

윤곽과 선분을 이용한 내용기반 화상정보 검색시스템의 구현

정 원 일[†] · 구 정 현^{††} · 최 기 호^{††}

요 약

본 논문에서는 화상내 객체의 윤곽정보를 얻고 객체내의 선분을 추출하여 이를 인덱싱하고 매칭하는 내용기반 화상정보 검색시스템을 구현하였다. 이를 위해 선분추출 방법으로서 기존의 SLHT(Straight Line Hough Transform)의 단점을 개선시켜 FSLHT(Flexible SLHT)를 제안하고 그 유효성을 입증하였으며, 선분에 의한 화상간의 유사도를 얻기 위해 DP(Dynamic Programming) 알고리즘을 적용하였다. 윤곽특징을 고려할 때 화상의 윤곽이 잡음에 민감하게 반응하는 문제를 보완하기 위하여 일정간격으로 샘플링한 영역의 대표값을 윤곽특징으로 취하였으며, Hough 변환에 의한 선분추출시 시각적인지에 기반한 방향성을 산출하여 변환함으로써 연속성상실 문제와 시간문제를 보완하였다.

The Implementation of Content-based Image Retrieval System Using Contours and Lines

Jeoung Won Ill[†] · Koo Jeong Hyun^{††} · Choi Ki Ho^{††}.

ABSTRACT

In this paper, we implemented the content-based image retrieval system that indexes and retrieves the images by acquiring contour information of images and by extracting lines from the object. For this purpose, we proposed the advanced line extraction algorithm called FSLHT(Flexible SLHT) which covers drawback of SLHT(Straight Line Hough Transform) and applied a DP(Dynamic Programming) algorithm to get a desirable similarity of images by lines. We estimated the contour features as a key value of sampled region to compensate for the problem that image contours are heavily depend on the noise. When performing the Hough transform we calculated the directionality based on the perceptual organization and transformed according to this direction to overcome the problem of time consuming and discontinuity.

1. 서 론

오늘날의 컴퓨터 기술의 발달로 인하여 텍스트뿐
만 아니라 정지화상, 비디오, 오디오와 같은 대규모의

† 정 회 원:대전산업대학교 전자공학과
 †† 정 회 원:광운대학교 컴퓨터공학과
 †† 정 회 원:광운대학교 컴퓨터공학과/신기술연구소
 논문접수:1996년 2월 8일, 심사완료:1996년 4월 1일

디지털 멀티미디어 정보를 손쉽게 획득하고 저장, 조작, 전송할 수 있게 되었다. 그러나 이러한 종류의 정보를 브라우징(browsing)하고, 질의, 검색하는 효율적인 방법은 아직 부족한 실정이다. 현재 화상정보의 관리와 검색을 지원하는 상용화된 데이터베이스 시스템들은 대부분 각 화상정보와 관련된 텍스트나 키워드의 사용을 기본으로 하고 있다. 이 방법은 각각의 화상에 입력된 텍스트정보에 전적으로 의존하며, 화상의 시각적인 속성을 직접 얻어내지는 못한다. 따라서 텍스트 기반 검색은 다음과 같은 몇몇 문제점이 있다[1]. 첫번째, 검색이 단지 키워드에 의존하므로 만약 질의가 초기에 묘사되지 않은 화상속성을 참조하게 되면 검색은 대부분 실패하게 된다. 두번째로, 텍스트나 형태와 같은 시각적인 속성들을 텍스트로서 표현하기란 매우 어렵다. 세번째로, 비록 화상의 모든 유용한 속성을 텍스트로 표현할 수 있다 하더라도 화상의 속성을 일관적이고 보편타당한 단어로 규정할 수는 없다. 즉, 사용되는 키워드는 서로 다른 언어나 문화적인 습성같은 예기치않은 상황에 대처할 수가 없으며, 컬러, 텍스트, 형태와 같은 시각적인 영역에 관하여 단일하고 일관된 텍스트의 표현이 부족하다는 것이다. 이에 따라 키워드와 관계없이 사용자의 시각적인 예를 질의로 하여 화상정보를 검색하는 내용기반 화상정보 검색시스템의 개발이 요구되고 있다.

그동안 화상의 내용을 기반으로 하는 정보검색은, 화상의 칼라, 모양, 질감, 구조등과 같은 특징을 추출하여 검색하는 방법들이 연구되어왔다[2, 3, 4, 5, 6, 7, 8, 9]. 그러나 과거의 이러한 연구들은 그 적용 대상이 국한되어지거나 이론적으로만 제안되어진 실정이다. 이중에서도 특히 화상의 형태에 기반한 검색은 특징을 추출하기가 매우 어렵기 때문에 연구가 미흡하였으며, 화상의 가장 확실하고 간단한 형태인 직선을 특징으로 추출하여 이를 매칭하는 방식의 검색을 생각해 보았다.

SLHT(Straight Line Hough Transform)는 화상으로부터 선분을 추출할 수 있는 알고리즘으로 이는 Duda와 Hart[10]에 의하여 화상을 x, y 평면으로 볼 때 x, y 좌표계상에 있는 하나의 화소를 기준점과 직교하는 선분으로 생각할 수 있으며, 이때, 화소와 기준점간의 거리 r 과 이루는 각도 θ 로 나타내는 r, θ 공간상의 파라미터로의 변환이 이루어졌다.

본 논문에서는 인간이 화상을 식별하는 주된 특징인 형태에 기반한 검색을 위하여, 화상내 객체의 윤곽정보를 얻고 객체내의 선분을 추출하여 이를 인덱스하고 매칭함으로써 화상정보를 검색해 내는 효율적인 방법을 제시하였다. 화상으로부터 선분을 추출해 내는 방법으로 기존의 SLHT를 개선시킨 FSLHT(Flexible SLHT) 알고리즘을 제안하며, 선분들의 유사도에 기반한 화상간의 매칭을 위하여 DP(Dynamic Programming) 알고리즘을 적용한다. 본 논문에서는 윤곽특징과 선분특징을 병합한 검색을 가능하게 함으로써, 사용자는 윤곽에 의한 검색 결과를 판단할 수 있으며 이로 인해 보다 광범위하고 능동적인 검색이 가능한 내용기반 화상정보 검색시스템을 구현한다.

본 논문의 구성을 보면, 2장에서는 내용을 기반으로 하는 정보검색의 소개, 3장에서는 제안된 화상정보 검색시스템에 대하여 서술하였으며, 4장에서는 제안된 방법에 의한 실험과 고찰을 하였고, 5장에서 결론 및 앞으로의 연구과제에 대해 기술하였다.

2. 내용기반 화상정보 검색

본 장에서는 내용기반 화상정보검색의 특징과 이러한 시스템을 이루는 기본요소에 대하여 설명하였다.

