모양 기반 영상검색을 위한 골격 나무 구조

박 종 승†

요 약

본 논문에서는 내용기반 영상검색 시스템에서 검출 정확성을 개선하기 위한 모양 기반 객체 표현 기법인 골격 나무 구조와 골격 나무 구조에 기반한 검색 기법을 제시한다. 모양의 표현을 위한 새로운 골격 나무 표현 기법은 객체의 모양을 계층적인 나무 구조로 표현한다. 나무 구조에서의 상위 레벨의 노드들은 객체 모양의 대략적인 몸체를 표현하고 하위 레벨의 노드들은 객체 모양의 세밀한 부분을 표현한다. 대부분의 영상 잡음 영향은 하위 레벨의 노드들에 국한되므로 외곽선 잡음의 영향은 하위 레벨에 대한 가중치를 조절하여 감소시킬 수 있다. 두 골격 나무의 유사성은 두 나무의 골격 노드들을 매칭하여 계산된다. 두 골격 노드들의 집합이 표현하는 영역의 유사성 계산을 위해서 Zernike 모멘트 불변량과 Fourier 기술자 불변량의 조합이 사용된다. 검출의 정확도를 다른 유사성 측도와 비교하고 그 결과를 제시하였다. 실험 결과 제안된 검색 기법의 검출 정확도가 증가하였음을 알 수 있었다.

키워드 : 모양 매칭, 모양 기반 영상검색, 골격 나무, Fourier 기술자, 모멘트 불변량

Skeleton Tree for Shape-Based Image Retrieval

Jong-Seung Park†

ABSTRACT

This paper proposes a skeleton-based hierarchical shape description scheme, called a skeleton tree, for accurate shape-based image retrieval. A skeleton tree represents an object shape as a hierarchical tree where high-level nodes describe parts of coarse trunk regions and low-level nodes describe fine details of boundary regions. Each node refines the shape of its parent node. Most of the noise disturbances are limited to bottom level nodes and the boundary noise is reduced by decreasing weights on the bottom levels. The similarity of two skeleton trees is computed by considering the best match of a skeleton tree to a sub-tree of another skeleton tree. The proposed method uses a hybrid similarity measure by employing both Fourier descriptors and moment invariants in computing the similarity of two skeleton trees. Several experimental results are presented demonstrating the validity of the skeleton tree scheme for the shape description and indexing.

Key Words : Shape matching, Shape-based image retrieval, Skeleton tree, Fourier descriptors, Moment invariants

1. 서 론

내용기반 영상검색 문제는 대형 영상 데이터베이스로부터 질의 영상(query image)과 가장 유사한 영상들의 목록을 얻는 문제이다. 모양에 기반한 영상검색에서의 가장 중요한 두 가지 문제는 객체의 특징 추출(feature extraction)과 객체의 매칭(object matching)이다. 특징 추출은 영상의 내용에 대해 자동으로 특성들을 묘사하는 작업이다. 특성들로 기술되는 영상 내용에 대해 유사성 측도를 적용하여 유사한 영상들을 검색한다.

영상검색 시스템 개발에 있어서 영상 이해 기술의 한계로 인해서 지능적인 영상검색이 어려움을 겪어왔다. 대부분의 기존의 영상검색 기법들은 칼라, 텍스처, 영역 등의 영상 특징 분석에 의존하고 있다. 이러한 특징들은 질의 객체의 모양 속성을 적절히 표현하지 못한다[1]. 따라서 질의 영상과의 관련성이 떨어지는 영상들이 자주 검색되는 단점을 지닌다[2]. 모양 표현을 위해서는 체인 코드(chain code), 시그니쳐(signature), 골격(skeleton) 등의 표현 방식에 사용된다. 골격 표현 방식은 골격의 계산이 안정적이지 못하고 모양의 구체적인 형태에 민감한 단점이 있다. 최근 모양 특징들을 효과적으로 다룰 수 있는 연구들이 진행되었다. 골격 추상화 기법[3]이 제안되었고 골격 유사도 측정 기법[4]이 제안되었다. 골격과 유사한 표현 기법으로 충격 그래프 (shock graph)[5,6], 골격 그래프(skeletal graph)[7] 등이 제안되었다. Mokhtarian [8] 등은 모양의 외곽선 표현을 위해 곡률

† 종신회원 : 인천대학교 컴퓨터공학과 조교수
논문접수 : 2005년 10월 31일, 심사완료 : 2007년 5월 25일

스케일 공간(curvature scale space) 표현을 제안하였다.

검출의 정확도는 어떤 특징을 선택하고 그들을 어떻게 표현하는가에 좌우된다. 모양의 유사성의 표현 방법으로 Sclaroff 등[9]은 모달 변형(modal deformation) 에너지의 양으로 모양의 유사성을 표현하는 방법을 제안하였다. 모양 기반 영상검색 시스템에서 불변량이 유용하게 사용된다. 모양 매칭에서 흔히 사용되는 불변량은 Fourier 기술자(Fourier descriptor)와 모멘트 불변량(moment invariant)이다[1,2,10]. 이론적으로 이들 불변량은 크기 변화나 회전에 불변인 특징이 있다. 그러나 실제 환경에서의 디지털 영상의 경우 불변 특정은 일반적으로 잘 유지되지 않는다.

본 논문에서는 새로운 모양 기반 영상검색 기법을 제안한다. 영상들은 이미 분할되어있거나 분할 가능하다고 가정한다. 본 논문에서는 모양의 표현을 위한 골격 나무를 제안하고 골격 나무의 유사도 측정 기법을 제안한다. 객체의 모양을 골격 나무로 표현하고 모양의 유사도 측정을 통해 모양 기반 검색을 실시한다. 골격 나무는 객체를 계층적으로 표현하는 모양 표현 기법으로써 상위 레벨의 노드들은 객체의 전체적이고 대략적인 모습을 표현하고 하위 레벨의 노드들은 상위 노드들의 모양을 구체화한다. 각 노드는 부모 노드의 모양을 세분화하므로 대부분의 잡음 간섭은 하위 노드에 치중되게 된다. 하위 노드의 가중치 조절을 통해서 잡음 감소가 이루어진다.

영상에 포함된 객체의 모양을 먼저 골격 나무 표현으로 변환하고 골격 나무 표현으로부터 유사도를 측정한다. 한 질의 영상이 주어질 때 제안하는 알고리즘의 처리 절차는 다음과 같다:

단계 1: 질의 영상의 분할 이진 영상을 골격화(skeletonize)한다.

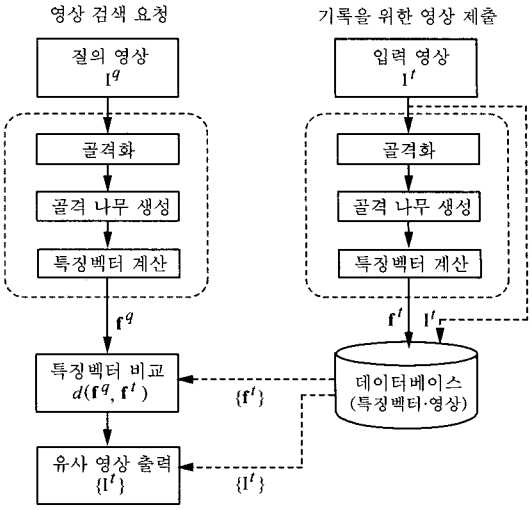
단계 2: 골격 영상(skeleton image)로부터 골격 나무(skeleton tree)를 만든다.

