

영상 데이터 검색을 위한 다계층 고유벡터 모양 정보 기술자

Multi-layer Eigenvector Shape descriptor for Image Retrieval Applications

김중득, 김해광

현대전자산업주식회사 정보통신연구소

요 약

멀티미디어 데이터의 증가로 사용자가 원하는 데이터의 신속하고 정확한 검색이 필요하게 되었다. 본 논문에서는 모양 정보를 기반으로 영상 데이터를 효과적이며 효율적으로 검색하기 위하여, 새로운 모양 정보 특징 및 검색 방법을 제안한다. 본 논문에서는 화소의 공간적분포로 나타나는 모양 정보를 covariance matrix 의 eigenvector 를 이용하여, 계층적으로 영역을 분할하고, 각 분할된 영역에서 크기 변화, 위치 이동, 회전에 불변하는 특징들을 추출한다. 영상 정보의 검색은 특징벡터 공간에서 질의 영상에서 추출된 특징과, 데이터베이스에 기록된 영상들의 특징 사이의 거리를 계산하여, 거리에 반비례하는 유사도가 높은 영상들을 출력한다. 제안된 모양 특징은 또한 계층수의 조정에 의해서 모양 정보를 표현할 수 있는 정도를 조절 할 수 있다는 장점이 있다.

1. 서 론

최근 들어서 대용량 저장 매체, 고속 통신망, 고성능 프로세서 기술의 발달로 방대한 양의 멀티미디어 데이터가 존재하게 되었다. 그리고 사용자는 방대한 양의 멀티미디어 데이터에서 사용자가 원하는 데이터를 검색하는 작업이 필요하게 되었다. 멀티미디어 데이터 검색을 위해서는 문자, 음향, 영상 등의 자료들을 기술 (Description) 과정을 거쳐 데이터를 특징적으로 대표하도록 고안된 데이터 모델로 표현하고, 표현된 데이터 모델 사이의 유사도를 계산함으로써 사용자가 원하는 데이터를 검색한다. 여기서 사용되는 멀티미디어 데이터의 데이터 모델을 기술자 (Descriptor)라고도 한다. 정지 영상을 표현하기 위한 특징으로 색상, 질감, 모양 정보 등의 특징을 사용할 수 있으며 본 논문에서는 영상 데이터의 특징으로 모양 정보를 이용하여 영상을 검색하는 방법에 대해서 기술한다.

모양 정보의 특징으로는 기하학적 특성 (둘레, 면적)이나 기하학적 모우먼트 (Geometric Moments)^[1], Zernike Moments^[2], 기하학적 해싱 (Geometric Hashing)^[3]을 등의 정보가 사용될 수 있다. 본 논문에서는 영역 기반의 모양 정보 검색을 목적으로 한 크기 변화, 위치 이동, 회전에 불변하는 특성을 가진 특징을 제안하고 있으며, 이 방법은 변형에 불변하기 위한 전처리가 불필요하고, 적은 데이터의 특징 정보를 사용하고, 특징 추출이 간단하고, 모양 정보 검색 결과의 정확도가 높다는 장점이

있다.

제 2 장에서는 모양 정보의 특징을 추출하는 방법과 추출된 특징으로 두 모양 정보의 유사도를 계산하는 방법에 대해서 기술한다. 제 3 장에서는 제안된 방법의 성능을 평가하기 위한 실험 결과에 대해서 설명하며, 마지막으로 제 4 장에서 결론을 내린다.