2.1 내용기반 정보검색의 특징

화상의 내용을 기반으로 하는 검색은 다음과 같은 특징을 갖는다[1]. 첫번째로, 이는 유사성에 근거한 검색으로써, 사용자는 시스템과 시각적으로 서로 상호작용함으로써 검색된 후보화상중에서 필요한 정보만을 취할 수 있다. 두번째로, 시스템과의 상호작용을 통해 매우 다양한 검색 가능성을 제공한다. 세번째로, 최초의 검색 결과는 사용자가 확인하고 최종선택을 할 수 있는 바람직한 특징을 갖는 화상들의 집합으로서, 멀티미디어 시나리오와 같은 이후의 응용단계에 사용될 수 있는 임시출력이라고 볼 수 있다. 네번째로, 사용자가 최종선택을 할 수 있도록 하기 위하여 여러장의 아이콘화된 후보화상을 동시에 화면에 표현할 수 있어야 한다.

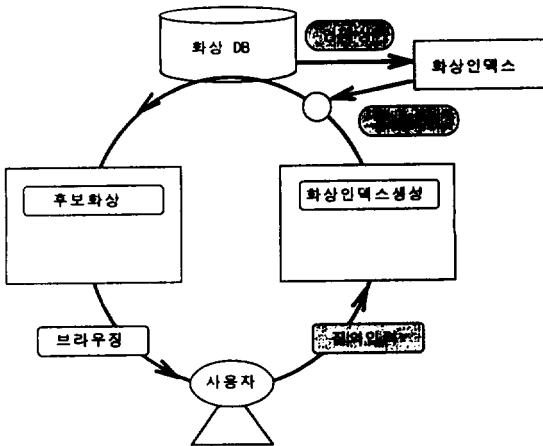
2.2 내용기반 화상정보 검색시스템의 기본요소

내용기반 화상정보 검색시스템이 갖추어야 할 기

본요소는 크게 인덱싱, 질의입력, 매칭, 브라우징 (browsing)의 4가지로 볼 수 있다[7].

- 인덱싱: 검색 대상이 되는 화상정보를 화상의 내부적인 표현(인덱스)으로 저장하는 단계로서, 화상처리단계가 필요하다.
- 질의입력: 사용자가 원하는 질의화상을 그래픽 툴등을 이용하여 손쉬운 형태로 입력한다.
- 매칭: 질의화상과 DB내의 화상을 매칭하는 단계로서 효율적인 매칭법이 요구된다. 시스템은 유사도에 기반한 매칭을 하여 후보화상들을 결정한다.
- 브라우징: 시스템은 사용자가 최종선택을 할 수 있도록 유사도에 따른 검색 결과를 보여줄 수 있어야 한다.

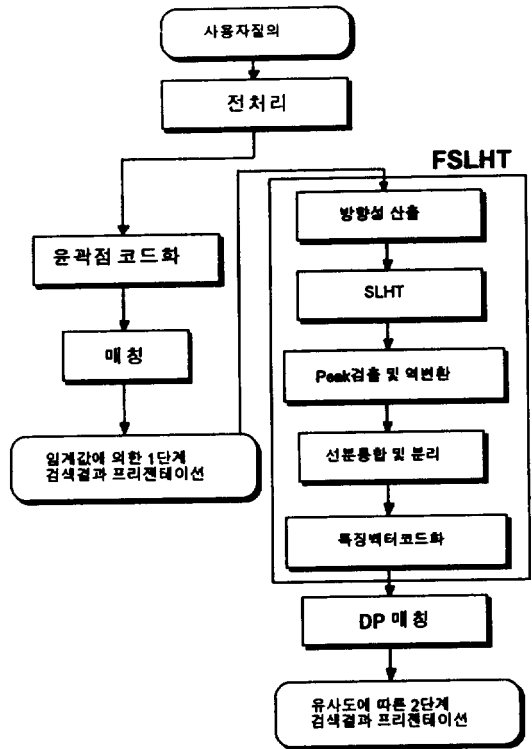
그림 1은 내용기반 화상정보 검색시스템을 나타낸다.



(그림 1) 내용 기반 화상정보 검색시스템의 기본요소 (Fig. 1) Basic elements of content based image retrieval system

3. 제안된 내용기반 화상정보 검색시스템

본 논문에서는 다양한 형태를 갖는 화상들을 대상으로 하여 사용자가 객체의 형태에 관련된 질의를 하였을때 이에 기반하여 화상을 검색할 수 있는 화상정보 검색시스템을 구성한다.



(그림 2) 제안된 내용기반 화상정보 검색시스템의 구성도 (Fig. 2) Block diagram of content based image retrieval system

그림 2에서는 제안된 화상정보 검색시스템의 구성도를 나타내었으며, 각 질에서 크게 전처리, 윤곽점코드화단계, 선분추출단계, 매칭단계에 대해 기술하였다. 선분추출단계에서는 개선된 SLHT를 행하였고, 선분들간의 유사도에 기반한 화상검색을 위하여 매칭단계에서 DP 알고리즘[11]을 이용하였다. 본 논문에서 질의할 수 있는 화상데이터는 직선성이 어느정도 보장되어지는 화상들로 제한한다.

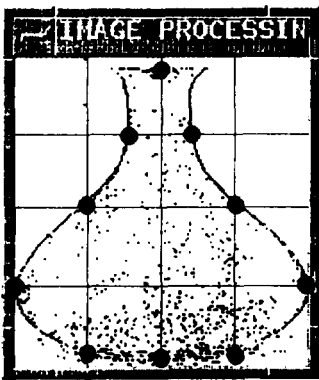
3.1 전처리

입력화상으로부터 윤곽점이나 선분을 추출하기 위해서는 전처리 단계가 반드시 필요하다. 이를 위해 본 연구에서는 에지검출, 잡음제거, 세선화, 객체영역추출, 정규화를 행하였다. 에지검출을 위한 마스크로는 컴퍼스연산자를 사용하였고, 정규화는 가로 x 세로의 폭을 120x120으로 선형정규화 하였다.

3.2 윤곽점 코드화

전처리 단계후의 이진화상은 주로 객체영역, 그중에서도 객체의 윤곽에 대한 정보를 주로 가지게 된다. 그러나 대부분의 경우에 객체의 윤곽은 추적이 가능한 한화소의 연속점으로서 추출되는 것이 아니라 불규칙하게 끊어진 점으로서 나타난다. 이 문제는 화상의 윤곽선을 결정적인 특징으로서 사용하고자 하는데 있어 커다란 장애가 되고 있다. 본 논문에서는 전처리를 통하여 얻어진 이진화상의 객체에 대한 개략적인 윤곽정보를 추출하여 이를 특징으로 하였다.

전처리된 이진 화상을 상, 좌, 우의 4방향에서 주사를 하였을 때, 최초로 만나는 점의 x 혹은 y의 좌표를 구하여 이를 코드화 한다. 이러한 방법은 화상처리 분야에서 널리 사용되는 방법인데 이는 잡음에 매우 민감하게 변한다. 이러한 문제점을 보완하기 위하여 4방향에서 주사할 때 그 주사구간을 일정하게 샘플링하여 구간내의 평균값을 취하는 방법을 사용하였다. 또한 불연속점에 대한 오차를 최소로 하기 위하여 평균값 계산시에 구간내의 n개의 최대값과 최소값을 제외하고 산출하였다. 그림 3에 이러한 방법을 나타내었다.