단계 3: 질의 영상의 골격 나무로부터 특징벡터를 계산하고 데이터베이스에 저장되어있는 골격 나무들에 대한 특징벡터의 유사도(similarity)를 계산한다.

단계 4: 유사도가 가장 높은 저장 영상들을 나열하여 출력한다.

영상 기록 과정(image archival process)에서는 골격 나무를 단계 1-2를 거친 후 골격 나무에 대해 특징벡터를 계산하고 골격 나무와 특징벡터를 데이터베이스에 저장한다. 영상검색 과정에서는 질의 영상에 대한 골격 나무에 대한 특징벡터를 계산하고 이를 데이터베이스의 골격 나무들의 특징벡터와 비교한다. 전체적인 시스템 흐름도는 (그림 1)과 같다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 골격 나무를 설명한다. 3장에서는 골격 나무에 대해 특징벡터를 계산하고 유사도를 측정하는 방법을 기술한다. 4장에서 실험 결과를 기술한다. 5장에서 결론을 맺는다.



(그림 1) 제안된 영상검색 시스템의 시스템 흐름도.

2. 모양 표현을 위한 골격 나무

모양 기반 영상검색에서 모양 표현을 위해 제안된 골격 나무(skeleton tree)를 기술한다. 영상의 각 객체 영역에 대해서 거리 변환(distance transform)으로 골격 영상을 얻는다. 골격 영상으로부터 골격 나무를 계산할 수 있다. 골격 나무는 객체를 계층적으로 표현하며 상위 레벨의 노드들은 객체의 대략적인 몸체의 형상을 표현하고 하위 레벨은 노드의 구체적인 세부 형상을 표현한다. 각 하위 노드들은 부모 노드의 모양을 구체화한다.

골격 표현은 모양 묘사의 자연스러운 방법이다. 특히 인체나 동물 등의 변형체의 표현에 적합하다. 골격 표현은 자연스러운 묘사뿐만 아니라 추후에 역 골격 변환(inverse skeleton transform)으로 골격 데이터로부터 원래의 모양을 그대로 재건할 수 있는 장점을 가진다. 골격 표현은 이러한 강력한 장점을 가지지만 특징 추출 시의 잡음의 영향에 민감하다는 단점이 있다. 예를 들어서 고르지 못하고 울퉁불퉁한 외곽선에 대해서 중심축(medial axis) 접근을 한다면 너무 많은 가지(branch)들과 모양 왜곡이 발생할 것이다. 본 연구에서는 이러한 문제점의 해결을 위해 잡음에 강건한 골격 표현 기법을 제시한다.

제안하는 골격 나무에서는 모양을 계층적 나무 구조로 표현하여 외곽선 잡음에 대한 민감성을 줄이도록 한다. 골격 나무 구조는 골격 프리미티브(skeletal primitive)로부터 생성한다. 나무의 최상위 레벨인 레벨 1의 노드들은 가장 전체적이고 대략적인 모양을 표현한다. 그 하위 레벨의 노드들은 상위 레벨 노드를 구체화한다. 따라서 외곽선 잡음에 대한 민감성은 하위 레벨의 노드들에 국한된다. 모양 구성에서 하위 레벨 노드들에 가중치를 두어서 잡음 민감성을 조절할 수 있다.

2.1 노드와 링크의 생성

각 객체 영역에 대해서 골격(skeleton)으로 바꾸고 이로부

터 골격 세그먼트를 추출한다. 객체 모양 *obj*의 골격 *skel(obj)*은 모양 내부에 포함되는 최대 디스크의 중심의 자취이다. 골격을 추출하는 방법으로는 다양한 기법들이 제안되었다. 거리 변환(distance transform) 방법[11], 형태학 연산자 (morphological operator) 방법[12], Voronoi 그래프 방법[13] 등이 대표적이다. 본 논문에서는 절차가 비교적 단순한 거리 변환 기법으로 골격을 계산한다. 거리 변환에서는 영역 내부의 각 점에 대해 해당 점에서의 최단거리의 경계 화소까지의 거리를 할당한다. 구현에서 골격 계산을 위한 빠른 두 패스 알고리즘[12]을 사용하였다. 거리 변환 맵에서의 거리가 국부 최대인 화소들이 골격 화소들이다. 과다한 골격 가지들의 생성을 방지하기 위해서 일정 문턱값보다 작은 거리의 가지들을 제거한다. 문턱값의 선택은 골격 나무의 복잡도에 영향을 미친다. 높은 문턱값은 골격 가지의 수를 크게 줄여 골격 나무를 단순화시킨다.

골격(skeleton)으로부터 세그먼트들을 추출하고 이를 사용하여 계층적나무 구조인 골격 나무(skeleton tree)로 변환한다. 세그먼트의 끝점들과 접합점들을 노드로 정의하고 세그먼트 상의 화소 리스트를 링크로 간주한다. 끝점과 접합점에서 모든 세그먼트들을 분리하도록 하여 각 세그먼트의 끝이 두 점이 되도록 한다. 이렇게 구성된 세그먼트를 *골격 세그먼트(skeletal segment)*라고 한다. 골격 세그먼트들로부터 골격 나무를 만든다. 골격 나무는 한 노드 집합과 한 링크 집합으로 표현된다. 각 골격 세그먼트의 두 끝점은 두 노드를 정의하고 두 노드 사이의 하나의 링크를 정의한다. 해당 노드가 이미 노드 집합에 포함되어 있다면 동일 노드로 간주하고 기존의 포함된 노드를 사용한다. 각 링크는 객체의 영역의 부분 영역인 *영향 영역(influence zone)*을 정의한다. 이 부분 영역은 링크에 해당하는 세그먼트 상의 화소들로부터 생성된 영역에 해당한다. 세그먼트 상의 화소들은 골격 화소들이다. 각 화소는 거리 변환 시의 외곽선까지의 거리값을 유지하고 있다. 각 화소는 이 거리값을 반지름으로 하는 디스크 영역을 정의한다. 세그먼트 상의 모든 화소들이 정의하는 디스크 영역들을 합한 영역이 링크가 정의하는 영향 영역이다.

한 객체 *obj* 에 대해서 한 노드 집합 $V(obj)$ 과 한 링크 집합 $E(obj)$ 이 얻어지면 골격나무를 구성할 수 있다. 하나의 링크에만 연결된 노드를 종단 노드(leaf node)라고 하고 종단 노드에 연결된 링크를 *스킨 링크(skin link)*라고 하자. 스킨 링크가 아닌 링크를 *뼈대 링크(bone link)*라고 하자. 링크 집합 $E(obj)$는 스킨 링크 집합 $E_S(obj)$와 뼈대 링크 집합 $E_B(obj)$의 합으로 생각할 수 있다.

2.2 골격 나무의 생성

노드 집합 $V(obj)$와 링크 집합 $E(obj)$로부터 골격 나무를 구성한다. 먼저 *최대 영향 경로(maximum influence path)*를 결정한다. 최대 영향 경로 $(v_1,\cdots,v_n)$는 경로상의 모든 노드들이 비종단 노드(non-leaf node)이며 v_1에서 v_n으로 모든 경로들 중에서 경로상의 모든 링크의 영향 영역의 합이 최대인 경로이다. 경로의 두 끝 노드 v_1과 v_n의 결정을 위해서 모든 비종단 노드들에 대해서 두 노드의 최대 영향 경로를 구하고 두 노드의 최대 영향 경로가 최대인 경로의 두 끝 노드를 v_1과 v_n으로 지정하고 이때의 경로를 골격 나무의 최대 영향 경로로 지정한다. 골격 나무의 최대 영향 경로 상의 링크들을 *스파인 링크(spine link)*라고 하자.

