

쇼크 그래프를 이용한 효과적인 영상 군집화

장석우[○], 솔리마 카남^{*}, 백우진^{*}

[○]안양대학교 디지털미디어학과

^{*}건국대학교 컴퓨터학과

e-mail: swjang@anyang.ac.kr, {solima,wjpaik}@kku.ac.kr

Effective Image Clustering Using Shock Graphsm

Seok-Woo Jang[○], Solima Khanam^{*}, Woojin Paik^{*}

[○]Dept. of Digital Media, Anyang University

^{*}Dept. of Computer Science, Konkuk University

● 요약 ●

본 논문에서는 쇼크(shock) 그래프 기반의 뼈대 특징을 이용하여 모양 정보를 분류하기 위해 그래프 편집 거리(edit cost) 기반의 k-means 군집화 알고리즘을 적용하는 방법을 제안한다. 본 논문에서 제안된 방법에서는 먼저 질의 영상과 대상 데이터베이스 영상으로부터 뼈대 기반의 쇼크 그래프를 추출한 후 종점(end points)과 분기점(branch points)을 가중치를 이용하여 적응적으로 선택한다. 그런 다음, 두 영상 사이의 편집 거리를 구하여 이를 k-means 군집화 알고리즘의 거리 척도로 적용함으로써 대용량의 영상을 보다 효과적으로 분류한다. 성능을 평가하기 위해서 제안된 알고리즘을 MPEG-7 데이터베이스에 적용하였으며, 그 결과 제안된 영상 분류 방법이 기존의 영상 분류 방법에 비해서 보다 효과적으로 모양 기반의 영상을 분류하였음을 확인하였다.

키워드: 영상 군집화(image clustering), 쇼크 그래프(shock graph), 편집 거리(edit distance)

I. 서론

모양 기반의 물체 인식은 윤곽선을 이용한 방법과 모양 내부(interior)를 이용한 방법으로 구분된다. 그리고 내부를 이용한 방법에는 뼈대(skeleton) 기반의 방법, 그리고 이를 확장한 중간축(medial axis)을 이용한 방법과 쇼크(shock) 그래프를 이용한 방법이 있다[1]. 그리고 뼈대기반의 모양검색 방법들은 비록 시간 복잡도가 높은 편이나 위상적(topological)이고 기하학적인 정보를 제공할 뿐만 아니라 시각적인 변화에 강건하기 때문에 윤곽선(contour) 기반의 검색 방법보다 성능이 우수하다고 알려져 있다.

시간적인 복잡도를 줄이기 위한 뼈대 기반의 방법은 Latecki에 의해 제안되었다[2]. 이 방법에서는 모양 정합의 불안정성이 뼈대의 분기점(branch points) 때문에 발생한다고 판단하고 종점(end points)만을 이용하여 정합을 수행한다. 이에 반하여 Zaboli[3]와 Goh[4]는 분기점만을 이용한 모양 정합을 통해 복잡도를 감소시켰다. 그러나 단지 분기점이나 종점의 한 종류만을 고려하는 것은 항상 정확하다는 보장이 없다. 따라서 가중치를 이용하여 분기점과 종점을 적응적으로 선택하는 알고리즘이 제안된 바 있다[5]. 이 방법에서는 샘플 포인트의 개수를 가중치를 이용하여 동적으로 축소하여 실행시간을 감소시켰다.

내용기반 영상 검색 분야에서 제기되는 또 다른 이슈 중의 하나는 최근 들어 급격하게 증가한 영상 데이터베이스의 크기이다.

이 이슈를 해결하기 위해서 군집화(clustering) 알고리즘과 같은 새로운 기법과 도구가 영상 브라우징(browsing)과 검색(retrieval) 등의 다양한 분야에서 폭넓게 사용되고 있다. 특히, 군집화 알고리즘은 대용량의 데이터베이스 내에 포함된 많은 영상들을 특징이 유사한 그룹별로 효과적으로 분류하기 위해 많이 사용되고 있다.

