

# DTW(Dynamic Time Warping)를 이용한 영상 정보 검색

하정요\* 이나영\*\*, 김계영\*\*\*, 최형일\*\*\*\*

## 요 약

영상을 검색하는 데에는 형태, 색상, 질감 등 여러 가지 특징을 사용 할 수 있다. 그 중 가장 활발한 연구가 이루어지고 있는 분야가 형태 정보와 색상 정보를 이용하는 분야이다. 형태 정보를 이용하는 검색 방법에는 2D 모멘트와 푸리에변환 등의 방법이 유명하다. 또 다른 방법으로는 CSS(Curvature Scale Space)가 있는데, 이는 윤곽선 정보를 이용하여 윤곽선의 굴곡을 2D 그래프로 표현하여 그 그래프의 극대점을 특징 값으로 사용하여 영상을 비교 검색하는 방법이다. 기존 CSS 방법에는 몇 가지 문제점이 있어서 본 논문에서는 기존 방법을 향상시킨 ICSS 방법을 사용하여 영상을 검색한다. 색상 정보를 이용하는 방법에는 RGB 색상정보를 이용하는 방법과 HSI 색상정보를 이용하는 방법 등이 있는데 본 논문에서는 HSI 색상정보를 이용하여 색상 히스토그램으로 표현한 후 영상의 비교 척도로 사용하였다. 영상의 유사도를 측정하는 방법으로는 유클리디언 디스턴스를 주로 사용하는데, 본 논문에서는 정확도와 검색 시간을 단축시키고자 DTW를 사용하여 영상의 유사도를 측정하였고, 유클리디언 디스턴스를 사용했을 때와 비교하여 성능 향상 결과를 보인다.

## Image Information Retrieval Using DTW(Dynamic Time Warping)

Jeong-Yo Ha\*, Na-Young Lee\*\*, Gye-Young Kim\*\*\*, Hyung-il Choi\*\*\*\*

## Abstract

There are various image retrieval methods using shape, color and texture features. One of the most active area is using shape and color information. A number of shape representations have been suggested to recognize shapes even under affine transformation. There are many kinds of method for shape recognition, the well-known method is Fourier descriptors and moment invariant. The other method is CSS(Curvature Scale Space). The maxima of curvature scale space image have already been used to represent 2-D shapes in different applications. Because preexistence CSS exists several problems, in this paper we use improved CSS method for retrieval image. There are two kinds of method, One is using RGB color information feature and the other is using HSI color information feature. In this paper we used HSI color model to represent color histogram before, then use it as comparison measure. The similarity is measured by using Euclidean distance and for reduce search time and accuracy, We use DTW for measure similarity. Compare with the result of using Euclidean distance, we can find efficiency elevated.

Keywords : Image Retrieval, shape, CSS, ICSS, DTW, Euclidean Distance

## 1. 서 론

최근 디지털 영상, 비디오, 음악 등 멀티미디어 정보의 공유가 가속화되면서 이들을 효과적으로 검색할 수 있는 시스템의 필요성이 대두되고 있다. 검색 시스템들은 주로 텍스트 주석 등을 검색에 이용하였지만 이러한 방식은 멀티미디어 정보를 검색하는데 비효율적이며 텍스트 주석 작성에 많은 시간이 소요된다. 때문에 근래 들어 자신이 원하는 그림, 영화의 한 장면, 특정 음악의 일부나, 효과음 등으로 질의하는 내용기반 검색 시스템

※ 제일저자(First Author) : 하정요

접수일:2009년 07월 21일, 완료일:2009년 9월 23일

\* 숭실대학교 대학원 미디어학과 박사과정

[saint@ssu.ac.kr](mailto:saint@ssu.ac.kr)

\*\* Carnegie Mellon University, Robotics Institute 방문연구원

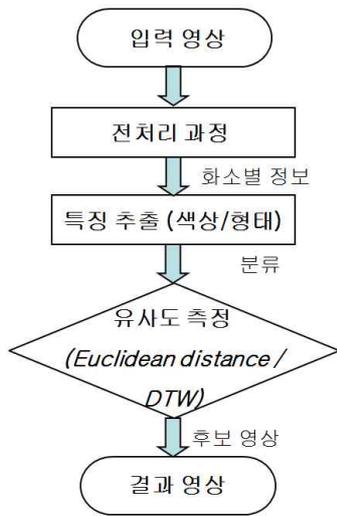
\*\*\* 숭실대학교 컴퓨터학과 교수

\*\*\*\* 숭실대학교 미디어학과 교수

▣ 본 연구는 서울시 R&BD 프로그램의 지원받아 수행하였음(10581cooperateOrg93112).

