

論文96-33B-7-7

내용기반 화상 검색시스템의 설계및 구현

(The Design and Implementation of a Content-based Image Retrieval System)

鄭元一*, 崔炫燮**, 崔基浩**

(Won Ill Jeoung, Hyun Sub Choi, and Ki Ho Choi)

요 약

멀티미디어정보중 화상과 같은 복잡한 데이터 검색은 키워드가 아닌 내용에 기반한 검색방법이 요구된다. 본 논문에서는 텍스트가 아닌 가지적으로 입력된 질의화상에 대해 칼라, 라인, 모양등의 특징을 이용하여 후보화상을 검색해 내는 내용기반 화상검색시스템을 설계하고 구현하였다. 화상검색을 위해 화상내 각 16개 영역별 대표칼라성분을 추출한 칼라 특징과, 전처리된 이진화상내의 4개의 대표선분을 추출한 라인 특징및 16개 영역별 방향밀도벡터를 추출한 모양 특징을 이용하였다. 화상검색을 위한 질의는 전체화상이나 스케치나 문양에 의한 질의방법을 사용가능케 하였으며, 실험결과, precision 91%, recall 33%및 평균순위 3.1 를 보임으로써 높은 검색효율을 보였고, 가중치에 따라 특징을 조합하여 검색함으로써 검색효율과 정확도를 높일 수 있었다.

Abstract

To retrieve complex data such as images in multimedia information, we need the content-based retrieval methods based on the visual properties rather than keywords. In this paper, a content-based image retrieval system is designed and implemented to retrieve images using the features of images such as colors, lines and intensity vector features when a visual query inputs. The contents for image retrievals are the color features extracted from the color component of 16 blocks of the image, the line features extracted from 4 lines in the image and the shape features extracted from the intensity vectors of the 16 blocks. We can either use a whole image or a sketch image for query. As the experimental results demonstrate the precision 91%, the recall 33% and the average rank 3.1, the retrieval performance is found to be high. The experimental results indicate that the retrieval using the weighted features have led to substantial improvement in the percision and performance of systsem.

I. 서 론

현재의 하드웨어의 기술은 사용자들로 하여금 대규모 디지털 화상을 획득하고 저장, 검색, 조작및 전송을

가능하게 하였으며, 멀티미디어정보시스템은 사용자가 텍스트, 화상, 그래픽, 비디오, 오디오와 같은 다양한 데이터 유형을 빠르고 쉽게 검색, 저장, 관리할 수 있게 날로 발전하고 있다. 이러한 멀티미디어시스템은 의학, 디지털 도서관, 교육, 학습, 비즈니스 의사결정, 주식동향, 오락, 군사, 예술, 패션, 카달로그등 광범위한 분야에 응용가능하여 지고 있다.

* 正會員, 大全産業大學校 電子工學科

(Dept.of Electronic Engineering, Daejeon National Univ. of Technology)

** 正會員, 光雲大學校 컴퓨터工學科

(Dept. of Computer Engineering, Kwangwoon Univ.)

接受日字:1996年2月10日, 수정완료일:1996年7月3日

현재 화상검색을 지원하는 상용 시스템들은 대부분 각 화상과 연관된 텍스트나 키워드의 사용을 기본으로 하여 검색및 관리기능을 수행하는 화상DB들로, 질의어는 SQL이나 텍스트 검색프로그램과 같은 표준질의언

어를 사용하고 있다. 이들 질의어는 논리조합인 AND, OR, NOT과 같은 키워드/텍스트 질의를 포함하며 synonyms, thesaurus나 논리적인 의미계층 기능을 지원하고 있다. 그러나 이들의 검색방법은 각 화상에 입력된 초기 텍스트정보에 전적으로 의존하므로 만약 질의가 초기에 묘사되지 않은 화상속성을 참조하게 되면 검색은 대부분 실패하게 된다. 또한, 특정 질감이나 모양같은 몇몇 복잡한 가지적인 속성을 텍스트로는 묘사가 불가능하며, 설사 텍스트로 화상의 모든 중요한 특성을 묘사한다 하더라도 화상 속성을 일관되고 공통된 단어로 기술할 수 없다.

따라서, 화상과 같은 복잡한 데이터 유형은 가지적인 속성에 근거한 보다 지능화된 검색방법이 요구되며, 키워드와 관계없이 사용자가 시각적인 예(visual example)로 질의어를 입력하여, 화상의 특징을 내용으로 검색을 할 수 있는 내용기반 화상검색시스템이 필요하다. [11][12]

그동안 칼라, 모양, 질감 등의 성분을 이용한 내용기반 검색 시스템에 관한 다양한 연구가 있었다. Hirata 등은 사용자가 그린 대략적인 스케치 또는 칼라나 복사본 화상(monochrome photo)으로 화상 데이터를 검색할 수 있는 방법을 보였다. 이들 연구에서는 아이콘 화상을 사용하여 화상 데이터의 특징값을 나타냈다. 이러한 접근 방식은 다양한 형태의 화상 데이터를 다룰 수 있다는 장점이 있으나 유사한 화상을 검색하기 위해서 모든 데이터를 화소 대 화소 비교로 순차적 검색방법을 사용하므로 매우 비효율적이 된다. 따라서 매우 많은 양의 화상 데이터를 저장하고 관리하는 데이터베이스 응용에는 적절치 않다. [13]

QBIC에서는 텍스트와 더불어 스케치, 사용자가 구성한 질의 화상, 칼라, 레이아웃이나 구조적 묘사, example 화상, 아이콘이나 그래픽 정보를 기반으로 한 대규모 데이터베이스 질의가 가능한 검색방법을 개발하였다. [9] 기존의 내용기반 화상검색시스템에 대한 연구들은 주로 화상내의 칼라 히스토그램이나, 곡률, 모멘트, 주변길이, 면적과 같은 모양특징 및 질감등을 이용한 검색방법들이 대부분이며, 물체의 위치나 크기에 매우 민감하며, 특정응용에만 적합하거나, 검증을 통하지 않은 이론으로만 제안된 상태의 연구들이다.

