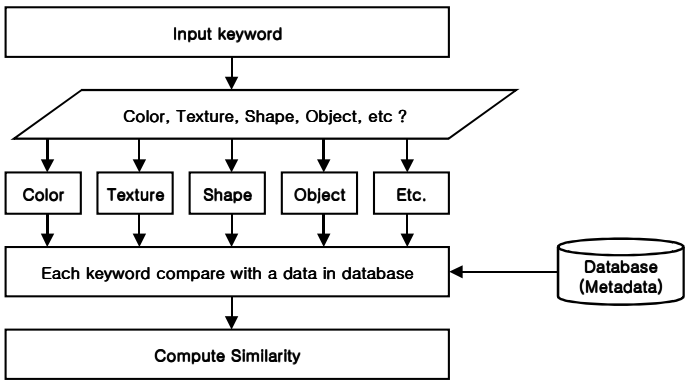


그림 5.13 본 논문에서 사용된 키워드 기반 검색 방법

Fig. 5.13 Keyword-Based Retrieval Technique using this Paper

VI. 결론 및 향후 연구과제

본 논문에서는 기존의 키워드 기반 검색 시스템의 단점을 보안하기 위하여 키워드 기반 검색에 멀티미디어 데이터 기술자 정보를 추가하여 검색하고 이를 관리하는 시스템을 설계하였다.

첫째, 기존의 단순한 키워드 기반 방식의 단점을 해결하기 위해 먼저, 단순한 키워드 방식이 아닌 의미기반 키워드 방식을 사용하였다. 이는 단순한 키워드 대신 기술자 정보와 부합되는 정보를 키워드로 입력하게 함으로써 구문론적 검색이 이루어지도록 하였다. 색상 키워드는 규정된 색만을 입력가능 하도록 하여 같은 의미를 갖는 키워드일지라도 검색에서 제외되는 문제를 제거하였다. 또한, 질감, 형태, 객체 등으로 입력을 구분하여 줌으로써 입력 시 좀 더 객관적인 키워드 도출 및 구문 검색이 쉽도록 하였다.

둘째, 제안된 시스템은 키워드 기반에 필요에 따라 기술자 정보를 포함한 검색이 가능하도록 구성하였다. 기술자는 색상, 질감, 형태를 사용할 수 있도록 하였고 사용자의 요구에 의해 다양한 알고리즘의 추가 또는 삭제가 가능하도록 하였다.

셋째, 제안된 시스템은 키워드 및 각 기술자에 대한 조합을 사용하여 통합 검색 기능을 제공한다. 사용자에 의해 설정된 비율에 의해 각 방법들은 통합 유사도를 도출하여 가장 유사도가 높은 이미지를 검색하게 된다.

기술자 정보를 색상 기술자와 질감 기술자를 조합하여 사용함으로써 색상이나 질감 등 하나만의 기술자를 사용함으로 발생하는 오류를 줄일 수 있었다. 그리고, 기존의 키워드 검색 방식을 1계층으로 하고 기술자 정보를 2계층, 각종 기술자 및 키워드 조합을 3계층 검색방식을 취함으로써 사용자로 하여금 폭넓은 검색 방식을 취하도록 하였다.

마지막으로 모든 이미지 정보는 데이터베이스에 저장되고 검색된 결과는 검색의 효율성 , 고속 검색, 다양한 출력 방식의 제공 등의 기능을 지원하기 위해 XML 데이터로 구성되어 사용자에게 제공된다.

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

또한, 향후 연구될 연구에 필요한 다양한 시스템을 테스트하기 위한 기반 시스템을

설계함으로써 추후 연구에서 손쉬운 알고리즘 테스트 및 기능 변경을 손쉽게 하였다.

이와같이 제안된 시스템은 멀티미디어의 데이터의 다양한 서술자 정보를 추출함으로써 최적의 데이터베이스를 구성하고 검색할 수 있는 시스템으로 활용할 수 있을 것이다.