템에 대해 많은 연구가 이루어지고 있다[1][2][3]. 정지 영상에 대한 일반적인 내용기반 검색 기법들은 영상의 색상이나 질감, 형태 등의 특징을 사용한다. 이러한 시스템에서는 처음에 영상의 특징들이 자동으로 추출되어 영상과 함께 색인되어 데이터베이스에 저장된다. 사용자는 질의로 원하는 색이나 질감 등을 포함하는 영상을 요청하거나 원하는 영상의 형태를 스케치하는 방식으로 내용 기반 질의를 할 수 있다. 질의 결과는 정확한 영상이 출력되기 보다는 요청한 영상과 비슷한 영상들의 집합이 출력된다. 본 논문에서는 사전에 미리 학습된 색상 정보와 이미지의 영상의 정보들을 추출하고 질의 영상과 데이터베이스에 저장된 영상의 유사도를 측정하는 방법으로 DTW 알고리즘을 사용하였다.

2장에서는 유사도를 측정하기 전에 유사하지 않은 영상을 제거하는 방법으로 HSI 색상 공간에서 H(Hue)값을 가지고 색상이 유사하지 않은 영상을 제거한 뒤, 3장에서 ICSS 방법을 이용하여 모양 정보를 추출한다. 4장에서는 DTW 알고리즘을 사용하여 유사도를 측정한다. 5장에서는 실험 및 결과를 보이며, 마지막 6장에서는 결론 및 향후 해결 방안에 대하여 기술한다.



(그림 1) 전체 시스템 구성도

(그림 1)은 본 논문에서 제안하는 시스템 구성도이다. 전처리 과정, 특징 추출, 유사도의 측정으

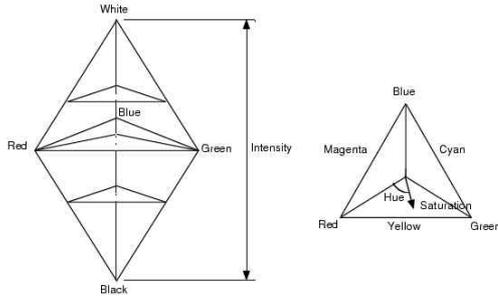
로 크게 3단계로 이루어진다. 전처리 과정에서는 후속 단계인 특징 추출을 위해 기본적인 영상 처리과정을 수행한다. 색상 특징을 위해서는 화소별 gray-level 정보를 추출한다. 특징 추출 단계에서는 시각적인 특징을 추출한다. 특징 추출 기법은 검색을 위한 유사성 정도를 측정 할 수 있는 기준으로 색상, 형태 등의 기본적인 특징들을 설정하여 특징 벡터를 구성한다. 마지막 단계인 유사도 측정단계에서는 유클리언 디스턴스 방법과 DTW(Dynamic Time Warping)알고리즘을 이용하여 주어진 조건에 가장 유사한 영상들을 추출하여 보여주게 되는데 사용자가 예제 영상을 가지고 질의를 하면 예제 영상의 특징 벡터와 영상 데이터베이스에 저장된 특징 벡터간의 유사도를 계산하여 유사도가 큰 순위대로 영상들을 출력하게 된다.

## 2. 색상정보 추출

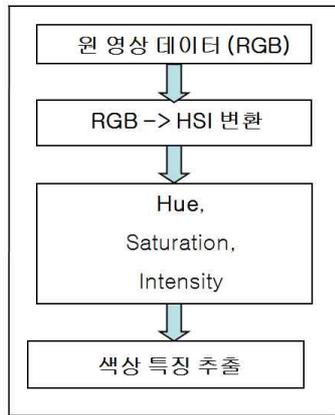
최근 색상정보의 가장 보편적인 이미지 정보 추출 방법은 영상의 색상 히스토그램을 사용함으로써 정보를 표현할 수 있다[7]. 이 방법은 각색상에 대해 이미지 픽셀 빈도를 나타내는 방법으로 색상 히스토그램은 색상에 대한 빈도수만으로 표현하므로 이미지의 회전이나 크기 및 위치 변화, 적은 시야변경에 대해 유사성을 잃지 않는 장점을 가진다. 그러나 색상 히스토그램 방식은 모든 색상에 대해 그 빈도수를 가지게 되므로 정보의 차원이 높다는 문제점을 가진다.

HSI 색상모델은 인간의 시각과 가장 유사하기 때문에 많이 사용되므로 본 논문에서는 HSI 색상 모델을 사용하였다. 그레이 영상을 사용할 경우에는 주위의 밝기 변화에 대하여 상당히 의존적이고 RGB 색상정보를 사용할 경우에는 주위의 밝기 변화에 대하여 상당히 의존적이고 RGB 색상정보를 사용할 경우에는 RGB 각 성분이 명암도와 밀접한 관련을 가지고 있으므로, 주위 밝기 변화에 대해서 같은 색상이라도 RGB 각 값들이 모두 변하게 되는 문제점이 있다. 그에 반해 HSI 색상모델은 밝기 변화에 대한 강인성을 보장한다. HSI 색상모델은 색채요소 (Hue, Saturation)와 명암요소(Intensity)를 분리하여 명암 요소를 제거함으로써 영상 획득시의 조명 변화의 영향을 줄일

수 있는 장점이 있다. H 값은 색상 집합을 통해서 0~360도의 각도를 가지고 S는 각도에 대한 색상의 세기, 즉 채도를 나타낸다. S 값이 높아질수록 높은 순도의 색상이 된다. I 값은 명도로 색상의 어두움과 밝기를 나타낸다.