2. Multi-layer Eigenvector (MLEV) Shape Descriptor

(1) Covariance Matrix Eigenvector

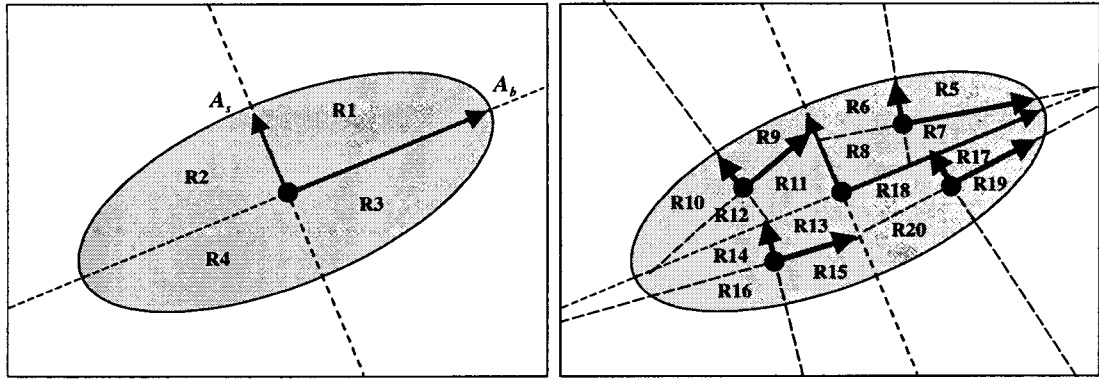
CMEV (Covariance Matrix eigenvector)^[11]를 이용하여 대략 적인 모양 정보를 표시할 수 있다. 모양 정보의 화소의 위치를 2 차원 (x,y) 좌표로 구성된 Random Vector 로 생각하고, 실제 모양의 화소들의 위치 정보로부터 Random Vector 의 Covariance matrix 를 구하고 이의 eigenvector 를 구한다. 이렇게 구해진 두 eigenvector 는 각각 모양 정보의 최대 분산 방향과 최소 분산 방향을 나타내고, eigenvalue 는 분산 정도를 나타낸다.

이와 같이 구한 두개의 eigenvalue 로도 모양 정보를 표현할 수 있으나, 세밀한 모양 정보를 표현하기 위해서는 정보가 불충분하다. 따라서, 본 논문에서는 두개의 eigenvector 를 연장하여 만들어지는 두개의 축을 재귀적으로 사용하여, 모양 정보를 계층별로 세분화하고, 각각의 세분화된 영역을 변형에 불변하는 특징으로 표현하는 다계층 고유벡터를 모양정보 기술자를 제안한다.

(2) Multi-layer Eigenvector (MLEV) Shape Descriptor

본 논문에서 제시하는 모양 정보 특징 추출 과정은 다음과 같다. 입력 모양 정보로부터 covariance matrix 를 계산하고, 계산된 covariance matrix 의 eigenvector 를 구한다. 구해진 두 eigenvector 를 두 축으로 입력 영상을 4 개의 하위영역 (sub-region)으로 분할한다. 그리고 분할된 각각의 하위 영역에 대하여 covariance matrix 및 그 eigenvector 를 구하여 영역을 분할하는 과정을 주어진 계층의 수 만큼 반복한다. 이와 같은 방법으로 구해진 각 영역에 대해 4 개의 변형에 불변하는 특징을 추출한다.

모양 정보 특징 추출 과정 중 영역 분할 과정의 예를 그림 1 에서 나타내었다. 그림 1 (a)는 첫번째 계층에서의 영역 분할 과정이다. 입력 영상으로부터 두개의 eigenvector (A_1, A_2)를 구하고, 두개의 eigenvector 로 구성되는 두 축으로 최초의 영역 R_0 를 4 개의 하위 영역 (R_1, R_2, R_3, R_4)으로 분할한다. 그림 1 (b)는 두 번째 계층에서의 영역 분할 과정이다. 첫번째 과정에서 분할된 각각의 하위 영역 R_1, R_2, R_3, R_4 에서 대해서 앞에서 설명한 방법과 동일한 방법으로 $R_5 \sim R_{20}$ 까지의 하위 영역으로 분할한다. 그리고 첫번째 계층의 영역을 루트 영역 (root region) 이라고 한다.



(a) 첫번째 계층 영역 분할

(b) 두 번째 계층 영역 분할

그림 1. 영역 분할 예.

영역 분할 과정이 완료된 후 분할된 각각의 영역으로부터 모양 정보를 표현할 수 있는 특징을 추출하는 과정을 수행한다. 본 논문에서 사용할 모양 정보의 특징으로 *eigenvalue*의 비, *eigenvector*의 방향, 영역의 무게 중심, 영역의 밀집도 (*compactness*)의 4 가지 성질을 이용한다. 다음은 각각의 성질에 대해서 자세한 설명이다.