(그림 3) 윤곽점 코드화
(Fig. 3) Contour point coding

샘플링 간격은 정규화된 이진화상의 가로, 세로폭 모두 10화소 간격으로 하였고 평균값 계산에서 제외할 최대, 최소값의 갯수는 3으로 하였다. 결과적인 코드는 한면에서 12개씩 등록되어 48개의 윤곽점을 얻을 수 있다. 상, 좌, 하, 우의 순으로 12개씩 코드화 하

여 상-하는 y좌표를, 좌-우는 x좌표를 갖게 된다.

3.3 FSLHT 선분추출 알고리즘

기존의 선분추출 알고리즘인 SLHT의 약점을 보완하기 위한 방법으로 몇가지 개선점을 갖는 FSLHT를 제안한다. 첫째로, 전역적인 접근법의 문제점인 계산시간이 오래 걸리는 문제를 시각적인지에 기반하여 후보가 될 수 있는 방향들만으로 제한하여 변환함으로써 크게 개선될 수 있고, 두번째로 이러한 영역에 가중치를 부여하여 보다 정확한 선분을 추출한다. 세번째로 선분통합을 함으로써 비슷한 선분들을 하나의 대표적인 선분으로 표현하고, 선분제거의 과정을 통해 배경이나 잡음에 의해 만들어진 선분을 제거할 수 있도록 한다.

3.3.1 시각적인지를 이용한 방향성 추출

문자에서 획의 방향성을 구하기 위한 방법인 방향 코드화[12]를 화상에 적용할 수 있도록 개선하여, 이 방향을 근거로하여 Hough 변환을 행하게 되면 처리 시간을 크게 감소시킬 수 있으며 보다 정확한 선분을 추출할 수 있다. 화상의 에지분포는 문자에 비해 매우 불규칙적이기 때문에, 하나의 화소와 인접해 있는 이웃화소들과의 관계만을 기준으로 방향성을 규정하기 매우 모호하다. 본 연구에서는 화소들간의 인접성과 연속성을 고려하여 시각적 인지에 기반한 방향코드화를 행하였다.

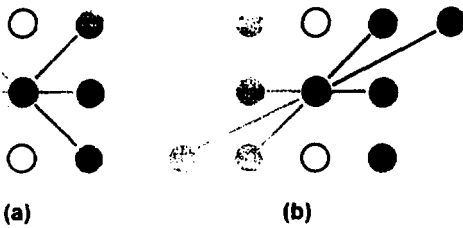
이진화상에서 에지를 이루는 화소를 흑화소라고 할때, 화상내의 모든 흑화소에 대해서 0, 22.5, 45, 67.5, 112.5, 135, 157.5도의 방향으로 흑화소의 존재를 추적하여 그림 4와 같이 현화소에서 8방향중 해당하는 방향에 흑화소가 존재하는 조건을 만족할 때마다 방향카운터값을 1씩 증가시켜 나가며, 하나의 화소마다 추적을 마치면 4개의 방향중 카운터 값이 가장 큰 값을 그 화소의 방향코드로 등록한다. 이 과정에서 연속성과 인접성의 원칙을 따라야 하며, 이에 관한 규칙은 다음과 같다.

규칙 1: 각각의 방향에 따라 흑화소를 추적해 나갈 때 불연속점은 k개 이하이어야 한다.

규칙 2: 하나의 화소에 대한 추적이 끝난 후 방향 카운터값이 $n(n=4)$ 이상이어야 한다.

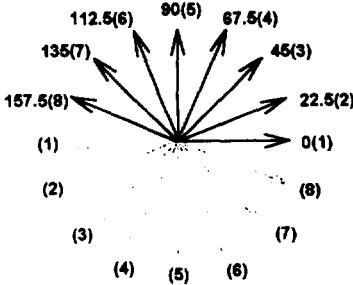
규칙 1에서 불연속점은 그림 4(a)에서 처럼 추적할

경우에서 규정한 위치에 흑화소가 존재하지 않을 때 마다 하나씩 증가시키며 이러한 점이 $k(k=2)$ 개 이상이 되는 순간에는 추적을 중지한다. 이렇게 하여 모든 흑화소에 대한 방향성이 정해지면 이 방향코드에 따라 가중치를 부여하여 이후의 변환에 적용한다. 그림 5는 방향코드를 나타낸다.



(그림 4) 시각적인지기반 방향코드화
(a) $\theta = 0$ (b) $\theta = 22.5$

(Fig. 4) Directional coding based on perceptual organization
(a) $\theta = 0$ (b) $\theta = 22.5$



(그림 5) 방향코드
(Fig.5) Directional code

3.3.2 SLHT

표1은 현화소에서 이웃하는 픽셀들의 인접성을 나타내기 위해 해당 각도 범위에 대해서 8방향으로 설정한 예를 보이고 있다. 3.3.1에서 구한 방향성에 근거하여 표 1과 같은 영역에서 SLHT를 행한다.

SLHT를 수행시에 이러한 방향에 대해서만 행하게 되면 처리시간을 최대 1/4로 단축시킬 수 있다. 원점은 화상의 왼쪽-상단으로 하였고 화상내의 모든 흑화소에 대하여 이를 r, θ 좌표계로 변환한다. 이때, 변환

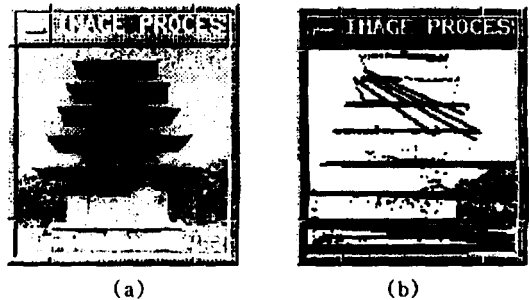
할 θ 의 영역은 방향코드에 제한하여 정하고, r, θ 배열의 증가분은 가중치에 따른다. SLHT의 처리속도를 위해 추출하는 선분의 각도 3도 간격으로 하였다.

<표 1> 방향코드에 기반한 SLHT의 변환 영역
(Table 1) Region of SLHT based on Directional Code

| 방향 | 범 위(각도) |
|----|--------------------|
| 1 | 67.5~90, -67.5~-90 |
| 2 | 45~90 |
| 3 | 22.5~67.5 |
| 4 | 0~45 |
| 5 | -22.5~0, 0~22.5 |
| 6 | 0~-45 |
| 7 | -22.5~67.5 |
| 8 | -45~-90 |

3.3.3 Peak선택 및 역변환

SLHT된 데이터는 화상내에서 그려질 수 있는 선분들을 r, θ 로서 표현하고 있으며, 그 값은 각각의 선분위에 몇개의 점이 존재하는가를 의미한다. 이를 L이라 표시하기 로하며, 그 값이 매우 큰 값들만을 특징으로 하기 위하여 임계치 이상되는 선분을 peak라 표시하며 이 peak들만을 골라내는 작업이 필요하다. 이때 임계치를 너무 크게 잡으면 추출할 수 있는 선분이 특정한 부분에 국한될 수 있고, 너무 작게 잡으면 이후의 선분통합의 결과로서 잡음에 의한 선분이 추출될 가능성이 높다. 이렇게 하여 얻은 선분들에 대해서 역변환을 하여 x, y 좌표값들을 구하고 시점 S



(그림 6) 역변환
(Fig. 6) Inverse transform

와 중점 E, 중점좌표 C, 선분의 길이 I을 또 다른 정보들로서 등록한다. 또한 M을 선분의 밀도라 하여 $M = L/I$ 로서 정의한다.