앞으로, 최적화된 특징 정보의 표현 알고리즘 개발, 최적화된 검색 엔진, 지속적인 이미지 데이터베이스 구축, 네트워크 부하 문제 해결 등에 대한 연구가 되어야 할 것이다. 또한, 동일한 데이터를 무선기기에서도 동시에 사용할 수 있는 유무선 연동 시스템도 구현해야 할 것이다. 마지막으로 이러한 연구결과를 직접 사용할 유무선 컨텐츠 개발 및 사업화 가능성에 대한 연구도 필요하다.

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] 하성욱, "U-Commerce의 현황과 전망", 2002

[5] Hagel, John III and Jeffery F. Rayport, "The New Infomediaries", The McKinsey Quarterly, 1997.

[6] Rappa, Michael, "Business Models on the Web"

[7] 이경전, "인터넷 비즈니스 모델의 발전방향", 추계MIS학회, 1999. 11.

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

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

[10] 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,

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

[12] 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.

[13] 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.

[14] 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.

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

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

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

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

[23] 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.

[24] Y. Gong, "Advancing content-based image retrieval by exploiting image color and region features", Multimedia Systems, Vol.7, pp.449~457, 1999.

[25] B.M. Mehtre, M.S. Kankanhalli and W.F. Lee, "Content-based Image Retrieval Using a Composite Color-Shape Approach", Information Processing & Management , Vol.34, No.1, pp.109~120, 1998.

[26] T. Gevers and A.W.M. Smeulders, "Image Indexing using Composite Color

and Shape Invariant Features", ICCV, Bombay, India. pp.576~581, 1998.

[27] R.J. Qian, P.J.L. Van Beek, and M.I. Sezan, "Image retrieval using blob histograms", ICME 2000, Vol.1, pp.125~128, 2000.

[28] T.K. Shih, C.S. Wang, A.Y. Chang and C.H. Kao, "Indexing and Retrieval Scheme of the Image Database Based on Color and Spatial Relations", ICME 2000. Vol.1, pp.129~132, 2000.

[29] M.A. Stricker and M. Orengo, "Color Indexing with Weak Spatial Constraints", Storage and Retrieval for Image and Video Databases (SPIE), pp.29~40, 1996.

[30] W. Hsu, T.S. Chua and H.K. Pung "An integrated color-spatial approach to content-based image retrieval,", ACM Multimedia Conference, Vol.2, pp.305~313, 1995.

[31] Z. Lei, L. Fuzong and Z. Bo, "A CBIR method based on color-spatial feature", TENCON 99. Proceedings of the IEEE Region 10 Conference, Vol.1, pp.166-169, 1999.

[32] Y. Liu, and S. Ozawa, "A integrated color-spatial image representation and the similar image retrieval", IEEE Southwest Symposium on Image Analysis and Interpretation, pp.283~287, 2000.

[33] A. Rao, R.K. Srihari and Z. Zhang, "Spatial color histograms for content-based image retrieval", IEEE Int. Conference on Tools with Artificial Intelligence, pp.183~186, 1999.

[34] G. Pass and Zabih R., "Histogram refinement for content-based image retrieval", IEEE Workshop on WACV, pp.96~102, 1996.

[35] R. Barber, W. Equitz, W. Flickner, W. Niblack, D. Petkovic and P. Yanker, "Efficient query by image content for very large image databases", Compcon Spring '93, Digest of Papers, pp.17~19, 1993.

[36] V. Kovalev and S. Volmer, "Color Co-Occurrence Descriptors for Querying-by-Example", IEEE conference on Multimedia Modeling, pp.32~38, 1998.

[37] J. Huang, S.R. Kumar, M. Mitra, W.J. Zhu and R. Zabih, "Image indexing

using color correlograms", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.762~768, 1997.

[38] J. Huang, "Color-Spactial Image Indexing and Application" , Ph.D. Thesis in Cornell Univ., 1998.

[39] ISO/IEC JTC 1/SC 29/WG 11: "Information Technology-Multimedia Content Description Interface-Part 1: Systems", MPEG document N4001, Singapore, March, 2001.

[40] ISO/IEC JTC 1/SC 29/WG 11: "Information Technology-Multimedia Content Description Interface-Part 2: Description Definition Language", MPEG document N4002, Singapore, March, 2001.

