

이진 영상의 골격을 이용한 형상 기술자

이종하*, 최양림**, 조남익*

*서울대학교 전기공학부, **삼성전자

Shape Descriptor using Skeleton of Binary Image

Jong Ha Lee, Yanglim Choi and Nam Ik Cho

e-mail : jhlee@ispl.snu.ac.kr, ychoi@mrrnd.sec.samsung.co.kr, nicho@snu.ac.kr

요약

본 논문에서는 영상처리 기법 중 하나인 이진 영상에 대한 골격화를 이용한 새로운 형상 기술자를 제안한다. 내용기반 영상 검색에서 형상을 이용하는 것은 가장 우수하고 인간이 가장 쉽게 인지할 수 있는 방법 중의 하나이다. 형상을 이용한 검색을 위해서는 인간이 인지하는 형상에 대한 정보를 간략화시킬 수 있는 기술자가 필요하다. 본 논문에서 제안하는 골격을 이용한 형상 기술자는 물체의 중요한 정보 중 하나인 골격을 영상 검색에 이용함으로써 기존의 물체가 갖고 있는 복잡한 형상 정보들을 여러개의 직선의 조합으로 간략하게 표현하고 이를 검색에 사용하는 것이다. 이를 위해서 단순한 골격화 외의 다른 형태학적 영상 처리를 이용하여 효과적인 직선 추출을 위한 여건을 마련한다. 그리고 근사화된 직선들이 추출되면 스케일에 대해 정규화하여 골격을 이루는 직선들의 양 끝점을 형상 기술자로 얻을 수 있다. 각 특징벡터에 대한 정합은 각각의 회전에 대해 유클리디안 거리를 이용한다. 실험 결과, 제안된 방법이 자체한 부분보다는 대략적인 형상 검색과 동일한 카테고리들의 데이터 집합에서 부분적인 변화에 대해 우수한 성능을 나타낸다는 것을 알 수 있다.

1 서론

형상 기술자(Shape Descriptor)는 자동으로 추출 가능한 보다 낮은 추상적인 단계의 기술(lower abstraction level description)로서 영상에서 인간이 매우 쉽게 인식할 수 있는 기본적인 기술자라 할 수 있다. 영상에서 주어지는 물체의 형상을 기술하고 이를 토대로 하여 정합 또는 유사도를 측정하는 알고리즘은 영상인식 분야에 많이 연구되고 있지만 특정 물체가 아닌 일반적인 영상인식에는 아직 많은 문제점들이 있다. 현재 MPEG-7 그룹에서 제안되고 있는 형상 기술자들도 그러한 문제를 극복하기 위해 주어진 물체의 다양한 변형을 통하여 특징점들을 찾는데 주력을 하고 있으며 그 종류도 매우 다양하다. MPEG-7의 표준화 단계인 XM(eXperiment Model)에서 채택되고 있는 형상 기술자는 첫째, 물체의 다양한 형상들에 대해 그 형상들의 영상 내의 분포를

알 수 있도록 여러 종류의 Zernike 기저함수를 준비하고 일정 크기의 영상을 각각의 기저 함수들에 투영시켜 그 값들을 기술자로 사용하는 Zernike Moment 형상 기술자(Zernike Moment Shape Descriptor)[1, 2, 3, 4]와 둘째, 이진 영상으로부터 추출된 외각선을 따라 저주파 통과 필터링을 하면서 외각선 상에 존재하는 변곡점들의 변화를 스케일 공간(scale space) 상에 내타내어 첨두치와 그 위치를 2차원 벡터 형태로 나타내는 곡률 스케일 공간 형상 기술자(Curvature Scale Space Shape Descriptor)[1, 5]가 대표적이다. 전자의 경우, 형상을 이루는 영상과 기저함수가 동일한 크기를 가져야 하며, 형상이 원의 형태를 가져야 하는 제한조건이 있다. 후자의 경우, 추출된 형상이 단 하나의 객체이어야 한다는 단점을 가진다. 본 논문에서 제안하고자 하는 형상 기술자도 위에서 언급한 후자의 경우와 마찬가지로 형상은 추출되어졌다는 가정 하에 이루어져야 하지만 객체의 수에 구애받지 않는다. 그리고, 추출된 기술자가 형상의 대략적인 모습을 그대로 유지함으로써 물체의 부분적인 움직임과 같은 다른 기술자를 기술하는 수단이 될 수 있다는 장점이 있다.

