



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2017년 8월
박사학위논문

질의 영상 특성에 기반한
적응적 웨이트드 CBIR 알고리즘 개발

조선대학교 대학원

정보통신공학과

오상언

2017년 8월

박사학위논문

영상특성 기반 정영자 웨이브 기반 압간검

오상언

질의 영상 특성에 기반한 적응적 웨이티드 CBIR 알고리즘 개발

Design of Adaptive Weighted CBIR Algorithm
based Features of Query Image

2017년 8월 25일

조선대학교 대학원

정보통신공학과

오 상 언

질의 영상 특성에 기반한 적응적 웨이티드 CBIR 알고리즘 개발

지도교수 박 종 안

이 논문을 공학박사학위신청 논문으로 제출함

2017년 4월

조 선 대 학 교 대 학 원

정보통신공학과

오 상 언

오상언의 박사학위 논문을 인준함

위원장 조선대학교 교수 한 승 조 (인)

위 원 조선대학교 교수 권 구 락 (인)

위 원 전남도립대학교 교수 천 종 훈 (인)

위 원 조선대학교 교수 안 영 은 (인)

위 원 조선대학교 교수 박 종 안 (인)

2017년 6월

조선대학교 대학원

목 차

목 차	i
그림 목 차	iii
표 목 차	vi
ABSTRACT	vii
I. 서론	1
A. 연구 배경 및 목적	1
B. 논문 구성	3
II. 내용기반 영상검색 방법	4
A. 내용기반 영상 검색 시스템의 구조	4
B. 질의 및 검색 방법	5
C. 내용기반 영상검색 시스템	10
D. 영상의 특징 추출 방법	15
III. 적응적 웨이티드 CBIR 알고리즘 제안	39
A. 전처리 과정	40
B. 레이블링 및 객체 판별과정	43
1. 레이블링 및 크기 필터링	43
2. 객체 판별	48
C. 객체 특징 추출	53
1. 조밀도 특징 추출	53
2. 코너점 특징 추출	54
3. 색상 및 농담 특징 추출	57
D. 적응적 웨이티드 부여	61

1. 면적 분석 및 웨이티드 부여	61
2. 코너점 분석 및 웨이티드 부여	63
3. 조밀도 분석 및 웨이티드 부여	64
4. 색상 분석 및 웨이티드 부여	66
5. 적응적 웨이티드	68
E. 유사도 측정 및 순위 결정	68
1. 유사도 측정	68
2. 순위 결정	70
IV. 실험 및 분석	72
A. 실험	72
B. 분석	84
V. 결론 및 향후 연구과제	92
참고문헌	94

그림 목 차

그림 2.1	내용기반 영상검색 시스템의 구조	4
그림 2.2	섬네일 브라우징 방법의 개요도	8
그림 2.3	2D와 3D로 표현된 검색영상	9
그림 2.4	QBIC 시스템	11
그림 2.5	Photobook 시스템	12
그림 2.6	VisualSEEK 시스템	13
그림 2.7	NETRA	14
그림 2.8	RGB 채널의 그레이-레벨 히스토그램	16
그림 2.9	RGB 색상공간의 구조	17
그림 2.10	CCV 기법을 활용한 색상 분포 연결	18
그림 2.11	CCV 기법을 활용한 연결 속성 테이블	19
그림 2.12	Munsell 시스템의 색상표현	21
그림 2.13	Munsell 색상 시스템	21
그림 2.14	HS계열의 색상공간	22
그림 2.15	형태 기반 검색 기법의 분류	28
그림 2.16	저나이크 다항식의 위상	30
그림 2.17	Harris 코너점 검출 결과	34
그림 3.1	제안하는 적응적 웨이트드 CBIR 알고리즘 흐름도	39
그림 3.2	Otsu 방법을 이용하여 이분화한 히스토그램	41
그림 3.3	Otsu 방법을 이용하여 이분화한 샘플이미지	41
그림 3.4	전처리 과정 이미지	43
그림 3.5	화소연결성의 이웃화소	44
그림 3.6	화소 연결성을 이용하여 연결된 화소들	45
그림 3.7	레이블링 된 영상 정보	46
그림 3.8	레이블링 후 분리된 영상들	47
그림 3.9	크기 필터링	48

그림 3.10	면적 특징 추출	49
그림 3.11	배경이 주된 영상의 레이블링 특성과 특징 벡터	51
그림 3.12	객체 판별에서 구분된 영상들	52
그림 3.13	둘레 계산을 위한 경계선 추출	53
그림 3.14	조밀도를 추가한 특징 벡터	53
그림 3.15	고유치 분포에 따른 특징점 분할	56
그림 3.16	코너 부분 검출	56
그림 3.17	코너점의 개수를 추가한 특징 벡터	57
그림 3.18	인덱스 이미지	59
그림 3.19	uint8을 이용한 화소값과 칼라맵	60
그림 3.20	최종 특징 벡터	60
그림 3.21	단일 객체 영상과 다중 객체 영상의 비교	62
그림 3.22	다른 형태의 객체를 가지는 객체 영상의 특징 벡터 비교	64
그림 3.23	다른 조밀도 값을 가지는 객체 영상의 특징 벡터 비교	65
그림 3.24	같은 색상의 분포가 다른 객체 비교	67
그림 3.25	유사영상 검색 결과창	71
그림 4.1	데이터베이스 영상들의 특징 벡터 데이터	72
그림 4.2	해변의 질의 영상	74
그림 4.3	데이터베이스 내 해변 영상의 유사영상들	74
그림 4.4	해변의 검색 결과	75
그림 4.5	산의 질의 영상	75
그림 4.6	데이터베이스 내 산 영상의 유사영상들	76
그림 4.7	산의 검색 결과	76
그림 4.8	질의 영상 1, 2	77
그림 4.9	데이터베이스 내 질의 영상1의 유사영상	78
그림 4.10	데이터베이스 내 질의 영상2의 유사영상	78
그림 4.11	공룡과 집의 검색 결과	79
그림 4.12	나뭇잎의 질의 영상	80
그림 4.13	데이터베이스 내 나뭇잎 영상의 유사 영상들	81

그림 4.14	나뭇잎의 검색 결과	81
그림 4.15	원주민의 질의 영상	82
그림 4.16	데이터베이스 내 원주민 영상의 유사영상들	83
그림 4.17	원주민 영상의 검색 결과	83
그림 4.18	Recall과 Precision	85
그림 4.19	평균 Recall과 Precision 비교 그래프	91

표 목 차

표 2.1	저나이크 모멘트의 목록	31
표 3.1	영상의 타입	58
표 3.2	면적의 웨이티드 부여	63
표 3.3	코너점의 웨이티드 부여	64
표 3.4	조밀도의 웨이티드 부여	66
표 3.5	색상의 웨이티드 부여	67
표 4.1	이미지 데이터베이스의 분류	73
표 4.2	이미지 그룹별 Recall 및 Precision	89

ABSTRACT

Design of Adaptive Weighted CBIR Algorithm based Features of Query Image

Sang-Eon Oh

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

Department of Information and communications Engineering
Graduate School of Chosun University

As smartphones became popular recently, image service users increased, and service quality, including storing, sharing, and syncing data, became high quality and high capacity. Various attempts are needed to process and service large amounts of image data efficiently, as such, production and consumption of large amounts of the image data are rapidly increasing. Because of the development of computer technology and multimedia information, it is easy to acquire and store various forms of image information as well as textual information. Additionally, imaging information has skyrocketed in many areas, while the management is becoming increasingly difficult. When searching for a specific data, the image information is larger than the character information, and it is not easy to find fast and efficient retrievals. In this regard, a new retrieval method is needed to effectively manage the images.

Therefore, in this paper, I propose a useful adaptive weighted CBIR algorithm based features of query image. Proposed algorithm is composed of four steps. First, it is a preprocessing to remove unnecessary background and trim object image information using Otsu method. Also erosion and dilation are used for removing unwanted information from the images. Second, it labels the images

and rearranges them in order of size, normalizing the area, performing size determination process of size filtering. And it is passing through the object discrimination process where the background is classified into the main image and the image where the object is discriminated. If the background is the main image, weights are assigned to use only color information. And if the object is a discriminated image, feature vectors are constructed by extracting feature information of area, compactness, number of corner points, and color value for each labeling. Third, based on the extracted feature vector, area, number of corner points, and compactness of the object are analyzed and the weighted value is adaptively applied according to each feature vector value. And the color pixel values are adaptively weighted using the distribution map of the color map of the index data matrix. Finally, similarity measurement is performed according to the weighted values applied adaptively according to the query image, and the image retrieval was performed.

Since adaptive weighted CBIR algorithm is possible to retrieve according to various query images, proposed algorithm complements the selection limits of the query image that existing algorithms have.

In order to evaluate the performance of the proposed algorithm in this paper, I used Recall and Precision, which are widely used in CBIR. The precision is 0.08 lower than the color histogram statistical method, but the recall is 0.16 higher. Also the precision is 0.05 and the recall is 0.03 higher than the shape template matching method. More over the precision is 0.02 and the recall is 0.07 higher than HSV color and uniform local binary pattern method. The results confirmed that the proposed algorithm has better retrieval performance.

Based on the above, I assume that the proposed algorithm will be able to extract the various features of multimedia data and utilize the optimal database as systems capable of constructing and retrieving.

I. 서 론

A. 연구 배경 및 목적

최근 스마트폰이 대중화됨에 따라 영상 서비스 사용자는 늘어나고, 데이터의 저장과 공유, 동기화를 비롯하여 서비스도 고품질, 대용량화 되었다. 대용량 영상 데이터를 효과적으로 빠르게 처리하고 서비스하기 위한 다양한 연구가 필요하게 되었다. 그리고 소셜 네트워크 서비스의 등장으로 많은 영상 데이터의 생산과 소비가 급증하고 있다. 컴퓨터 기술의 발달과 멀티미디어 정보화에 따라 문자 정보와 더불어 다양한 형태의 영상 데이터를 획득하기 쉽고 저장이 간편하게 되었다. 또한 영상 정보는 여러 분야에서 사용이 급증하였지만 그 관리는 더욱 더 어려워지고 있다[1].

영상 데이터는 찾고자하는 자료를 검색할 경우 문자 정보에 비하여 용량이 크며, 비정형적이어서 빠르고 효율적인 검색이 가능하지 않았다[2][3]. 이런 점에서 영상 데이터를 효율적으로 관리할 수 있는 새로운 검색 방법이 필요하게 되었다.

멀티미디어 정보 검색 분야는 새롭게 주목받고 있다. 멀티미디어 데이터가 가지는 형태적이고 의미적인 내용 정보의 추출과 해석을 기반으로 하여 정보 저장 및 검색 분야는 자연어 등을 통한 인간 중심의 사용자 인터페이스 기능을 포함하며 이에 대한 연구 개발이 활발하다[4][5].

정보 검색은 여러 사용자들을 위한 유저 인터페이스 기술과 멀티미디어 정보 검색 기술 등의 다양한 정보 기술 분야를 포함하고 있다. 그 중 영상인식 기술은 멀티미디어 데이터를 사용자에게 효과적인 제공을 위한 기술이다. 정보 구축 기술은 정보 인프라의 형성에 사용되고 광역 분산 데이터베이스를 구축하며 문자 정보와 멀티미디어 정보를 저장 및 관리한다. 또한 디지털 문서 기술은 디지털화되어 있는 문서들을 유연적으로 이용할 수 있게 한다.

효과적인 영상 검색을 위한 방법들 중 하나는 문자기반(text based) 검색으로 키워드를 이용하여 질의 영상을 검색하는 방법이다[6]. 그러나 이는 대용량의 정보에 대하여 사람이 키워드를 일일이 부여해야한다는 점과 동일한 영상에서도 작성자의 주관적

인 판단에 따라 다른 키워드가 부여될 수 있어 정보의 일관적이고 객관적이지 못하다 [7]. 또한 정보 검색 시 사용자가 질의 영상에 대한 고유 키워드를 파악하고 있어야 하는 단점도 있다.

의미론 기반(semantic based) 검색은 정보를 추상화된 특징으로 나타내는 방법으로 질의어의 구분이나 의미를 분석하여 검색을 하는 방법이다. 그러나 이 방법은 특징을 추출하기 위한 방법의 기발이 적용분야에 따라 달라질 수 있어 제한된 환경에서만 사용이 가능하다[8][9].

내용 기반(content based) 검색은 영상의 내용요소들인 색상, 질감, 물체의 모양 및 공간적 관계 등과 같은 영상의 특징과 속성을 이용하여 영상자료간의 유사성을 계산하고, 계산된 유사도를 바탕으로 영상을 검색하는 방법이다. 이 방법은 영상의 특징과 속성을 자동으로 추출한 특징벡터를 색인정보로 저장하는 시스템을 구현할 수 있다는 장점이 있다[10-14].

내용 기반 영상 검색 시스템에서 사용하는 검색 방법으로는 1992년 Kato[15]에 의해서 제시된 이후 널리 사용되고 있으며, 영상이 가진 시각적인 속성들을 대표하는 특징값을 이용하여 사용자가 원하는 영상을 검색할 수 있도록 하는 것이다. 그러므로 영상을 대표하는 색상, 모양, 질감 등의 저단계 시각 특징들의 효과적인 추출이 무엇보다 중요시된다[16].

색상 정보를 이용하여 특징을 추출하는 대표적인 방법은 컬러 히스토그램을 이용하는 방법이다. 컬러 히스토그램은 계산이 용이하고, 카메라 위치의 작은 변화에도 강건함을 보인다. 이러한 특징으로 인하여 컬러 히스토그램은 장면 검출과 내용 기반 영상 검색과 같은 다양한 멀티미디어 응용 분야에서 사용되고 있다[17].

질감은 색상이나 명암도의 존재만으로 산출되지 않은 유사한 성질을 가지고 있는 시각적 패턴이며 실질적인 모든 표면의 고유한 성질이다[18]. 또한 질감은 표면의 구조적인 배열과 주위 환경에 대한 관계 등의 중요한 정보를 담고 있다. 패턴 인식과 컴퓨터비전 분야에서의 질감 특징의 중요성과 유용성이 부각되어 많은 연구가 진행되고 있다[19].

모양을 이용하는 방법은 주로 윤곽선, 체인코드, 불변모멘트 등을 사용한다[20]. 그

러나 모양을 표현하는 방법의 어려움으로 인해 색상이나 질감보다는 그 이용이 제한된다. 따라서 이 방법은 상표나 로고 검색과 같은 한정된 응용에 주로 사용된다.

현재는 색상, 질감, 모양 등과 같은 특징 추출을 각각 이용하기 보다는 여러 가지의 방법을 결합하여 각각의 단점을 보완하는 방법이 많이 사용되고 있다[21-23].

본 논문에서는 내용 기반 영상 검색 시스템에서 영상의 내용을 기반으로 형태, 색상, 코너점, 레이블링, 조밀도 등의 검색하고자 하는 데이터베이스 영상들의 특징 벡터들을 추출하고, 질의 영상의 특징 벡터를 자동으로 분석하여 영상을 분류한다. 분류한 질의 영상에 따라서 적응적으로 서로 다른 웨이트드를 부여하여 유사 영상을 검색하는 효율적인 적응적 웨이트드 CBIR 알고리즘을 제안한다.

B. 논문 구성

본 논문의 구성은 다음과 같다. II장에서는 내용 기반 검색 방법의 구조와 시스템들 그리고 특성 추출의 방법에 대해서 기술하며, III장에서는 제안하는 영상 특성 기반 적응적 웨이트드 CBIR 알고리즘을 설계하고 IV장에서는 이를 실험을 통해 분석을 하며 마지막으로 V장에서 결론을 맺는다.

II. 내용기반 영상검색 방법

A. 내용기반 영상검색 시스템의 구조

내용기반 영상검색 시스템은 기존의 문자기반 영상 검색의 단점을 극복하기 위해 1998년 Kato[15]에 의해 제시되었다. 내용기반 영상 검색은 특성 요소들인 색상, 질감, 물체의 모양과 같은 낮은 수준의 시각 특성 값을 영상 데이터 간의 유사도를 계산하고, 계산된 유사도를 기준으로 영상을 자동적으로 검출하여 영상의 시각적 특성 요소를 요약하여 자동적으로 색인화한다[24]. 그리고 질의 검색 자료로 영상 또는 영상의 관련 특성 정보를 이용하여 분석한다. 영상 데이터베이스의 구성에 이용한 특징 추출 방법을 사용하여 특징 정보를 추출한 후에 영상 데이터베이스에서 가장 유사한 영상을 검색하여 보여 준다. 그림 2.1은 내용기반 영상검색 시스템의 일반적인 구조를 나타내고 있다[25].

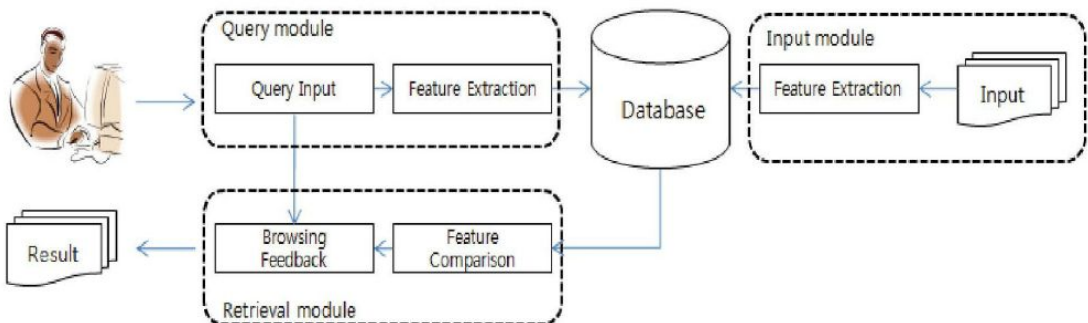


그림 2.1. 내용기반 영상검색 시스템의 구조[25]

영상검색 시스템은 크게 영상의 입력, 질의와 검색의 과정으로 구성되며, 검색결과를 사용자에게 시각적으로 보여줄 수 있는 영상들로 구성된다. 영상의 입력과정은 입력된 영상이 가지고 있는 특징을 추출하여 색인화를 거쳐 영상 데이터베이스를 구축하는 데 이용한다[26]. 사용자가 검색하고자 하는 질의 영상에 대한 특징을 추출하고 데

이터베이스의 구축 시 이용하는 방법과 동일한 방법을 사용하여 같은 종류의 특징을 추출해야 한다.

전체적인 내용기반 영상검색 시스템의 처리 과정은 사용자 인터페이스와 프로세싱으로 구성된다. 사용자 인터페이스는 일반적으로 GUI(Graphic User Interface)로 구성된 시스템에서 질의 영상을 입력받는다. 이 후 영상에 관련된 정보에 대해 프로세싱 단계에서 특징 추출 데이터베이스의 구축에 이용한 특징 추출 방법에 의해 동일하게 특징 벡터를 생성한다[27-29].

검색과정은 특징 추출 데이터베이스 내에 있는 특징 벡터들과 사용자 인터페이스에서 추출한 질의 영상의 특징 벡터를 비교하고 특징 벡터간의 거리 및 유사도를 계산하여 질의 영상의 특징과 데이터베이스 내의 영상들의 특징을 비교하여 가장 유사도가 높은 특징을 가지는 영상들을 우선적으로 출력한다. 특징들은 문서기반의 특징들인 키워드나 주석을 비롯하여 색상, 질감, 모양과 같은 내용기반 특징들을 모두 포함한다 [30][31].

B. 질의 및 검색 방법

질의 영상을 원활히 검색하기 위해선 데이터베이스에 있는 영상자료의 접근이 편하여야 한다. 데이터베이스의 영상자료 접근 위한 질의 및 검색하는 형식에는 다양한 방법이 있다. Gudiva 및 Jain 등은 이러한 여러 종류의 질의 및 검색방법을 잘 구분하고 있다. "영상검색 전략"이라는 주제로 접근하며, 질의와 브라우징 및 네비게이션과 같은 용어를 사용한다[32][33].

1. 질의방법

주제어에 대한 질의는 영상검색시스템에서 가장 많이 사용하고 있는 방법이다. 디지털 카메라를 제작 및 판매하는 제작자 또는 디지털 앨범과 같은 영상을 취급하는 관련 업자들은 사진을 관리하기 위해 주제어에 기반한 문자기반의 검색도구를 사용하고 있다. 이러한 방법들은 미리 주제어를 선정해 놓고 사진을 분류할 때 그 주제어에 따라

서 저장되도록 한다. 주제어 분류 방법은 통상적으로 다층적인 구조를 형성하고 있다. 사용자가 "풍경화"라는 메뉴를 선택하면 사막, 호수, 강, 바다와 같은 서브 메뉴를 보여 주어 계층적으로 그룹화된 관리가 가능하도록 하고 있다.

사용자 인터페이스에서 다양한 예시를 준비해 두고 사용자가 선택하면 데이터베이스 내에 있는 영상들과 서로 비교한다. 사용자가 편리하게 사용하도록 다양한 유사도 측정방법들의 선택이 가능하도록 질의방법을 개선한 시스템들도 있다. 예시에 의한 질의 방식은 전반적으로 다음의 3가지 형식을 취한다. 즉, 첫째로 영상의 특성 요소들(색상, 질감, 모양 등)을 선택할 수 있는 기능 (예, QBIC), 둘째로 특성 간에 가중치를 조절할 수 있는 기능 (예, Excalibur), 셋째로 특성 요소 및 가중치를 시스템이 전부 결정 (예, Image Surfer)하여 검색하는 방법이다. 예시로 주어진 영상들은 데이터베이스 내에 저장되어 있는 영상들 중 임의로 선정하거나 대표적인 특성 요소를 갖는 영상들로 미리 나타내져 있어 질의 영상을 선택하기만 하면 특성추출 과정을 별도로 거치지 않아도 되기 때문에, 시스템에서는 매우 효율적인 방법이라고 할 수 있다. 그러나 사용자가 검색하고자 하는 질의영상들과 비슷한 영상이 예시로 주어진 영상 중에서 찾을 때에는 계속 다른 예시 영상들을 받아서 유사한 영상을 선택하거나 질의영상과 다소 차이가 있는 영상들을 선택할 수밖에 없기 때문에 한 번에 만족스러운 결과를 얻기가 어려운 단점이 있다.

사용자가 영상의 특징이나 속성을 직접 설정하여 질의를 수행하는 방법이 있다. 이 방법은 영상 처리에 대한 사전지식이 필요하며 영상 내에서 어떠한 특성들이 구성되어 있는 지에 대한 이해가 필요하다. 이러한 방법을 이용하고 있는 QBIC시스템을 예를 들면, 색상 히스토그램을 이용한 질의 중에서 사용자가 “17 %의 푸른색과 15 %의 노란색을 가진 영상을 찾아라.”하는 식으로 사용할 수 있다. 현재 QBIC에서는 색상 히스토그램 방법에서 사용자의 편리성을 위해 색상도표와 색상 점유율을 도식적으로 나타내는 방법을 사용하고 있다[34][35].

사용자가 예시 영상 대신 간단하게 영상의 내용을 스케치한 정보를 질의로 사용한다. 사용자 인터페이스에는 그리기 도구를 제공하여 편리하게 스케치할 수 있도록 한다[36-39]. 스케치하는 방법에는 사용자가 윤곽을 직접 스케치하고, 윤곽의 내부를

인터페이스에서 색상으로 채우는 방법과, 제공된 그리드를 이용하여 원하는 형태의 그리드를 선택하고 시스템에서 스케치를 완성하는 기법을 사용한다.

사용자가 직접 작성한 질의영상은 각 물체의 색상, 모양, 윤곽선, 배치 등 다양한 특성 요소들을 표현할 수 있기 때문에 사용자가 원하는 특성을 편리하고 자유롭게 표현할 수 있다는 장점이 있다. 그러나 이러한 방식의 문제점은 사용자가 직접 작성한 질의영상에서 표현한 특성 요소가 사용자의 기억 왜곡이나 작성과정에서의 잘못으로 인해 찾고자하는 영상과 많은 차이가 있을 수 있다는 점이다. 이 방식을 이용한 시스템에서는 사용자가 작성하는 질의의 형태를 제한하거나[39], 사용자가 직접 작성하는 질의가 어느 수준 이상으로 찾고자하는 영상에 근접한 형태로 입력된다는 것을 전제로 하여 개발을 진행하고 있으며, 질의에 오류가 있는 경우를 대비한 처리문제는 향후의 연구과제로 남겨놓고 있다[37].

인터넷의 발달로 웹에서 정보들, 특히 영상정보를 검색하기 위한 요구가 급증하고 있기 때문에 전 세계에 분포되어 있는 웹사이트들로부터 영상 검색을 위한 메타 검색 엔진이 필요해진다. 콜럼비아 대학에서 개발된 MetaSeek는 이러한 메타 검색엔진의 대표 시스템으로 VisualSEEK, WebSEEK, QBIC 및 Virage 등의 내용기반 영상검색 시스템에 의해 운용되고 있다.

2. 브라우징

영상자료를 화면에 표시하기 위한 방법으로 가장 쉽게 이용할 수 있는 방법은 영상을 한 장씩 나타내 주는 방법이 있으나, 이는 비효율적이라는 점 때문에 잘 사용되지 않는다. 가장 일반적으로 사용되고 있는 브라우징 방법은 섬네일(thumbnail) 브라우징과 슬라이드 쇼인 두 가지 형태로 나타낼 수가 있다.

섬네일 브라우징은 하나의 화면에 다수의 영상을 나타내기 위해 작은 크기로 축소된 섬네일 영상을 사용한다. 섬네일 브라우징은 화면에 나타난 영상을 클릭하게 되면 실제크기의 영상이나 화면 크기로 재설정된 영상을 표시해 준다.

슬라이드 쇼는 실제크기의 영상이나 화면크기로 재설정된 영상을 한 장씩 화면에 보여 주는 방법으로 화면이 바뀌는 영상들 사이의 슬라이드 시간을 조절할 수 있도록 한

다. 일반적인 영상검색시스템에서 채택하여 사용하고 있는 브라우징 방법은 대부분 섬네일 브라우징 방법을 사용하고 있는데 반해, 슬라이드 쇼 방법은 특별한 용도 외에는 거의 사용되지 않는다.

내용기반 영상검색시스템에서 채택되어 많이 사용하고 있는 섬네일 브라우징 방법을 표시하면 그림 2.2와 같다. 이는 통상적으로 화면의 이동을 편리하게하기 위해 이전 또는 다음과 같은 마우스 버튼을 제공한다. 축소된 영상들 중에서 관심 있는 영상을 확대하고 싶은 경우에는 해당하는 영상화면 부분을 클릭 하면 확대된 영상(원 영상)을 볼 수 있다[40].

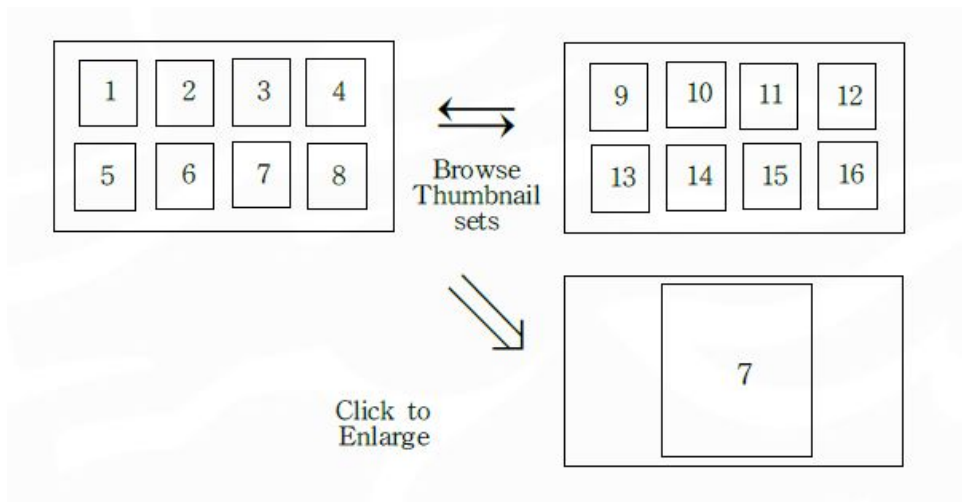


그림 2.2. 섬네일 브라우징 방법의 개요도[40]

3. 네비게이션

우리들은 인터넷을 ‘정보의 바다’로 비유하는 것을 자주 접할 수 있다. 인터넷에서 정보 탐색을 위해서는 이곳저곳을 탐색해야 하는데, 이러한 의미에서 ‘네비게이션’이라는 용어를 사용한다. 사용자가 영상 검색을 위해 특정시스템을 이용할 경우 인터넷의 영상정보공간을 향해한다는 점에서 고안된 것이다.

현재 사용되고 있는 방식의 인터넷 브라우저는 직관에 의한 기하학적인 성질들을 자연스럽게 유도해 내지 못하고 있으며, 유사 영상과의 유사도 수치에 의해 검색 순위를

결정하고 있다. 이런 점들의 보완을 위해 네비게이션이라는 용어가 탄생되었으며, 이에 대한 새로운 이론이 다양하게 제안되고 있다. 그림 2.3에서 보는 바와 같이 데이터베이스 내의 자료나 검색된 유사 영상들을 나타내기 위해 이차원 또는 삼차원의 방법을 사용하여 표현한다[40].