[41] ISO/IEC JTC 1/SC 29/WG 11: "Information Technology-Multimedia Content Description Interface-Part 3: Visual", MPEG document N4003, Singapore, March, 2001

[42] ISO/IEC JTC 1/SC 29/WG 11: "Information Technology-Multimedia Content Description Interface-Part 4: Audio", MPEG document N4004, Singapore, March, 2001.

[43] 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.

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

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

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

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

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

[49] 이동호, 유광석, 김희율, "컬러와 모양 정보를 이용한 캐릭터 이미지 검색", 방송공학회

논문지, Vol.5, No.1, pp.50~60, 2000.
[50] 이상미, "MPEG-7 기술을 이용한 내용 기반 정보검색의 활용", 통신소프트웨어 학술대회, Vol.1 No.1, pp.253~256, 1997.
[51] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann, pp.81~113, 2001.
[52] Jain A. K, "Fundamentals of Digital Image Processing", Englewood Cliffs. NJ: Prentice-Hall, 1989.
[53] Otha Y, T. Kanade and T. Sakai, "Colour Information for Region Segmentation", Computer Graphics and Image Processing, pp.222~241, 1980.
[54] Itten. J, "색체의 예술", 1976.
[55] Alberto Del Bimbo, "Visual Information Retrieval", Morgan Kaufmann Publishers, pp.97~99, 2001.
[56] M.A. Stricker and M. Orengo, "Similarity of Color Images". Storage and Retrieval for Image and Video Databases (SPIE), pp.381~392, 1995.
[57] 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.
[58] 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
[59] 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
[60] Smith J. R. and S. F. Chang, "VisualSEEK: A fully automated content-based image query system", ACM multimedia'96, Boston, MA. 1996. 11
[61] Alberto Del. Bimbo, "Visual Information Retrieval", Morgan Kaufmann, pp.117~131, 2001.
[62] 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

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

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

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

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

[67] 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

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

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

[70] 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.

[71] 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.

[72] 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.

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

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

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

[76] Randy crane, "A simplified approach to image processing", Prentice-Hall, pp.18

~29, 1997.

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

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

[79] 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.

[80] C. Schmid and R. Mohr, "Matching by Local Invariants", Technical report N2644, IN R IA , 1995.

[81] C. Schmid and R. Mohr, "Local Gray value Invariants for Image Retrieval", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.19, No.5, pp.530 ~ 535, 1997.

[82] http://www.w3.org/XML/

[83] http://www.w3.org/DOM/DOMTR/

[84] http://www.w3.org/TR/xlink/

[85] 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.

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

[87] 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.

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

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

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

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

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

[93] 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.

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

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

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

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

[98] 김재건, "웹 영상 검색 시스템", 경상대학교대학원 석사학위 논문, 2001.

[99] Van Rijsbergen, C. J., "A non-classical logic for information retrieval", Computer journal, 29(6), pp.481~485, 1996.

[100] Niblack, W. R. Barber, W. Equitz, M. Flickner, D. Petkovic and P. Yanker, "The QBIC project Querying images by content using colour, texture and shape", In SPIE, Storage and Retrieval for Image and Video Database, 1993.

[101] Ogle V. E. and M. Stonebraker, "Chabot: Retrieval from a relational database of images", IEEE Computer, 28(9), pp.40~48, 1995.

[102] Zhong Y, Zhang HJ & Jain AK,, "Automatic caption localization in compressed video". ÆEE Trans on PAMI, 22(4), pp.385~392, 2000.

(별 지)

저작물 이용 허락서

학 과	전자공학과	학 번	19941080	과 정	박사
성 명	한글: 강성관 한문 : 姜成官 영문 : Kang Sung Kwan				
주 소	광주광역시 광산구 월곡1동 596-3				
연락처	E-MAIL : arome@netian.net				
논문제목	한글 : 특징 기술자와 의미론적 키워드를 이용한 웹 기반 이미지 검색 시스템 설계 영문 : Design of the Web based Image Retrieval System Using Feature Descriptors and Semantic Keyword				

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.

2005 년 6 월 일

저작자: 강 성 관 (서명 또는 인)

조선대학교 총장 귀하