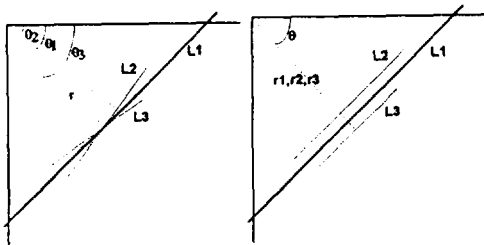
그림 6은 Hough 변환후에 역변환된 30개 선분을 나타낸다.

3.3.4 선분 통합

3.3.3의 결과로서 구한 선분들은 비슷한 각과 거리를 가지면서 r, θ 배열의 값이 커서 추출되는 경우가 많다. 따라서 비슷한 선분들을 하나의 대표적인 선분으로 표현할 수 있어야 한다. 본 연구에서는 선분의 특징인 r, θ , C를 이용하여 선분통합을 행하였다. C가 어떤 차이 이하가 되는 선분들에 대해서 통합 대상 선분들의 유형을 두가지로 분류하였다.

- 유형 1: θ 가 같으면서 비슷한 r을 가지는 경우
- 유형 2: r가 같으면서 비슷한 θ 를 가지는 경우

유형 1과 유형 2의 경우를 같이 고려하여, r, θ 의 값이 어느정도의 차이로 비슷한 경우에, 이들을 최대 L을 갖는 선분으로 통합하였다. 본 논문에서는 r은 10, θ 은 9이하의 차이가 있는 선분을 통합의 대상으로 하였다. 그림 7은 선분통합의 두가지 규칙을 나타내며 이러한 패턴은 화상의 특성으로 미루어 보아 가장 잘 나타나는 경우이다. 그림 8과 같은 통합의 문제가 남아있는데 이는 DP 매칭과정에서 높은 유사도를 가지므로 이 단계에서는 고려하지 않는다.

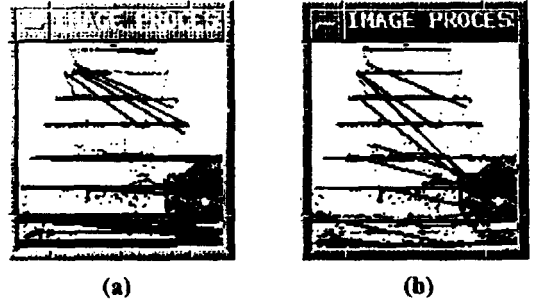


(그림 7) 선분통합의 두가지 규칙
(Fig. 7) Two laws of line integration



(그림 8) 남아있는 선분통합의 문제
(Fig. 8) Reserved line integration problems

그림 9(b)는 9(a)를 선분통합한 후의 18개의 선분들을 나타낸다.

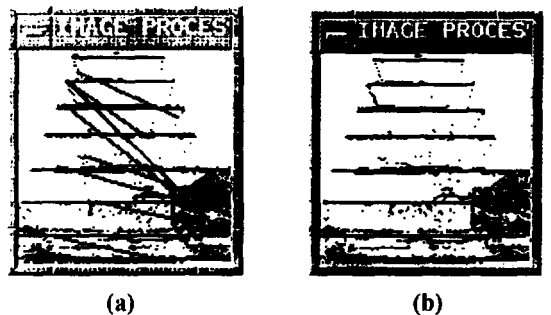


(그림 9) 선분통합
(Fig. 9) Line integration

3.3.5 선분삭제

선분통합후 얻은 선분이라 할지라도 선분위에 실제 존재하는 점들간의 조밀성에 대한 정보는 가지고 있지 않다. 즉 그림 9(b)에서와 같이 실제 선분이라고 인정할 수 없는 선들이 잘못 추출되는 경우가 존재한다. 이는 전역적인 정보에 근거하여 선분추출시 약점으로 지적되는 문제인데, 이는 선분의 밀도 M을 이용하여 임계치 이하의 선분들을 제거함으로써 보정할 수 있다.

그림 10(b)는 그림 10(a)를 선분삭제한 후의 10개의 선분들을 나타낸다.

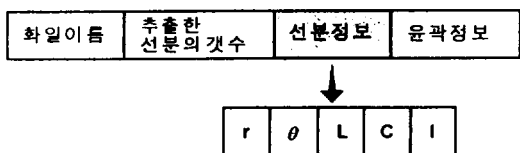


(그림 10) 선분삭제
(Fig. 10) Line deletion

3.4 매칭 단계

화상간의 유사도 산출을 위하여 매칭에 관한 일정

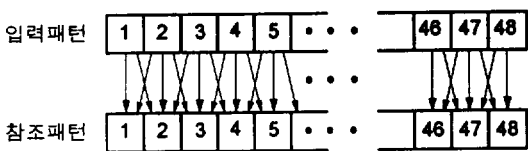
한 기준이 필요하다. 내용기반 검색의 특징중 하나인 유사도는 시스템의 성능에 직접적인 영향을 미치므로 특징에 맞는 효과적인 매칭방법이 필요하다. 본 논문에서는 윤곽점과 선분의 매칭으로 구분하여 각각의 방법에 따라 유사도를 산출하였으며 이들을 그림 11처럼 화상 인덱스 구조로 저장한다.



(그림 11) 화상인덱스의 구조
(Fig. 11) Structure of image index

3.4.1 윤곽점의 매칭

3.2절에서 구한 윤곽점코드를 이용하여 화상간의 유사도를 산출할 때에는 48개의 코드간의 값의 차이를 계산하여 어느 임계치 이하이면 카운트수를 1씩 증가시켜 나감으로써 결과적인 매칭카운트수를 산출할 수 있다. 다음에 이 카운트수를 백분율비로 나타내어 화상에 대한 유사도로 삼는다. 48개의 코드를 비교시에 잡음이나 불연속점에 대한 영향을 고려하여 그림 12와 같은 형태로 비교한다.

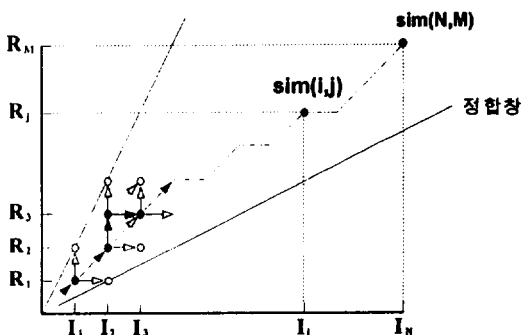


(그림 12) 윤곽점에 의한 비교
(Fig. 12) Comparison with contour points

3.4.2 선분의 매칭

입력패턴과 참조패턴을 비교할때, 두 패턴의 길이가 서로 다를 경우에는 선형신축으로 두 패턴의 길이를 같게 할 수 있다. 그러나 패턴내의 특징의 변화가 선형이 아닐 경우에는 선형신축은 의미가 없게 된다. 이렇게 패턴내의 특징의 변화가 단일하지 않은 경우에 두개의 패턴을 비교하기 위한 이론으로 DP(Dynamic Programming)를 이용한 정합법이 있다. DP는 다단

계 처리에 있어서 global optimum을 구하는 방법으로서, 주어진 두점간의 최적의 경로는 그 경로위에 있는 어떠한 두점간에 있어서도 최적의 경로를 이룬다는 Bellman의 최적화의 원칙[13]을 기본으로 한다. 본 논문에서는 서로 다른 개수의 선분들을 갖는 화상간의 유사도를 산출하기 위하여 특징추출단계에서 얻은 특징벡터 r, θ, L, C 를 화상의 매칭대상으로 하여 DP 알고리즘을 적용하였다.