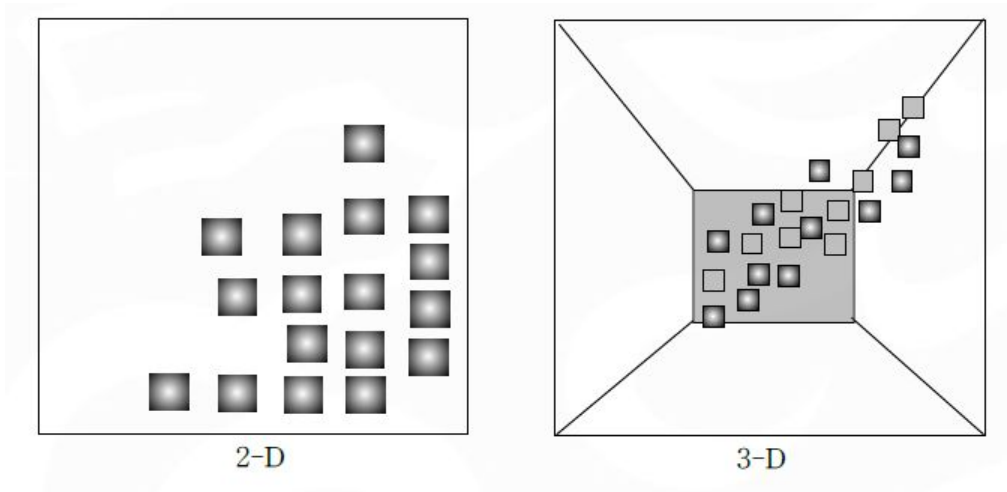


그림 2.3. 2D와 3D로 표현된 검색영상[40]

C. 내용기반 영상검색 시스템

1. QBIC(Query by Image Content)

QBIC 시스템은 IBM 연구소에서 개발한 최초의 상업적인 내용기반 영상검색 시스템으로 현재까지 다른 내용기반 영상검색 시스템들에 많은 영향을 주고 있다. QBIC 시스템의 질의 방법은 위에 언급된 질의 방법들을 모두 지원한다. 이외에도 영상의 컬러 히스토그램이나 레이아웃, 색상 및 질감 등에 의한 검색이 가능한 시스템이다 [41][42].

QBIC에서 사용한 색상 정보는 RGB로 계산하여 64단계의 양자화를 하여 컬러 히스토그램을 계산하여 구한다. 질감 정보는 통계적 특징인 거침, 대비, 방향 특징들을 사용하여 구한다. 모양 특징은 원, 이심율 등을 사용하여 구한다. 특히 컬러 히스토그램 차를 위해 사용하는 측정 기준은 식(2.1)과 같이 정의한다.

$$\begin{aligned}
 D^2 &= z^T A_z \\
 A &= a_{ij} = \left(1 - \frac{d_{ij}}{d_{\max}} \right)
 \end{aligned}
 \tag{2.1}$$

여기서, A 는 컬러-거리 매트릭스로 정의되며 d_{ij} 는 3차원 색의 차이다. 또한 Z 는 히스토그램 차분 벡터를 의미하고 T 는 차원을 나타낸다.

데이터베이스 내에 있는 영상들에 사람이 의미 정보를 직접 부여할 수 있도록 주제어에 의한 검색 또한 제한적으로 지원한다. 하지만 기본적인 특징 요소들과 의미 정보를 표현하는 논리적인 특징 요소들이 체계적이지 못하며, 복잡하고 다양한 질의 방법들에 대한 상호 인터페이스가 통합되어 있지 않기 때문에 측정에 어려움이 있다[43].

그림 2.4는 QBIC 시스템의 주요 질의 방법인 색상을 이용한 시스템을 나타내고 있다. 검색은 주로 사용된 질의 영상을 가지고 하는 방법이 아니라 색상이나 모양 특징 요소를 임의로 정해주고 영상 데이터베이스 내에서 일치하는 영상을 보여주는 방식을 사용하고 있다. 그러나 기본적인 특징 요소들과 의미 정보를 표현하는 논리적인 특징 요소들이 체계적이지 못하고, 질의 인터페이스가 복잡하기 때문에 전문가를 제외한 일

반 사용자들이 이용하기 어렵다는 단점도 있다[44-46].



그림 2.4. QBIC 시스템[46]

2. Photobook

Photobook은 MIT공대에서 개발한 내용기반 영상검색 시스템으로 KL(Karhunen-Loe-ve) 변환을 사용하여 주성분 값으로 표현한다. 그림 2.5은 Photobook 시스템을 나타낸다[47]. 이 시스템은 모양, 질감 그리고 얼굴 특징 등을 각각 추출하여 3개의 서브 시스템을 구성하고, 질의 시에 사용자는 3개의 서브시스템들의 각각에 대응하는 특징 요소에 기반을 두고 검색을 한다. 또한 영상 브라우징인 모티프 인터페이스(Motif Interface)를 사용하여 영상의 카테고리 선택이 가능하며, 객체지향 메모리 기반 AI 데이터베이스를 사용하여 주석 검색을 한다[48-50].

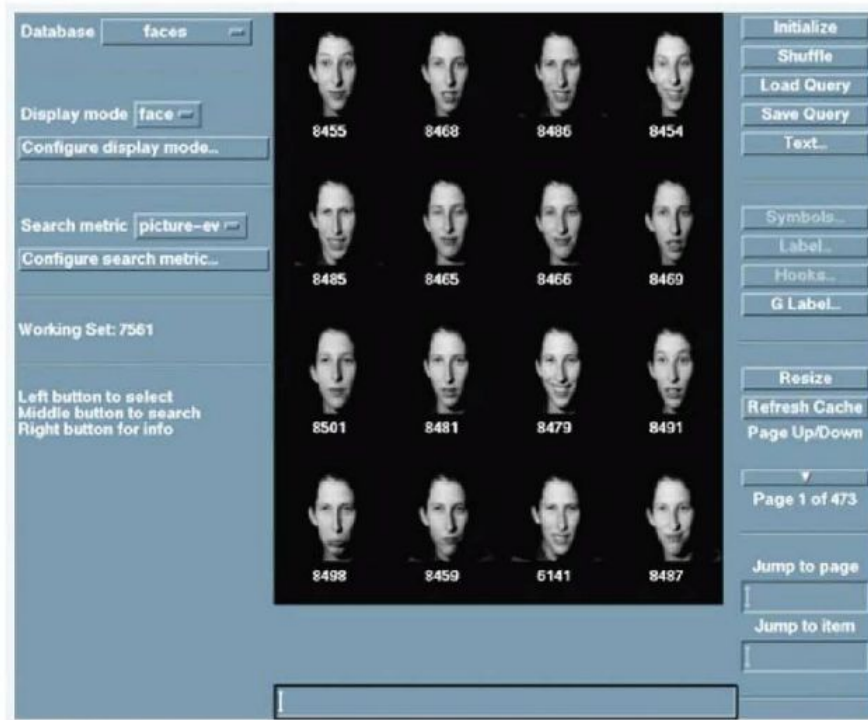


그림 2.5. Photobook 시스템[47]

3. VisualSEEK와 WebSEEK

VisualSEEK[51]와 WebSEEK[52]는 콜럼비아 대학에서 개발한 내용 기반 영상 검색 시스템과 웹 기반 문서의 영상 검색 시스템이다. 그림 2.6은 VisualSEEK 시스템의 예이다.

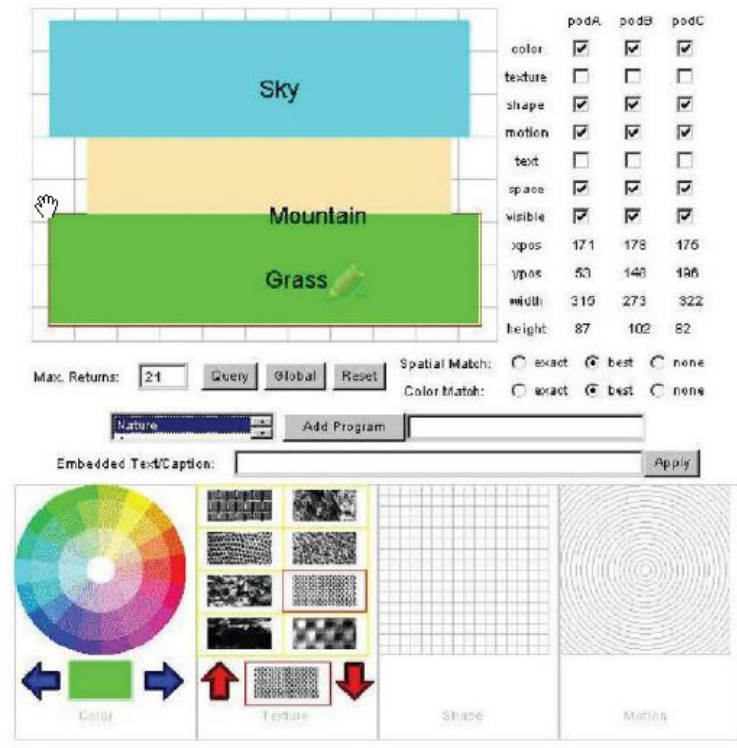


그림 2.6. VisualSEEK 시스템 [51]

초기 인터페이스에서는 저장되어 있는 데이터들을 분류한 메뉴 화면을 이용하여 주제별로 영상 검색을 할 수 있도록 하였다. 문자를 넣어 검색이 가능한 키워드 기반 검색 기능과 영상 공간 정보를 이용한 검색 기능을 제공하는 SAFE 엔진을 포함한다. 색상 특징을 이용하여 영상 검색을 하며 데이터베이스 내에 있는 영상들은 가장 우세한 색상 값에 의해 영역이 분리되며 각 영역별로 위치 벡터와 특징 벡터를 추출한다. 이후 양자화된 컬러 집합을 이용한 역 투영 기술(back-projection)을 이용하여 영역으로 분리한다[53-55].

4. NETRA

미국 캘리포니아 산타바바라 대학교에서 개발한 시스템으로 영상 데이터베이스 내에 있는 영상들을 동일한 색상 공간 영역으로 나누고 각 영역별로 색상, 모양, 공간 레이

아웃 그리고 질감의 특징값들을 추출하여 영역별 유사성을 비교한다.

영상은 RGB 색상 공간으로 양자화되고 256개의 색상을 갖는 코드북으로 표현된다. 영상 영역별 색상 역시 양자화되고 색상 코드북에 의해서 특징 벡터로 표현된다. 질감 특징은 영상에 대한 가버 웨이블릿 변형을 이용하여 정규화된 평균값과 표준 편차에 의해 구해진다. 영역의 형태를 표현하기 위해서는 세 가지의 특징 벡터인 영역의 윤곽선 곡률 함수와 각각의 윤곽점에 주어진 중심 거리 함수, 그리고 마지막으로 x축 좌표에 일치하는 윤곽선의 점들과 y축 좌표에 일치하는 윤곽선의 점들로 되어있는 복합 위치 좌표 함수를 이용한다. 영상의 유사도는 RGB 색상 공간에서 가중치를 부여한 유클리디안 거리값을 이용하여 구해진다.

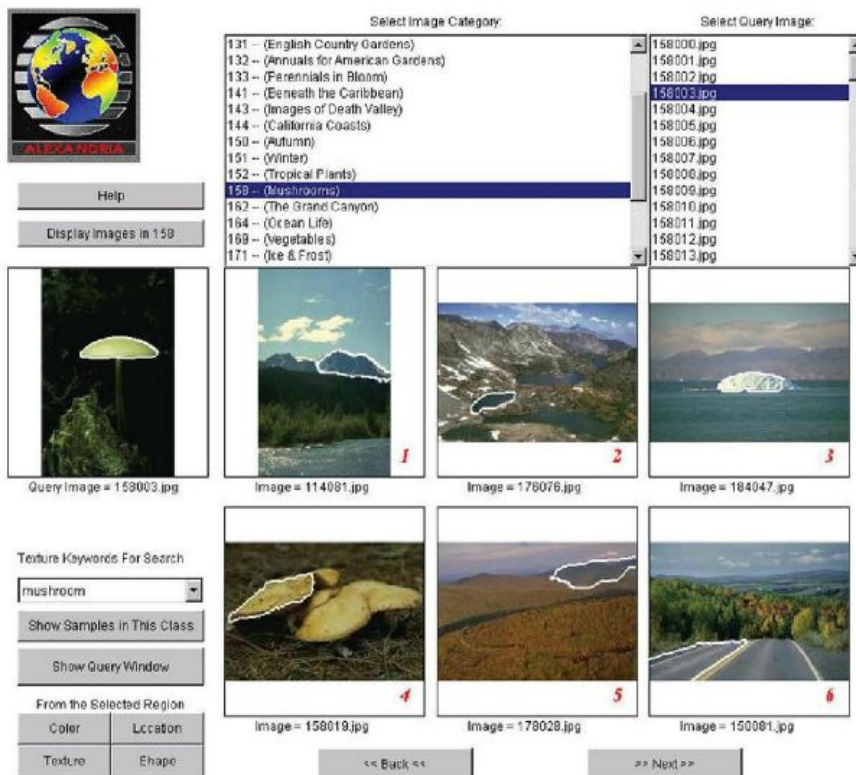


그림 2.7. NETRA[56]

NETRA 시스템은 영상을 유사한 성질을 가지는 영역들로 나누어, 영역별로 유사도를

비교하는 영역기반 영상 검색을 바탕으로 개발되었다. 하지만 영상 분할 시 너무 많은 개수의 영역을 생성하게 되어 질의 영상과 데이터베이스 영상들에 대해서 유사도를 비교할 때 효율적이지 못하다는 단점이 있다[56].

5. Chabot

버클리 대학에서 개발하였으며 관계형 데이터베이스 시스템을 기반으로 내용기반 영상검색 시스템의 특징 요소인 문자와 색상 특징을 사용하여 영상을 검색한다[57]. Chabot에서 영상은 고해상도부터 저해상도까지 다섯 단계의 영상 형태로 서비스되므로 방대한 저장 용량이 필요하다. 이 시스템의 특징은 질의영상에 대해 색상 히스토그램을 구한 후 색상에 대해 적절한 키워드를 기술하여 영상을 검색한다. 키워드를 이용하여 기존의 관계형 데이터베이스 시스템에 있는 키워드들과 연결하여 검색할 수 있는 장점도 있다. 그러나 내용기반 검색에서는 색상 정보만을 사용하여 단조롭다는 단점이 있다[58-60].

D. 영상의 특징 추출 방법

1. 색상을 이용한 특징 추출

색상은 영상의 특징을 가장 쉽게 표현할 수 있는 특성이며 영상검색 시스템에 있어서 가장 널리 사용된다. 색상 특징은 가장 중요하며 낮은 수준의 정보를 가지고 있다. 또한 영상의 이동 및 회전에 불변하며 영상의 크기나 각도의 변화 등에 영향을 많이 받지 않는다[61].

영상을 표현하기 위해 전체적인 색상특성과 물체의 형태를 조합하여 표현하는 국부적인 색상특성을 통하여 나타낼 수 있다. 질의 영상을 색상으로 나타내기 위한 가장 통상적인 방법은 색상의 밝기를 히스토그램으로 나타내는 것이다. 예를 들면, 그레이스케일 영상은 그림 2.8에서 보는 바와 같이 일정 수준의 그레이레벨을 갖는 화소들의 수를 누적하여 나타낸 히스토그램으로 변환할 수 있다. 즉, 원래의 색상 영역을 RGB 각 채널의 그레이레벨 영상으로 분리한 후, 각각의 채널별로 0에서 255까지의 밝기를

가진 화소들을 누적 계수하여 히스토그램으로 표시한다.

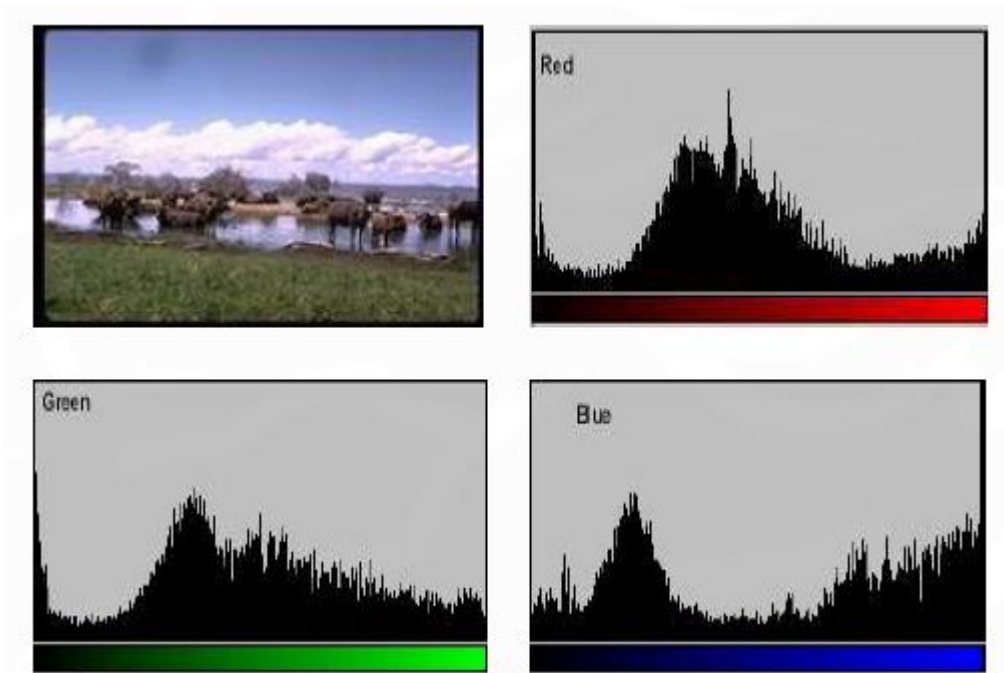


그림 2.8. RGB 채널의 그레이-레벨 히스토그램

히스토그램을 영상검색기법으로 최초로 도입한 Swain에 의하면, "히스토그램은 영상의 이동과 회전에 불변하며, 영상의 크기, 카메라 각도의 변화, 물체의 흡수(occlusion) 등에 덜 민감하다. 그러므로 히스토그램은 영상검색을 위한 기법으로 영상간의 유사도를 표현하는데 사용하는 것이 가능하다."고 주장하며, opponent RGB공간에서 색상 히스토그램에 관한 다양한 실험을 하고 있다[62]. 여기에서 Swain이 영상검색에서 이용한 opponent RGB공간은 인간의 시각체계와 유사한 rg , by , ub 축을 사용한다. 이러한 공간을 식으로 표현하면 식 (2.2)가 된다.

$$rg = r - g$$

$$by = 2 \times b - r - g \tag{2.2}$$

$$wb = r + g + b$$

식에서 r, g, b 는 RGB공간의 각 채널을 표시한다. 여기에서 RGB 색상공간이란 매우 간단한 정육면체의 구조를 이루며, 정육면체의 각 축은 색상의 밝기를 나타내는 영역으로서 그림 2.9와 같다.

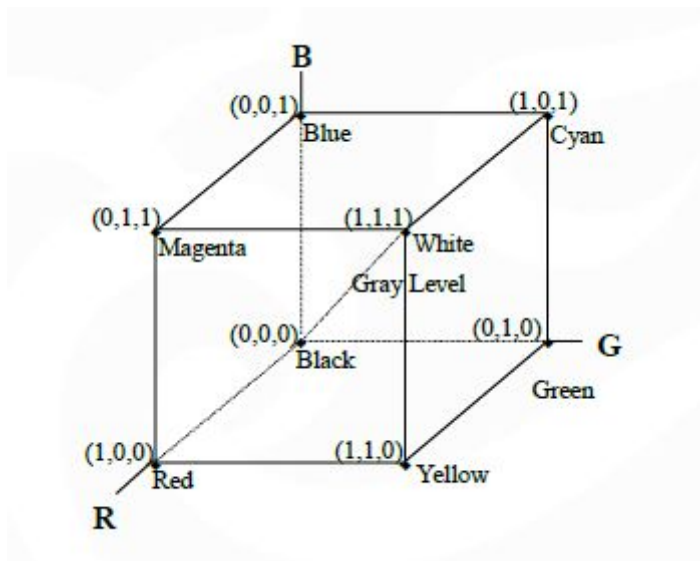


그림 2.9. RGB 색상공간의 구조

영상간의 유사도를 표현하기 위해 색상 히스토그램을 사용하는 경우 조명조건이나 해상도의 변화가 있으면 히스토그램에 영향을 미치게 되므로, 이런 영향들을 줄이고자 두 히스토그램간의 비교는 식 (2.3)과 같은 히스토그램 인터섹션을 사용한다.

$$H(I,M) = \sum \min(I_j, M_i) \tag{2.3}$$

여기에서 I , M 은 각각 질의영상과 목표영상, j 는 히스토그램 빈이다. 위의 식에서 알 수 있듯이 히스토그램 인터섹션의 결과는 두 영상간의 색상 빈에서 같은 색상의 화소수를 나타낸다. 영상의 크기와 관계없이 결과가 0과 1 사이의 값을 가지게 되도록 정규화하기 위한 식을 나타내면 식 (2.4)와 같다.

$$H(I, M) = \frac{\sum \min(I_j, M_j)}{\sum_{j=1}^n M_j} \tag{2.4}$$

히스토그램 인터섹션방법은 질의영상의 색상 히스토그램을 데이터베이스 내의 유사 영상들의 히스토그램과 비교해 보는 검색방법이기 때문에 히스토그램의 계산만 필요하여 간단하다는 장점을 지니고 있다. 색상 히스토그램에 대한 연구들 중에는 Stricker에 의한 히스토그램의 용량에 관한 연구가 있다. 이 논문에서는 누적 히스토그램을 제안한다[63]. 히스토그램은 색상의 밝기가 히스토그램으로 변환될 때 영상에서 가지고 있는 공간정보는 쓰지 않게 되므로 히스토그램 비교방법을 영상검색에서 이용할 경우 공간정보를 필요로 하는 영상의 검색결과는 좋지 않게 된다.

공간정보 문제를 해결하는 방법으로 영상의 밀집과 분산관계를 이용하여 화소들을 분리하는 CCV방법을 Pass 등이 제안하였다. 그들은 "유사하지 않는 두 개의 영상일 지라도 히스토그램 상으로 비슷한 결과가 나타날 수 있다"고 주장하며, CCV기법을 활용하여 이를 보완하고자 하였다[64]. CCV 방법에서는 색상별로 양자화하여 영상 내의 색상이 뭉쳐져 있는지 아니면 분산되어 있는지를 구분하여 색상의 응집 및 분산 상태를 구분한다. 예를 들면, 그림 2.10에서 보는 바와 같이 임의의 화소값 10과 19사이를 1, 20에서 29까지를 2 같이 나타내어 진행하며, 임의의 화소의 상하좌우에 있는 화소들간의 연결여부를 문자로 표시하면, 화소들의 응집여부를 그림 2.11와 같이 작성할 수 있다.

22 10 21 22 15 16	2 1 2 2 1 1	B C B B A A
24 21 13 20 14 17	2 2 1 2 1 1	B B C B A A
23 17 38 23 17 16	2 1 3 2 1 1	B C D B A A
25 25 22 14 15 21	2 2 2 1 1 2	B B B A A E
27 22 12 11 21 20	2 2 1 1 2 2	B B A A E E
24 21 10 12 22 23	2 2 1 1 2 2	B B A A E E

그림 2.10. CCV 기법을 활용한 색상 분포 연결

Label	A	B	C	D	E	⇒	Color	1	2	3
Color	1	2	1	3	2		Coherence	12	20	0
Size	12	15	3	1	5		Incoherence	3	0	1

그림 2.11. CCV 기법을 활용한 연결 속성 테이블

그림 2.11에서 색상 (1, 2, 3)에 대한 화소들의 분포는 [(12, 3),(20, 0),(0, 1)]과 같이 표현되며, 이를 해당되는 영상의 특성으로 이용이 가능하다. 여기서 집합요소인 숫자들의 조합 중 앞의 수는 응집을, 뒤의 수는 분산을 나타낸다. 그리고 응집여부를 판단하기 위한 값은 임의의 값인 4로 한 경우를 예로 들었다. 즉, 1로 구분된 자료들 중 A 항목에 포함되는 자료의 수는 4보다 크기 때문에 응집으로 분류하고, C 항목에 분류된 자료는 4보다 작은 3의 값을 가지기 때문에 분산으로 분류한다.

공간정보를 고려하기 위한 다른 방법으로 Huang은 Color Correlogram을 제안하고 있다. 이는 임의의 한 화소 i 가 k 거리만큼 떨어진 화소 j 를 찾을 확률을 상관관계로 작성한 테이블을 특성벡터로 이용한다. 그들에 의하면, Color Correlogram은 다음과 같은 성질을 가진다.

- i) 색상의 공간 상관관계를 포함
- ii) 계산하기 쉬움
- iii) 특성벡터의 크기가 매우 작은 장점이 있음[65][66].

또한 Stricker는 영상을 다섯 개의 영역으로 분할하여 영역들이 각각 최소한의 공간 정보를 포함할 수 있도록 한다. 이러한 영역들로부터 색상 모멘트인 평균, 분산, 왜곡을 계산하여 특징벡터로 사용함으로써 영상의 이동과 회전에 강한 알고리즘을 제안하였다[67]. 이들이 모멘트를 계산하는 식은 다음과 같다.

$$E_{l,i} = \frac{1}{NR_l^{tot}} \sum_p R_l(P) C_i(P) \quad (2.5)$$

$$\sigma_{l,i} = \left(\frac{1}{NR_l^{tot}} \sum_p R_l(P) (C_i(P) - E_{l,i})^2 \right)^{\frac{1}{2}} \quad (2.6)$$

$$S_{l,i} = \left(\frac{1}{NR_l^{tot}} \sum_p R_l(P) (C_i(P) - E_{l,i})^3 \right)^{\frac{1}{3}} \quad (2.7)$$

여기에서 N은 전체영상의 화소수이며, R_l 은 fuzzy 영역의 멤버십 함수, $C_i(P)$ 는 임의의 화소 P에서 i 번째 채널, $R_l^{tot} = \sum_{j=1}^N R_l(j)$ 이며, $E_{l,i}$, $\sigma_{l,i}$, $S_{l,i}$ 는 각 영역별 평균 색상, 분산 및 왜곡을 나타내고 있다. 영역의 경계에 가까운 위치의 화소들은 중앙에 위치한 화소들 보다 가중치를 더 적게 주어 함수를 조절하여 영상의 회전과 이동에 각 모멘트가 완만하게 변하도록 하였다.

Gong은 RGB의 색상공간이 가지는 좌표계가 인간에게 의미정보를 직관적으로 전달하여 주지 못하기 때문에 인간의 시각인식체계와 비슷한 Munsell색상을 사용하는 것을 제안하였으며, 11개의 색상으로 특징벡터를 구성한 히스토그램에서 조명의 변화 또는 잡음의 영향을 최소화하는데 노력하였다[68]. Cox는 이와 비슷하게 HSV색상으로 변환한 영상에서 12개의 색상으로 특징벡터를 구성하여 영상검색에 적용하였다.[69][70]. 이러한 Munsell 및 HS계열의 공간은 사람의 시각인식체계와 유사하기 때문에 영상처리분야에서 많이 사용되고 있다. Munsell의 색상표현방법은 색도, 명도,

채도의 세 가지 색상속성을 그림 2.12와 그림 2.13과 같이 삼차원의 좌표계에 표시한다.