(그림 2) HSI 컬러 공간에서의 컬러표현



(그림 3) 색상정보 추출과정

색상 특징 추출단계에서는 전처리 단계에서 추출된 화소별 RGB 색상 정보 및 그레이 정보를 가지고 영상들의 시각적인 특징들을 추출한다. RGB 색상모델을 HSI 색상모델로 변환 후 색도(H), 채도(S), 명도(I)를 추출하여 히스토그램의 교집합을 사용한 히스토그램 인터섹션(Histogram Intersection)을 사용하였다. HSI 색상모델에서 특징 추출방법은 밝기 값과 색상을 주로 이용한다.

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}, \quad \theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(r-g) + (r-b)]}{[(r-g)^2 + (r-b)(g-b)]^{\frac{1}{2}}} \right\} \quad (1)$$

$$S = 1 - \frac{3}{(r+g+b)} [\text{Min}(r, g, b)] \quad (2)$$

$$I = \frac{1}{3}(r+g+b) \quad (3)$$

본 논문에서는 색채 히스토그램을 구성하기 위해서는 HSI 요소 중 색상요소(H)를 사용한다. Hue 값과 Saturation 값만을 사용함으로써 명암(I)변화에 둔감할 수 있다는 장점이 있다. 또한 이차원 평면상의 히스토그램을 사용할 수 있으므로 메모리와 계산상 효율적이다.

### 3. 형태정보 추출

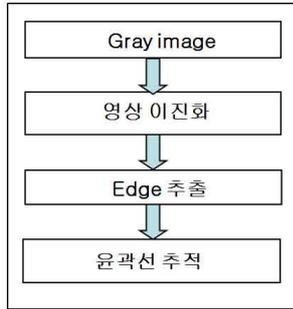
#### 3.1 윤곽선 추출

윤곽선은 인간이 시각적으로 가장 쉽게 판별할 수 있는 영상 특징으로 색인, 검색, 유사성 비교 등에 사용되는 주요 특징이다. 본 논문에서는 복잡한 배경에서 얼굴 영역의 윤곽선을 안정적으로 추출하기 위해 스네이크라 불리는 능동적 윤곽선 모델을 사용하며 에너지 함수는 다음 수식과 같다.

$$E_{snake}^* = \sum_{i=1}^N (\alpha E_{cont}(v_i) + \beta E_{curv}(v_i) + \gamma E_{image}(v_i)) ds \quad (4)$$

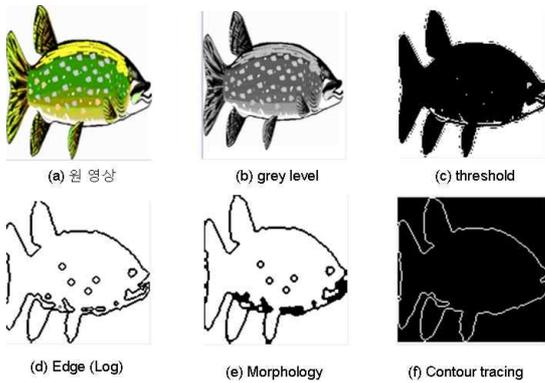
$$v_i = (x_i, y_i)$$

식 (4)의 첫 번째와 두 번째 항은 스네이크의 연속성과 완만성을 제어하는 함수이다. 즉, 스네이크를 이루는 정점들 간의 간격이 작고 완만한 곡선을 이룰수록 작은 값을 가지게 된다. 이러한 특징을 이용하여 객체에 대한 윤곽선을 추출할 때 영상 에너지 함수로써 화소의 밝기 값을 이용한다. 아래 (그림 4)는 형태정보의 전체적인 추출과정을 나타낸다.



(그림 4) 형태정보 추출과정

(그림 5)는 형태 특징을 추출하기 위한 전처리 과정과 윤곽선의 추출결과를 보여준다.