본 논문에서는 화상 데이터에서 검색 대상이 되는 객체의 윤곽선을 방향벡터화하여 화상의 특징값으로 사용하였다. 이러한 방법으로 얻은 벡터화된 화상의 특

징값들은 거리, 방향, 인접 요소와 이루고 있는 각도등을 평가하여 화상의 유사성을 비교할 수 있기 때문에 여러 개의 객체들이 있을 경우 각각을 따로 식별할 수 있고 여러 응용대상에 적용할 수 있는 장점이 있다.

따라서 본 논문에서는 기존의 여러 방법들의 장점을 이용하고 기존 연구의 문제점을 극복하고자 하였으며 칼라, 라인, 벡터특징의 가중치에 의한 검색으로 융통성과 검색효율을 높일 수 있는 내용기반 화상검색 시스템을 제안한다.

이를 위해 본 논문에서는 사용자가 질의어로서 RGB 256 칼라 화상을 입력하게 되면 각 영역별대표 칼라성분과 인접 영역간의 칼라값 차이를 산출함으로써 칼라 특징을 추출하며, 전처리한 이진화상을 골격선 추출하고 인접4방향 성분 에 따라 픽셀을 방향벡터화함으로써 모양 특징을 추출하고, 전처리한 이진화상을 Hough 변환을 통해 평균적으로 한화상당 4개의 라인성분을 추출함으로써 칼라, 벡터, 라인에 의한 새로운 내용기반 화상검색시스템을 개발하고 구현한다.

II. 내용기반 화상 검색시스템의 설계

내용기반 화상검색은 화상의 내용에 의한 화상 DB의 질의가 연구의 핵심이 되며, 머신 비전과는 달리 화상표현 및 유사도 측정은 인간의 인지 및 지각과정과 일치하는 속성을 가져야 한다. 기존의 DB와 비교할때 화상DB는 매우 큰 데이터규모, 부가적인 텍스트 검색과 더불어 텍스트로 적절하게 기술할 수 없는 오브젝트에 대한 내용 및 유사도기반 검색의 요구, 가지적이고 직관적인 사용자 인터페이스 제공의 필요등과 같은 특성을 가지고 있다. [14][15]

따라서 본 내용기반 화상검색시스템은 그림 1과 같이 사용자인터페이스를 통해 가지적으로 입력한 질의 화상에 대해서 질의어 형성모듈을 통해 비전루틴에 의해 칼라, 라인, 벡터 특징을 자동으로 추출하여 인덱스 키를 생성하고 특징 DB에 저장된 인덱스 테이블과 유사도검색을 통해서 질의화상과 유사한 화상을 화상 DB로부터 사용자에게 프리젠테이션하도록 구성한다. 원화상과 각특징을 DB에 삽입시에는 사용자 인터페이스의 삽입모듈을 통해 화상처리루틴에 의해 자동특징추출 및 인덱싱되어 원화상은 화상DB에 저장되고, 특징인덱스는 특징 DB에 별도로 저장되게 한다.

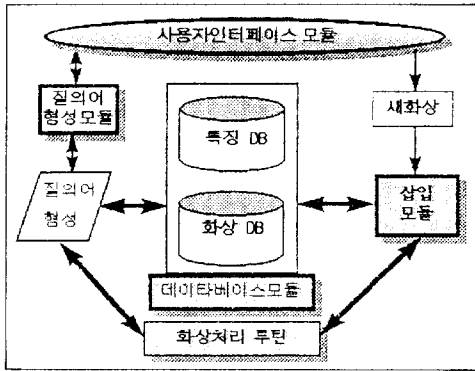


그림 1. 내용기반 화상검색시스템
Fig. 1. Content-based image retrieval system.

1. 화상 특징추출을 위한 전처리 과정

RGB full 컬러 원화상이 입력되면 에지를 검출하기 위해서 256 그레이 레벨로 변환시킨 후 메디안 필터링을 통해 잡음을 제거시킨다. 정규화 단계에서는 필터링된 원화상에 대해서 폭과 높이를 고려하여 원화상의 비율을 그대로 유지하면서 $N \times M$ 으로 정규화 한다. 에지검출 단계에서는 Sobel 에지 검출 방법^[6]을 사용하여 0도, 45도, 90도, 135도 4방향으로 에지를 검출하기 위해 3×3 mask(template)를 각 화상 픽셀에 적용하여 이진화상을 생성해 낸다. 잡음제거 단계에서는 불필요한 에지와 고립된 픽셀은 잡음으로 간주하여 제거한다.

세션화 및 골격선화 단계에서는 에지성분의 분포에 따른 각 영역별 방향성분과 에지분포밀도를 벡터로서 표현하여 특징값을 얻을 수 있는 방향밀도벡터 추출을 위한 핵심 전처리 단계로써, 중앙픽셀을 중심으로 인접 픽셀들간에 $P_2 \rightarrow P_3, P_3 \rightarrow P_4, \dots, P_9 \rightarrow P_2$ 순으로 0→1의 전이수를 카운트하고 나서, $(P_2, P_4, P_6), (P_4, P_6, P_8), (P_2, P_6, P_8), (P_2, P_4, P_8)$ 1인 경우 버퍼 화상에 중앙 픽셀값을 1로 세트한다. 그리고 나서 원 이진화상의 픽셀값과 임시버퍼 화상의 픽셀값을 비교하여 1이면 원 이진화상의 해당 픽셀값을 0으로 지워나감을 반복함으로써 골격선화 시킨다. 그림 2는 원화상의 전처리 결과 및 골격선 추출 3×3 마스크를 보이고 있다.