골격을 이용한 형상 기술자는 주어진 형상으로부터 인간의 인식에서 가장 기초적인 선을 추출하여 이를 간략화하여 형상 기술자로 표현하고자 하는 것이다. 특히 에지보다는 골격을 추출함으로써 기술자를 최대한 간략화하여 정합의 과정을 보다 간단하게 할 수 있다. 형상으로부터 골격을 추출하기 위한 방법은 여러가지가 있다. 대표적인 것으로는 형상의 외각으로부터 반복적인 에로전(erosion)을 이용한 방법, 거리 변환으로부터 얻은 거리맵(distance map)에서 최대 크기의 원 중심들을 추출하는 방법[6], 그리고 마찬가지로 거리맵으로부터 국부최대(local maximum)를 추출하는 방법[7] 등이 있다. 본 논문에서는 구현이 어느 정도 용이하고 추출된 골격의 연결성(connectivity)이 보장되는 세번째 방법을 사용한다. 국부최대를 얻기 위한 방법은 Nevatia-Babu의 에지검출 방법[8]의 변형된 형태로서 흑백 영상에서 에지를 검출하는 방법과 비슷하게 검출 마스크를 거리변환된 영상에 적용하여 얻는다. 검출된 골격영상으로부터 근사화된 직선을 얻기 위해서는 골격을 이루는 픽셀들의 연결성과 방향성이 고려되어야 한다. 이

를 보장하기 위해 세선화(thinning) 작업을 거친 후, 필셀을 따라서 연결(linking) 작업을 수행한 후, 각 선들에 대한 시작점과 끝점에 대한 정보를 얻는다. 이를 이용하여 선의 길이와 선들 사이의 거리, 기울기 등에 대한 문턱치를 조절하면서 선들을 병합하여 직선 근사화된 선들의 목록을 작성함으로써 간략화된 골격 정보를 얻게 된다. 그리고 스케일 불변의 기술자를 얻기 위해 스케일에 대해 정규화(normalization)하여 최종적으로 골격에 대한 선들의 목록을 얻게 되며, 이것을 형상 기술자로서 사용하게 된다. 그리고 Housdorff 거리 개념을 사용하여 중심이 되는 선을 찾고 이를 중심으로 일정 간격으로 회전을 시켜가면서 각각의 형상에 대한 기술자들 사이의 유사도를 측정하는 간단한 병합 방법을 사용한다.

본 논문의 2절에서는 이진영상으로부터 거리맵을 얻고 이를 이용하여 골격의 추출과 직선 근사화하는 방법, 3절에서는 유사도를 측정하는 방법에 대해서 서술하고, 4절에서는 실험 결과에 대해서 논의한다. 그리고 마지막으로 5절에서 결론을 맺는다.

2 골격 추출 및 직선 근사화

골격 추출 및 근사화 과정에 대한 간략한 순서도가 그림 1에 나타나 있다. 그림 1의 과정에 따라 이진영상의 골격을 이용한 형상 기술자를 기술할 수 있다.

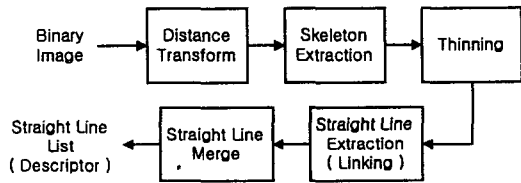


그림 1: 이진 영상의 골격을 이용한 형상 기술자 추출 알고리즘

2.1 골격 추출

거리맵의 국부 최대가 골격을 이루는 점이라는 것은 잘 알려진 사실이다[7]. 거리맵을 얻기 위한 거리변환(distance transform)은 물체 내부의 각 점들을 배경으로부터 최단거리의 값으로 나타내는 함수이다. 그림 2는 임의의 형상에 대한 거리맵을 나타낸 것이다. 거리변환으로부터 얻어진 그림 2 (b)의 거리맵에서 국부 최대를 얻는 방법은 수정된 Nevatia-Babu 에지 검출 방법[8]을 사용한다. 우선 그림 3과 같은 네 방향의 국부 최대 검출 마스크를 마련하고 각각의 마스크에 대하여 컨벌루션을 수행한다. 그 결과, 크기가 가장 큰 결과의 방향에 해당하는 라벨을 기록하는 방향맵(direction map)과 크기맵(magnitude map)을 얻는다. 여기서 얻어진 방향맵에서 해당되는 방향의 90도 회전된 방향에서 크기가 가장 큰 픽셀만 남기고 나머지 픽셀은 제거

함으로써 세선화 작업을 수행하면 그림 2 (d)와 같은 골격화 영상을 얻을 수 있다.

