creative commons

COMMONS DEED

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

dCollection

[UCI]I804:24011-200000265101
조선대학교
CHOSUN UNIVERSITY

2015년 8월
석사학위 논문

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

2015년 8월 석사학위논문

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

김 국 정

조선대학교
CHOSUN UNIVERSITY

2015년 8월

석사학위 논문

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

Multi-object Image Retrieval based on Normalization Labeling Invariant Moments

2015년 8월 25일

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

지도교수 박 종 안

이 논문을 공학석사학위 신청 논문으로 제출함

2015년 4월

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

김국정의 석사학위논문을 인준함

위원장 조선대학교 교수 김 판 구 (인)

위 원 조선대학교 교수 안 영 은 (인)

위 원 조선대학교 교수 박 종 안 (인)

2015년 5월

조선대학교 대학원

조선대학교
CHOSUN UNIVERSITY

목 차

Ⅰ. 서론 ····· 1
1. 연구 배경 및 목적 ····· 1
2. 연구 내용 ····· 3
3. 논문 구성 ····· 3

Ⅱ. 관련연구 ····· 4

Ⅲ. 제안하는 영상검색 방법 ····· 13
1. 전처리 ····· 13
1) Otsu 방법 및 워터쉐드 변환을 이용한 이진화 ····· 15
2. 레이블링 ····· 19
3. 불변 모멘트 특성벡터 추출 ····· 21
1) 레이블링 영역 형태 정보 추출 ····· 21
2) 불변 모멘트 ····· 21
3) 영상 변환에 강인한 정규화 ····· 25
4) 특성 정보 저장 ····· 26
4. 유사도 계산 ····· 26
5. 성능평가 척도 ····· 27

Ⅳ. 실험 및 분석 ····· 29
1. 실험 환경 ····· 29
2. 시뮬레이션 ····· 29
3. 분석 및 고찰 ····· 36

조선대학교
CHOSUN UNIVERSITY

Ⅴ. 결론 ····· 42

【참고문헌】 ····· 43

조선대학교
CHOSUN UNIVERSITY

그 림 목 차

그림 2.1 내용기반 영상검색 시스템의 구조 ····· 4
그림 3.1 제안한 알고리즘의 구성도 ····· 13
그림 3.2 전처리 과정 ····· 14
그림 3.3 RGB 이미지 ····· 14
그림 3.4 그레이스케일 이미지 ····· 15
그림 3.5 Otsu 방법 및 워터쉐드 변환을 이용한 이진화 과정 ····· 19
그림 3.6 4-연결성과 8-연결성 ····· 20
그림 3.7 4-연결성과 8-연결성을 이용한 레이블링 ····· 21
그림 3.8 상위 10개의 대표 레이블링과 그에 해당하는 모멘트를 이용한 특성벡터 추출 ····· 25
그림 3.9 영상변환에 강인한 정규화 ····· 26
그림 4.1 실험 영상 ····· 30
그림 4.2 질의 영상 20.jpg ····· 31
그림 4.3 2진화 영상 ····· 31
그림 4.4 질의 영상의 특징벡터 추출 ····· 32
그림 4.5 질의 영상의 정규화벡터 추출 ····· 32
그림 4.6 특징벡터를 이용한 영상 번호별 유사도 측정 결과 ····· 33
그림 4.7 특징벡터를 이용한 유사도 순위별 측정 결과 ····· 34
그림 4.8 유사도 값에 의한 영상 검색 결과 ····· 35
그림 4.9 질의 영상 Q1~Q4 ····· 37
그림 4.10 Q1, Q2의 기존 알고리즘과 제안한 알고리즘의 검색 결과 ····· 38
그림 4.11 Q3, Q4의 기존 알고리즘과 제안한 알고리즘의 검색 결과 ····· 39

조선대학교
CHOSUN UNIVERSITY

표 목 차

표 3.1 데이터베이스 테이블 구성 ····· 26
표 4.1 영상의 오차율 ····· 35
표 4.2 기존 알고리즘과 제안한 알고리즘의 성능 분석 ····· 40
표 4.2 단일 객체 영상들의 질의 영상에 따른 실험 결과 ····· 41
표 4.4 다중 객체 영상들의 질의 영상에 따른 실험 결과 ····· 41

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

Multi-object Image Retrieval based on Normalization Labeling Invariant Moments

Kim Gukjeong
Advisor : Prof. Park Jongan, Ph.D.
Dept, of Software Convergence Engineering,
Graduate School of Industry Technology Convergence,
Chosun University

In CBIR with multi-objects, It shows an interest in the image retrieval method based on object features. In this paper, we have proposed an image retrieval method based on normalized labelling features of multi-objects. First, it extracts, normalizes and rearranges the object-labels which present image shape information. Secondly, it measures image similarity using moment features invariant to transition, rotation and scaling of the objects. The simulation results using Multi-object database showed that our proposed method is superior to the edge histogram method and the wavelet shape extraction method. The experimental results demonstrate the high retrieval efficiency i.e Recall 0.23, Precision 0.85. Consequently, our retrieval system shows better performance than the conventional system that cannot express the shape of objects exactly.

조선대학교
CHOSUN UNIVERSITY

Ⅰ. 서론

1. 연구 배경 및 목적

인터넷과 모바일 단말기의 발달로 웹 사이트나 블로그, SNS 등을 통해 많은 사람들이 디지털 영상을 쉽게 생성하고 저장할 수 있게 되었다. 그러나 그 영상량이 방대하여 관리와 검색은 점점 어려워지고 있다. 이에 따라 영상을 입력으로 이용하는 다양한 응용 분야들이 최근 증가하고 있다. 특히 영상 검색은 최근 상당히 주목을 받고 있고 많은 연구가 이루어지고 있는 분야이다.

영상 검색을 위해서는 데이터베이스에 저장되는 각 영상에 대한 분류 및 주제어에 대한 결정 과정이 필수적이다. 초기 영상 검색에서는 수동으로 인덱스를 지정하는 방식을 이용하였으나 입력 작업의 비효율성과 작업자의 주관에 따라 입력 내용의 결정되는 등의 많은 어려움을 내포하고 있었다. 이에 최근에는 영상을 분석함으로써 칼라 분포나 질감 등과 같은 영상의 특징 정보를 자동으로 추출하여 검색에 이용하는 내용기반 영상 검색(content-based image retrieval, CBIR)이 등장하게 되었다[1-7].

내용 기반 영상 검색은 컬러(color), 형태(shape), 그리고 질감(texture) 등의 내용 표현 요소들을 통하여 얻어진 특징 값을 사용하여 영상들 간의 유사도를 계산하고 검색을 수행한다. 그러므로 효율적인 영상 검색을 위해서는 무엇보다도 각 영상을 대표하는 표현 요소들의 효과적인 특징 추출이 중요하다. 영상의 표현 요소들 중에서 컬러 특징(color feature)은 영상을 표현하는 주된 특징으로 내용 기반 영상 검색의 초기부터 사용되어 온 방법이다. 컬러 특징은 영상 내의 조명 변화나 관측 위치, 크기 변화 등에 어느 정도 무관하게 적용 가능하다는 장점을 가지고 있으며, 방법이 간단하여 내용 기반 영상 검색에서 가장 많이 사용되는 방법 중의 하나이다. 컬러 특징을 추출하는 대표적인 방법이 컬러 히스토그램(color histogram)을 이용하는 것이다[8]. 컬러 히스토그램을 특징 정보로 하면 컬러의 수 만큼에 해당하는 크기의 공간이 필요하므로 보통 유사한 컬러 별로 군집하는 컬러 집합을 사용하게 된다. 초기에는 히스토그램을 균등한 비율로 나누어 히스토그램 구간(bin)을 결정하는 방식으로 컬러를 군집했지만, 효율적인 검색 시스템을 위해서 유사한 컬러끼리 군집하는 방법들이 제안되었다. 단, 컬러

조선대학교
CHOSUN UNIVERSITY

히스토그램이나 컬러 집합을 이용하면 영상 내 객체들의 배치 즉 컬러 분포의 공간적인 정보를 전혀 이용할 수 없다는 단점이 있다.

질감은 영상 내부의 동일한 패턴을 가지는 영역에 대한 특징으로 하나의 컬러 정보로만은 표현할 수 없다. 질감은 영상 내의 객체의 표면이 가지는 영상에게는 매우 중요한 특징이 된다. 특히 항공사진을 분석하는 경우에 질감 정보가 많이 사용되어 왔다. Gabor 필터, wavelet 변환 등을 통해서 질감 특징을 분석하는 방법이 연구되고 있는데[9], 이러한 방법은 인공적인 질감 즉 건물의 벽돌 무늬나 의류에서 나타나는 무늬 등의 분석에 큰 효과를 보이고 있다.

형태(shape) 특징값은 영상의 회전, 스케일 변화 등에 매우 민감하며 배경과 객체를 자동 또는 수동으로 분리해내는 알고리즘이 필요하다는 어려움을 가지고 있다. 이러한 점을 극복하기 위하여 Jain과 Vailaya[10]는 별도의 객체 분리 과정 없이 영상의 에지 성분들에 대하여 에지 히스토그램을 구성하고 회전과 스케일 변화 등을 고려하기 위해서 히스토그램 정규화나 히스토그램 smoothing 방법 등을 적용하였다. 그러나, 이 방법은 영상의 스케일 변화에는 강인하지만, 히스토그램 smoothing 에 의해 다른 영상도 같다고 판단할 문제점을 가지고 있다. 이밖에 사용자가 스케치한 영상을 가지고 2차 spline 곡선을 구성하여 검색하는 방법[11] 등이 제안 되었지만 이러한 방법들은 이론 자체가 어렵고, view point 변화와 객체 영역을 미리 설정해주어야 하는 단점을 가지고 있다.

영상의 내용을 파악하기 위해서는 무엇보다도 영상 내의 객체(object)의 형태(shape)에 관심을 가질 필요가 있다. 많은 영상에서는 비교적 크기가 크고 주변 배경과는 대비되는 특징을 가지며 또한 경계가 뚜렷하게 구별되는 객체를 가지고 있다. 형태 분석 및 표현 방법은 크게 형태의 외곽선 정보를 이용한 것과 영역 내부를 이용한 것으로 분류가 된다. 외곽선 정보를 이용하는 방법은 영상에서 외곽선을 추출하여 이의 분석을 통해 그 변화를 특징 정보를 삼는 것이다. 역 내부 정보를 이용하는 것은 영상 전체의 화소의 분포를 분석하는 것으로 영상에 대한 모멘트를 구하는 방법과 같이 형태의 분포 정도를 특징으로 삼는 것이다[12]. 객체를 추출하는 과정은 객체 특징의 다양성을 효과적으로 반영하며 진행되어야 한다. 이에 따라 영상에 포함된 관심 객체를 추출한 후, 영상의 핵심정보에 대한 효과적인 형태특성을 추출하여 영상을 인식하고자 하는 여러 연구결과가 발표되고 있다[13-16].

조선대학교
CHOSUN UNIVERSITY

위에서 설명한 여러 가지 특징들을 추출하여 효과적으로 사용함으로써 내용 기반 영상 검색의 효율성을 높일 수 있다. 본 논문에서는 다양한 객체를 갖는 영상으로부터 영상에 내포된 핵심적인 영상 정보를 표현할 수 있는 의미 있는 관심 객체를 정규화 레이블링을 통해 추출하고 객체의 위치 이동이나 회전, 크기 변화 등과 같은 각종 변환에 강인한 불변 모멘트 특징 값을 이용하여 유사도를 측정하는 영상검색 알고리즘을 제안하고자 한다.

2. 연구 내용

본 논문에서는 다중 객체 영상에서의 정규화 레이블링 모멘트기반 형태검색 알고리즘을 제안한다. 제안한 형태 속성 기반 검색 알고리즘은 크게 4단계로 구성된다. 즉,

제 1단계는 전처리 과정으로 잡음제거 및 수축과 팽창에 의한 홀잡음제거
제 2단계는 레이블링 이용 다중 객체 추출 및 불변 모멘트 특성벡터 추출
제 3단계는 특성벡터 정규화 처리
제 4단계는 유사도 측정 단계로 이루어진다.

컴퓨터 시뮬레이션에서 제안한 알고리즘의 유사도 측정은 Wang 데이터베이스를 이용하여 기존의 형태검색 기법과 비교하여 정확도 성능을 나타낸다.

3. 논문 구성

이 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 연구를 고찰하고, 3장에서는 다중 객체를 갖는 영상의 정규화 레이블링 모멘트 특성을 추출하여 형태를 검색하는 새로운 알고리즘을 제안한다. 그리고 4장에서는 제안된 알고리즘에 대한 시뮬레이션을 수행하고 그 결과를 분석한다. 그리고 5장에서 결론을 짓는다.

조선대학교
CHOSUN UNIVERSITY

Ⅱ. 관련연구

1. 내용기반 영상 검색

내용기반 영상검색은 영상을 대표할 수 있는 특징값을 자동으로 추출하여, 이를 이용하여 질의자가 원하는 영상을 검색하는 것을 말한다. 이러한 내용기반 영상검색 시스템은 일반적으로 영상의 입력, 질의, 검색 세 가지의 과정으로 나뉘며, 그 구조는 아래와 같다[17].

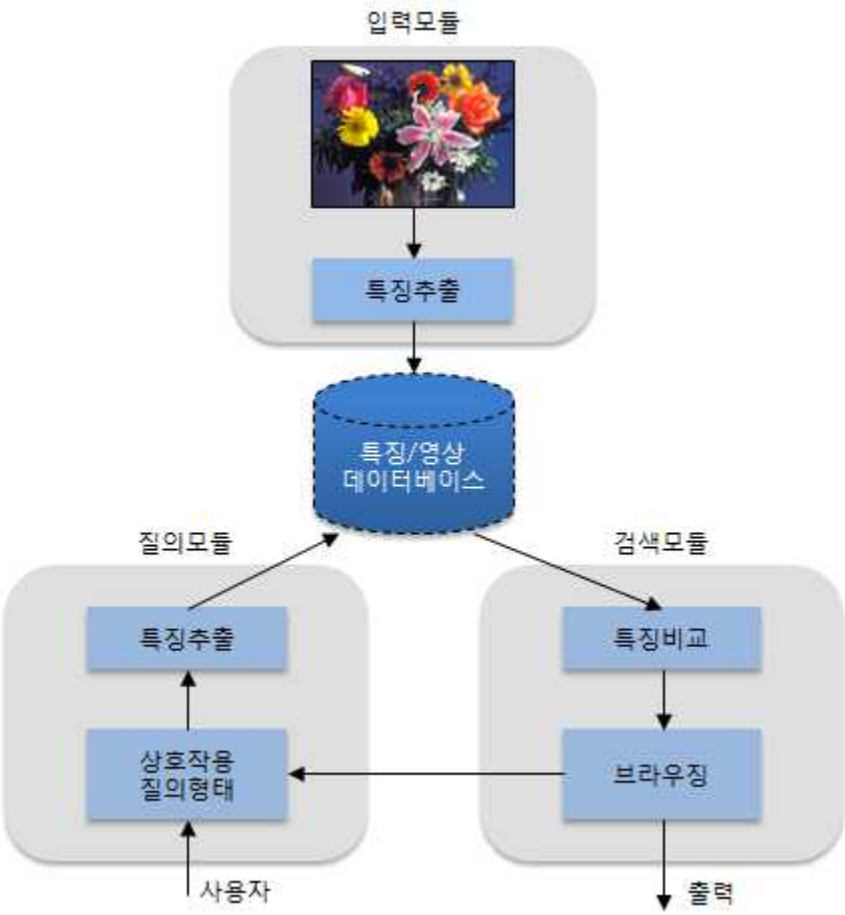


그림 2.1 내용기반 영상검색 시스템의 구조

조선대학교
CHOSUN UNIVERSITY

영상의 입력과정은 영상이 입력되었을 때에 영상을 대표할 수 있는 특징을 추출하여 영상 데이터베이스 내에 저장하는 단계이고, 영상의 질의 과정은 사용자, 질의자가 찾고자 하는 영상의 입력이 주어지면 특징을 추출하는 과정이고, 질의 영상의 특징이 추출되어 검색과정에서 미리 구성되어진 영상 데이터베이스내의 특징과 질의 영상의 특징을 비교하여 가장 근사한 영상을 출력하는 단계가 검색 과정이다.

내용 기반 검색을 효율적으로 수행하기 위해서는 의미 있는 특징 추출이 매우 중요하다. 의미 있는 특징들은 크게 키워드나 주석과 같은 의미와 관련 있는 특징들과, 컬러(color), 질감(texture) 및 형태(shape)등의 시각적인 특징들로 나눌 수 있다. 이와 같은 특징들은 인식대상을 유일하고, 명확하게 표현할 수 있고, 회전(rotation), 이동(translation), 크기변화(scaling)와 같은 기하학적인 변환에 대해서 불변하는 성질을 가져야만 한다. 내용기반 영상검색에서는 위에서 언급한 특징의 조건들을 만족하기 위해서 색상, 형태, 질감을 단독 혹은 병합하여 특징 값으로 이용한다.

1992년 kato[18]에 의하여 내용기반 영상 검색이라는 용어가 현재까지 널리 사용되고 있다. 내용기반 영상 검색은 영상의 내용을 분석하여 시각적 특징 정보를 자동으로 추출하고 색인(Index)을 추가하여 저장하는 과정과 데이터베이스에서 사용자가 원하는 영상을 자동으로 검색하는 과정을 포함한다. 원 영상으로부터 정확한 특징 정보를 추출하기가 어려운 단점이 있으나. 특징 정보를 자동으로 추출하여 색인 과정을 통한 데이터베이스구축에 필요한 시간 및 인력의 소모를 줄일 수 있다는 장점이 있다.

내용기반 영상 검색은 영상으로부터 컬러, 질감, 형태 등과 같은 특징을 이용한다. 컬러를 이용한 방법에는 1991년 swain과 ballard[8]이 제안한 컬러 히스토그램을 이용하는 방법이 있다. 이 방법은 각 컬러에 대해 영상 픽셀 빈도를 나타내는 방법으로 질의 영상과 데이터베이스 영상의 히스토그램을 비교하여 공유되는 최솟값을 구하는 히스토그램 인터섹션 기법을 사용하였다. 이 방법은 현재도 많은 내용기반 영상 검색 시스템에서 사용되고 있다.

김성희(2004)는 이미지 및 동영상 내용 기반 검색 시스템의 개념과 유형을 조사한 후, 상용중인 내용 기반 검색 시스템 5개를 선정하여 성능을 분석하였

다. 현재 상업용으로 이용 가능한 시스템인 QBIC, Excalibur Visual Retrieval

조선대학교
CHOSUN UNIVERSITY

Ware, ImageFinder & VideoFinder, VIR Image Engine, IMatch의 일반적인 시스템 특성, 제공되는 검색 기법, 운영 체제 등을 분석하였다[19].

김현술(2001)은 영상 데이터베이스에서 자동으로 영상을 검색할 수 있는 내용 기반 영상 검색 방법을 제안하였다. 내용 기반 영상 검색 시스템에서는 주로 영상의 특징으로 색상, 형태 및 질감 등의 원시적인 특징들이 사용되고 있으며 이 연구에서도 이러한 특징들을 검색 정보로 사용하여 검색을 수행하였다[20].

이지민(2010)은 영상 내의 특정 객체의 위치와 형태에 대한 정보를 가지는 interest point로서 에지영상에서 라인을 추출하여 얻어낸 코너점을 이용한 영상 검색 알고리즘으로 재배열된 코너 패치 속성의 히스토그램 빈도수를 사용한 검색기법을 제안하였다[21].

황춘화(2011)는 영상을 색상히스토그램 방법으로 색상의 분포가 비슷한 영상들을 질감 특징인 EHD(Edge Histogram Descriptor)과 이웃 화소들의 질감을 고려한 GLCM(Gray Level Co-occurence)방법을 이용하여 색상 및 질감특징들에 가중치를 주어서 특징 값들을 비교함으로써 질의 영상과 비슷한 영상들을 색인 및 검색함으로써 가장 비슷한 영상들을 보여주는 영상검색 방법을 제안하였다[22].

검색을 위한 내용기반 영상 질의의 목적은 사용자의 질의 영상과 유사도 기준을 만족하는 영상을 데이터베이스로부터 효율적으로 검색하는 것이다.

영상의 내용을 기반으로 하는 검색은 전통적인 데이터베이스 검색과 비교하여 몇 가지 중요한 차이점을 갖는다.

첫째, 데이터베이스 검색은 정확성(exact)을 목표로 하지만, 내용기반 검색은 유사성(similarity)에 기반을 둔 검색이다. 따라서 사용자의 검색 절차를 매우 단순화 할 수 있다.

둘째, 사용자와 시스템과의 상호작용을 통해 다양한 검색 가능성을 제공한다.

셋째, 시각적인 질의를 필요로 하는 멀티미디어 시나리오와 같은 응용 단계에서 사용될 수 있다.

넷째, 유사도(similarity value)에 따른 후보 영상을 최종 사용자가 선택할 수 있도록 영상들을 시각적으로 표현과 출력할 수 있다.

이러한 영상 데이터베이스 특징 중 유사성 검색은 주어진 표본과의 유사한 정보를 나타내는 유사성 척도(similarity measure)에 의해 검색하는 것으로 사용자의 입장에서 보다 자연스럽게 데이터베이스와 대화할 수 있는 기능을 제공한다.

조선대학교
CHOSUN UNIVERSITY

유사성 척도는 유사성 검색의 성능과 밀접한 관련이 있으며 문제 영역의 특성에 따라 다양한 척도를 가질 수 있다. 그리고 유사성 척도를 이용하여 논리적으로 유사한 영상들을 함께 모아 저장함으로써 검색 시 탐색 시간을 줄이고 시스템 효율을 증대시킬 수 있다. 뿐만 아니라, 유사도를 기준으로 퍼지 질의가 가능해진다. 내용기반 영상 검색은 사용자의 질의 영상과 비교하여 유사도가 높은 영상을 검색하는 유사도 검색 방법을 사용하며, 이미지 검색 엔진, 지리 정보 시스템, 의료 분야 등의 다양한 분야에 응용될 수 있다.

2. 영상의 내용 표현 요소

1) 색상 정보

영상의 색상정보를 이용한 영상검색 기법은 영상내의 각 화소의 색상값을 특징값으로 하며, Swain과 Ballard가 제안한 색상히스토그램 정합을 기반으로 하는 영상 검색 기법이 소개된 후 색상을 이용한 내용기반 영상검색에서 색상 히스토그램은 가장 일반적으로 색상정보를 대표하는 특징값이 되었다. 이런 색상 히스토그램과 같은 색상정보를 이용한 방법은 칼라 영상의 경우 그레이 영상에 비해 정보량이 많음을 이용하여 영상의 통계적 특성을 이용한 3차원 히스토그램을 생성한 후에 히스토그램 공간상에서 질의영상과 데이터베이스 영상의 정합(Matching)을 시도하는 인덱싱 알고리즘이다. 이 알고리즘은 간단하고, 히스토그램의 특성으로 이동(translation)이나 크기변화(scaling)에 비교적 강건하고, 물체의 변형이나, 일부의 절취에 대해서 비교적 강건한 장점을 가지고 있으나, 특징값이 영상내의 각 객체에 대한 세부정보를 표현할 수 없고, 빛의 밝기와 영상내의 물체의 크기에 민감하여, 전혀 다른 영상일지라도 같은 색상분포를 가질 수 있는 단점을 가진다. 때문에 색상과 더불어 질감, 형태와 같은 특징을 병합하여 이용하거나, 영상에서 영역별로 색상 히스토그램을 이용하는 방법과 같이 색상 히스토그램과 함께 다른 추가정보를 이용하여 사용한다.

색상 정보를 이용한 대표적인 연구로는 색상 모델의 선택이 어떻게 히스토그램에 매칭 되는가를 분석한 연구[23], 이미지의 지역적 색상 속성을 특성화하기 위해 컬러히스토그램을 집합을 제안한 연구[24], VisualSEEK 시스템에서 컬러 이미지의 위치영역에 사용한 색상 집합에 대한 연구[25]등이 있다.

조선대학교
CHOSUN UNIVERSITY

2) 질감 정보

질감(Texture)은 화소 값들의 공간적인 배열로 표현되어지는 속성으로 부드러움, 거침, 규칙적임 등으로 묘사할 수 있다. 이러한 질감으로 이루어진 영상을 질감 영상이라 하며, 모피, 곡물류, 위성사진과 의료 영상 등이 여기에 속한다. 최근 이러한 질감 영상들을 분석하기 위하여 질감 영상 분할이 많이 연구되어지고 있다. 질감 영상 분할이란 질감 영상에서 비슷한 질감 별로 분할하는 것을 말한다. 질감영상 분할은 질감에 대한 모든 속성을 정확히 정의할 수 없기 때문에 분할하는데 많은 어려움이 있다. 이러한 이유로 질감을 묘사하기 위한 많은 특징들이 제안되었고, 여러 가지 질감 영상을 분할하기 위한 방법들이 제안되어 왔다[26-30].

질감 영상을 분할하기 위해서는 먼저 영상에서 질감을 구분할 수 있는 특징을 추출하여야 한다. 질감의 특징을 추출하는 방법으로는 크게 통계적인 방법, 구조적인 방법 그리고 스펙트럼을 이용한 방법 등이 있다[26][30]. 통계적인 방법은 일정 영역 내 명암도의 통계적인 특징을 이용하는 방법으로 질감을 부드러움, 거침, 알갱이 모양 등으로 표현한다. 이 방법은 계산이 단순하다는 장점이 있지만, 일정 크기의 영역만 묘사할 수 있다는 문제점이 있다. 구조적인 방법은 화소 값들의 규칙적인 배열 상태를 다루는 방법으로 가까운 이웃 화소 간의 공간적인 상호작용만을 표현할 수 있고 전반적인 특징을 알지 못한다는 단점이 있다. 스펙트럼을 이용한 방법은 주파수 영역에서 질감의 특징을 추출하는 방법으로 푸리에 스펙트럼의 성질을 이용하여 영상의 전반적인 주기성을 검출하는데 주로 사용된다. 그러나 이 방법은 영상의 국부적인 묘사를 할 수 없다는 단점이 있다. 최근 필터링을 이용하여 질감의 특징을 추출하는 방법들이 제안되고 있는데, 이는 질감마다 주파수 성분을 가지는 성질을 이용한 것으로, 질감 영상을 주파수 대역 별로 분할하여 질감의 특징을 추출하는 방법이다[31][32]. 이 중 웨이브렛 변환을 이용한 방법이 가장 많이 사용된다[33-36].

웨이브렛 변환은 정규직교기저(orthonormal basis)의 저대역 통과필터와 고대역 통과필터로 구성된 필터 뱅크(filter bank)를 이용하여 영상을 다중해상도(multi-resolution)로 분해하는 것이다[37]. 이는 국소적인 영역과 전체적인 영역의 정보를 가지면서 공간과 주파수 영역에서의 묘사가 모두 가능하다는 장점이 있

조선대학교
CHOSUN UNIVERSITY

다[16-18]. 웨이브렛 변환을 이용하여 질감 영상을 분할하기 위해서는 변환 후 웨이브렛 계수(wavelet coefficient) 간의 연관성을 추출해야 한다[34]. 최근 간단하면서도 효과적으로 질감의 특징을 추출하기 위해 웨이브렛 계수들의 포락선(envelop)이 이용되고 있다. 이는 질감의 특징을 추출하기 위하여 이산 웨이브렛 패킷 프레임(DWPF: Discrete Wavelet Packet Frame)[38]을 수행하고, 얻어진 계수들의 포락선을 사용하여 질감의 특징을 추출하는 방법이다. 이산 웨이브렛 패킷 프레임은 주파수의 전영역을 동일한 대역으로 분해하는 것으로 계산량이 많고, 특징 벡터가커져 처리 시간이 길어진다는 단점이 있다[39].