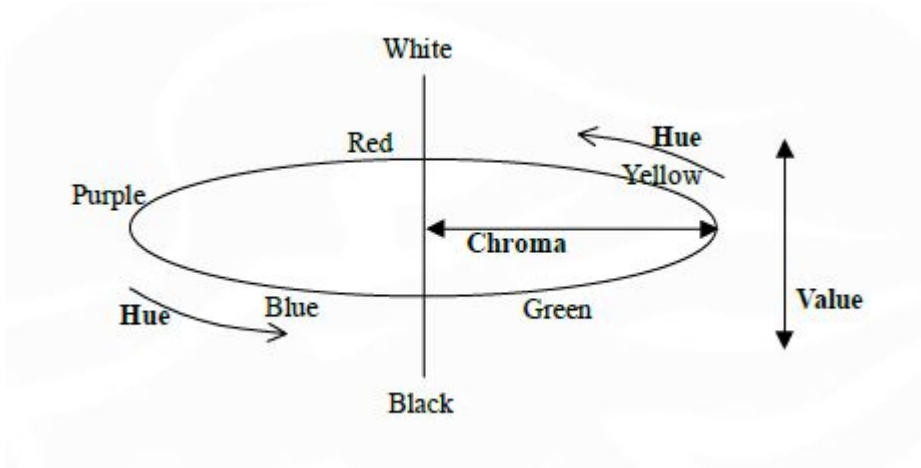


그림 2.12. Munsell 시스템의 색상표현

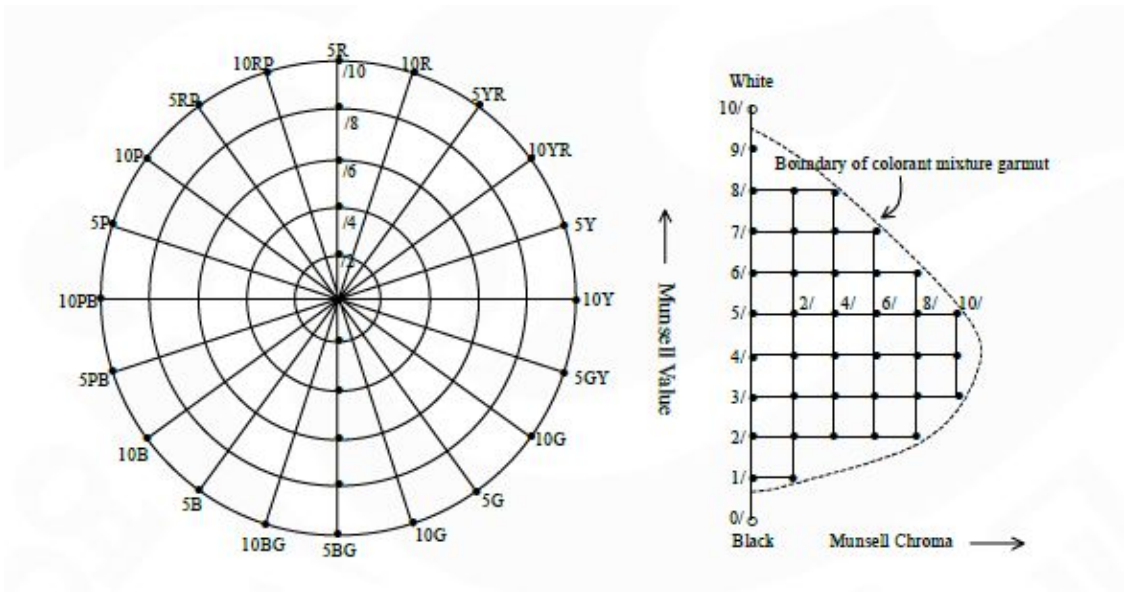


그림 2.13. Munsell 색상 시스템

그리고 HS계열의 색상공간을 그림 2.14에 나타내었다. 그림에서 보는 바와 같이 중앙의 선은 밝기의 정도를 나타내고, 채도(saturation)는 중앙 축과의 거리, 그리고 색도는 중앙을 둘러싸고 있는 원 또는 육각형으로 표시한다.

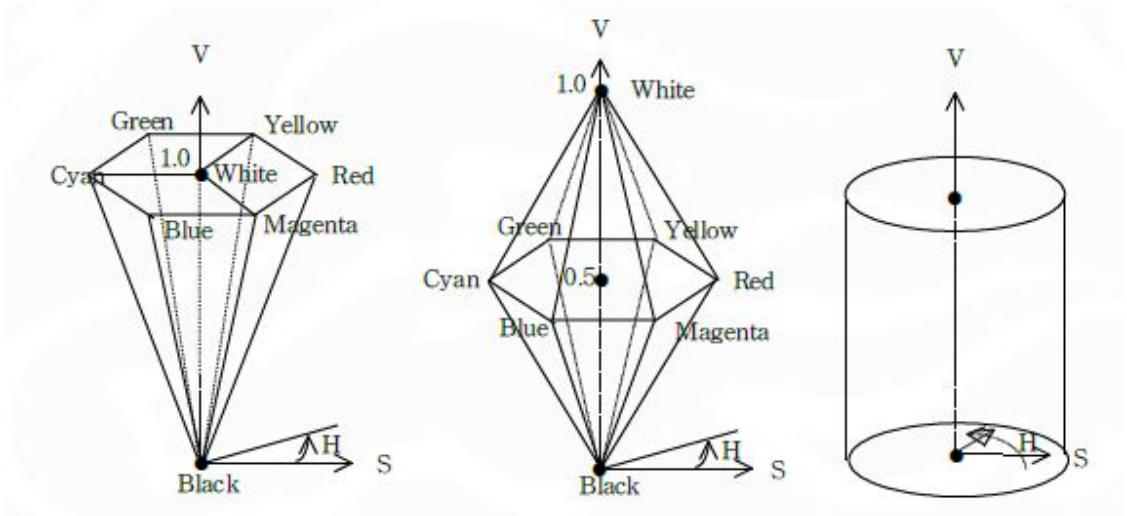


그림 2.14. HS계열의 색상공간

HS계열의 색상들은 RGB공간을 변형하여 사용상의 간편성을 도모함으로써 직관적인 면은 개선되었으나, 인식상 균등한 색상분포를 제공하지 않는 단점(비직선성)이 있다. 다시 말하면, RGB가 R, G, B의 비율에 의해 색상을 결정하는 수학적 모델이라면, HS 모델은 직관적인 모델이다.

앞에서 살펴 본 바와 같이 색상은 영상에서 보이고 있는 가장 대표적인 속성으로서 영상검색에서 특징벡터로 가장 많이 사용하고 있으며, 현재까지 이용되고 있는 영상검색시스템에서도 다른 특징요소들 보다 우선적으로 사용되고 있다.

2. 질감을 이용한 특징 추출

질감은 특정 패턴이 유사적 규칙성을 보이는 지역적인 영역의 특징을 말한다. 질감 특징의 유사성에 기반하여 검색하는 방법은 다양한 상황에 적용될 수는 없다. 하지만

유사한 색상을 가진 영상들의 영역을 구분해내는데 유용하게 사용될 수 있다. 다양한 방법들이 질감 정보의 유사성 계산을 위해서 제안되었다. 그 중 많은 방법들이 질의 영상과 데이터베이스의 영상들로부터 계산된 2차 통계(second-order statistics) 자료를 비교하여 검색을 수행한다. 이러한 방법으로 1978년 Tamura 등은 대비(contrast), 조약함(coarseness), 방향성(directionality), 규칙성(regularity), 선형성(line-likeness) 및 거친 정도(roughness) 등을 이용하여 질감 정보를 표현하였다[71]. 1996년 Liu와 Picard는 주기성(periodicity)과 방향성(directionality) 그리고 임의성(randomness) 등을 이용하여 질감 정보를 계산하였다[72]. 다른 방법으로는 Manjunath와 Ma에 의한 Gabor 필터를 이용하는 방법[73], Kaplan 등의 프랙탈을 이용하는 방법[74] 등이 있다. 질감 정보를 이용한 영상 검색 시 질의는 원하는 텍스처를 선택하거나 샘플 영상을 제시하는 것으로 검색이 가능하다. 개발된 시스템으로는 1998년 Ma와 Manjunath에 발표된 "texture thesaurus"가 있다.

Tamura는 질감 영상의 분석과 분류를 위해 질감 특징들을 6가지로 정의하고 실험과 비교를 통해 그 특징의 효과를 입증하였다. 정의된 특징 요소들은 IBM의 QBIC 등의 다양한 시스템에서 유용하게 사용되고 있다. 제안된 질감 특징들은 다음과 같다[71].

조약함은 가장 기본적인 질감 정보를 나타내며 반복되는 규칙적인 패턴의 미세한 정도를 나타낸다. 영상 $f(x, y)$ 에 대해 조약함을 계산하는 과정은 다음과 같다.

i) step 1 : 영상내의 모든 픽셀들에 대해 2의 지수의 크기 즉, $1 \times 1, 2 \times 2, \dots, 32 \times 32$ 의 범위 내의 픽셀들에 대한 평균을 구한다. (x, y) 위치에서 크기가 $2^k \times 2^k$ 인 범위에 대한 평균은 식 (2.8)과 같다.

$$A_{k(x, y)} = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} \frac{f(i, j)}{2^{2k}} \quad (2.8)$$

ii) step 2 : 각 위치에 대해 수직, 수평 방향으로 평균값들의 차를 계산한다. 예를 들면, 수평 방향으로의 차는 식 (2.9)와 같이 계산할 수 있다.

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (2.9)$$

iii) step 3 : 각 위치에 대해 다음 식 (2.10)의 조건을 만족시키는 best size를 결정한다.

$$S_{best}(x, y) = 2^k \quad (2.10)$$

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

여기서, k 는 수직 또는 수평 방향 중에서 E 를 최대로 만드는 값이다.

iv) step 4 : 조약함 F_{crs} 를 얻기 위해 S_{best} 를 평균한다.

$$F_{crs} = \frac{1}{m \times n} \sum_i^m \sum_j^n S_{best}(i, j) \quad (2.11)$$

여기서, m, n 은 조약함 계산 시 영상의 유효 너비 및 높이이다.

contrast는 영상 내 패턴에서의 화소들의 대비 정도를 나타내며 다음과 같은 방법으로 얻어질 수 있다.

먼저 영상내의 흰색, 검은색의 분포 정도를 고려하기 위해 식 (2.12)와 같이 kurtosis α_4 를 계산한다.

$$\alpha_4 = \frac{\mu_4}{\sigma^2} \quad (2.12)$$

여기서 μ_4 는 평균에 대한 4차 모멘트이고 σ^2 은 분산이다. 여기에 영상내의 화소 값들의 동적인 범위를 고려하여 대비는 식 (2.13)과 같이 정의된다.

$$F_{con} = \frac{\sigma}{(\alpha_4)^n} \quad (2.13)$$

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

$$\begin{aligned}
 |\Delta G| &= \frac{(\Delta_H + |\Delta_V|)}{2} \\
 \theta &= \tan^{-1}\left(\frac{\Delta_V}{\Delta_H}\right) + \frac{\pi}{2}
 \end{aligned} \quad (2.14)$$

이 얻어진다. $|\Delta G|$ 가 일정 문턱값 이상인 픽셀에 대해 θ 를 양자화하여 식 (2.15)와 같은 히스토그램을 만든다.

$$H_D(k) = \frac{N_\theta(k)}{\sum_{i=0}^{n-1} N_\theta(i)}, \quad k = 0, 1, 2, \dots, n-1 \quad (2.15)$$

여기서 $N_\theta(k)$ 는 $|\Delta G| \geq t$, $(2k-1)\pi/2n \leq \theta < (2k+1)/2n$ 조건을 만족하는 포인트의 수이다. 위의 $H_D(k)$ 의 피크들의 sharpiness를 고려하여 식 (2.16)와 같이 방향성을 계산한다.

$$F_{dir} = 1 - r \cdot n_p \sum_p^{n_p} \sum_{\phi \in w_p} (\phi - \phi_p)^2 \cdot H_D(\phi) \quad (2.16)$$

여기서, n_p 는 피크의 수, ϕ_p 는 H_D 에서 p_{th} 피크의 위치이다. 또한 w_p 는 p_{th} 피크의 범위를 나타내며 r 은 ϕ 의 양자화 레벨에 관한 정규화 계수이고 ϕ 는 양자화된 방향값이다.

선형성은 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 i 를 가지고 또 다른 에지 픽셀이 j 방향 코드를 가지며 서로 거리가 d 만큼 떨어져 있는 경우의 빈도 수를 나타낸다. 이러한 $n \times n$ 크기의 co-occurrence 행렬을 이용하여 선형성을 구하는 식은 식 (2.17)과 같다.

$$F_{lin} = \frac{\sum_i^n \sum_j^n P_{Dd}(i, j) \cos\left(\frac{(i-j)2\pi}{n}\right)}{\sum_i^n \sum_j^n P_{Dd}(i, j)} \quad (2.17)$$

규칙성은 영상 내의 질감 특성들의 규칙성 정도를 나타낸다. Tamura는 영상 전체에 걸쳐서 질감 특징들이 변화한다면 불규칙적이라 가정하고 따라서 영상을 분할하여 각 분할된 부분에 대한 질감 특성을 추출하여 특징들의 변화도를 고려하였다. 규칙성을 정의하는 식은 식 (2.18)과 같다.

$$F_{reg} = 1 - r(\sigma_{ers} + \sigma_{con} + \sigma_{dir} + \sigma_{lin}) \quad (2.18)$$

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

거친 정도는 시각에 대한 실험에 바탕을 두어 질감 특징 중에서 가장 중요한 요소인 조약함 및 대비의 영향을 고려하기 위해 식 (2.19)와 같이 거친 정도를 정의하였다.

$$F_{rgh} = F_{crs} + F_{con} \quad (2.19)$$

위와 같은 질감 특징들은 현재 IBM의 QBIC 등의 시스템에서 사용되고 있으며 영상의 질감 특징들의 일부 또는 6가지 모두를 추출하여 영상들간에 서로 비교함으로써 질감 정보 기반 영상 검색이 가능하다.

3. 형태를 이용한 특징 추출

물체의 형태 정보는 시각적으로 영상의 매우 다양한 정보를 가지고 있다. 형태 정보 기반 검색 방법은 구현하기 어렵지만 형태 정보가 물체를 가장 잘 표현하는 방법이기 때문에 내용 기반 검색 시스템에서 가장 중요한 요소이다. 형태 표현 방법은 그림 2.15와 같이 크게 2가지 부류로 나눌 수 있다[75]. 하나는 윤곽선 기반 방법이고 다른 하나는 영역 기반 방법이다. 윤곽선 기반 방법은 영상 내의 물체의 외부 윤곽선에 기반하여 특징을 추출하는 방법으로 체인 코드(chaincode), 푸리에 묘사(Fourier descriptor), UNL 푸리에 특징, 물체의 윤곽선의 반경, 특징 점들을 이용한 방법 등의 다양한 방법들이 존재한다. 영역 기반 방법은 물체 영역의 기하학적인 특징을 추출하는 방법으로 면적(area), 신장율(elongation), 오일러 수(Euler number), 불변 모멘트(invariant moments), 저나이크 다항식을 이용한 저나이크 모멘트(Zernike moments) 등을 비롯하여 많은 방법들이 있다.

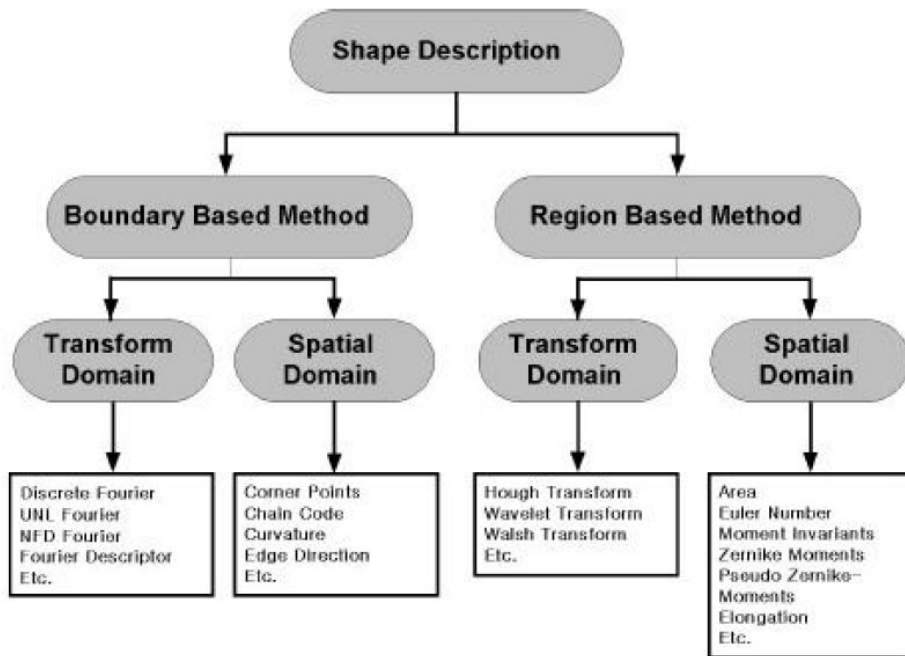


그림 2.15. 형태 기반 검색 기법의 분류[75]

에지의 방향 히스토그램을 이용한 방법은 Jain에 의해 제안된 방법으로 영상 내의 물체의 에지의 방향 히스토그램을 형태 정보로 이용한다. 즉, 에지의 방향 히스토그램을 구한 후에 히스토그램 인터섹션 방법을 이용하여 각 영상의 히스토그램을 비교한다 [76].

먼저 Canny 에지 연산자[77]를 이용하여 에지의 위치를 검출한 후 에지 위치에서 그레디언트의 방향을 구해서 에지의 방향 정보를 얻어낸다. 즉, 픽셀 (m, n) 에서 가로 수평 방향으로의 변화량을 $g_h(m, n)$ 이라 하고, 수직 방향으로의 변화량을 $g_v(m, n)$ 라고 하면 방향은 식 (2.20)을 이용하여 계산할 수 있다.

$$\theta = \tan^{-1} \frac{g_v(m, n)}{g_h(m, n)} \quad (2.20)$$

이렇게 에지 위치에서 얻어진 방향들을 이용하여 히스토그램을 생성한 후 히스토그램 인터섹션을 이용하여 유사도를 비교한다. 하지만 에지의 방향을 히스토그램화하는 것은 영상의 이동에는 불변하지만 축척이나 회전에는 변화한다는 단점을 가지고 있다. 따라서 Jain은 다음과 같은 방법을 이용하여 개선하였다. 축척에 대한 문제는 영상의 총 에지 개수를 이용하여 히스토그램을 정규화함하여 해결하였다. 하지만 회전에 대한 변화는 축척에 대한 것보다 다소 어려운데 물체가 회전하면 에지의 방향 히스토그램에서는 이동이 일어나게 된다. 즉, 에지 방향의 히스토그램의 그룹이 45° 의 각도로 나뉘어져 있다면, 30° 나 40° 의 각도를 가지는 에지는 같은 그룹에 포함되게 된다. 그러나 10° 의 각도를 회전할 경우 각기 서로 다른 그룹에 포함되게 됨으로써 원래의 히스토그램과는 다른 모양의 히스토그램이 만들어지게 된다. 이런 문제를 해결하기 위해 히스토그램 평활화(smoothing) 연산을 수행하였다. 하지만 히스토그램을 평활화하여 매끄럽게 만들어도 급격한 회전에는 적용되지 못한다는 단점을 가지고 있기 때문에 이동 축척의 변화에는 좋은 결과가 나왔지만 회전의 변화에는 다소 좋지 못한 결과를 보였다.

저나이크 모멘트는 단위원의 내부에서 완전하게 직교집합을 형성하는 복소 저나이크 다항식으로부터 유도된 것으로 저나이크 모멘트를 이용한 영상의 인식에 대한 연구가 Khontanzad에 의해 제안되었다[78]. 저나이크 다항식의 집합은 식 (2.21)의 $V_{nm}(x, y)$ 에 의해서 나타내어진다.

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \quad (2.21)$$

여기서, n 은 0과 양의 정수를 나타내며, m 은 $|m| \leq n, n - |m| = \text{even}$ 조건을 만족하는 정수이다. 그리고 ρ 와 θ 는 중심까지의 거리와 시계 반대 방향의 각도를 의미한다. 또한 $R_{nm}(\rho)$ 는 방사형 다항식이다.

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2}-s\right)! \left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s} \quad (2.22)$$

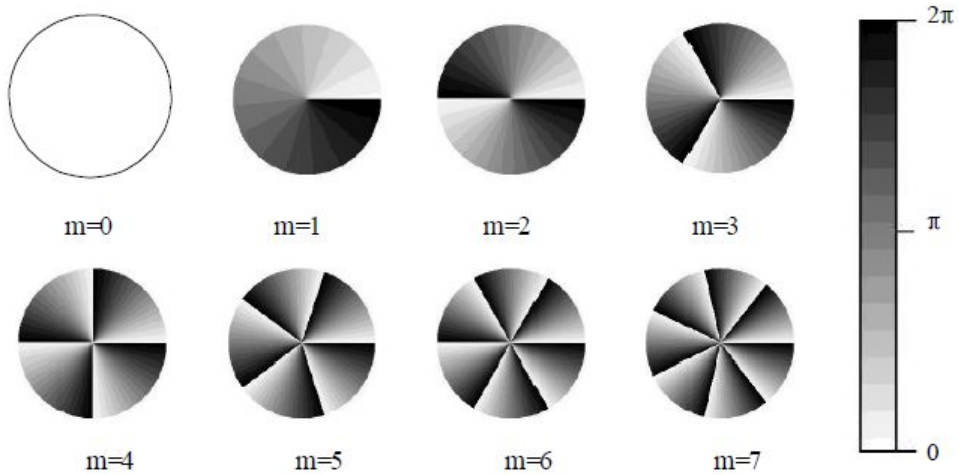


그림 2.16. 저나이크 다항식의 위상[78]

그림 13은 반복계수 m 에 따른 저나이크 다항식의 위상을 나타내며 이로부터 저나이크 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 영상 $f(x, y)$ 에 대한 n 차, 반복계수 m 의 저나이크 모멘트는 저나이크 다항식을 이용하여 식 (2.23)과 같이 정규 모멘트와 유사하게 정의된다.

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(x, y), \quad x^2 + y^2 \leq 1 \quad (2.23)$$

차수에 따라 생성되는 저나이크 모멘트의 목록은 표 2.1과 같다. 각 영상에 대해서

추출된 모멘트들을 비교하는 것으로 영상간의 형태의 유사성을 비교할 수 있다.

표 2.1. 저나이크 모멘트의 목록

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	$A_{100} A_{102} A_{104} A_{106} A_{108} A_{1010}$
4	$A_{40} A_{42} A_{44}$	11	$A_{111} A_{113} A_{115} A_{117} A_{119} A_{1111}$
5	$A_{51} A_{53} A_{55}$	12	$A_{120} A_{122} A_{124} A_{126} A_{128} A_{1210} A_{1212}$
6	$A_{60} A_{62} A_{64} A_{66}$	13	$A_{131} A_{133} A_{135} A_{137} A_{139} A_{1311} A_{1313}$

저나이크 모멘트의 경우 회전에 의한 모멘트의 크기가 불변이며 영상의 재구성이 쉽다는 특징이 있다. 저나이크 모멘트를 적용하기 위해선 반드시 영상에 대해서 크기와 이동에 대한 정규화 과정이 필요하다. 즉, 영상 내의 물체가 영역 $x^2 + y^2 \leq 1$ 내부에 위치하도록 조정되어야 한다는 것이다. 저나이크 모멘트는 완전한 직교집합으로, 이를 이용하여 원 영상을 복원할 수 있다. 또한 영상의 기하학적 변형(이동, 회전, 축척)과 조명변화에 매우 강인하며 영상 전체의 형태 정보를 효과적으로 표현할 수 있다는 장점도 있다. 그러나 영상의 열화에 매우 민감하고 영상의 세부 화소 값은 무시되는 단점이 있다.

지역적인 미분 불변치를 이용하는 방법은 영상에서 많은 정보를 포함하고 있는 부분들이나 변화가 큰 부분들을 따로 추출하고 그 부분들에 대한 지역적인 특징들을 검색에 이용한다. 이를 위해 영상에 있어서 특징점으로서 관심지점(interest point)을 정의하고 그 특징점에 대해 미분 불변치를 구하는 것으로 지역적인 특징을 추출하고 각 영

상간의 유사한 지역적인 특징을 가지는 특징점들을 비교하는 것으로 유사성을 측정할 수 있다.

관심지점은 영상에서 신호가 2차원적으로 변하는 위치를 나타내며, 코너 포인트, T 교차점, 텍스처가 급격히 변하는 위치 등이 해당된다. 관심지점은 회전, 이동, 축척, 영상 열화, 조명 변화, 관점(viewpoint) 변화 등에 안정적이며, 높은 정보량을 가지기 때문에 영상 검색에 효과적으로 이용된다[79-81]. 관심지점의 성능은 검출기의 성능에 따라 좌우되며, 이는 다양한 영상변화에도 반복적인 검출결과를 나타내는 반복성과 분별성을 나타내는 정보가 비교의 기준이 된다. 현재 검출기는 Moravec, Harris, Heitge, Forstner, Horaud, Cottie 등에 의해 고안되었다. 이중 반복성과 정보량 측면에서 Harris 검출기가 가장 우수하다고 판명이 되었다[80]. Harris 검출기는 자기 상관함수와 관련된 행렬을 구하고, 자기 상관함수의 중요 곡률(curvature)이 되는 이 행렬의 고유값(eigenvalue)을 척도로 인터레스트 포인트를 검출한다. 그리고 이 검출기는 Moravec 검출기의 문제점을 개선시킨 검출기로 지역 윈도우 영상을 다양한 방향으로 이동하였을 때 영상 세기의 변화량($E_{x,y}$)을 기준으로 영상의 영역을 분류한다.

$$E_{x,y} = \sum_u \sum_v W_{u,v} (I_{u+x,v+y} - I_{u,v})^2 \quad (2.24)$$

여기서, W 는 단위 직각 윈도우를 나타내며, x, y 는 이동(shift) 값을 나타내며 u, v 는 윈도우 영역을 의미한다.

윈도우 내의 영상 세기가 대략적으로 비슷한 경우 영상 세기의 변화량은 작다. 윈도우가 에지 성분에 걸쳐 있다면 에지 성분을 따라 이동할 경우 영상 세기의 변화는 적지만, 수직으로 이동시에는 변화량이 크게 된다. 윈도우 영역이 코너 포인트 등의 인터레스트 포인트에 해당하는 영역은 영상세기의 변화량은 항상 크게 된다. Moravec의 검출기는 변위가 $\{(1, 0), (1, 1), (0, 1), (-1, 1)\}$ 일 때의 변화량의 최소 값이 문턱 값 이상이고, 지역의 최대 값일 때 인터레스트 포인트로 판별한다.

Harris 검출기는 Moravec 검출기의 세 가지 문제점을 개선하였다. Moravec 검출기

는 단위 직각 원도구를 사용해서 잡음에 민감하였으나 Harris는 원형인 가우시안 원도구를 사용하여 잡음문제를 해결하였고, 원도구의 이동이 45° 마다 이루어지므로 변화량 결과가 이방성(anisotropic)이라는 문제를 해결하기 위해 모든 방향으로의 이동을 고려해 해석적인 팽창(analytic expansion)을 적용하였고, 검출 성능 향상을 위해 자기상관함수 행렬 M 의 고유값을 이용하였다.

$$E_{x,y} = \sum_u \sum_v W_{u,v} (I_{u+x,v+y} - I_{u,v})^2 = \sum_u \sum_v W_{y,v} (xX + yY + O(x^2, y^2))^2 \quad (2.25)$$

원도구 영역의 작은 이동에서 영상 세기의 변화량은 식 (2.26)과 같다.

$$\begin{aligned}
 E(x, y) &= Ax^2 + 2Cxy + By^2 = (x, y)M(x, y)^T \\
 M &= \begin{bmatrix} A & C \\ C & B \end{bmatrix} \\
 X &= I \otimes (-1, 0, 1) = \frac{dI}{dx} \\
 Y &= I \otimes (-1, 0, 1)^T = \frac{dI}{dy} \\
 A &= X^2 \otimes W \\
 B &= Y^2 \otimes W \\
 C &= (XY) \otimes W
 \end{aligned} \quad (2.26)$$

M 은 회전에 불변하며, x, y 의 변화 시 원점에서 분포상태의 척도가 된다. M 의 고유값 α, β 는 자기상관함수의 중요 곡률에 비례하므로 이를 기준으로 영상에서의 영역을 분류한다. 두개의 고유값이 모두 작으면 원도구 영역은 거의 동일한 픽셀 값을 갖고, 고유값의 하나는 크고 하나는 작으면 원도구 영역은 에지 성분에 해당되고, 고유값이 모두 크면 원도구 영역은 인터레스트 포인트에 해당된다.

자기상관함수 매트릭스의 고유값은 다음 식들과 같은 성질을 가진다.

$$Tr(M) = \alpha + \beta = A + B \quad (2.27)$$

$$Det(M) = \alpha\beta = AB - C^2 \quad (2.28)$$

$$R = Det - kTr^2 \quad (2.29)$$

응답함수 R 은 A, B, C 의 함수이며 k 값은 임의의 작은 양수이다. R 값을 척도로 R 값이 양수면 인터리스트 포인트, 음수면 에지 영역, 작으면 유사한 픽셀 값을 갖는 영역으로 판별한다[80].

그림 2.17는 영상에 대해 Harris 코너점 검출기를 통해 특징점을 추출한 결과를 보여준다.