(그림 5) 전처리 및 형태정보 추출 결과

### 3.2 ICSS를 이용한 형태 특징 추출

스네이크를 이용하여 추출된 윤곽선과 특정 모양간의 유사성을 비교할 수 있는 거리 측정법이 필요한데 이를 위한 대표적인 방법으로 체인코드, 푸리에 변환, 2D 모멘트, CSS(Curvature Scale Space) 등의 방법이 있다. 체인코드는 가장 기본적인 방법으로 주어진 방향이 갖는 숫자의 시퀀스로 객체의 윤곽선을 표현하는 것이다. 체인코드는 잡음에 민감한 단점이 있다[4][5][6]. 유도코드는 일반 체인 코드에서 시퀀스 요소간의 방향 숫자 값 차이로 시퀀스를 형성하며 이는 회전에 불변한다. 그러나 90도 단위의 회전에만 가능하며 임의의 각도에 대해서는 불가능하다. 윤곽선 기반 방법으로 대표적으로 많이 사용되는 푸리에 기술자(Fourier Descriptor)는 푸리에 변환을 통해 생

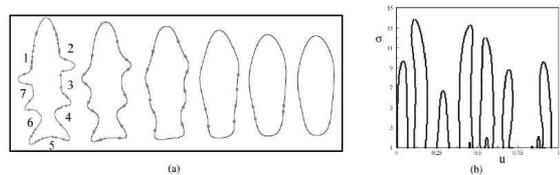
성된 복소수 형태의 상관계수들을 모양 기술자로 이용한다. 2D 모멘트는 2차원 랜덤 변수의 확률 밀도로써 정규화 된 그레이 레벨의 영상 함수를 나타낸다. 본 논문에서의 ICSS모델은 가우시안 분포 함수를 이용하여 윤곽선을 평활화 하면서 곡선 상에 존재하는 영교차점(zero-crossing)을 계산하며, 이들 영교차점의 변화를 통해 윤곽선의 시각적 특징을 대표하는 CSS영상을 추출하는 기존 CSS방법의 문제점을 보완한 모델이다.

#### 3.2.1 CSS(Curvature Scale Space)

CSS는 형태를 기반으로 영상을 검색하는 여러 가지 방법 중의 하나이다. 영상의 객체에서 추출된 윤곽선 정보를 사용해서 다른 영상과의 유사도를 비교한다. 다른 여러 가지 방법에 비해 크기, 회전, 이동 등의 변화에 강건하다는 장점이 있다. 영상에서 추출된 윤곽선은 보통 몇 백에서 몇 천개의 점으로 이루어진다. 그 점을 대략 200개 정도로 정규화 시킨다. 정규화 된 각 점들은 x, y의 좌표로 표현되며 아래의 식(5)의 Curvature function 에 모든 좌표 값을 대입하여 영교차점이 없어질 때까지 시그마 값을 0.1 씩 증가시켜 가면서 윤곽선을 평활화 시킨다.

$$\kappa(u, \sigma) = \frac{\dot{X}(u, \sigma)\ddot{Y}(u, \sigma) - \ddot{X}(u, \sigma)\dot{Y}(u, \sigma)}{(\dot{X}^2(u, \sigma) + \dot{Y}^2(u, \sigma))^{3/2}} \quad (5)$$

(Curvature Function)



(그림 6) 객체의 평활화 과정

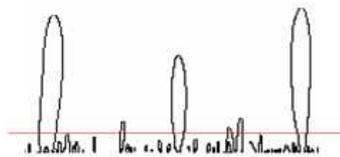
(그림 6)의 (a)는 시그마 값이 증가함에 따라 윤곽선의 영교차점이 없어지면서 점점 평활화 되는 과정을 보여주며 (b)의 그래프는 윤곽선이 평활화 되면서 나타난 CSS영상으로 y축은 시그마 값이고, x축은 arc length로 윤곽선을 200개의 점으로 정규화 시켜 펼쳐놓은 것이다. 그래프는 시

그마의 값이 상승함에 따라 영교차점이 이동하는 궤적을 나타낸 것으로 극대점은 한 커브의 영교차점이 서로 만나서 없어지는 점을 의미한다. 객체가 회전, 이동되거나 크기가 다르더라도 CSS영상은 시작점의 위치만 바뀔 뿐 전체적인 모양은 거의 변함이 없으므로 어파인 변형에 강건하다고 할 수 있다. (b)의 그래프에서 나타난 극대점(Maxima)의 좌표 값을 가지고 다른 영상과의 유사도를 측정한다.

3.2.2 ICSS(Improved Curvature Scale Space)

기존 CSS방법에는 몇 가지 보완할 부분이 있다. Curvature Function을 사용하여 얻어진 CSS영상을 보면 시그마 값이 낮은 부분에서 작은 극대점을 많이 볼 수 있다. 이 부분은 윤곽선에서 나타난 부분적인 왜곡이나 잡음인데, CSS에선 이 부분을 제거하기 위해서 가장 큰 극대점의 1/6미만은 잡음으로 간주해서 사용하지 않는다고 정의했다. 하지만 1/6미만을 잡음으로 제거하면 잡음이나 지역적인 왜곡이 완전히 제거되지 않아 정확도가 떨어지고 계산시간이 오래 걸리는 문제점이 발생한다. 또 다른 문제점은 만약 객체가 원(circle)이라면 영교차점이 발생하지 않으므로 CSS영상에서 극대점이 나타나지 않아야 한다. 그렇지만 원형인 경우에도 지역적인 왜곡이나 잡음 등으로 인하여 극대점이 나타나게 된다. 그러한 경우 위에서 말한 잡음 제거 방법으로는 잡음이 제거되지 않는다. 이러한 경우 객체가 원형임을 사전에 판별하게 되면 굳이 CSS영상을 만들지 않아도 된다. 이 두 가지 문제점을 보완한 방법이 ICSS 방법이다.