3) 형태 정보

형태정보는 영상과 영상내의 객체의 중요한 특징이다. 이러한 형태정보에 의한 방법은 크게는 영상내 객체의 윤곽선 특징을 이용하는 boundary- based 기법과 영상내의 객체의 지역을 이용하는 region- based 기법으로 구분되고, 영상내 객체의 면적과 길이, 원형성(circularity), 신장률(elongation), 곡률(curvature), 주축의 회전각(major axisorientation)등을 특징으로 이용한다. 이러한 형태 특징의 이용은 위치나 크기 등에 영향을 받지 않고, 높은 수준의 검색환경을 제공하는 장점이 있다. 그러나 객체를 이루는 윤곽선이 형태의 변환이나 방향에 따라 민감하게 반응하기 때문에 형태 정보를 표현하기 위해서는 일반적으로 기하학적 변환에 무관한 불변 모멘트(invariant moment)를 정의하여야하고, 영상내의 다양한 객체들을 표현하는 수학적인 모델이 없으며, 형태간의 유사도 측정 기준이 부족하다는 문제점이 있다.

각각의 객체에 대하여 유일하게 정의되는 형태 정보는 객체의 윤곽을 구분 짓는 특성으로 하나의 객체를 인식하기 위한 구조적인 속성을 제공한다. 형태 정보는 객체의 윤곽 특징, 다각형 근사, 유한 요소 모델 등으로 표현할 수 있으며 객체를 이루는 윤곽선은 객체가 놓여 있는 위치나 크기 등에 영향을 받지 않는다는 장점이 있다. 하지만 형태 정보는 객체의 모양 유사성에 대한 주관적인 해석이 가능하므로 객체의 형태를 정확히 정의한다는 것은 어렵다. 그리고 객체를 이루는 윤곽선이 형태의 변환이나 방향에 따라 민감하게 반응하기 때문에 형태 모양의 추출 및 모델링이 어렵다는 단점이 있다.

본 연구에서는 대표적인 형태기반 검색 기법인 Jain과 Vailaya 등의 에지의 방

조선대학교
CHOSUN UNIVERSITY

향 히스토그램을 이용한 방법[10], 저나이크 모멘트를 이용한 방법[40]및 지역적인 미분 불변치(local differential invariants)를 이용한 방법들에 대해 알아본다[41][42].

(1) 에지의 방향 히스토그램을 이용한 기법

Jain에 의해 제안된 방법으로 영상 내의 물체의 에지의 방향 히스토그램을 형태 정보로 이용하였다. 즉 에지의 방향 히스토그램을 구한 후 히스토그램 인터섹션 방법을 이용하여 각 영상의 히스토그램을 비교하였다[10].

먼저 Canny 에지 연산자를 이용하여 에지의 위치를 찾은 후 에지 위치에서 그레디언트의 방향을 구하면 에지의 방향 정보를 얻어낼 수 있다. 즉 픽셀(m,n)에서 가로 수평 방향으로의 변화량을 (m,n)이라 하고, 수직 방향으로의 변화량을 $g_v(m,n)$라고 하면 방향은 식 (2-1)을 이용하여 계산할 수 있다.

$$\theta = \tan^{-1} \frac{g_{v(m,n)}}{g_{h(m,n)}} \tag{2-1}$$

이렇게 에지 위치에서 얻어진 방향들을 이용하여 히스토그램을 생성한 후 앞절의 컬러 히스토그램 인터섹션을 이용하여 유사도를 비교한다. 하지만 에지의 방향을 히스토그램화 하는 것은 영상의 이동에는 불변하지만 축척이나 회전에는 변화한다는 단점을 가지고 있다. 따라서 Jain은 이러한 다음과 같은 방법을 이용하였다. 축척에 대한 문제는 영상내의 총 에지 수를 이용하여 히스토그램을 정규화 함으로써 해결할 수 있다. 하지만 회전에 대한 변화는 축척에 대한 것보다 다소 어렵다. 물체가 회전하면 에지의 방향 히스토그램에서는 이동(shift)이 일어나게 된다. 하지만 다음과 같은 문제가 발생한다. 즉, 에지 방향의 히스토그램의 그룹이 45。의 각도로 나뉘어져 있을 경우, 30, 40。의 에지는 같은 그룹에 포함되게 된다. 그러나 10。의 회전이 이루어질 경우 각기 서로 다른 그룹에 속하게 됨으로써 원래의 히스토그램과는 다른 모양의 히스토그램이 만들어지게 된다. 이러한 문제를 해결하기 위해서 히스토그램에 대해 평활화(smoothing)연산을 수행하였다. 하지만 히스토그램을 매끄럽게 만들어도 급격한 회전에는 적용되지 못하는

조선대학교
CHOSUN UNIVERSITY

단점을 가지고 있으므로, 이동 축척의 변화에는 어느 정도 좋은 결과가 나왔지만, 회전의 변화에는 좋지 못한 결과를 보였다.

(2) 저나이크 모멘트를 이용한 기법

저나이크 모멘트는 단위원의 내부에서 완전한 직교집합을 형성하는 복소 저나이크 다항식으로부터 유도되며 저나이크 모멘트를 이용한 영상의 인식에 대한 연구가 Khontanzad에 의해 제안되었다[40]. 저나이크 다항식의 집합은 식 (2-2)의 $_m\ x,y)$에 의해서 나타내어진다.

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

여기서,

n : 0과 양의 정수

m : $|m| \le n, n-|m| = even$ 조건을 만족하는 정수

ρ,θ : 중심까지의 거리와 시계반대방향의 각도

$R_{nm}(\rho)$: 방사형 다항식

$$R_{nm}(\rho) = \sum_{i=1}^{n-|m|} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \rho^{n-2s} \tag{2-3}$$

디지털 영상 f(x,y)에 대한 n차, 반복계수 m의 저나이크 모멘트는 저나이크 다항식을 이용하여 식 (2-4)와 같이 정규 모멘트와 유사하게 정의 된다.

$$A_{nm} = \frac{n+1}{\Pi} \sum_x \sum_y f(x,y) V^s{}_m(x,y), \quad \text{기서}\ x^2+y^2 \le 1 \tag{2-5}$$

각 영상에 대해서 이렇게 추출된 모멘트들을 비교함으로서 영상간의 형태 유사성을 비교할 수 있다. 저나이크 모멘트의 경우 회전에 의한 모멘트의 크기가

조선대학교
CHOSUN UNIVERSITY

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

조선대학교
CHOSUN UNIVERSITY

Ⅲ. 제안하는 영상검색 방법

본 논문에서는 영상 내의 특정 객체의 위치와 형태에 대한 정보를 가지는 특성벡터를 이용한 영상검색 알고리즘으로 불변모멘트를 이용한 특성추출 및 축차 근사형 유사도 검색방법을 제안하였다. 아래 그림 3.1은 제안한 알고리즘의 순서도이다.

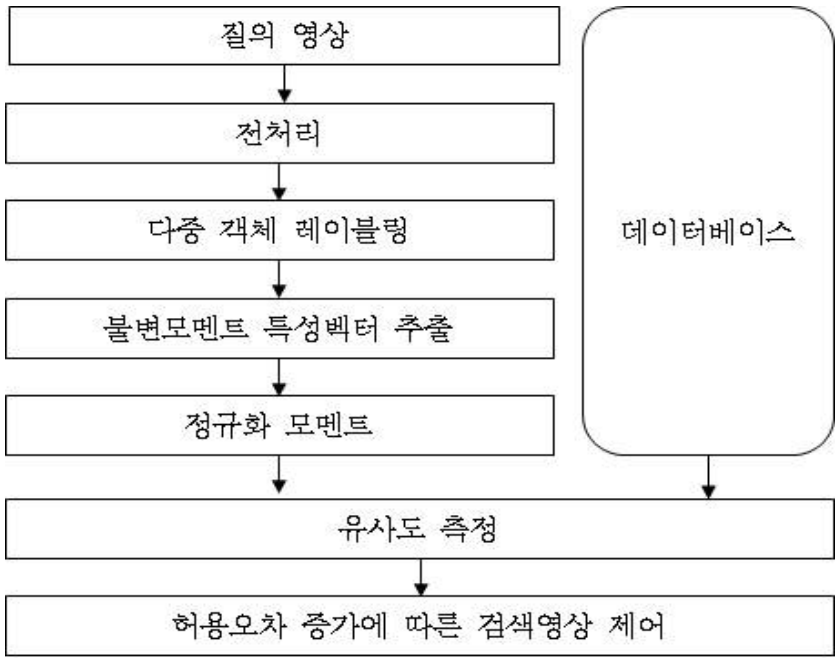


그림 3.1 제안한 알고리즘의 구성도

1. 전처리

전처리 과정으로 첫 번째 잡음제거 및 수축과 팽창에 의한 홀잡음을 제거 한다. 그 다음 Otsu 방법을 이용하여 최적 전역적 문턱처리를 거치고 마지막으로 워터쉐드를 이용하여 분할한다. 그림 3.2는 전처리 과정, 그림 3.3은 질의 영상, 그림 3.4는 그레이스케일 이미지이다.

조선대학교
CHOSUN UNIVERSITY

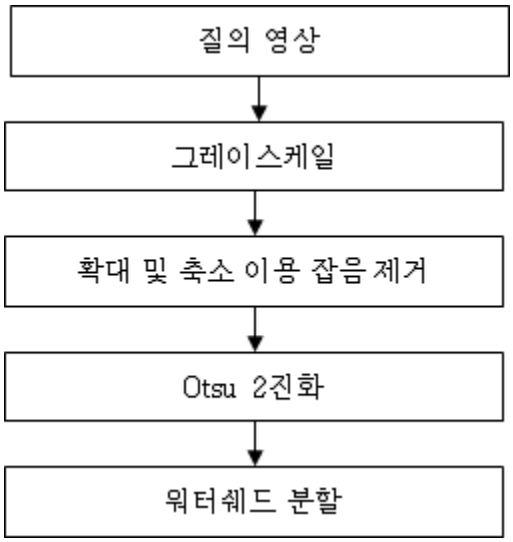


그림 3.2 전처리 과정

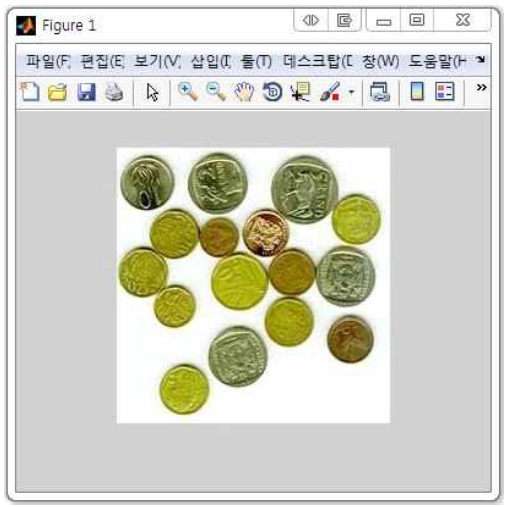


그림 3.3 RGB 이미지

조선대학교
CHOSUN UNIVERSITY

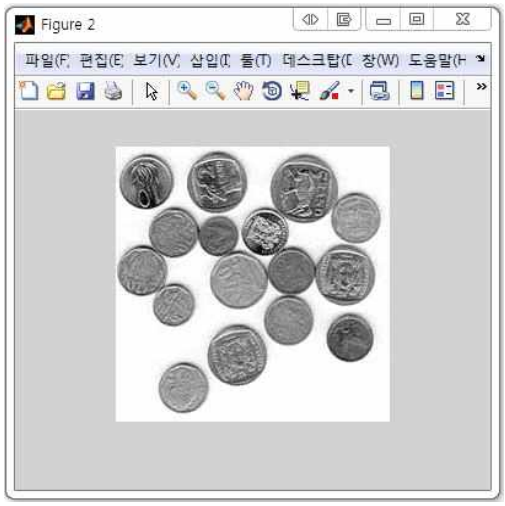


그림 3.4 그레이스케일 이미지

1) Otsu 방법 및 워터쉐드 변환을 이용한 이진화

입력 영상은 개체와 배경들의 정보가 포함되어 있다. 일반적으로 이진화 과정은 입력된 영상으로부터 개체와 배경을 분리하기 위한 전처리 과정으로 사용된다. 이진화는 영상의 픽셀값(pixelvalue)을 0또는 255색으로 변환하여 원하는 개체를 추출하기 위한 과정이다. 이진화를 수행하기 위해서는 영상 내 모든 픽셀에 대하여 그레이(gray)값이 임계값보다 크면 255,작으면 0으로 바꾸는 방법을 사용한다. 이때 그레이 스케일 값의 임계값은 그레이 스케일 범위인 0~255 사이의 정수값을 사용한다.

Otsu의 방법은 주어진 영상의 히스토그램을 작성하고, 이를 이용하여 유사 밝기 값을 갖는 객체들을 분리, 추출하는 방법이다. 즉 Otsu의 방법는 영상 밝기에 대하여, 분할되는 두 클래스의 분리를 클래스사이의 분산을 최대화시키는 임계값을 찾는 방법이다. 주어진 영상이 개 그레이레벨 $1,2,\cdots,$]의 화소로 되어 있다고 하면, 밝기 값이 인 화소의 수를 f, 영상의 총 화소 수를 N이라고 할 때,$N=\underset{i=1}{\ }f_i$이고, 화소가 i 밝기 값을 가질 확률은

$$p_i = f_i / N \tag{3-1}$$

조선대학교
CHOSUN UNIVERSITY

이 된다. 그리고

$$_{=1}\ _{i}\quad 1, P_i \geq 0 \tag{3-2}$$

이다. 입력 영상이 배경과 객체 두 개로 나누어 졌다면, 화소들은 그레이레벨 $[1,\cdots,t]$에 속하는 클래스 C_1과 그레이레벨 $[t+1,\cdots,L]$에 속하는 C_2 클래스로 나눌 수 있다. 따라서 두 클래스에 대한 그레이레벨 확률분포는 $c_1 : P_1/w_1(t)$, $P_2/w_1(t)$ $\cdots$, $P_t/w_1(t)$ 와 $c_2 : P_{t+1}/w_2(t)$, $P_{t+2}/w_2(t)$ $\cdots$, $P_L/w_2(t)$ 이 된다. 여기서

$$(t) = \sum_{i=1}^{t}\ _{i} \tag{3-3}$$

$$w_2(t) = \sum_{i=t+1} P_i \tag{3-4}$$

이고, 클래스 C_1과 C_2의 평균은 각각

$$\mu_1 = \sum_{i=1}^{t} iP_i/w_1(t) \tag{3-5}$$

$$\mu_2(t) = \sum_{i=t+1}^{L} iP_i/w_2(t) \tag{3-6}$$

이다. 또한, 전체 영상의 평균을 μ_T라고 하면,

조선대학교
CHOSUN UNIVERSITY

$$\mu_1 + w_2\mu_2 = \mu \tag{3-7}$$

$$w_1 + w_2 = 1 \tag{3-8}$$

이다. 클래스 과 C_2에 대하여 클래스내의 분산(within-class variance)을 v_w^2, 클래스사이의 분산(between-class variance)을 v_B^2, 그리고 전체 구간에서의 분산(total variance)을 v_T^2 라고 하면, 최적의 임계값은 다음의 판별기준식들 가운데서 하나를 최대화하는 t로 선택할 수 있다.

$$\lambda = v_B^2 / v_w^2 \tag{3-9}$$

$$\lambda = v_T^2 / v_w^2 \tag{3-10}$$

$$\eta = v_B^2 / v_T^2 \tag{3-11}$$

여기서

$$v_w^2 = w_1\sigma_1^2 + w_2\sigma_2^2 \tag{3-12}$$

$$v_B^2 = w_1(\mu_1 - \mu_T)^2 + w_2(\mu_2 - \mu_T) \tag{3-13}$$

$$v_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2 p(i) \tag{3-14}$$

이다. 여기서 σ_1^2과 σ_2^2 은 클래스 C_1과 C_2의 분산이다. 판별기준 식 (3-12)-(3-14)에서 v_T^2는 t의 선택과 무관하므로 η의 계산이 가장 간단하다. 따라서

조선대학교
CHOSUN UNIVERSITY

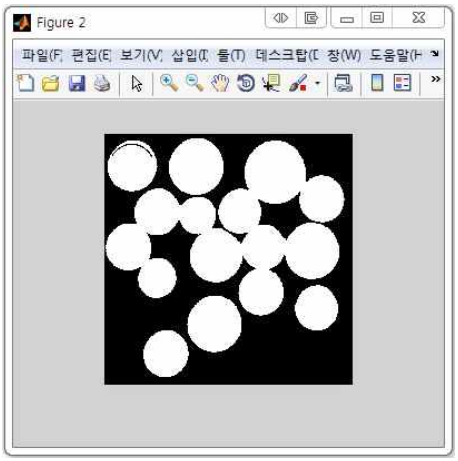
$$(t^*) = \arg max_{1 \le t \le L} \; v_B^2(t) \tag{3-15}$$

를 만족하는 t^*를 최적의 임계값으로 결정할 수 있다.

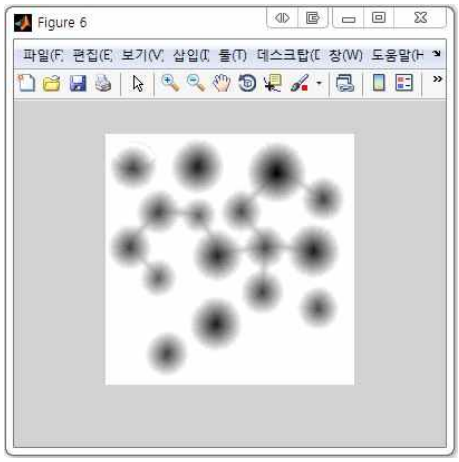
워터쉐드 알고리즘은 수리 형태론(mathematical morphological)을 이용한 영상 분할의 한 방법이다. 영상 화소의 밝기 값에 대한 기울기(gradient) 값을 고도(altitude) 정보로 변환시켜 영역을 구분한다. 이 알고리즘은 담수 영역(catchment basin)을 구분하는 분수령(watershed)을 찾음으로써 각 담수 지역에 결합된 영역을 찾아내는 알고리즘이다.

분할을 위해서 워터쉐드 변환과 함께 자주 사용되는 도구는 거리 변환(distance transform)이다. 이진 영상의 거리 변환은 상대적으로 간단한 개념이다. 각 화소로부터 가장 가까운 영이 아닌 화소까지의 거리이다. 그림 3.5 (a)는 작은 이진 영상 매트릭스를 보여주며, 그림 (b)는 그에 해당하는 거리 변환을 보여준다. 모든 1의 값 화소는 0이 아닌 가장 가까운 화소가 그 자신이므로 거리변환 값이 0임을 주목한다.

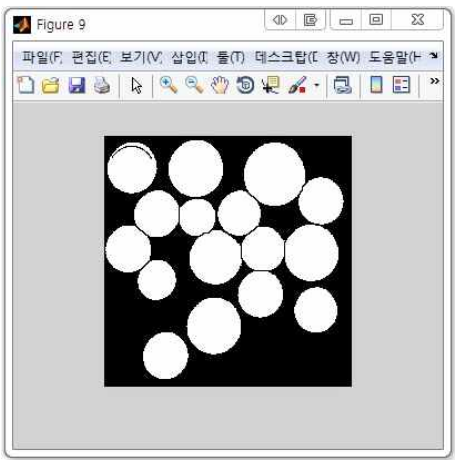
조선대학교
CHOSUN UNIVERSITY



(a) 이진영상



(b) 거리변환



(c) 분할결과

그림 3.5 Otsu 방법 및 워터쉐드 변환을 이용한 이진화 과정

2. 레이블링

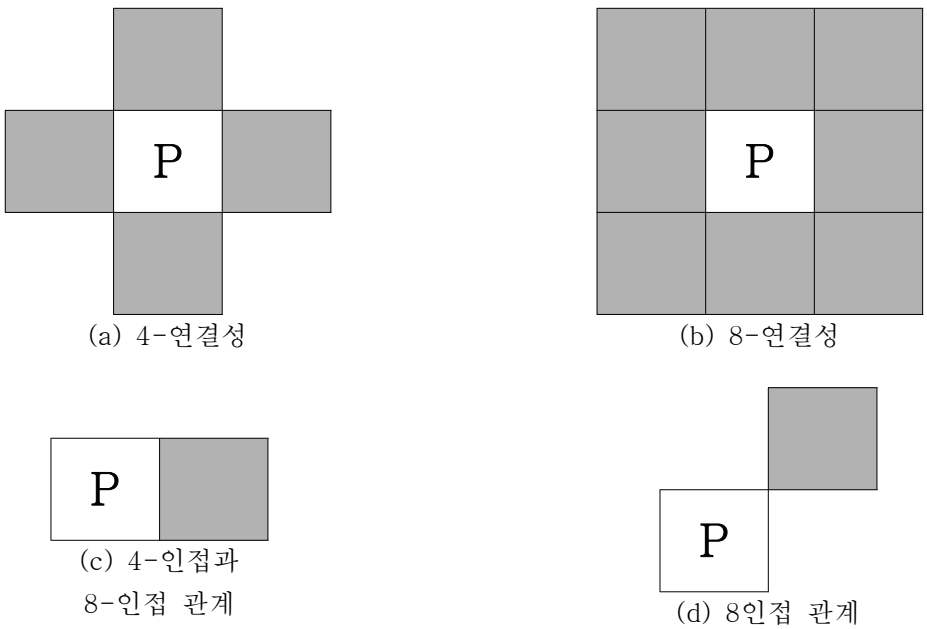
레이블링은 가버 필터링 후 이진화 된 이미지의 모든 화소를 탐색하여 4-연결성(4-connected)이나 8-연결성(8-connected)등으로 연결 관계를 이용하여 번호를 매기는 방법으로 각 객체를 구분하는 방법이다.

레이블링시 4-연결성과 8-연결성은 그림 3.6과 같이 나타낼 수 있다. 이미지상의 (x,y)에 있는 한 화소 p는 좌표가 (x + 1,y), (x - 1,y), (x,y+1), (x,y-1)인 두 개

조선대학교
CHOSUN UNIVERSITY

의 수평 이웃과 두 개의 수직이웃을 갖는데 p의 4이웃 집합을 N4(p)로 표기하고 그림 2.13(a)와 같이 나타낸다. 화소 p의 네 개의 대각선 이웃은 (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)이고 수직 및 수평성분과 함께 p의 8이웃 집합이라 하고 N8(p)와 같이 표기하고 그림 3.6과 같이 나타낸다.

한편, 이미지 상의 두 화소 p와 주변화소가 만일 주변화소 ∈ N4(p)이면 4-인접(4-adjacent)한다고 하고 마찬가지로,p와 주변화소가 만일 주변화소 ∈ N8(p)이면 8-인접(8-adjacent)한다고 한다. 그림 3.6와 같이 4-인접관계는 4-인접과 동시에 8-인접한다고 할 수 있지만 그림 3.6는 4-인접하지 않고 8-인접한다고 할 수 있다.



(a) 4-연결성 (b) 8-연결성

(c) 4-인접과 8-인접 관계 (d) 8인접 관계

그림 3.6 4-연결성과 8-연결성

4-인접관계와 8-인접관계에 따라 레이블링 결과가 달라지는데 4-인접관계를 이용하여 레이블링을 하게 되면 수직과 수평 성분에 관해서만 연결이 유지되므로 그림 3.7 (a)와 같이 레이블링이 1부터 4까지 4단계로 나타내어지고, 8-인접관계를 이용하여 레이블링을 하면 수직, 수평 및 대각선 방향의 화소까지 인접관계에 포함되므로 그림 3.7 (b)와 같이 2단계로 나타내어진다.

조선대학교
CHOSUN UNIVERSITY

				2		
		2	2	2	2	
1	1					
1			3		4	
				4	4	
				4		

(a) 4-연결성

				1		
		1	1	1	1	
1	1					
1			2		2	
				2	2	
				2		

(b) 8-연결성

그림 3.7 4-연결성과 8-연결성을 이용한 레이블링

3. 불변 모멘트 특징벡터 추출

1) 레이블링 영역 형태 정보 추출

영상이 레이블링으로 영역이 분할되면 각각의 레이블링에 대한 특징 정보를 추출한다. 하나의 이미지는 수많은 레이블링으로 구성되므로 모든 레이블링에 대한 특징 정보 추출은 오버 헤드가 발생할 가능성이 높다. 따라서 이 논문에서는 레이블링 된 영역 중에서 크기 비율 10위 이내의 레이블링에 대해서만 특징 정보를 추출한다. 형태에 대한 특징 정보로는 이동, 회전, 크기 변화에 강인한 불변 모멘트(Invariant Moment)를 사용한다.

2) 불변 모멘트

일반적으로 2차원의 연속 함수 f(x, y)에 대한 (p+q)차수의 모멘트는 식 (3-16)과 같다.

$$_{x)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \quad , \quad (p,q=0,1,2,\ldots) \qquad (3\text{-}16)$$

조선대학교
CHOSUN UNIVERSITY

식 (3-16)은 디지털 영상에서 식 (3-17)과 같이 정의된다.

$$_{q} = \sum_{x}\sum_{y} x^{p} y^{q} f(x,y) \quad , \quad (p,q=0,1,2,\ldots) \tag{3-17}$$

임의의 영역에 대한 함수 (x,y)에 대해서 차수 (p, q)에 대한 중심 모멘트 (central moment)의 정의는 식 (3-18)과 같으며 이동에 불변인 특성을 갖는다.

$$\mu_{pq} = \sum_{x}\sum_{y} (x-x)^{p}(y-y)^{q} f(x,y) \tag{3-18}$$

식 (3-18)에서 인덱스 x, y는 각각 수평 x축, 수직 y축의 위치를 나타낸다. 모멘트 m_{0}은 객체 영역을 나타내며 3차 모멘트까지의 중심 모멘트는 식 (3-19)과 같은 관계를 갖는다.

$$\begin{aligned}
\mu_{00} &= m_{00} = \mu \\
\mu_{10} &= \mu_{01} = 0 \\
\mu_{11} &= m_{11} - \mu xy \\
\mu_{20} &= m_{20} - \mu(\overline{x})^{2} \\
\mu_{02} &= m_{02} - \mu(\overline{y})^{2} \\
\mu_{21} &= m_{21} - m_{20}\overline{y} - 2m_{11}\overline{x} + 2\mu(\overline{x})^{2}(\overline{y})^{2} \\
\mu_{12} &= m_{12} - m_{02}\overline{x} - 2m_{11}\overline{y} + 2\mu(\overline{xy})^{2} \\
\mu_{03} &= m_{03} - 3m_{02}\overline{y} + 2\mu(\overline{y})^{3} \\
\mu_{30} &= m_{30} - 3m_{20}\overline{x} + 2\mu(\overline{x})^{3}
\end{aligned} \tag{3-19}$$

중심 모멘트의 의미는 다음과 같다.