2. 칼라 특징 추출

1) 색상 양자화(color quantization)

본 논문에서는 256 RGB 칼라를 16 칼라로 양자화하고 각 영역별(4×4 block) 대표칼라성분을 추출하여

검색의 특징값으로 사용한다. 원화상에서 최선의 색상을 선택하는 방법은 popularity방법, mediantcut, 옥트리 양자화(octree quantization)방법, 분산 최소화(variance minimization)방법등이 있으며,^[7] 본 논문에서는 옥트리 양자화방법을 사용하여 원화상을 16($k=16$)칼라로 양자화하였다. 먼저 화상을 순차적으로 읽어 나가면서 새로운 색상을 색상 참조표에 저장한다. 계속 화상을 읽어나가다가 $k+1$ 개의 색상이 되면 가장 가까운 이웃 색상과 통합해 그 평균값을 새로운 색상으로 한다. 만일 새로운 색상이 나타나 $k+1$ 색상이 될 때마다 이웃 색상과 통합해 색상을 k 보다 적은 개수로 유지하면서 화상을 다 읽을 때까지 반복한다.

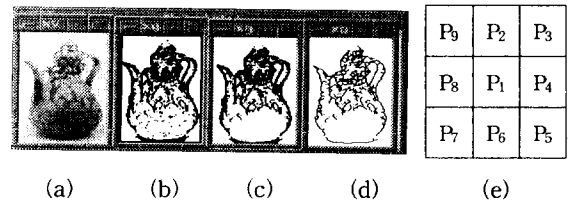


그림 2. 전처리 결과 및 골격선 추출 3×3 마스크
(a) 정규화화상 (b) 에지검출화상 (c)잡음제거 화상 (d)골격선추출화상 (e) 3×3 mask
Fig. 2. Result of preprocessing and 3×3 mask to skeleton.

2) 대표 색상에 의한 칼라특징 추출

입력된 원화상을 150×150 으로 정규화시킨 후에 각 4×4 의 영역으로 나눈 다음, 칼라히스토그램을 구하여 각각의 영역에 대한 대표값을 측정한다. 이때 대표되는 색상성분은 최우선 순위로 대표값을 설정하는 것이 아니라 차순위가 대표값이 된다. 왜냐하면 스캐너에 의해 입력받은 화상의 대부분은 흰색성분을 너무 많이 포함하고 있기 때문에 최악의 경우 모든 영역에 대한 대표색상값이 흰색으로 표현될 수가 있기 때문이다. 또한 16개의 영역을 0~15로 레이블하고 각 인접 영역간의 색상대비에 의해 칼라색상차를 구한다.

3) 칼라 성분에 의한 인덱스키

그림 3은 원화상의 칼라성분에 대한 인덱스키를 보이고 있다. 처음 16개의 칼라성분은 각영역별 대표칼라이고 그다음 15개의 칼라색상차는 인접영역간의 칼라색상차이다. 칼라 성분만의 매칭 효율은 그렇게 좋지 못한 편이지만, 다른 성분과의 다중 인덱싱을 하여 칼

라성분이 필터역할을 수행할 수 있다.

칼라 성분(16개)	칼라 색상차(15개)
------------	-------------

그림 3. 칼라 성분의 인덱스키 형식

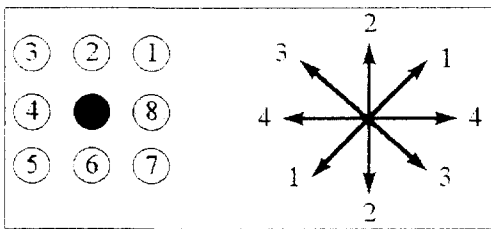
Fig. 3. Index key format of color component.

3. 방향밀도벡터 특징추출

1) 방향밀도벡터 코드화

내용기반 화상 검색을 위한 화상특징으로 사용한 방향밀도벡터는 사용자가 질의어로 입력한 화상의 여러 특징량, 즉 위상정보, 밀도정보, 위치정보등을 포함한 특징량으로써, 에지의 기울기와 위치 및 양과 같은 정보를 전부 포함하는 정보량이다. 화상내에서 방향밀도벡터의 추출은 그림 4와 같이 위상방향 마스크(mask)를 가진 전체 화상의 패턴 영역을 주사하여, 각 위상방향에서 방향선분별로 방향선소를 합한 것이 된다.

전처리가 끝난 골격선택된 이진화상내의 모든 object 들은 수평, 수직, 대각, 역대각의 에지성분으로 이루어진 형태를 가지며 이러한 특성을 검색시 특징으로 활용하기 위해서는 각 픽셀을 주변의 픽셀위치 및 분포에 따라 마스크내의 중앙픽셀을 4방향코드 벡터화를 행한다.



(a) (b)

그림 4. 방향밀도벡터의 4방향 마스크 (a) 3x3 mask (b) 4방향벡터코드

Fig. 4. Four direction mask of density vectors. (a) 3 x 3 mask (b) four direction vector code

2) 방향밀도벡터 추출

150xN 형태로 정규화된 전체화상영역에 대해서 4x4 block으로 영역을 나눈 후 각 영역별로 방향밀도 벡터 성분비율을 구하여 인덱스키값으로 변환시킨다. 먼저 전체화상영역에 대해 산출한 총방향밀도벡터의 수를 1,2,3,4 각 방향별로 산출한 방향밀도벡터의 수로 나눈 백분율로 화상의 1,2,3,4 대표방향성분 비율 4개를 구한다. 그리고, 전체 화상을 16개(4x4block) 영역으로

분할한 다음 각 영역별로 산출한 총방향밀도벡터의 수를 1,2,3,4 각 방향별 총방향밀도벡터 수로 나눈 백분율로 16개 영역별 방향밀도벡터 성분비율을 구한다. 마지막으로, 각 영역에 대해 픽셀의 분포비를 산출하기 위해서는 전체 4방향 밀도벡터수를 각 영역별 총 방향 밀도벡터수로 나누어서 각 영역별 픽셀의 분포비를 구하였다. 그림 5는 영역 #6의 방향밀도벡터추출의 일부를 예로 보이고 있다.