최대 영향 경로가 결정되면 골격 나무의 최상위 레벨인 레벨 1의 노드와 링크를 구성한다. 최대 영향 경로의 노드들인 $(v_1,\cdots,v_n)$ 노드들을 왼쪽에서 오른쪽의 순서대로 레벨 1 노드로 추가한다. 레벨 1이 구성되면 레벨 1의 각 노드에 연결된 노드들을 레벨 1의 노드의 자식 노드로 추가한다. 이 추가 과정을 $V(obj)$의 모든 노드들이 추가될 때까지 반복한다.

노드 집합 $V(obj)$와 링크 집합 $E(obj)$로부터 골격 나무 $T(V,E)$를 생성하기 위한 전체적인 골격 나무의 생성 절차는 다음과 같다:

단계 1: 골격 세그먼트의 추출: 한 객체 *obj*에 대해서 영역 정보를 얻고 영역에 대해 거리 변환 기법으로 골격(skeleton)을 만든다. 골격으로부터 골격 세그먼트를 추출한다.

단계 2: 노드 집합과 링크 집합의 생성: 골격 세그먼트로부터 노드 집합 $V(obj)$와 링크 집합 $E(obj)$를 계산한다.

단계 3: 골격 나무의 레벨 1을 생성: 최대 영향 경로를 결정하고 이로부터 골격 나무의 레벨 1을 만든다.

단계 4: 골격 나무의 모든 레벨을 생성: 골격 나무에 포함되지 않은 모든 노드 v에 대하여, 노드 v가 골격 나무에 포함된 한 노드 v'에 연결되어 있다면 v를 v'의 자식 노드로 추가한다. 모든 노드가 골격 나무에 포함될 때까지 추가 과정을 반복한다.

여기에서 새로 추가되는 노드는 항상 이전에 추가된 노드의 자식 노드로 되므로 골격에 구멍이 있는 경우라고 하더라도 사이클이 없는 나무 구조가 생성된다.

2.3 골격 나무의 단순화

영역 외곽선의 이산적인 표현과 외곽선 추출시의 부정확성에 의해 노드들이 과도하게 생성될 수 있다. 표현의 단순화 및 빠른 계산을 위해서 객체의 표현에 영향이 없는 범위 내에서 노드 수를 줄이는 것이 바람직하다. 이를 위해서 골격 나무의 생성 이전에 불필요한 노드와 링크를 제거하는 절차를 둔다. 골격 표현에서는 외곽선에서 약간의 굴곡의 변화가 있더라도 한 노드가 여러 노드로 나누어지는 특성이 있다. 이러한 과대한 노드 생성을 방지하기 위해서 *Cut 연산*, *Delete 연산*, *Merge 연산*의 세 연산을 정의하고 이들 연산을 나무의 생성 단계에 적용한다. 연산자의 적용 규칙은 다음과 같다:

- **Cut 연산:** 한 스킨 링크 상의 화소들에 대해 거리 영상에서의 그라디언트 크기의 합이 매우 작으면 그 링크를 $E_S(obj)$에서부터 제거하고 그 링크에 연결된 종단 노드를 $V(obj)$로부터 제거한다. 그라디언트 크기의 합이 매우 작은 스킨 링크는 외곽선의 작은 변화로부터 발생된 것으로 그 링크의 대부분의 영향 영역은 이미 다른 링크의 영향 영역과 중복되어 있어서 링크 삭제가 전체 모양을 크게 변화시키지 않는다.
- **Delete 연산:** 두 스킨 링크의 영향 영역의 대부분이 겹치면 두 스킨 링크 중 영향 영역이 작은 스킨 링크와 그의 종단 노드를 제거한다. 많은 스킨 링크들이 밀접하게 나타나는 경우 이를 적절히 제거하여 스킨 링크의 수를 줄이는 역할을 한다.
- **Merge 연산:** 두 비종단 노드(non-leaf node)가 너무 가까이 인접해 있으면 두 노드를 하나의 노드로 합치고 두 비종단 노드에 연결된 모든 링크들을 합친 후의 노드까지 확장한다.

이러한 세 연산자는 노드의 수와 링크의 수를 크게 감소시켜서 매우 간단한 골격 나무가 생성되도록 한다.

3. 골격 나무의 모양 매칭

3.1 골격 나무의 유사도

두 객체 모양의 유사도는 해당하는 두 골격 나무의 유사도로 정의한다. 두 객체 obj와 obj'에 대해서 해당하는 두 골격 나무를 $T(V,E)$와 $T(V', E')$라 하자. 각 골격 나무로부터 모양 영역을 계산할 수 있다. 모양 영역은 골격 나무의 모든 링크들의 영향 영역이 합이다. 골격 나무의 유사도는 골격 나무가 표현하는 영역의 유사도를 계산하여 구해진다. 모양의 유사도를 측정하는 방법으로 대표적으로 모멘트 불변량(moment invariants)과 Fourier 기술자(Fourier descriptors)가 사용된다. 두 방법 모두 이차원 공간에서 이동, 크기변화, 그리고 회전에 대한 불변량을 제공한다. 두 가지 불변량의 계산 방법에는 큰 차이가 있다. 모멘트 불변량은 영역 기반 측정값이며 Fourier 기술자는 외곽선 기반 측정값이다. Mehtre 등[2]은 여러 가지 표현에 따른 검색 효율성을 비교하였다. 축소된 체인 코드, Fourier 기술자, 모멘트 불변량 등을 비교한 결과 Fourier 기술자와 모멘트 불변량의 동시 사용이 가장 우수한 평균 검출 효율성을 나타냄을 보였다. 이러한 결과는 인간의 인지 체계가 모양의 영역 특성과 외곽선 특성을 동시에 고려하여 모양 유사도를 측정하는 사실과 상통한다. 모멘트 불변량은 외관(appearance)의 분류에 효과적이고 Fourier 기술자는 객체의 자세(pose)의 구별에 효과적이다. 본 논문에서는 이러한 두 가지 다른 종류의 유사도 측정 기법을 도입하여 골격 나무의 유사도 계산에 적용하고자 한다. 모멘트 불변량과 Fourier 기술자의 각각에 대해서 모든 이차원 변환에 대해서 불변 특성을 유지하는 특징값들을 정의하고 이들의 조합으로 골격 나무의 유사도를 계산하는 방법을 사용한다.

3.2 모멘트 불변량

모멘트 불변량(moment invariants)는 일정 변환에 대해 불변인 모멘트에 대한 함수이다. 모멘트는 연속적인 영상 밝기 함수에 기반하고 있다. 이산적인 영상에 대해서는 적분을 합으로 계산하는 단순한 연산으로 근사 값을 계산할 수 있다. 이진 디지털 영상 행렬을 b라고 하고, 이차원 모양은 영역의 화소들의 집합 $S=\{(x,y)|b(x,y)=1\}$ 으로 표시할 수 있다. 모양 S의 (p,q)차 모멘트는 (식 1)과 같이 주어진다.

$$m_{pq}(S)=\sum_{(x,y)\in S}x^p y^q \tag{1}$$

이러한 모멘트에 대해서 낮은 차수의 모멘트들의 비선형 조합으로 모멘트 불변량을 유도할 수 있다. Hu는 연속 함수들에 대해서 모멘트의 불변 특성을 증명하였다[14].