μ_{20} : 수평축 분산

μ_{02} : 수직축 분산

μ_{11} : 수평축과 수직축의 상호분산(Covariance)

조선대학교
CHOSUN UNIVERSITY

$_2$: 수평축에서 오른쪽에 비해 왼쪽으로 퍼져 있는 정도

μ_{21} : 수직축에서 위쪽에 비해 아래쪽으로 퍼져 있는 정도

μ_{30} : 수평축의 비대칭 정도(Skew)

μ_{03} : 수직축의 비대칭 정도(Skew)

또한 정규 모멘트는 중심 모멘트를 일정한 크기의 값으로 나눠준 것으로 식 (3-20)과 같이 정의되며, 이는 크기에 불변인 특성을 부여한다.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{3-20}$$

여기서 $\gamma = \frac{p+q+2}{2}$ 이고, $p+q=0,1,2,3$ 이다.

정규화 된 중심 모멘트를 이용하여 식 (3-21)에 의해서 2차, 3차 정규 모멘트로 구성되고 그 영역의 불변 모멘트를 계산 할 수 있다.

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} + \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{3-21}$$

식 (3-19)에서 정의한 불변 모멘트의 의미는 다음과 같다.

조선대학교
CHOSUN UNIVERSITY

: 수평, 수직 방향 분산의 합. 수직과 수평축으로 많이 퍼져 있을수록 값이 커짐.

ϕ_2 : 수직, 수평축의 분산 값이 비슷할 경우 수직, 수평축에 대한 상호 분산 값.

ϕ_3 : 좌우, 상하로 치우친 값을 강조하는 결과.

ϕ_4 : 좌우, 상하로 치우친 값을 상쇄하는 결과.

ϕ_5, ϕ_6, ϕ_7 : 크기, 회전, 위치에 불변한 특징 값을 추출.

본 연구에서는 특성벡터로 연산 복잡도를 감소하도록 레이블링을 정규화 함으로써 1차~4차 모멘트만을 사용하도록 하였다.

아래의 그림 3.8은 앞에서 설명된 식의 과정을 표현한 그림이다. 레이블링된 객체 영역에 대한 불변 모멘트를 이용하여 특성벡터를 생성하기 된다.

조선대학교
CHOSUN UNIVERSITY

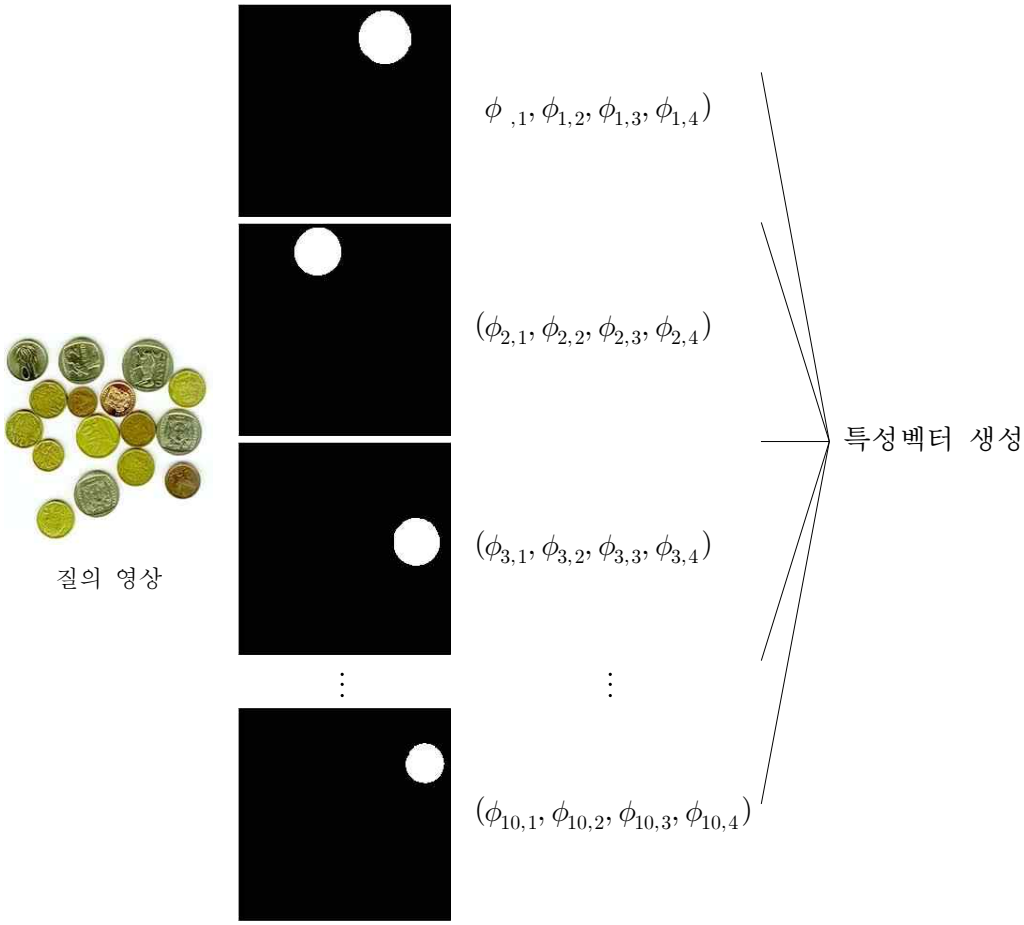


그림 3.8 상위 10개의 대표 레이블링과
그에 해당하는 모멘트를 이용한 특성벡터 추출

3) 영상 변환에 강인한 정규화

본 논문에서는 찌그러짐, 기울기, 크기변환 등으로 영상의 특성이 잡음에 따라 달라질 수 있으므로 정규화 하는 작업이 필요하다. 앞서 추출한 특성벡터에서 영상 변환에 강인한 정규화 벡터를 식 (3-22)에 의해서 추출한다.

$$_{pq} = \frac{^{i+2}X}{^{i}N} \qquad (3\text{-}22)$$

조선대학교
CHOSUN UNIVERSITY

	1	2	3	4
1	0.1612	7.6560e-05	4.3010e-05	1.5877e-06
2	0.1611	1.1945e-04	2.7259e-05	5.4229e-07
3	0.1599	5.6906e-05	1.6793e-05	5.8992e-08
4	0.1597	1.3646e-04	6.6313e-06	1.4940e-08
5	0.1596	4.3296e-05	1.6938e-05	1.2193e-08
6	0.1596	7.8768e-05	3.2401e-06	8.4975e-09
7	0.1595	7.8118e-05	8.1867e-06	9.9510e-09
8	0.1595	8.1437e-05	4.9231e-06	9.9089e-09
9	0.1594	7.4428e-05	1.1823e-06	1.5487e-09
10	0.1594	2.3288e-05	1.0719e-06	5.7343e-09
11	0.1594	6.0334e-05	2.2884e-06	2.1863e-09
12	0.1594	1.4745e-05	7.7614e-06	8.4920e-09
13	0.1594	5.1019e-05	4.1455e-07	1.1075e-09
14	0.1593	1.5746e-05	2.1453e-06	2.7414e-10
15	0.1593	5.1442e-05	2.0208e-06	1.6760e-09
16	0.1593	3.8388e-05	2.6777e-07	2.2842e-10

	1	2	3	4
1	0.1608	8.4304e-05	2.9020e-05	7.2967e-07
2	0.1603	1.0427e-04	1.6894e-05	2.0541e-07
3	0.1598	7.8886e-05	1.3454e-05	2.8708e-08
4	0.1596	8.6173e-05	8.9365e-06	1.1877e-08
5	0.1596	6.6727e-05	9.4550e-06	1.0214e-08

그림 3.9 영상변환에 강인한 정규화

4) 특성 정보 저장

비교 대상이 되는 이미지들을 앞에서 기술한 방법에 의하여 특징 정보들이 추출되어 데이터베이스에 색인되어 있어야 한다. 각 이미지의 특징 정보를 저장하기 위하여 [표 3.1]과 같이 데이터베이스의 테이블을 설계한다.

표 3.1 데이터베이스 테이블 구성

field 1	field 2	field 3	field 4	field 5
특성벡터	정규화벡터	레이블링수	이미지주소	이미지이름
MomentV	NormV	LabelNum	FileWithPath	FileName

4. 유사도 계산

영상 데이터베이스에 저장된 여러 영상들 중에서 사용자가 질의한 영상과 가장 유사한 영상을 검색하는 시스템을 구현하기 위해서는 영상들 간의 유사성 척도(Similarity measure)에 대한 정의와 이를 효율적으로 활용하는 방법이 필요하

조선대학교
CHOSUN UNIVERSITY

다. 영상들 간의 유사성 척도는 특징 벡터로 사용된 영상의 속성이 무엇이냐에 따라서 다양한 방법이 있을 수 있다. 유사성 측정은 두 영상 데이터 사이의 유사성을 수치로 나타낸 것이다. 데이터 분석, 색인, 검색 전반에 걸쳐서 밀접한 관계가 있으며, 내용기반 검색 성능 전체에 영향을 주는 핵심기술이라 할 수 있다.

본 논문에서는 질의 영상과 데이터베이스 영상 사이의 내용기반 영상검색의 시각 유사성을 계산하였다. 따라서, 검색 결과는 질의 영상과 함께 영상의 유사성에 의해 배열하게 되는 영상의 단일 영상이 아니라 유사도 순으로 뿌려주는 검색 리스트이다. 많은 유사성 측정은 최근 특징 분포의 경험적인 추정에 의거하여 영상검색을 위해 개발되었으며 각각의 유사성/거리 측정은 영상 검색 시스템의 검색 성능에 현저하게 영향을 미칠 것이다. 본 논문에서는 질의 영상과 데이터베이스의 영상들 간의 형태 특징유사도를 측정하기 위해서 다음과 같은 간단한 거리척도 함수를 이용한다.

$$Q, I) = \sum_{=1}^{k} |F_i - F'| \qquad (3\text{-}23)$$

식 (3-23)에서 Q는 질의 영상, I는 데이터베이스내의 영상이고, F_i와 F_i'는 각각 두 영상의 특징 벡터이다. 영상들 간의 유사도는 두 특징 값들 사이의 절대적인 거리를 계산함으로써 쉽게 얻을 수 있고 계산된 유사도에 따라 내용 기반 영상 검색 시스템에서 질의 영상과 유사한 영상들이 영상 데이터베이스로부터 검색된다.

5. 성능평가 척도

내용 기반 영상 검색의 효율성을 분석하기 위해, 일반적으로 Recall과 Precision의 두 가지 성능 평가 척도를 이용한다. Recall은 영상 데이터베이스에서 질의와 관련된 영상 중 검색된 영상의 비율이고, Precision은 검색된 영상 중에서 질의와 관련된 영상의 비율을 나타낸다. 즉, A를 영상 데이터베이스 내에서 관련된 영상의 집합이고, B는 검색된 영상의 집합이라고 하자. 그러면 Recall과 Precision은 다음 아래의 식과 같은 조건부 확률로 정의할 수 있다.

조선대학교
CHOSUN UNIVERSITY

$$call = P(\frac{B}{A}), \qquad \mathrm{P}recision = P(\frac{A}{B}) \tag{3-24}$$

실제적인 실험에서는 다음 식 (3-25)와 같은 산출식을 사용한다.

$$R = \frac{R_r}{T}, \quad P_r = \frac{R_r}{T_r} \tag{3-25}$$

여기서 T는 검색대상 데이터베이스에서 질의와 관련된 항목의 총수를 나타내고, R_r은 검색된 항목 중에서 질의와 관련된 항목의 수, T_r은 검색된 항목의 총수를 나타낸다. Recall이 크다는 것은 매칭도가 높다는 것을 뜻하며, 반대로 Precision이 높다는 것은 잘못 검출된 영상이 많지 않다는 것을 뜻한다.

조선대학교
CHOSUN UNIVERSITY

Ⅳ. 실험 및 분석

1. 실험 환경

본 논문에서 제안한 알고리즘은 다양한 영상에 대하여 MATLAB 2014a 소프트웨어를 사용하여 실험하였다. 본 실험은 질의 영상이 모두 같은 상태에서 분석된다. 영상 검색 방법을 설명하기 위해 데이터베이스 영상을 각각의 영상 집합 마다 50~1000개를 놓고 비교되는 영상에 따라 20개를 찾도록 정의하여 놓았다. 실험에서 질의 영상과 비슷한 영상은 알맞은 영상의 개수로 만들어놓았다. 저장해 놓은 데이터는 나중에 다른 질의 영상을 입력 시켰을 때 따로 계산을 하지 않고 질의 영상만 계산되고 데이터의 특징 값들을 이용해 유사도 측정을 비교하여 결과 영상을 도출하게 된다.

2. 시뮬레이션

그림 4.1의 실험 영상은 본 시뮬레이션을 구현하기 위한 실험 영상으로 사용되었다. 먼저 본 알고리즘들의 이해를 돕기 위해 데이터베이스를 그림 4.1과 같이 실험 영상을 30개로 정한 후 30개의 영상에서 질의 영상을 선택한다. 선택된 질의 영상은 본 알고리즘들을 이용하여 데이터베이스 영상들과 비교되어 검색 결과를 나타낸다. 검색 결과는 데이터베이스 30개중에 검색 결과를 20개로 나타내었으며 유사도 순으로 출력 되도록 하였다.

조선대학교
CHOSUN UNIVERSITY

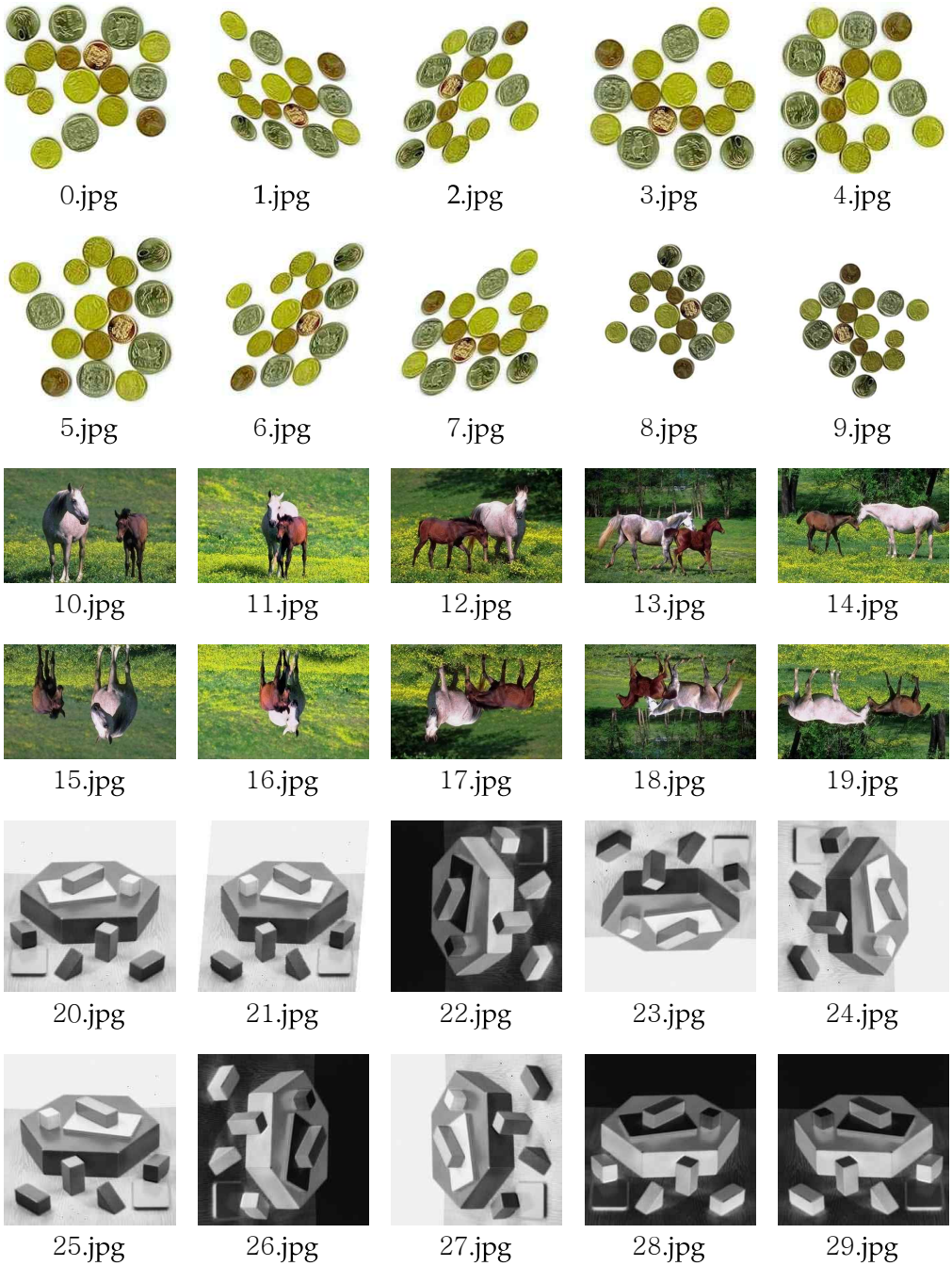


그림 4.1 실험 영상

먼저 그림 4.1을 실험 영상으로 두고 그림 4.1의 20.jpg를 질의 영상으로 정한

조선대학교
CHOSUN UNIVERSITY

후 실험을 수행하였다. 그림 4.2는 질의 영상이다.

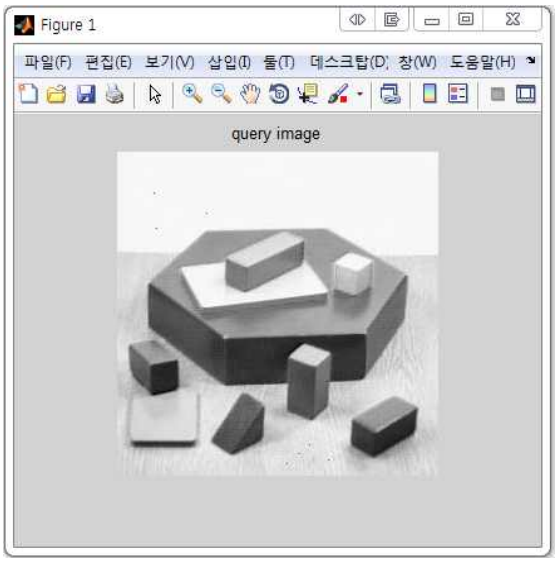


그림 4.2 질의 영상 20.jpg

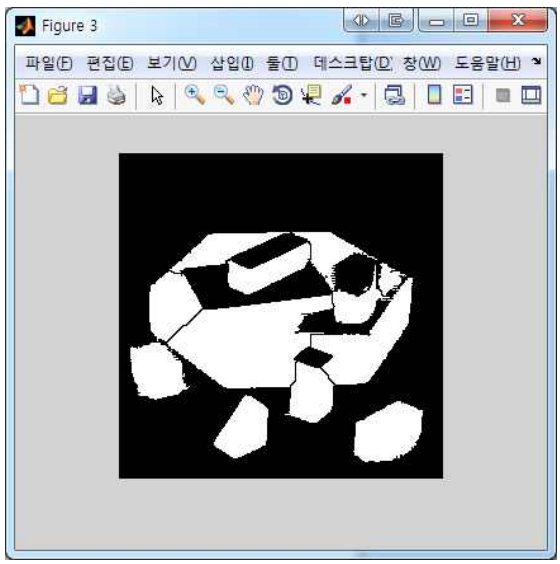


그림 4.3 2진화 영상

조선대학교
CHOSUN UNIVERSITY

	1	2	3	4
1	0.4423	0.0421	0.0142	0.0167
2	0.3064	0.0020	0.0259	4.0060e-04
3	0.2545	0.0315	0.0013	1.0753e-04
4	0.2443	0.0133	0.0058	5.3333e-04
5	0.2383	0.0217	2.3678e-04	4.4487e-05
6	0.2140	0.0036	4.9422e-04	1.0369e-04
7	0.1864	0.0025	7.2033e-04	5.7414e-06
8	0.1864	0.0047	7.5975e-04	3.5017e-05
9	0.1825	0.0053	1.1800e-04	9.3472e-06
10	0.1776	0.0041	2.6199e-04	6.0142e-06

그림 4.4 질의 영상의 특징벡터 추출

	1	2	3	4
1	0.7215	0.3640	0.1180	0.0087
2	0.5286	0.1561	0.0731	0.0058
3	0.4326	0.1294	0.0072	0.0038
4	0.3629	0.0789	0.0126	0.0066
5	0.3414	0.0537	0.0118	0.0059

그림 4.5 질의 영상의 정규화벡터 추출

그림 4.4에서 보듯이 불변모멘트를 이용하여 특징값들이 추출된다. 그림 4.5에서는 추출된 특징벡터를 이용하여 정규화 벡터를 생성한다.

그림 4.6에서는 제안한 알고리즘을 이용한 특징벡터의 유사도 값들을 그래프로 나타내었다. 그림 4.6의 순위는 데이터베이스 영상의 번호에 맞추어 유사도의 순위가 정해졌다.

조선대학교
CHOSUN UNIVERSITY

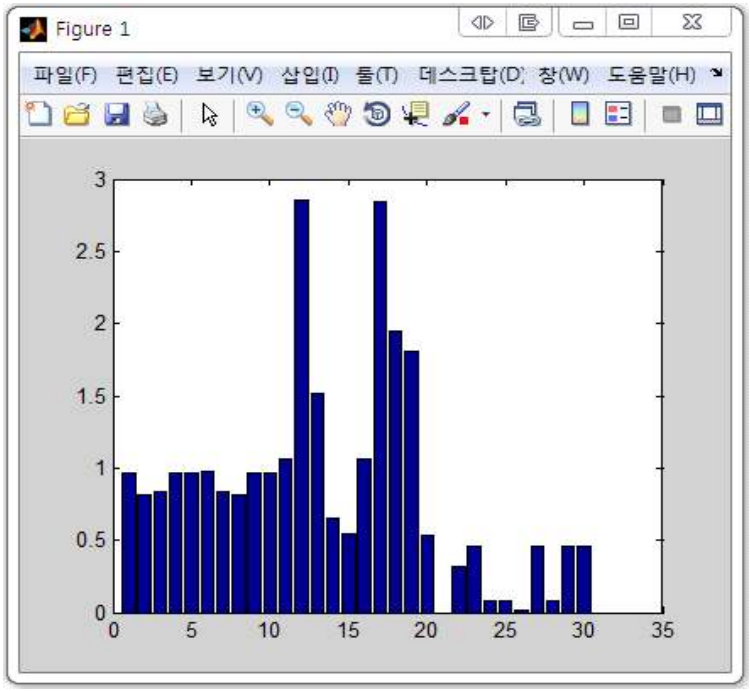


그림 4.6 특징벡터를 이용한 영상 번호별 유사도 측정 결과

마찬가지로 그림 4.7에서도 그림 4.6와 같은 방식으로 나타내었지만 30개의 영상검색은 유사도 측정값의 순위대로 재배열되어 영상검색이 이루어졌다.

조선대학교
CHOSUN UNIVERSITY

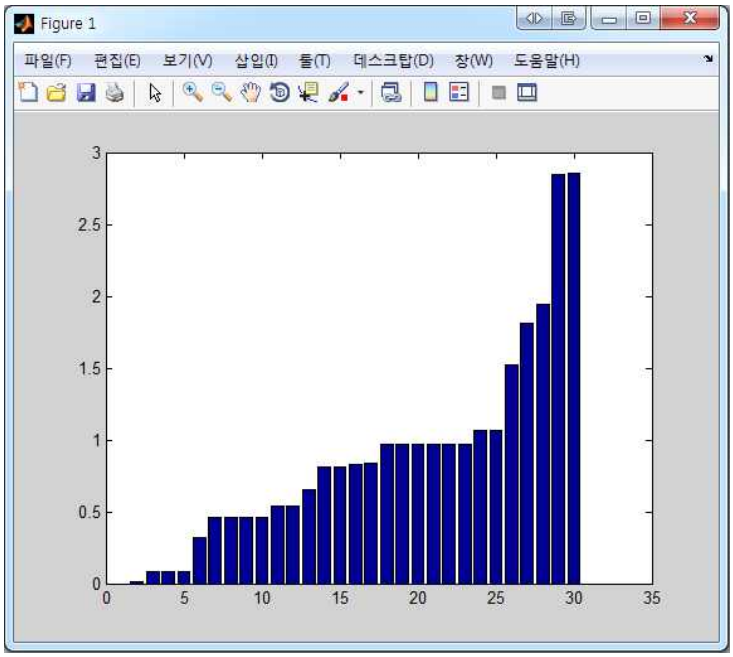


그림 4.7 특징벡터를 이용한 유사도 순위별 측정 결과

그림 4.8에서 보면 질의 영상을 가장 먼저 찾는 것을 확인할 수 있다. 그 다음의 영상들은 질의 영상들과 비슷한 영상들의 순으로 찾아 졌다. 30개의 데이터베이스 영상 중 6개의 유사 영상이 유사도 순으로 찾아지는 것을 확인 할 수 있다. 표 4.1의 오차율은 유클리디언 오차거리를 이용하여 질의 영상과 유사 영상들 간의 특징값의 오차를 나타낸 것이다. 유사도 순으로 나열되었기 때문에 오차율의 순서 역시 유사도 순서와 같게 나온다. 영상마다 유사 영상 값 오차율의 임계값이 다르므로 각 영상의 정보에 따라 가중치를 주어야 한다.