(그림 13) DP 알고리즘
(Fig. 13) DP algorithm

특징추출후에 하나의 화상은 특징벡터들의 집합으로 볼 수 있고 각 특징벡터(r_i, L_i, θ_i, C_i)는 이 화상의 선분들을 표현한다. DP 매칭의 복잡성을 줄이기 위하여 이 특징벡터들을 12방향($\theta = n \cdot \pi / 12, n = 1, 2, \dots, 12$)에 따라서 12개의 영역으로 양자화한 후 각 영역에 있는 특징벡터들을 r 에 따라 정렬함으로써 하나의 문자에 대해서 최대 12개의 정렬된 특징벡터의 열들을 얻을 수 있다. 이렇게 되면 각각의 영역에 있는 선분들은 최대차이 15도를 갖는 비슷한 방향과, r 에 대해 차이를 갖는 각기 다른 위치에 놓이게 된다. 그리고 나서 그림 13과 같이 모든 입력화상의 특징벡터의 열들은 DB에 있는 대응하는 화상의 특징벡터들과 각각 매칭작업을 하게 된다. I_r 과 R_r 을 $\theta = r \cdot \pi / 12, r = 1, 2, \dots, 12$ 일때의 입력패턴과 참조패턴을 나타내는 특징벡터의 열이라 하자. 여기서 $I_r = \{I_{r_1}, I_{r_2}, I_{r_3}, I_{r_4}\} | i = 1, 2, \dots, N$ 이고 $R_r = \{R_{r_1}, R_{r_2}, R_{r_3}, R_{r_4}\} | j = 1, 2, \dots, M$ 이다. DP 매칭의 알고리즘은 아래와 같다.

STEP 1: $sim(0, 0) = 0$

```

STEP 2: for i = 1, N do
    sim(i, 0) = 0
end
for j = 1, M do
    sim(0, j) = 0
end
STEP 3: for i = 1, N do
    for j = 1, M do
        s1 = sim(i-1, j-1) + d(i, j)
        s2 = sim(i, j-1)
        s3 = sim(i-1, j)
        sim(i, j) = max(s1, s2, s3)
    end
end
STEP 4: Sr = sim(N, M)
    
```

sim(i, j), I_r과 R_r간의 유사도를 계산하기 위하여 사용된 누적 배열
d(i, j), I_r과 R_r간의 특징간의 유사도를 저장하고 있는 배열

$$\begin{aligned}
 d(i, j) &= (1000 - (\alpha * \theta_Dif + \beta * r_Dif + \gamma * L_Dif + \delta * C_Dif)) / 10 \quad (1) \\
 \theta_Dif &= abs(I\theta_i - R\theta_j) \\
 r_Dif &= abs(Ir_i - Rr_j) \\
 L_Dif &= abs(IL_i - RL_j) \\
 C_Dif &= abs(IC_i - RC_j)
 \end{aligned}$$

하나의 화상에 대해서 최대 12개의 특징벡터의 열이 있으므로 최종적인 유사도 S를 얻기 위해서 위의 알고리즘은 최대 12개의 각기 다른 θ 에 대해서 선택적으로 수행되어야 한다. 그러면 S는 다음과 같이 계산된다.

$$S'_r = \frac{S_r}{n} \cdot (I_r + R_r) \quad n \text{은 Backtrack된 경로의 갯수} \quad (2)$$

$$S = \frac{1}{avr} \cdot \sum_{r=1}^{12} S'_r \quad (3)$$

$$avr = \sum_{r=1}^{12} \left(\sum_{i=1}^{N_i} I_i + \sum_{j=1}^{M_j} R_j \right) \quad (4)$$

4. 실험 결과 및 고찰

본 장에서는 내용기반 화상검색시스템의 구현환경에 대해 설명하고 실험을 통해 실험결과를 입증하고 분석하였으며 문제점과 개선방안등을 고찰하였다.

4.1 구현환경

본 연구에서는 제안된 내용기반 화상정보 검색시스템을 PC와 LAN으로 연결된 SUN Sparc-II workstation을 사용하여 X-Window상에서 C언어와 motif 구현하였다. 검색에 사용된 화상의 수는 280개로서 탐, 도자기, 불상등과 같은 문화재 관련화상 100개와 탐에서 배경을 지운 화상 30개, 스케치로 그려넣은 50개의 화상과 그의 80개의 화상으로서 실험하였다. 실험 데이터의 유형을 표2로 나타내었다.

〈표 2〉 화상데이터별 분류
 〈Table 2〉 Class of image data

| 화상의 유형 | 갯 수 |
|---------|-----|
| 탐 | 30 |
| 탐(배경없음) | 30 |
| 도자기 | 40 |
| 불상 | 30 |
| TV. 세트 | 20 |
| 면허시험안내도 | 20 |
| 자동차 | 20 |
| 스케치화상 | 50 |
| 기타 | 40 |
| 총계 | 280 |

4.2 실험결과 및 분석

4.2.1 특징추출부

이진화상을 입력으로 하여 선분추출을 하는 경우에, 추출되는 선분은 이진화상에서의 에지의 갯수와 밀접한 관계가 있다. 선분추출에 민감하게 영향을 끼치는 또 하나의 요소는 Hough 변환후에 Peak를 찾아내는 과정으로 전처리에 어느정도 종속되나, 이전의 과정에서 방향성에 근거하여 구한 정보이므로 화상 내의 선분에 대한 보다 믿음만한 파라미터가 될 수

있다. 본 논문에서는 최적화된 선분추출을 위하여 배경을 가지고 있고, 탐침이 어느정도 강조되어지는 10개의 화상에 대해서 Peak값에 따른 추출선분의 갯수와 오검출된 선분에 대한 실험을 하였다. 추출하는 최대 선분의 갯수는 10개로 잡았고, 선분이라고 볼 수 있는 기준은 탐침과 밀단의 갯수라 하였다. 표 3은 Peak 값을 여러가지로 변화시키면서 행한 선분추출의 결과이며, 변환할 방향의 영역은 45도간으로 한 경우이다.