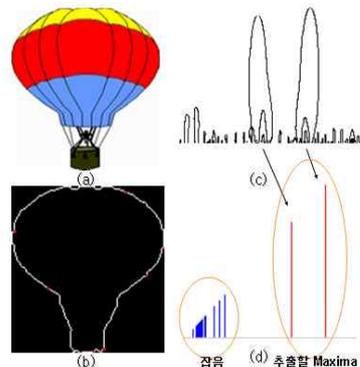
3.2.3 클러스터링을 이용한 잡음 제거 방법



(그림 7) CSS 영상에서의 극대점

(그림 7)을 보면 다른 극대점들보다 확연하게 큰 3개의 극대점을 볼 수 있다. 가로 줄 아래 부분

극대점들은 CSS방법에서 잡음으로 제거한 점들이다. 하지만 그 줄 위로도 작은 3개의 극대점을 볼 수 있는데 이는 윤곽선 추출과정에서 엘리머싱 현상으로 인하여 나타난 극대점들이다. 따라서 제거되지 않을 경우 유사도 비교의 정확성을 떨어뜨리고 계산시간을 증가시키게 된다. 이 문제를 보완하기 위해 본 논문에서는 클러스터링을 이용하여 추출해야 할 극대점과 제거해야 할 극대점을 구분하는 방법을 제안한다. 먼저 가장 큰 극대점과 가장 작은 극대점 값을 기준으로 그 사이 값들이 어느 극대점과 가까운지 클러스터링 한다. 그렇게 해서 두 그룹으로 나눈 후, 각 그룹의 평균을 계산한다. 이 계산된 두 평균값을 기준으로 다시 기존 극대점을 클러스터링 한다. 이 과정을 두 그룹의 평균의 변화가 없을 때까지 반복한다. 최종적으로 두 그룹으로 나누어지게 되는데 이때 왼쪽부분은 잡음, 오른쪽 부분은 극대점으로 분류할 수 있다. 이렇게 했을 때 최대 극대점을 1/6 부분 언저리에 걸려있던 극대점들을 확실하게 잡음으로 분류할 수 있게 된다.



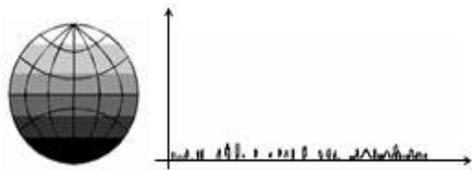
(그림 8) 잡음 제거 영상

(그림 8)의 (a)는 원본 영상, (b)는 윤곽선 정보 영상, (c)는 CSS영상, (d)는 클러스터링 된 영상으로 클러스터링 후에 잡음과 극대점으로 나뉘어진 것을 볼 수 있다.

3.2.4 원형객체 판별법

객체가 원형인 경우 영교차점이 없으므로 평활화 시킨 경우 CSS영상이 나오지 않아야 한다. 하지만 아래의 (그림 9)를 보면 객체가 원형이지

만 CSS영상에서 많은 극대점이 나오는 것을 볼 수 있다. 이 같은 원인으로서는 전처리 과정에서 생긴 잡음이나 엘리어싱 현상이라고 볼 수 있다. 따라서 Curvature Function을 사용하기 전에 원형임을 판별해 낼 수 있다면 검색의 정확도 및 수행 속도 면에서도 향상될 것이고 계산시간도 줄어들 것이다. 원형임을 판별해 내는 방법으로는 먼저 완전하게 원인 영상을 여러 개 준비하고 아래 식(6)을 이용하여 원형의 정도(Circularity)를 계산한다. border length는 객체둘레의 길이이며, area는 객체의 넓이를 나타낸다. 이렇게 계산된 원형의 정도를 분석해서 어느 값 이하를 원형으로 볼 것인지 임계값을 정한다. 두 번째는, 원형을 평활화 시켜서 CSS영상을 구한 다음, 최대 극대점의 시그마 값을 계산한다.



(그림 9) 원형 객체와 CSS영상

$$Circularity = \frac{(border\ length)^2}{area} \tag{6}$$

객체가 원이라면 위 두 방법을 통하여 결정된 임계값을 넘기지 않을 것이다. 본 논문에서 실험을 통하여 얻은 임계값은 원형의 정도가 12.9, 최대극대점 시그마 값은 2.4 이다. 이 두 임계값을 이용하여 입력되는 영상의 객체가 원형인지의 여부를 판별해 낼 수 있다.