1) *Eigenvalue*의 비 : *eigen_ratio*

두 축을 이루는 *eigenvalue*의 비를 *eigen_ratio*라 하며, 식 (1)과 같이 계산된다.

$$eigen_ratio = \text{단축의 } eigenvalue / \text{장축의 } eigenvalue \quad (1)$$

장축 또는 단축의 *eigenvalue*를 사용하지 않고 장축과 단축의 *eigenvalue*의 비를 사용하는 이유는 크기 변화에 일관된 성질을 가지도록 하기 위해서 이다.

2) *Eigenvector*의 방향 : *normal_angle*

본 논문에서 사용하는 *eigenvector*의 방향은 회전에 대해서 일관된 성질을 갖기 위해서 현재 영역에서 구해진 장축의 *eigenvector*와 루트 영역에서 구해진 장축의 *eigenvector*로 이루어진 두 축의 각도 차로 정의한다. 그리고 두 축의 각도차는 그림 2의 θ, θ' 와 같이 두개의 각이 나오는데, 그 중 적은 값을 *normal_angle*이라고 정의하고 모양 정보의 특징으로 사용한다.

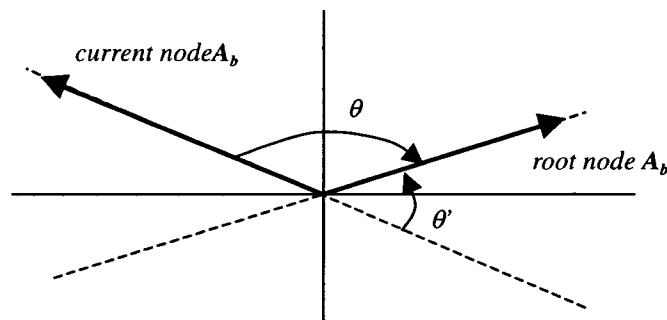


그림 2. *normal_angle* 계산 예제.

3) 영역의 무게 중심 : *center_point*

분할된 영역의 무게 중심은 루트 영역의 무게 중심으로부터 현재 영역의 무게 중심까지의 거리로 정의한다. 그리고 크기 변화에 대해서 일관된 성질을 갖기 위해서 루트 영역의 장축의 eigenvalue (λ_b) 로 정규화 시킨다. 따라서 영역의 무게 중심은 식 (2)로 계산된다.

$$center_point(i) = \frac{D(i)}{\lambda_b} \quad (2)$$

여기서 $D(i)$ 는 루트 영역의 무게 중심과 i 번째 영역의 무게 중심 사이의 거리이다.

4) 영역의 밀집도 : *compactness*

영역의 밀집도는 전체 영역에 대한 물체 내부에 속하는 화소수의 비를 나타낸다. 본 논문에서 밀집도는 각 영역의 참조 영역에 대한 해당 영역의 물체 내부 화소수의 비로 나타낸다. 우선 첫번째 계층에서의 참조 영역은 루트 영역에서 구해지는 eigenvector 의 두 축과 평행하고 물체를 완전히 포함하는 직사각형 영역을 참조영역으로 한다. 그리고 두 번째 이상의 계층에서의 참조 영역을 첫번째 계층에서의 참조영역을 물체 영역의 영역분할 방법과 동일한 방법으로 분할한 영역이다. 따라서 각 영역의 밀집도는 식 (3)으로 계산된다.

$$compactness(i) = \frac{Ri}{Ti} \quad i: \text{Region Number} \quad (3)$$

여기서, Ri, Ti 는 각각 i 번째 영역에서 물체 영역의 화소수와 참조 영역의 화소수 이다.