조선대학교
CHOSUN UNIVERSITY

표 4.1 영상의 오차율

Image Number	20	25	27	24	23
Relative error	0.000	0.0110	0.0818	0.0829	0.0845
Image Number	21	29	28	26	22
Relative error	0.3163	0.4583	0.4583	0.4583	0.4583
Image Number	19	14	13	7	1
Relative error	0.5391	0.5404	0.6500	0.8127	0.8152
Image Number	2	6	9	8	0
Relative error	0.8328	0.8412	0.9667	0.9667	0.9694

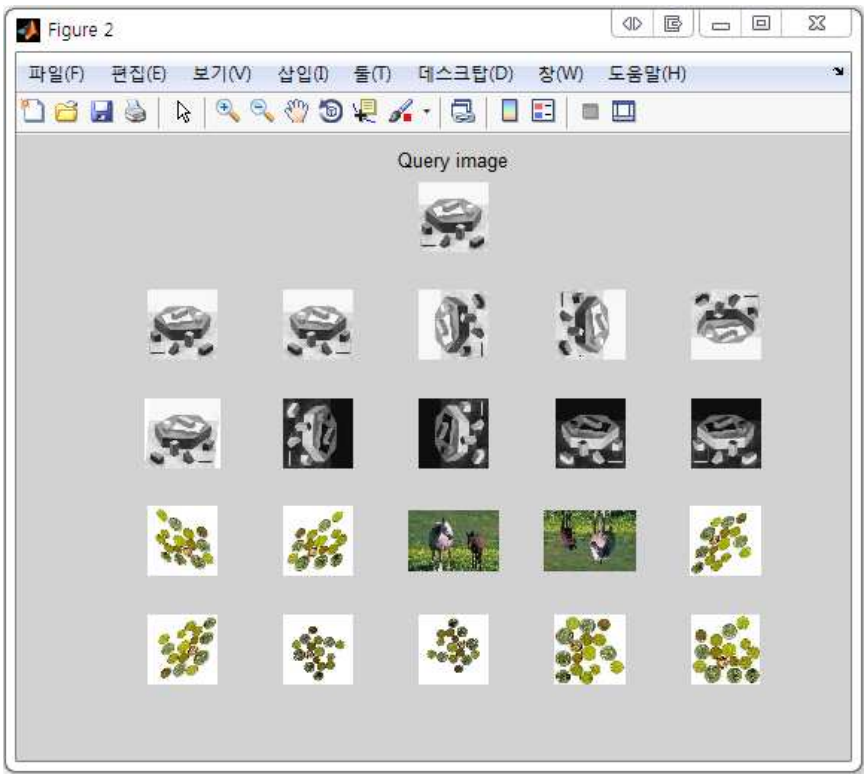


그림 4.8 유사도 값에 의한 영상 검색 결과

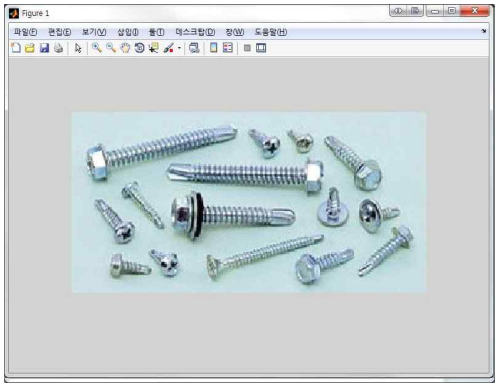
조선대학교
CHOSUN UNIVERSITY

3. 분석 및 고찰

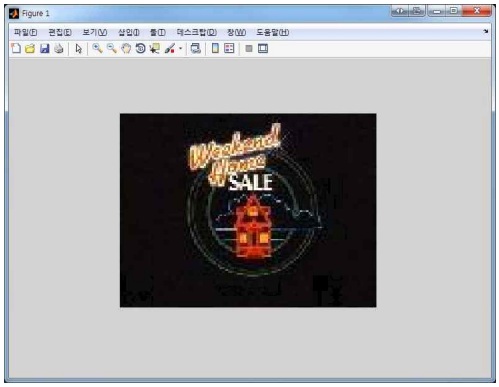
본 논문에서는 제안된 알고리즘의 성능 비교를 위해 Jain과 Vailaya가 제안한 에지 히스토그램을 이용한 방법과 Lu와 Chung 등이 제안한 웨이블릿을 근거한 모양 특징 정보 추출 방법과 비교하고 실험 결과를 제시한다. 본 실험은 질의 영상이 모두 같은 상태에서 분석된다. 영상 검색의 정확성을 실험하기 위해 각 질의 영상을 회전(R), 스케일(S), 노이즈(Z) 변형시킨다. 본 논문에서 제안한 알고리즘들의 성능 측정을 위해 그림 4.9의 질의 영상 Q1~Q4를 갖는 기존 알고리즘과 제안한 알고리즘의 검색 결과의 성능 측정을 비교하기 위해 실험에 사용하였다. 질의 영상을 4개로 사용하였으며 각각의 질의 영상 Q1~Q4는 3000개의 데이터베이스 안에 질의 영상에 맞는 유사 영상을 가지고 있다. 제안한 내용기반 영상 검색 시스템 평가를 위하여 기존의 알고리즘들을 이용하여 제안한 시스템과 Precision과 Recall을 서로 비교 평가하였다.

조선대학교
CHOSUN UNIVERSITY

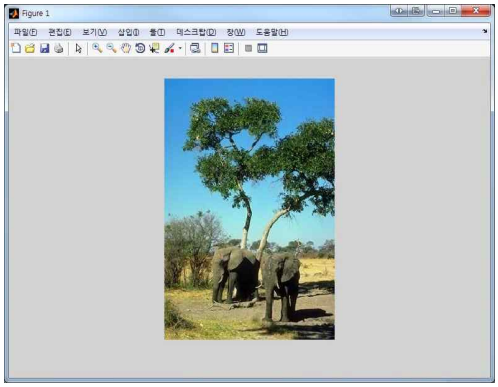
Q1



Q2



Q3

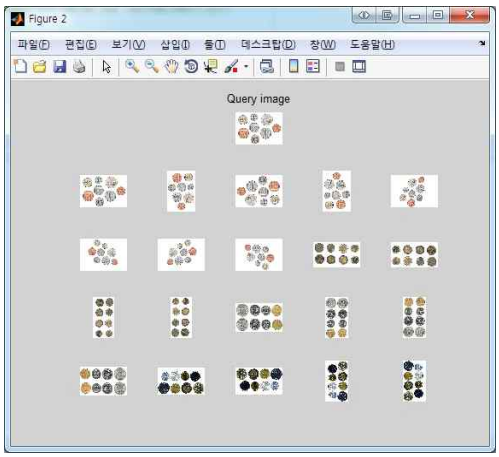


Q4

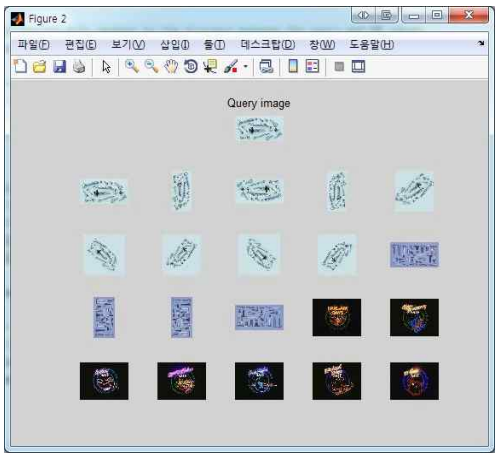
그림 4.9 질의 영상 Q1~Q4

먼저 질의 영상 Q1~Q4를 이용하여 기존 알고리즘들과 제안한 알고리즘의 검색 성능을 비교해 보면 그림 4.10, 그림 4.11과 같다.

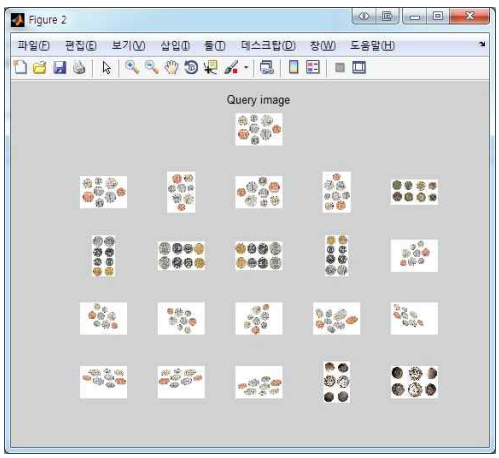
조선대학교
CHOSUN UNIVERSITY



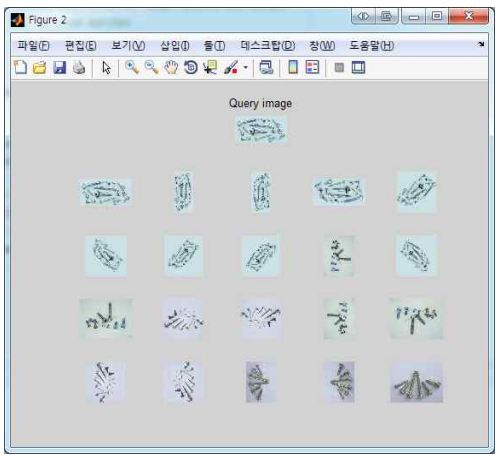
Q1의 에지 히스토그램 알고리즘



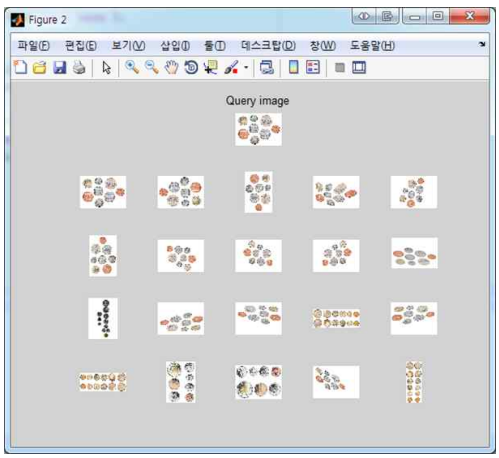
Q1의 에지 히스토그램 알고리즘



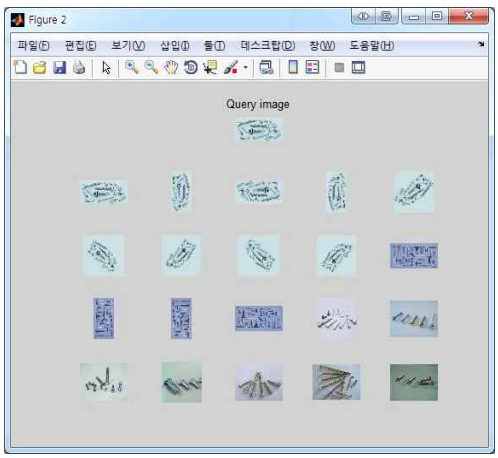
Q1의 웨이블릿 모양 추출 알고리즘



Q2의 웨이블릿 모양 추출 알고리즘



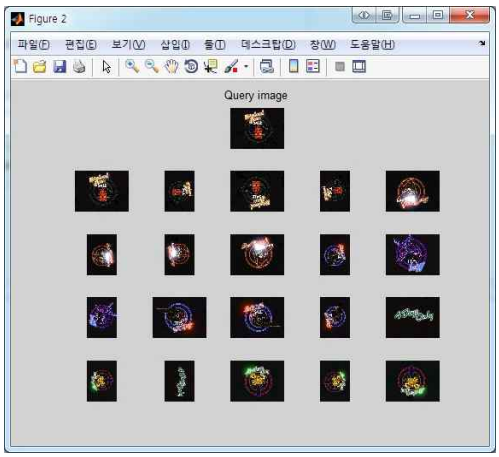
Q2의 제안한 알고리즘



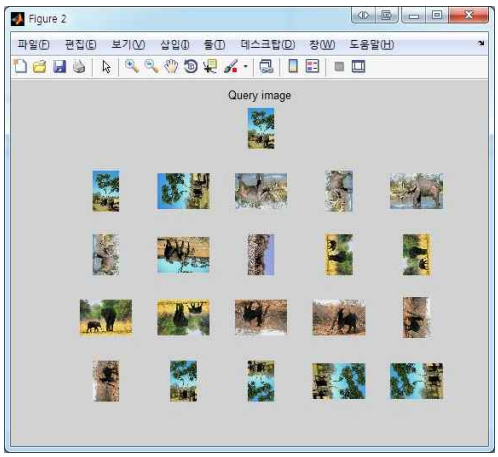
Q2의 제안한 알고리즘

그림 4.10 Q1, Q2의 기존 알고리즘과 제안한 알고리즘의 검색 결과

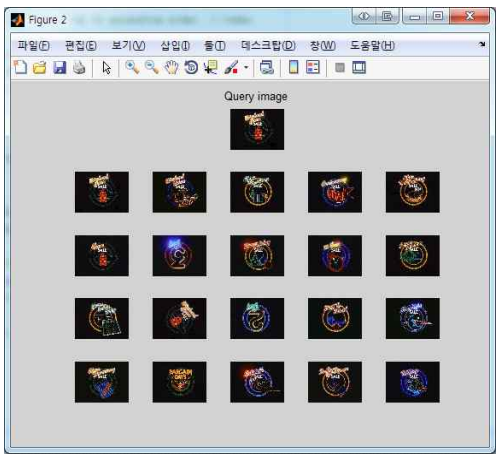
조선대학교
CHOSUN UNIVERSITY



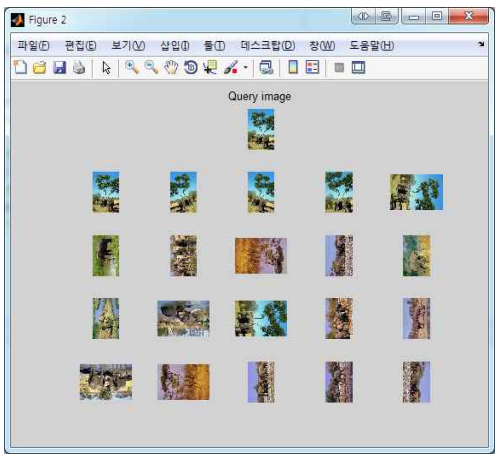
Q3의 에지 히스토그램 알고리즘



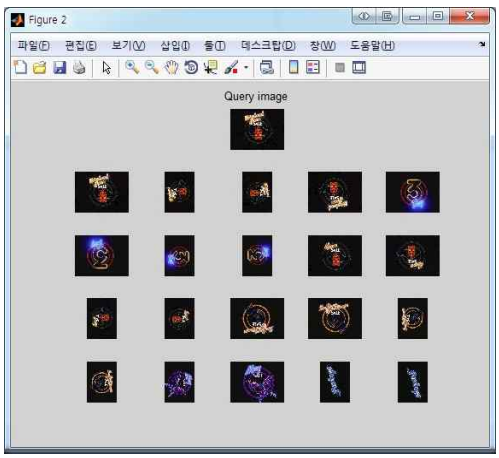
Q4의 에지 히스토그램 알고리즘



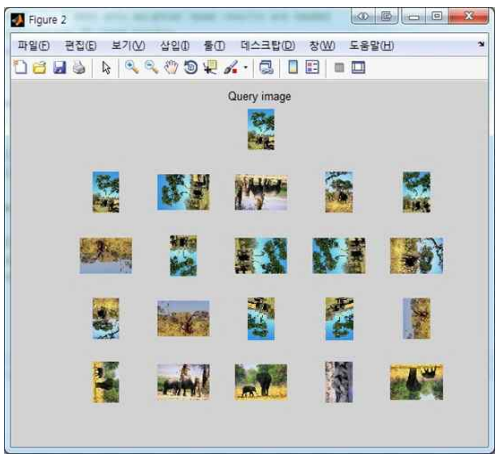
Q3의 웨이블릿 모양 추출 알고리즘



Q4의 웨이블릿 모양 추출 알고리즘



Q3의 제안한 알고리즘



Q4의 제안한 알고리즘

그림 4.11 Q3, Q4의 기존 알고리즘과 제안한 알고리즘의 검색 결과

조선대학교
CHOSUN UNIVERSITY

그림 4.10과 그림 4.11에서와 같이 기존의 에지 히스토그램을 이용한 알고리즘보다 제안한 알고리즘이 육안 상 회전이나 확대 및 축소된 영상을 더 잘 찾아내는 것을 확인할 수 있었다. 그러나 질의 영상 Q1과 같이 객체가 뚜렷한 영상에 비해 질의 영상 Q4와 같은 영상은 기존 에지 히스토그램 알고리즘이 다른 영상이 검색되는 순위가 낮은 반면 제안한 알고리즘은 다른 영상을 먼저 찾아내는 것을 확인할 수 있었다. 그러나 표 4.2과 같이 검색 성능을 Recall과 Precision을 이용하여 비교해 보면 두 알고리즘 모두 Recall은 0.60, Precision은 0.3 으로 결과적으로 같은 검색 성능을 보이는 것을 확인할 수 있었다.

표 4.2 기존 알고리즘과 제안한 알고리즘의 성능 분석

Method / Simulation	에지 히스토그램		웨이블릿 모양 추출		제안한 알고리즘	
	Recall	Precision	Recall	Precision	Recall	Precision
Image Q1	0.32	0.4	0.52	0.65	0.56	0.7
Image Q2	0.9	0.45	0.9	0.45	0.9	0.45
Image Q3	0.26	0.2	0.2	0.15	0.53	0.4
Image Q4	0.6	0.3	0.6	0.3	0.9	0.45
평균	0.52	0.34	0.56	0.39	0.72	0.50

표 4.3은 하나의 객체만 존재하는 단일객체 영상들만을 대상으로 했을 때의 실험결과로서, 에지 히스토그램을 이용한 방법과 웨이블릿을 근거한 모양 특징 정보 추출 방법과 제안된 시스템의 검색 결과를 비교한 것이다. 이들 각각의 특징으로 검색한 결과를 살펴보면 제안된 시스템은 Recall 0.21, Precision 0.8 로 높은 검색 정확도를 나타낸 반면, 에지 히스토그램에서는 Recall 0.15, Precision 0.56, 웨이블릿 모양 추출에서는 Recall 0.19, Precision 0.7로 나타내었다.

조선대학교
CHOSUN UNIVERSITY

표 4.3 단일 객체 영상들의 질의 영상에 따른 실험 결과

Simulation / Method	DB size		T	R_r	Recall	Precision
에지 히스토그램	300	220	60	34	0.15	0.56
웨이블릿 모양 추출	300	220	60	42	0.19	0.7
제안한 알고리즘	300	220	60	48	0.21	0.8

표 4.4는 단일 객체를 갖는 영상들뿐만 아니라 다중 객체를 갖는 영상들도 포함한 전체 영상을 가지고 실험한 결과를 나타내고 있다. 여기서도 표 4.3에서와 같은 조건을 주고 에지 히스토그램을 이용한 방법과 웨이블릿을 근거한 모양 특징 정보 추출 방법과 제안된 시스템간의 검색 결과를 보여주고 있다. 결과 값을 비교해 보면 에지 히스토그램에서는 Recall 0.13, Precision 0.5, 웨이블릿 모양 추출 방법에서는 Recall 0.17, Precision 0.63으로 단일 객체 영상보다 다중 객체에서 검색 효율이 떨어지는 걸 확인할 수 있다. 반면, 제안된 시스템의 Recall과 Precision은 0.23과 0.85로 높은 검색효율을 나타내고 있다. 이러한 결과는 본 논문에서 제안하고 적용한 정규화 레이블링 불변모멘트 알고리즘을 통해 다중 객체를 잘 표현했기 때문에 가능했다. 또한 불변 모멘트의 특성을 이용하여 다양한 형태의 객체를 효과적으로 표현하였기 때문에 기존의 알고리즘보다 높은 정확도를 얻을 수 있었다.

표 4.4 다중 객체 영상들의 질의 영상에 따른 실험 결과

Simulation / Method	DB size	T	T_r	R_r	Recall	Precision
에지 히스토그램	300	220	60	30	0.13	0.5
웨이블릿 모양 추출	300	220	60	38	0.17	0.63
제안한 알고리즘	300	220	60	51	0.23	0.85

조선대학교
CHOSUN UNIVERSITY

Ⅴ. 결론

본 연구에서는 형태정보의 속성으로 정규화 레이블링 불변모멘트 기반으로 다중 객체영상에서의 형태검출을 하는 간단하면서도 효율적인 검색 기법을 제안하였다. 동작 순서는 질의 영상이 주어지면 영상 내 객체를 레이블링하여 최대 면적부터 재배열을 하고 불변모멘트를 이용하여 특징벡터를 추출한다. 추출한 특성벡터에서 영상 변환에 강인한 정규화 벡터를 생성한다. 이로 얻은 특징벡터 값을 이용하여 유사도를 측정하게 된다.

제안된 알고리즘은 시뮬레이션을 통하여 알고리즘의 간단함과 함께 검색의 효율성을 입증하였다. 결과값을 비교해 보면 에지 히스토그램을 사용하는 방식에 비해 Recall과 Precision을 비교해 본 결과 기존 알고리즘 보다 다중 객체 영상에서 Recall은 0.1, Precision은 0.35 높은 것으로 보아 검색 성능이 더 좋은 것으로 확인되었다. 또, 웨이블릿 모양 추출 알고리즘에 비해 Precision이 0.22 높은 것으로 확인되었다.

이와 같이 다중 객체 레이블링 보정과 속성 재배열 기법은 급속도로 발전하고 있는 멀티미디어 데이터의 다양한 정보를 추출함으로써 형태를 기반으로 하는 영상검색에서 정확도와 검색시간 단축 및 최적의 데이터베이스를 구성하고 검색할 수 있는 시스템이나 앞으로 내용 기반 검색 시스템 구축을 위한 기초 자료로 활용할 수 있을 것이라 생각한다.

조선대학교
CHOSUN UNIVERSITY

【참고문헌】

[1] Y. Rui and T. S. Huang, "Image retrieval: current techniques, promising directions, and open issues," J. of Visual Communication and Image Representation, vol. 10, pp. 39-62, 1999.

[2] A. W. M Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[3] S. Michael, "Next-generation web searches for visual content," IEEE Computer, pp. 46-52, Nov. 2000.

[4] Y. H. Ang, Zhao Li and S. H. Ong, "Image Retrieval based on Multimensional Feature Properties", The International Society for Optical Engineering(SPIE), vol. 2420, pp. 47-57, 1995.

[5] R. Brunelli and O. Mich, "Image Retrieval by Examples", IEEE Trans. on Multimedia, vol. 2, no. 3, pp. 164-171, Sep. 2000.

[6] A. Yoshitaka and T. Ichikawa, "A Survey on Content-Based Retrieval for Multimedia Databases", IEEE Trans. on Knowledge and Data Engineering, vol. 11, no. 1, pp. 81-93, 1999.

[7] J.K. Wu, "Content-based Indexing of Multimedia Databases", IEEE Trans. on Knowledge and Data Engineering, Vol.9, No.6, pp.978-989, 1997.

[8] M. J. Swain and D. H. Ballar, "Color index", Int. J. Computer Vision, vol. 7, no, 1, pp. 11-32, July 1991.

[9] Gross M. H. and Koch R., "Visualization of multidimensional shape and texture features in laser range data using complex-valued Gabor wavelets", IEEE Transactions on Visualization and Computer Graphics, vol. 1, pp. 44-59, March 1995.

[10] A. K. Jain and A. Vailaya, “Image retrieval using color and shape,” Pattern Recognition, vol. 29, no. 8, pp. 1233-1244, 1996.

[11] A. D. Bimbo and P. Pala, “Visual image retrieval by elastic matching of user sketches,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, no. 2, pp. 121-132, 1997.

[12] Willian I. Grosky, Peter Neo, and Rajiv Mehrotra, "A Pictorial Index

조선대학교
CHOSUN UNIVERSITY

Mechanism for Model-based Matching", Data and Knowledge Engineering, vol. 8, pp. 309-327, 1992.

[13] Jong-An Park, Young-Eun An, and Min-Hyuk Chang, "2nd Chain code method based on the maximum histogram for shape retrieval", JKIIT, vol. 8, no. 11, pp. 239-245, Dec. 2010.

[14] Seon-Jong Kim, and Young In Kim, "An Extraction of Region-based Information for Image Retrievals", JKIIT, vol. 3, no. 4, pp. 31-37, Aug. 2005.

[15] Juntaek Oh, Hyunwook Kwak, and Wookhyun Kim, "Recognition of Traffic Signs using Wavelet Transform and Shape Information", IEIE, pp.125-134, 2004.

[16] Gwang-won Kang, Ji-min Lee, and Jong-an Park, "Image retrieval using the maximum sequence of the rearranged feature vector based on object corners", JKIIT, vol. 8, no. 12, pp. 247-254, Dec. 2010.

[17] Taek Jin Kwon, "Content-Based Image Retrieval Using Shape and Texture Feature", Master thesis, The Graduate School Kangwon National University, pp. 3-6, 2001.

[18] K. Hirata and T. Kato, "Query by Visual Example", Proc. EDBT conference, pp56-71, 1992.

[19] Seong-Hee Kim, "A Study on the Performance Analysis of Content-based Image & Video Retrieval Systems", Journal of The Korean Biblia Society for Library and Information Science, vol. 15 no. 2, pp. 97-116, 2004.

[20] Hyun-Sool Kim, "Study on the Content-based Image Retrieval using Color, Shape and Texture Information", Doctoral dissertation, The Graduate School Yonsei University, 2001.

[21] Ji-Min Lee, "CBIR based on the Rearrangement of Corner Patch Feature", Master thesis, The Graduate School Chosun University, 2010.

[22] Chun-Hua Huang, "Content-based Image Retrieval using Color Histogram and Multi-texture Analysis", Master thesis, The Graduate School Soongsil University, 2011.

[23] T. Gevers and A. Smeulers, "A Comparative Study of Several Colour Models for Colour Image Invariant Retrieval", Proc. of First Int.

조선대학교
CHOSUN UNIVERSITY

Workshop on Image Database and Multimedia Search, Amsterdam, pp.17-27, Aug. 1996.

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

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

[26] H. Wechsler, "Texture Analysis - A Survey", Signal Processing. vol. 2, no. 3. pp. 271-282, 1980.

[27] J. Y. Hsiao and A. A. Sawchuk, "Texture Feature Performance for Image Segmentation", Pattern Recognition, vol. 23, no. 3, pp. 291-309, 1990.

[28] J. M. Coggins, "A Framework for Texture Analysis Based on Spatial Filtering", University Microfilms International, 1983.

[29] C. H. Chen, L. F. Pau, and P. S. P. Wang, "Handbook of Pattern Recognition and Computer Vision", World Scientific Publishing Co., 1998.

[30] R.M Haralick, "Statistical and Structural Approaches to texture", Proc. IEEE, vol. 67, no. 5, pp. 786-804, May 1979.

[31] T. Randen, "Filter and Filter Design for Image Recognition", Ph.D thesis, Norwegian Univ., Oct. 1997.

[32] T. Randen and J. H. Husoy, "Filtering for Texture Classification: A Comparative Study", IEEE Trans. on Patt. 43Anal. Machine Intell., vol. 21, no. 4, pp. 291-310, Apr. 1999.

[33] T. Chang and C. j. Kuo, "Texture Analysis and Classification with Tree-Structured Wavelet Transform", IEEE Trans. on Image Proc., vol. 2, no. 4, pp. 429-441, 1993.

[34] M. Unser and M. Eden, "Multiresolution Feature Extraction and Selection for Texture Segmentation", IEEE Trans. on Patt. Anal. Machine Intell., vol. 11, no. 7, pp. 717-728, July 1989.

[35] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signatures", IEEE Trans. on Patt. Anal. Machine Intell., vol. 15, no. 11,

조선대학교
CHOSUN UNIVERSITY

pp. 1186-1191, Nov. 1993.

[36] Y. Y. Tang, L. H. Yang, J. Liu, and H. Ma, "Wavelet Theory and Its Application to Pattern Recognition", World Scientific, 2000.

[37] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Trans. on Patt. Anal. Machine Intell., vol. 11, no. 7, pp. 674-693, July 1989.

[38] J. C. Goswami and A. K. Chan, "Fundamentals of Wavelets: Theory, Algorithms, and Application", Wiley Inter-science, 1999.

[39] Byung-Gil Ye, "Texture Segmentation Using Wavelet Frame and Spatial Information", Master thesis, Korea Maritime University, 2002.

[40] Nishat Ahmad, "Object Retrieval Approach with invariant Features Based on corner Shapes," IEEE Communication Society, vol. 1, pp. 840-845, 2007.

[41] 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. 486-4971, 1990.

[42] C. Schmid and R. Mohr, "Matching by Local Invariants," Technical report N2644, INRIA, 1995.

[43] Sxhmid and R. Mohr, "Local Gray value Invariant for Image Retrieval," IEEE Trans. on pattern Analysis and Machine Intelligence, vol. 19, no. 5, pp. 550-535, 1997.

[44] A. K. Jain and A. Vailaya, “Shape-based retrieval: A case study with trademark image databases,” Pattern Recognition, vol. 31, no. 9, pp. 1369-1390, 1998.