#### 4. DTW를 이용한 유사도 측정

앞장에서 추출된 특징을 바탕으로 질의 영상과 DB영상과 비교를 위해서, DB영상에 대한 학습 과정을 거치게 된다. 학습 과정을 바탕으로 DB의 특징을 파악하며 비교, 분석이 가능하게 되어 진다. 이 과정에서는 DTW(Dynamic Time Warpin g)알고리즘[8]이 사용 되었다. DTW는 기준이 되

는 특징의 패턴과 입력된 특징의 패턴간의 유사도를 동적 프로그래밍(Dynamic programming)을 이용해 구하는 방법이다.

이전 단계까지의 여러 장의 학습 결과를 DTW에 넣게 되는데 DTW는 서로 다른 길이의 특징 벡터의 비교 연산이 용의하기 때문에 학습단계에서 서로 다른 모양 특징에 대한 비교 연산이 가능하게 된다. 두 영상에 투영된 데이터 특징을 각각  $Q=q_1, q_2, \dots, q_m$ ,  $C=c_1, c_2, \dots, c_n$  라고 했을 때, 이들 사이의 DTW 매칭 경로의 값을 식(7)과 같이 정의 할 수 있다.

$$d(uk) = d(qk, ck) = \left| [Q(qk) - \mu_Q] - [C(ck) - \mu_C] \right| \tag{7}$$

이렇게 얻어진 매칭경로를 최소화 하는 경로 값을 식(8)과 같이 표현한다.

$$DTW(Q, C) = \min \frac{1}{K} \sqrt{\sum_{k=1}^K W_k} \tag{8}$$

$$W_k = \{w_1, w_2, \dots, w_k\}$$

이때의 누적거리는 식(9)로 구하게 된다.

$$\gamma(i, j) = d(q_i, c_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\} \tag{9}$$

DTW는 두 패턴간의 누적거리를 최적화하는 (m, n)평면의 최적경로  $m=w(n)$ 를 찾는 방법이다. 이 방법을 이용하여 학습 영상(히스토그램) 중 대표 패턴을 하나 정하여 참조 영상으로 지정을 한 후 이것과 학습 영상간의 패턴을 학습하고, 참조 영상과 학습 영상간의 최적경로( $w_n$ )를 계산하여, 특징으로서 사용하게 된다. DTW를 통하여 추출된 경로를 이용하여 학습된 특징 분석 단계의 파라미터로 활용하게 되는데, 식(10)과 같이 표현하게 된다.

$$V_k = [w_k(1) - G_1, w_k(2) - G_2, \dots, w_k(L_1) - G_n]^T \tag{10}$$

여기서 각각 G, k, t는 식(11)과 같다.

$$G_k = \frac{p}{n} t \quad (11)$$

$$k = 1, \dots, N-1$$

$$t = 1, \dots, n$$

마지막으로 학습된 특징의 분석을 위하여 상관 관계에 대한 정보를 분석에 사용하기 위해 공분산 계산을 하게 되는데, 이때 학습 파라미터들의 평균은 식(12)으로 표현가능하며,

$$\mu = \frac{1}{N-1} \sum_{k=1}^{N-1} V_k \quad (12)$$

파라미터와 평균의 차를 이용한 공분산은 식(13)과 같다.

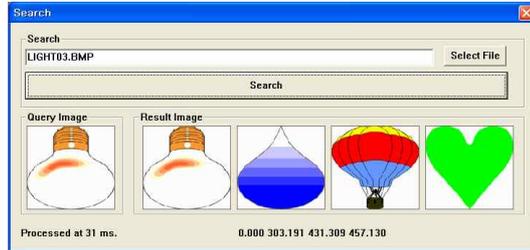
$$\Sigma = \frac{1}{N-1} \sum_{k=1}^{N-1} (V_k - \mu)(V_k - \mu)^T \quad (13)$$

위와 같은 계산을 입력되는 학습 영상 히스토그램에 대하여 위 단계를 반복하여 수행하여 학습 영상의 특징을 학습하게 된다. 본 논문에서는 이 DTW학습 단계와 최적 경로 비교 단계를 형태 정보와 색상 정보에 각각 적용하여 이미지 검색의 척도로서 사용하게 된다.

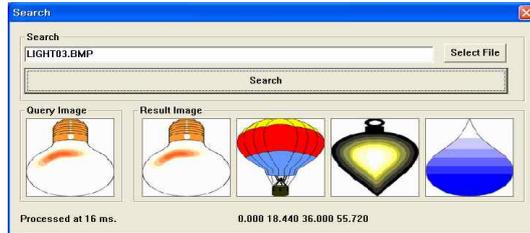
### 5. 실험결과

본 논문에서는 제안된 방법을 이용하여 영상 객체를 검색하고, 영상을 고의적으로 변형하였을 경우 잘 검출되는지 실험하였다.