(3) 유사도 측정

본 논문에서 제안하는 모양 정보의 특징은 계층별로 세분화된 각 영역별로 *eigen_ratio*, *normal_angle*, *center_point*, *compactness* (단, 첫번째 계층은 *eigen_ratio*, *compactness* 만 사용)의 4 개의 특징 값을 가지므로 두 영상의 유사도는 각 영역의 유사도의 합으로 계산된다. 따라서 질의 영상과 대상 영상의 유사도는 식 (4)로 계산된다. *Total_diff* 는 질의 영상과 대상 영상의 상이도 (difference)로 *Total_diff* 가 적을수록 유사도가 높고, *Total_diff* 가 클수록 유사도가 낮음을 나타낸다.

$$Total_diff = w_{er} \cdot eigen_ratio_diff(0) + w_{comp} \cdot compact_diff(0) + \sum_{n=1}^{N-1} \left(w_{er} \cdot eigen_ratio_diff(n) + w_a \cdot normal_angle_diff(n) + w_{cp} \cdot center_diff(n) + w_{comp} \cdot compact_diff(n) \right) \quad (4)$$

여기서, N 은 분할된 영역의 수이고, $w_{er}, w_a, w_{cp}, w_{comp}$ 은 *eigen_ratio*, *normal_angle*, *center_point*, *compactness* 에 대한 가중치이며, *eigen_ratio_diff(n)*, *normal_angle_diff(n)*, *center_diff(n)*, *compact_diff(n)* 는 각각 질의 영상과 대상 영상의 n 번째 영역에서의 *eigen_ratio*, *normal_angle*, *center_point*, *compactness* 의 상이도로 식 (5)~(8)으로 계산된다.

$$eigen_ratio_diff(n) = |eigen_ratio(Qn) - eigen_ratio(Tn)| \quad (5)$$

$$normal_angle_diff(n) = |normal_angle(Q_n) - normal_angle(T_n)| \times (1 - eigen_ratio(Q_n)) \quad (6)$$

$$center_diff(n) = |center_point(Q_n) - center_point(T_n)| \quad (7)$$

$$compact_diff(n) = |compactness(Q_n) - compactness(T_n)| \quad (8)$$

여기서, Q_n 은 질의 영상의 n 번째 영역이며, T_n 은 대상 영상의 n 번째 영역이다.

3. 실험 결과

(1) 실험 데이터

본 논문에서 제시하는 이진 모양 정보 기술자의 성능을 평가 하기 위한 실험 데이터로 MPEG7 테스트 영상인 트레이드마크¹⁴⁾를 사용한다. 총 2810 장의 영상으로 구성되어 있으면 9 개의 유사한 영상들의 그룹으로 나누어져 있다.

그림 3 은 9 개의 그룹과 그룹에 포함되어 있는 영상들이다.