〈표 3〉 여러가지 Peak에 대한 선분추출의 결과
 〈Table 3〉 Results of line extraction for various Peak

| Peak 값 | 정확도 | 에러율 |
|-----------|-----|-----|
| Peak = 35 | 73% | 15% |
| Peak = 30 | 80% | 18% |
| Peak = 25 | 83% | 20% |
| Peak = 20 | 85% | 30% |
| Peak = 15 | 85% | 33% |

정확도와 에러율은 다음과 같이 정의하였다.
 정확도 = 추출한 탐침의 갯수/존재하는 갯수
 에러율 = 탐침이 아닌 선분의 갯수/추출한 선분의 갯수

Hough 변환을 수행시에 기본적으로는 모든 방향에 대하여 행하나 본 논문에서는 방향성에 근거한 Hough 변환을 행하였다. 8방향성에 대하여 Hough 변환을 행할 영역을 그 방향을 포함한 30도간으로 한 경우, 45도간으로 한 경우, 60도간으로 한 경우, 모든 방향에 대하여 행한 경우로 나누어 이에 대한 속도비

〈표 4〉 방향성에 근거한 선분추출 실험
 〈Table 4〉 Experiment of line extraction based on Direction

| 영역 | 정확도 | 에러율 | 속도비 |
|------|-----|-----|------|
| 30도 | 78% | 17% | 0.10 |
| 45도 | 83% | 20% | 0.17 |
| 60도 | 83% | 23% | 0.24 |
| 90도 | 76% | 28% | 0.40 |
| 180도 | 70% | 35% | 1 |

와 정확도, 에러율에 대한 실험을 함으로써 방향코드에 대한 신뢰도를 측정하였다. 표 4는 이를 나타내며, 역변환시의 Peak는 25도로 한 경우이다.

다음은 결과적으로 이러한 최적의 파라미터를 가지고 기존의 Hough 변환과 본 연구에서 제안한 FSLHT의 성능을 평가하였다. 실험 대상은 위와 같으며 표 5는 이를 나타내며 역변환시의 Peak는 30, 변환 영역은 45도간으로 한 경우이다.

〈표 5〉 SLHT와 FSLHT의 비교
 〈Table 5〉 Comparison with SLHT and FSLHT

| | 정확도 | 에러율 | 속도비 |
|-------|-----|-----|------|
| SLHT | 70% | 38% | 1 |
| FSLHT | 85% | 18% | 0.17 |

4.2.2 정보검색부

일반적인 정보검색 시스템은 precision과 recall을 사용하여 평가한다. 주어진 질의에 대해서 T를 검색 대상중 관련된 항목의 총 수라 하고, T_r을 검색된 항목의 총 수, R_r을 검색된 항목 중 관련된 항목의 수라고 할 때, precision은 R_r/T_r, recall은 R_r/T로서 정의한다. 그러나 유사도에 의한 매칭을 하는 시스템에서도 정규화된 precision과 recall로서 평가한다[14]. 이들 파라미터는 유사도 계산으로 정렬된 결과검색열중에서, 관련된 항목이 나타나는 위치를 참조하게 된다. 예를 들면, 정규화된 recall은 관련된 T개의 항목 0순위로부터 차례로 T-1의 순위에 나타나는가를 평가하는 exact match와 비교하여, 관련된 항목들이 검색된 항목 리스트의 0순위에 얼마나 근접하여 나타나는가를 측정함으로써 구할 수 있다.

본 논문에서는 시스템의 검색 효율을 평가하기 위하여 precision과 recall과 함께 다음과 같은 평가파라미터를 정의하였다[1]. 화면에 디스플레이 하는 화상의 갯수는 10개이다.

- NR : Number of relevant image, 디스플레이 된 화상중 관련된 화상의 수
- AVRR : 디스플레이 된 화상중 관련된 화상들의 평균 순위

위의 파라미터를 가지고 스캔된 화상을 질의로 한 경우와 스케치한 화상을 가지고 질의를 한 경우에 대

하여 실험을 하였으며 각각에 대하여 윤곽특징에 의한 실험과 FSLHT에 의한 실험, 윤곽과 선분에 의한 병합검색 실험을 하였다. recall과 precision에서 검색된 화상은 유사도가 50이상인 경우이다. 선분에 의한 검색시 식(3.1)에서의 파라미터는 각각 $\alpha=10$, $\beta=3$, $\gamma=3$, $\delta=2$ 로 하였으며 윤곽에 의한 검색시의 임계치는 20으로 하였다.

표 6은 윤곽에 의한 검색 결과를 나타낸다.

〈표 6〉 윤곽에 의한 검색 결과
 〈Table 6〉 Results of retrieval by Contours

| 평가파라미터 | 탑 | TV세트 | 도자기 | 불상 | 스케치화상 |
|-----------|------|------|------|------|-------|
| 질의횟수 | 5 | 5 | 5 | 5 | 5 |
| recall | 0.48 | 0.62 | 0.88 | 0.6 | 0.8 |
| precision | 0.55 | 0.47 | 0.59 | 0.31 | 0.52 |
| NR | 6.0 | 7.4 | 8.8 | 6.2 | 5.8 |
| AVRR | 3.5 | 4.1 | 3.9 | 4.3 | 4.7 |
| 검색속도(초) | 5.6 | | | | |

표 7은 선분에 의한 검색 결과를 나타낸다.

〈표 7〉 선분에 의한 검색 결과
 〈Table 7〉 Results of retrieval by lines

| 평가파라미터 | 탑 | TV세트 | 스케치화상 |
|-----------|------|------|-------|
| 질의횟수 | 5 | 5 | 5 |
| recall | 0.8 | 0.85 | 0.66 |
| precision | 0.91 | 0.21 | 0.73 |
| NR | 8.6 | 5.2 | 7.2 |
| AVRR | 4.3 | 4.7 | 4.0 |
| 검색속도(초) | 17.3 | | |

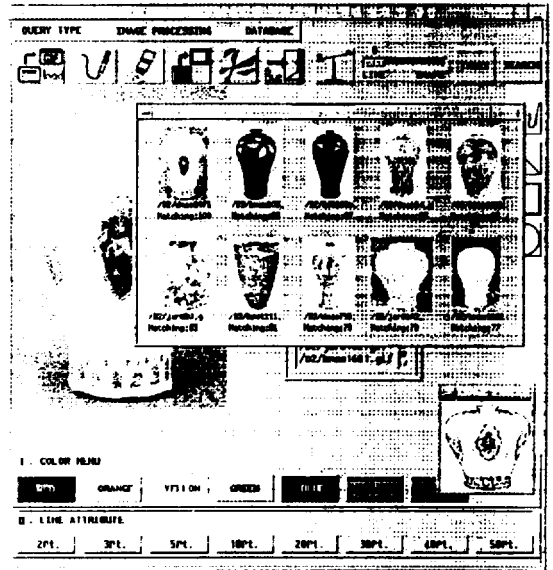
표 8은 윤곽과 선분에 의한 병합검색 결과를 나타낸다.

〈표 8〉 윤곽과 선분에 의한 검색 결과
 〈Table 8〉 Results of retrieval by contours and lines

| 평가파라미터 | 탑 | TV세트 | 스케치화상 |
|--------|------|------|-------|
| 질의횟수 | 5 | 5 | 5 |
| recall | 0.65 | 0.95 | 0.54 |

| | | | |
|-----------|------|-----|-----|
| precision | 0.81 | 0.4 | 0.5 |
| NR | 8.0 | 6.4 | 7.2 |
| AVRR | 4.3 | 3.9 | 4.7 |

그림 14는 윤곽에 의한 검색 예를 보여주고 있다.