본 논문에서는 쇼크 그래프 특징을 사용하여 모양을 그룹별로 분류하기 위해 그래프 편집 거리(edit distance) 기반의 k-means 군집화 알고리즘을 적용하는 방법을 제안한다. 본 논문에서 분류하고자 하는 영상 데이터베이스는 클래스의 개수가 정해져 있으며, 대용량의 영상 데이터베이스를 효율적으로 클러스터링 해야 하므로 k-means 알고리즘[6]이 가장 적합하다고 판단하였다. 제안된 영상 분류 방법에서 거리 함수로는 시각적인 변화에 강인성을 보여주며 축소된 시간 복잡도를 가진 쇼크 그래프 기반의 편집 거리를 사용한다.

II. 쇼크 그래프

쇼크 그래프는 물체의 윤곽선 내부에 적어도 두 지점에서 접하는 최대 원의 중심점들의 궤적으로, 모양 검색 분야에서 복잡한 물체를 표현하기 위해 사용된다[5]. 일반적으로, 쇼크 점은 중간축(medial axis) 점의 한 종류이다. 그림 1은 쇼크 그래프를 도식적

으로 간단히 보여주고 있다.

2차원 쇼크 점은 분기점으로 흐름의 시작을 나타내고, 4차원 쇼크 점은 종점으로 흐름의 끝을 나타낸다. 제안된 방법은 시작 부분에서 주어진 질의 영상의 중간축으로부터 종점과 분기점을 적응적으로 선택한다[5]. 그리고 노드 사이의 정합을 위해서 쇼크 포인트 정합과 편집 연산(edit operations)을 사용한다. 제안된 방법에서는 시간 복잡도가 높은 쇼크 그래프의 경계 부분인 연결 커브(joint curve)를 사용하는 대신 질의 영상과 대상 영상 사이의 쇼크 포인트 정합을 사용할 것이다. 또한, 시각적인 변환을 처리하기 위해 축소(contract)나 병합(merge)과 같은 편집 연산으로부터 유도되는 변형 거리(deform cost)를 고려할 것이다. 따라서 두 영상 사이의 정합 거리는 쇼크 포인트의 정합 거리와 변형 거리의 합이 될 것이다.

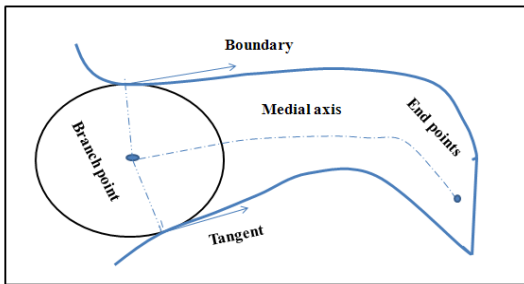


그림 1. 쇼크 그래프
Fig. 1. Shock Graph

쇼크 그래프 사이의 정합 거리를 구하기 위해서는 먼저 쇼크 그래프를 노드(분기점과 종점)로 구분한다. 만일 두 개의 쇼크 그래프 g 와 g^* 가 각각 샘플 노드 (g_1, g_2, \dots, g_n) 와 $(g_1^*, g_2^*, \dots, g_m^*)$ 로 구분된다고 가정하자. 그러면 이 노드들은 $m \times n$ 행렬의 요소로 고려될 수 있다. 만일 $C(i, j)$ 와 $d([k, j], [l, j])$ 를 각각 (g_1, g_2, \dots, g_n) 와 $(g_1^*, g_2^*, \dots, g_m^*)$ 의 정합 거리, 그리고 그래프의 세그먼트(segment) (s_k, \dots, s_i) 와 (s_l, \dots, s_j^*) 의 정합 거리라고 하면 쇼크 포인트 정합 거리 $C(i, j)$ 는 식 (1)과 같이 정의된다. 그리고 식 (1)은 다이나믹 프로그래밍에 의해 해결될 수 있다.

$$C(i, j) = \min_{k, l} [C(i-k, j-l) + d([i-k, i], [j-l, j])] \quad (1)$$

변형 거리의 정의에 의하면 접합(splice) 연산은 쇼크 그래프의 종점을 제거하는 것과 관련이 있으며, 축소(contract) 연산은 분기점을 제거하는 것과 관련이 있다. 따라서 접합 거리는 질의 영상과 대상 영상 사이의 종점의 개수의 차이이다. 유사하게, 축소 거리는 분기점 개수의 차이이다. 만일 D 를 쇼크 포인트(종점 또는 분기점)의 차이라고 가정하면 질의 영상과 대상 영상 사이의 전체 그래프의 편집 거리 M 은 식 (2)와 같이 정의될 수 있다.