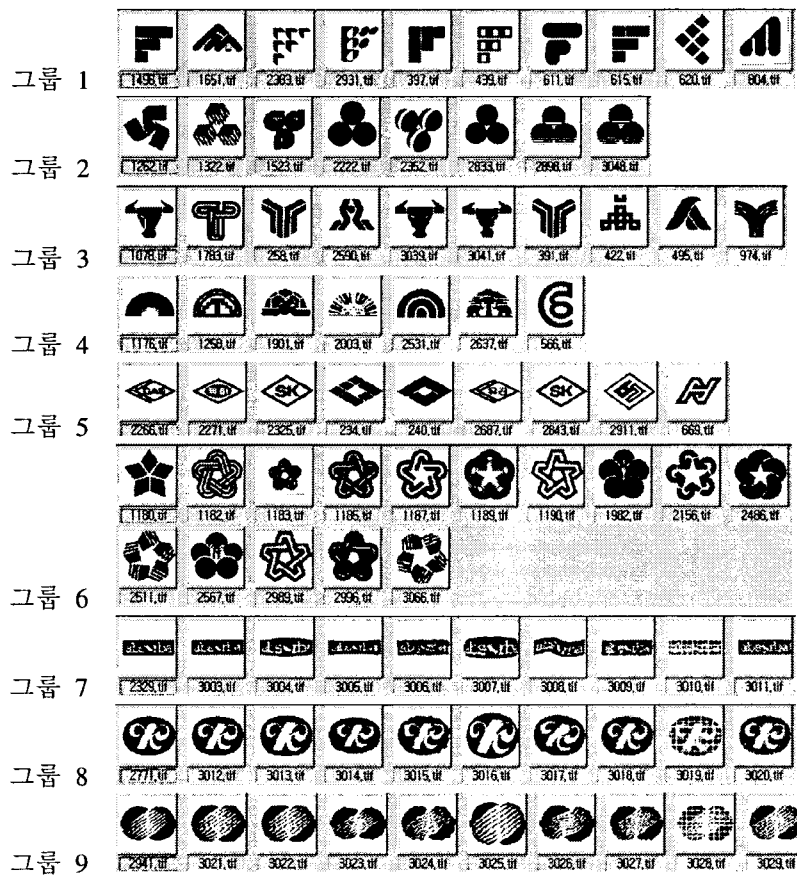


그림 3. 실험 영상 그룹

(2) 성능 평가 방법

이진 영상 정보 검색의 성능을 평가 하는 방법으로 본 논문에서는 Bull's Eye Performance (BEP)를

사용한다. BEP는 N 개의 유사 영상이 존재하는 그룹 내의 한 영상을 질의 했을 경우 상위 $2N$ 개의 검색 결과 중 동일 그룹의 영상이 존재할 확률로 정의한다.

(3) 실험 결과

본 논문의 성능을 평가하기 위한 실험 조건으로 3개 계층을 사용하였으며, 가중치 w_{er} , w_{ar} , w_{cp} , w_{comp} 는 각각 1.0, 10, 0.8, 0.2로 설정하였다.

표 1은 본 논문에서 제시하는 이진 모양 정보 기술자를 이용한 검색의 결과를 나타내었다. 표 1의 결과는 각 그룹 별로 각 그룹의 모든 영상에 대한 BEP의 평균값이다. 본 논문에서 제시하는 이진 모양 정보 기술자의 성능을 평가하기 위해 invariant moments^[5]를 기술자로 사용하였을 경우와 비교하였다.

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Group9
MLEV(%)	87.00	70.31	81.00	69.39	71.39	51.11	76.00	83.00	91.00
Moments(%)	24.00	26.56	32.00	24.49	16.05	16.44	35.00	34.00	32.00

표 1. 트레이드마크 검색 결과.

4. 결 론

본 논문에서는 영상 정보의 특징을 표현하는 기술자로 모양 정보의 화소 분포에 대한 covariance matrix eigenvector를 계층적으로 적용하는 방법을 사용하였다. 즉, 계층별로 입력 모양 정보를 분할하고 분할된 각 영역에 대해서 특징을 추출함으로써, 기존의 두개의 eigenvector로 전체 모양 정보를 표현하는 방법에서 복잡한 모양 정보를 표현하지 못한다는 단점을 보완하였다. 그리고 입력 모양 정보의 크기 변화, 회전, 이동에 대해서 일관된 성질을 가진다는 장점과 특징 추출 과정이 covariance matrix 계산 과정과, eigenvector 추출 과정 등 비교적 간단한 계산 과정을 반복적으로 사용하므로 낮은 계산 복잡도로 특징을 추출할 수 있다는 장점이 있다. 그리고 실험 결과에서 보인 것과 같이 모양 정보 검색에서도 기존의 기술자에 비하여 우수한 성능을 보임을 알 수 있다.

참고 문헌

- [1] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Addison-Welsey Publishing company, Inc., 1993.
- [2] Alireza Khotanzad, Yaw Hua Hong, "Invariant Image Recognition by Zernike Moments", *IEEE Trans. Pattern Anal. Machine Intell.*, vol 12, No. 5, pp 489-pp497, May 1990.
- [3] S. D. Cohen and L. J. Guibas. "Shape-based Image Retrieval Using Geometric Hashing," *Proc. ARPA Image Understanding Workshop*, May 1997.
- [4] MPEG Requirements Group, "Description of MPEG-7 Content Set" *Doc, ISO/MPEG N2467*, October 1998.
- [5] M. K. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. On Info. Theory*, pp179-pp187, February 1962.