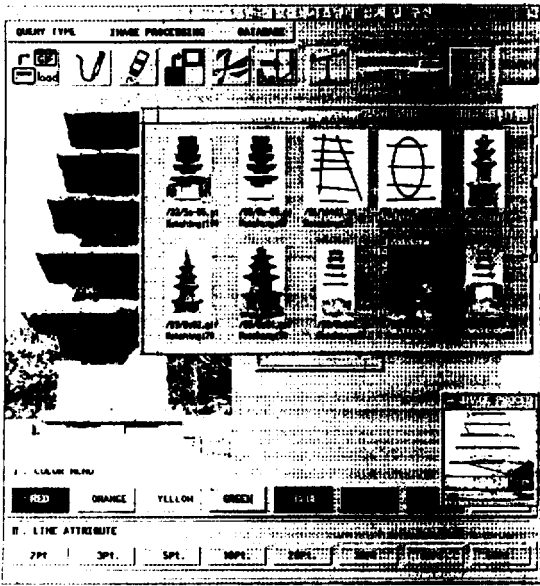


(그림 14) 윤곽에 의한 검색 예(임계치 = 15)
 (Fig. 14) An example of retrieval by contours

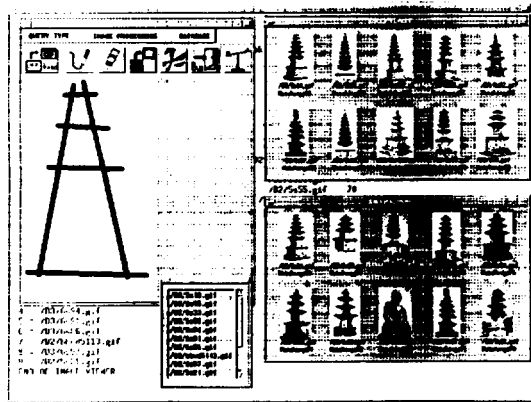
그림 15는 선분에 의한 검색 예를 보여주고 있다.

그림 16은 윤곽과 선분에 의한 병합검색 예를 보여주고 있다.

본 논문에서는 검색의 효율을 향상시키기 위하여 클러스터링 단계를 두었다. 즉 질의화상과 참조화상 간의 비교시 어떠한 키(key)값을 두어 이에 따라 분기, 검색함으로써 시스템의 속도에 있어 향상을 이루었다. 화상으로부터 추출한 선분의 갯수는 화상의 직선성을 잘 반영하므로 이를 분기조건으로 하여 검색을 하게 되면 관계없는 화상들간의 극단적인 비교를 방지할 수 있다. 본 논문에서는 질의화상과 참조화상으로부터 추출한 선분의 갯수차를 구하여 어떤 값 이하가 되는 화상들만을 질의화상에 대한 검색 대상으로 하였다. 표 9는 임의의 7개의 화상질을 5번씩 하였을때, 추출선분의 갯수차에 따른 검색 결과이다. 표



(그림 15) 선분에 의한 검색 예($\alpha = 10, \beta = 2, \gamma = 1, \delta = 1$)
 (Fig. 15) An example of retrieval by lines



(그림 16) 윤곽과 선분에 의한 검색 예($\alpha = 15, \beta = 3, \gamma = 2, \delta = 2$)
 (Fig. 16) An example of retrieval by contours and lines

9에서 N은 질의화상과 참조화상으로부터 추출한 선분의 갯수차이다.

4.3 고찰

제안된 시스템에서는 FSLHT를 이용하여 선분을 추출하였고 개개선분들의 유사도 계산에 DP를 적용

〈표 9〉 클러스터링을 고려한 검색 결과
 (Table 9) Results of retrieval consider with clustering

| 평가파라미터 | N = 10 | N = 5 | N = 3 |
|-----------|--------|-------|-------|
| recall | 0.72 | 0.63 | 0.47 |
| precision | 0.67 | 0.7 | 0.72 |
| NR | 7.3 | 6.3 | 5.8 |
| AVRR | 4.1 | 4.7 | 3.8 |
| 검색속도비 | 1 | 0.77 | 0.59 |

하였다. 선분추출시에 오검출되는 선분의 원인인 전처리 과정에서 객체의 경계 부분을 이진화상으로 제대로 표현하지 못하는 문제가 가장 크다고 할 수 있으며 이에 따라 보다 효율적인 전처리기술이 필요하다. 표 4에서 변환하는 영역이 증가함에 따라 정확도가 떨어지는 반면 에러율이 높아지는 원인은 방향성이 부여되지 않는 경우에 배경에 의한 선분을 추출할 가능성이 커짐에 따라 생기는 결과이다. 유사도의 계산에 DP알고리즘을 적용한 것은 유사도 산출에 매우 융통성을 제공할 수 있었다. 즉 화상마다 추출되는 선분의 갯수가 다른 상황에서도 바람직한 유사도를 산출할 수 있었고, 극단적인 비교를 없앴으로써 속도와 신뢰도면에서 효율적이라고 할 수 있다. 그러나 매칭을 위하여 각도에 따라 12개의 영역으로 분할함에 따라, 이로 인한 오차도 있을 수 있다. 예를 들어 15도의 선분과 16도의 선분이 각각 다른 클래스로 분리됨으로써 유사도를 떨어뜨릴 수 있는 문제가 있다. 또한 선분들간의 유사도는 3.4절의 식 3.1에 의해 계산되므로 4개의 요소에 곱해지는 $\alpha, \beta, \gamma, \delta$ 는 각각 그 요소에 대한 가중치를 부여하는 의미를 가지며, 따라서 이 요소들을 변화시켜 감에 따라 계산되어지는 유사도는 민감하게 반응한다. 이것은 사용자가 선분에 의한 검색을 할 때 어떤 요소에 중점을 두어 검색할 것이냐 하는 문제로서 검색에 있어서 융통성을 부여할 수 있는 부분이다.

검색에 관한 실험결과에서 보는 것처럼 본 시스템에서 검색할 수 있는 화상은 특징에 따라 분류되어진다. 즉 선분특징은 주로 탑이나 TV 세트와 같은 직선성이 강한 화상검색에 주로 쓰이고 윤곽특징은 배경의 영향이 약한 도자기등의 검색, 혹은 필터로서 적합하다. 표 5의 선분에 의한 검색 결과에서 보는 것처럼

럼 탐과 같이 선분의 특성이 두드러지는 화상의 precision은 거의 1에 근접해 가는 걸을 볼 수 있으나, TV 세트의 경우에는 선분의 패턴이 매우 단순하여 recall은 매우 높은 반면 precision은 상당히 낮다. 윤곽과 선분에 의한 병합검색의 경우에는, 이러한 경우에 윤곽에 의한 검색에서 TV의 외곽에 대한 필터 역할을 해 줌으로써 precision을 어느정도 향상시킬 수 있었다. 검색의 시간효율을 향상시키기 위하여 클러스터링 단계를 둬으로써 표 9과 같이 검색속도의 개선을 이루었으며, 이는 검색 대상이 되는 화상의 수가 증가함에 따라 더욱 효과적일 수 있다. 기존의 국외의 시스템들은 칼라, 모양, 질감 특징들을 조합한 화상 검색 시스템들으로써, 상용화된 시스템이 아니고 실험적인 연구가 대부분이며, 예로써 IBM의 QBIC [Query By Image Content][15] 시스템과 비교해 보면 동일한 대상에 대해 상호 비교 평가는 할 수 없었으나 recall, precision, AVRR 등의 수치값은 QBIC에 비해 비슷하거나 보다 우수한 값을 보였다.