실험을 위하여 구성된 컴퓨터는 Intel Pentium 듀얼코어 2.2Ghz CPU 와, 2GByte 의 메모리를 탑재한 Desktop PC를 사용하였고, OS는 Microsoft의 Windows XP professional을 사용하였으며, Visual C++ 6.0 사용하여 영상 검색 시스템을 제작하였다. 그리고 Microsoft의 Access를 이용하여 DB를 제작하였다. 또한 특징추출 편의를 위해 128 \* 128 의 동일한 크기의 정규화 된 약 100여 개의 BMP 형태(format)의 영상으로 수정하여 특징추출을 통한 검색과 분류를 시행하였다. 실험은 같은 질의 영상으로 유클리디언 디스턴스(UD)와 DTW로 각각 매칭 했을 때의 비교 검색 결과를 보이고, 검색의 처리 과정도 간단하게 소개한다.

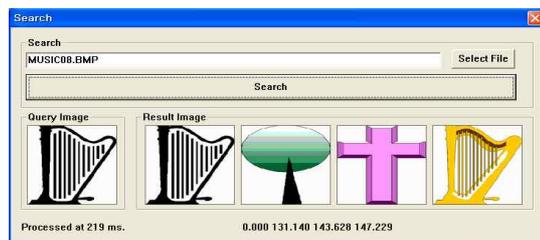


(그림 10) UD방법을 사용한 검색 결과

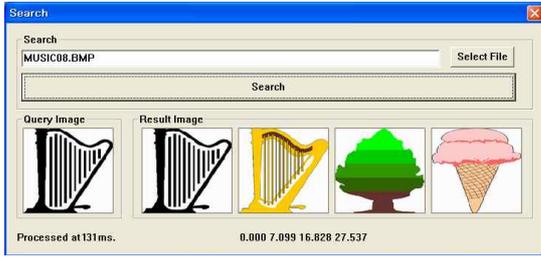


(그림 11) DTW방법을 사용한 검색 결과

(그림 10)은 UD방법을 이용했을 때의 검색 결과를 보인 것이고, (그림 11)은 DTW방법을 이용했을 때의 검색 결과를 보인 것이다. 결과 영상은 검색 순위가 높은 순서대로, 왼쪽 영상이 가장 유사하고 오른쪽으로 갈수록 유사도가 낮아지는 영상을 보여준다. 검색창 아래의 수치는 검색시간, 결과 영상 4개의 유사도 수치를 나타낸다. DTW방법이 UD방법을 사용했을 때보다 검색시간 수치가 낮아지는 것을 볼 수 있다. 이는 DTW방법이 UD방법보다는 계산 시간이 덜 걸려서 검색시간을 줄임을 의미한다. 다음은 이전 실험 영상보다 다소 복잡한 영상의 실험 결과이다.



(그림 12) UD방법을 사용한 검색 결과

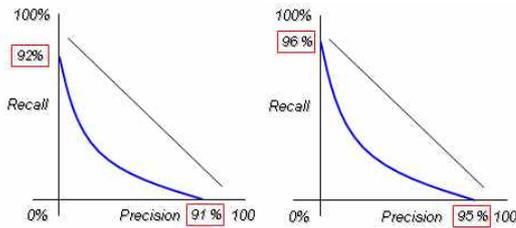


(그림 13) DTW방법을 사용한 검색결과

UD방법을 사용하였을 때 (그림 12)와 같이 질의 영상과 같은 모양의 하프가 결과 영상의 맨 뒤에 검색됨을 볼 수 있다.

(그림 13)은 DTW방법을 사용하여 하프를 검색한 결과 인데, (그림 12)의 UD방법을 사용하여 검색한 결과와 비교하여 볼 때 좋은 결과와 빠른 수행속도를 얻었음을 알 수 있다.

다음은 설문지를 통하여 사람이 검색한 결과와 시스템에서 검색한 결과가 얼마나 유사한지를 측정된 Precision / Recall의 결과를 보인다.



(a) UD방법의 Precision / Recall      (b) DTW방법의 Precision / Recall  
(그림 14) 각 방법의 Precision / Recall

(그림 14)의 (a)와 (b)를 비교하여 보면 DTW 매칭방법이 UD 매칭방법보다 Precision / Recall이 다소 높게 나온 것을 보여준다. 실험 결과를 종합하여 보면 객체가 간단한 모양이었을 때는 UD방법이나 DTW방법 모두 거의 비슷하게 검색이 된 반면, 객체의 모양이 복잡해지고 맥시마의 개수가 많아질수록 UD방법보다는 DTW방법을 사용하여 객체를 검색한 경우가 보다 좋은 결과를 나타낸 것을 볼 수 있었다.