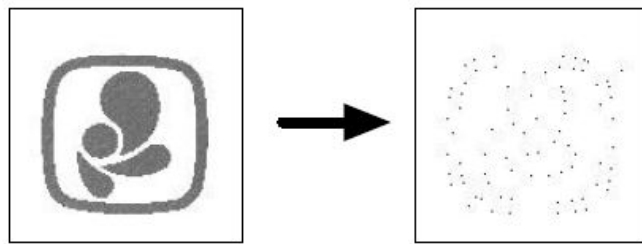


그림 2.17. Harris 코너점 검출 결과

영상으로부터 추출된 특징점에서 회전, 이동, 크기 변형에도 강인하게 영상 검색에 적용할 수 있는 미분 불변치 벡터는 다음과 같이 계산할 수 있다. 미분 불변치 계산은 원영상에 작은 크기의 잡음이 첨가된 경우 원영상과 잡음이 첨가된 영상의 미분치를 비교하면 매우 다르다. 심지어 작은 잡음이 첨가된 경우라도 고주파 잡음을 포함하는 경우 미분치는 상당히 달라지는데, 고차항의 미분일수록 값은 더욱 상이하게 된다. 원 신호를 $f(x)$ 라하고 잡음이 첨가된 신호를 $h(x)$ 라 할 때,

$$h(x) = f(x) + \epsilon \sin(wx) \quad (2.30)$$

$$h(x)' = f(x)' + w\epsilon \cos(wx) \quad (2.31)$$

$$h(x)'' = f(x)'' - w^2\epsilon \sin(wx) \quad (2.32)$$

고주파 잡음인 경우 w 값이 매우 크므로 $h(x)'$ 와 $f(x)'$ 는 매우 상이하며, 고차항 미분일수록 더욱 상이하게 된다. 그러므로 원신호를 미분하기 전에 신호를 평활화(smoothing)를 할 필요가 있다. 평활화 함수로 가우시안 함수 G 를 이용하여 원 신호로부터 잡음을 제거한 다음 미분을 수행한다. 또한 미분 연산과 콘볼루션은 교환법칙이 성립하므로 식 (2.33)과 같이 나타낼 수 있다.

$$\partial_i(G * F) = G * \partial_i F = \partial_i G * F \quad (2.33)$$

미분 계산을 안정시키는 가장 간단한 방법은 평활화 함수인 가우시안 함수를 미분한 것과 영상을 콘볼루션시키는 것이다.

가우시안 함수 $G(x, \sigma)$, 가우시안 함수의 미분치 $G_{i_1, i_2, \dots, i_D}(x, \sigma)$ 일 때

$$G(x, \sigma) = \frac{1}{(\sqrt{2\pi\sigma^2})^D} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (2.34)$$

$$G_{i_1, i_2, \dots, i_D}(x, \sigma) = \frac{\partial^D}{\partial i_1 \partial i_2 \dots \partial i_D} G(x, \sigma) \quad (2.35)$$

여기서, $x = (x_1, x_2, \dots, x_D)$ (영상에서는 $D=2$) 이며, 가우시안 함수내의 σ 는 평활화의 정도를 결정한다. N 차까지 미분값을 계산하여 벡터로 표시할 수 있다. 이는 Koenderink에 의해 처음으로 로컬 젯(Local Jet)이라는 벡터로 표현되었다[82].

영상의 특징점들을 중심으로 하는 지역 윈도우 영상에 대한 특징들을 미분값들의 집

합으로 나타낼 수 있다. 미분값 구할 때 가우시안 미분값과 지역 윈도우 영상의 콘볼루션을 이용해 안정적으로 구할 수 있고 이러한 미분값의 집합을 로컬 젯이라 한다 [82]. 지역 윈도우 영상을 I 라고 하고, 주어진 크기가 σ 일 때, 임의의 한점 $x = (x_1, x_2)$ 에서 N 차 로컬젯은 식 (2.36)와 같다.

$$J^N[I](x, \sigma) = L_{i_1, i_2, \dots, i_n}(x, \sigma) | (x, \sigma) \in I^* R ; n = 0, 1, 2, \dots, N \quad (2.36)$$

σ 는 가우시안 함수의 크기를 나타내며 영상의 크기가 변하는 경우에 영상 검색에 적용할 때 중요하게 사용된다.

미분 불변치는 로컬 젯으로부터 계산되며, 이러한 특징 벡터로 영상 신호를 표현할 수 있다. 특징 벡터는 여러 가지 영상변형에서도 강인하게 특징 값을 유지한다. 2차 미분항을 고려한 미분 불변치 벡터 DI 를 Einstein 합방식과 직교좌표계에서 표현하면 식 (2.37)와 같다.

$$DI[0, \dots, 4] = \begin{bmatrix} L \\ L_i L_i \\ L_i L_{ij} L_i \\ L_{ii} \\ L_{ij} L_{ji} \end{bmatrix} = \begin{bmatrix} L \\ L_x L_x + L_y L_y \\ L_{xx} L_x L_x + 2L_{xy} L_x L_y + L_{yy} L_y L_y \\ L_{xx} + L_{yy} \\ L_{xx} L_{xx} + 2L_{xy} L_{yx} + L_{yy} L_{yy} \end{bmatrix} \quad (2.37)$$

L_i 는 위에서 표시한 로컬 젯의 성분이다. 즉, L 은 지역 윈도우 영상과 가우시안 함수를 콘볼루션한 결과 성분이다. 다른 크기의 영상들에 대해서도 적용하기 위해 가우시안 함수의 크기를 변형하는 것으로 미분 불변치 벡터를 구할 수 있다. 영상의 경우 벡터 $x = (x, y)$ 로 2차원이며, L_{xy} 에서 첨자 x, y 는 x, y 방향으로 각각의 편미분에 해당된다. Einstein 합방식에서 i 는 변수들에 대한 미분의 합을 나타내며, L_i 와 L_{ij} 는 식 (2.38), (2.39)과 같이 표현될 수 있다.

$$L_i = \sum_i L_i = L_x + L_y \quad (2.38)$$

$$L_{ij} = \sum_i \sum_j L_{ij} = L_{xx} + L_{xy} + L_{yx} + L_{yy} \quad (2.39)$$

3차항의 미분을 고려한 경우에 미분 불변치 벡터 DI 를 Einstein 합방식으로 표현하면 식 (2.40)과 같다.

$$DI[5, \dots, 8] = \begin{bmatrix} \varepsilon_{ij}(L_{jkl}L_iL_kL_l - L_{jkk}L_iL_lL_l) \\ (L_{iij}L_jL_kL_k - L_{ijk}L_iL_jL_k) \\ -\varepsilon_{ij}L_{jkl}L_iL_kL_l \\ L_{ijk}L_iL_jL_k \end{bmatrix} \quad (2.40)$$

단, $\varepsilon_{12} = -\varepsilon_{21}$ 이고, $\varepsilon_{11} = -\varepsilon_{22} = 0$ 이다.

미분 불변치 벡터 DI 를 직교 좌표계에서 표현한 식을 간략히 하면 식 (2.41)와 같다.

$$\begin{aligned}
 DI[5] &= L_{xxx}L_yL_yL_y - L_{yyy}L_xL_xL_x + 3L_{xyy}L_xL_xL_y - 3L_{xxy}L_xL_yL_y \\
 DI[6] &= L_{xxx}L_xL_yL_y + L_{yyy}L_xL_xL_y + L_{xxy}L_y(L_yL_y - 2L_xL_x) + L_{xyy}L_x(L_xL_x - 2L_yL_y) \\
 DI[7] &= L_{xxx}L_xL_xL_y - L_{yyy}L_xL_yL_x + L_{xxy}L_x(2L_yL_y - L_xL_x) + L_{xyy}L_y(L_yL_y - 2L_xL_x) \\
 DI[8] &= L_{xxx}L_xL_xL_x + L_{yyy}L_yL_yL_y + 3L_xL_y(L_{xxy}L_x + L_{xyy}L_y)
 \end{aligned} \quad (2.41)$$

영상들에 대해 Harris의 코너점 추출기를 이용하여 특징점들을 추출한 후 각각의 특징점들에 대해 미분 불변치를 계산한다. 그리고 영상간의 유사도 비교는 각각의 특징점들간의 거리를 비교하는 것으로 수행된다. 비교되는 특징점들간의 미분 불변치 특징 벡터의 거리가 미리 정한 문턱 값을 넘지 못할 경우 두 특징점들을 동일한 포인트로

간주하며, 두 영상간의 일치하는 특징점들의 수를 이용하여 두 영상간의 유사도를 계산한다. 유사도 계산 방법은 다음과 같이 나타낼 수 있다. N 개의 특징점들을 가지고 있는 두 영상 I_1, I_2 에 대해 I_1 내의 특징점의 집합을 $\{p_i\}, i=1, 2, \dots, N$, I_2 내의 특징점의 집합을 $\{p_j\}, j=1, 2, \dots, N$ 라 할 때 식 (2.42)과 같이 계산할 수 있다[83].

$$\begin{aligned}
 & \text{for } i = 1, 2, \dots, N \\
 & \quad \text{if } (d_{\min}(I_1(p_i), I_2(p_j)), \text{for } j = 1, 2, \dots, N) < \textit{Threshold}) \\
 & \quad \quad \textit{count} = \textit{count} + 1
 \end{aligned} \tag{2.42}$$

유사도는 \textit{count}/N 으로 계산되며 유사도가 클수록 두 영상간의 형태 정보가 유사함을 나타낸다.

III. 적응적 웨이트드 CBIR 알고리즘 제안

본 논문에서는 영상 특성 기반 적응적 웨이트드 CBIR 알고리즘을 제안한다. 제안된 알고리즘의 흐름은 그림 3.1에서 볼 수 있다. 영상의 내용기반 특성 중 레이블링, 색상, 코너점, 조밀도, 면적 등을 이용하여 특징 벡터를 구성하고 입력 영상을 객체가 있는 영상과 배경이 주된 영상을 분류한다.

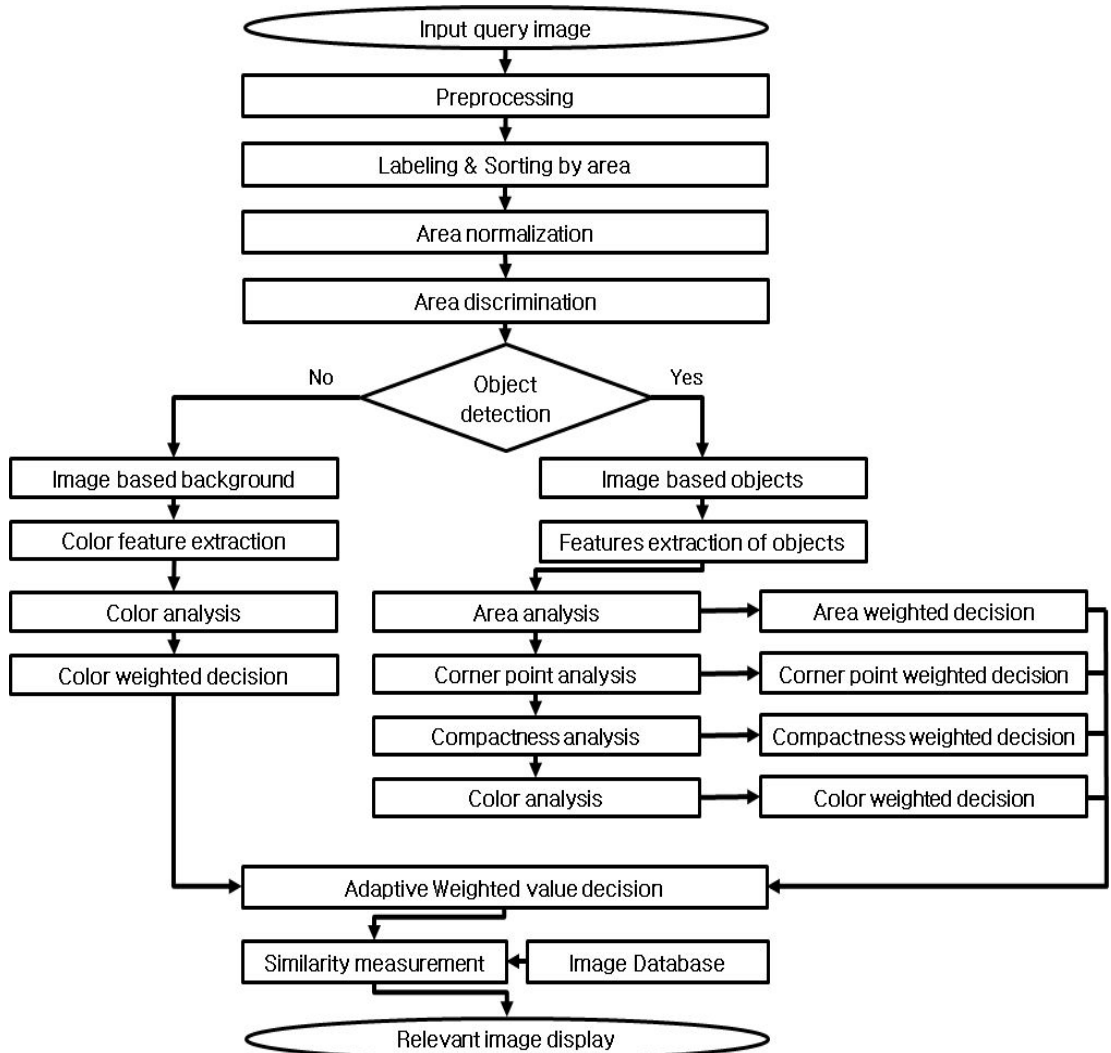


그림 3.1. 제안하는 적응적 웨이트드 CBIR 알고리즘 흐름도

분류된 입력 영상 중에서 배경이 주된 영상은 색상의 특징만 이용하도록 웨이트드를 부여하며, 객체가 있는 영상은 객체별로 색상, 면적, 코너점, 조밀도등의 특징 벡터의 값에 따라서 적응적으로 웨이트드를 부여하여 유사 영상을 검색한다.

알고리즘은 첫 번째로 영상에서 불필요한 정보들을 제거하기 위해 Otsu 방법과 침식, 팽창연산을 이용하여 불필요한 배경과 객체 영상 정보를 다듬는 전처리 과정과 두 번째, 영상에 레이블링하여 크기순으로 재배열하고 면적을 정규화하여 크기 필터링을 하는 면적 판단 과정을 거친 후 배경이 주된 영상과 객체가 판별된 영상으로 분류하는 객체 판별 과정을 지난다. 그 후 배경이 주된 영상이면 색상 정보만을 이용하도록 웨이트드를 부여하고 객체가 판별된 영상이면 각각의 레이블링마다 면적, 조밀도, 코너점의 개수, 농담값, 색상값의 특성 정보를 추출하여 특징 벡터를 구성하였다. 그리고 세 번째로 추출한 특징 벡터를 기반으로 객체의 면적과 코너점의 개수와 조밀도를 분석하여 각각의 특징 벡터 값에 맞춰서 적응적으로 웨이트드를 부여하며, 색상 화소값은 인덱스 데이터 행렬의 칼라맵의 분포도를 이용하여 적응적으로 웨이트드를 부여한다. 마지막으로 질의 영상에 따라서 적응적으로 부여된 웨이트드에 따라서 유사도 측정을 하고 영상검색을 진행하였다.

A. 전처리 과정

영상들을 특징에 따라 분류하기 위해서는 불필요한 정보들을 제거하는 과정이 필요하다. 본 논문에서는 각각의 영상마다 최적의 레이블링을 가지기 위한 전처리 과정으로 Otsu 방법을 사용한다.

Otsu 방법은 다음의 두 가지 가정으로 시작한다. 첫 번째는 일반적인 이미지는 히스토그램 영역에서 두 가지의 산 모양으로 나타난다는 것이고, 두 번째는 두 산은 각각 정규 분포의 모양을 따른다는 것이다 [84]. 이런 가정을 바탕으로 정규 분포의 분산과 평균값을 이용하여 두 산을 구분하는 적절한 경계 값을 찾는 것이 Otsu 방법이다. 그림 3.2은 그림 3.3의 (a)를 Otsu 방법으로 이진화한 히스토그램을 나타내며 (b)는 이진화 결과 이미지를 나타낸다.

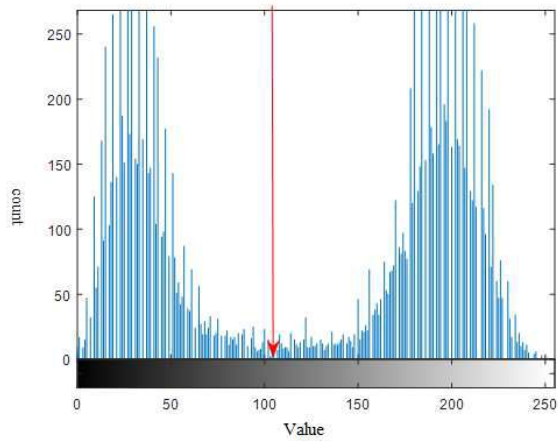


그림 3.2. Otsu 방법을 이용하여 이분화한 히스토그램



(a) 샘플이미지



(b) 결과 이미지

그림 3.3. Otsu 방법을 이용하여 이분화한 샘플이미지

Otsu 방법은 정규 분포의 분산과 평균을 이용하여 적절한 임계값을 자동으로 찾는 방법이다. 일반적으로 이미지에서 물체와 배경은 각각 유사한 밝기와 분포를 가지고 있기 때문에 히스토그램에서 두 개의 산이 나타난다. 분산은 픽셀들이 얼마나 퍼져있는 지에 대한 정보이기 때문에 분산 값이 작을수록 특정 픽셀에 몰리게 된다. 이런 원리를 이용하여 각 영역 사이의 분산이 최대가 되는 지점이나 각 영역 내의 분산이 최소가 되는 지점을 나눔으로써 가장 좋은 경계 값을 얻을 수 있다. 본 논문에서는 각

영역 내의 분산이 최소가 될 때를 이용하였으며 이를 식으로 나타내면 다음과 같다.

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \quad (3.1)$$

여기서 $q_1(t)$ 와 $u_1(t)$ 와 $\sigma_1^2(t)$ 는 다음과 같이 나타낼 수 있다.

$$q_1(t) = \sum_{i=1}^t p(i) \quad (3.2)$$

$$u_1(t) = \sum_{i=1}^t \frac{ip(i)}{q_1(t)} \quad (3.3)$$

$$\sigma_1^2(t) = \sum_{i=1}^t [i - u_1(t)]^2 \frac{p(i)}{q_1(t)} \quad (3.4)$$

를 의미한다. 여기서 $p(i)$ 는 i 에서의 히스토그램의 확률을 나타내며, $q_1(t)$ 는 t 보다 작거나 같은 그룹의 확률의 합을 나타낸다. $u(t)$ 는 그룹의 평균을 의미하며, $\sigma^2(t)$ 는 그룹의 분산을 나타낸다.

식 (3.1)은 각 영역 내의 분산 합을 의미하고 최종적으로는 $\sigma_w^2(t)$ 가 최소가 되는 t 를 찾아 해당하는 t 값을 임계값으로 설정한다.

적절한 임계값으로 이진화를 한 영상의 미세한 부분을 제거하기 위해 수축과 팽창 연산을 거쳐 그림 3.4과 같이 전처리를 한다.

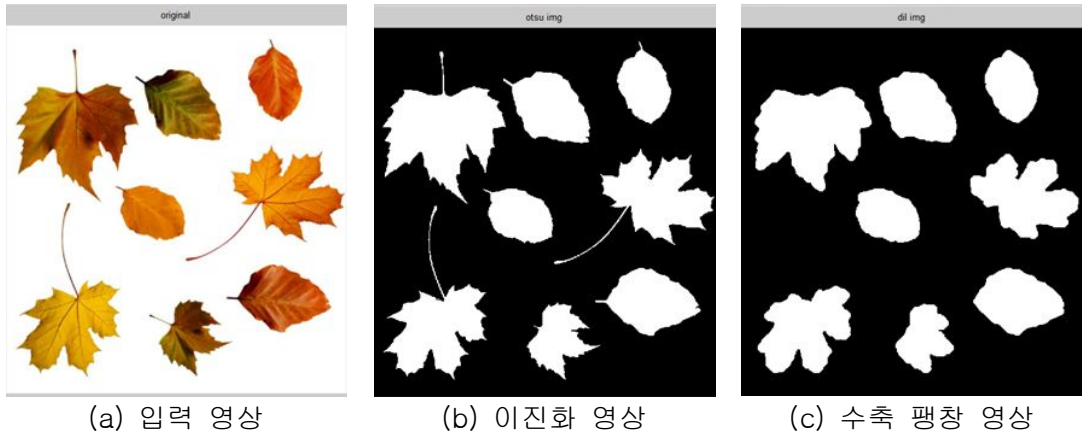


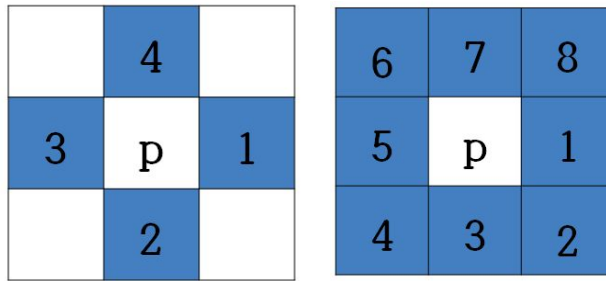
그림 3.4. 전처리 과정 이미지

B. 레이블링 및 객체 판별과정

1. 레이블링 및 크기 필터링

제안한 알고리즘에서는 레이블링마다 색상, 면적, 조밀도, 코너점 등의 특성을 각각 추출해내기 때문에 정확한 레이블링을 하는 과정이 필요하다. 이에 화소 연결성을 이용하여 레이블링의 번호를 부여하고 면적의 크기순으로 재배열하는 과정을 추가하였다.

화소 연결성은 에지 기반과 영역 기반의 영역 분할 모두에서 핵심적인 개념이다. 직사각형의 표본 패턴을 가진 일반 영상에서 한 화소를 둘러싸고 있는 이웃 화소들의 정의는 4-방향과 8-방향 이웃화소의 두 가지 방법으로 한다. 그림 3.5에서 보는 바와 같이 4-방향 이웃화소는 중심화소의 위, 아래, 좌, 우의 화소들만을 포함하고 8-방향 이웃 화소는 4-방향 이웃화소에서 모든 화소와 대각선 방향의 이웃 화소들을 추가하여 포함한다.



(a) 4-방향 이웃화소 (b) 8-방향 이웃화소

그림 3.5. 화소연결성의 이웃화소

어떤 화소 p_1 에서 다른 화소 p_n 으로의 4-방향 연결 경로는 $\{p_1, p_2, \dots, p_n\}$ 화소들로 이어지는데, 여기서 p_{i+1} 은 모든 $i=1, 2, \dots, n-1$ 에 대한 p_i 의 4-방향 이웃 화소이다. 이러한 경로가 만일 p_i 의 8-방향 이웃이면 8-방향으로 연결된 것이다. 만일 어떤 화소 집합이 있고 그 화소 집합의 화소쌍들이 적어도 4-방향 연결 경로를 갖는다면, 이 집합은 4-방향으로 연결된 영역이다. 8-방향으로 연결된 영역도 같은 방법으로 정의한다. 4-방향 연결성과 8-방향 연결성의 구분은 중요하다. 그림 3.6의 음영 처리된 부분은 두개의 8-방향 연결 영역 또는 네개의 4-방향 연결 영역으로 이루어져 있다.

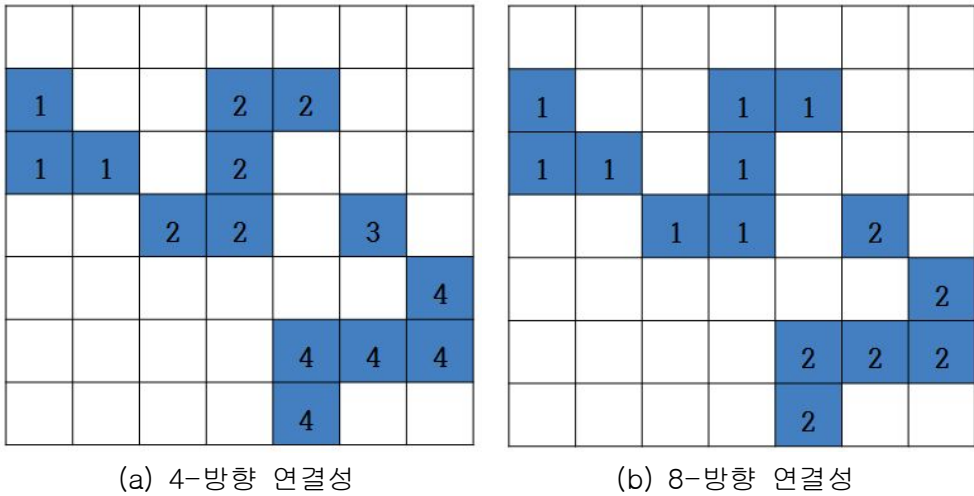


그림 3.6. 화소 연결성을 이용하여 연결된 화소들

영상에서 각각의 물체들을 구분하기 위해서는 화소들의 연결 유무를 알아야 한다. 한 픽셀에 이웃한 상, 하, 좌, 우의 4방향에 대한 연결성을 평가하기 위하여 그림 18의 (a)와 같은 형태의 연결성을 사용한다. 4-연결성 배열을 영상의 전체 영역에 대하여 대각 방향으로 적용하여 주변 픽셀과 연결된 픽셀은 연결성 배열의 가운데 픽셀에 레이블 번호를 지정한다.

영상을 읽는 과정에서 그림 3.5의 (a)와 같이 한 픽셀을 p 라 하고 1, 2, 3, 4를 각각 우측, 하단, 좌측, 상단의 이웃 픽셀이라고 할 때, p 가 배경이 아닌 경우에 대해서 4-연결성 배열을 적용해 검사한다.

먼저, 왼쪽 상단에서 오른쪽 하단 방향으로 적용하는 경우를 고려해 보면, 4와 3이 모두 배경이면 새로운 레이블을 p 에 지정한다. 4와 3 중 하나의 이웃 픽셀만이 개체이면 1인 픽셀의 레이블을 p 에 지정한다.

4와 3이 둘 다 개체이고, 같은 레이블이면 그 레이블을 p 에 지정하고, 레이블이 다르면 둘 중 작은 레이블을 p 에 지정한다.

위의 과정을 전체 영상에 대하여 수행하면 배경이 아닌 모든 개체에 레이블이 지정되지만, 동일한 개체에 둘 이상의 레이블이 지정되는 경우가 생긴다. 그러므로 다음

단계로 오른쪽 하단으로부터 왼쪽 상단 방향으로 배열을 한번 더 적용함으로써 하나의 개체에 둘 이상의 레이블이 지정되는 것을 최소화해 준다. 이 과정은 다음과 같다.

- i) p가 배경이 아닌 경우에 대하여 2와 1이 모두 배경이면 현재 지정된 레이블을 그대로 유지시킨다.
- ii) 2가 배경이고 1이 개체인 경우, p의 레이블이 1의 레이블보다 크면 1의 레이블을 p에 지정한다.
- iii) 2가 개체이고 1이 배경인 경우, p의 레이블이 2의 레이블보다 크면 2의 레이블을 p에 지정한다.
- iv) 2와 1이 모두 개체인 경우, 2의 레이블과 1의 레이블 중에서 더 작은 레이블을 p에 지정한다.

위의 과정을 거쳐 그림 3.7와 같이 화소들의 연결 성분을 찾아내어 구분 짓는 것을 레이블링이라 한다.

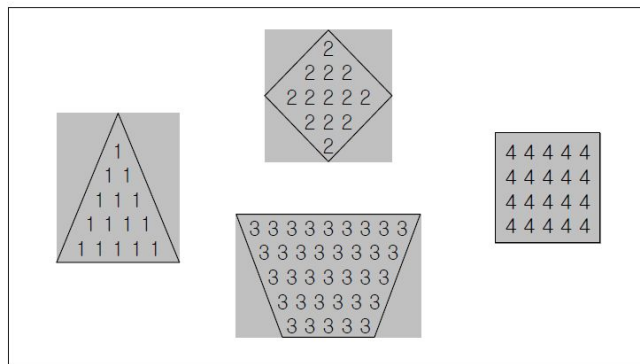


그림 3.7. 레이블링된 영상 정보

레이블링 알고리즘에는 회귀와 순차 알고리즘이 있다. 순차 알고리즘이 회귀 알고리즘에 비해 메모리가 적게 들고 수행시간이 짧아 많이 사용된다. 이 외에 스택(stack)

을 이용하여 레이블링을 구현하는 경우가 있다. 이는 순차 연결 알고리즘보다 수행시간이 짧고, 쉽게 구현이 가능하다는 장점이 있지만 영상의 크기가 정해지지 않았을 경우에는 오히려 수행 시간이 오래 걸린다는 단점이 있다. 이에 본 논문에서는 특정한 크기의 영상을 사용하지 않고 다양한 크기의 영상들을 사용하기 때문에 순차 연결 알고리즘을 사용하여 레이블링을 구현하였다. 다음 그림 3.8은 제안한 알고리즘에서 레이블링되어 각각 분리된 영상들이다.