5. 결 론

본 연구에서는 인간이 화상을 식별하는 주된 특징인 형태에 기반한 검색을 위하여, 화상내 객체의 윤곽정보를 얻고 객체내의 선분을 추출하여 이를 인덱스하고 매칭하는 방식의 화상정보 검색시스템을 구현하였다. 화상으로부터 선분을 추출해 내는 방법으로 기존의 SLHT를 개선시킨 FSLHT를 개발하였으며, 선분들의 유사도에 기반한 화상간의 매칭을 위하여 DP(Dynamic Programming) 알고리즘을 적용하였다. 본 시스템에서는 윤곽특징과 선분특징에 의한 병합검색을 가능하게 함으로써, 사용자는 윤곽특징에 의한 첫번째 검색 결과를 판단할 수 있으며 이로 인해 보다 광범위하고 능동적인 검색을 가능하게 하였다. 윤곽특징을 고려할 때 화상의 윤곽정보가 잡음에 민감하게 반응하는 문제를 보완하기 위하여 일정간격으로 샘플링한 영역의 대표값을 윤곽특징으로 잡았다. Hough 변환의 가장 큰 문제점인 시간문제와 연속성의 상실문제는 시각적인지에 기반한 방향성을 산출하여 이 방향에 근거한 Hough 변환을 행함으로써 극복하였다. 또한 선분간의 매칭과정에서 각기 다른 갯수의 선분을 비교하기 위하여 DP알고리즘을 적

용함으로써 유사도를 산출하였다. 논문에서 구현한 시스템에서는, 탐과 TV 세트, 자동차등 직선성이 어느정도 존재하는 화상에 대해 윤곽과 선분에 의한 검색을 한 결과 평균 80% 가량의 recall과 60%이상의 precision을 얻을 수 있었다.

앞으로의 과제로 본 시스템의 질의가능한 유형을 확장하는 것이 필요하다. 이 문제는 곡선이 직선의 결합으로 이루어진다는 것을 감안할 때 화상을 분할하여 그 영역내에서 FSLHT를 행하여 보다 작은 직선을 영역당 찾아냄으로써 어느정도 개선되리라고 본다. 객체의 보다 구체적인 형태를 반영할 수 있도록 윤곽점을 코드화하는 단계에서 이전 윤곽점과의 좌표에 대한 편차를 추가시킬 수 있겠다. 배경에 민감한 문제를 어느정도는 보완하였으나 배경의 영향이 매우 강하게 나타나는 화상에서는 역시 문제가 있다. 현재의 수준에서 화상의 형태만을 가지고 검색하는 것은 매우 어려운 문제이므로 보다 효율적인 검색을 위해서는 선분이나 윤곽특징이외에 컬러나 텍스처(texture), 혹은 텍스트등을 이용한 중간 필터단계를 추가시키는 것이 바람직하다.

참 고 문 헌

- [1] Batber, W.Equitz, C. Faloutsos, Query by Content for Large On-Line Image Collections' IEEE, 1995
- [2] William I. Grosky, Rajiv Mehrotra, "Index-Based Object Recognition in Pictorial Data Management", Computer vision, Graphics, and Image Processing, vol.52, pp.416-436, 1990.
- [3] William I. Grosky, Peter Neo, Rajiv Mehrotra, "A Pictorial Index Mechanism for Model-based Matching", pp.309-327, vol.8, 1992.
- [4] Tat-Seng Chua, Swee-Kiew Lim and Hung-Keng Pung, "Content-based Retrieval of Segmented Images", ACM Multimedia '94, pp.211-218, 1994.
- [5] Kyoji Hirata, Toshikazu Kato, "Query by Visual Example", Extending database Technology '92, pp.56-71, 1992.
- [6] Kyoji Hirate, Toshikazu, "Rough Sketch-Based Image Information Retrieval", NEC Res. &

Develop., vol.34, no.2, 1993.

[7] Kyoji Hirata, Yoshinori Hara, Naoki Shibata, Fusako Hirabayashi, "Media-based Navigation for Hypermedia Systems", Hypertext '93 proceedings, pp.159-172, 1993.

[8] Micheal J. Swain, "Interactive Indexing into Image Database", SPIE vol.1908, pp.95-103, 1993.

[9] W.Niblack, R.Barber, W.Equitz, M.Flickner, Giasman, D.Petkovic, P.Yanker, "The QBIC Project: Querying Images by Content Using Color, Texture, Shape", SPIE vol. 1908, pp. 173-187, 1993.

[10] R. O. Duda and P. E. Hart, "Using of the Hough transform to detect lines and curves in pictures", Commun. ACM, pp.11-15, 1972.

[11] FANG-HSUAN, WEN-HSING HSU, "Recognition of Handwritten Chinwse Characters by Modified Hough Transform Techniques", IEEE Transactions on Pattern Recognition and Machine Intelligence, 1989.

[12] 장희돈, "적응형 규칙망에 의한 개선된 한글 문서인식시스템의 구현", 광운대학교 박사학위 논문, pp.72-73, 1994.

[13] R.E.Bellman and S.Dreyfus., Applied Dynamic Programming. Princeton, N.J.:Princeton University Press, 1962.

[14] Salton, G., and M.J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, New York, N.Y., 1983.

[15] IBM Almaden Research Center, "Query by Image and Video Content: The QBIC System", IEEE Multimedia, pp.23-32, September 1995.

[16] Virginia E.Ogle and Michael Stonebeaker, "Chabot: Retrieval from a Relational DataBase of Images", IEEE Multimedia, pp.40-48, September 1995.



정 원 일

1964년 한국항공대학교 전자공학과(공학사)

1979년 건국대학교 대학원 전자공학과(공학석사)

1986년~현재 광운대학교 전자계산기공학과(박사수료)

1873년~현재 대전산업대학교 전자공학과 교수
 관심분야: 멀티미디어, 하이퍼미디어, 멀티미디어 정보검색, 회로 및 시스템



구 정 현

1994년 광운대학교 전자계산기공학과 졸업(공학사)

1996년 광운대학교 대학원 전자계산기공학과(공학석사)

1996년 2월~현재 현대중공업 시스템제어 개발실

관심분야: 멀티미디어 정보검색 (화상 검색)



최 기 호

1973년 한양대학교 전자공학과(공학사)

1977년 한양대학교 대학원 전자공학과(공학석사)

1987년 한양대학교 대학원 전자공학과(공학박사)

1977년~1979년 한국과학기술연구소 연구원

1979년~현재 광운대학교 컴퓨터공학과 교수/신기술연구소

1989년~1990년 Univ. of Michigan 전기및 전산과 Visiting Scholar

관심분야: 멀티미디어 정보 검색, 하이퍼미디어, 멀티모달