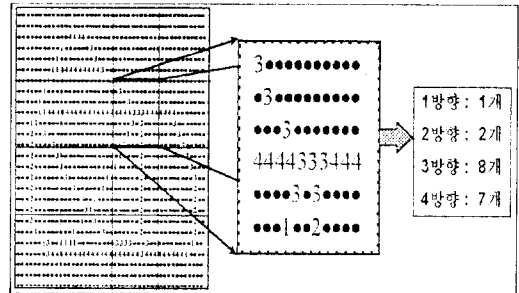


그림 5. 방향밀도벡터 추출결과

Fig. 5. Result of density vector extraction.

3) 방향밀도벡터에 의한 인덱스키 형식

전체화상의 1, 2, 3, 4 각 방향별 대표방향밀도벡터 성분 비율 4개와 각 16개 영역에 대한 1, 2, 3, 4 방향별 대표방향밀도벡터 성분비율 64개에 각 영역별 픽셀 분포비율 16개를 각각 곱함으로써 각 영역의 4방향 인덱스값을 계산해 내었다. 따라서, 각 영역의 4방향 벡터성분의 인덱스값은 총픽셀에 대한 방향분포비라고 할 수 있으며, 이는 각 영역별 4방향 밀도벡터 성분비율(%)에 각 영역의 픽셀분포비율(%)을 곱하여 계산한다. 방향밀도벡터에 의한 인덱스키의 형식은 그림 6과 같다. 전체대표방향은 화상의 1,2,3,4각 방향별 대표방향 밀도 벡터비율을 나타내고, block#은 각영역별 총픽셀에 대한 각방향별 성분비율을 포함한 픽셀의 분포비율을 나타낸다.

전체대표방향				block #1				block #2				...	#block 16			
1	2	3	4	1	2	3	4	1	2	3	4	...	1	2	3	4

그림 6. 방향밀도벡터 성분의 인덱스키 형식

Fig. 6. Index key format of density vector components.

4. 라인 특징 추출

1) Hough 변환에 의한 라인성분 추출

M x N의 크기를 갖는 이진화상은 x, y의 2차원좌

표평면으로 생각할 수 있다. 따라서 화상의 어느 한점 (x, y) 를 지나는 임의의 직선을 $y = mx + c$ 라고 하는 직선의 방정식으로 나타낼 수가 있다. 여기서 m 은 직선의 기울기이며 c 는 직선의 y 절편이라고 보면 하나의 점이 하나의 직선을 나타내는 다른 파라미터로서 일대일 대응이 되는 것을 알 수 있다. 그림 7은 Hough 변환의 예를 보이고 있다.

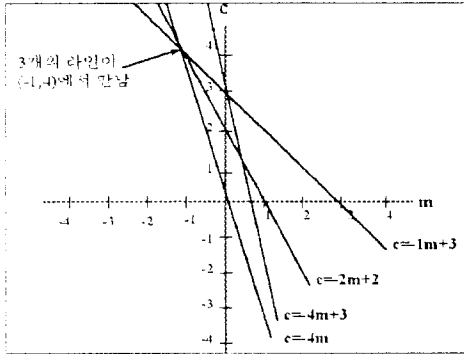


그림 7. Hough 변환의 원리
Fig. 7. Principle of hough transform.

이와 같이 이진화상의 모든점에 대해서 (x, y) 에 위치하는 점 (m, c) 값을 갖는 또 다른 평면으로 변환하면, 그 교점은 (x, y) 좌표계에 있는 점들이 공통으로 가지는 (m, c) 로 볼 수 있으며 만나는 직선의 수가 많을 수록 그 직선위에 위치하는 점의 수가 많음을 나타낸다. 그런데 각각의 pixel마다 변환 결과로서 m (기울기)와 c (y 절편)이 무한대의 값을 가질 수 있다. 이는 좌표축에 대해서 수직적인 직선을 표현하는 경우 나타나는데 이렇게 되면 파라미터의 값을 무한대로 잡아야 하는데 실제 값을 저장하는 구조가 이런 형태를 가질 수는 없으므로 이 문제를 보완하는 방법으로써, 그림 8과 같이 Duda & Hart가 제안한 polar coordinate (r, θ) 를 사용하였다. (x, y) 좌표를 (m, c) 파라미터가 아닌 (r, θ) 파라미터로 변환한다. 따라서 모든 (x, y) 좌표에 대해서 변환을 하면 위의 곡선들이 만나는 점을 찾을 수 있고 이는 직선에 해당하는 점들의 r, θ 값이 된다. (r, θ) accumulator array의 크기로서 r 은 화상 대각선의 길이만큼, θ 는 $-90 \sim 90$ 만큼 설정하였다.

2) 라인성분에 의한 인덱스키 형식

전처리된 이진화상을 입력으로 하여 각 점에 대해서 Polar Coordinate Hough transform을

수행한 다음, (r, θ) 배열중 그 값이 일정한 임계값을 넘는 것에 대해서만 역변환(inverse hough transform)하여 직선성분이 큰 점들을 얻을 수 있으며, 이들 직선중에서 유사한 직선들을 하나로 통합하기 위하여 선분통합의 과정을 거치게 되면 최종적으로 4개의 직선만을 골라 직선의 양끝좌표를 화상의 인덱스키로 하였다. Hough변환에 의한 인덱스 key 형식은 다음 그림 9와 같다.

$$x \cos \theta + y \sin \theta = r$$

$$\begin{aligned} r &= \frac{x}{\cos \theta} + (y - x \tan \theta) \sin \theta \\ &= \frac{x}{\cos \theta} + y \sin \theta - x \frac{\sin^2 \theta}{\cos \theta} \\ &= x \left(\frac{1 - \sin^2 \theta}{\cos \theta} \right) + y \sin \theta \\ &= x \cos \theta + y \sin \theta \end{aligned}$$

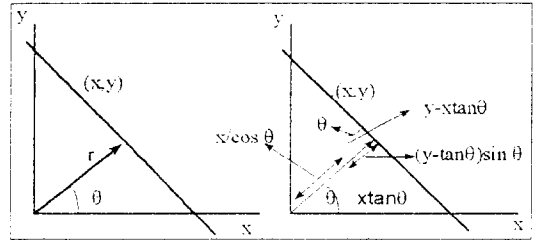


그림 8. Polar Coordinate Hough 변환의 개념도
Fig. 8. Principle of polar coordinate hough transform.