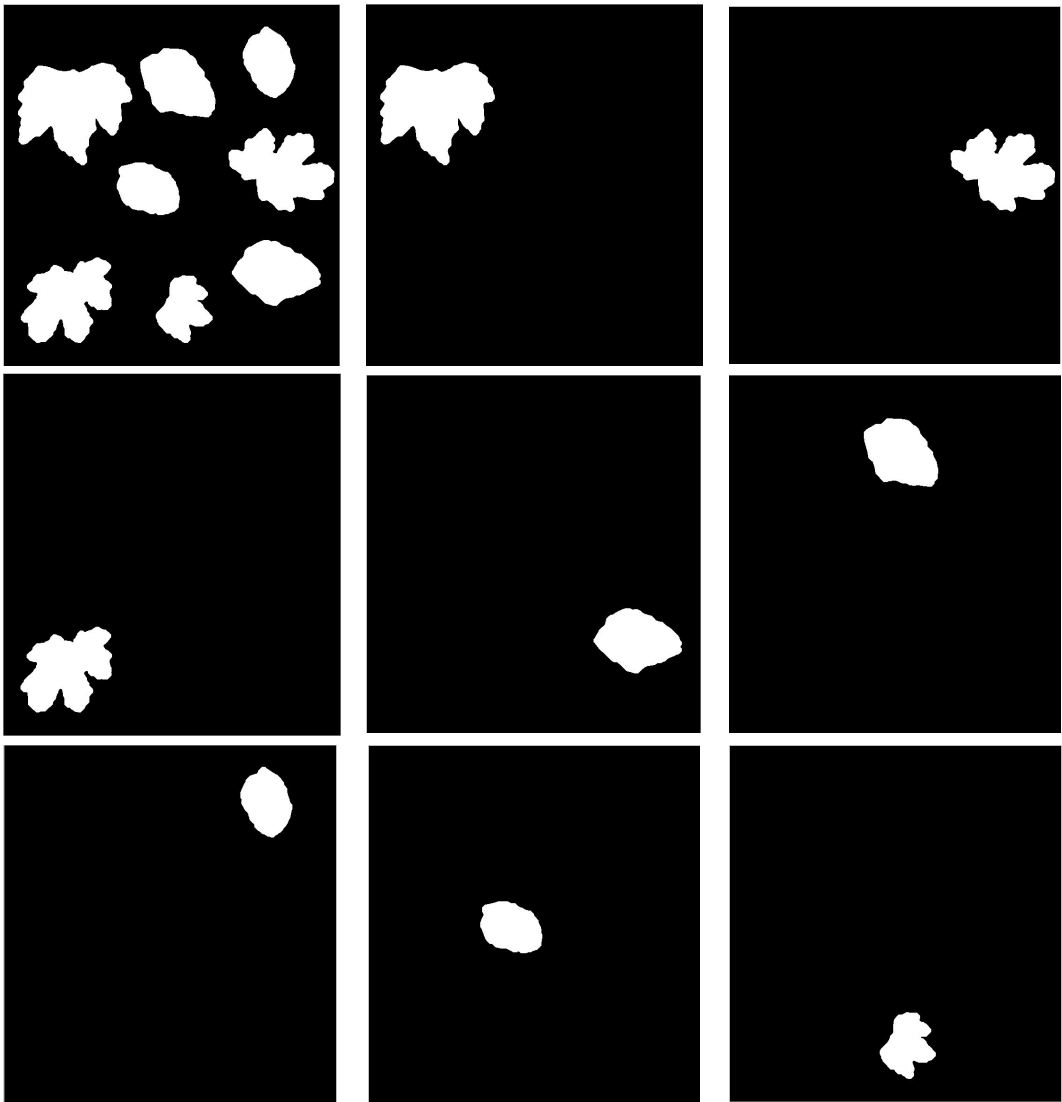


그림 3.8. 레이블링 후 분리된 영상들

분리한 레이블링들에는 영상 정보의 많은 부분을 가지고 있는 레이블링도 있지만 영상의 작은 부분의 정보만을 가지고 있는 레이블링도 있다. 또한 영상 정보에 원치 않는 잡음이 끼어있는 경우도 있을 수 있다. 이런 경우 그림 3.9와 같이 특정한 크기 이하의 성분들은 무시할 수 있는 데 이를 크기 필터링이라 한다. 또한 특정 크기 이상의 물체가 영상 정보에 포함되어 있을 경우 알고리즘 수행 시간을 크게 증가시킬 수 있다. 이러한 경우를 방지하기 위해 크기 필터링을 하는 경우도 있다. 본 논문에서는 레이블링들을 면적의 크기순으로 재배열하여 크기 필터링을 하였다.

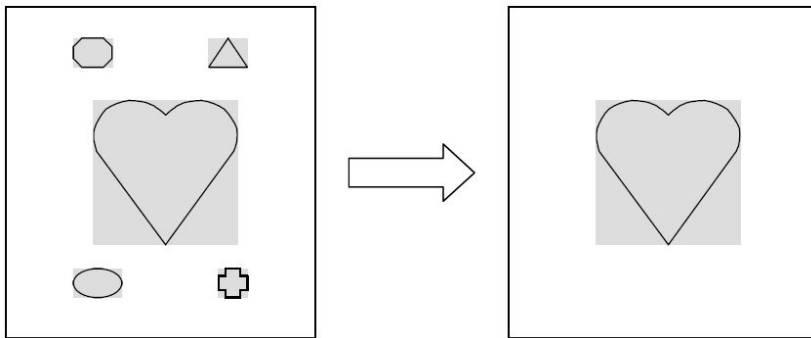


그림 3.9. 크기 필터링

2. 객체 판별

객체는 레이블링을 거친 영상의 면적을 분석함으로써 판단할 수가 있는데 영상에서 객체의 특징을 가장 쉽게 판단할 수 있는 특성 중의 하나는 면적이다. 디지털 영상처리 분야에서 현재 사용되는 물체의 면적 측정 방법으로는 경계선 내부에 존재하는 화소를 판별하고 판별된 화소의 수를 계산하여 면적을 측정하는 방법으로 컴퓨터 단층 영상의 면적 및 체적 측정에 활용되고 있다. 이 방법은 화소가 물체 영역에 속하는지를 판단해야 하며 노이즈에 약한 단점이 있고 한 평면에 같은 질감을 가진 2 개 이상의 물체가 존재하는 경우 화소의 밝기 값이 유사해 구별이 곤란하여 인위적인 분할 과정을 거쳐야 측정이 가능하다. 또 다른 방법으로 화면을 분할하여 면적을 계산하는 방법이 있다. 이 방법은 화면을 점차적으로 세분화시켜 각 세분화된 영역의 면적을 측정하여 전체 영역에 대해 누적함

로써 계산 속도를 향상시킬 수 있으나 이 방법의 경우에도 한 평면에 같은 질감을 가진 2 개 이상의 물체가 존재하는 경우 구별이 곤란하다. 본 논문에서는 레이블링을 통하여 면적을 계산하였다. 즉, 레이블링으로 각각 부여된 번호를 이용하여 화소의 수를 계산하여 면적을 나타내었다. 그림 3.10의 (a)는 레이블링된 이미지의 일부분에서 부여된 번호를 보여주며, (b)는 각각의 레이블링의 면적을 계산하여 나타낸 면적과 레이블링 번호를 크기순으로 정렬한 특징을 보여준다.

	41	42	43	44	45	46	47	48	49
161	0	0	0	0	0	1	1	1	1
162	0	0	0	0	1	1	1	1	1
163	0	0	0	1	1	1	1	1	1
164	0	0	0	1	1	1	1	1	1
165	0	0	0	1	1	1	1	1	1
166	0	0	0	1	1	1	1	1	1
167	0	0	0	1	1	1	1	1	1
168	0	0	0	1	1	1	1	1	1
169	0	0	0	1	1	1	1	1	1
170	0	0	0	1	1	1	1	1	1
171	0	0	0	1	1	1	1	1	1
172	0	0	0	0	1	1	1	1	1
173	0	0	0	0	1	1	1	1	1
174	0	0	0	0	1	1	1	1	1
175	0	0	0	0	1	1	1	1	1
176	0	0	0	0	1	1	1	1	1
177	0	0	0	1	1	1	1	1	1
178	0	0	0	1	1	1	1	1	1
179	0	0	1	1	1	1	1	1	1
180	0	1	1	1	1	1	1	1	1
181	0	1	1	1	1	1	1	1	1
182	1	1	1	1	1	1	1	1	1
183	1	1	1	1	1	1	1	1	1

(a) 레이블링 이미지의 일부분

	1	2	3	4	5	6	7	8
1	33702	23392	23120	17208	17189	11616	11538	10515
2	1	6	2	7	4	8	3	5
3	0.2273	0.1578	0.1559	0.1161	0.1159	0.0783	0.0778	0.0709

(b) 크기순으로 재배열된 면적 특징

그림 3.10. 면적 특징 추출

그림 3.10의 (b)에서 1행은 각 레이블링의 크기를 나타내며, 2행은 레이블링의 번호를 나타낸다. 면적이 클수록 영상의 많은 정보를 담고 있기 때문에 크기 순으로 재배열하였다. 그리고 3행은 정규화 면적을 나타낸 것으로 영상의 크기의 변화나 찌그러짐 등의 영상 손실에 강인함을 가지기 위해 면적을 정규화하여 특징에 추가하였다. 면적 정규화는 다음 식을 통해 진행되었다.

$$N_{L_i} = \frac{A_{L_i}}{\sum_{k=1}^n A_{L_k}}, \quad i = 1, 2, \dots, n \quad (3.5)$$

여기서, N_{L_i} 는 i 번째 레이블링의 정규화 면적을 의미하며, A_{L_i} 는 i 번째 레이블링의 면적을 나타낸다. n 은 이미지 내에서 마지막으로 구분된 레이블링 번호이다.

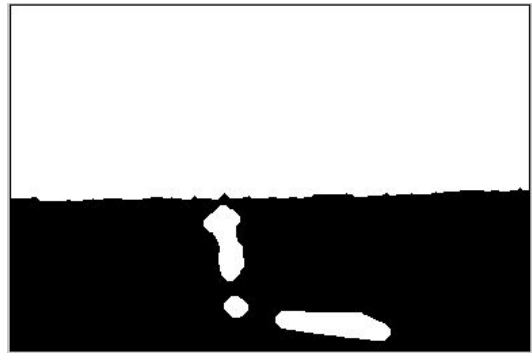
영상에서 객체 판별은 배경이 주된 영상인지 객체를 판별할 수 있는 영상인지를 구분하는 것으로 판단할 수 있다. 배경을 주로 가지고 있는 영상은 대개 색상이 단순하여 면적이 가장 큰 레이블링이 영상의 거의 대부분을 차지하는 영상들이다. 이는 레이블링을 할 때 영역의 화소값에 영향을 받기 때문이다. 이를 이용하여 레이블링의 정규화 면적이 다른 레이블링보다 일정 임계값 이상 차이가 나는 영상들을 분류해 낼 수가 있다. 다음의 식 (3.6)의 조건에 맞는 영상들이 분류가 되며 정규화면적의 차이가 Th 의 임계값보다 큰 영상들을 찾는다. Th 는 큰 그룹들의 경계가 되는 지점의 값을 가진다.

$$\left| NS_{L_n} - NS_{L_{n+1}} \right| > Th \quad (3.6)$$

분류된 영상은 색상에 가장 많은 웨이트드를 주어서 유사도 측정을 할 수가 있다. 그림 3.11은 배경이 주된 영상의 특성과 그 특징 벡터를 보여준다.



(a) 배경이 주된 영상



(b) 레이블링 영상

LV_inf <9x3 double>			
	1	2	3
1	53664	1319	966
2	1	4	2
3	0.9592	0.0236	0.0173

(c) 특징 벡터

그림 3.11. 배경이 주된 영상의 레이블링 특성과 특징 벡터

그림 3.11의 (c)에 표시된 정규화 면적을 보면 가장 큰 레이블링이 배경 정보를 담고 있기 때문에 영상에서 판별된 다른 객체들보다 매우 큰 면적을 가지고 있는 것을 확인할 수 있다. 이와 같은 면적 특징을 가지고 있는 영상들을 배경이 주된 영상들로 구분하였고, 이 외의 영상들은 객체들이 있는 영상으로 구분하였다. 배경이 주된 영상들은 객체가 없거나 객체의 정보가 미약하기 때문에 배경의 색상 특징만을 이용하며, 객체들이 있는 영상들은 레이블링된 객체들의 특징을 추출하고 분석하는 과정으로 이어진다. 그림 3.12는 객체가 판별된 영상들과 배경이 주된 영상으로 구분되어진 샘플 이미지들이다.



(a) 배경이 주된 샘플이미지



(b) 객체가 판별된 샘플이미지

그림 3.12. 객체 판별에서 구분된 영상들

위의 그림에서와 같이 배경이 주된 영상들은 객체가 없는 영상들이 대부분으로 산이나 바다와 같이 확실하게 구분되어지는 영역이 레이블링되어지며, 객체가 판별된 이미지들은 배경이 있더라도 객체의 모양이 확실하게 구분되어지면 영상의 복잡함에 상관없이 객체가 우선적으로 레이블링되어지는 것으로 객체를 판별할 수 있다.

C. 객체 특징 추출

1. 조밀도 특징 추출

조밀도는 영상 내에서 객체의 영역이 얼마나 조밀하게 모여있는지를 판별한다. 영상의 특징 벡터의 빠른 추출을 위해 객체마다 틀린 모양을 정확하게 판별하기 보다는 객체의 뭉쳐짐 정도를 판별하였다. 조밀도는 단위 둘레당 면적량으로 나타낼 수가 있다. 이를 위해 레이블링된 객체마다의 둘레길이를 구하는 것이 선행이 되어야한다. 이에 이진화된 레이블링 객체에 서브샘플링(sub-sampling)을 하고 경계를 얻어 그 둘레의 화소 수를 계산하는 것으로 측정하였다. 그림 3.13은 가장 큰 면적을 가지는 레이블링 객체의 서브샘플링과 경계선을 나타낸 영상이다. 그리고 그림 3.14는 단위 둘레당 면적을 나타낸 조밀도를 추가한 특징벡터를 나타낸다.

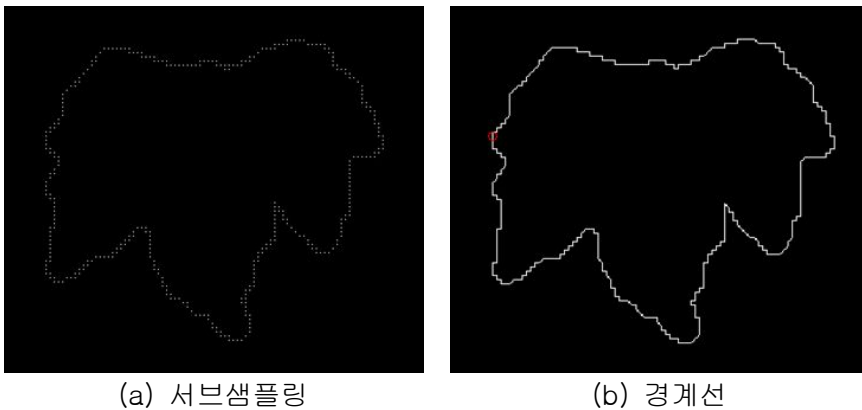


그림 3.13. 둘레 계산을 위한 경계선 추출

LV_info <4x8 double>								
	1	2	3	4	5	6	7	8
1	33702	23392	23120	17208	17189	11616	11538	10515
2	1	6	2	7	4	8	3	5
3	0.2273	0.1578	0.1559	0.1161	0.1159	0.0783	0.0778	0.0709
4	41.6074	31.8259	32.3810	39.9258	37.9448	32.9065	31.2683	25.8990

그림 3.14. 조밀도를 추가한 특징 벡터

2. 코너점 특징 추출

영상 처리에서 코너 정보는 영상의 윤곽선들과 함께 중요한 정보들 중의 하나이다. 디지털 컴퓨터 및 프로세서의 발달로 인해 데이터의 처리 속도 및 연산 속도가 향상되었다. 하지만 영상데이터들은 엄청난 양의 데이터를 가지고 있어 그 많은 양의 데이터들을 모두 고려 대상으로 간주하는 것은 불가능하다. 또한 컴퓨터는 물체를 검출하고 인식하는 것을 프로그램에 의해 순차적으로 진행하지만, 인간은 형태학적 분석과 물체의 위치변화를 동시에 처리할 수 있어 물체인식 능력이 월등히 앞선다.

물체의 위치를 검출하는 컴퓨터 비전 분야에 있어서 많은 양의 영상 데이터들 중 물체의 특징점들을 간략히 정리할 수 있고, 또한 중요한 특징점들을 추출하는 방법은 영상 데이터의 처리속도 향상과 더불어 연산량을 축소시킬 수 있는 중요한 요소가 된다.

이러한 특징점들 중 영상의 객체 윤곽선 전체를 검출하는 윤곽선 추출 방법을 이용하는 것보다 영상에 있어 명암분포가 급격히 변할 수 있는 특징점들을 추출하는 코너 추출방법이 많은 양의 영상 데이터의 특징점들을 축소하면서 패턴의 특징을 대표할 수 있어 입력 영상의 처리를 보다 빠르고 쉽게 할 수 있는 변수가 된다. 본 논문에서는 형태적인 정보로 코너점을 추출하는 방법으로 Harris 코너 연산자를 이용하였다.

Harris 코너 연산자는 1988년 Harris, C. 와 Stephens, M에 의하여 제안되었다 [85]. 이 연산자는 Moravec 코너연산자와 윤곽선 추출 연산자의 결합형태로 영상 신호의 국부적 자기상관법(Local Auto-correlation function)에 기초하여 영상의 천이에 따른 명암의 공간적 변화를 분석하여 코너를 검출하도록 하였다. Harris 코너 연산자의 적용방법은 다음과 같다.

- i) 명암의 공간적 변화를 계산한다.

$$\begin{aligned}
 E(u, v) &= \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \\
 &= [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}
 \end{aligned}
 \tag{3.6}$$

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_{xy} \\ I_{xy} & I_y^2 \end{bmatrix} \quad (3.7)$$

여기서 $I(x, y)$ 는 영상의 밝기값을 나타내며, $E(u, v)$ 는 영상의 명암값의 평균을 나타낸다. 그리고 $w(x, y)$ 는 영상에 수직성분과 수평성분에 강한 Prewitt 원도우이다.

- ii) 구한 행렬 M 의 고유치를 λ_1 과 λ_2 라고 한다면, 결정 변수 R 의 값 변화로 코너를 결정한다.

$$R = \det(M) - k(\text{Trace}(M))^2 \quad (3.8)$$

여기서 $\det(M)$ 은 고유치 λ_1 과 λ_2 의 곱을 의미하며, $\text{Trace}(M)$ 은 두 고유치 λ_1 과 λ_2 의 합을 의미한다. 또한 k 는 상수를 의미하며 그 변화폭은 [0.4 ~ 0.6] 정도이다.

- iii) 구한 결정 변수 R 에 문턱값을 적용하여 국부 최대값으로 코너를 검출한다. 고유치 λ_1 과 λ_2 사이의 변화와 코너의 관계는 다음 3가지 경우로 나타낼 수 있다.

- (1) 만약 고유치 λ_1 과 λ_2 이 모두 작은 값을 가진다면, 이는 결정 변수 R 이 거의 상수값을 가지는 것을 의미하고 영상에서 명암의 변화가 작은 평평한 상태를 의미한다.
- (2) 만약 고유치 λ_1 과 λ_2 중 어느 하나의 고유치가 다른 고유치 보다 상당히 크다면, 이는 영상에서의 윤곽선에 가깝다는 것을 의미한다.
- (3) 그림 4에 나타낸 것처럼 만약 고유치 λ_1 과 λ_2 모두가 상당히 크다면 이는 결정 변수 R 이 모든 방향으로 증가되는 것을 의미하며 코너에 가깝다는 것을 의미한다.

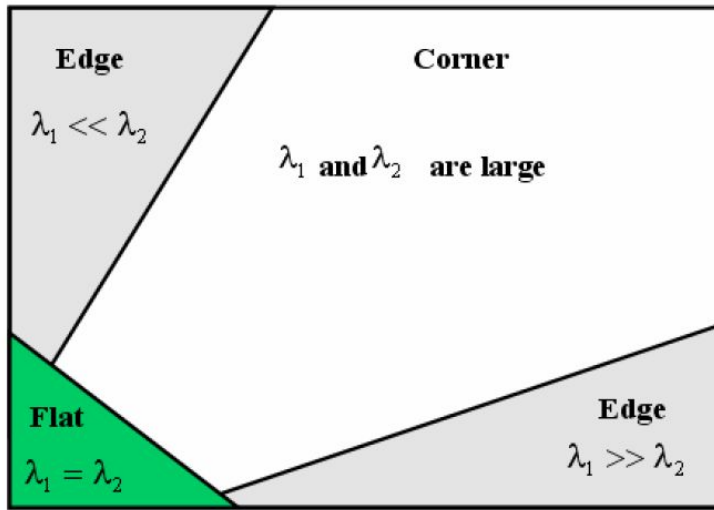


그림 3.15. 고유치의 분포에 따른 특징점 분할

그림 3.16은 Harris 코너 연산자를 이용하여 입력영상 중 일부 레이블링의 코너 부분을 검출한 결과이다.



(a) 레이블링 이미지



(b) 코너 검출

그림 3.16. 코너 부분 검출

유사도 측정에서 비교 데이터의 양이 많은 위치 정보는 이용하지 않으며 코너점의 개수로 비교하여 데이터의 양을 축소하였다. 다음 그림 3.17은 코너점의 개수를 추가한 특징 벡터를 나타낸 것으로 5행에 값이 입력되었다.

LV_info <5x8 double>								
	1	2	3	4	5	6	7	8
1	33702	23392	23120	17208	17189	11616	11538	10515
2	1	6	2	7	4	8	3	5
3	0.2273	0.1578	0.1559	0.1161	0.1159	0.0783	0.0778	0.0709
4	41.6074	31.8259	32.3810	39.9258	37.9448	32.9065	31.2683	25.8990
5	9	10	10	17	24	16	17	8

그림 3.17. 코너점의 개수를 추가한 특징 벡터

3. 색상 및 농담 특징 추출

본 논문에서는 기본적인 색상을 판단해 내기 위하여 RGB 칼라모델을 사용하여 uint8의 인덱스 영상(Indexed Image)을 만들고 각 레이블링의 평균 색상을 계산하여 특징 벡터로 이용하였다.

영상 처리에서는 다음 표 3.1과 같이 영상의 네가지 기본 타입을 정의한다.

표 3.1. 영상의 타입

타입	설명
Binary	영상의 화소 값을 흑(0)과 백(1)으로만 나타냄
Indexed	영상에서 사용되는 모든 칼라를 colormap을 사용하여 나타냄(logical, uint8, uint16, single, double)
Grayscale	영상에서 화소 값을 그레이 음영으로 나타냄 uint8 - values range [0, 255] uint16 - values range [0, 65535]
Truecolor	영상에서 화소 값을 R, G, B 성분으로 나타냄

인덱스 영상은 인덱스 화소값과 인덱스 화소값에 대한 칼라맵의 조합으로 구성된 영상을 의미한다. 여기서 인덱스 화소값은 데이터의 유형에 따라 결정되며, 칼라맵은 R, G, B 순서로 나타낸다.

인덱스 이미지는 데이터 행렬 X 와 칼라맵 행렬 map 으로 구성된다. map 은 범위 $[0, 1]$ 의 부동 소수점 값을 포함하는 $double$ 형 클래스로 구성된 $m \times 3$ 행렬이다. map 의 각 행은 하나의 색을 나타내는 빨간색, 녹색, 파란색의 구성요소를 지정한다. 인덱스 이미지는 픽셀 값과 칼라맵 값 간에 '직접 매핑'을 사용한다. 각 이미지 픽셀의 색은 X 의 대응값을 map 에 대한 인덱스로 사용하여 결정된다. 따라서 X 의 값은 정수여야 한다.

값 1이 map 의 첫 번째 행을, 값 2가 두 번째 행을 가리키는 식으로 진행된다. 칼라맵은 대개 인덱스 이미지와 함께 저장되어 해당 이미지와 함께 자동으로 불러지는 방식이지만 디폴트 칼라맵만 사용하도록 제한되어 있지 않으며, 선택하는 어떤 칼라맵이든 사용할 수 있다. 다음 그림 3.18은 입력 영상의 인덱스 이미지를 보여준다. 여기서 픽셀값은 $uint8$ 을 이용하여 구성하였고, 인덱싱된 칼라맵의 일부분은 그림 3.19에서 볼 수 있다.



그림 3.18. 인덱스 이미지

X <768x714 uint8>											
	614	615	616	617	618	619	620	621	622	623	
103	6	6	6	5	5	7	7	4	7	4	
104	6	6	6	6	5	6	7	4	7	7	
105	6	6	6	6	6	5	7	7	7	7	
106	1	6	6	6	6	7	7	7	7	7	
107	1	6	6	6	6	4	7	7	7	7	
108	1	6	6	6	6	7	7	7	7	7	
109	1	1	6	6	6	4	7	7	7	7	
110	1	1	6	6	6	6	6	7	7	7	
111	1	1	6	6	6	6	4	7	7	7	
112	1	1	6	6	6	6	6	6	7	7	
113	2	1	1	6	3	6	6	4	7	7	
114	1	1	1	6	6	6	6	6	6	7	
115	2	1	1	6	6	6	6	6	7	7	
116	2	1	1	1	6	6	6	6	6	7	
117	2	2	1	1	6	6	6	6	6	4	
118	2	2	1	1	6	6	6	6	5	4	
119	2	2	1	1	6	6	6	6	5	4	
120	2	2	2	1	6	6	6	5	5	6	
121	2	2	2	1	1	6	6	6	6	6	
122	2	2	2	1	1	6	6	6	6	6	
123	2	2	2	2	1	1	6	6	6	6	
124	2	2	2	2	1	1	6	6	6	6	
125	2	2	2	2	2	1	1	6	6	3	

(a) 인덱싱된 칼라맵

map <8x3 double>			
	1	2	3
1	0.4902	0.2588	0.0314
2	0.9059	0.8118	0.7020
3	0.8196	0.4667	0.0471
4	0.9843	0.9529	0.8980
5	0.9882	0.9843	0.9569
6	0.9608	0.9922	0.9922
7	0.9882	0.9608	0.9843
8	0.9961	0.9961	0.9961

(b) uint8을 이용한 화소값

그림 3.19. uint8을 이용한 화소값과 칼라맵

입력 영상이 칼라 영상이 아니라 그레이 영상일 경우 농담 값을 비교하도록 하기 위해 농담 레벨에서도 똑같이 인덱스 화소값과 맵을 구성하였다. 다음 그림 3.20은 각 레이블링의 평균 R, G, B 화소값과 평균 농담값을 추가한 특징벡터를 보여준다. 화소값과 농담값은 전부 double 형으로 처리가 되었다.

LV_info <9x8 double>									
	1	2	3	4	5	6	7	8	레이블링의 개수
1	33702	23392	23120	17208	17189	11616	11538	10515	
2	1	6	2	7	4	8	3	5	
3	0.2273	0.1578	0.1559	0.1161	0.1159	0.0783	0.0778	0.0709	정규화 면적
4	41.6074	31.8259	32.3810	39.9258	37.9448	32.9065	31.2683	25.8990	조밀도
5	9	10	10	17	24	16	17	8	코너점의 개수
6	0.3780	0.5522	0.6123	0.3882	0.3600	0.4403	0.5194	0.3036	평균 농담값
7	0.3876	0.8778	0.8305	0.5577	0.1740	0.8564	0.8866	0.1056	R
8	0.3456	0.7846	0.7420	0.4895	0.1513	0.7637	0.7928	0.0917	평균 색상값 G
9	0.2971	0.6764	0.6393	0.4123	0.1260	0.6564	0.6838	0.0762	B

그림 3.20. 최종 특징 벡터

이는 영상마다 추출되는 최종 특징 벡터가 되며 데이터베이스의 모든 영상에 적용되어 유사도 측정의 기본 자료가 된다.

D. 적응적 웨이티드 부여

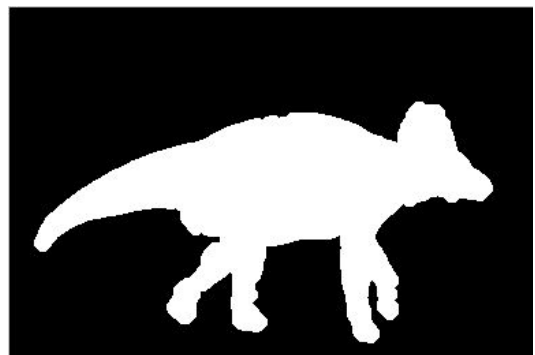
1. 면적 분석 및 웨이티드 부여

앞에서와 같이 레이블링된 영상들에서 특성을 계산한 특징벡터를 사용하여 질의 영상 객체들의 특징들을 분석하고 그 값들에 따라서 적응적으로 웨이티드를 부여해 주는 작업이 필요하다. 면적은 객체의 크기를 나타내기 때문에 객체의 정보량과 관련이 매우 깊기 때문에 먼저 정규화된 면적의 분석으로 시작된다.

정규화 면적의 총 누적량은 1의 값을 가지고 있는데, 단일 객체 영상이라면 영상의 모든 정보를 하나의 객체가 가지게 되어 그 객체의 정규화 면적도 1의 값을 가지게 된다. 그리고 다중 객체 영상은 객체들마다 영상의 정보를 나누어서 가지게 되므로 그 정규화 면적도 정보량에 따라 값이 정해진다. 이에 근거하여 0에서 1의 사이의 정규화 면적 값의 크기가 클수록 많은 정보량을 가지기 때문에 높은 웨이티드가 부여되도록 설정하였다. 그림 3.21은 단일 객체 영상과 다중 객체 영상의 특징을 보여준다.