모멘트 변환은 $x^p y^q$ 항을 Zernike 다항식으로 대치하여 일반화된 형식으로 확장할 수 있다. Zernike 다항식은 단위 디스크에 직교하는 복소수 함수들의 완전 집합이다[15]. 단위 원 외부에서 소멸하는 반복횟수 m의 n차원 Zernike 모멘트는 (식 2)와 같이 주어진다.

$$A_{n,m}=\frac{n+1}{\pi}\sum_{(x,y)\in S,x^2+y^2\le 1}V^*_{n,m}(x,y) \tag{2}$$

여기서 $V^*_{n,m}=V_{n,-m}$ 이고, n과 m은 정수로서 $n\ge 0$이고, $n-|m|$은 짝수이며, $|m|\le n$ 이다. 또한 $V_{n,m}$ 은 (식 3)과 같이 구할 수 있다.

$$V_{n,m}(x,y)=R_{n,m}(x,y)\exp(jm\tan^{-1}(y/x)),$$
$$R_{n,m}(x,y)=\sum_{s=0}^{(n-|m|)/2}\frac{(-1)^s(x^2+y^2)^{n/2-s}(n-s)!}{s!((n+|m|)/2-s)!((n-|m|)/2-s)!} \tag{3}$$

영상 분석에서 Zernike 모멘트의 사용은 Taegue[15]에 의해서 제안되었고, 그 이후로 Zernike 모멘트가 많은 영상 분석 작업에서 사용되기 시작하였다[1,16~18]. Zernike 모멘트의 크기(magnitude) $|A_{n,m}|$은 회전에 불변이다. 이동에 불변이 되도록 하기 위해서는 원점을 무게중심으로 이동하면 된다[19]. 크기변화에 불변이 되도록 하기 위해서는 원점에서 최대 거리가 1이 되도록 점들을 스케일링하면 된다[17].

정규화는 처음의 두 특징값 $|A_{0,0}|$과 $|A_{1,1}|$에 영향을 미친다. 따라서 2차 모멘트부터 5차 모멘트까지의 10개의 특징값을 선택을 하였다. 이들 10개의 특징값으로 10차원 특징벡터 f_Z 를 (식 4)와 같이 구성한다.

$$f_Z=[|A_{2,0}|,|A_{2,2}|,|A_{3,1}|,|A_{3,3}|,|A_{4,0}|,|A_{4,2}|,|A_{4,4}|,|A_{5,1}|,|A_{5,3}|,|A_{5,5}|] \tag{4}$$

3.3 Fourier 기술자

Fourier 기술자는 모양의 외곽선 점들로부터 구할 수 있는 이차원 불변 특성값이다. 한 모양의 외곽선이 N개의 화소를 가진다면 윤곽선은 (식 5)과 같이 두 파라메터화된 수식으로 표현할 수 있다.

$$x(n)=x_n,\ y(n)=y_n,\ n=0,\cdots,N-1 \tag{5}$$

이 식들은 복소수 평면에서 표현한다면 (식 6)으로 쓸 수 있다.

$$z(n) = x(n)+jy(n) \tag{6}$$

한 윤곽선에 대한 Fourier 기술자 $Z(k)$는 윤곽선 식 $z(n)$의 계수들의 이산적 Fourier 변환이고 (식 7)과 같다.

$$Z(k)=\frac{1}{N}\sum_{n=0}^{N-1} z(n)\exp(-j2\pi kn/N), k=-N/2,...,N/2-1 \tag{7}$$

$Z(k)$를 간단히 정규화하면 샘플링의 시작 위치, 회전, 크기변화, 이동에 불변인 Fourier 기술자를 얻을 수 있다. 이동 불변에 대해서는, $z(n)$를 z_0만큼 이동하면 첫 Fourier 기술자인 $Z(0)$에만 영향을 미치고 그 외의 Fourier 기술자에는 변화가 없다. 따라서, 이동에 대해 불변성을 가지도록 하기 위해서 첫 번째 Fourier 기술자 $Z(0)$을 0으로 하여 모양의 중심(centroid)를 원점으로 이동하도록 한다. 크기변환 불변에 대해서는, $z(n)$을 인자 s만큼 크기변환하면 Fourier 기술자도 동일한 인자만큼 크기변환한 효과가 있다. Fourier 기술자의 특성에서 외곽선을 반시계방향으로 따라간다면 $Z(1)$이 0이 아니게 되고, 시계방향으로 따라간다면 $Z(-1)$이 0이 아니게 된다. 본 논문에서는 항상 반시계방향으로 외곽선을 따라간다고 가정한다. 따라서 두 번째 Fourier 기술자 $Z(1)=r_1\exp(j\phi_1)$는 0이 아니다. 따라서, 크기변화에 불변한 특성을 얻기 위해서 모든 Fourier 기술자들을 $|Z(1)|$로 나눈다. 회전 불변에 대해서는, $z(n)$를 각도 θ_0만큼 회전하면 Fourier 기술자는 위상을 상수 θ_0만큼 시프트한 효과가 있다. 따라서, 회전 불변 특성은 크기변환 불변 특성을 적용하고 그 크기(magnitude)를 취함으로써 얻어진다. 시작위치 불변에 대해서는, $z(n)$의 시작 위치를 0에서 n_0 위치로 움직이는 것은 Fourier 기술자에서 $2\pi n_0k/N$ 만큼의 선형적 위상 시프트를 만든다. 따라서, 외곽선의 시작위치에 불변한 특성을 가지도록 하기 위해서는 각 Fourier 기술자의 위상에서 두 번째 Fourier 기술자의 위상을 가중치 k로 하여 빼면 된다. 즉 $Z(k)=Z(k)\exp(-j\phi_1k)$로 대치하면 시작 위치 불변 특성을 얻을 수 있다.

Fourier 기술자로부터 불변량 벡터를 정의하기 위해서 m개의 Fourier 기술자로부터 m차원 특징벡터 f_F를 (식 8)과 같이 구성한다.

$$f_F=[\ Z(-m/2),\cdots,Z(-1),Z(2),\cdots,Z(m/2+1)] \tag{8}$$

여기에서 m은 8이나 16과 같은 2의 지수승으로 하여 변환이 FFT로 빠르게 수행되도록 한다. 특징값의 수는 m과 같으며 실험적으로 m이 8 이상이면 분류에 충분히 이용될 수 있다. 본 논문에서는 m을 8로 두어 8차원 특징벡터인

$$f_F=[\ Z(-4),\ Z(-3),\ Z(-2),\ Z(-1),\ Z(2),\ Z(3),\ Z(4),\ Z(5)]$$

를 계산한다.

3.4 골격 나무의 유사도 계산

모멘트 불변량과 Fourier 기술자로 두 특징벡터 f_Z와 f_F를 정의하였다. 골격 나무의 유사도 측정을 위해서 두 특징벡터를 사용한다. 먼저 각 골격 나무로부터 두 특징벡터를 계산한다. 골격 나무로부터 모든 링크들의 영향 영역을 더하면 모양의 영역을 얻을 수 있고 영역으로부터 영역 외곽선도 쉽게 구할 수 있다. 영역과 외곽선으로부터 모멘트 불변량 f_Z와 Fourier 기술자 f_H를 계산한다.