## 6. 결론 및 향후 연구방향

본 논문에서는 ICSS를 사용해서 비교적 좋은 결과의 모양 정보를 추출할 수 있었으며, 색상추출에서도 HSI색상모델을 이용하여 RGB방법을 이용한 것보다 좋은 색상 정보를 추출할 수 있었다. 유사도 계산 방법에서는 유클리디언 디스턴스를 사용하였을 때보다 DTW 방법을 이용한 매칭이 Precision/Recall의 결과에서 볼 수 있듯이 검색의 정확도와 수행시간 면에서 더욱 효과적이라는 사실을 실험으로 입증하였다. 향후 연구해야 할 과제로는 같은 크기의 정규화 된 영상 비교뿐만 아니라 보다 다양한 형태와 종류의 이미지 매칭을 시도하고 그리고 데이터베이스에서의 영상의 수가 방대해질수록 시간이 걸리는 현상이 발생할 것이다. 따라서 서버기반 영상 검색 시스템을 구축하기 위해서는 데이터베이스의 저장방법을 개선하여 더욱 빠른 검색 속도와 정확도를 향상시키는 연구도 필요하다. 또한 색상, 형태 특징 말고도 질감 특징 등 여러 가지 특징을 더 적용하면 좀 더 효과적으로 검색할 수 있을 것으로 기대된다.

## 참고 문헌

- [1] S. Abbasi, Curvature scale space in shape similarity retrieval, Ph.D. thesis, Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, GU2 5XH, England, 1998
- [2] Donna J. Williams and Mubarak Shah, "A Fast Algorithm for Active Contours and Curvature Estimation", CVGIP: Image Understanding, vol.55, No.1, pp.14-26, 1992
- [3] K. Arbter, W.E. Snyder, H. Burkhardt, G. Hirzinger, Applications of affine-invariant Fourier descriptors to recognition of 3-D object, IEEE trans. Pattern Anal. Mach. Intell. 12, pp.640-646, 1990
- [4] A. Zhao, J. Chen, Affine curve moment invariants for shape recognition, Pattern Recognition 30, pp.895-901, 1997
- [5] Hu MK, Visual pattern recognition by moment invariants. IRE Trans Inf Theory IT-8: 179-187, 1962
- [6] Wallace TP, Wintz P, An efficient three-dimensional aircraft recognition algorithm using normalised Fourier Descriptors. Computer Graphics Image Process 13: 99-126, 1980

- [7] 김성만, 김재원, 최효성, 이양원, “색상정보와 모양정보를 이용한 상표 검색 시스템의 설계 및 구현”, 한국멀티미디어학회 추계학술발표 논문집 784-788, 1999
- [8] Mizuhara, Yuko ; Hayashi, Akira ; Suematsu, Nobuo, “Embedding of time series data by using dynamic time warping distances”, Systems and computers in Japan, v.37 no.3, pp.1-9, 2006
- [9] 하정요, 최미영, 최형일, “색상과 형태를 이용한 내용기반 영상 검색”, 한국 컴퓨터정보학회 논문집 제13권 제1호, pp.117-124, 2008

**하 정 요**



2004년 :가톨릭대학교  
컴퓨터공학과 (공학사)  
2006년 :송실대학교 대학원  
미디어학과 (공학석사)  
2006년~현재 :송실대학교 대학원  
미디어학과 박사과정

관심분야 : 컴퓨터비전, 영상정보검색, 이동물체추적 등



**최 형 일**

1979년 :연세대학교  
전자공학과 (공학사)  
1982년 :미시간대학교  
전자공학과 (공학석사)  
1987년 :미시간대학교  
전자공학과 (공학박사)  
1987년~현 재 : 송실대학교 IT대학 미디어학과교수  
2003년~2004년 : 한국정보과학회 이사  
2005년~현 재 : 한국디지털미디어학회 이사  
2008년~현 재 : 송실대학교 문화콘텐츠기술(CT)  
연구소 소장

관심분야 : 컴퓨터비전, 퍼지 및 신경망이론, 패턴인식, 인터페이스에이전트, 지식기반시스템, 비디오검색, 등

**이 나 영**



2000년 :청운대학교  
컴퓨터학과 (공학사)  
2002년 :송실대학교 대학원  
컴퓨터학과 (공학석사)  
2007년 :송실대학교 대학원  
컴퓨터학과 (공학박사)

2008년 : 한국원자력연구원  
동위원소이용기술개발부 (Post-Doc)

현 재 : Carnegie Mellon University,  
Robotics Institute, (방문연구원)

관심분야 : 컴퓨터비전, 영상복원, 패턴인식, 얼굴인식, 의료/영상처리 등

**김 계 영**



1990년 :송실대학교  
전자계산학과 (공학사)  
1992년 :송실대학교  
컴퓨터학과 (공학석사)  
1996년 :송실대학교  
컴퓨터학과 (공학박사)

2001년~현 재 : 송실대학교 컴퓨터학과교수

관심분야 : 컴퓨터비전, 형태인식, 생체인식, 증강현실, 영상처리 등