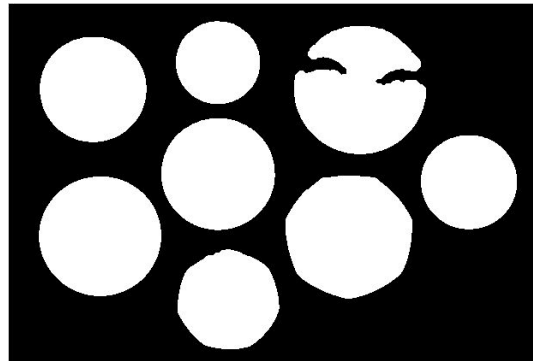
(a) 단일 객체 영상



(b) 레이블링 영상1



(c) 다중 객체 영상



(d) 레이블링 영상2

LV_inf <9x1 double>	
	1
1	24997
2	1
3	1
4	21.6612
5	6
6	0.3107
7	0.5610
8	0.5402
9	0.4728

(e) 단일 영상의 특징 벡터

LV_inf <9x8 double>								
	1	2	3	4	5	6	7	8
1	16754	16701	15701	13658	12094	10734	9749	7489
2	6	7	1	3	2	5	8	4
3	0.1628	0.1623	0.1526	0.1328	0.1176	0.1043	0.0948	0.0728
4	40.3711	27.7425	39.3509	36.7151	34.5543	32.4290	31.0478	27.2327
5	5	10	8	9	7	6	6	9
6	0.6814	0.7141	0.6297	0.6605	0.6829	0.6236	0.6038	0.6389
7	0.6354	0.6818	0.8539	0.6403	0.7075	0.6510	0.8166	0.6177
8	0.5589	0.5938	0.8465	0.5753	0.6316	0.5822	0.8094	0.5576
9	0.5053	0.5433	0.8446	0.5260	0.5907	0.5328	0.8082	0.5097

(f) 다중 객체 영상의 특징 벡터

그림 3.21. 단일 객체 영상과 다중 객체 영상의 비교

단일 객체 영상의 특징 벡터에서 표시되어 있는 부분에서 볼 수 있듯이 정규화 면적은 1의 값을 가지게 되며, 다중 객체 영상에서는 객체의 영역들이 부분으로 나누어져 있기 때문에 정규화 면적의 값들이 객체별로 골고루 분포되어 있다는 것을 확인할 수 있다. 그리고 객체 판별 과정에서 하나의 레이블링이 영상 정보의 절반 이상을 차지할 경우 배경 영상으로 분류하였기 때문에 정규화 면적은 0에서 0.5의 사이 값을 가지게 되며 예외적으로 단일 객체 영상은 1의 값을 가진다. 이에 따라 실험을 통해 다음의 표 3.2와 같이 면적의 웨이트드를 부여하였다.

표 3.2. 면적의 웨이트드 부여

면적	0~0.1	0.1~0.2	0.2~0.3	0.3~0.4	0.4~0.5	0.5~0.99	1
웨이트드	1	2	3	4	5	0	5

2. 코너점 분석 및 웨이트드 부여

코너점은 형태적인 정보의 복잡함을 판단하는 데에 중요한 정보이다. 이는 영상의 윤곽선 전체를 검출하는 윤곽선 추출 방법보다 영상에 있어 명암분포가 급격히 변할 수 있는 특징점의 추출인 코너 추출방법이 영상 데이터의 특징점들을 축소하면서 패턴의 특징을 대표할 수 있어 입력 영상의 처리를 보다 빠르고 쉽게 할 수 있는 변수가 된다. 이에 따라 코너점의 위치정보를 전부 이용하기보다는 코너점의 개수로 객체 형태의 복잡함 정도를 판단할 수 있도록 코너점의 개수를 이용하였다. 즉, 코너점의 개수가 많을수록 그 객체의 형태가 복잡하다는 것을 의미하며, 반대로 코너점의 개수가 적을수록 그 객체의 형태가 각이 적은 원의 형태에 가깝다는 것을 의미한다. 그림 3.22는 객체가 원에 가까운 영상과 복잡한 형태를 가지는 영상의 특징 벡터를 비교한 것이다.

LV_inf <9x8 double>								
	1	2	3	4	5	6	7	8
1	16754	16701	15701	13658	12094	10734	9749	7489
2	6	7	1	3	2	5	8	4
3	0.1628	0.1623	0.1526	0.1328	0.1176	0.1043	0.0948	0.0728
4	40.3711	27.7425	39.3509	36.7151	34.5543	32.4290	31.0478	27.2327
5	5	10	8	9	7	6	6	9
6	0.6814	0.7141	0.6297	0.6605	0.6829	0.6236	0.6038	0.6389
7	0.6354	0.6818	0.8539	0.6403	0.7075	0.6510	0.8166	0.6177
8	0.5589	0.5938	0.8465	0.5753	0.6316	0.5822	0.8094	0.5576
9	0.5053	0.5433	0.8446	0.5260	0.5907	0.5328	0.8082	0.5097

(a) 원에 가까운 객체를 가지는 영상의 특징 벡터

LV_inf <9x11 double>

	1	2	3	4	5	6	7	8	9	10	11
1	21826	20896	17238	16812	15733	14876	13358	11499	5097	3914	1473
2	14	2	13	3	5	1	12	4	10	6	7
3	0.1529	0.1464	0.1208	0.1178	0.1102	0.1042	0.0936	0.0806	0.0357	0.0274	0.0103
4	34.0499	23.7725	26.3578	22.8735	18.4227	17.7306	16.5322	23.6119	19.7558	18.9082	10.4468
5	21	17	15	14	15	13	27	19	20	19	13
6	0.3026	0.4171	0.5840	0.5703	0.5877	0.5019	0.6128	0.3549	0.3541	0.6641	0.7102
7	0.0590	0.3096	0.6679	0.6326	0.7222	0.4483	0.7771	0.1498	0.1359	0.8733	0.8727
8	0.0532	0.2924	0.6260	0.5939	0.6691	0.4159	0.7218	0.1447	0.1301	0.8109	0.8111
9	0.0511	0.2817	0.6023	0.5716	0.6433	0.3998	0.6938	0.1397	0.1256	0.7797	0.7797

(b) 복잡한 형태의 객체를 가지는 영상의 특징 벡터

그림 3.22. 다른 형태의 객체를 가지는 객체 영상의 특징 벡터 비교

위의 그림의 특징벡터에 표시된 부분에서 볼 수 있듯이 원에 가까운 객체를 가질수록 코너점의 개수는 적은 값을 가지게 되며, 객체가 복잡한 형태를 가질수록 코너점의 개수는 더 높은 값을 가지게 된다. 객체의 형태정보인 코너점을 분석하는 데 있어서 복잡함 정도에 더 비중을 두었으며, 코너점 개수의 절대적인 총량이 정해져 있지 않기 때문에 영상들의 특징벡터들을 각각 추출하여 코너점 개수의 값들을 분석하였다. 이에 따라 다음의 표 3.3과 같은 코너점 웨이티드 부여 표를 작성하였다.

표 3.3. 코너점의 웨이티드 부여

코너점의 개수	배경 영상	0~8	8~15	15~22	22~30	30~
웨이티드	0	1	2	3	4	5

3. 조밀도 분석 및 웨이티드 부여

조밀도의 특징도 코너점의 개수와 동일하게 형태적인 정보의 복잡함을 판단하는 데 있어서 중요한 정보이다. 영상을 비교함에 있어서 특징 벡터의 빠른 비교를 위해 객체마다 윤곽선의 틀린 모양을 정확하게 구별해 내기 보다는 객체의 뭉쳐짐 정도를 구별해 내는 것이 더 좋은 방법이다.

조밀도는 단위 둘레당 면적량으로 나타내어지므로 영상 내에서 객체의 영역이 얼마나 조밀하게 모여 있는지를 판별할 수가 있다. 그림 3.23은 형태가 단순하여 조밀도가

높은 객체를 가지는 영상의 특징벡터와 형태가 복잡하여 조밀도가 낮은 객체를 가지는 특징벡터를 비교한 것이다.

LV_inf <9x8 double>								
	1	2	3	4	5	6	7	8
1	16732	16690	15703	13663	12106	10740	9755	7491
2	6	7	1	3	2	5	8	4
3	0.1626	0.1622	0.1526	0.1328	0.1177	0.1044	0.0948	0.0728
4	40.4155	27.5868	39.4548	36.7285	34.4900	32.4471	31.0669	27.2400
5	5	9	5	6	6	5	6	8
6	0.6812	0.7140	0.6298	0.6606	0.6831	0.6238	0.6039	0.6391
7	0.6453	0.6867	0.8504	0.6499	0.7076	0.6545	0.8128	0.6252
8	0.5660	0.5993	0.8426	0.5824	0.6356	0.5876	0.8055	0.5635
9	0.5098	0.5476	0.8406	0.5311	0.5951	0.5377	0.8043	0.5139

(a) 높은 조밀도의 객체를 가지는 영상의 특징 벡터

LV_inf <9x7 double>							
	1	2	3	4	5	6	7
1	37065	4933	3836	833	804	725	573
2	2	3	13	12	7	1	8
3	0.7600	0.1012	0.0787	0.0171	0.0165	0.0149	0.0117
4	17.3607	13.2252	13.4596	8.4141	8.5532	6.7130	7.5395
5	24	19	24	29	26	22	27
6	0.2545	0.2977	0.3994	0.2801	0.3306	0.3846	0.2615
7	0.4004	0.5079	0.3480	0.4940	0.5483	0.1485	0.5385
8	0.3267	0.4231	0.3312	0.4108	0.4588	0.0957	0.4063
9	0.3106	0.4014	0.3241	0.3941	0.4332	0.0916	0.3871

(b) 낮은 조밀도의 객체를 가지는 영상의 특징 벡터

그림 3.23. 다른 조밀도 값을 가지는 객체 영상의 특징 벡터 비교

위의 그림에서 특징벡터에 표시된 부분에서 볼 수 있듯이 단순한 형태의 객체를 가질수록 조밀도는 높은 값을 가지게 되며, 객체가 복잡한 형태를 가질수록 조밀도는 더 낮은 값을 가지게 된다. 이는 전에 분석한 형태 정보인 코너점의 개수와 동일하게 적용할 수 있다. 객체의 형태정보인 코너점과 동일하게 조밀도를 분석하는 데 있어서 복잡함 정도에 더 비중을 두었다.

조밀도는 코너점과는 다르게 영상의 객체가 복잡할수록 낮은 값을 가지게 되므로 웨이티드는 낮은 값을 가지는 쪽에 더 많이 부여가 되도록 설정하였으며, 영상들의 특징 벡터들을 분석해 본 결과 가장 단순한 형태의 객체가 40에서 45 사이의 값을 가진다는 것을 확인할 수 있었다. 이에 따라 다음의 표 3.4와 같은 조밀도 웨이티드 부여 표를 작성하였다.

표 3.4. 조밀도의 웨이티드 부여

조밀도	배경 영상	0~10	10~18	18~28	28~37	37~
웨이티드	0	5	4	3	2	1

4. 색상 분석 및 웨이티드 부여

색상은 면적이거나 형태적인 정보들을 판단하는 데에 특징벡터를 이용한 것과는 다르게 색상 정보를 추출할 때 사용된 인덱스 이미지의 인덱스 번호의 분포를 이용한다. 이는 특징 벡터에서 객체의 색상 평균값으로 색상의 웨이티드를 판단할 수가 없기 때문이다.

인덱스 이미지는 데이터 행렬과 칼라맵 행렬로 구성되는데 여기서 데이터 행렬을 이용하게 된다. 데이터 행렬은 uint8을 사용하여 만들어진 칼라맵 8X3 행렬의 픽셀 매핑을 위한 주소라고 볼 수 있다. 즉, 데이터 행렬은 질의 영상의 크기와 같으며 칼라맵과의 매핑을 위해 픽셀마다 1에서 8사이의 값을 가지고 있다. 값 1이 map의 첫 번째 행을, 값 2가 두 번째 행을 가리키는 식으로 처리가 된다. 칼라맵은 대개 인덱스 이미지와 함께 저장되어 해당 이미지와 함께 자동으로 불러지는 방식이지만 디폴트 칼라맵만 사용하도록 제한되어 있지 않으며, 선택하는 어떤 칼라맵이든 사용할 수 있다. 다음 그림 3.24는 인덱스 이미지에서 같은 색상의 분포가 높은 객체와 낮은 객체를 비교하여 보여준다.



(a) 같은 색상 분포가 높은 객체

(b) 같은 색상 분포가 낮은 객체

그림 3.24. 같은 색상의 분포가 다른 객체 비교

이에 따라 객체마다 데이터 행렬의 칼라맵 번호의 분포도를 파악할 수가 있다. 색상의 분포도를 판단하는 것은 식 (3.9)로 판단한다.

$$distribution = \frac{\sum ID_{main}}{A_{total}} \times 100 \quad (3.9)$$

ID_{main} 은 인덱스 데이터 행렬에서 객체의 주요 번호를 뜻하며, A_{total} 은 객체 면적의 총량을 의미한다. 즉, 객체에서 가장 많은 분포를 가지는 칼라맵 번호의 개수를 객체의 총면적에 나누어 백분율로 나타낸 것이다.

이를 이용하여 가장 많은 분포를 가지는 칼라맵 번호의 개수와 객체의 총 면적을 비교하여 60%가 넘을 경우 색상 정보의 중요성이 높은 것으로 설정하였다. 그리고 60% 이하의 값들은 다른 객체 영상들의 색상 분포를 분석하여 표 3.5와 같은 웨이티드 표를 작성하였다.

표 3.5. 색상의 웨이티드 부여

색상 분포	배경 영상	0~20	20~33	33~45	45~60	60~
웨이티드	5	1	2	3	4	5

5. 적응적 웨이티드

본 논문에서 질의 영상에서 추출된 특징 벡터를 이용하여 각각의 특성을 분석한 결과에 따라 웨이티드를 통합하는 과정이 필요하다. 각 특성에 따라 웨이티드를 부여한 후 적응적으로 웨이티드를 부여하는 식은 다음과 같다.

$$Aw_{total} = \sum_{i=1}^n w_i (area, corner, compactness, color) \quad (3.10)$$

여기서 Aw_{total} 는 질의 영상에 따른 적응적 웨이티드를 의미하며 각각 정규화 면적, 조밀도, 코너점의 개수, 색상의 특성을 의미한다. n 은 최대 4의 값을 가지고, w 는 각 특징벡터의 웨이티드를 뜻하며 이는 객체 분석을 통해 결정된 값이다. 웨이티드는 질의 영상의 주요 특징에 따라서 다른 값을 줘서 중요한 특징을 더 부각하여 비교를 하게끔 하는 역할을 한다.

E. 유사도 측정 및 순위 결정

1. 유사도 측정

질의영상에 의해 추출된 특성벡터는 데이터베이스 내에 있는 영상들의 색인벡터와 비교함으로써 영상검색이 시작된다. 일반적인 검색시스템과는 달리 내용기반 영상검색은 정확한 검색을 요구하는 것이 아니고, 질의영상과의 유사성의 순서에 따라 데이터베이스 내에 있는 영상을 순위별로 검색하여야 한다. 질의벡터와 색인벡터간의 유사한 정도는 유사도 측정 또는 유사도 계산과정으로 수행한다[86]. 즉, 일반적인 검색시스템에서는 정확히 일치하는 자료를 검색하지만, 내용기반 영상검색에서는 질의영상과 특성값이 정확히 일치하는 하나의 영상을 찾는 것이 아니라, 특성값이 유사한 다수의 영상들을 검색결과로 제시하여 사용자가 최종적으로 선택할 수 있도록 한다.

이와 같이 검색결과를 계산하는 방법은 지난 수십 년간 정보검색 분야에서 문헌검색에 사용된 다양한 방법을 찾아 볼 수 있다. 이러한 다양한 방법 중에서 대표적인 ‘단순한 정합’, ‘Dice 계수’, 그리고 ‘Jaccard 계수’에 의한 방법은 다음과 같이 나타낸다 [87].

$$S_{sm}(Q, T) = |Q \cap T| \quad (\text{단순 정합}) \quad (3.11)$$

$$S_{Dice}(Q, T) = 2 \frac{|Q \cap T|}{|Q| + |T|} \quad (\text{Dice의 방법}) \quad (3.12)$$

$$S_{Jaccard}(Q, T) = \frac{|Q \cap T|}{|Q \cup T|} \quad (\text{Jaccard의 방법}) \quad (3.13)$$

여기서 Q 는 질의벡터, T 는 데이터베이스내의 색인벡터를 말하며, 식 (3.13)은 질의벡터와 인덱스 벡터에서 같은 특성을 가지는 인수들의 개수를 나타낸다. 다시 말하면, 영상에서 시각적으로 같은 색상을 가지는 화소들의 개수를 표시하는 것과 같다. 식은 각각 표현한 정합의 개념을 포함하고는 있지만 다른 방법임을 알 수 있다. 단순한 정합은 같은 자료만 계수하며, Dice 및 Jaccard 방법은 비교대상간의 정합계수는 최대 1이하가 되도록 정규화하고 있다.

영상검색시스템에서 가장 많이 사용되는 유사도 측정방법은 식 (3.14)에서와 같이 벡터간의 거리를 계산하는 L_p norm 계산방법이다.

$$L_p(\bar{Q}, \bar{T}) = \left| \sum_i^n (Q_i - T_i)^p \right|^{1/p} \quad (3.14)$$

두 영상간의 유사성을 측정하기 위한 함수는 유클리드 거리함수를 이용한다, 유클리드 거리함수의 특징은 정규직교 변환 하에서 그 거리가 항상 보존된다는 것이다. 유클

리드 거리를 두 벡터간의 거리로 사용하면, 위 식에서 P 의 값은 2로 할 수 있다. 그리하여 다음 식과 같이 유사도를 계산한다.

$$Distance = \left| \sum_i^n (Q_i - T_i)^2 \right|^{1/2} \quad (3.15)$$

$$Similarity = 1 - Distance \quad (3.16)$$

위의 식에서 알 수 있는 바와 같이 질의영상의 벡터와 색인벡터간의 차이를 식 (3.15)에서 구한 후, 그 값의 역으로 식 (3.16)과 같이 유사도를 측정한다. 특징벡터의 요소가 복합적인 요소로 구성되어 있으므로 앞의 웨이티드 부여 방법에 따라 벡터간의 웨이티드 값을 주어 정규화하여 유사도에 의한 검색순위를 결정한다.

2. 순위 결정

순위결정이라 함은 검색된 영상이 유사도 측정에 의해 계산된 수치가 작은 순서에서 큰 순서로 나열하는 것이며, 검색결과에는 통상적으로 검색된 영상과 함께 순위에 따라 영상들이 표시된다. 이러한 순위는 시스템의 성능을 평가할 때 매우 중요한 자료로 사용된다. 본 논문에서는 검색결과를 쉽게 찾아보며, 순위를 시각적으로 확인할 수 있는 영상표시 윈도우창을 그림 3.24와 같이 제작하였다. 그림에서 보는 바와 같이 가장 위에 있는 영상이 질의영상이며, 두 번째 줄에서부터 우측으로 가장 유사한 순으로 검색결과를 20장씩 보여 준다.

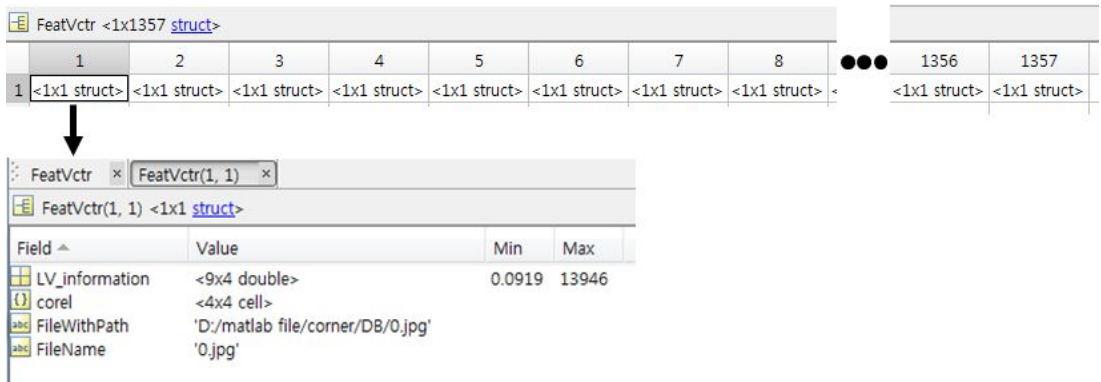


그림 3.24. 유사영상 검색 결과창

IV. 실험 및 분석

A. 실험

본 논문에서 제안한 알고리즘은 다양한 영상에 대하여 MATLAB 2013 소프트웨어를 사용하여 실험하였다. 본 실험은 다양한 질의 영상들에 대해서 수행되었으며, 영상 검색 방법을 설명하기 위해 Wang 데이터베이스를 기본으로 하여 house, people, leap 등의 데이터베이스를 추가한 1357장의 영상들에서 가장 유사한 20개를 찾도록 정의하여 놓았다. 실험에서 질의 영상과 비슷한 영상은 알맞은 영상의 개수로 만들어 놓았다. 본 알고리즘의 시뮬레이션은 데이터베이스의 영상들을 대상으로 특징 벡터들의 데이터를 미리 생성해 놓았다. 저장해 놓은 데이터는 나중에 다른 질의 영상을 입력하여도 따로 계산 과정을 거치지 않고 질의 영상만 계산하여 데이터의 특징 벡터들을 이용해 유사도 측정을 비교하여 결과 영상을 보여주게 된다. 그림 4.1은 저장해 놓은 데이터베이스 영상들의 특징 벡터 데이터를 보여준다.



The image shows a MATLAB interface. At the top, a variable named 'FeatVctr' is defined as a 1x1357 struct array. Below this, a detailed view of the first element, 'FeatVctr(1, 1)', is shown. This view includes a table with the following data:

Field	Value	Min	Max
LV_information	<9x4 double>	0.0919	13946
corel	<4x4 cell>		
FileWithPath	'D:/matlab file/corner/DB/0.jpg'		
FileName	'0.jpg'		

그림 4.1. 데이터베이스 영상들의 특징 벡터 데이터

사용된 데이터베이스의 영상들의 종류를 분류하면 표 4.1과 같다.

표 4.1. 이미지 데이터베이스의 분류

영상 분류	관련 영상
원주민	98
해변	97
사원	93
버스	100
공룡	100
코끼리	100
꽃	100
말	100
산	100
음식	100
사람	200
집	72
나뭇잎	96
총 영상 수	1357

영상의 특성에 기반한 적응적인 웨이티드 CBIR 알고리즘의 실험을 위하여 위의 데이터베이스 영상들의 특징 벡터들을 앞에서 설명한 알고리즘의 방식으로 분석하였다. 그리고 데이터베이스의 이미지 그룹에서 각각 질의 영상들을 선정하여서 실험하여 그 결과를 다음과 같이 나타내었다. 그림은 질의 영상, 찾고자 하는 유사영상들, 검색 결과 이미지의 순으로 보여준다.

먼저 객체 판별에 따라서 분류된 배경이 주된 영상인 해변영상과 산 영상 중에서 질의영상을 선정하여 실험한 검색 결과를 나타낸 이미지들이다.



그림 4.2. 해변의 질의 영상

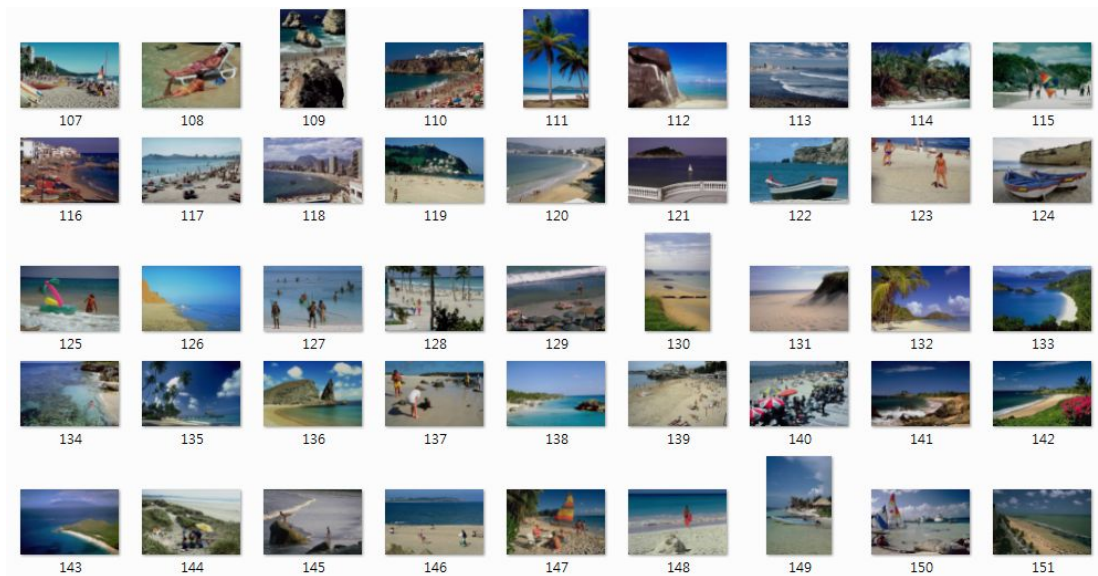


그림 4.3. 데이터베이스 내 해변 영상의 유사영상들

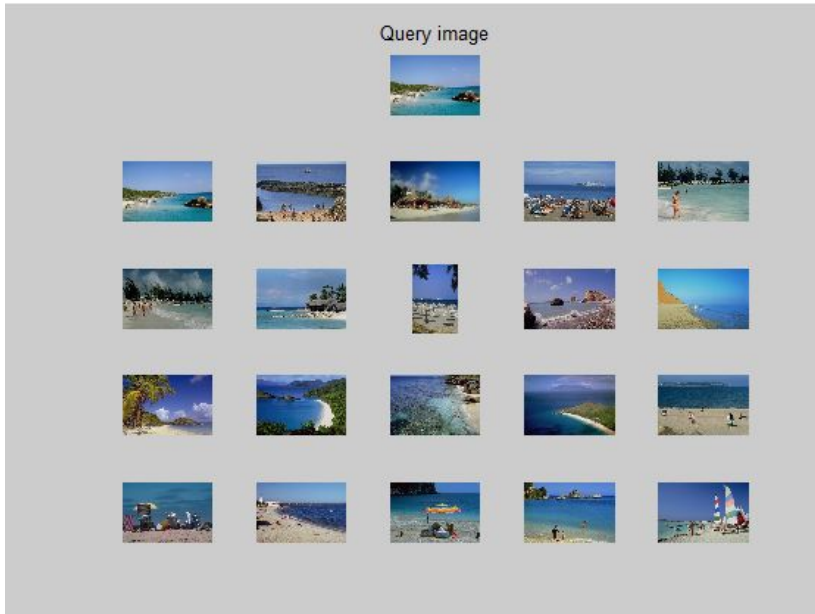


그림 4.4. 해변의 검색 결과(w -면적:0, 코너:0, 조밀:0, 색상:5)



그림 4.5. 산의 질의 영상

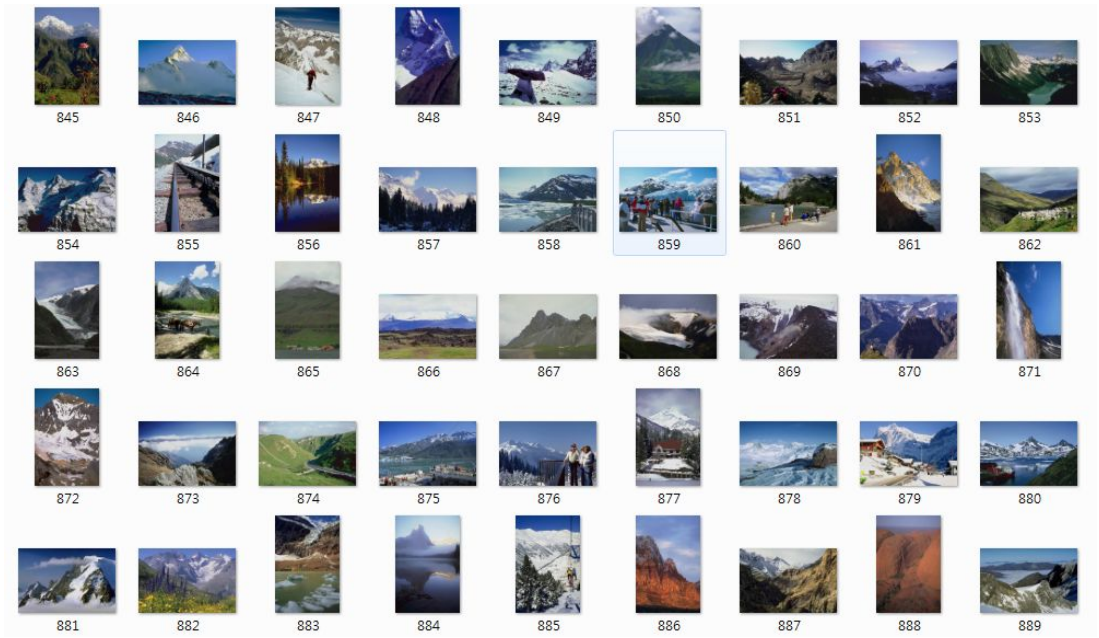


그림 4.6. 데이터베이스 내 산 영상의 유사영상들

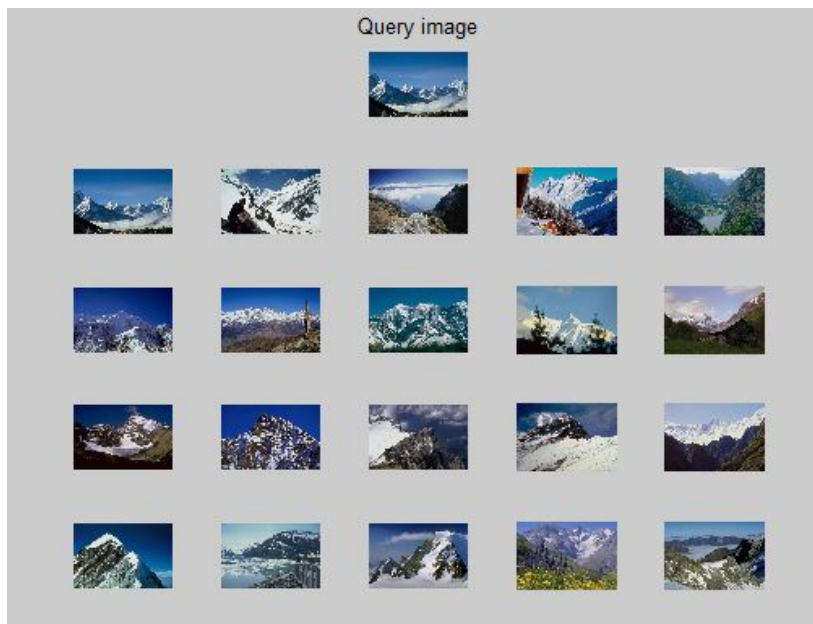


그림 4.7. 산의 검색 결과(w -면적:0, 코너:0, 조밀:0, 색상:5)

배경이 주된 영상들은 가장 큰 객체의 면적이 다른 객체들보다 월등히 크다는 특징이 있다. 이는 영상의 레이블링 과정 중에서 색상이 단순한 부분들 하나의 큰 객체로 인식하여 분할하기 때문이다. 이를 바탕으로 색상 정보에만 높은 웨이트드를 부여하여 영상검색을 실시하였다. 위의 해변과 산의 검색결과와 유사 영상들 중에서도 색상이 비슷한 영상들이 먼저 검색이 되는 것을 시각적으로 확인이 가능하였다.