두 골격 나무의 유사도는 각 골격 나무의 특징벡터들의 유사도로 측정된다. 질의 물체obj^q의 특징벡터들을 f_Z와 f_F라고 하고 등록 물체 obj^t의 특징벡터들을 f_Z'와 f_F'라고 하자. f_Z와 f_Z'의 유클리드 거리 d_Z와 f_F와 f_F'의 유클리드 거리 d_F를 계산한다. 유사도 d_C는 (식 9)와 같이 두 거리의 평균으로 계산된다.

$$d_C(obj^q,obj^t)=\frac{1}{2}d_Z(obj^q,obj^t)+\frac{1}{2}d_F(obj^q,obj^t) \tag{9}$$

여기서 $d_Z(obj^q,obj^t)=d(f_Z(obj^q),f_Z'(obj^t))$이고 $d_F(obj^q,obj^t)=d(f_F(obj^q),f_F'(obj^t))$이다. 여기서 d는 두 벡터의 거리를 나타낸다. (식 9)는 영역 특징의 차이인 d_Z와 외곽선 특징의 차이인 d_F에 동일한 가중치를 준 경우이다.

외곽선이 긴 경우에 외곽선 불변량에 높은 가중치를 두고 영역의 면적이 넓은 경우에 영역 불변량에 높은 가중치를 두면 보다 정확한 유사성 측정이 가능하다. 두 특징에 가중치를 달리 하여 (식 10)과 같이 가중치가 적용된 유사도를 정의할 수 있다.

$$d_W(obj^q,obj^t)=w_Zd_Z(obj^q,obj^t)+w_Fd_F(obj^q,obj^t) \tag{10}$$

식 (9)에서 w_Z와 w_F는 0에서 1 사이의 값으로 각각 영역 모습의 유사성과 외곽선 모습의 유사성에 대한 가중치이다. 두 가중치 w_Z와 w_F는 (식 11)과 같이 골격 나무가 표현하는 모양과 원래 모양의 넓이의 비율과 외곽선 길이의 비율로부터 구해진다.

$$w_Z=\alpha(T^q)/\alpha(obj^q),\ w_F=\varphi(T^q)/\varphi(obj^q) \tag{11}$$

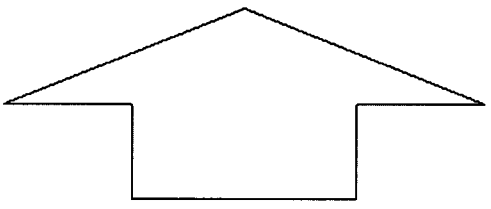
여기서 α는 모양의 면적을 구하는 함수이고 φ는 모양의 경계선 둘레(perimeter)를 구하는 함수이다. 모양의 면적은 모양의 영상 영역 내의 화수의 수로 계산한다. 모양의 경계선 둘레는 경계선에 해당하는 화소의 수로 계산한다. 질의 모양(query shape) obj^q에 대한 골격 나무 T^q는 그 모양이 유사하므로 대부분 1에 근접한 값을 가지나 추상화가 된 경우에는 1보다 작은 값을 가진다. 골격 나무가 표현하는 모양의 넓이는 원래의 넓이와 유사하지만 외곽선이 많이 추상화된 경우에는 w_F가 상대적으로 작아진다. 영역의 일부가 변형되어서 넓이가 많이 줄었지만 그 이외의 부분에서의 외곽선은 유사한 경우에는 w_Z가 상대적으로 작아진다. 이러한 방식으로 영역 기반 기준과 외곽선 기반 기준이 적정하게 균형을 이루게 함으로써 보다 정확한 검출이 가능하다.

4. 실험 결과

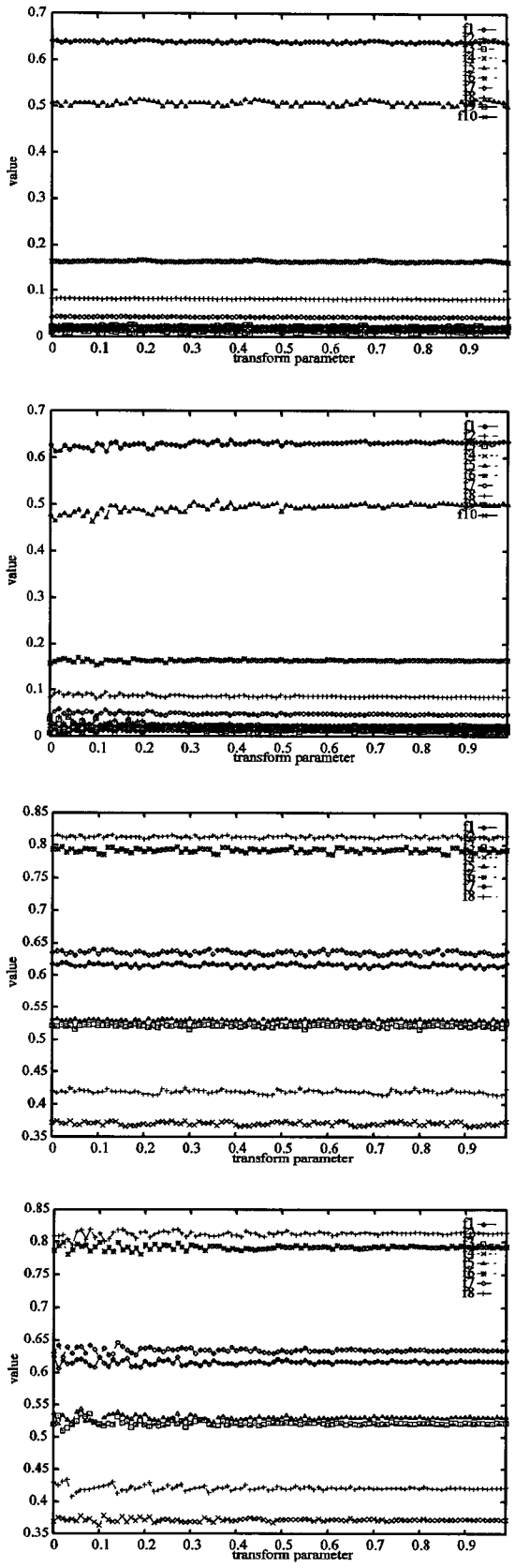
먼저 두 특징벡터 f_Z와 f_H의 불변 특성에 대해서 실험하였다. 테스트 패턴은 (그림 2)와 같은 크기 337×145의 화살표 모양 영상을 사용하였다. 이동 불변 특성의 테스트를 위해서 패턴을 화소 단위로 -400에서 400으로 8 화도 단위로 이동하면서 측정하였다. 회전 불변 특성의 테스트를 위해서 패턴을 3.6도 단위로 0도에서 360도까지 회전하면서 측정하였다. 크기변환 불변 특성의 테스트를 위해서 4% 단위로 50%에서 200%까지 크기 변화하면서 측정하였다.

이동 변환에 대해서는 모든 경우에서 모든 불변량이 동일한 값을 가졌다. Fourier 기술자 불변량의 경우에 m=8로 하여 f_H가 8차원 벡터가 되도록 하였다. 회전과 크기변화에 대한 각 불변량의 측정 결과가 (그림 3)에 있다. Zernike 모멘트 불변량의 10개의 값과 Fourier 기술자 불변량의 8개의 값에 대해서 각 변환 인자에 대해서 일정한 값을 유지함을 알 수 있다. 위의 그래프는 Zernike 모멘트 불변량 f_Z의 불변 특성이고 아래의 그래프는 Fourier 기술자 불변량 f_H의 불변 특성이다. 각각에 대해 왼쪽 그래프는 회전에 대한 불변 특성이고 오른쪽 그래프는 크기변화에 대한 불변 특성이다.