$$M = C + D \quad (2)$$

III. 영상 균집화

EM(expectation-maximization) 알고리즘은 문제 속에 정보가 숨어 있을 때 그 정보의 최적 해를 찾아내는 데 매우 유용한 방법이다[7]. 예를 들어, 어떠한 데이터를 클러스터링 하는 문제에서 만약 각 데이터 점들이 어떠한 집합 혹은 클러스터에 속하는 것을 인지할 경우에 최적의 코드 벡터들은 간단히 이들 클러스터의 중심이 될 것이다. 그러나 불행하게도 우리는 데이터 점들이 어느 그룹에 속해 있는지 알 수 없으며, 결국 이것이 감춰진 정보가 된다. EM 알고리즘은 현존하는 최상의 해결법으로 문제 속에 숨겨진 정보를 일단 추정(expectation 단계)하여 얻은 해에서 이 추정값을 개선하기 위하여 다시 이것을 이용하는 절차(maximization 단계)를 해가 수렴할 때까지 번갈아 가며 수행하는 방법이다. EM 최적화 알고리즘 중에서 가장 간단한 예가 k-means 알고리즘이다.

EM 알고리즘의 경우 초기값 선택이 최적화에 결정적인 요소가 되며, k-means 알고리즘은 임의의 초기값에서 정해진 수렴값에 수렴할 때까지 추정(E)과 최대화(M) 과정을 반복하면서 중심을 찾는다. 즉, E 단계에서는 중심에서 클러스터를 선택하여 결정하고, M 단계에서는 거꾸로 클러스터에서 중심을 결정하는 과정을 반복하는 추정 알고리즘이다.

k-means 알고리즘에서 사용하는 거리척도 d 는 식 (3)과 같이 정의되는 유클리디안 거리를 사용한다. 식 (3)에서 x 와 y 는 특정 벡터를 의미한다.

$$d(x, y) = \|x - y\| = \sqrt{(x - y)^T(x - y)} \quad (3)$$

k-means 알고리즘에서 각 클래스의 중심 c 의 계산은 다음 식 (4)와 같이, 유클리디안 거리를 척도로 사용할 경우 단순히 집합 중심의 평균이 된다.

$$c(x) = \frac{1}{N} \sum_{n=1}^N x_n \quad (4)$$

k-means 알고리즘에서 왜곡(distortion)이 안정되었는지를 확인하는 효과적인 방법은 이전 반복 시의 왜곡과 현재 왜곡에 대한 상대적인 감소치를 점검하는 것이다. 이를 수식으로 표현하면 식 (5)와 같다.

$$\Delta D = \frac{D_{prev} - D_{curr}}{D_{prev}} < 10^{-4} \quad (5)$$

이 알고리즘은 일반적으로 100번을 반복하기 이전에 수렴해야 한다. 위에서 언급했듯이 k-means 알고리즘의 목적은 입력된 객체, 그리고 해당 클러스터의 평균 사이의 거리를 기반으로 입력된 자료를 가장 유사한 클러스터에 할당함으로써 입력 자료를 k개의 클러스터로 효과적으로 분류하는 것이다.

본 논문에서는 모양을 k개의 클래스로 분류하기 위해서 쇼크 그래프 특징을 가지고 있는 N개의 모양 객체 $O = \{O_1, O_2, \dots, O_N\}$ 가 존재할 경우 각 객체가 가지고 있는 특징에 따라 입력 객체 O_i 를 j번째 클러스터에 할당함으로써 객체들을 k개의 클러스터에 할당한다. 본 논문에서 사용하는 모양 분류를 위한 k-means 군

집화 알고리즘은 식 (2)에서 정의한 그래프 편집 거리를 거리 척도로 사용한다.