← line #1 →		← line #2 →		...	← line #4 →	
라인 시작 좌표	라인 끝 좌표	라인 시작 좌표	라인 끝 좌표		라인 시작 좌표	라인 끝 좌표
					라인 시작 좌표	라인 끝 좌표

그림 9. 라인성분의 인덱스키 형식
Fig. 9. Index key format of line components.

5. 내용기반 화상 검색시스템의 매칭 및 검색기법

데이터베이스에 저장되는 화상인덱스는 일련의 화상 특징추출과정을 거치면서 16개영역별 대표칼라16개 및 인접영역간의 칼라색상차이 15개로 구성된 31개의 칼라 인덱스키와, 화상의 1,2,3,4 대표방향 4개 및 16개 각영역별 1,2,3,4 방향벡터성분 64개로 구성된 68개의 방향벡터 인덱스키, 그리고 화상당 4개의 라인에 대한 시작좌표 및 끝좌표로 구성된 16개의 라인성분에 대한 인덱스키를 합한 화상당 총 115개의 인덱스키를 생성하여 저장한다. 만약 사용자가 화상을 질의로 줄 경우 질의화상 역시 같은 화상 특징추출과정을 거치게 되고 115개의 특징값을 인덱스로써 얻게 된다. 따라서 유사

화상을 검색하기 위해서는 완전 비교(exact matching)가 아닌 유사도비교(similarity matching)를 통해 일정 범위내에서 주어진 값이 만족하는지를 평가해야만 한다. 따라서 질의로 주어진 질의화상 인덱스키와 데이터베이스내에 저장된 원화상의 인덱스키의 각 필드를 모두 비교하게 된다. 각 필드에 대한 매칭방법은 그림 10과 같다.

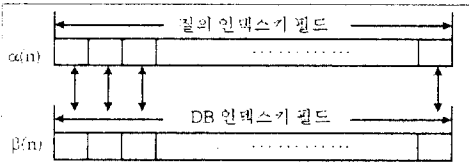


그림 10. 인덱스 키의 유사도 매칭
Fig. 10. Similarity matching of index key.

$\alpha(n)$ 이 질의화상 인덱스키값중 n 번째 필드라하고 $\beta(n)$ 을 데이터 베이스내에 저장된 원화상의 인덱스키값 n 번째 필드라 하면 threshold τ 에 대해서 $\tau \leq |\alpha(n) - \beta(n)|$ 를 만족하면 match counter를 증가 시킨다. 따라서 115 개의 필드가 모두 위의 조건을 만족할 경우 match counter는 115가 된다. 결국 모든 비교가 끝나면 원화상 인덱스키에서 해당 화상의 위치와 match count 값의 쌍으로 구성된 결과를 얻게 된다.

검색결과는 칼라, 벡터 및 라인성분에 대한 비교가 되므로 검색시 특성성분에 대한 검색가중치를 부여할 수 있도록 하였다. 이는 사용자가 질의한 화상에 대해서 추출된 벡터, 칼라, 라인성분에 의하여 검색할 경우 특정 한가지 성분의 비중을 높여서 검색할 수 있도록 융통성을 부여한 것이다. 따라서 만약 사용자가 벡터성분 위주로 검색하고자 한다면, 벡터성분의 가중치 α 를 높게 지정하고 라인성분의 가중치 β 와 칼라성분 γ 가중치를 낮게 지정함으로써 검색할 수 있게 하였다. 이들 가중치의 범위는 0에서 1.0 사이의 값을 갖는다. 이를 위하여 칼라, 벡터, 라인의 3가지 성분에 대해 다음과 같이 연산한다.

a) 각 index 성분의 최대길이를 고려하여 match count수를 백분율로 나타낸다. 이를 normalized score S이라고 하면 각특징들의 normalized score 는 다음과 같이 계산한다.

- 벡터의 normalized score : $V_s = \frac{\text{matched count수}}{\text{총인덱스수}} \times 100$

- 라인의 normalized score : $L_s = \frac{\text{matched count수}}{\text{총인덱스수}} \times 100$

- 칼라의 normalized score : $C_s = \frac{\text{matched count수}}{\text{총인덱스수}} \times 100$

b) 백분율로 나타낸 각 index 성분의 normalized score S에 가중치를 곱한다.

- 벡터의 유사도 : αV_s (α : 벡터 성분의 가중치를 나타내는 계수값)

- 라인의 유사도 : βL_s (β : 라인 성분의 가중치를 나타내는 계수값)

- 칼라의 유사도 : γC_s (γ : 칼라 성분의 가중치를 나타내는 계수값)

c) 각 index별로 계산된 유사도 $\alpha V_s, \beta L_s, \gamma C_s$ 을 더함으로써 최종 검색유사도를 측정한다.

- 검색 유사도(similarity) $SIM = \alpha V_s + \beta L_s + \gamma C_s$

d) 이를 match count의 값으로 decending sort 할 경우 유사도가 가장 높은 화상순으로 검색하여 사용자에게 프리젠테이션하게 된다.