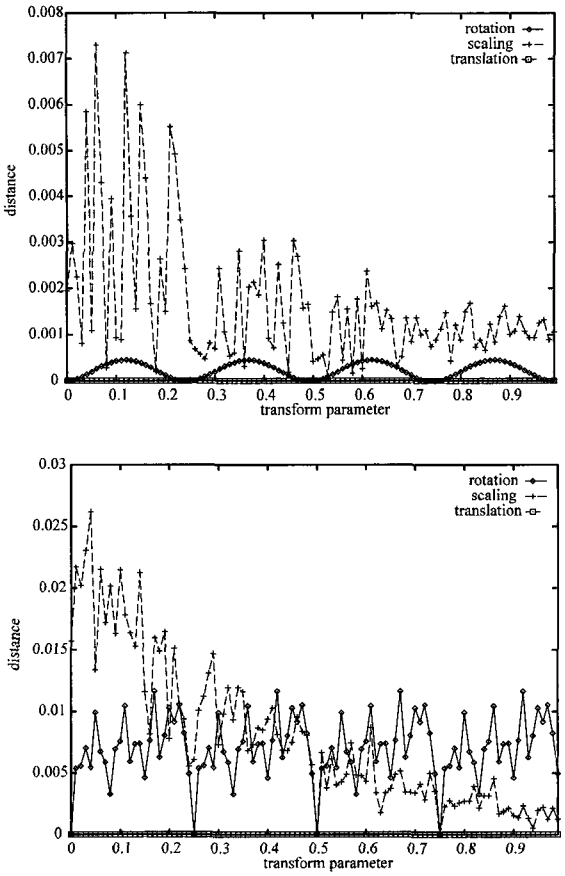
변환 인자의 변화에 따른 f_Z와 f_H의 변화를 측정한 결과가 (그림 4)에 있다. 수직축은 원 패턴의 특징벡터와 변환 패턴의 특징벡터의 Euclid 거리값이다. 수평축은 각 변환의 변환 인자를 0에서 1 사이로 정규화한 값이다. 변환에 따른 특징벡터의 변화가 f_Z의 경우 0.0075 이하로 유지되었고 f_H의 경우 0.026 이하로 유지되었다.



(그림 2) 불변 특성 실험을 위한 테스트 패턴



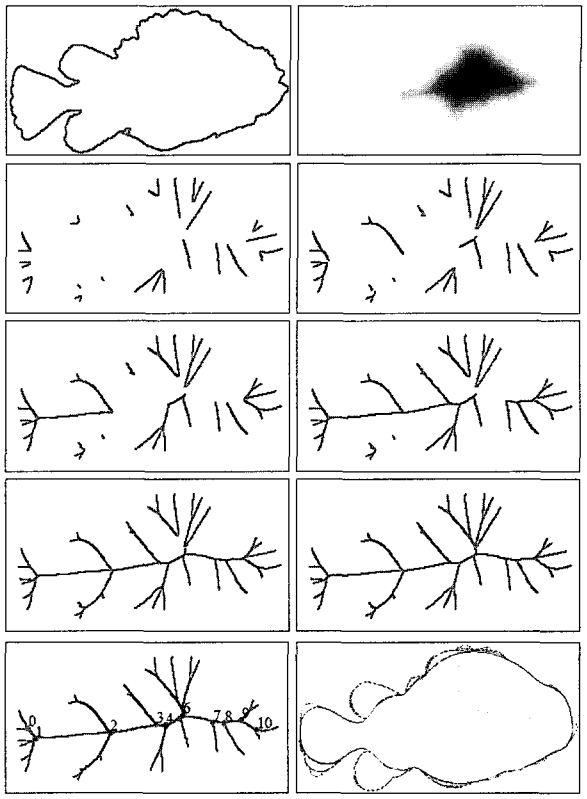
(그림 3) Zernike 모멘트 불변량 fZ의 불변 특성(위)과 Fourier 기술자 불변량 fH의 불변 특성(아래)



(그림 4) Zernike 모멘트 불변량 벡터 fZ(왼쪽)와 Fourier 기술자 불변량 벡터 fH(오른쪽) 변화

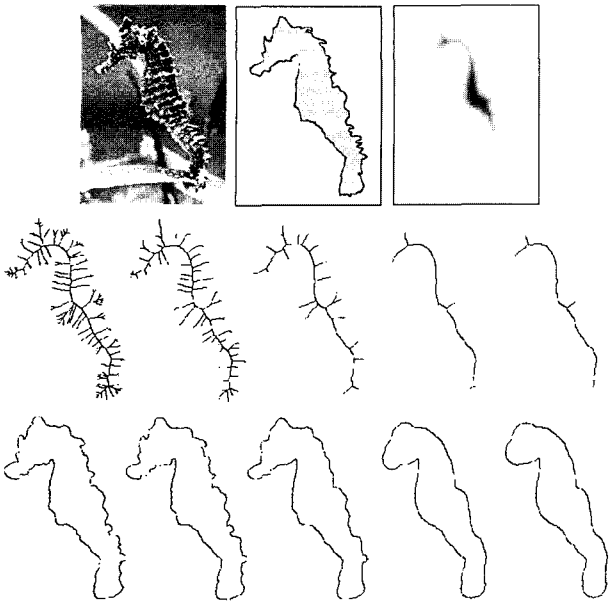
Zernike 모멘트와 Fourier 기술자는 이차원 모양 변환에 대해서 좋은 불변 특성을 보인다. 다만 완전한 불변이 아니고 약간의 변화를 보이는데 이는 연속 함수에 대해서 정의된 불변 특성을 이산적 특성을 가지는 디지털 영상에 적용해서 발생된 결과이다. 이동 변환은 이산적이어도 불변 특성을 동일하게 유지하지만 회전 변환이나 스케일 변화의 경우에는 샘플링이나 양자화로 인해 불변 특성에 변화가 있다.

특징벡터들이 완전한 불변 특성을 유지하더라도 실 영상에 적용하는 데에는 어려움이 있다. 이는 실 영상은 다양한 종류의 가려짐이나 조명에 따른 밝기 변화 등의 요소를 포함한다. 따라서 모멘트를 적용하면 가려짐이나 조명변화에 의해 균일한 모양의 표현이 어렵다. 실 영상에서는 밝기 변화나 경계 부분의 불분명함으로 인해 모양 영역의 외곽선의 정확한 검출이 힘들다. 따라서 Fourier 기술자 값들이 실 영상에서의 환경의 변화에 의해 다양하게 바뀌어질 수 있다. 불변량의 적용을 더욱 어렵게 하는 것은 객체의 움직임으로 인해서 객체의 모양이 바뀌어지는 경우이다. 객체의 일정 부분의 영역이 객체의 움직임으로 인해서 영역의 모습이 변하게 되면 변화 후의 객체의 모습을 동일 객체로 인식하기가 어렵게 된다.