그리고 객체가 판별된 영상들 중에서 먼저 단일 객체를 가지는 영상들에 대해서 질의 영상을 선정하고 다음과 같이 실험하였다.



(a) 질의 영상1



(b) 질의 영상2

그림 4.8. 질의 영상 1, 2

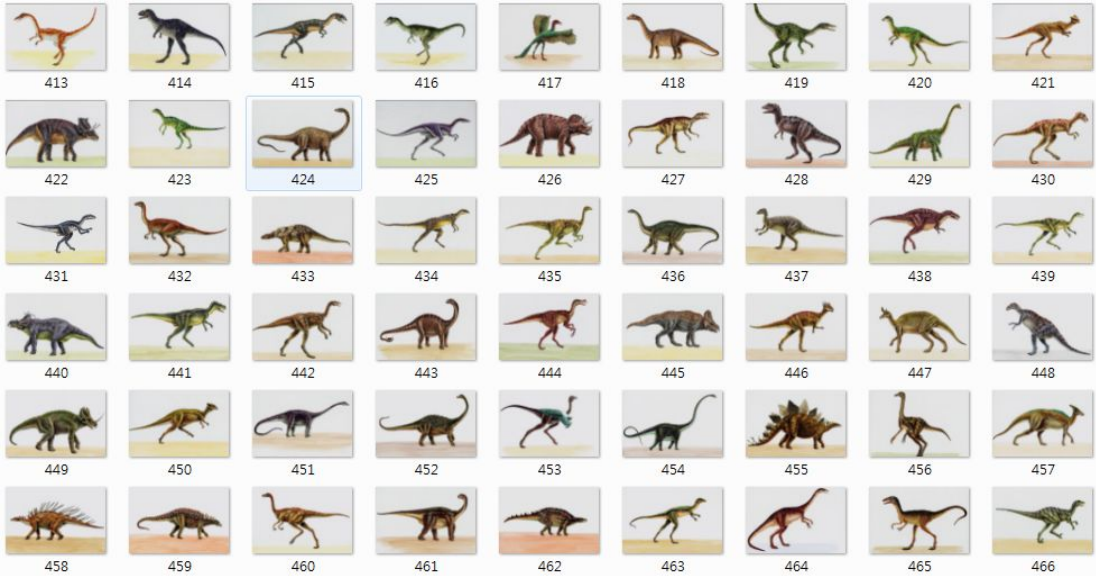
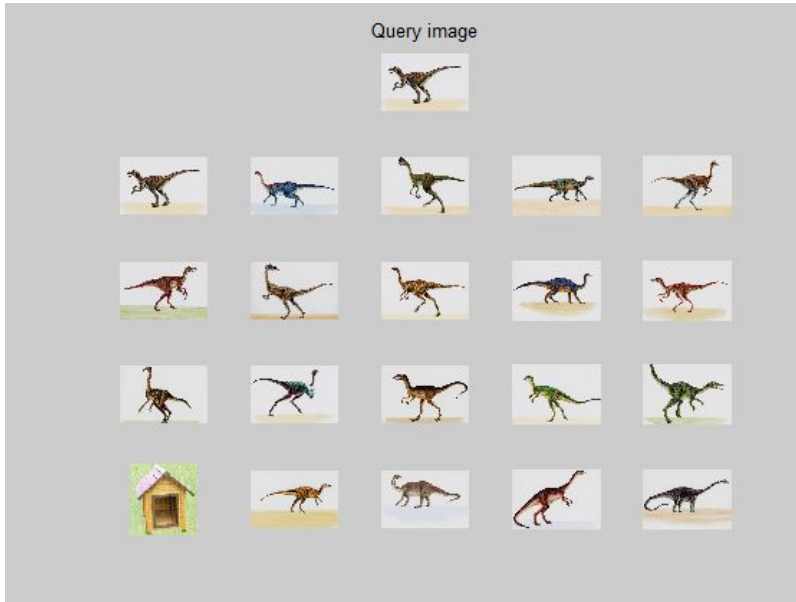


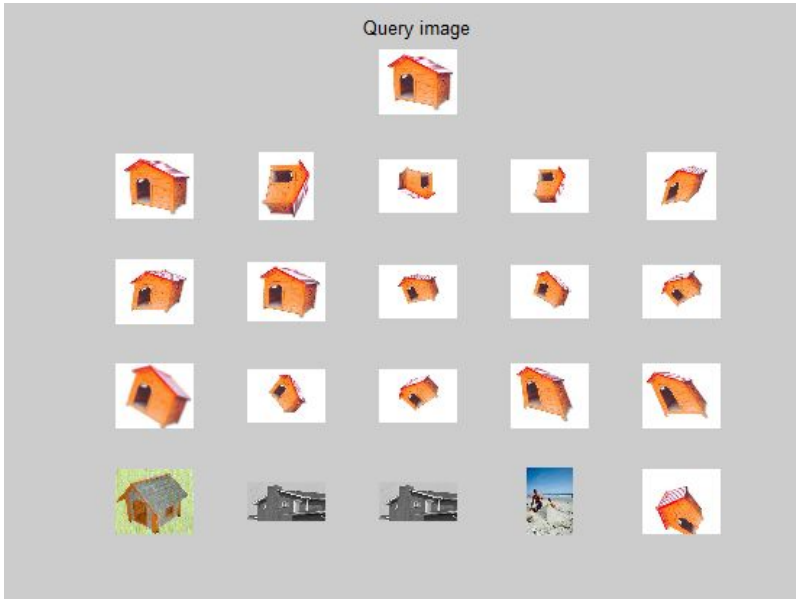
그림 4.9. 데이터베이스 내 질의 영상1의 유사영상



그림 4.10. 데이터베이스 내 질의 영상2의 유사영상



(a) 질의 영상1의 검색 결과(w -면적:5, 코너:4, 조밀:4, 색상:4)



(b) 질의 영상2의 검색 결과(w -면적:5, 코너:3, 조밀:3, 색상:5)

그림 4.11. 공룡과 집의 검색 결과

각각의 질의 영상들에 대한 각각의 유사영상들에서 제안된 알고리즘의 검색 결과는 그림 4.11에서 나타내었으며 이는 단일 영상의 특징을 쉽게 판별하여 분류가 가능했기 때문에 시각적으로 검색의 성능이 우수하다는 것을 확인하였으며, 영상의 회전, 확대 및 축소, 찌그러짐 등의 변형에서도 검색 성능이 우수하다는 것을 확인할 수 있었다.

그리고 다중 객체를 가지는 영상들 중에서 명확하게 객체가 분류되는 나뭇잎 영상을 질의 영상으로 선정하여 실험한 결과이다.



그림 4.12. 나뭇잎의 질의 영상

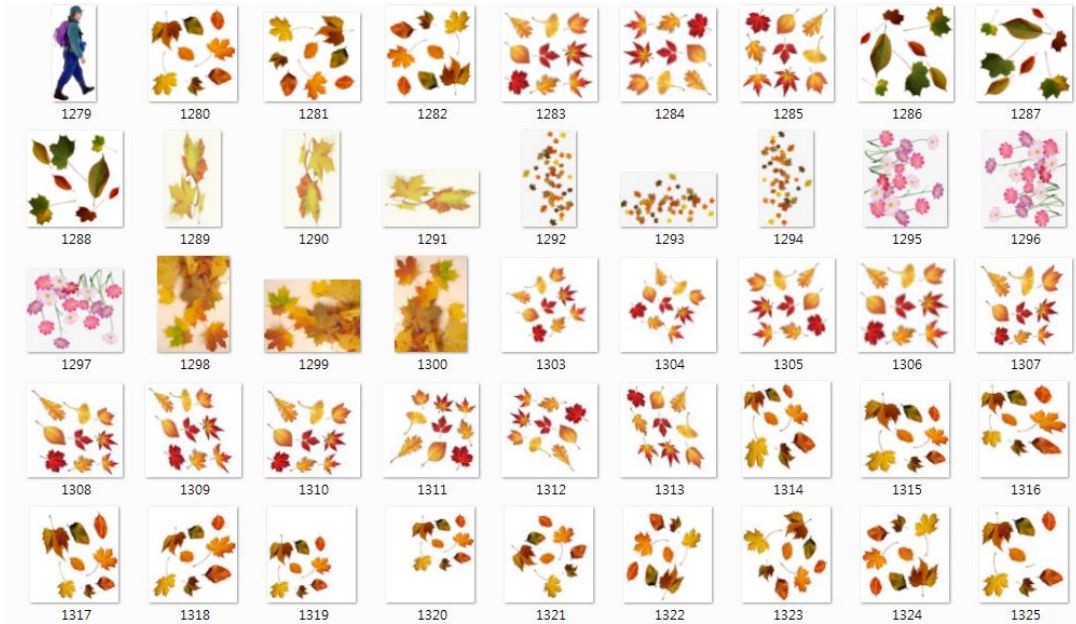


그림 4.13. 데이터베이스 내 나뭇잎 영상의 유사 영상들



그림 4.14. 나뭇잎의 검색 결과(w -면적:2, 코너:4, 조밀:4, 색상:2)

나뭇잎이나 동전과 같이 레이블링이 명확한 다중 객체 영상들은 객체들의 크기가 거의 비슷하다는 특징이 있다. 이는 영상 내에서 명확한 객체를 가지기 위해서는 어느 정도의 모양과 크기를 가져야하기 때문이다. 이와 같은 이유로 정규화 면적과 형태적인 특징 정보인 조밀도와 코너점의 개수, 색상값에 적응적으로 웨이트드가 부여되어 영상 검색을 한 결과는 형태가 비슷한 다중 객체 영상들이 먼저 검색되면서 영상의 찌그러짐이나 회전 등의 영상 변화에서도 강인함을 시각적으로 확인할 수 있었다.

복잡한 영상들은 명확하게 객체가 추출이 되거나 확연하게 특징이 추출되어지지 않는다. 이에 따라 복잡한 영상을 대상으로 질의 영상을 선정하여 실험할 경우 검색하고자하는 유사 영상에 포함되어 있지 않은 영상들이 검출될 확률이 높았다. 아래의 결과는 복잡한 원주민 영상들 중 질의 영상을 선정하여 영상 검색을 실시한 것이다.



그림 4.15. 원주민의 질의 영상

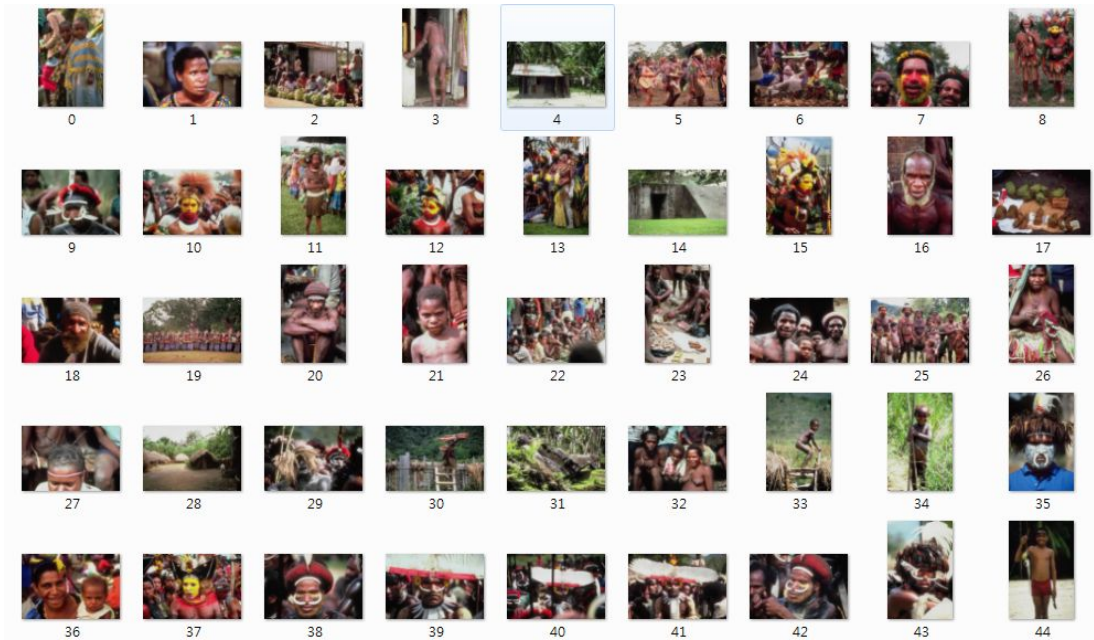


그림 4.16. 데이터베이스 내 원주민 영상의 유사영상들

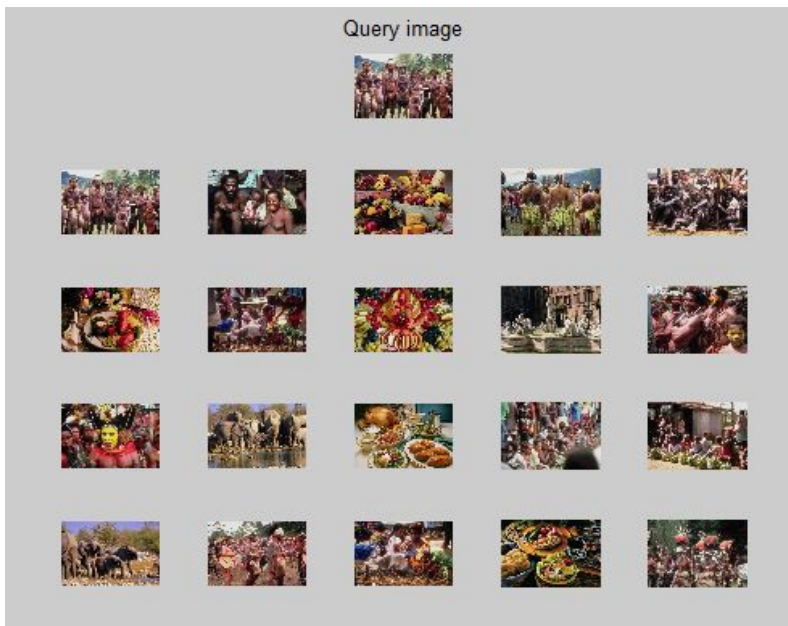


그림 4.17. 원주민 영상의 검색 결과(w -면적:3, 코너:3, 조밀:4, 색상:2)

실험한 결과 배경이 주된 영상이나 객체가 판별된 영상들 중 명확하게 객체를 가지고 있는 질의 영상들의 경우 시각적으로 좋은 유사 영상의 검출률을 가지는 결과를 보여주었으나, 영상의 복잡도가 심하여 객체의 특징이 정확히 파악이 되지 않은 질의 영상의 경우 유사 영상의 검출율이 다른 질의 영상에서보다 다소 떨어지는 것을 보여주었다.

B. 분석

영상검색시스템을 개발하는 과정에서 시스템의 구성과 영상내의 특징을 추출하는 방법의 개발도 중요하지만 이에 못지않게 시스템의 성능을 분석할 수 있는 성능평가과정은 매우 중요하다. 그러나 성능평가를 위해 사용하고 있는 평가도구가 문헌마다 달라서 일관성 있는 기준을 설정하지 못하고 있으며, 성능을 평가하기 위해 공통적으로 사용할 수 있는 실험영상도 존재하지 않아 기준이 되는 데이터베이스 영상을 시험용으로 사용하고 있으므로, 이의 개선을 위한 많은 연구가 요구되고 있다. 그리하여 본 논문에서는 이러한 성능평가에 대한 개념을 정리하고 적절한 평가도구를 사용함으로써 제안된 시스템의 평가에 대한 객관성을 확보하고자 한다.

현실적으로 영상검색기술의 적용과 검색능력을 향상시키는데 장애물로 작용하는 여러 요인들이 존재하고 있어, 이 분야를 연구하는 연구자들은 다음과 같은 의문을 항상 가지게 된다.

- i) 영상검색 방법들의 성능은 어떻게 비교하는가?
- ii) 영상특징들은 어떻게 선택하며, 추출된 특성은 어떻게 평가하는가?
- iii) 새로운 알고리즘을 발표하는 논문은 어떻게 조사하는가?
- iv) 영상검색기술과 시스템능력을 향상하기 위해 상호협조는 어떻게 이루어지는가?

위와 같은 의문들은 지난 수십 년간에 걸쳐 문자검색, 데이터베이스 및 컴퓨터비전과 같은 다양한 분야에서 일어나고 있었다. 이러한 의문에 답하기 위해 문자검색용 테스트-베드를 준비하기 위한 노력을 TREC(Text Retrieval Conference)에서 수행하고 있으며, 또한 문자검색시스템의 성능향상과 평가를 위한 포럼을 운영하기도 한다[88]. 이러한 연구노력의 결과는 영상검색분야에도 많은 영향을 끼쳐, 표준영상과 벤치마킹

할 수 있는 질의영상이 가까운 미래에 출현하리라 본다.

영상검색시스템은 영상들을 순위를 부여하여 검색하는 것이 목적이지만, 순위를 결정함에 있어 수작업으로 질의영상과의 관련성 여부를 결정해야 하는 주관적인 면이 있었다. 그럼에도 불구하고 시스템의 전체적인 검색효율성은 이러한 관련 있는 영상들을 모두 알고 있을 경우에만 측정이 가능하게 된다. 일반적으로 정보검색시스템의 테스트-베드는 다음과 같은 조건들이 구비되고 있어야 한다.

- i) 검색대상의 표준 자료 (문서, 영상 등)
- ii) 표준 질의 자료
- iii) 질의와 관련 있는 자료목록
- iv) 평가도구

정보검색에서 사용하고 있는 검색효율성 평가방법은 영상검색에서도 적용할 수가 있으며, 이에 대한 기본적인 개념을 그림 4.18에 표시한다.

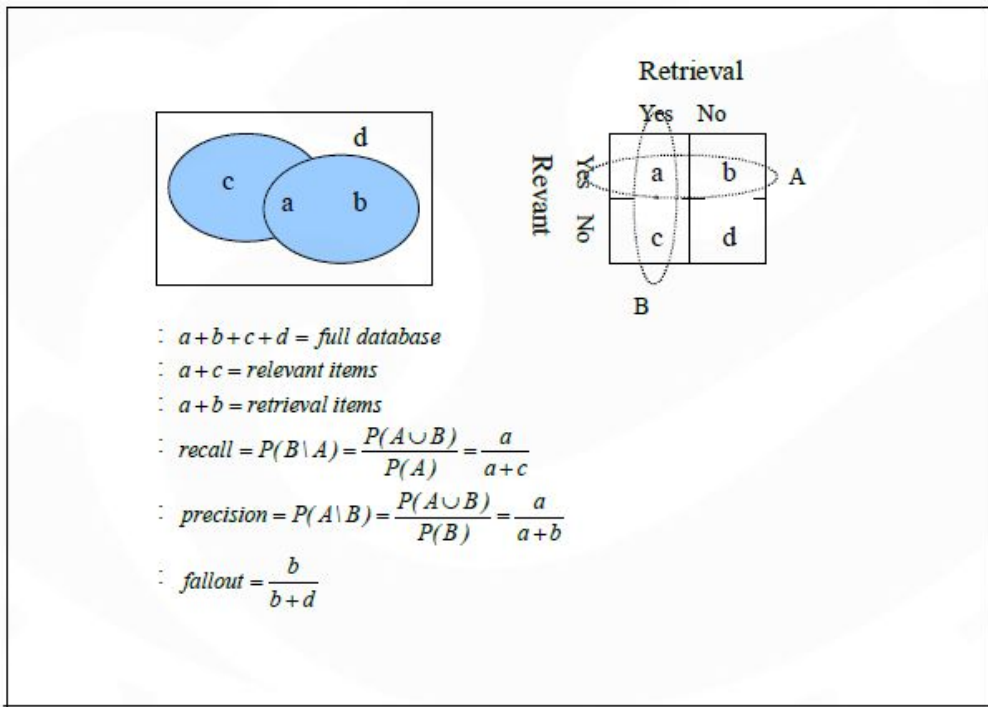


그림 4.18. Recall 과 Precision

그림 4.15에서 a는 정확하게 검색된 정보의 집합, b는 검색은 되었으나 관련이 없는 것, c는 관련은 있으나 검색에 실패한 정보, 그리고 d를 검색되지도 않고 관련도 없는 정보의 집합이라고 하면, A는 검색된 정보의 집합(a+b), B는 관련이 있는 정보들의 집합(a+c)으로 볼 수 있다. 이러한 벤 다이어그램을 기본으로 재현율과 검출율 및 실패율(fallout)을 계산하면 그림과 같다.

정보검색을 위하여 시스템에 질의하면 순위를 부여하여 검색결과를 보여주는데, 이때 k 번째 되는 순위까지를 검색결과로 한다면 다음과 같은 값들을 구할 수 있다. 여기서 순위 n 에 들어가는 정보의 집합을 V_n 이라면, $V_n \in \{0,1\}$ 이며, 그림 4.15에서

$$a, b, c, d \text{를 각각 } A_k, B_k, C_k, D_k \text{로 다시 나타내면, } A_k = \sum_{n=0}^{k-1} V_n, B_k = \sum_{n=0}^{k-1} (1 - V_n)$$

$$C_k = \sum_{n=0}^{N-1} V_n - A_k, D_k = \sum_{n=0}^{N-1} (1 - V_n) - B_k \text{로 표시할 수 있다. 이러한 값들로부터 표준}$$

정보검색척도로 사용되는 재현율, 검출율 및 실패율을 계산할 수 있으므로 이를 식으로 나타내면 다음과 같다.

$$Recall : R_k = \frac{A_k}{A_k + C_k} \quad (4.1)$$

$$Precision : P_k = \frac{A_k}{A_k + B_k} \quad (4.2)$$

$$Fallout : D_k = \frac{B_k}{B_k + D_k} \quad (4.3)$$

위의 식에 의하여 정보검색시스템의 효율성을 측정할 수 있으며, 다른 시스템과도 비교할 수 있다.

정보검색분야에서 사용하고 있는 개념을 영상검색시스템의 효율성을 평가하기 위해

도입한 다양한 기법들이 있다. 즉, 본 논문에서 다루고 있는 내용기반 영상검색은 정보검색에서와 같이 정확한 검색을 요구하는 것이 아니라 유사한 영상을 검색하게 되므로 앞에서 언급한 유사도 측정과 관련하여 평가 방법을 고려해야 한다. 재현율, 검출율 및 실패율을 사용하여 영상검색시스템의 성능을 평가하는 방법들은 다음과 같이 다양한 방법이 있다.

- i) P_k vs. R_k : 검색시스템의 효율성을 나타내며 가장 많이 사용되는 도구로써 신뢰성이 높음.
- ii) A_k vs. B_k : 검색 성공율.
- iii) A_k vs. F_k : 상대적 검색 성공율.
- iv) P_k at cut-off $k = \int \left(\sum_{n=0}^{N-1} V_n \right)$: 관련되는 영상을 최하위까지 고려.
- v) average P_k at $R_k=0.2, 0.5, 0.8$: recall 0.2, 0.5, 0.8인 경우.
- vi) average P_k at eleven recall point : 11개의 recall점.
- vii) AVRR(average rank of all relevant items) : 유사한 영상이 검색된 순위들의 평균값으로 질의영상과 가장 유사한 영상의 순위는 “0”임.
- viii) AVRR/IAVRR(ideal average relevant rank) : IAVRR은 유사한 영상들이 상위에서 차례로 검색되는 경우를 말한다.
- ix) $\frac{A_k}{B_k}$ vs. A_k : 응답율

이상과 같은 평가도구들은 대부분 재현율, 정확율과 실패율을 이용하여 계산하게 된다. 또한 이러한 도구들은 공통적인 척도와 함께 상호 연관성을 가지고 평가해야만 올바른 성능평가가 이루어 질 수 있다. 그런 의미에서 정보 검색에서 표준으로 사용되고 있는 평가방법인 재현율 대 검출율(Recall vs. Precision)은 많은 장점을 지니고 있다. 재현율에 대한 정확율의 값이 높으면 상대적으로 더 나은 시스템이라 평가한다. 즉, 모든 재현율이 가지는 값에 대하여 검출율의 값이 높으면 비교하는 시스템보다 성능이 좋다고 평가한다.

제한한 알고리즘을 평가하기 위한 방법으로, 주관적인 평가와 객관적인 평가를 병행

하였다. 주관적인 평가는 검색된 결과를 영상표시 윈도우창에서 시각적으로 직접 확인하는 방법을 사용하여 검색된 영상들이 질의 영상과 얼마나 유사한지를 직접 판단할 수 있도록 하였으며, 객관적인 평가는 앞서 설명한 객관적 평가 척도인 Recall과 Precision을 사용하였다. 또한 평가의 객관성을 유지하기 위해 동일한 영상자료로서 제안한 방법과 색상을 이용한 영상 검색 기법인 Varish의 색상 히스토그램의 통계적인 특성을 이용한 방법[89], 그리고 형태를 이용한 영상 검색 기법인 Nikkam의 형태 템플릿 매칭 기반 분석 방법[90]과 Song의 HSV색상과 균일국부이진패턴을 이용한 방법[91]들을 비교 분석하였다. 제안한 방법은 레이블링, 색상과 농도의 화소값, 코너점, 정규화 면적, 조밀도 등의 특성을 기반으로 특징 벡터를 생성하였으며, 질의 영상에 따라서 자동으로 분류하고 적응적으로 웨이트드를 부여하여 영상을 검색하게 하는 반면, Varish에 의한 색상 히스토그램의 통계적인 특성을 이용한 방법은 색상 특성의 RGB 세가지 히스토그램을 여러개의 bin으로 분류하고 편차와 왜도, 첨도의 통계적인 값을 기반으로 특징을 분류하는 방법이다. 그리고 Nikkam에 의한 형태 템플릿 매칭 기반 분석 방법은 템플릿 기반으로 형태의 전역적인 특성을 추출하고 이 데이터를 수평적인 분석으로 윤곽선의 형태를 매칭하는 방법이다. 마지막으로 Song의 HSV색상과 균일국부이진패턴을 이용한 방법은 영상을 HSV 색상 모델로 변환하여 Hue를 추출하고 영상을 회색 음영으로 변환하여 ULBP(Uniform Local Binary Patterns)를 구한다. 그리고 추출된 Hue의 값에 따른 ULBP를 10개로 분류하여 만들어지는 히스토그램을 특징으로 사용하는 방법이다.