III. 내용기반 화상검색시스템의 구현

본 내용기반 화상검색시스템의 구현환경은 Sun-Sparc Classic의 X-window상에서 C와 Motif를 사용하였고, 화상저장을 위한 데이터베이스 시스템으로 Exodus를 사용하였다.¹⁸⁾ Exodus는 위스콘신 대학에서 개발된 확장가능한 DBMS로 저장객체(storage object)라는 임의의 크기를 갖는 가변 레코드를 지원하는 특징을 갖고 있다. 본 논문에서 사용된 Exodus storage manager는 gcc 2.5.8컴파일러를 사용하는 버전3.1으로 SunOs 4.1.3에서 운용된다. 그리고 사용자 인터페이스 개발도구로서 UIL(User Interface Language)를 사용하였고, 사용자의 질의나 화상 데이터를.gif형식으로 입력을 받을 수 있도록 256 컬러 스캐너를 사용하였다.

그림 11은 내용기반 화상검색시스템의 비전루틴 구성도를 보이고 있다. 내용기반검색을 위해 본 논문에서 사용한 특징들은 칼라, 벡터, 라인성분들으로써, 원화상이 입력되면 칼라성분의 경우는 16칼라로 양자화하여 화상의 대표칼라와 16개영역별 대표칼라 및 인접영역간의 색상차를 추출하여 인덱싱하게 되며, 방향밀도벡터의 경우는 전처리된 이진화상에 대해서 골격선화한다

음 방향코드화를 행하고, 방향코드벡터를 추출하여 화상의 전체 및 16개영역별로 대표방향밀도성분비율과 픽셀의 분포비율을 추출하여 인덱싱한다.

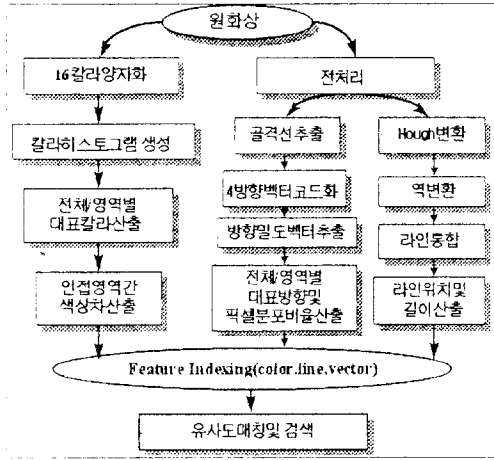


그림 11. 내용기반 화상검색시스템의 비전루틴의 구성도
 Fig. 11. The configuration of vision routine in the content-based image retrieval system.

여기서 전처리된 이진화상이란 256칼라로된 원화상에 대해 그레이레벨 변환과 메디안 필터링 및 정규화를 거쳐서 에지검출된 배경을 0으로 윤곽선을 1로 표시한 이진화상이다. 라인성분의 경우에는 전처리한 이진화상에 대하여 Hough 변환 및 역변환을 행하여 라인을 통합한 다음, 화상당 4개라인을 추출하여 인덱싱을 하였다. 추출된 벡터, 라인, 칼라의 각 특징에 대한 부여된 검색가중치를 통해 특징 DB내 인덱싱 table과 유사도 매칭 및 검색이 수행되며 상위 10순위내의 후보자를 순서적으로 축소된 크기(100x100)로 사용자에게 디스플레이한다.

IV. 실험결과 및 고찰

검색에 사용된 화상은 청자 65개, 백자 50개, 불상 30개, 석탑 40개, 건축물 등 기타 15개 등 총 200여개의 화상이다. 사용자는 전체화상과 스케치를 사용하고, 전체화상의 일부분양을 마우스로 선택하여 질의어로 사용하였다. 실험조건은 총 200개의 DB내 화상중에서 임의의 질의화상 10개를 무작위로 추출해서 각각 10번 씩 검색 테스트 하였다. 검색된 화상수는 10개가 넘을

수도 있고 10개 미만일 수도 있으며, 디스플레이는 유사도가 높은 순으로 최대 10개까지 화면에 출력하도록 하였다.

내용기반 멀티미디어 정보 검색시스템에서는 기존의 정확한 매칭이 아닌 유사매칭을 하므로 시스템의 평가도 다르게 계산해야 한다. 일반적으로 정확한 매칭이 아닌 유사매칭을 수행하는 시스템의 평가는 normalized precision와 recall statistics를 근거로 하고 있다.^[9] 주어진 질의와 관련된 총 항목수를 T라 하고, T로부터 검색된 항목의 총수를 T_r이라 하며, 검색된 항목중 질의와 유사한 항목의 수를 R_r이라고 할때 precision은 R_r / T_r 이고, recall statistics는 R_r / T 이 된다. 평균 순위는 검색된 10개 화상중 질의화상과 유사한 화상의 수를 rank된 순위의 합으로 나눈 것이 된다.

표 1은 검색가중치를 각각 0.8V_s + 0.1L_s + 0.1C_s로 하였을 경우 질의 방법에 따른 실험결과를 나타내고 있다. 검색시 벡터성분의 가중치를 높게한 이유는 칼라는 16칼라에 의한 특징추출이기 때문에 화상의 칼라특징을 정확하게 반영하지 못하고 있고, 라인의 경우에는 직선성분을 많이 가지고있는 검색대상화상들이 적기 때문에 모든종류의 화상에 적용가능한 벡터성분의 비중을 높게함으로써 정확성을 높이기 위함이었다. 표 1의 실험결과를 살펴보면 전체화상과 스케치의 검색결과를 비교해보면 전체화상이 precision 94%, recall 40%로 스케치의 각 87%, 36%보다 좋은 검색결과를 보임을 알 수 있으며, 문양의 경우는 같은 문양을 갖는 화상만 정확하게 검색됨을 알 수 있다.

표 1. 질의방법에 따른 실험결과
 Table 1. Experiment results by query types.