IV. 실험 결과

제안된 방법의 시작 부분에서는 가중화된 알고리즘[5]을 사용하여 샘플 포인트를 선택한다. 예를 들어, 본 논문에서는 표 1과 같이 사과(apple), 박쥐(bat), 그리고 벌레(beetle)의 3개의 그룹으로 구성된 MPEG-7의 영상을 사용하였다. 표 1에서 편집 거리 척도는 쇼크 그래프의 종점(end points)과 분기점(branch point)을 이용하여 각 질의영상(사과, 박쥐, 벌레)에 따라 제시된다. 거리 척도의 값들은 영상을 개별적인 그룹으로 분류하기 위해 제안된 k-means 군집화 알고리즘의 거리 척도로 사용된다.

그림 2는 본 논문에서 제안된 방법에 의해서 15개의 영상이 그룹 2와 그룹 3의 5번째 영상을 제외하고는 3개의 그룹(apple, bat, beetle)으로 올바르게 분류되는 결과를 보여준다. 그러나 bat와 beetle은 동물 부류(category)에 속하므로 이것은 부류 내의 에러이다. 또한, k-means 군집화 알고리즘은 대용량의 영상 데이터베이스의 처리에 매우 효율적이다. 그리고 이 알고리즘의 계산상의 복잡도도 $O(Nkt)$ 인데, N은 물체의 전체 개수를 의미하고, k는 클러스터의 개수를 나타내며, t는 알고리즘의 반복 회수를 의미한다. 더욱이, 그래프 편집 거리 척도는 다른 때대 기반의 방법보다 감소된 복잡도를 가지고 수행된다. 결과적으로, 제안된 영상 분류 시스템은 간단하면서도 효율적인 방식으로 수행된다.

그림 3은 본 논문에서 제안한 방법과 MPEG-7 데이터베이스를 사용하는 다른 모양 기반의 분류 방법들과의 성능을 비교하여 보여준다. 다른 때대 기반의 분류 방법들은 MPEG-7 데이터베이스를 이용하지 않으므로 본 논문에서는 제안된 방법과 이들을 비교할 수는 없다. 일반적으로, MPEG-7 데이터베이스는 서로 다른 종류(약 70 부류)의 많은 영상들로 구성되어 있다. 그리고 각 부류는 20개의 서로 다른 영상으로 이루어져 있다. 이 데이터베이스에서 동일한 그룹의 어떤 영상들은 시각적으로 일정부분 동일하지 않을 수 있으며, 반대로 서로 다른 그룹의 어떤 영상들은 시각적으로 유사성을 가질 수 있다. 본 논문에서는 윤곽선 기반[8]의 방법과 영역 기반의 방법[9]을 구현한 후, 이들과 제안된 방법의 성능을 비교 분석하였다. 그 결과 표 1에서와 같이 본 논문에서 제안된 영상 분류 방법이 보다 우수함을 확인하였다. 윤곽선 기반의 방법은 분류 성공률 측면에서는 매우 우수한 것처럼 보이나 다양한 종류의 영상에 적용할 수 없다는 단점을 가진다. 즉, 상표나 로고, 엠블럼, 클립아트, 문자 등과 같이 연결되지 않은 영역을 가진 복잡한 물체나 작은 틈을 가진 윤곽선의 경우에는 적용하여도 성능이 매우 좋지 않은 결과를 가진다.

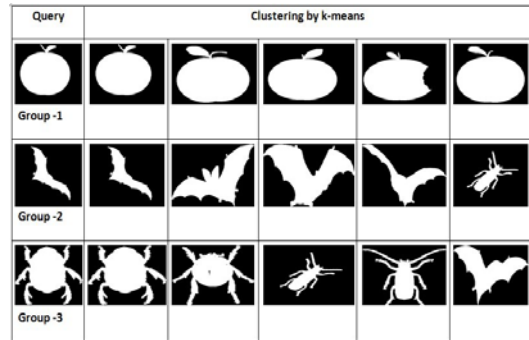


그림 2. 영상 분류
Fig. 2. Image Classification

표 1. 거리 척도
Table 1. Distance Measure

No	영상	종점 개수	분기점 개수	Apple1 거리	Bat3 거리	Beetle10 거리
1	Apple1	5	7	0	40	89
2	Apple3	6	8	17	39	88
3	Apple9	7	9	14	40	89
4	Apple14	7	9	17	39	88
5	Apple15	8	10	23	37	86
6	Bat3	25	27	26	38	93
7	Bat4	40	40	45	38	84
8	Bat8	160	138	40	0	83
9	Bat10	65	64	49	45	81
10	Bat20	32	35	41	32	70
11	Beetle3	154	129	89	83	0
12	Beetle7	80	72	90	56	45
13	Beetle10	94	83	97	74	75
14	Beetle11	127	117	113	78	54
15	Beetle20	148	107	124	89	62