조선대학교
CHOSUN UNIVERSITY

creative commons
COMMONS DEED

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

dCollection

조선대학교
CHOSUN UNIVERSITY

2015년 8월
석사학위 논문

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

2015년 8월 석사학위논문

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

김 국 정

조선대학교
CHOSUN UNIVERSITY

2015년 8월

석사학위 논문

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

Multi-object Image Retrieval based on Normalization Labeling Invariant Moments

2015년 8월 25일

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

정규화 레이블링 불변모멘트에 기반한 다중객체 영상검색

지도교수 박 종 안

이 논문을 공학석사학위 신청 논문으로 제출함

2015년 4월

조선대학교 산업기술융합대학원

소프트웨어융합공학과

김 국 정

조선대학교
CHOSUN UNIVERSITY

김국정의 석사학위논문을 인준함

위원장 조선대학교 교수 김 판 구 (인)

위 원 조선대학교 교수 안 영 은 (인)

위 원 조선대학교 교수 박 종 안 (인)

2015년 5월

조선대학교 대학원

조선대학교
CHOSUN UNIVERSITY

목 차

Ⅰ. 서론 ····· 1
1. 연구 배경 및 목적 ····· 1
2. 연구 내용 ····· 3
3. 논문 구성 ····· 3

Ⅱ. 관련연구 ····· 4

Ⅲ. 제안하는 영상검색 방법 ····· 13
1. 전처리 ····· 13
1) Otsu 방법 및 워터쉐드 변환을 이용한 이진화 ····· 15
2. 레이블링 ····· 19
3. 불변 모멘트 특성벡터 추출 ····· 21
1) 레이블링 영역 형태 정보 추출 ····· 21
2) 불변 모멘트 ····· 21
3) 영상 변환에 강인한 정규화 ····· 25
4) 특성 정보 저장 ····· 26
4. 유사도 계산 ····· 26
5. 성능평가 척도 ····· 27

Ⅳ. 실험 및 분석 ····· 29
1. 실험 환경 ····· 29
2. 시뮬레이션 ····· 29
3. 분석 및 고찰 ····· 36

조선대학교
CHOSUN UNIVERSITY

Ⅴ. 결론 ····· 42

【참고문헌】 ····· 43

조선대학교
CHOSUN UNIVERSITY

그 림 목 차

그림 2.1 내용기반 영상검색 시스템의 구조 ····· 4
그림 3.1 제안한 알고리즘의 구성도 ····· 13
그림 3.2 전처리 과정 ····· 14
그림 3.3 RGB 이미지 ····· 14
그림 3.4 그레이스케일 이미지 ····· 15
그림 3.5 Otsu 방법 및 워터쉐드 변환을 이용한 이진화 과정 ····· 19
그림 3.6 4-연결성과 8-연결성 ····· 20
그림 3.7 4-연결성과 8-연결성을 이용한 레이블링 ····· 21
그림 3.8 상위 10개의 대표 레이블링과 그에 해당하는 모멘트를 이용한 특성벡터 추출 ····· 25
그림 3.9 영상변환에 강인한 정규화 ····· 26
그림 4.1 실험 영상 ····· 30
그림 4.2 질의 영상 20.jpg ····· 31
그림 4.3 2진화 영상 ····· 31
그림 4.4 질의 영상의 특징벡터 추출 ····· 32
그림 4.5 질의 영상의 정규화벡터 추출 ····· 32
그림 4.6 특징벡터를 이용한 영상 번호별 유사도 측정 결과 ····· 33
그림 4.7 특징벡터를 이용한 유사도 순위별 측정 결과 ····· 34
그림 4.8 유사도 값에 의한 영상 검색 결과 ····· 35
그림 4.9 질의 영상 Q1~Q4 ····· 37
그림 4.10 Q1, Q2의 기존 알고리즘과 제안한 알고리즘의 검색 결과 ····· 38
그림 4.11 Q3, Q4의 기존 알고리즘과 제안한 알고리즘의 검색 결과 ····· 39

조선대학교
CHOSUN UNIVERSITY

표 목 차

표 3.1 데이터베이스 테이블 구성 ····· 26
표 4.1 영상의 오차율 ····· 35
표 4.2 기존 알고리즘과 제안한 알고리즘의 성능 분석 ····· 40
표 4.2 단일 객체 영상들의 질의 영상에 따른 실험 결과 ····· 41
표 4.4 다중 객체 영상들의 질의 영상에 따른 실험 결과 ····· 41

조선대학교
CHOSUN UNIVERSITY

ABSTRACT

Multi-object Image Retrieval based on Normalization Labeling Invariant Moments

Kim Gukjeong
Advisor : Prof. Park Jongan, Ph.D.
Dept, of Software Convergence Engineering,
Graduate School of Industry Technology Convergence,
Chosun University

In CBIR with multi-objects, It shows an interest in the image retrieval method based on object features. In this paper, we have proposed an image retrieval method based on normalized labelling features of multi-objects. First, it extracts, normalizes and rearranges the object-labels which present image shape information. Secondly, it measures image similarity using moment features invariant to transition, rotation and scaling of the objects. The simulation results using Multi-object database showed that our proposed method is superior to the edge histogram method and the wavelet shape extraction method. The experimental results demonstrate the high retrieval efficiency i.e Recall 0.23, Precision 0.85. Consequently, our retrieval system shows better performance than the conventional system that cannot express the shape of objects exactly.

조선대학교
CHOSUN UNIVERSITY

Ⅰ. 서론

1. 연구 배경 및 목적

인터넷과 모바일 단말기의 발달로 웹 사이트나 블로그, SNS 등을 통해 많은 사람들이 디지털 영상을 쉽게 생성하고 저장할 수 있게 되었다. 그러나 그 영상량이 방대하여 관리와 검색은 점점 어려워지고 있다. 이에 따라 영상을 입력으로 이용하는 다양한 응용 분야들이 최근 증가하고 있다. 특히 영상 검색은 최근 상당히 주목을 받고 있고 많은 연구가 이루어지고 있는 분야이다.

영상 검색을 위해서는 데이터베이스에 저장되는 각 영상에 대한 분류 및 주제어에 대한 결정 과정이 필수적이다. 초기 영상 검색에서는 수동으로 인덱스를 지정하는 방식을 이용하였으나 입력 작업의 비효율성과 작업자의 주관에 따라 입력 내용의 결정되는 등의 많은 어려움을 내포하고 있었다. 이에 최근에는 영상을 분석함으로써 칼라 분포나 질감 등과 같은 영상의 특징 정보를 자동으로 추출하여 검색에 이용하는 내용기반 영상 검색(content-based image retrieval, CBIR)이 등장하게 되었다[1-7].

내용 기반 영상 검색은 컬러(color), 형태(shape), 그리고 질감(texture) 등의 내용 표현 요소들을 통하여 얻어진 특징 값을 사용하여 영상들 간의 유사도를 계산하고 검색을 수행한다. 그러므로 효율적인 영상 검색을 위해서는 무엇보다도 각 영상을 대표하는 표현 요소들의 효과적인 특징 추출이 중요하다. 영상의 표현 요소들 중에서 컬러 특징(color feature)은 영상을 표현하는 주된 특징으로 내용 기반 영상 검색의 초기부터 사용되어 온 방법이다. 컬러 특징은 영상 내의 조명 변화나 관측 위치, 크기 변화 등에 어느 정도 무관하게 적용 가능하다는 장점을 가지고 있으며, 방법이 간단하여 내용 기반 영상 검색에서 가장 많이 사용되는 방법 중의 하나이다. 컬러 특징을 추출하는 대표적인 방법이 컬러 히스토그램(color histogram)을 이용하는 것이다[8]. 컬러 히스토그램을 특징 정보로 하면 컬러의 수 만큼에 해당하는 크기의 공간이 필요하므로 보통 유사한 컬러 별로 군집하는 컬러 집합을 사용하게 된다. 초기에는 히스토그램을 균등한 비율로 나누어 히스토그램 구간(bin)을 결정하는 방식으로 컬러를 군집했지만, 효율적인 검색 시스템을 위해서 유사한 컬러끼리 군집하는 방법들이 제안되었다. 단, 컬러

조선대학교
CHOSUN UNIVERSITY

히스토그램이나 컬러 집합을 이용하면 영상 내 객체들의 배치 즉 컬러 분포의 공간적인 정보를 전혀 이용할 수 없다는 단점이 있다.

질감은 영상 내부의 동일한 패턴을 가지는 영역에 대한 특징으로 하나의 컬러 정보로만은 표현할 수 없다. 질감은 영상 내의 객체의 표면이 가지는 영상에게는 매우 중요한 특징이 된다. 특히 항공사진을 분석하는 경우에 질감 정보가 많이 사용되어 왔다. Gabor 필터, wavelet 변환 등을 통해서 질감 특징을 분석하는 방법이 연구되고 있는데[9], 이러한 방법은 인공적인 질감 즉 건물의 벽돌 무늬나 의류에서 나타나는 무늬 등의 분석에 큰 효과를 보이고 있다.

형태(shape) 특징값은 영상의 회전, 스케일 변화 등에 매우 민감하며 배경과 객체를 자동 또는 수동으로 분리해내는 알고리즘이 필요하다는 어려움을 가지고 있다. 이러한 점을 극복하기 위하여 Jain과 Vailaya[10]는 별도의 객체 분리 과정 없이 영상의 에지 성분들에 대하여 에지 히스토그램을 구성하고 회전과 스케일 변화 등을 고려하기 위해서 히스토그램 정규화나 히스토그램 smoothing 방법 등을 적용하였다. 그러나, 이 방법은 영상의 스케일 변화에는 강인하지만, 히스토그램 smoothing 에 의해 다른 영상도 같다고 판단할 문제점을 가지고 있다. 이밖에 사용자가 스케치한 영상을 가지고 2차 spline 곡선을 구성하여 검색하는 방법[11] 등이 제안 되었지만 이러한 방법들은 이론 자체가 어렵고, view point 변화와 객체 영역을 미리 설정해주어야 하는 단점을 가지고 있다.

영상의 내용을 파악하기 위해서는 무엇보다도 영상 내의 객체(object)의 형태(shape)에 관심을 가질 필요가 있다. 많은 영상에서는 비교적 크기가 크고 주변 배경과는 대비되는 특징을 가지며 또한 경계가 뚜렷하게 구별되는 객체를 가지고 있다. 형태 분석 및 표현 방법은 크게 형태의 외곽선 정보를 이용한 것과 영역 내부를 이용한 것으로 분류가 된다. 외곽선 정보를 이용하는 방법은 영상에서 외곽선을 추출하여 이의 분석을 통해 그 변화를 특징 정보를 삼는 것이다. 역 내부 정보를 이용하는 것은 영상 전체의 화소의 분포를 분석하는 것으로 영상에 대한 모멘트를 구하는 방법과 같이 형태의 분포 정도를 특징으로 삼는 것이다[12]. 객체를 추출하는 과정은 객체 특징의 다양성을 효과적으로 반영하며 진행되어야 한다. 이에 따라 영상에 포함된 관심 객체를 추출한 후, 영상의 핵심정보에 대한 효과적인 형태특성을 추출하여 영상을 인식하고자 하는 여러 연구결과가 발표되고 있다[13-16].

조선대학교
CHOSUN UNIVERSITY

위에서 설명한 여러 가지 특징들을 추출하여 효과적으로 사용함으로써 내용 기반 영상 검색의 효율성을 높일 수 있다. 본 논문에서는 다양한 객체를 갖는 영상으로부터 영상에 내포된 핵심적인 영상 정보를 표현할 수 있는 의미 있는 관심 객체를 정규화 레이블링을 통해 추출하고 객체의 위치 이동이나 회전, 크기 변화 등과 같은 각종 변환에 강인한 불변 모멘트 특징 값을 이용하여 유사도를 측정하는 영상검색 알고리즘을 제안하고자 한다.

2. 연구 내용

본 논문에서는 다중 객체 영상에서의 정규화 레이블링 모멘트기반 형태검색 알고리즘을 제안한다. 제안한 형태 속성 기반 검색 알고리즘은 크게 4단계로 구성된다. 즉,

제 1단계는 전처리 과정으로 잡음제거 및 수축과 팽창에 의한 홀잡음제거
제 2단계는 레이블링 이용 다중 객체 추출 및 불변 모멘트 특성벡터 추출
제 3단계는 특성벡터 정규화 처리
제 4단계는 유사도 측정 단계로 이루어진다.

컴퓨터 시뮬레이션에서 제안한 알고리즘의 유사도 측정은 Wang 데이터베이스를 이용하여 기존의 형태검색 기법과 비교하여 정확도 성능을 나타낸다.

3. 논문 구성

이 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 연구를 고찰하고, 3장에서는 다중 객체를 갖는 영상의 정규화 레이블링 모멘트 특성을 추출하여 형태를 검색하는 새로운 알고리즘을 제안한다. 그리고 4장에서는 제안된 알고리즘에 대한 시뮬레이션을 수행하고 그 결과를 분석한다. 그리고 5장에서 결론을 짓는다.

조선대학교
CHOSUN UNIVERSITY

Ⅱ. 관련연구

1. 내용기반 영상 검색

내용기반 영상검색은 영상을 대표할 수 있는 특징값을 자동으로 추출하여, 이를 이용하여 질의자가 원하는 영상을 검색하는 것을 말한다. 이러한 내용기반 영상검색 시스템은 일반적으로 영상의 입력, 질의, 검색 세 가지의 과정으로 나뉘며, 그 구조는 아래와 같다[17].

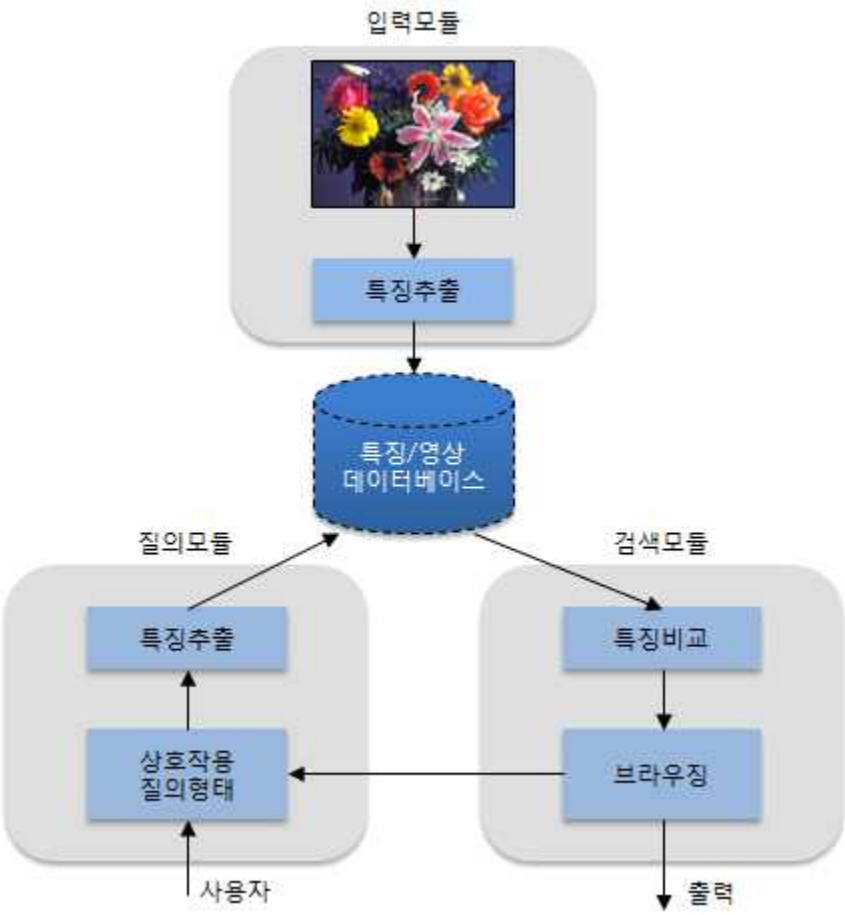


그림 2.1 내용기반 영상검색 시스템의 구조

조선대학교
CHOSUN UNIVERSITY

영상의 입력과정은 영상이 입력되었을 때에 영상을 대표할 수 있는 특징을 추출하여 영상 데이터베이스 내에 저장하는 단계이고, 영상의 질의 과정은 사용자, 질의자가 찾고자 하는 영상의 입력이 주어지면 특징을 추출하는 과정이고, 질의 영상의 특징이 추출되어 검색과정에서 미리 구성되어진 영상 데이터베이스내의 특징과 질의 영상의 특징을 비교하여 가장 근사한 영상을 출력하는 단계가 검색 과정이다.

내용 기반 검색을 효율적으로 수행하기 위해서는 의미 있는 특징 추출이 매우 중요하다. 의미 있는 특징들은 크게 키워드나 주석과 같은 의미와 관련 있는 특징들과, 컬러(color), 질감(texture) 및 형태(shape)등의 시각적인 특징들로 나눌 수 있다. 이와 같은 특징들은 인식대상을 유일하고, 명확하게 표현할 수 있고, 회전(rotation), 이동(translation), 크기변화(scaling)와 같은 기하학적인 변환에 대해서 불변하는 성질을 가져야만 한다. 내용기반 영상검색에서는 위에서 언급한 특징의 조건들을 만족하기 위해서 색상, 형태, 질감을 단독 혹은 병합하여 특징 값으로 이용한다.

1992년 kato[18]에 의하여 내용기반 영상 검색이라는 용어가 현재까지 널리 사용되고 있다. 내용기반 영상 검색은 영상의 내용을 분석하여 시각적 특징 정보를 자동으로 추출하고 색인(Index)을 추가하여 저장하는 과정과 데이터베이스에서 사용자가 원하는 영상을 자동으로 검색하는 과정을 포함한다. 원 영상으로부터 정확한 특징 정보를 추출하기가 어려운 단점이 있으나. 특징 정보를 자동으로 추출하여 색인 과정을 통한 데이터베이스구축에 필요한 시간 및 인력의 소모를 줄일 수 있다는 장점이 있다.

내용기반 영상 검색은 영상으로부터 컬러, 질감, 형태 등과 같은 특징을 이용한다. 컬러를 이용한 방법에는 1991년 swain과 ballard[8]이 제안한 컬러 히스토그램을 이용하는 방법이 있다. 이 방법은 각 컬러에 대해 영상 픽셀 빈도를 나타내는 방법으로 질의 영상과 데이터베이스 영상의 히스토그램을 비교하여 공유되는 최솟값을 구하는 히스토그램 인터섹션 기법을 사용하였다. 이 방법은 현재도 많은 내용기반 영상 검색 시스템에서 사용되고 있다.

김성희(2004)는 이미지 및 동영상 내용 기반 검색 시스템의 개념과 유형을 조사한 후, 상용중인 내용 기반 검색 시스템 5개를 선정하여 성능을 분석하였

다. 현재 상업용으로 이용 가능한 시스템인 QBIC, Excalibur Visual Retrieval

조선대학교
CHOSUN UNIVERSITY

Ware, ImageFinder & VideoFinder, VIR Image Engine, IMatch의 일반적인 시스템 특성, 제공되는 검색 기법, 운영 체제 등을 분석하였다[19].

김현술(2001)은 영상 데이터베이스에서 자동으로 영상을 검색할 수 있는 내용 기반 영상 검색 방법을 제안하였다. 내용 기반 영상 검색 시스템에서는 주로 영상의 특징으로 색상, 형태 및 질감 등의 원시적인 특징들이 사용되고 있으며 이 연구에서도 이러한 특징들을 검색 정보로 사용하여 검색을 수행하였다[20].

이지민(2010)은 영상 내의 특정 객체의 위치와 형태에 대한 정보를 가지는 interest point로서 에지영상에서 라인을 추출하여 얻어낸 코너점을 이용한 영상 검색 알고리즘으로 재배열된 코너 패치 속성의 히스토그램 빈도수를 사용한 검색기법을 제안하였다[21].

황춘화(2011)는 영상을 색상히스토그램 방법으로 색상의 분포가 비슷한 영상들을 질감 특징인 EHD(Edge Histogram Descriptor)과 이웃 화소들의 질감을 고려한 GLCM(Gray Level Co-occurence)방법을 이용하여 색상 및 질감특징들에 가중치를 주어서 특징 값들을 비교함으로써 질의 영상과 비슷한 영상들을 색인 및 검색함으로써 가장 비슷한 영상들을 보여주는 영상검색 방법을 제안하였다[22].

검색을 위한 내용기반 영상 질의의 목적은 사용자의 질의 영상과 유사도 기준을 만족하는 영상을 데이터베이스로부터 효율적으로 검색하는 것이다.

영상의 내용을 기반으로 하는 검색은 전통적인 데이터베이스 검색과 비교하여 몇 가지 중요한 차이점을 갖는다.

첫째, 데이터베이스 검색은 정확성(exact)을 목표로 하지만, 내용기반 검색은 유사성(similarity)에 기반을 둔 검색이다. 따라서 사용자의 검색 절차를 매우 단순화 할 수 있다.

둘째, 사용자와 시스템과의 상호작용을 통해 다양한 검색 가능성을 제공한다.

셋째, 시각적인 질의를 필요로 하는 멀티미디어 시나리오와 같은 응용 단계에서 사용될 수 있다.

넷째, 유사도(similarity value)에 따른 후보 영상을 최종 사용자가 선택할 수 있도록 영상들을 시각적으로 표현과 출력할 수 있다.

이러한 영상 데이터베이스 특징 중 유사성 검색은 주어진 표본과의 유사한 정보를 나타내는 유사성 척도(similarity measure)에 의해 검색하는 것으로 사용자의 입장에서 보다 자연스럽게 데이터베이스와 대화할 수 있는 기능을 제공한다.

조선대학교
CHOSUN UNIVERSITY

유사성 척도는 유사성 검색의 성능과 밀접한 관련이 있으며 문제 영역의 특성에 따라 다양한 척도를 가질 수 있다. 그리고 유사성 척도를 이용하여 논리적으로 유사한 영상들을 함께 모아 저장함으로써 검색 시 탐색 시간을 줄이고 시스템 효율을 증대시킬 수 있다. 뿐만 아니라, 유사도를 기준으로 퍼지 질의가 가능해진다. 내용기반 영상 검색은 사용자의 질의 영상과 비교하여 유사도가 높은 영상을 검색하는 유사도 검색 방법을 사용하며, 이미지 검색 엔진, 지리 정보 시스템, 의료 분야 등의 다양한 분야에 응용될 수 있다.

2. 영상의 내용 표현 요소

1) 색상 정보

영상의 색상정보를 이용한 영상검색 기법은 영상내의 각 화소의 색상값을 특징값으로 하며, Swain과 Ballard가 제안한 색상히스토그램 정합을 기반으로 하는 영상 검색 기법이 소개된 후 색상을 이용한 내용기반 영상검색에서 색상 히스토그램은 가장 일반적으로 색상정보를 대표하는 특징값이 되었다. 이런 색상 히스토그램과 같은 색상정보를 이용한 방법은 칼라 영상의 경우 그레이 영상에 비해 정보량이 많음을 이용하여 영상의 통계적 특성을 이용한 3차원 히스토그램을 생성한 후에 히스토그램 공간상에서 질의영상과 데이터베이스 영상의 정합(Matching)을 시도하는 인덱싱 알고리즘이다. 이 알고리즘은 간단하고, 히스토그램의 특성으로 이동(translation)이나 크기변화(scaling)에 비교적 강건하고, 물체의 변형이나, 일부의 절취에 대해서 비교적 강건한 장점을 가지고 있으나, 특징값이 영상내의 각 객체에 대한 세부정보를 표현할 수 없고, 빛의 밝기와 영상내의 물체의 크기에 민감하여, 전혀 다른 영상일지라도 같은 색상분포를 가질 수 있는 단점을 가진다. 때문에 색상과 더불어 질감, 형태와 같은 특징을 병합하여 이용하거나, 영상에서 영역별로 색상 히스토그램을 이용하는 방법과 같이 색상 히스토그램과 함께 다른 추가정보를 이용하여 사용한다.

색상 정보를 이용한 대표적인 연구로는 색상 모델의 선택이 어떻게 히스토그램에 매칭 되는가를 분석한 연구[23], 이미지의 지역적 색상 속성을 특성화하기 위해 컬러히스토그램을 집합을 제안한 연구[24], VisualSEEK 시스템에서 컬러 이미지의 위치영역에 사용한 색상 집합에 대한 연구[25]등이 있다.

조선대학교
CHOSUN UNIVERSITY

2) 질감 정보

질감(Texture)은 화소 값들의 공간적인 배열로 표현되어지는 속성으로 부드러움, 거침, 규칙적임 등으로 묘사할 수 있다. 이러한 질감으로 이루어진 영상을 질감 영상이라 하며, 모피, 곡물류, 위성사진과 의료 영상 등이 여기에 속한다. 최근 이러한 질감 영상들을 분석하기 위하여 질감 영상 분할이 많이 연구되어지고 있다. 질감 영상 분할이란 질감 영상에서 비슷한 질감 별로 분할하는 것을 말한다. 질감영상 분할은 질감에 대한 모든 속성을 정확히 정의할 수 없기 때문에 분할하는데 많은 어려움이 있다. 이러한 이유로 질감을 묘사하기 위한 많은 특징들이 제안되었고, 여러 가지 질감 영상을 분할하기 위한 방법들이 제안되어 왔다[26-30].

질감 영상을 분할하기 위해서는 먼저 영상에서 질감을 구분할 수 있는 특징을 추출하여야 한다. 질감의 특징을 추출하는 방법으로는 크게 통계적인 방법, 구조적인 방법 그리고 스펙트럼을 이용한 방법 등이 있다[26][30]. 통계적인 방법은 일정 영역 내 명암도의 통계적인 특징을 이용하는 방법으로 질감을 부드러움, 거침, 알갱이 모양 등으로 표현한다. 이 방법은 계산이 단순하다는 장점이 있지만, 일정 크기의 영역만 묘사할 수 있다는 문제점이 있다. 구조적인 방법은 화소 값들의 규칙적인 배열 상태를 다루는 방법으로 가까운 이웃 화소 간의 공간적인 상호작용만을 표현할 수 있고 전반적인 특징을 알지 못한다는 단점이 있다. 스펙트럼을 이용한 방법은 주파수 영역에서 질감의 특징을 추출하는 방법으로 푸리에 스펙트럼의 성질을 이용하여 영상의 전반적인 주기성을 검출하는데 주로 사용된다. 그러나 이 방법은 영상의 국부적인 묘사를 할 수 없다는 단점이 있다. 최근 필터링을 이용하여 질감의 특징을 추출하는 방법들이 제안되고 있는데, 이는 질감마다 주파수 성분을 가지는 성질을 이용한 것으로, 질감 영상을 주파수 대역 별로 분할하여 질감의 특징을 추출하는 방법이다[31][32]. 이 중 웨이브렛 변환을 이용한 방법이 가장 많이 사용된다[33-36].

웨이브렛 변환은 정규직교기저(orthonormal basis)의 저대역 통과필터와 고대역 통과필터로 구성된 필터 뱅크(filter bank)를 이용하여 영상을 다중해상도(multi-resolution)로 분해하는 것이다[37]. 이는 국소적인 영역과 전체적인 영역의 정보를 가지면서 공간과 주파수 영역에서의 묘사가 모두 가능하다는 장점이 있

조선대학교
CHOSUN UNIVERSITY

다[16-18]. 웨이브렛 변환을 이용하여 질감 영상을 분할하기 위해서는 변환 후 웨이브렛 계수(wavelet coefficient) 간의 연관성을 추출해야 한다[34]. 최근 간단하면서도 효과적으로 질감의 특징을 추출하기 위해 웨이브렛 계수들의 포락선(envelop)이 이용되고 있다. 이는 질감의 특징을 추출하기 위하여 이산 웨이브렛 패킷 프레임(DWPF: Discrete Wavelet Packet Frame)[38]을 수행하고, 얻어진 계수들의 포락선을 사용하여 질감의 특징을 추출하는 방법이다. 이산 웨이브렛 패킷 프레임은 주파수의 전영역을 동일한 대역으로 분해하는 것으로 계산량이 많고, 특징 벡터가커져 처리 시간이 길어진다는 단점이 있다[39].