질의유형 파라메타	전체화상	스케치	문 양
DB size	200	200	50
Test query 수	10	10	10
Precision	0.94(80/85)	0.87(72/82)	1.0(12/12)
Recall	0.40(80/200)	0.36(72/200)	0.24(12/50)
평균 순위	3.0	3.2	1.0

그림 12는 검색가중치를 각각 0.8V_s + 0.1L_s + 0.1C_s로 하여 유사도매칭한 전체화상의 검색결과를 보이고 있으며, 그림 13은 동일한 조건에서 유사도매칭한

스케치의 검색결과를 보이고 있다. 그리고 그림 14는 문양에 대한 검색결과이다. 문양의 경우에는 벡터성분만을 이용하여 검색실험을 하였다.

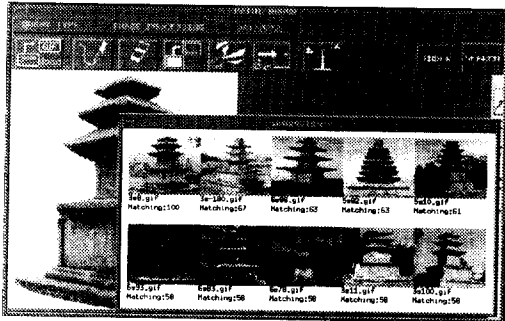


그림 12. 전체화상에 의한 화상검색결과
Fig. 12. Result of image retrieval by whole image.

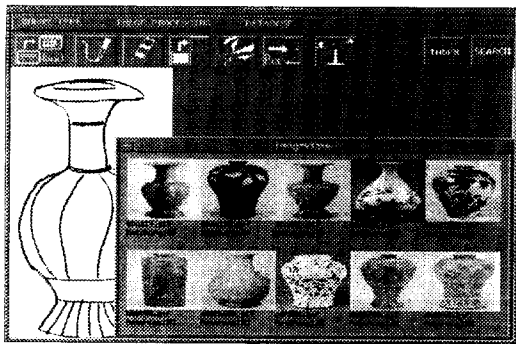


그림 13. 스케치에 의한 화상검색 결과
Fig. 13. Result of image retrieval by sketch.

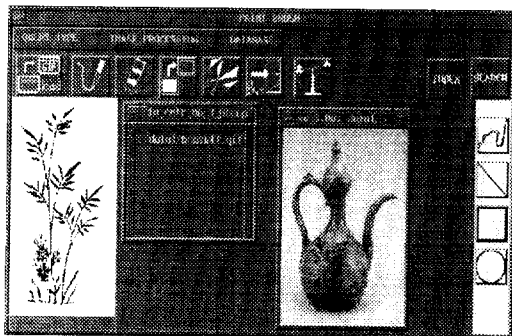


그림 14. 문양에 의한 화상검색 결과
Fig. 14. Result of image retrieval by object.

표 2는 검색에 사용된 특징성분에 따른 검색실험 결과를 나타내고 있으며, 다음과 같은 측정파라미터를 사

용하였다. 평균 순위는 검색된 10개 화상중 질의화상과 유사한 화상의 수를 rank된 순위의 합으로 나눈 것이 된다. 따라서 만약 검색된 10개의 화상중에 유사한 화상이 7개이며 각각 1순위, 2순위, 3순위, 4순위, 5순위, 7순위, 9순위에 위치해 있다면 평균순위는 $(1+2+3+4+5+7+9)/10$ 가 된다. 검색순위는 질의화상이 몇순위로 검색되는가를 나타낸 것이며, 검색실패는 검색 후 디스플레이된 화상중에서 질의화상과 다른 화상의 개수이다.

표 2. 특징성분에 따른 실험결과
Table 2. Experiment results of feature component.

질의유형 / 파라메타	칼라	벡터	라인
DB size	200	200	200
Test query 수	10	10	10
평균 순위	5.5	3.2	4.5
검색 순위	1.0	1.0	1.0
검색 실패	0/10	0/10	0/11

표 2의 결과를 보면, 평균순위에서 칼라는 5.5이고 라인은 4.5로 평균순위가 벡터의 3.2인 평균순위에 훨씬 못미치고 있으며, 이는 칼라의 경우는 16칼라 양자화에서 비롯된 칼라특징 추출의 정확성 결여로 판단되며, 라인의 경우는 도자기등 직선성분이 적은 화상에 대해 오검출된 라인특징추출 때문으로 판단된다. 그러나 벡터의 경우는 윤곽성분의 분포를 방향벡터화하였으므로 물체의 위치나 모양, 각도에 따라 정확한 특징추출이 가능했기 때문에 검색효율이 좋은 것으로 나타났다.

표 3. 특징성분에 따른 QBIC의 실험결과
Table 3. Experiment results of feature component in QBIC.

질의유형 / 파라메타	칼라	모양
DB size	1,000	295
Test query 수	10	10
평균순위	5.4	8.6
검색순위	1.1	n/a
검색실패	3/72	11/64

표 3은 IBM의 QBIC시스템 [9]에서의 1,000개 DB크기, 검색테스트 10회, 디스플레이 20개의 실험조건에서 칼라와 모양특징을 사용하여 검색효율을 실험한 결과를 나타내고 있다.

본 논문과 QBIC의 성능비교는, DB의 크기와 후보자 디스플레이의 개수가 다르므로 정확한 비교평가는 불가능하지만, 전체적으로 볼 때 칼라의 경우 검색효율이 다소 떨어짐을 제외하면 QBIC에 비해 본 논문의 검색효율이 다소 우수함을 알 수 있었다.

본 논문에서 구현한 화상검색시스템은 입력한 RGB 256 칼라 화상에 대해 칼라, 라인의외도 벡터성분에 의한 모양특징을 추출함으로써 물체의 위치나 크기에 영향을 받지 않고 모든 화상에 적용할 수 있는 시스템이다. 왜냐하면, 방향벡터성분은 화상에서 검색대상이 되는 물체의 윤곽선을 방향벡터화한 것으로써, 이렇게 방향벡터화한 특징값들은 물체의 거리, 방향, 인접요소와의 이루는 각도등을 평가하여 화상특징간의 유사성을 비교할 수 있으므로 물체의 위치나 크기에 영향을 받지 않고 검색할 수 있다는 장점과 여러개의 물체들이 있는 경우 따로 식별할 수 있기 때문이다.