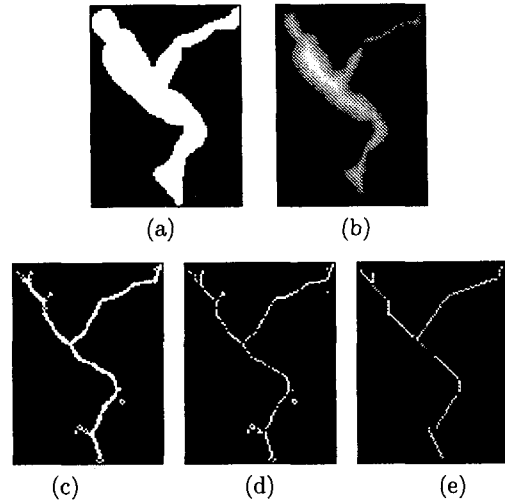


그림 2: (a) 이진영상, (b) 흑백영상으로 스케일링된 거리맵, (c) 골격화 영상, (d) 세선화된 골격화 영상, (e) 10개 이하의 직선으로 근사화 된 결과.

2.2 직선 근사화

여러 형태의 선들로 이루어진 영상에서 직선을 추출한다는 것은 그 영상을 구성하고 있는 각 픽셀들을 하나의 방향을 따라 연결하여 그 시작점과 끝점에 대한 목록을 작성한다는 것을 의미한다. 그러기 위해서 우선 2.1절에서 얻은 골격화 영상을 구성하고 있는 모든 직선 성분을 추출해야 한다. 본 논문에서는 앞 절에서 구한 네가지 방향을 가지는 방향맵을 사용한다. 방향맵에서 동일한 라벨을 갖는 픽셀들끼리 연결하여 각각의 선 세그먼트들의 시작점과 끝점에 대한 목록을 작성한다. 위와 같은 선 연결 작업 후에 여기에서 얻어진 직선들에 대한 목록으로부터 각 직선들 사이의 각도, 거리, 직선의 길이에 대한 문턱치를 변화시켜 가면서 원하는 직선의 개수 이하가 될 때까지 반복 수행을 통해 직선 병합을 이루게 된다. 그 결과가 그림 2 (e)에 나타나 있다. 마지막으로, 각 직선들의 끝점 사이의 최대 거리에 대해 직선의 목록을 정규화함으로써 이진 영상의 골격을 이용한 형상 기술자를 얻게 된다.

3 유사도 측정

본 논문에서 사용한 정합 기법은 골격을 이루는 직선들의 끝점들 간의 거리를 측정하여 최소거리의 합으로 비유사도(dissimilarity) 값을 얻어서 이 값을 비교한다. 비유사도 측정 함수는 다음의 식 (1)-(4)와 같다.

-33 -17 100 -17 -33	100 -15 -25 -35 -45
-33 -17 100 -17 -33	-15 100 -15 -25 -35
-33 -17 100 -17 -33	-25 -15 100 -15 -25
-33 -17 100 -17 -33	-35 -25 -15 100 -15
-33 -17 100 -17 -33	-45 -35 -25 -15 100
(a) 0도	(b) 45도
-33 -33 -33 -33 -33	-45 -35 -25 -15 100
-17 -17 -17 -17 -17	-35 -25 -15 100 -15
100 100 100 100 100	-25 -15 100 -15 -25
-17 -17 -17 -17 -17	-15 100 -15 -25 -35
-33 -33 -33 -33 -33	100 -15 -25 -35 -45
(c) 90도	(d) 135도

그림 3: 국부최대 검출 마스크

$$N = \min\{N_Q, N_M\} \quad (1)$$

$$D_{1k} = \min_{i,j} \{ \|Q_{S_i} - M_{S_j}\| + \|Q_{E_i} - M_{E_j}\| \} \quad (2)$$

$$D_{2k} = \min_{i,j} \{ \|Q_{S_i} - M_{E_j}\| + \|Q_{E_i} - M_{S_j}\| \} \quad (3)$$

$$D = \sum_{k=0}^{L-1} \min\{D_{1k}, D_{2k}\} \quad (4)$$

여기서, Q 는 검색하고자 하는 직선, M 은 검색되는 직선, S 는 직선의 시작점, E 는 직선의 끝점, 그리고 N 은 해당되는 기술자가 가지는 직선의 총 개수이다. 식(2)와 (3)의 두 경우 중에서 최소가 되는 거리의 총합이 비교되는 두 기술자의 비유사도가 된다. 따라서 이 값이 작을수록 두 물체가 서로 더 유사하다는 것을 의미한다. 그리고 일정한 회전각도 간격으로 위와 같은 비유사도 측정함수를 적용하여 비유사도가 최소가 되는 값을 취함으로써 회전 불변의 특성을 얻는다. 이를 위해 Hausdorff 거리[9] 개념을 이용하여 각 직선의 중점에 대해 Hausdorff 거리가 가장 작은 직선을 찾고, 이 직선의 중점을 회전 중심으로 정한다.