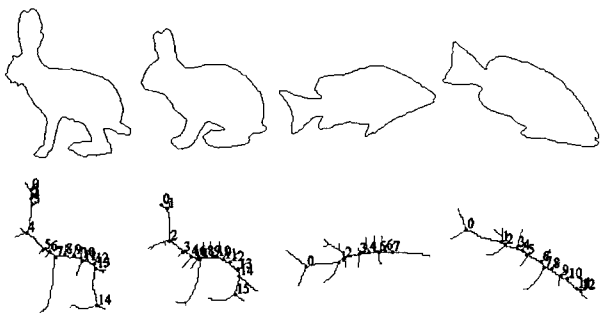
(그림 5) 물고기 영상으로부터의 골격 나무 생성 과정. 왼쪽 위로부터 (a)-(j)임: 물고기 모양 분할 영상(a)과 거리 영상(b), 반복횟수 1,2,3,5,7,13에 대한 골격 세그먼트 탐색 과정(c-h), 최종 형성된 골격 나무 구조(i), 골격 나무 구조로부터 복원된 모양 모습(j)

한 물고기 모양의 이진 영상에 대해서 골격 나무 생성을 실험하였다. (그림 5)는 물고기 모양으로부터 골격 나무를 형성하는 예를 보여준다. 물고기 영상은 (그림 5)(a)에 표시된 299×155 크기의 이진 영상이다. 이로부터 (그림 5)(b)와 같은 거리영상을 구하고 1,223개의 골격 화소들을 생성하였다. 한 끝점 골격 화소로부터 출발하여 교차점 또는 끝점을 만날 때까지 연결된 골격 화소들을 따라간다. 끝점이나 교차점을 만나면 지금까지 따라온 골격 화소들을 하나의 골격 세그먼트로 저장한다. 이러한 방식으로 모든 골격 화소들을 처리할 때까지 반복한다. 골격 세그먼트를 찾는 과정이 (그림 5)(c-h)에 있다. 추출된 골격 세그먼트의 수는 총 71개이고 이 중 31개의 골격 세그먼트가 첫 번째 반복에서 얻어지고 나머지 골격 세그먼트들이 이 후의 12번의 반복에서 얻어진다. 모든 골격 세그먼트가 얻어지면 최대 영향 경로를 구해서 골격 나무를 생성한다. 생성된 골격나무의 모습이 (그림 5)(i)에 있다. 골격나무가 형성되면 객체의 모양은 즉시 복원될 수 있다. (그림 5)(j)에서 가장 안쪽 영역은 최대 영향 경로 상의 링크들인 스파인 링크들로부터만 복원된 모양이고, 중간 바깥쪽 영역은 비종단 노드들의 링크인 뼈대 링크들로부터만 복원된 모양이고, 가장 바깥쪽 영역은 종단 노드들의 링크인 스킨 링크까지 합친 모든 링크들로부터 복원된 모양이다.

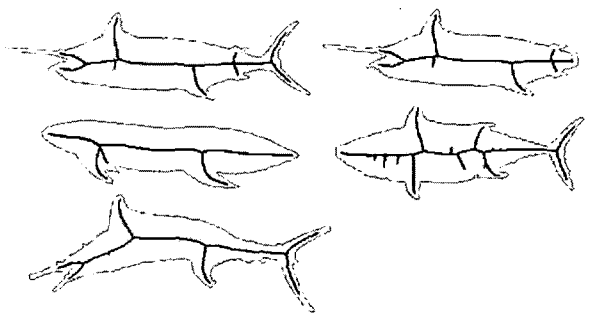


(그림 6) 해마 영상으로부터의 골격 세그먼트 생성 과정. 왼쪽 위로부터 (a)-(m)임: 해마 원 영상(a), 모양 영상(b), 거리영상(c), 문턱값을 3,5,10,20,26로 달리 한 경우의 골격 영상 (d-h) 및 해당하는 영역 복원 (i-m)

골격 화소를 계산할 때 거리 영상에서 일정 거리값 이하인 화소들을 제거함으로써 골격 세그먼트들이 과도하게 많이 생성되는 것을 방지할 수 있다. (그림 6)(a)는 192×252 크기의 해마 영상을 보여준다. 이로부터 얻어진 이진 영상이 (b)이고 거리 영상이 (c)이다. 이진 영상에서의 영역 외곽선에서 뾰족뾰족한 굴곡들이 빈번히 일어나고 이로 인해서 과도하게 많은 골격 세그먼트들이 생성된다. 거리 영상에서 문턱값을 조절하여 불필요한 골격 세그먼트의 생성을 막을 수 있다. (그림 6)(d-h)의 그림이 문턱값을 달리한 경우의 골격 화소들을 보여주고 있다. 문턱값을 높여갈수록 골격 세그먼트의 수가 줄어든다. 그러나 문턱값을 너무 높이면 과도한 모양 추상화로 인해 모양 정보가 손실된다. (그림 6)(i-m)의 영상이 해당 문턱값으로 얻어진 골격 화소들로부터 복원된 모양 정보를 보여준다. 실험적으로 문턱값은 7 화소에서 10 화소 정도의 값이 가장 적절함을 알 수 있었다.



(그림 7) 여러 모양으로부터의 골격 나무의 생성 예



(그림 8) 여러 물고기 모양으로부터 생성한 골격 나무의 모습. 왼쪽 위부터 A,B,C,D,E

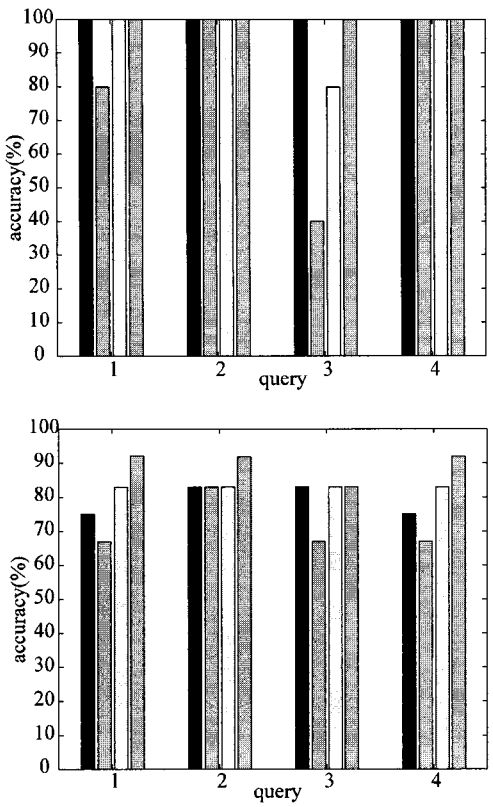
〈표 1〉 유사도 측정의 비교

모양 / 방법	A	B	C	D	E
d_Z	0.000000	0.884384	1.000000	0.080693	0.729096
d_F	0.000000	1.000000	0.554456	0.670575	0.217260
d_C	0.000000	0.942192	0.777228	0.375634	0.473178
d_W	0.000000	0.357106	0.176234	0.821043	0.382429

골격 나무의 생성은 골격 화소를 처리하는 작업이므로 매우 빠르게 수행된다. 또한 임의의 구조에 대하여 항상 올바른 골격 나무가 생성됨을 알 수 있었다. 다만 하나의 이진 영상이 여러 물체에 대한 영역을 포함하는 경우는 배제하였다. 하나의 이진 영상은 하나의 물체 영역만 가진다는 가정을 두었고 만약 여러 영역을 포함하고 있는 경우에는 처음의 하나의 영역에 대해서만 처리를 하였다. 골격 나무의 생성 예가 (그림 7)에 있다.

모양 검색 실험을 위해서 물고기 모양 데이터베이스를 사용하였다. 데이터베이스는 Mokhtarian 등 [8]의 이진 영상들을 사용하였고 각 영상에 대해 골격 나무를 생성하였다. 그림 8에 사용된 영상의 일부가 있다. 모양 A를 질의 영상으로 하였을 때 검색 결과 B,C,E가 유사한 모양으로 검색되었다. 이 중에서 C가 가장 유사도가 높았다. D는 유사도가 떨어져 검색되지 않았다. 유사도 측정 기준을 Zernike 모멘트 불변량만 사용한 경우(d_Z), Fourier 기술자 불변량만 사용한 경우(d_F), 단순 혼합 기준을 사용한 경우(d_C), 가중치 혼합 기준을 사용한 경우(d_W)의 네 경우를 비교하였다. 각 경우에 대한 유사도가 <표 1>에 있다. 값이 0에 가까울수록 유사함을 의미하고 1에 가까울수록 유사하지 않음을 의미한다.