3) 형태 정보

형태정보는 영상과 영상내의 객체의 중요한 특징이다. 이러한 형태정보에 의한 방법은 크게는 영상내 객체의 윤곽선 특징을 이용하는 boundary- based 기법과 영상내의 객체의 지역을 이용하는 region- based 기법으로 구분되고, 영상내 객체의 면적과 길이, 원형성(circularity), 신장률(elongation), 곡률(curvature), 주축의 회전각(major axisorientation)등을 특징으로 이용한다. 이러한 형태 특징의 이용은 위치나 크기 등에 영향을 받지 않고, 높은 수준의 검색환경을 제공하는 장점이 있다. 그러나 객체를 이루는 윤곽선이 형태의 변환이나 방향에 따라 민감하게 반응하기 때문에 형태 정보를 표현하기 위해서는 일반적으로 기하학적 변환에 무관한 불변 모멘트(invariant moment)를 정의하여야하고, 영상내의 다양한 객체들을 표현하는 수학적인 모델이 없으며, 형태간의 유사도 측정 기준이 부족하다는 문제점이 있다.

각각의 객체에 대하여 유일하게 정의되는 형태 정보는 객체의 윤곽을 구분 짓는 특성으로 하나의 객체를 인식하기 위한 구조적인 속성을 제공한다. 형태 정보는 객체의 윤곽 특징, 다각형 근사, 유한 요소 모델 등으로 표현할 수 있으며 객체를 이루는 윤곽선은 객체가 놓여 있는 위치나 크기 등에 영향을 받지 않는다는 장점이 있다. 하지만 형태 정보는 객체의 모양 유사성에 대한 주관적인 해석이 가능하므로 객체의 형태를 정확히 정의한다는 것은 어렵다. 그리고 객체를 이루는 윤곽선이 형태의 변환이나 방향에 따라 민감하게 반응하기 때문에 형태 모양의 추출 및 모델링이 어렵다는 단점이 있다.

본 연구에서는 대표적인 형태기반 검색 기법인 Jain과 Vailaya 등의 에지의 방

조선대학교
CHOSUN UNIVERSITY

향 히스토그램을 이용한 방법[10], 저나이크 모멘트를 이용한 방법[40]및 지역적인 미분 불변치(local differential invariants)를 이용한 방법들에 대해 알아본다[41][42].

(1) 에지의 방향 히스토그램을 이용한 기법

Jain에 의해 제안된 방법으로 영상 내의 물체의 에지의 방향 히스토그램을 형태 정보로 이용하였다. 즉 에지의 방향 히스토그램을 구한 후 히스토그램 인터섹션 방법을 이용하여 각 영상의 히스토그램을 비교하였다[10].

먼저 Canny 에지 연산자를 이용하여 에지의 위치를 찾은 후 에지 위치에서 그레디언트의 방향을 구하면 에지의 방향 정보를 얻어낼 수 있다. 즉 픽셀(m,n)에서 가로 수평 방향으로의 변화량을 (m,n)이라 하고, 수직 방향으로의 변화량을 $g_v(m,n)$라고 하면 방향은 식 (2-1)을 이용하여 계산할 수 있다.

$$\theta = \tan^{-1} \frac{g_{v(m,n)}}{g_{h(m,n)}} \tag{2-1}$$

이렇게 에지 위치에서 얻어진 방향들을 이용하여 히스토그램을 생성한 후 앞절의 컬러 히스토그램 인터섹션을 이용하여 유사도를 비교한다. 하지만 에지의 방향을 히스토그램화 하는 것은 영상의 이동에는 불변하지만 축척이나 회전에는 변화한다는 단점을 가지고 있다. 따라서 Jain은 이러한 다음과 같은 방법을 이용하였다. 축척에 대한 문제는 영상내의 총 에지 수를 이용하여 히스토그램을 정규화 함으로써 해결할 수 있다. 하지만 회전에 대한 변화는 축척에 대한 것보다 다소 어렵다. 물체가 회전하면 에지의 방향 히스토그램에서는 이동(shift)이 일어나게 된다. 하지만 다음과 같은 문제가 발생한다. 즉, 에지 방향의 히스토그램의 그룹이 45。의 각도로 나뉘어져 있을 경우, 30, 40。의 에지는 같은 그룹에 포함되게 된다. 그러나 10。의 회전이 이루어질 경우 각기 서로 다른 그룹에 속하게 됨으로써 원래의 히스토그램과는 다른 모양의 히스토그램이 만들어지게 된다. 이러한 문제를 해결하기 위해서 히스토그램에 대해 평활화(smoothing)연산을 수행하였다. 하지만 히스토그램을 매끄럽게 만들어도 급격한 회전에는 적용되지 못하는

조선대학교
CHOSUN UNIVERSITY

단점을 가지고 있으므로, 이동 축척의 변화에는 어느 정도 좋은 결과가 나왔지만, 회전의 변화에는 좋지 못한 결과를 보였다.

(2) 저나이크 모멘트를 이용한 기법

저나이크 모멘트는 단위원의 내부에서 완전한 직교집합을 형성하는 복소 저나이크 다항식으로부터 유도되며 저나이크 모멘트를 이용한 영상의 인식에 대한 연구가 Khontanzad에 의해 제안되었다[40]. 저나이크 다항식의 집합은 식 (2-2)의 $_m\ x,y)$에 의해서 나타내어진다.

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

여기서,

n : 0과 양의 정수

m : $|m| \le n, n-|m| = even$ 조건을 만족하는 정수

ρ,θ : 중심까지의 거리와 시계반대방향의 각도

$R_{nm}(\rho)$: 방사형 다항식

$$R_{nm}(\rho) = \sum_{i=1}^{n-|m|} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \rho^{n-2s} \tag{2-3}$$

디지털 영상 f(x,y)에 대한 n차, 반복계수 m의 저나이크 모멘트는 저나이크 다항식을 이용하여 식 (2-4)와 같이 정규 모멘트와 유사하게 정의 된다.

$$A_{nm} = \frac{n+1}{\Pi} \sum_x \sum_y f(x,y) V^s{}_m(x,y), \quad \text{기서}\ x^2+y^2 \le 1 \tag{2-5}$$

각 영상에 대해서 이렇게 추출된 모멘트들을 비교함으로서 영상간의 형태 유사성을 비교할 수 있다. 저나이크 모멘트의 경우 회전에 의한 모멘트의 크기가

조선대학교
CHOSUN UNIVERSITY

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

조선대학교
CHOSUN UNIVERSITY

Ⅲ. 제안하는 영상검색 방법

본 논문에서는 영상 내의 특정 객체의 위치와 형태에 대한 정보를 가지는 특성벡터를 이용한 영상검색 알고리즘으로 불변모멘트를 이용한 특성추출 및 축차 근사형 유사도 검색방법을 제안하였다. 아래 그림 3.1은 제안한 알고리즘의 순서도이다.

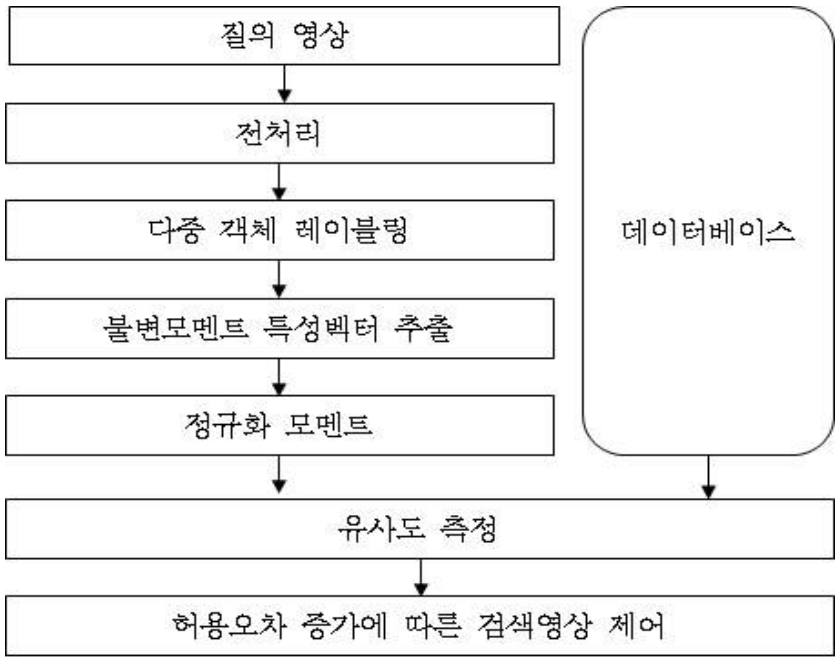


그림 3.1 제안한 알고리즘의 구성도

1. 전처리

전처리 과정으로 첫 번째 잡음제거 및 수축과 팽창에 의한 홀잡음을 제거 한다. 그 다음 Otsu 방법을 이용하여 최적 전역적 문턱처리를 거치고 마지막으로 워터쉐드를 이용하여 분할한다. 그림 3.2는 전처리 과정, 그림 3.3은 질의 영상, 그림 3.4는 그레이스케일 이미지이다.

조선대학교
CHOSUN UNIVERSITY

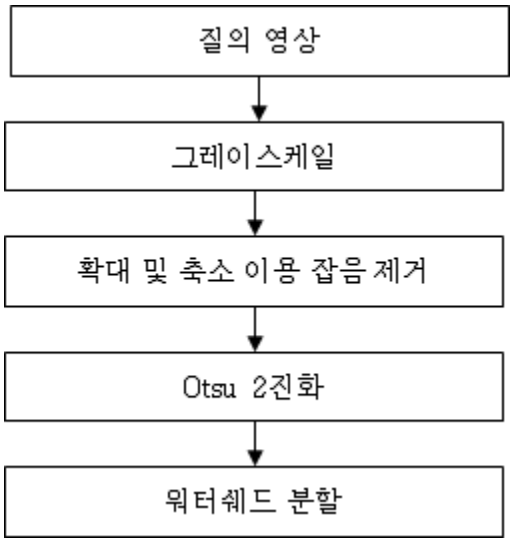


그림 3.2 전처리 과정

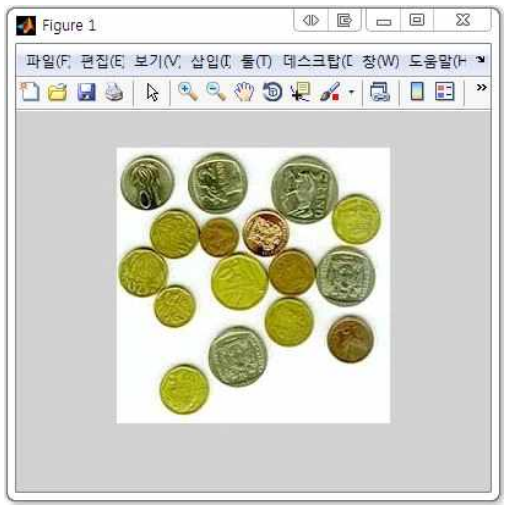


그림 3.3 RGB 이미지

조선대학교
CHOSUN UNIVERSITY

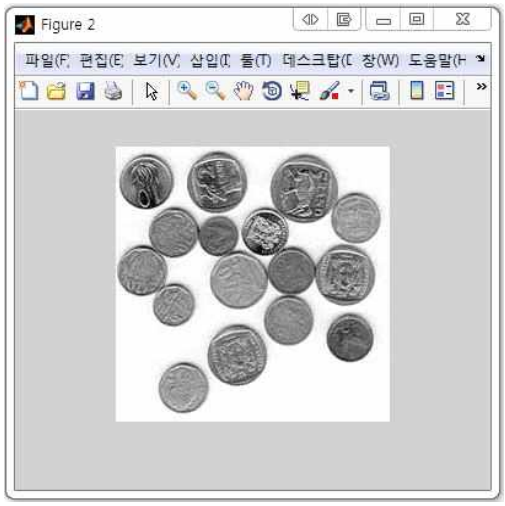


그림 3.4 그레이스케일 이미지

1) Otsu 방법 및 워터쉐드 변환을 이용한 이진화

입력 영상은 개체와 배경들의 정보가 포함되어 있다. 일반적으로 이진화 과정은 입력된 영상으로부터 개체와 배경을 분리하기 위한 전처리 과정으로 사용된다. 이진화는 영상의 픽셀값(pixelvalue)을 0또는 255색으로 변환하여 원하는 개체를 추출하기 위한 과정이다. 이진화를 수행하기 위해서는 영상 내 모든 픽셀에 대하여 그레이(gray)값이 임계값보다 크면 255,작으면 0으로 바꾸는 방법을 사용한다. 이때 그레이 스케일 값의 임계값은 그레이 스케일 범위인 0~255 사이의 정수값을 사용한다.

Otsu의 방법은 주어진 영상의 히스토그램을 작성하고, 이를 이용하여 유사 밝기 값을 갖는 객체들을 분리, 추출하는 방법이다. 즉 Otsu의 방법는 영상 밝기에 대하여, 분할되는 두 클래스의 분리를 클래스사이의 분산을 최대화시키는 임계값을 찾는 방법이다. 주어진 영상이 개 그레이레벨 $1,2,\cdots,$]의 화소로 되어 있다고 하면, 밝기 값이 인 화소의 수를 f, 영상의 총 화소 수를 N이라고 할 때,$N=\underset{i=1}{\ }f_i$이고, 화소가 i 밝기 값을 가질 확률은

$$p_i = f_i / N \tag{3-1}$$

조선대학교
CHOSUN UNIVERSITY

이 된다. 그리고

$$_{=1}\ _{i}\quad 1, P_i \geq 0 \tag{3-2}$$

이다. 입력 영상이 배경과 객체 두 개로 나누어 졌다면, 화소들은 그레이레벨 $[1,\cdots,t]$에 속하는 클래스 C_1과 그레이레벨 $[t+1,\cdots,L]$에 속하는 C_2 클래스로 나눌 수 있다. 따라서 두 클래스에 대한 그레이레벨 확률분포는 $c_1 : P_1/w_1(t)$, $P_2/w_1(t)$ $\cdots$, $P_t/w_1(t)$ 와 $c_2 : P_{t+1}/w_2(t)$, $P_{t+2}/w_2(t)$ $\cdots$, $P_L/w_2(t)$ 이 된다. 여기서

$$(t) = \sum_{i=1}^{t}\ _{i} \tag{3-3}$$

$$w_2(t) = \sum_{i=t+1} P_i \tag{3-4}$$

이고, 클래스 C_1과 C_2의 평균은 각각

$$\mu_1 = \sum_{i=1}^{t} iP_i/w_1(t) \tag{3-5}$$

$$\mu_2(t) = \sum_{i=t+1}^{L} iP_i/w_2(t) \tag{3-6}$$

이다. 또한, 전체 영상의 평균을 μ_T라고 하면,

조선대학교
CHOSUN UNIVERSITY

$$\mu_1 + w_2\mu_2 = \mu \tag{3-7}$$

$$w_1 + w_2 = 1 \tag{3-8}$$

이다. 클래스 과 C_2에 대하여 클래스내의 분산(within-class variance)을 v_w^2, 클래스사이의 분산(between-class variance)을 v_B^2, 그리고 전체 구간에서의 분산(total variance)을 v_T^2 라고 하면, 최적의 임계값은 다음의 판별기준식들 가운데서 하나를 최대화하는 t로 선택할 수 있다.

$$\lambda = v_B^2 / v_w^2 \tag{3-9}$$

$$\lambda = v_T^2 / v_w^2 \tag{3-10}$$

$$\eta = v_B^2 / v_T^2 \tag{3-11}$$

여기서

$$v_w^2 = w_1\sigma_1^2 + w_2\sigma_2^2 \tag{3-12}$$

$$v_B^2 = w_1(\mu_1 - \mu_T)^2 + w_2(\mu_2 - \mu_T) \tag{3-13}$$

$$v_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2 p(i) \tag{3-14}$$

이다. 여기서 σ_1^2과 σ_2^2 은 클래스 C_1과 C_2의 분산이다. 판별기준 식 (3-12)-(3-14)에서 v_T^2는 t의 선택과 무관하므로 η의 계산이 가장 간단하다. 따라서

조선대학교
CHOSUN UNIVERSITY

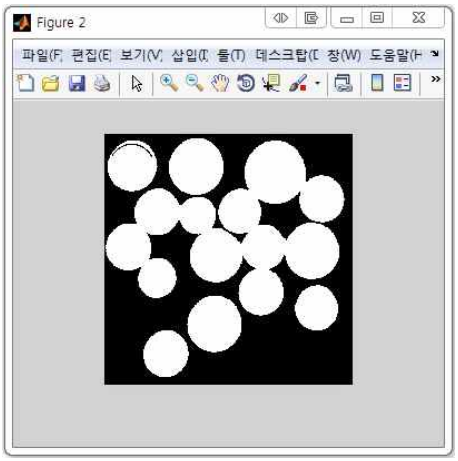
$$(t^*) = \arg max_{1 \le t \le L} \; v_B^2(t) \tag{3-15}$$

를 만족하는 t^*를 최적의 임계값으로 결정할 수 있다.

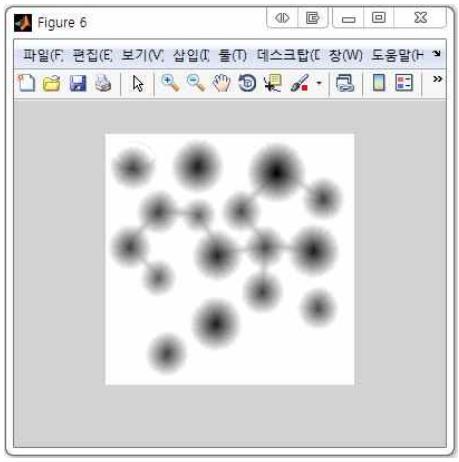
워터쉐드 알고리즘은 수리 형태론(mathematical morphological)을 이용한 영상 분할의 한 방법이다. 영상 화소의 밝기 값에 대한 기울기(gradient) 값을 고도(altitude) 정보로 변환시켜 영역을 구분한다. 이 알고리즘은 담수 영역(catchment basin)을 구분하는 분수령(watershed)을 찾음으로써 각 담수 지역에 결합된 영역을 찾아내는 알고리즘이다.

분할을 위해서 워터쉐드 변환과 함께 자주 사용되는 도구는 거리 변환(distance transform)이다. 이진 영상의 거리 변환은 상대적으로 간단한 개념이다. 각 화소로부터 가장 가까운 영이 아닌 화소까지의 거리이다. 그림 3.5 (a)는 작은 이진 영상 매트릭스를 보여주며, 그림 (b)는 그에 해당하는 거리 변환을 보여준다. 모든 1의 값 화소는 0이 아닌 가장 가까운 화소가 그 자신이므로 거리변환 값이 0임을 주목한다.

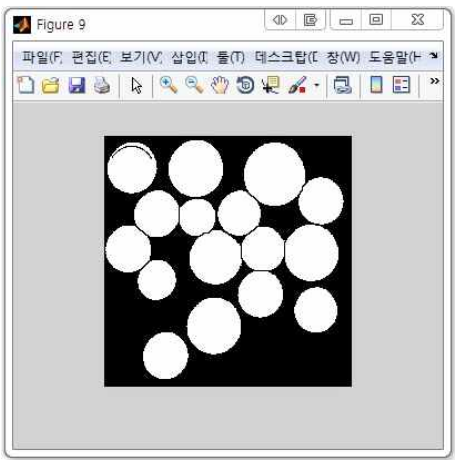
조선대학교
CHOSUN UNIVERSITY



(a) 이진영상



(b) 거리변환



(c) 분할결과

그림 3.5 Otsu 방법 및 워터쉐드 변환을 이용한 이진화 과정

2. 레이블링

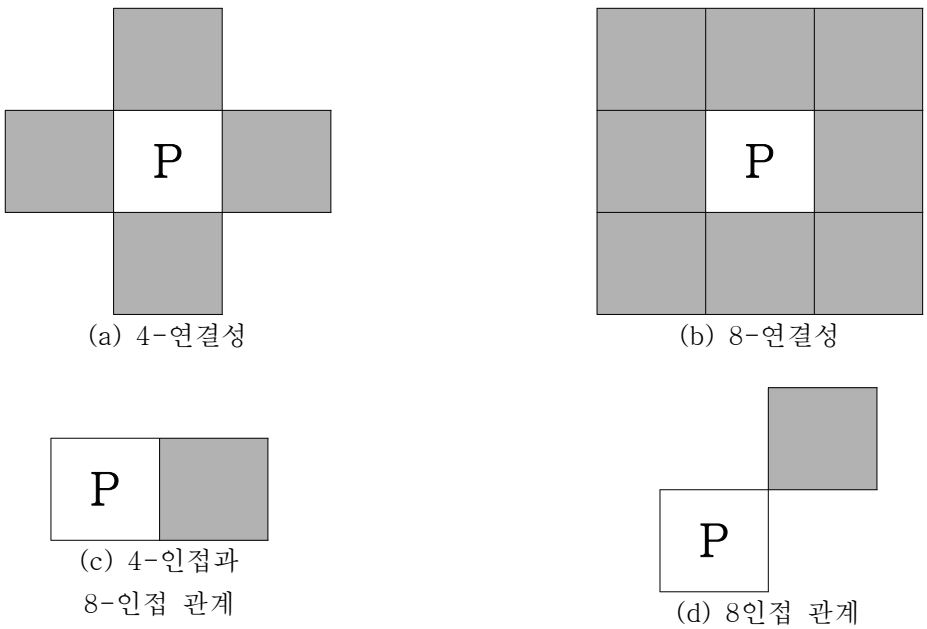
레이블링은 가버 필터링 후 이진화 된 이미지의 모든 화소를 탐색하여 4-연결성(4-connected)이나 8-연결성(8-connected)등으로 연결 관계를 이용하여 번호를 매기는 방법으로 각 객체를 구분하는 방법이다.

레이블링시 4-연결성과 8-연결성은 그림 3.6과 같이 나타낼 수 있다. 이미지상의 (x,y)에 있는 한 화소 p는 좌표가 (x + 1,y), (x - 1,y), (x,y+1), (x,y-1)인 두 개

조선대학교
CHOSUN UNIVERSITY

의 수평 이웃과 두 개의 수직이웃을 갖는데 p의 4이웃 집합을 N4(p)로 표기하고 그림 2.13(a)와 같이 나타낸다. 화소 p의 네 개의 대각선 이웃은 (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)이고 수직 및 수평성분과 함께 p의 8이웃 집합이라 하고 N8(p)와 같이 표기하고 그림 3.6과 같이 나타낸다.

한편, 이미지 상의 두 화소 p와 주변화소가 만일 주변화소 ∈ N4(p)이면 4-인접(4-adjacent)한다고 하고 마찬가지로,p와 주변화소가 만일 주변화소 ∈ N8(p)이면 8-인접(8-adjacent)한다고 한다. 그림 3.6와 같이 4-인접관계는 4-인접과 동시에 8-인접한다고 할 수 있지만 그림 3.6는 4-인접하지 않고 8-인접한다고 할 수 있다.



(a) 4-연결성 (b) 8-연결성

(c) 4-인접과 8-인접 관계 (d) 8인접 관계

그림 3.6 4-연결성과 8-연결성

4-인접관계와 8-인접관계에 따라 레이블링 결과가 달라지는데 4-인접관계를 이용하여 레이블링을 하게 되면 수직과 수평 성분에 관해서만 연결이 유지되므로 그림 3.7 (a)와 같이 레이블링이 1부터 4까지 4단계로 나타내어지고, 8-인접관계를 이용하여 레이블링을 하면 수직, 수평 및 대각선 방향의 화소까지 인접관계에 포함되므로 그림 3.7 (b)와 같이 2단계로 나타내어진다.

조선대학교
CHOSUN UNIVERSITY

				2		
		2	2	2	2	
1	1					
1			3		4	
				4	4	
				4		

(a) 4-연결성

				1		
		1	1	1	1	
1	1					
1			2		2	
				2	2	
				2		

(b) 8-연결성

그림 3.7 4-연결성과 8-연결성을 이용한 레이블링

3. 불변 모멘트 특징벡터 추출

1) 레이블링 영역 형태 정보 추출

영상이 레이블링으로 영역이 분할되면 각각의 레이블링에 대한 특징 정보를 추출한다. 하나의 이미지는 수많은 레이블링으로 구성되므로 모든 레이블링에 대한 특징 정보 추출은 오버 헤드가 발생할 가능성이 높다. 따라서 이 논문에서는 레이블링 된 영역 중에서 크기 비율 10위 이내의 레이블링에 대해서만 특징 정보를 추출한다. 형태에 대한 특징 정보로는 이동, 회전, 크기 변화에 강인한 불변 모멘트(Invariant Moment)를 사용한다.

2) 불변 모멘트

일반적으로 2차원의 연속 함수 f(x, y)에 대한 (p+q)차수의 모멘트는 식 (3-16)과 같다.

$$_{x)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \quad , \quad (p,q=0,1,2,\ldots) \qquad (3\text{-}16)$$

조선대학교
CHOSUN UNIVERSITY

식 (3-16)은 디지털 영상에서 식 (3-17)과 같이 정의된다.

$$_{q} = \sum_{x}\sum_{y} x^{p} y^{q} f(x,y) \quad , \quad (p,q=0,1,2,\ldots) \tag{3-17}$$

임의의 영역에 대한 함수 (x,y)에 대해서 차수 (p, q)에 대한 중심 모멘트 (central moment)의 정의는 식 (3-18)과 같으며 이동에 불변인 특성을 갖는다.

$$\mu_{pq} = \sum_{x}\sum_{y} (x-x)^{p}(y-y)^{q} f(x,y) \tag{3-18}$$

식 (3-18)에서 인덱스 x, y는 각각 수평 x축, 수직 y축의 위치를 나타낸다. 모멘트 m_{0}은 객체 영역을 나타내며 3차 모멘트까지의 중심 모멘트는 식 (3-19)과 같은 관계를 갖는다.

$$\begin{aligned}
\mu_{00} &= m_{00} = \mu \\
\mu_{10} &= \mu_{01} = 0 \\
\mu_{11} &= m_{11} - \mu xy \\
\mu_{20} &= m_{20} - \mu(\overline{x})^{2} \\
\mu_{02} &= m_{02} - \mu(\overline{y})^{2} \\
\mu_{21} &= m_{21} - m_{20}\overline{y} - 2m_{11}\overline{x} + 2\mu(\overline{x})^{2}(\overline{y})^{2} \\
\mu_{12} &= m_{12} - m_{02}\overline{x} - 2m_{11}\overline{y} + 2\mu(\overline{xy})^{2} \\
\mu_{03} &= m_{03} - 3m_{02}\overline{y} + 2\mu(\overline{y})^{3} \\
\mu_{30} &= m_{30} - 3m_{20}\overline{x} + 2\mu(\overline{x})^{3}
\end{aligned} \tag{3-19}$$

중심 모멘트의 의미는 다음과 같다.