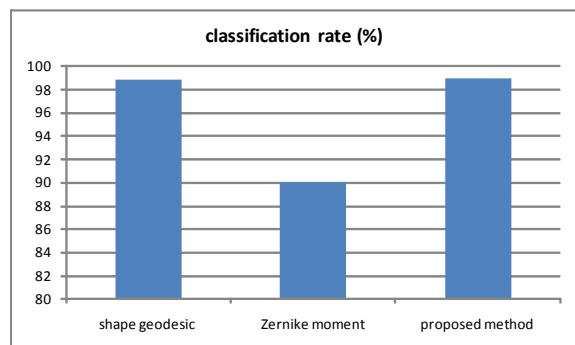


그림 3. 분류 비율
Fig. 3. Classification Rate

IV. 결론

본 논문에서는 모양 기반의 영상 검색 분야에서 매우 유용하게 사용되는 방법 중에 하나인 뼈대 특징을 이용한 쇼크 그래프를 기반으로 모양을 분류하기 위해서 k-means 군집화 알고리즘을 추가적으로 적용하였다. 그리고 이 알고리즘에서 정합 거리 함수로는 시각적으로 변화하는 환경에서 강인성을 보여주는 그래프 편집 거리를 사용하였다. 본 논문에서 제안한 쇼크 그래프 기반의 k-means 군집화 알고리즘의 접근 방법은 MPEG-7과 같은 대용량의 모양 분류를 위해서는 아직까지 시도되지 않은 새로운 방법으로 파악되었다. 성능평가 실험에서는 본 논문에서 제안한 방법의 영상 분류 정확도가 매우 높았으며, 계산 복잡도도 일정 부분 감소하였다.

향후 연구로는 보다 다양한 종류의 모양 정보를 포함하고 있는 영상을 이용하여 제안된 쇼크 그래프 특징 기반의 군집화를 이용한 모양 분류 방법의 성능을 비교 평가할 예정이다. 또한, 제안된 방법에서 사용된 여러 가지 파라미터의 조율을 수행할 예정이다.

참고문헌

[1] T. B. Sebastian, P. N. Klein, B. B. Kimia, "Recognition of Shapes by Editing Their Shock Graphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 5, pp. 550~571, 2004.

[2] X. Bai and L. Jan Latecki, "Path Similarity Skeleton Graph Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 7, pp. 1282~1292, July 2008.

[3] H. Zaboli, M. Rahmati, and A. Mirzaei, "Shape Recognition by Clustering and Matching of Skeletons," *Journal of Computers*, Vol. 3, No. 5, pp. 24~33, May 2008.

[4] W.-B. Goh, "Strategies for Shape Matching Using Skeletons," *Computer Vision and Image Understanding*, Vol. 110, Issue 3, pp. 326~345, 2008.

[5] S. Khanam, S.-W. Jang, W.-J. Paik, "An Improved Shock Graph-based Edit Distance Approach Using an Adaptive Weighting Scheme," *International Conference on FGCN*, pp. 501~508, 2009.

[6] M.-C. Chiang, C.-W. Tsai, and C.-S. Yang, "A Time-Efficient Pattern Reduction Algorithm for K-Means Clustering," *Information Sciences*, Vol. 181, No. 4, pp. 716~731, February 2011.

[7] Y. Bazi, L. Bruzzone, and F. Melgani, "Image Thresholding based on the EM Algorithm and the Generalized Gaussian Distribution," *Pattern Recognition*, Vol. 40, No. 2, pp. 619~634, 2007.

[8] K. Nasreddine, A. Benzinou, and R. Fablet, "Variational Shape Matching for Shape Classification and Retrieval," *Pattern Recognition Letters*, Vol. 31, No. 12, pp. 1650~1657, September 2010.

[9] W.-Y. Kim and Y.-S. Kim, "A Region-based Shape Descriptor Using Zernike Moments," *Signal Processing: Image Communication*, Vol. 16, No. 1~2, pp. 95~102, September 2000.