유사도 계산의 정확성을 검증하기 위해 250개의 동물 영상들에 대해서 검색을 시도하였다. 영상들 중에 12개의 토끼 영상을 포함시켰다. 다른 방법들과의 정확성 비교를 위해서 Hu의 모멘트 불변량[16]과 비교하였다. (그림 9)는 가장 유사한 5개의 영상을 검색한 결과와 가장 유사한 10개의 영상을 검색한 결과에 대한 검색 정확도를 보여주고 있다. 각 질의에 대하여 4개의 막대가 표시되어 있는데 가장 왼쪽부터 d_F, d_H, d_Z, d_W에 해당한다. d_F는 Fourier 기술자만을 사용한 결과이고, d_H는 Hu의 모멘트 불변량만을 사용한 결



(그림 9) 검색 정확성의 비교: 가장 유사한 5개의 영상을 검색 (왼쪽), 가장 유사한 10개의 영상을 검색 (오른쪽)

과이고, d_Z는 Zernike 모멘트 불변량만을 사용한 결과이고, d_W는 가중치화된 혼합 특징벡터를 사용한 결과이다. 그림에서와 같이 두 경우 모두 d_H가 가장 떨어지는 성능을 보인 반면 제안된 d_W가 가장 우수한 성능을 보였다.

5. 결 론

본 논문에서는 객체를 계층적으로 표현할 수 있는 모양 표현 기법인 골격 나무 표현을 제안하였다. 골격 나무는 객체의 외곽선에 대한 변형에 대해 민감하지 않은 특징을 보인다. 또한 골격 나무만으로 원래 모양의 복원이 가능하다. 다른 모양과의 비교에 있어서 골격 나무의 일정 수준만 비교하는 방법으로 대략적인 전체 매칭이 가능하고 골격 나무의 특정 부분만 비교하는 방법으로 부분 매칭도 가능하다.

골격 나무는 모양 표현에 있어서 융통성있는 표현 방식으로 모양 검색을 위한 응용 시스템에 활용될 수 있다. 객체의 등록 시에 골격 나무를 생성하고 해당하는 특징벡터를 계산하여 골격나무와 함께 등록한다. 추후 질의 영상에 대해 질의 객체에 해당하는 골격 나무를 생성하고 생성된 골격 나무의 다양한 수준에서의 부분나무들의 특징벡터를 계산하고 이들 특징벡터를 데이터베이스의 특징벡터와 비교하는 방식으로 검색이 이루어진다.

특징벡터는 불변 특성을 가지는 Zernike 모멘트와 Fourier 기술자에 기반하였다. 두 특징벡터의 각 불변량이 이차원 변환에 대해 우수한 불변 특성을 가지지만 실제 적용에 있어서는 영상에서의 다양한 부정확성으로 인해 불변 특성이 잘 유지되지 않는다. 이러한 제한점을 극복하기 위해서 두 불변량의 가중치화된 혼합 사용 방식을 제안하였다. 제안한 골격 나무 표현과 유사도 측정 방식을 다수의 모양 영상에 대해 실험하였다. 실험 결과, 영역 외곽선의 변형이나 훼손에 대해서도 올바른 검색이 가능하였고 단일 불변량을 사용하는 경우보다 높은 검색 정확도를 보였다.

참 고 문 헌

[1] D. Zhang, G. Lu, "Review of shape representation and description techniques," Pattern Recognition, Vol. 37, pp. 1-19, 2004.

[2] M. S. K. B. M. Mehtre, W. F. Lee, "Shape measures for content based image retrieval: a comparison," Information Processing and Management, Vol. 33(3), pp. 319-337, 1997.

[3] A. Telea, C. Sminchisescu, S. J. Dickinson, "Optimal inference for hierarchical skeleton abstraction," ICPR'2004, pp. 19-22. 2004.

[4] A. Torsello, E. R. Hancock, "A skeletal measure of 2D shape similarity," Computer Vision and Image Understanding, Vol. 95(1), pp. 1-29, 2004.

[5] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, S. W. Zucker, "Shock graphs and shape matching," Int. J. Comput. Vision, Vol. 35(1), pp. 13-32, 1999.

[6] T. Sebastian, P. Klein, B. Kimia, "Recognition of shapes by editing their shock graphs," IEEE Trans. Pattern Anal. Machine Intell., Vol. 26(5), pp. 550-571, 2004.

[7] C. D. Ruberto, "Recognition of shapes by attributed skeletal graphs," Pattern Recognition, Vol. 37(1), pp. 21-31, 2004.

[8] F. Mokhtarian, S. Abbasi, J. Kittler, "Robust and efficient shape indexing through curvature scale space," in: Proceedings of British Machine Vision Conference, pp. 53-62, 1996.

[9] S. Sclaroff, A. Pentland, "Modal matching for correspondence and recognition," IEEE Trans. Pattern Anal. Machine Intell., Vol. 17(6), pp. 545-561, 1995.

[10] A. Folkers, H. Samet, "Content-based image retrieval using Fourier descriptors on a logo database," ICPR'2002, Vol. 3, Quebec City, Canada, pp. 521-524, 2002.

[11] C. Arcelli, G. S. di Baja, "Euclidean skeleton via centre-of-maximal-disc extraction," Image and Vision Computing, Vol. 11, pp. 163-173, 1993.

[12] R. M. Haralick, L. G. Shapiro, "Computer and Robot

Vision I," Addison-Wesley, 1992.

[13] D. Attali, A. Montanvert, "Modeling noise for a better simplification of skeletons," ICIP'1996, Vol. 3, pp. 13-16, 1996.

[14] M. K. Hu, "Visual pattern recognition by moment invariants," IRE Transactions in Information Theory, Vol. 8, pp. 179-187, 1962.

[15] M. R. Teague, "Image analysis via the general theory of moments," Journal of Optical Society of America, Vol. 70, pp. 920-930, 1980.

[16] C.-H. Teh, R. T. Chin, "On image analysis by the methods of moments," IEEE Trans. Pattern Anal. Machine Intell., Vol. 10(4), pp. 496-513, 1988.

[17] A. Khotanzad, Y. H. Hong, "Invariant image recognition by Zernike moments," IEEE Trans. Pattern Anal. Machine Intell., Vol. 12(5), pp. 489-497, 1990.

[18] E. A. N. Ezer, B. Sankur, "A comparative study of moment invariants and Fourier descriptors in planar shape recognition," in: Proceedings of 7th Mediterranean Electrotechnical Conference, Vol. 1, pp. 242-245, 1994.

[19] C. W. Chong, P. Raveendran, R. Mukundan, "Translation invariants of Zernike moments," Pattern Recognition, Vol. 36, pp. 1765-1773, 2003.

박 종 승

e-mail : jong@incheon.ac.kr

1992년 경북대학교 전자계산학과 (이학사)

1994년 포항공과대학교 컴퓨터공학과 (공학석사)

1999년 포항공과대학교 컴퓨터공학과 (공학박사)

1999년~2001년 한국전자통신연구원 가상현실연구부 선임연구원

2001년~2003년 디지탈아리아 시각정보처리연구소 책임연구원

2004년~현재 인천대학교 컴퓨터공학과 조교수

관심분야 : 영상처리, 3D 비전, 증강현실, 게임공학