μ_{20} : 수평축 분산

μ_{02} : 수직축 분산

μ_{11} : 수평축과 수직축의 상호분산(Covariance)

조선대학교
CHOSUN UNIVERSITY

$_2$: 수평축에서 오른쪽에 비해 왼쪽으로 퍼져 있는 정도

μ_{21} : 수직축에서 위쪽에 비해 아래쪽으로 퍼져 있는 정도

μ_{30} : 수평축의 비대칭 정도(Skew)

μ_{03} : 수직축의 비대칭 정도(Skew)

또한 정규 모멘트는 중심 모멘트를 일정한 크기의 값으로 나눠준 것으로 식 (3-20)과 같이 정의되며, 이는 크기에 불변인 특성을 부여한다.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{3-20}$$

여기서 $\gamma = \frac{p+q+2}{2}$ 이고, $p+q=0,1,2,3$ 이다.

정규화 된 중심 모멘트를 이용하여 식 (3-21)에 의해서 2차, 3차 정규 모멘트로 구성되고 그 영역의 불변 모멘트를 계산 할 수 있다.

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} + \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{3-21}$$

식 (3-19)에서 정의한 불변 모멘트의 의미는 다음과 같다.

조선대학교
CHOSUN UNIVERSITY

: 수평, 수직 방향 분산의 합. 수직과 수평축으로 많이 퍼져 있을수록 값이 커짐.

ϕ_2 : 수직, 수평축의 분산 값이 비슷할 경우 수직, 수평축에 대한 상호 분산 값.

ϕ_3 : 좌우, 상하로 치우친 값을 강조하는 결과.

ϕ_4 : 좌우, 상하로 치우친 값을 상쇄하는 결과.

ϕ_5, ϕ_6, ϕ_7 : 크기, 회전, 위치에 불변한 특징 값을 추출.

본 연구에서는 특성벡터로 연산 복잡도를 감소하도록 레이블링을 정규화 함으로써 1차~4차 모멘트만을 사용하도록 하였다.

아래의 그림 3.8은 앞에서 설명된 식의 과정을 표현한 그림이다. 레이블링된 객체 영역에 대한 불변 모멘트를 이용하여 특성벡터를 생성하기 된다.

조선대학교
CHOSUN UNIVERSITY

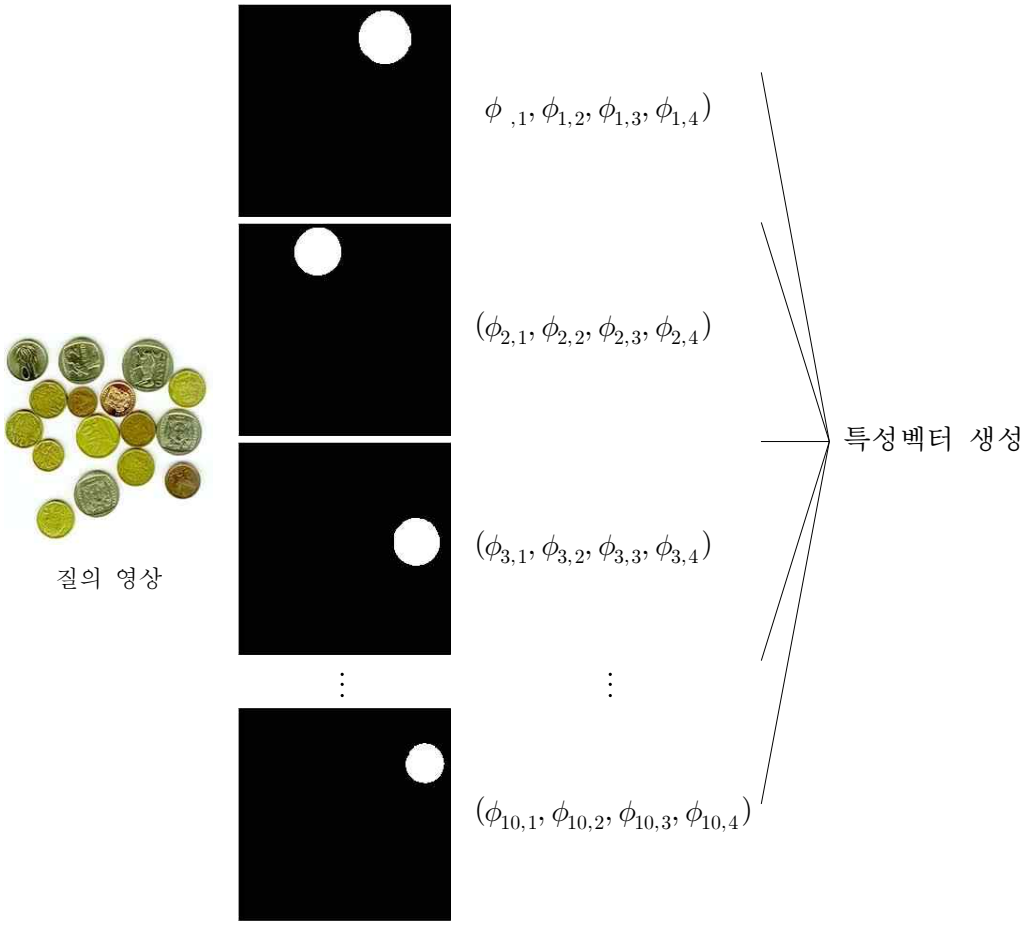


그림 3.8 상위 10개의 대표 레이블링과
그에 해당하는 모멘트를 이용한 특성벡터 추출

3) 영상 변환에 강인한 정규화

본 논문에서는 찌그러짐, 기울기, 크기변환 등으로 영상의 특성이 잡음에 따라 달라질 수 있으므로 정규화 하는 작업이 필요하다. 앞서 추출한 특성벡터에서 영상 변환에 강인한 정규화 벡터를 식 (3-22)에 의해서 추출한다.

$$_{pq} = \frac{^{i+2}X}{^{i}N} \qquad (3\text{-}22)$$

조선대학교
CHOSUN UNIVERSITY

	1	2	3	4
1	0.1612	7.6560e-05	4.3010e-05	1.5877e-06
2	0.1611	1.1945e-04	2.7259e-05	5.4229e-07
3	0.1599	5.6906e-05	1.6793e-05	5.8992e-08
4	0.1597	1.3646e-04	6.6313e-06	1.4940e-08
5	0.1596	4.3296e-05	1.6938e-05	1.2193e-08
6	0.1596	7.8768e-05	3.2401e-06	8.4975e-09
7	0.1595	7.8118e-05	8.1867e-06	9.9510e-09
8	0.1595	8.1437e-05	4.9231e-06	9.9089e-09
9	0.1594	7.4428e-05	1.1823e-06	1.5487e-09
10	0.1594	2.3288e-05	1.0719e-06	5.7343e-09
11	0.1594	6.0334e-05	2.2884e-06	2.1863e-09
12	0.1594	1.4745e-05	7.7614e-06	8.4920e-09
13	0.1594	5.1019e-05	4.1455e-07	1.1075e-09
14	0.1593	1.5746e-05	2.1453e-06	2.7414e-10
15	0.1593	5.1442e-05	2.0208e-06	1.6760e-09
16	0.1593	3.8388e-05	2.6777e-07	2.2842e-10

	1	2	3	4
1	0.1608	8.4304e-05	2.9020e-05	7.2967e-07
2	0.1603	1.0427e-04	1.6894e-05	2.0541e-07
3	0.1598	7.8886e-05	1.3454e-05	2.8708e-08
4	0.1596	8.6173e-05	8.9365e-06	1.1877e-08
5	0.1596	6.6727e-05	9.4550e-06	1.0214e-08

그림 3.9 영상변환에 강인한 정규화

4) 특성 정보 저장

비교 대상이 되는 이미지들을 앞에서 기술한 방법에 의하여 특징 정보들이 추출되어 데이터베이스에 색인되어 있어야 한다. 각 이미지의 특징 정보를 저장하기 위하여 [표 3.1]과 같이 데이터베이스의 테이블을 설계한다.

표 3.1 데이터베이스 테이블 구성

field 1	field 2	field 3	field 4	field 5
특성벡터	정규화벡터	레이블링수	이미지주소	이미지이름
MomentV	NormV	LabelNum	FileWithPath	FileName

4. 유사도 계산

영상 데이터베이스에 저장된 여러 영상들 중에서 사용자가 질의한 영상과 가장 유사한 영상을 검색하는 시스템을 구현하기 위해서는 영상들 간의 유사성 척도(Similarity measure)에 대한 정의와 이를 효율적으로 활용하는 방법이 필요하

조선대학교
CHOSUN UNIVERSITY

다. 영상들 간의 유사성 척도는 특징 벡터로 사용된 영상의 속성이 무엇이냐에 따라서 다양한 방법이 있을 수 있다. 유사성 측정은 두 영상 데이터 사이의 유사성을 수치로 나타낸 것이다. 데이터 분석, 색인, 검색 전반에 걸쳐서 밀접한 관계가 있으며, 내용기반 검색 성능 전체에 영향을 주는 핵심기술이라 할 수 있다.

본 논문에서는 질의 영상과 데이터베이스 영상 사이의 내용기반 영상검색의 시각 유사성을 계산하였다. 따라서, 검색 결과는 질의 영상과 함께 영상의 유사성에 의해 배열하게 되는 영상의 단일 영상이 아니라 유사도 순으로 뿌려주는 검색 리스트이다. 많은 유사성 측정은 최근 특징 분포의 경험적인 추정에 의거하여 영상검색을 위해 개발되었으며 각각의 유사성/거리 측정은 영상 검색 시스템의 검색 성능에 현저하게 영향을 미칠 것이다. 본 논문에서는 질의 영상과 데이터베이스의 영상들 간의 형태 특징유사도를 측정하기 위해서 다음과 같은 간단한 거리척도 함수를 이용한다.

$$Q, I) = \sum_{=1}^{k} |F_i - F'| \qquad (3\text{-}23)$$

식 (3-23)에서 Q는 질의 영상, I는 데이터베이스내의 영상이고, F_i와 F_i'는 각각 두 영상의 특징 벡터이다. 영상들 간의 유사도는 두 특징 값들 사이의 절대적인 거리를 계산함으로써 쉽게 얻을 수 있고 계산된 유사도에 따라 내용 기반 영상 검색 시스템에서 질의 영상과 유사한 영상들이 영상 데이터베이스로부터 검색된다.

5. 성능평가 척도

내용 기반 영상 검색의 효율성을 분석하기 위해, 일반적으로 Recall과 Precision의 두 가지 성능 평가 척도를 이용한다. Recall은 영상 데이터베이스에서 질의와 관련된 영상 중 검색된 영상의 비율이고, Precision은 검색된 영상 중에서 질의와 관련된 영상의 비율을 나타낸다. 즉, A를 영상 데이터베이스 내에서 관련된 영상의 집합이고, B는 검색된 영상의 집합이라고 하자. 그러면 Recall과 Precision은 다음 아래의 식과 같은 조건부 확률로 정의할 수 있다.

조선대학교
CHOSUN UNIVERSITY

$$call = P(\frac{B}{A}), \qquad \mathrm{P}recision = P(\frac{A}{B}) \tag{3-24}$$

실제적인 실험에서는 다음 식 (3-25)와 같은 산출식을 사용한다.

$$R = \frac{R_r}{T}, \quad P_r = \frac{R_r}{T_r} \tag{3-25}$$

여기서 T는 검색대상 데이터베이스에서 질의와 관련된 항목의 총수를 나타내고, R_r은 검색된 항목 중에서 질의와 관련된 항목의 수, T_r은 검색된 항목의 총수를 나타낸다. Recall이 크다는 것은 매칭도가 높다는 것을 뜻하며, 반대로 Precision이 높다는 것은 잘못 검출된 영상이 많지 않다는 것을 뜻한다.

조선대학교
CHOSUN UNIVERSITY

Ⅳ. 실험 및 분석

1. 실험 환경

본 논문에서 제안한 알고리즘은 다양한 영상에 대하여 MATLAB 2014a 소프트웨어를 사용하여 실험하였다. 본 실험은 질의 영상이 모두 같은 상태에서 분석된다. 영상 검색 방법을 설명하기 위해 데이터베이스 영상을 각각의 영상 집합 마다 50~1000개를 놓고 비교되는 영상에 따라 20개를 찾도록 정의하여 놓았다. 실험에서 질의 영상과 비슷한 영상은 알맞은 영상의 개수로 만들어놓았다. 저장해 놓은 데이터는 나중에 다른 질의 영상을 입력 시켰을 때 따로 계산을 하지 않고 질의 영상만 계산되고 데이터의 특징 값들을 이용해 유사도 측정을 비교하여 결과 영상을 도출하게 된다.

2. 시뮬레이션

그림 4.1의 실험 영상은 본 시뮬레이션을 구현하기 위한 실험 영상으로 사용되었다. 먼저 본 알고리즘들의 이해를 돕기 위해 데이터베이스를 그림 4.1과 같이 실험 영상을 30개로 정한 후 30개의 영상에서 질의 영상을 선택한다. 선택된 질의 영상은 본 알고리즘들을 이용하여 데이터베이스 영상들과 비교되어 검색 결과를 나타낸다. 검색 결과는 데이터베이스 30개중에 검색 결과를 20개로 나타내었으며 유사도 순으로 출력 되도록 하였다.

조선대학교
CHOSUN UNIVERSITY

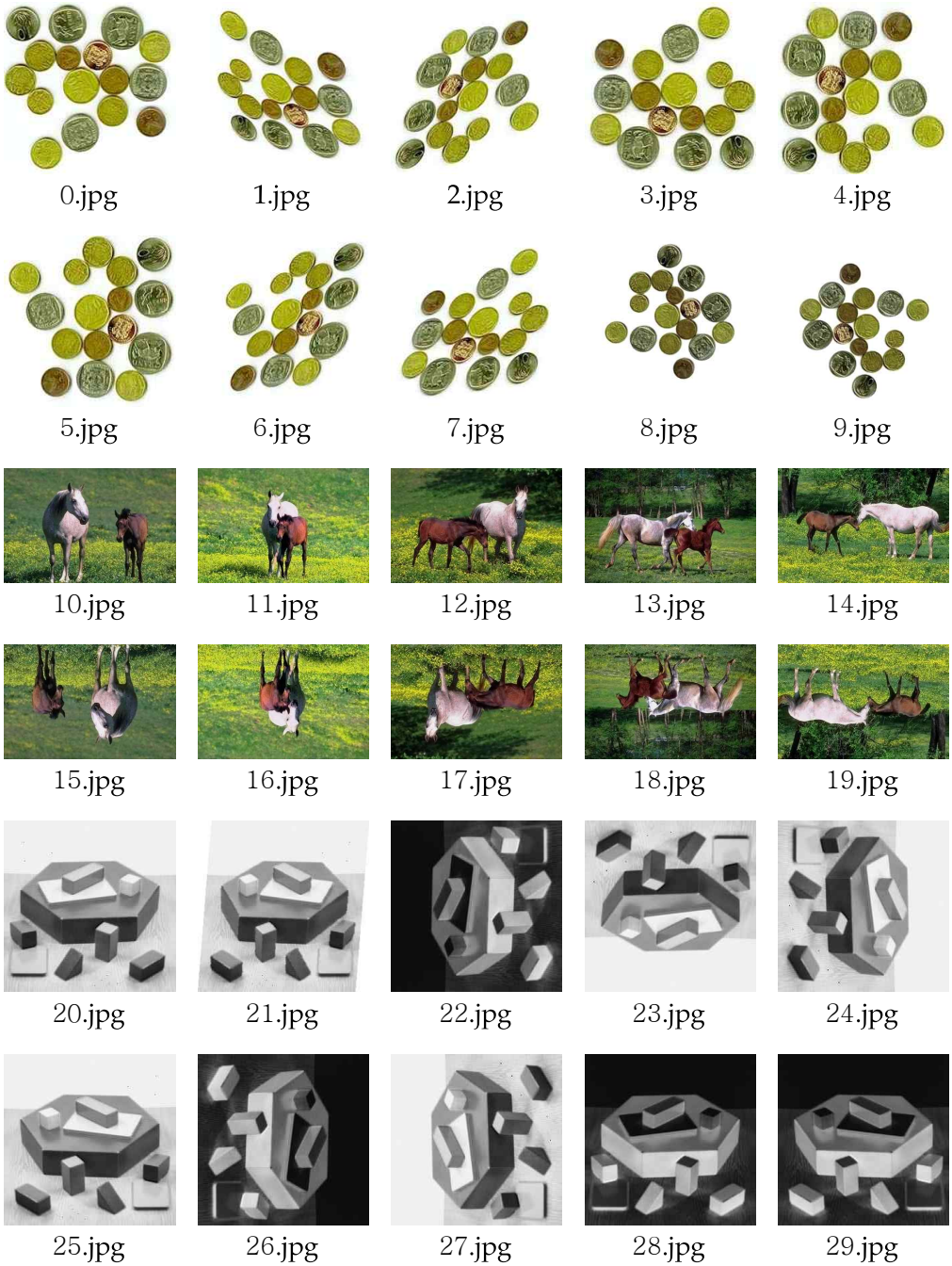


그림 4.1 실험 영상

먼저 그림 4.1을 실험 영상으로 두고 그림 4.1의 20.jpg를 질의 영상으로 정한

조선대학교
CHOSUN UNIVERSITY

후 실험을 수행하였다. 그림 4.2는 질의 영상이다.

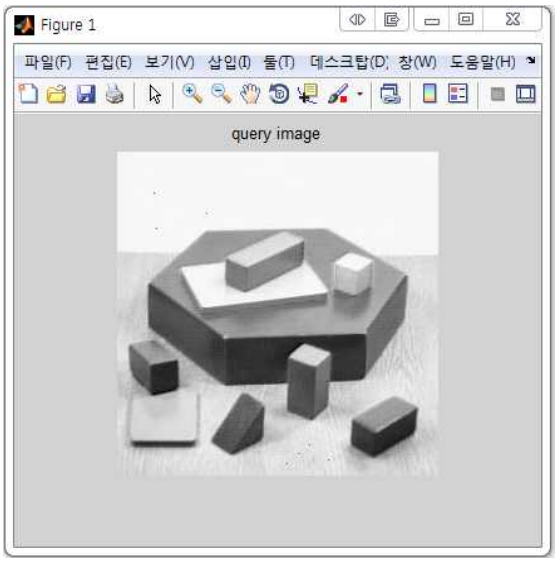


그림 4.2 질의 영상 20.jpg

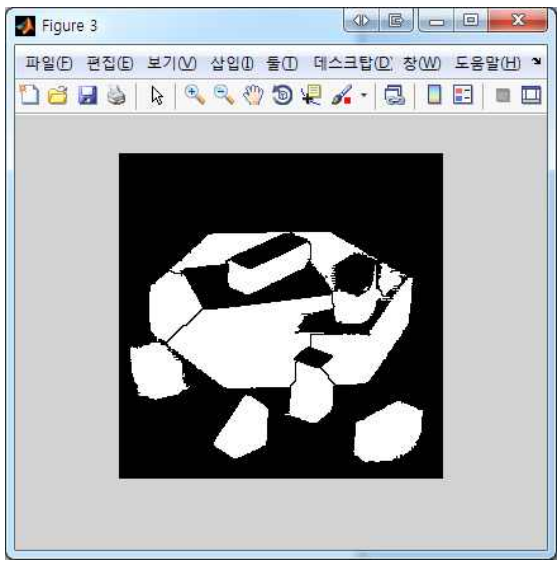


그림 4.3 2진화 영상

조선대학교
CHOSUN UNIVERSITY

	1	2	3	4
1	0.4423	0.0421	0.0142	0.0167
2	0.3064	0.0020	0.0259	4.0060e-04
3	0.2545	0.0315	0.0013	1.0753e-04
4	0.2443	0.0133	0.0058	5.3333e-04
5	0.2383	0.0217	2.3678e-04	4.4487e-05
6	0.2140	0.0036	4.9422e-04	1.0369e-04
7	0.1864	0.0025	7.2033e-04	5.7414e-06
8	0.1864	0.0047	7.5975e-04	3.5017e-05
9	0.1825	0.0053	1.1800e-04	9.3472e-06
10	0.1776	0.0041	2.6199e-04	6.0142e-06

그림 4.4 질의 영상의 특징벡터 추출

	1	2	3	4
1	0.7215	0.3640	0.1180	0.0087
2	0.5286	0.1561	0.0731	0.0058
3	0.4326	0.1294	0.0072	0.0038
4	0.3629	0.0789	0.0126	0.0066
5	0.3414	0.0537	0.0118	0.0059

그림 4.5 질의 영상의 정규화벡터 추출

그림 4.4에서 보듯이 불변모멘트를 이용하여 특징값들이 추출된다. 그림 4.5에서는 추출된 특징벡터를 이용하여 정규화 벡터를 생성한다.

그림 4.6에서는 제안한 알고리즘을 이용한 특징벡터의 유사도 값들을 그래프로 나타내었다. 그림 4.6의 순위는 데이터베이스 영상의 번호에 맞추어 유사도의 순위가 정해졌다.

조선대학교
CHOSUN UNIVERSITY

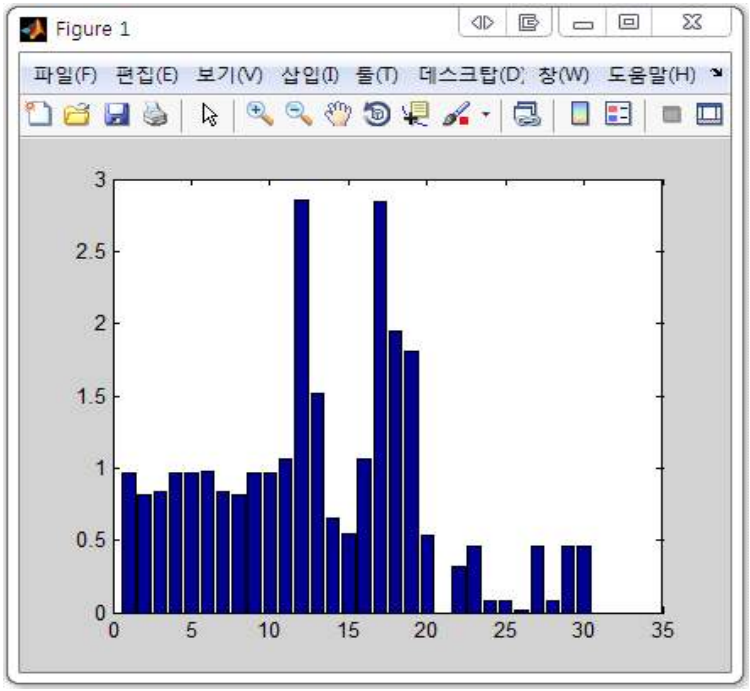


그림 4.6 특징벡터를 이용한 영상 번호별 유사도 측정 결과

마찬가지로 그림 4.7에서도 그림 4.6와 같은 방식으로 나타내었지만 30개의 영상검색은 유사도 측정값의 순위대로 재배열되어 영상검색이 이루어졌다.

조선대학교
CHOSUN UNIVERSITY

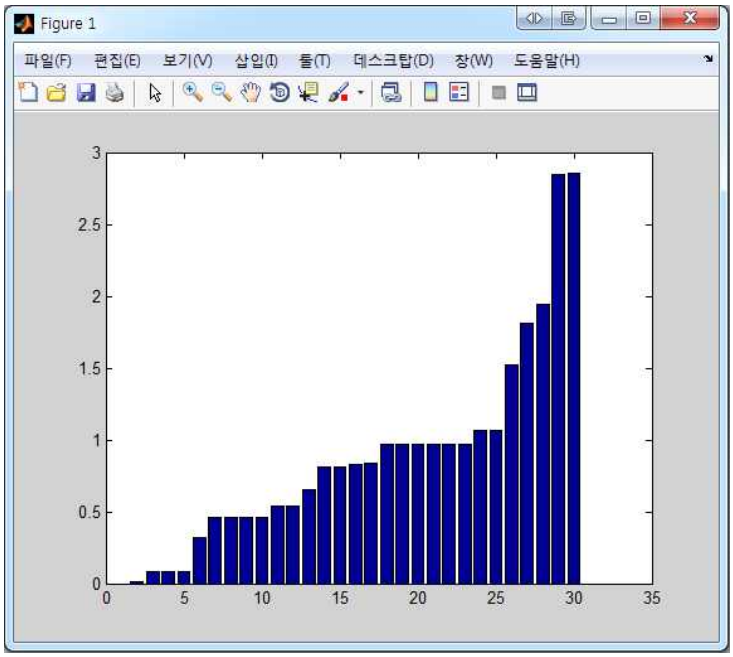


그림 4.7 특징벡터를 이용한 유사도 순위별 측정 결과

그림 4.8에서 보면 질의 영상을 가장 먼저 찾는 것을 확인할 수 있다. 그 다음의 영상들은 질의 영상들과 비슷한 영상들의 순으로 찾아 졌다. 30개의 데이터베이스 영상 중 6개의 유사 영상이 유사도 순으로 찾아지는 것을 확인 할 수 있다. 표 4.1의 오차율은 유클리디언 오차거리를 이용하여 질의 영상과 유사 영상들 간의 특징값의 오차를 나타낸 것이다. 유사도 순으로 나열되었기 때문에 오차율의 순서 역시 유사도 순서와 같게 나온다. 영상마다 유사 영상 값 오차율의 임계값이 다르므로 각 영상의 정보에 따라 가중치를 주어야 한다.