제안한 방법의 성능 비교를 위해 색상의 특성을 이용한 Varish의 색상 히스토그램의 통계적인 특성을 이용한 방법[89]과 형태의 특성을 이용한 Nikkam의 형태 템플릿 매칭 기반 분석 방법[90] 그리고 Song의 HSV색상과 균일국부이진패턴을 이용한 방법[91]을 각각 실험하여 각 그룹의 질의 영상에 대한 평균 Recall과 Precision을 계산하여 표 4.2에 나타내었다.

표 4.2. 이미지 그룹별 Recall 및 Precision

	색상 히스토그램 통계 특성 [89]		형태 템플릿 매칭 분석 [90]		균일국부 이진패턴 [91]		제안한 방법	
	Re.	Pre.	Re.	Pre.	Re.	Pre.	Re.	Pre.
원주민	0.39	0.45	0.51	0.35	0.42	0.46	0.59	0.47
해변	0.67	0.54	0.69	0.41	0.58	0.51	0.73	0.48
사원	0.62	0.52	0.66	0.39	0.60	0.49	0.65	0.52
버스	0.58	0.69	0.83	0.76	0.81	0.58	0.89	0.78
공룡	0.70	0.91	0.97	0.77	0.91	0.72	0.94	0.75
코끼리	0.61	0.62	0.68	0.50	0.73	0.54	0.71	0.49
꽃	0.72	0.82	0.87	0.64	0.88	0.65	0.91	0.74
말	0.58	0.69	0.81	0.43	0.75	0.61	0.73	0.47
산	0.61	0.65	0.58	0.47	0.69	0.51	0.66	0.49
음식	0.27	0.40	0.51	0.36	0.47	0.51	0.60	0.42
사람	0.51	0.63	0.66	0.51	0.43	0.49	0.61	0.46
집	0.68	0.93	0.94	0.71	0.89	0.67	0.92	0.77
나뭇잎	0.62	0.79	0.83	0.58	0.79	0.57	0.88	0.67
평균	0.58	0.66	0.73	0.53	0.69	0.56	0.76	0.58

제안한 알고리즘은 색상 히스토그램 통계 특성을 이용한 방법과 비교하여 꽃 영상의 경우 재현율이 0.19이 더 높았지만 검출율은 0.08이 더 낮았으며, 배경 영상인 산 영상의 경우에도 재현율이 0.05가 더 높고 검출율은 0.16이 더 낮았다. 위의 실험 결과를 토대로 제안한 알고리즘은 색상 히스토그램 통계 특성을 이용한 방법보다 재현율은 높지만 검출율이 낮은 방법이라는 것을 확인할 수 있었다. 그리고 형태 템플릿을 이용

한 방법과 비교하여 버스 영상의 경우 재현율이 0.06이 더 높고 검출율도 0.02가 더 높았으며, 복잡한 영상인 원주민 영상에서도 재현율이 0.08이 더 높고 검출율도 0.12가 더 높다는 것을 확인할 수 있었다. 그러나 단순한 형태를 비교해야하는 공룡 영상의 경우 재현율이 0.03이 더 낮았으며 검출율도 0.02가 더 낮은 것을 확인할 수 있었다. 이는 형태 속성을 이용한 알고리즘에 더 적합한 질의 영상을 사용한 경우에도 제안한 알고리즘도 그에 못지않은 실험 결과를 가진다는 것을 알 수 있다. 마지막으로 HSV색상과 균일국부이진패턴을 이용한 방법과 비교한 결과, 단순한 영상인 집 영상의 경우 재현율이 0.03이 더 높고 검출율은 0.08이 더 높았으며, 복잡한 영상인 음식 영상의 경우에는 재현율이 0.13이 더 높았지만 검출율은 0.09가 더 낮았다. 그리고 배경이 주된 영상인 산 영상에서는 재현율이 0.03이 더 낮았고 검출율도 0.03이 낮았다. 대체적인 질의 영상의 경우 제안한 알고리즘이 더 좋은 재현율과 검출율을 보였지만 색상이나 패턴의 속성 분석이 주로 필요한 질의 영상의 경우에는 HSV색상과 균일국부이진패턴을 이용한 방법이 더 좋은 재현율과 검출율을 보였다.

또한 제안한 알고리즘의 성능 향상을 직관적으로 확인이 가능하도록 막대그래프로 표시하여 비교하였다. 다음의 그림 4.19는 색상 히스토그램 통계 특성 방법, 형태 템플릿 매칭 방법, 그리고 균일국부 이진패턴 방법의 기존 알고리즘들과 제안한 알고리즘을 막대그래프로 비교하여 나타낸 것이다.

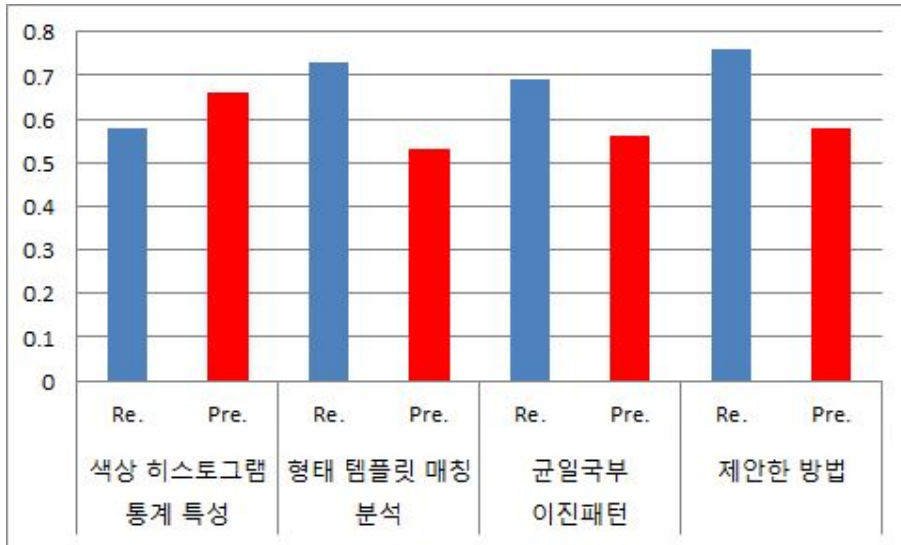


그림 4.19. 평균 Recall과 Precision 비교 그래프

평균 Recall과 Precision 막대그래프에서 확인해 볼 수 있듯이 기존의 알고리즘인 HSV색상과 균일국부 이진패턴 방법이나 형태 템플릿 매칭 분석 방법에 비해 검출율이나 재현율이 더 높은 것을 확인할 수 있었다. 그리고 색상 히스토그램 통계 특성 방법은 검출율이 더 좋은 방법으로 제안한 방법보다 검출율은 상당히 좋은 성능을 보였지만 재현율은 제안한 방법이 훨씬 성능이 좋다는 것을 확인하였다. 앞에서 Recall과 Precision을 이용하여 검색 성능을 비교한 분석 결과 제안한 방법은 색상 히스토그램 통계 특성 방법보다 검출율은 0.08이 낮았지만 재현율은 0.16이 높았으며, 형태 템플릿 매칭 분석 방법보다 검출율은 0.05, 재현율은 0.03이 높은 것을 확인하였다. 또한 균일국부 이진패턴 방법에 비해 검출율은 0.02, 재현율은 0.07이 높아 성능이 뛰어남을 확인하였다.

V. 결론 및 향후 연구과제

본 논문에서는 내용 기반 영상 검색 시스템에서 영상의 내용을 기반으로 형태, 색상, 코너점, 레이블링, 조밀도 등을 기반으로 검색 영상의 특징들을 분석하고, 적응적으로 가중치를 부여하여 영상을 검색하는 효율적인 적응적 웨이티드 CBIR 알고리즘을 제안하였다.

알고리즘은 첫 번째로 영상에서 불필요한 정보들을 제거하기 위해 Otsu 방법과 침식, 팽창연산을 이용하여 불필요한 배경과 객체 영상 정보를 다듬는 전처리 과정과 두 번째, 영상에 레이블링하여 크기순으로 재배열하고 면적을 정규화하여 크기 필터링을 하는 면적 판단 과정을 거친 후 배경이 주된 영상과 객체가 판별된 영상으로 분류하는 객체 판별 과정을 지난다. 배경이 주된 영상이면 색상 정보만을 이용하도록 웨이티드를 부여하고 객체가 판별된 영상이면 각각의 레이블링마다 면적, 조밀도, 코너점의 개수, 농담값, 색상값의 특성 정보를 추출하여 특징 벡터를 구성하였다. 그리고 세 번째로 추출한 특징 벡터를 기반으로 객체의 면적과 코너점의 개수와 조밀도를 분석하여 각각의 특징 벡터 값에 맞춰서 적응적으로 웨이티드를 부여하며, 색상 화소값은 인덱스 데이터 행렬의 칼라맵의 분포도를 이용하여 적응적으로 웨이티드를 부여한다. 마지막으로 질의 영상에 따라서 적응적으로 부여된 웨이티드에 따라서 유사도 측정을 하고 영상검색을 진행하였다. 그리고 제안한 알고리즘의 성능 평가를 위해 기존 기법들과의 비교를 통해 다음과 같은 결과를 확인하였다.

- 실험 평가를 위하여 표준화된 Wang 데이터베이스를 이용해 제안한 알고리즘의 검색 결과를 시각적으로 판단 및 분석한 결과는 크기 변화, 회전 등의 영상 변형에 강인함을 보임
- 실험 평가인 재현율과 검출율을 이용하여 분석한 결과, 제안한 알고리즘은 색상 정보를 이용한 알고리즘인 Varish의 색상 히스토그램 통계 분석 방법보다 꽃 영상의 경우 재현율이 0.19이 더 높았지만 검출율은 0.08이 더 낮았으며, 배경 영상인 산 영상의 경우에도 재현율이 0.05가 더 높고 검출율은 0.16이 더 낮았다.

실험 결과를 토대로 제안한 알고리즘은 색상 히스토그램 통계 특성을 이용한 방법보다 재현율은 높지만 검출율이 낮은 방법이라는 것을 확인할 수 있었다.

- 형태 정보를 이용한 알고리즘인 Nikkam의 형태 템플릿 매칭 분석 방법보다 버스 영상의 경우 재현율이 0.06이 더 높고 검출율도 0.02가 더 높았으며, 복잡한 영상인 원주민 영상에서도 재현율이 0.08이 더 높고 검출율도 0.12가 더 높다는 것을 확인할 수 있었다. 그러나 단순한 형태를 비교해야하는 공통 영상의 경우 재현율이 0.03이 더 낮았으며 검출율도 0.02가 더 낮은 것을 확인할 수 있었다. 이는 형태 속성을 이용한 알고리즘에 더 적합한 질의 영상을 사용한 경우에도 제안한 알고리즘도 그에 못지않은 실험 결과를 가진다는 것을 알 수 있다.
- 색상과 형태 정보를 동시에 이용한 Song의 HSV색상과 균일국부 이진패턴 방법보다 단순한 영상인 집 영상의 경우 재현율이 0.03이 더 높고 검출율은 0.08이 더 높았으며, 복잡한 영상인 음식 영상의 경우에는 재현율이 0.13이 더 높았지만 검출율은 0.09가 더 낮았다. 그리고 배경이 주된 영상인 산 영상에서는 재현율이 0.03이 더 낮았고 검출율도 0.03이 낮았다. 대체적인 질의 영상의 경우 제안한 알고리즘이 더 좋은 재현율과 검출율을 보였지만 색상이나 패턴의 속성 분석이 주로 필요한 질의 영상의 경우에는 HSV색상과 균일국부이진패턴을 이용한 방법이 더 좋은 재현율과 검출율을 보였다.
- 질의 영상에 따라 객체의 특징인 색상, 조밀도, 면적, 레이블링, 코너점을 분석하여 적응적으로 웨이트드를 부여하여 유사도 측정이 가능

위와 같은 사실을 통해 제안한 알고리즘은 멀티미디어 데이터의 다양한 특성 정보를 추출함으로써 최적의 데이터베이스를 구성하고 검색할 수 있는 시스템으로 활용할 수 있을 것이라 예상된다.

본 연구를 바탕으로 복잡한 영상에서의 검색율이 비교적 취약한 부분을 보완하고 검색 알고리즘을 개선한다면 보다 완벽하게 질의 영상에 따라 적응적으로 영상을 검색할 수 있을 것이다. 또한 멀티미디어 서비스의 발전에 발맞추어 다양한 영상 변형에서도 강인한 영상 검색 방법을 개발해야 할 것이다.

참고 문헌

- [1] K. Hirata and T. Kato, "Query by Visual Example," Proceedings of EDBT conference, pp. 56-71, Mar. 1992.
- [2] V. Castelli and L. D. Bergman, "Image Databases," WILEY press, 2002.
- [3] Y. Rui, T. S. Hang, and S. Fu Chang, "Image Retrieval : Current Technique, Promising Directions, and Open Issues," Journal of Visual Communication and Image Representation, Vol. 10, No. 4, pp. 39-62, Apr. 1999.
- [4] S. Chang, A. Eleftheriadis and R. McChlintock, "Next-generation Content Representation, Creation and Searching for New-media Applications in Education," Proceedings of the IEEE, Vol. 86, No. 5, pp.884-904, May 1998.
- [5] Y. Rui and T. Huang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," Journal of Communication and Image Representation, Vol. 10, pp. 36-62, 1999.
- [6] S. J. Song, "Combining Method of Multiple Image Features for the Content-Based Image Retrieval," Doctorate Thesis of Electronic Engineering Science in Pusan University, 2003.
- [7] K. Wong, K. Cheung and L. Po, "MIRROR: An Interactive Content Based Image Retrieval System," IEEE International Symposium on Circuits and Systems, Vol. 2, pp. 1541-1544, May 2005.
- [8] R. Schettini, G. Ciocca and S. Zuffi, "A Survey of Methods for Colour Image Indexing and Retrieval in Image Databases," Color Imaging Science: Exploiting Digital Media, Wiley, J & Sons Ltd, 2001.
- [9] N. Chang and K. Fu, "A Relational Database System for Images,"

- Technical Report RT-EE, Purdue University, pp. 28-79, 1979.
- [10] G. H. Lyu and H. S. Gwak, "A Study on Efficient Feature-Vector Extraction for Content-Based Image Retrieval System, The KIPS Transactions. Part B, Vol. 13b No. 3, pp. 309-314, June 2006.
 - [11] G. H. Lyu, J. H. Choi and H. S. Gwak, "Content-Based Retrieval Method using Wavelet Transform and Correlation," ICEIC 98, pp.132-153, 1998.
 - [12] M. S. Park, G. H. Lyu, H. S. Gwak, "Content-Based Image Retrieval using Histogram Area Calculation," Journal of The Korean Association of Computer Education, Vol. 6, No. 2, pp. 265-270, 2005.
 - [13] G. H. Lyu, J. M. Park, G. S. Lyu, S. S. Lyu and H. S. Gwak, "Content-Based Image Retrieval using Adaptive Color Histogram," The Journal of KICS, Vol. 30, No. 9, pp. 949-954, 2005.
 - [14] G. H. Lyu, "Histogram Matching Algorithm for Image Retrieval," Doctorate Thesis of Mechanical Engineering Science in Chonbuk University, 2007.
 - [15] Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner and D. Petkovic, "Indexing for Complex Queries on a Query-by-Content Image Database," Proceeding of the 12nd IAPR Conference on Pattern Recognition, Vol. 1, pp. 142-146, Oct. 1994.
 - [16] A. Pentland, R. Picard and S. Sclaroff, "Photobook: Content-based Manipulation of Image Databases," Proceedings SPIE Storage and Retrieval for Image and Video Databases, Vol. 2185, Feb. 1994.
 - [17] J. R. Smith and S. F. Chang, "Querying by Color Regions using the VisualSEEK Content-Based Visual Query System," In Intelligent Multimedia Information Retrieval, MIT Press, 1996.
 - [18] C. H. Kuo, T. C. Chou, N. L. Tsao and Y. H. Lan, "CanFind-semantic Image Indexing and Retrieval System," Proceedings of ISCAS 2003, Vol. 2, pp. 25-28, 2003.

- [19] J. R. Smith and S. F. Chang, "Visually Searching the Web for Content," IEEE Multimedia Magazine, Vol. 4, No. 3, pp. 12–20, July 1997.
- [20] J. K. Wu, "Content-Based Indexing of Multimedia Databases," IEEE Trans. on Knowledge and Data Engineering, Vol. 9, No. 6, pp. 978–989, 1997.
- [21] M. K. Mandal, T. Aboulnaser and S. Panchnathan, "Image Indexing using Moments and Wavelets," IEEE Trans. on Computer Electronics, Vol. 42, No. 3, pp. 557–565, 1996.
- [22] K. Hirata and T. Katp, "Query by Visual Example," Proceeding of EDBT Conference, pp. 56–71, 1992.
- [23] S. K. Chang, C. W. Yan, D. C. Dimitorff and T. Arndt, "An Intelligent Image Database System," IEEE Trans. Software Engineering, Vol. 14, No. 5, pp. 681–688, May 1998.
- [24] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision, Vol. 7, No. 1, pp. 11–31, 1991.
- [25] IBM Almaden Research Center, "Query by Image and Video Content: The QBIC System," IEEE Multimedia, pp. 23–32, 1995.
- [26] Elif Albuz, "Scalable Image Indexing and Retriever using Wavelet," Technical Report, University of Delaware, 1998.
- [27] P. P. Ohanian and R. C. Dubes, "Performance Evaluation for four Classes of Textural Features," Pattern Recognition, Vol. 25, No. 8, pp. 819–833, 1992.
- [28] R. Bajcsy, S. W. Lee and A. Leonardis, "Color Image Segmentation and Color Constancy," Proc. of SPIE Conference, Santa Clara, CA, USA, Feb. 1990.
- [29] I. Biederman, "Recognition-by-Components: A Theory of Human Image Understanding," Psychological Review, pp. 115–147, 1994.

- [30] M. A. Stricker and M. Orengo, "Similarity of Color Images," Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 2420, pp. 381–392, 1995.
- [31] Y. Deng, D. Mukherjee and B. S. Manjunath, "NETRA-V: Toward and Object-based Video Representation," Proceedings of SPIE Storage and Retrieval for Image and Video Databases, pp. 202–215, 1998.
- [32] V. Gudivada and V. Raghavan, "Content-Based Image Retrieval Systems," IEEE Computers, Vol. 28, No. 9, pp. 18–22, 1995.
- [33] R. Jain and A. Gupta, "Computer Vision and Visual Information Retrieval," IEEE Computer Society, 1996.
- [34] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic and W. Equiz, "Efficient and Effective Querying by Image Content," Journal of Intelligent Information System," pp. 231–262, July 1994.
- [35] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC Project: Querying Image by Content using Color, Texture and Shape," Proceedings of SPIE Storage and Retrieval for Image and Video Databases, pp. 173–187, Feb. 1993.
- [36] A. Del Bimbo and P. Pala, "Visual Image Retrieval by Elastic Matching of User Sketches," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 2, Feb. 1997.
- [37] C. E. Jacobs, A. Finkelstein and D. H. Salesin, "Fast Multiresolution Image Query," Proceedings of ACM SIGGRAPH, New York, 1995.
- [38] T. Kato, T. Kurita, N. Otsu and K. Hirata, "A Sketch Retrieval Method for Full Color Image Database," Proceedings of 11th ICPR, IEEE, pp.530–533, 1992.
- [39] A. Vailaya, Y. Zhong and A. K. Jain, "A Hierarchical System for Efficient

- Image Retrieval,” Proceedings of 13th ICPR, pp.356–359, IEEE, 1996.
- [40] T. S. Lai, “CHROMA: A Photographic Image Retrieval System,” Doctorate Thesis in University of Sunderland, Sunderland, United Kingdom, Jan. 2000.
- [41] Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, T. Humphrey, R. Jain and C. Shu, “The Virage Image Search Engine: An Open Framework for Image Management,” Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 2670, pp. 76–87, 1996.
- [42] J. Dowe, “Content-Based Retrieval in Multimedia Imaging,” Proceeding of SPIE Storage and Retrieval for Image and Video Databases, Vol. 1908, pp. 164–167, 1993.
- [43] N. S. Chang and K. S. Fu, “A Relational Database System for Images,” Technical Report, Purdue University, May 1979.
- [44] N. S. Chang and K. S. Fu, “Query-by Pictorial Example,” IEEE Reans. on Software Engineering, Vol. 6, No. 6, pp. 519–524, Nov. 1980.
- [45] S. K. Chang, “Pictorial Database Systems,” IEEE Trans. on Computer, Vol. 14, No. 11, pp.13–21, Nov. 1981.
- [46] M. Safar, C. Chahabi and X. Sun, “Image Retrieval by Shape: A Comparative Study,” IEEE International Conference on Multimedia and Expol, Vol. 1, pp. 141–144, 2000.
- [47] I. Biederman, “Recognition by Components: A Theory of Human Image Understanding,” Psychological Review, Vol. 94, No. 2, pp. 115–147, Apr. 1987.
- [48] M. Stricker and M. Orengo, “Similarity of Color Images,” Proceedings SPIE Storage and Retrieval Image and Video Databases, Vol. 2420, pp. 381–392, Feb. 1996.
- [49] C. Faloutsos, W. Equitz, M. Flickner, W. Nkblack, D. Petkovic and R.

- Barber, "Efficient and Effective Querying by Image Content," *Journal of Intelligent Information System*, Vol. 3, No. 3, pp. 231–262, July 1994.
- [50] B. V. Funt and G. D. Finlayson, "Color Constant Color Indexing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 5, pp. 522–529, 1995.
- [51] V. N. Gudivada and J. V. Raghavan, "Special Issue on Content-Based Image Retrieval System," *IEEE Computer Magazine*, Vol. 28, No. 9, pp. 18–22, Sep. 1995.
- [52] A. D. Narasimhalu, "Special Section on Content-Based Retrieval," *ACM Multimedia Systems*, Vol. 3, No. 1, pp. 1–41, Feb. 1995.
- [53] Y. Rui, T. S. Hang and S. Fu Chang, "Image Retrieval: Current Technique, Promising Directions and Open Issues," *Journal of Visual Communication and Image Representation*, Vol. 10, No. 4, pp. 39–62, Apr. 1999.
- [54] M. K. Mandal, T. Aboulnasr and S. Panchanathan, "Image Indexing using Moments and Wavelets," *IEEE Trans. on Computer Electronics*, Vol. 42, No. 3, pp. 557–565, 1996.
- [55] A. K. Jain and A. Vailaya, "Shape-Based Retrieval: A Case Study with Trademark Image Database," *Pattern Recognition*, Vol. 31, No. 9, pp. 1369–1390, 1998.
- [56] H. Wan, H. Hu and Z. Shi, "Image Similarity Measurement using Max Weighted Bipartite Matching," *Proceedings of ICII 2001*, Vol. 13, pp. 811–820, 2002.
- [57] G. D. Guo, A. K. Jain, W. Y. Ma and H. J. Zhang, "Learning Similarity Measure for Natural Image Retrieval with Relevance Feedback," *IEEE Reans. Neural Networks*, Vol. 13, pp. 811–820, 2002.
- [58] J. Smith and S. F. Chang, "Single Color Extraction and Image Query,"

- Proceedings of IEEE International Conference on Image Processing, Vol. 3, pp. 528–531, Oct. 1995.
- [59] J. Z. Wang, “Wavelet-Based Image Indexing Techniques with Partial Sketch Retrieval Capability,” Proceedings of the 4th Forum on Research and Technology Advances in Digital Libraries, pp. 13–24, May 1997.
- [60] Y. Tui, T. S. Hang and S. Fu Chang, “Image Retrieval from a Relational Database of Images,” IEEE Computer, Vol. 28, No. 9, 1995.
- [61] E. J. Ryu, Y. J. Song, W. B. Park and J. H. Ahn, “Content Base Image Retrieval using Color Feature of Contral Region and Comparing Bin,” Journal of KIPS Trans. Vol. 11b, No. 5, pp. 581–586, Aug. 2004.
- [62] M. J. Swain and D. H. Ballard, “Color Indexing,” International Journal of Computer Vision, Vol. 7, No. 1, pp. 11–31, 1991.
- [63] M. A. Stricker and M. J. Swain, “The Capacity of Color Histogram Indexing,” Proceedings of the Conference on Computer Vision and Pattern Recognition, pp. 704–708, 1994.
- [64] G. Pass, R. Zabih and J. Miller, “Comparing Images using Color Coherence Vectors,” ACM Conference on Multimedia, Boston, Massachusetts, Nov. 1996.
- [65] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu and R. Zabih, “Image Indexing using Color Correlograms,” IEEE Computer Vision and Pattern Recognition Conference, San Juan, Puerto Rico, June 1997.
- [66] J. Huang and R. Zabih, “Combining Color and Spatial Information for Content-Based Image Retrieval,” European Conference on Digital Libraries, Sep. 1998.
- [67] M. A. Stricker and A. Dimai, “Color Indexing with Weak Spatial Constraints,” Storage and Retrieval for Still Image and Video Databases, Vol. 2670, pp. 29–40, 1996.

- [68] Y. Gong, H. C. Chua and X. Guo, "Image Indexing and Retrieval Based on Color Histograms," *Multimedia Tools and Application*, pp. 133–156, 1996.
- [69] I. J. Cox, M. L. Miller, S. M. Omohundro and P. N. Yianilos, "PicHunter: Bayesian Relevance Feedback for Image Retrieval," *Proceedings of International Conference on Pattern Recognition*, Vienna, Austria, pp. 361–369, Aug. 1996.
- [70] I. J. Cox, M. L. Miller, S. M. Omogundro and P. N. Yianilos, "Target Testing and the PicHunter Bayesian Multimedia Retrieval System," *Proceedings of 3rd Forum on Research & Technology Advances in Digital Libraries*, IEEE, pp. 66–75, 1996.
- [71] 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, June 1978.
- [72] F. Liu and R. W. Picard, "Periodicity, Directionality and Randomness: Word Features for Image Modeling and Retrieval," *IEEE Trans. on PAMI*, Vol. 18, No. 7, pp. 722–733, July 1996.
- [73] 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, Aug. 1996.
- [74] L. M. Kaplan, R. Murenzi and K. R. Namuduri, "Fast Texture Database Retrieval using Extended Fractal Features," *Proceedings of SPIE, Storage and Retrieval for Image and Video Databases*, Vol. 3312, pp. 162–173, Jan. 1998.
- [75] M. Safar, C. Shahabi and X. Sun, "Image Retrieval by shape: a Comparative Study, ICME 2000, Vol. 1 pp. 141–144, 2000.
- [76] A. K. Jain and A. Vailaya, "Image Retrieval using Color and Shape,"

- Pattern Recognition, Vol. 29, No. 8, pp. 1233–1244, Aug. 1996.
- [77] J. Canny, “A Computational Approach to Edge Detection,” IEEE Tans. on Pattern Recognition and Machine Intelligence, Vol. 8, pp. 679–698, Nov. 1986.
- [78] 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.
- [79] C. Harris and M. Stephens, “A Combined Corner and Edge Detector,” Alvey Vision Conference, pp.147–151, 1988.
- [80] R. Deriche and G. Giraudon, “A Computational Approach for Corner and Vertex Detection,” Journal of Computer Vision, Vol. 10, No. 2, pp. 101–124, 1993.
- [81] C. Schmid, R. Mohr and C. Bauckhage, “Comparing and Evaluating Interest Point,” International Conference on Computer Vision, pp. 230–235, 1988.
- [82] C. Schmid and R. Mohr, “Matching by Local Invariants,” Technical Report N2644 in INDIA, Aug. 1995.
- [83] D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos, “A Novel Vector-Based Approach to Color Image Retrieval using a Vector Angular-Based Distance Measure,” Computer Vision and Image Understanding, Vol. 75, No. 1, pp. 46–58, Aug. 1999.
- [84] N. Otsu, “A Threshold Selection Method from Gray-level Histogram,” IEEE Trans. on Systems, Man and Cybernetics, Vol. 9, No. 1, pp. 62–66, Jan. 1979.
- [85] C. Harris and M. Stephens, “A Combined Corner and Edge Detector,” Proceedings of 4th Alvey Vision Conference, pp. 147–151, Aug. 1988.
- [86] V. N. Gudivada and V. V. Raghavan, “Content-Based Image Retrieval

- Systems,” IEEE Computer Society, Vol. 28, No. 9, pp. 18–22, Sep. 1995.
- [87] S. G. Shaila and A. Vadivel, “Indexing and Encoding based Image Feature Representation with Bin Overlapped Similarity Measure for CBIR Applications,” Journal of Visual Communication and Image Representation, Vol. 36, pp. 40–55, Apr. 2016.
- [88] J. R. Smith, “Image Retrieval Evaluation,” Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries, June 1998.
- [89] N. Varish and A. Kumar Pal, “Content Based Image Retrieval using Statistical Features of Color Histogram,” International Conference on Signal Processing, Communication and Networking, Mar. 2015.
- [90] P. S. Nikkam, N. P. Hegde and B. E. Reddy, “Decomposition-based Shape Template Matching for CBIR System,” International Conference on Computational Intelligence and Computing Research, pp. 1–5, Dec. 2015.
- [91] J. H. Song, “Content-based Image Retrieval using HSV Color and Uniform Local Binary Patterns,” Journal of KIIT, Vol. 12, No. 6, pp. 169–174, June 2014.