그러나, 향후 보완해야 할 문제점으로는 가중치에 의한 검색시 칼라특징의 경우는 16칼라로 양자화하여 각 영역별 대표칼라와 인접영역간의 칼라차이를 이용하여 검색하기 때문에, 칼라성분 검색가중치를 높였을 경우 검색효율이 낮아지므로 인덱스키 크기를 줄이면서 256칼라로 인덱싱할 수 있는 방법 연구가 필요하다. 라인 특징의 경우에는 도자기처럼 화상자체가 직선성분이 없는 둥근 모양일 경우에는 정확한 라인성분이 추출되지 않기 때문에, 역시 라인성분 검색가중치를 높였을 경우 검색효율이 낮아지므로 선분통합후 선분제거의 후처리가 필요하며, 라인보다는 곡선성분을 추출하여 인덱싱하는 연구가 필요하다.

또한 칼라, 라인, 벡터의 특징뿐 만 아니라, 화상의 질감과 영역, 원형성, 이심률, 중심축, 모멘트등의 구체적인 모양특징들을 경험적으로 조합하여 검색할 경우 검색효율을 향상시킬 수 있다.

V. 결 론

본 논문에서는 내용기반 검색을 위해 키워드가 아닌 시각적인 예(visual example)로써 질의하여 원하는 화상을 쉽게 검색할 수 있는 내용기반 화상검색시스템

을 설계하고 구현하였다. 검색을 위한 전처리 모듈에서는 기존의 화상처리기술을 응용한 필터링, 정규화, 예지검출, 세션화기법을 사용하였고 방향밀도 벡터추출을 위한 에지의 골격선추출 알고리즘을 개발하였으며, 라인성분을 추출하기 위해 Hough 변환 알고리즘을 수정, 응용하였다. 칼라성분을 추출하기 위해서 RGB공간상의 256칼라를 16칼라로 양자화시킴으로써 화상의 영역별 대표칼라 성분을 추출하였다.

본 논문에서 개발한 화상검색시스템은 질의종류에 따라 전체화상, 스케치와 문양으로 나누어 질의가능케 하였으며, 칼라, 라인, 방향밀도벡터에 의한 검색방법을 사용하므로 모든 종류의 화상검색에 적용할 수 있다. 또한, 사용자가 원하는 화상을 검색시 각 특징들을 가중치에 따라 조합하여 검색가능하므로, 단일 특징을 사용한 검색보다 검색효율과 정확도를 높일 수 있었다.

참 고 문 헌

- [1] Tzi-cker Chiueh, "Content-based image indexing," Proceedings of the 20th VLDB Conference, San Francisco, pp. 582-593, 1994.
- [2] M.J. Swain, "Interative indexing into image databases," SPIE, San Jose, vol. 1908, pp. 95-103, 1993.
- [3] Kyoji Hirata, Toshikazu Kato, "Query by visual example," Extending database Technology '92, pp. 56-71, 1992.
- [4] W.Niblack, R.Barber, et al., "The QBIC Project: Querying images by content using color, texture, and shape," SPIE, San Jose, vol. 1908, pp. 173-187, 1993.
- [5] R. Barber, et al., "Efficient query by image content for very large image databases," COMPCON, San Francisco, 1993.
- [6] Ioannis Pitas, "Digital image processing algorithms," Prentice Hall, 1993.
- [7] Haruhiko Nishiyama, Sumi Kin, et al., "An image retrieval system considering subjective perception," CHI94 Human Factors in Computing Systems, Boston, pp 30-36, 1994.
- [8] "Using the EXODUS storage manager V3.1," 1993.
- [9] R.Barber, W.Equitz, et al., "Query by con-

tent for large on-line image collections," A Guided Tour of Multimedia Systems and

Applications, IEEE Computer Society press, pp 357-378, 1995.

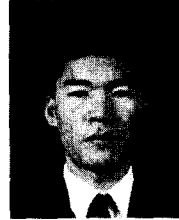
저 자 소 개



鄭元一(正會員)

1941년 7월 31일생. 1964년 2월 한국항공대학교 전자공학과 졸업. 1979년 8월 전국대학교 대학원 전자공학과(공학석사). 1996년 8월 광운대학교 대학원 전자계산기공학과 박사. 1973년 4월 ~ 현

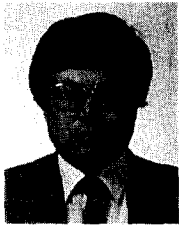
재 대전산업대학교 전자공학과 교수. 주관심분야는 멀티미디어, 하이퍼미디어, 멀티미디어정보검색, 회로 및 시스템등임



崔炫燮(正會員)

1964년 5월 16일생. 1987년 2월 목포대학교 전산통계학과 졸업. 1989년 2월 광운대학교 대학원 전자계산기공학과(공학석사). 1991년 9월 ~ 1993년 2월 목포대학교 전산통계학과 조교 1993년 3월 ~ 현재 광운대

학교 대학원 전자계산기공학과 박사과정중. 주관심분야는 하이퍼미디어, 멀티미디어정보시스템, 멀티미디어정보검색등임



崔基浩(正會員)

1949년 2월 17일생. 1973년 2월 한양대학교 전자공학과 졸업. 1977년 2월 한양대학교 대학원 전자공학과(공학석사). 1987년 2월 한양대학교 대학원 전자공학과(공학박사). 1977년 3월 ~

1979년 2월 한국과학기술연구소 연구원. 1979년 3월 ~ 현재 광운대학교 컴퓨터공학과 교수. 1989년 2월 ~ 1990년 2월 Univ. of Michigan 전기 및 전산과 Visiting Scholar. 주관심분야는 하이퍼미디어, 멀티미디어정보시스템, 멀티미디어정보검색, 멀티모달등임.