조선대학교
CHOSUN UNIVERSITY

표 4.1 영상의 오차율

Image Number	20	25	27	24	23
Relative error	0.000	0.0110	0.0818	0.0829	0.0845
Image Number	21	29	28	26	22
Relative error	0.3163	0.4583	0.4583	0.4583	0.4583
Image Number	19	14	13	7	1
Relative error	0.5391	0.5404	0.6500	0.8127	0.8152
Image Number	2	6	9	8	0
Relative error	0.8328	0.8412	0.9667	0.9667	0.9694

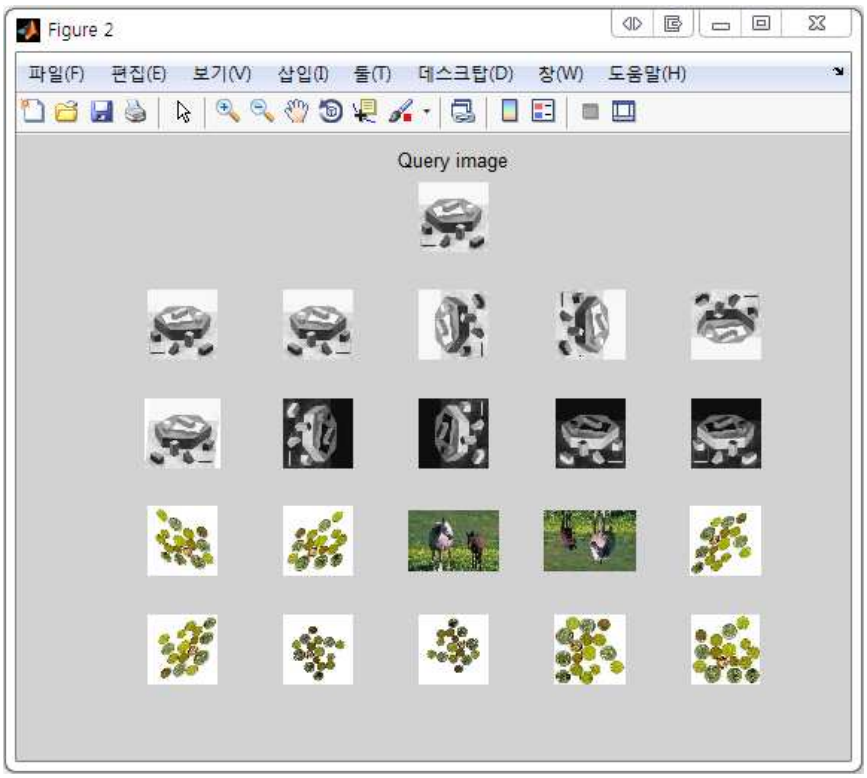


그림 4.8 유사도 값에 의한 영상 검색 결과

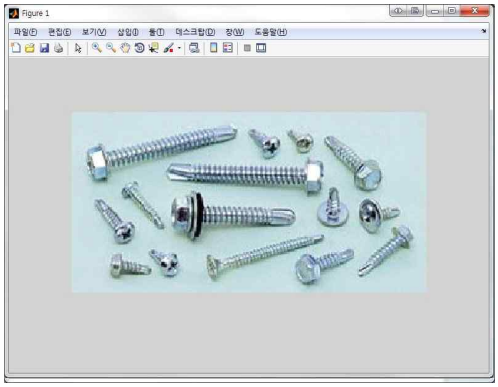
조선대학교
CHOSUN UNIVERSITY

3. 분석 및 고찰

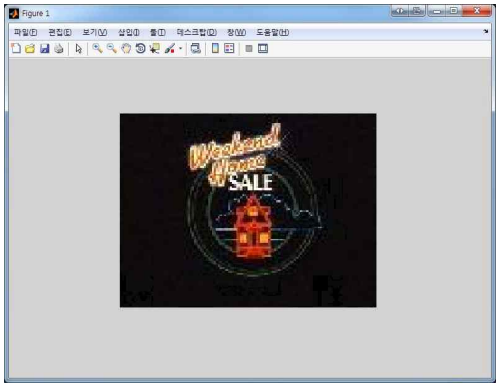
본 논문에서는 제안된 알고리즘의 성능 비교를 위해 Jain과 Vailaya가 제안한 에지 히스토그램을 이용한 방법과 Lu와 Chung 등이 제안한 웨이블릿을 근거한 모양 특징 정보 추출 방법과 비교하고 실험 결과를 제시한다. 본 실험은 질의 영상이 모두 같은 상태에서 분석된다. 영상 검색의 정확성을 실험하기 위해 각 질의 영상을 회전(R), 스케일(S), 노이즈(Z) 변형시킨다. 본 논문에서 제안한 알고리즘들의 성능 측정을 위해 그림 4.9의 질의 영상 Q1~Q4를 갖는 기존 알고리즘과 제안한 알고리즘의 검색 결과의 성능 측정을 비교하기 위해 실험에 사용하였다. 질의 영상을 4개로 사용하였으며 각각의 질의 영상 Q1~Q4는 3000개의 데이터베이스 안에 질의 영상에 맞는 유사 영상을 가지고 있다. 제안한 내용기반 영상 검색 시스템 평가를 위하여 기존의 알고리즘들을 이용하여 제안한 시스템과 Precision과 Recall을 서로 비교 평가하였다.

조선대학교
CHOSUN UNIVERSITY

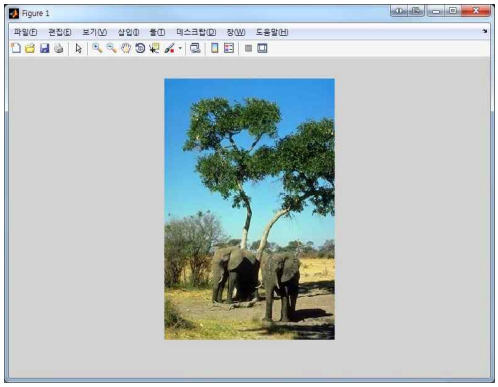
Q1



Q2



Q3

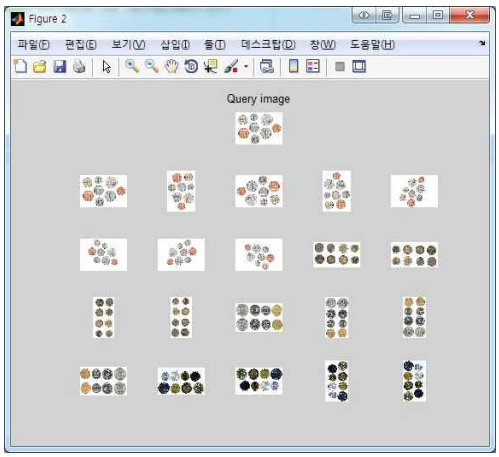


Q4

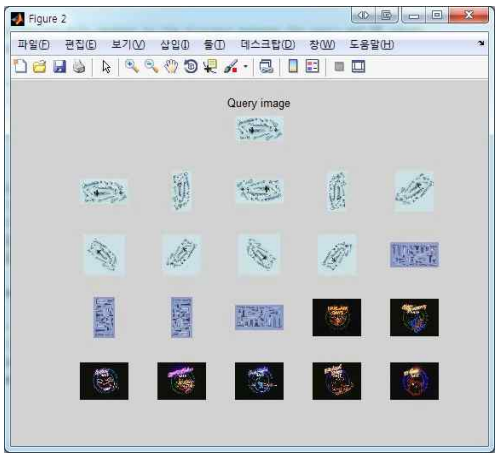
그림 4.9 질의 영상 Q1~Q4

먼저 질의 영상 Q1~Q4를 이용하여 기존 알고리즘들과 제안한 알고리즘의 검색 성능을 비교해 보면 그림 4.10, 그림 4.11과 같다.

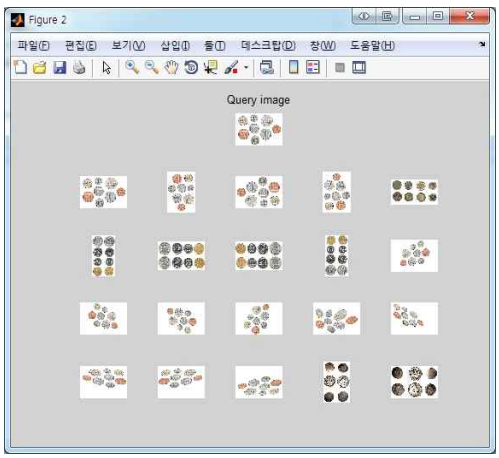
조선대학교
CHOSUN UNIVERSITY



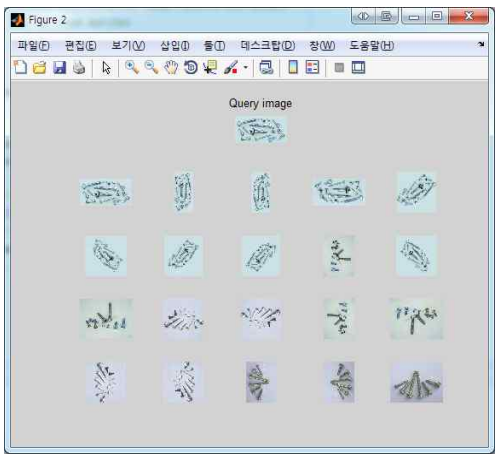
Q1의 에지 히스토그램 알고리즘



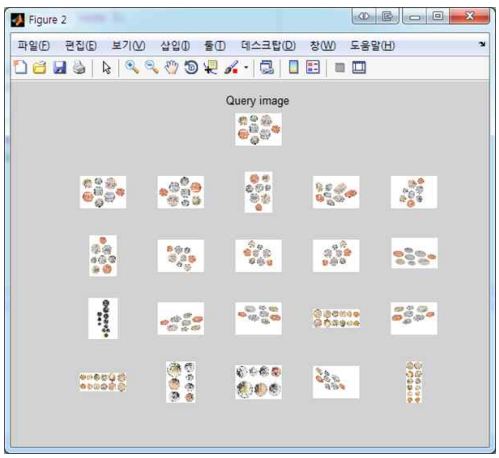
Q1의 에지 히스토그램 알고리즘



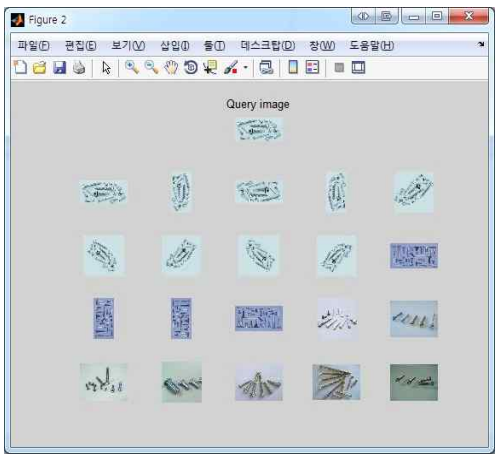
Q1의 웨이블릿 모양 추출 알고리즘



Q2의 웨이블릿 모양 추출 알고리즘



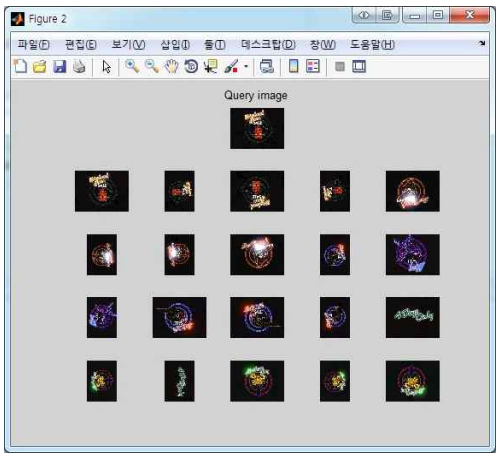
Q2의 제안한 알고리즘



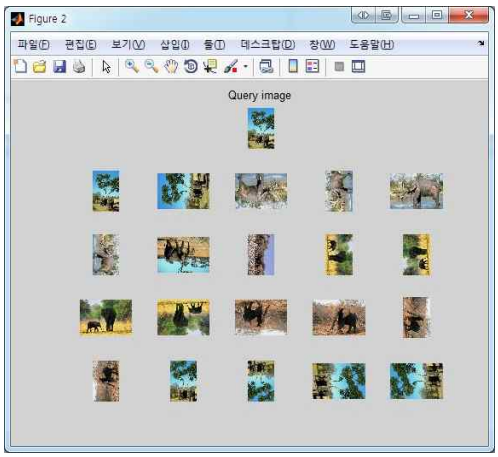
Q2의 제안한 알고리즘

그림 4.10 Q1, Q2의 기존 알고리즘과 제안한 알고리즘의 검색 결과

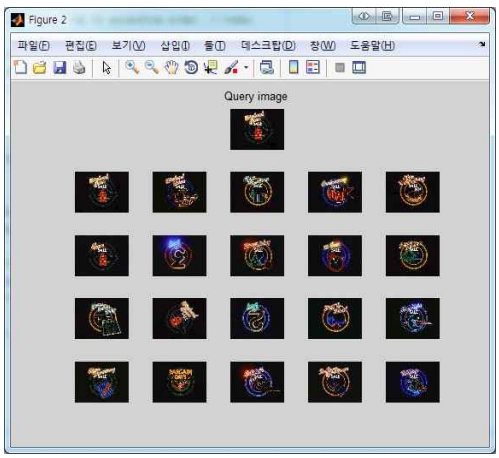
조선대학교
CHOSUN UNIVERSITY



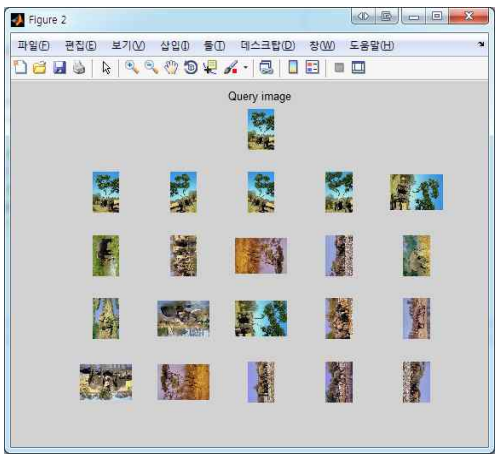
Q3의 에지 히스토그램 알고리즘



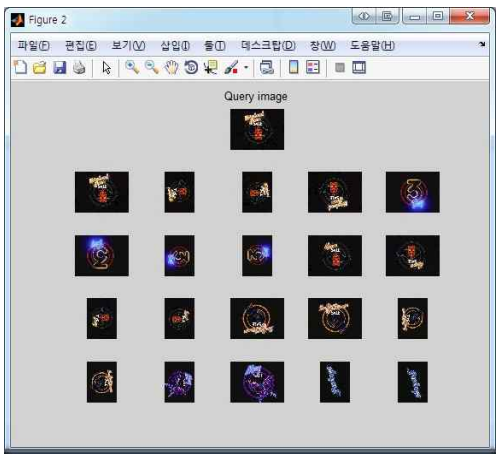
Q4의 에지 히스토그램 알고리즘



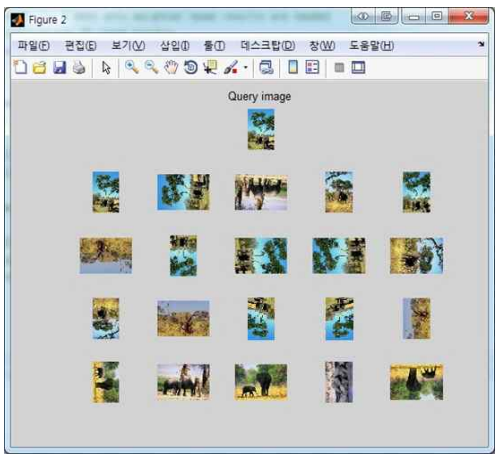
Q3의 웨이블릿 모양 추출 알고리즘



Q4의 웨이블릿 모양 추출 알고리즘



Q3의 제안한 알고리즘



Q4의 제안한 알고리즘

그림 4.11 Q3, Q4의 기존 알고리즘과 제안한 알고리즘의 검색 결과

조선대학교
CHOSUN UNIVERSITY

그림 4.10과 그림 4.11에서와 같이 기존의 에지 히스토그램을 이용한 알고리즘보다 제안한 알고리즘이 육안 상 회전이나 확대 및 축소된 영상을 더 잘 찾아내는 것을 확인할 수 있었다. 그러나 질의 영상 Q1과 같이 객체가 뚜렷한 영상에 비해 질의 영상 Q4와 같은 영상은 기존 에지 히스토그램 알고리즘이 다른 영상이 검색되는 순위가 낮은 반면 제안한 알고리즘은 다른 영상을 먼저 찾아내는 것을 확인할 수 있었다. 그러나 표 4.2과 같이 검색 성능을 Recall과 Precision을 이용하여 비교해 보면 두 알고리즘 모두 Recall은 0.60, Precision은 0.3 으로 결과적으로 같은 검색 성능을 보이는 것을 확인할 수 있었다.

표 4.2 기존 알고리즘과 제안한 알고리즘의 성능 분석

Method / Simulation	에지 히스토그램		웨이블릿 모양 추출		제안한 알고리즘	
	Recall	Precision	Recall	Precision	Recall	Precision
Image Q1	0.32	0.4	0.52	0.65	0.56	0.7
Image Q2	0.9	0.45	0.9	0.45	0.9	0.45
Image Q3	0.26	0.2	0.2	0.15	0.53	0.4
Image Q4	0.6	0.3	0.6	0.3	0.9	0.45
평균	0.52	0.34	0.56	0.39	0.72	0.50

표 4.3은 하나의 객체만 존재하는 단일객체 영상들만을 대상으로 했을 때의 실험결과로서, 에지 히스토그램을 이용한 방법과 웨이블릿을 근거한 모양 특징 정보 추출 방법과 제안된 시스템의 검색 결과를 비교한 것이다. 이들 각각의 특징으로 검색한 결과를 살펴보면 제안된 시스템은 Recall 0.21, Precision 0.8 로 높은 검색 정확도를 나타낸 반면, 에지 히스토그램에서는 Recall 0.15, Precision 0.56, 웨이블릿 모양 추출에서는 Recall 0.19, Precision 0.7로 나타내었다.

조선대학교
CHOSUN UNIVERSITY

표 4.3 단일 객체 영상들의 질의 영상에 따른 실험 결과

Simulation / Method	DB size		T	R_r	Recall	Precision
에지 히스토그램	300	220	60	34	0.15	0.56
웨이블릿 모양 추출	300	220	60	42	0.19	0.7
제안한 알고리즘	300	220	60	48	0.21	0.8

표 4.4는 단일 객체를 갖는 영상들뿐만 아니라 다중 객체를 갖는 영상들도 포함한 전체 영상을 가지고 실험한 결과를 나타내고 있다. 여기서도 표 4.3에서와 같은 조건을 주고 에지 히스토그램을 이용한 방법과 웨이블릿을 근거한 모양 특징 정보 추출 방법과 제안된 시스템간의 검색 결과를 보여주고 있다. 결과 값을 비교해 보면 에지 히스토그램에서는 Recall 0.13, Precision 0.5, 웨이블릿 모양 추출 방법에서는 Recall 0.17, Precision 0.63으로 단일 객체 영상보다 다중 객체에서 검색 효율이 떨어지는 걸 확인할 수 있다. 반면, 제안된 시스템의 Recall과 Precision은 0.23과 0.85로 높은 검색효율을 나타내고 있다. 이러한 결과는 본 논문에서 제안하고 적용한 정규화 레이블링 불변모멘트 알고리즘을 통해 다중 객체를 잘 표현했기 때문에 가능했다. 또한 불변 모멘트의 특성을 이용하여 다양한 형태의 객체를 효과적으로 표현하였기 때문에 기존의 알고리즘보다 높은 정확도를 얻을 수 있었다.

표 4.4 다중 객체 영상들의 질의 영상에 따른 실험 결과

Simulation / Method	DB size	T	T_r	R_r	Recall	Precision
에지 히스토그램	300	220	60	30	0.13	0.5
웨이블릿 모양 추출	300	220	60	38	0.17	0.63
제안한 알고리즘	300	220	60	51	0.23	0.85

조선대학교
CHOSUN UNIVERSITY

Ⅴ. 결론

본 연구에서는 형태정보의 속성으로 정규화 레이블링 불변모멘트 기반으로 다중 객체영상에서의 형태검출을 하는 간단하면서도 효율적인 검색 기법을 제안하였다. 동작 순서는 질의 영상이 주어지면 영상 내 객체를 레이블링하여 최대 면적부터 재배열을 하고 불변모멘트를 이용하여 특징벡터를 추출한다. 추출한 특성벡터에서 영상 변환에 강인한 정규화 벡터를 생성한다. 이로 얻은 특징벡터 값을 이용하여 유사도를 측정하게 된다.

제안된 알고리즘은 시뮬레이션을 통하여 알고리즘의 간단함과 함께 검색의 효율성을 입증하였다. 결과값을 비교해 보면 에지 히스토그램을 사용하는 방식에 비해 Recall과 Precision을 비교해 본 결과 기존 알고리즘 보다 다중 객체 영상에서 Recall은 0.1, Precision은 0.35 높은 것으로 보아 검색 성능이 더 좋은 것으로 확인되었다. 또, 웨이블릿 모양 추출 알고리즘에 비해 Precision이 0.22 높은 것으로 확인되었다.

이와 같이 다중 객체 레이블링 보정과 속성 재배열 기법은 급속도로 발전하고 있는 멀티미디어 데이터의 다양한 정보를 추출함으로써 형태를 기반으로 하는 영상검색에서 정확도와 검색시간 단축 및 최적의 데이터베이스를 구성하고 검색할 수 있는 시스템이나 앞으로 내용 기반 검색 시스템 구축을 위한 기초 자료로 활용할 수 있을 것이라 생각한다.

조선대학교
CHOSUN UNIVERSITY

【참고문헌】

[1] Y. Rui and T. S. Huang, "Image retrieval: current techniques, promising directions, and open issues," J. of Visual Communication and Image Representation, vol. 10, pp. 39-62, 1999.

[2] A. W. M Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[3] S. Michael, "Next-generation web searches for visual content," IEEE Computer, pp. 46-52, Nov. 2000.

[4] Y. H. Ang, Zhao Li and S. H. Ong, "Image Retrieval based on Multimensional Feature Properties", The International Society for Optical Engineering(SPIE), vol. 2420, pp. 47-57, 1995.

[5] R. Brunelli and O. Mich, "Image Retrieval by Examples", IEEE Trans. on Multimedia, vol. 2, no. 3, pp. 164-171, Sep. 2000.

[6] A. Yoshitaka and T. Ichikawa, "A Survey on Content-Based Retrieval for Multimedia Databases", IEEE Trans. on Knowledge and Data Engineering, vol. 11, no. 1, pp. 81-93, 1999.

[7] J.K. Wu, "Content-based Indexing of Multimedia Databases", IEEE Trans. on Knowledge and Data Engineering, Vol.9, No.6, pp.978-989, 1997.

[8] M. J. Swain and D. H. Ballar, "Color index", Int. J. Computer Vision, vol. 7, no, 1, pp. 11-32, July 1991.

[9] Gross M. H. and Koch R., "Visualization of multidimensional shape and texture features in laser range data using complex-valued Gabor wavelets", IEEE Transactions on Visualization and Computer Graphics, vol. 1, pp. 44-59, March 1995.

[10] A. K. Jain and A. Vailaya, “Image retrieval using color and shape,” Pattern Recognition, vol. 29, no. 8, pp. 1233-1244, 1996.

[11] A. D. Bimbo and P. Pala, “Visual image retrieval by elastic matching of user sketches,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, no. 2, pp. 121-132, 1997.

[12] Willian I. Grosky, Peter Neo, and Rajiv Mehrotra, "A Pictorial Index

조선대학교
CHOSUN UNIVERSITY

Mechanism for Model-based Matching", Data and Knowledge Engineering, vol. 8, pp. 309-327, 1992.

[13] Jong-An Park, Young-Eun An, and Min-Hyuk Chang, "2nd Chain code method based on the maximum histogram for shape retrieval", JKIIT, vol. 8, no. 11, pp. 239-245, Dec. 2010.

[14] Seon-Jong Kim, and Young In Kim, "An Extraction of Region-based Information for Image Retrievals", JKIIT, vol. 3, no. 4, pp. 31-37, Aug. 2005.

[15] Juntaek Oh, Hyunwook Kwak, and Wookhyun Kim, "Recognition of Traffic Signs using Wavelet Transform and Shape Information", IEIE, pp.125-134, 2004.

[16] Gwang-won Kang, Ji-min Lee, and Jong-an Park, "Image retrieval using the maximum sequence of the rearranged feature vector based on object corners", JKIIT, vol. 8, no. 12, pp. 247-254, Dec. 2010.

[17] Taek Jin Kwon, "Content-Based Image Retrieval Using Shape and Texture Feature", Master thesis, The Graduate School Kangwon National University, pp. 3-6, 2001.

[18] K. Hirata and T. Kato, "Query by Visual Example", Proc. EDBT conference, pp56-71, 1992.

[19] Seong-Hee Kim, "A Study on the Performance Analysis of Content-based Image & Video Retrieval Systems", Journal of The Korean Biblia Society for Library and Information Science, vol. 15 no. 2, pp. 97-116, 2004.

[20] Hyun-Sool Kim, "Study on the Content-based Image Retrieval using Color, Shape and Texture Information", Doctoral dissertation, The Graduate School Yonsei University, 2001.

[21] Ji-Min Lee, "CBIR based on the Rearrangement of Corner Patch Feature", Master thesis, The Graduate School Chosun University, 2010.

[22] Chun-Hua Huang, "Content-based Image Retrieval using Color Histogram and Multi-texture Analysis", Master thesis, The Graduate School Soongsil University, 2011.

[23] T. Gevers and A. Smeulers, "A Comparative Study of Several Colour Models for Colour Image Invariant Retrieval", Proc. of First Int.

조선대학교
CHOSUN UNIVERSITY

Workshop on Image Database and Multimedia Search, Amsterdam, pp.17-27, Aug. 1996.

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

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

[26] H. Wechsler, "Texture Analysis - A Survey", Signal Processing. vol. 2, no. 3. pp. 271-282, 1980.

[27] J. Y. Hsiao and A. A. Sawchuk, "Texture Feature Performance for Image Segmentation", Pattern Recognition, vol. 23, no. 3, pp. 291-309, 1990.

[28] J. M. Coggins, "A Framework for Texture Analysis Based on Spatial Filtering", University Microfilms International, 1983.

[29] C. H. Chen, L. F. Pau, and P. S. P. Wang, "Handbook of Pattern Recognition and Computer Vision", World Scientific Publishing Co., 1998.

[30] R.M Haralick, "Statistical and Structural Approaches to texture", Proc. IEEE, vol. 67, no. 5, pp. 786-804, May 1979.

[31] T. Randen, "Filter and Filter Design for Image Recognition", Ph.D thesis, Norwegian Univ., Oct. 1997.

[32] T. Randen and J. H. Husoy, "Filtering for Texture Classification: A Comparative Study", IEEE Trans. on Patt. 43Anal. Machine Intell., vol. 21, no. 4, pp. 291-310, Apr. 1999.

[33] T. Chang and C. j. Kuo, "Texture Analysis and Classification with Tree-Structured Wavelet Transform", IEEE Trans. on Image Proc., vol. 2, no. 4, pp. 429-441, 1993.

[34] M. Unser and M. Eden, "Multiresolution Feature Extraction and Selection for Texture Segmentation", IEEE Trans. on Patt. Anal. Machine Intell., vol. 11, no. 7, pp. 717-728, July 1989.

[35] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signatures", IEEE Trans. on Patt. Anal. Machine Intell., vol. 15, no. 11,

조선대학교
CHOSUN UNIVERSITY

pp. 1186-1191, Nov. 1993.

[36] Y. Y. Tang, L. H. Yang, J. Liu, and H. Ma, "Wavelet Theory and Its Application to Pattern Recognition", World Scientific, 2000.

[37] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Trans. on Patt. Anal. Machine Intell., vol. 11, no. 7, pp. 674-693, July 1989.

[38] J. C. Goswami and A. K. Chan, "Fundamentals of Wavelets: Theory, Algorithms, and Application", Wiley Inter-science, 1999.

[39] Byung-Gil Ye, "Texture Segmentation Using Wavelet Frame and Spatial Information", Master thesis, Korea Maritime University, 2002.

[40] Nishat Ahmad, "Object Retrieval Approach with invariant Features Based on corner Shapes," IEEE Communication Society, vol. 1, pp. 840-845, 2007.

[41] 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. 486-4971, 1990.

[42] C. Schmid and R. Mohr, "Matching by Local Invariants," Technical report N2644, INRIA, 1995.

[43] Sxhmid and R. Mohr, "Local Gray value Invariant for Image Retrieval," IEEE Trans. on pattern Analysis and Machine Intelligence, vol. 19, no. 5, pp. 550-535, 1997.

[44] A. K. Jain and A. Vailaya, “Shape-based retrieval: A case study with trademark image databases,” Pattern Recognition, vol. 31, no. 9, pp. 1369-1390, 1998.