4 실험 결과

실험은 MPEG-7의 XM에서 사용되는 이진영상들에 대해 수행했다. 그리고 직선 병합을 위한 각종 문턱치는 경험적으로 정했다. 직선의 병합은 30도 이하의 각에서만 이루어지게 했고, 병합이 이루어지는 두 직선의 끝점 사이의 거리는 실제 영상의 가로,세로 길이 중 작은 값의 5%, 직선 병합 후 무시되는 직선의 길이는 가로,세로 길이 중 큰 값의 1%로 정했다. 그리고 매 반









복 수행마다 그 문턱치를 10%씩 증가시키면서 직선의 수를 10개 이하가 되도록 했다. 실험 결과는 그림 4에 나타나 있다. 다른 종류들의 데이터들에 대해서는 대체로 우수한 성능을 나타내었으나, 인간의 인식에서 다른 종류의 물체라 하더라도 근사화 과정에서 골격이 유사할 경우 일부 잘못된 검색결과를 나타내기도 했다. 이러한 결과에서 알 수 있듯이 골격 추출과 이의 근사화가 유사도 측정에 가장 중요한 문제가 된다. 하지만 동일한 카테고리 내의 데이터 집합에서 부분적인 변화에 대해 우수한 성능을 나타내었다. 이는 본 논문에서 제안된 형상 기술자가 동일한 카테고리의 데이터 집합에서 국부적인 움직임을 추출하는데 유리하다는 것을 알 수 있다.

5 결론



본 논문에서는 이진영상의 골격을 형상 기술자로 기술하여 이를 병합에 이용하는 방법을 제안하였다. 이를 위해 이진영상으로부터 골격을 추출하는 방법과 이를 직선으로 근사화하여 기술자로 나타내는 방법들에 대해 서술하였다. 효과적인 직선 추출을 위해 이진영상을 거리변환하여 국부최대를 구함으로써 골격을 추출하고, 기존의 에지추출 방법을 수정하여 사용하였다. 근사화된 직선을 일정 개수로 제한함으로써 보다 빠른 정합을 수행할 수 있었다. 다양한 형상에 적용한 결과, 자세한 부분보다는 대략적인 형상의 검색과 동일한 카테고리의 데이터 집합에서 형상의 국부적인 변화를 검출하는데 효과적인 결과를 나타냈다. 이는 직선으로의 근사화 과정에서 자세한 부분의 정보를 잃어버리는 단점을 가져옴으로써 나타나는 효과이다. 하지만, 본 논문에서 제안한 기술자 그 자체가 형상의 대략적인 모습을 그대로 가지고 있어서 동일한 물체의 부분적인 움직임을 검출하는데 유용하였으며 따라서 제안된 방법이 물체의 부분적인 움직임을 기술하는 움직임 기술자(motion descriptor)에 대한 연구에 효과적인 성능을 나타낼 것으로 기대된다.

참조 서적

- [1] S. Jeannin, "MPEG-7 Visual part of eXperimentation Model Version 4.0," ISO/IEC JTC1/SC29/WG11/N3068, Dec. 1999.
- [2] W. Kim and Y. Kim, "A rotation invariant geometric shape descriptor using Zernike moment," ISO/IEC JTC1/SC29/WG11 P687, 46th MPEG meeting, Lancaster, Feb. 1999.
- [3] C. Teh and R. T. Chin, "On Image Analysis by the Methods of Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496-513, July 1988.
- [4] Y. Kim and W. Kim, "Content-Based Trademark Retrieval System using Visually Salient

Query	Retrieval Result
	
	
	
	

(a) 임의의 데이터 집합에 대해 수행한 결과

Query	Retrieval Result
	

(b) 분류된 데이터 집합에 대해 수행한 결과

그림 4: 검색 결과

- Feature," *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 307-312, 1997.
- [5] F. Mokhtarian and A. K. Mackworth, "A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 8, pp. 789-805, Aug. 1992.
- [6] Y. Ge and J. M. Fitzpatrick, "On the Generation of Skeletons from Discrete Euclidean Distance Maps," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 11, pp. 1055-1066, Nov. 1996.
- [7] M. W. Wright and F. Fallside, "Skeletonisation as model-based feature detection," *IEE Proceedings-I, Comm., Speech and Vision*, vol. 140, no. 1, pp. 7-11, Feb. 1993.
- [8] R. Nevatia and K. R. Babu, "Linear Feature Extration and Description," *Computer Graphics and Image Processing*, vol. 13, pp. 257-269, 1980.
- [9] D. P. Huttenlocher, C. A. Klanderma and W. J. Rucklidge, "Comparing Images Using the Hausforff Distance," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850-863, Sep. 1993.