

특징점 추출과 영상분할을 이용한 영역기반 영상검색 방법

A Region-Based Image Retrieval Algorithm Using Salient Point Extraction and Image Segmentation

이희경, 호요성

Hee-Kyung Lee, Yo-Sung Ho

Kwangju Institute of Science and Technology

{lhkyung, hoyo}@kjist.ac.kr

Abstract: 대부분의 영상색인 기법에서는 영상의 전역 특징값을 이용하지만, 이러한 방법은 영상의 지역적인 변화들을 담아내지 못하기 때문에 만족할 만한 결과를 제공하지 못한다. 본 논문에서는 이러한 문제점을 해결하기 위해 영상의 특징점(salient point)과 영상분할을 이용하여 중요영역(important region)을 추출하는 새로운 영역기반 영상검색 방법을 제안한다. 본 논문에서 제안하는 특징점 추출기법은 기존의 방법에 비해 빠르고 정확한 추출 결과를 보여준다. 선택된 영역에서 추출된 칼라나 질감 정보의 전역 특징값을 이용한 검색 방법의 결과보다 크게 향상됨을 알 수 있었다.

I. 서론

최근 인터넷을 포함한 여러 종류의 통신망을 통해 전달되는 멀티미디어 정보의 양이 급속하게 증가함에 따라 영상검색에 대한 연구가 활발해지고 있다[1]. 영상검색의 초기 연구에서는 사용자가 멀티미디어 데이터를 직접 분석하여 텍스트 형태로 찾고자 하는 내용을 입력하는 방법을 사용하였다. 그러나 이러한 텍스트에 기반한 영상검색 방법은 많은 시간과 노력을 요구하고 데이터를 저장하기가 어려워 실제 구현하기가 어렵다. 따라서 다양한 사용자들의 질의를 충족시켜 줄 수 있도록 칼라, 질감, 모양과 같은 여러 특징들을 포함하는 내용기반 영상검색 시스템이 필요하게 되었다.

대개의 경우, 영상색인은 영상의 전체 영역에서 특징들을 추출한다. 그러나 자연 영상은 몇 개의 다른 특성을 가진 영역으로 구성되어 있기 때문에 전역적 특징 데이터만으로 자연 영상의 특징을 추출하기는 어렵다. QBIC[2], Netra[3], VisualSEEK[4], Blobworld[5]와 같은 내용기반 시스템에서도 영상의 객체나 영역에 기반한 영상검색에 초점을 맞추고 있다.

본 논문에서는 우선 영상의 관심 객체를 포함하는 중요영역을 찾아낸다. 찾아진 영역안의 칼라와 질감 정보를 계산하여 질의 영상과 유사한 영상을 검색하는 시스템을 제안한다. 첫 번째 과정으로 칼라 영상분할 방법을 이용하여 영상의 의미있는 영역을 찾는 영상분할 기법은 다양하게 연구되고 있지만, 의미있는 객체를 정확하게 분할하는 일은 현재의 영상처리 기술로는 많은 한계가 있다. 이러한 영상분할의 한계를 극복하기 위해 영상분할의 결과가 만족스럽지 못한 경우에 대안적인 영역분할 방법을 제안한다. 만약 찾아진 중요영역의 넓이가 경계값(threshold)을 넘지 못하는 경우, 제안한 특징점 추출 알고리즘을 이용하여 관심 객체 주위에 Region of interest (ROI)를 만들어 중요영역을 뽑아낸다.

이렇게 추출된 영역에서 색상정보와 특징점을 이용하여 질감 정보를 계산한 후 검색에 이용한다.

II. 웨이브렛 기반 특징점 추출 기법

1. 기존의 특징점 추출 기법

이전에 제안된 특징점 추출 알고리즘에서는 웨이브렛 변환을 적용했을 때 부모와 자식 사이의 상관관계가 높으며 고주파 부대역에서 큰 계수값을 갖는 영역이 원 영상의 에지 부분에 해당된다는 사실을 이용하여 고주파 부대역 영역에서 전역적인 변화가 많은 부분을 찾고자 한다. 자식들 중에서 웨이브렛 계수가 가장 큰 자식 노드를 선택하여 그때의 웨이브렛 계수를 더해준다. 이러한 과정을 반복적으로 적용해서 주파수가 가장 높은 영역까지 이르렀을 때, 선택된 화소들에 gradient를 적용하여 웨이브렛 계수의 절대값이 가장 큰 화소를 최종적으로 선택한다[6].

2. 제안하는 특징점 추출 기법

특징점을 추출하는 기존의 알고리즘은 선택된 child의 웨이브렛 계수보다 선택되지 않은 다른 child들의 웨이브렛 계수가 더 큰 경우를 고려하지 않았다. 그 결과 배경과 같은 상대적으로 덜 중요한 영역에 특징점이 추출되는 한계성을 나타내고 있다. 본 논문에서는 대표적인 정지영상 부호화 방법인 SPIHT[7]에서 사용한 것과 같은 set partitioning과 linked list를 이용하였다.

본 논문에서 제안한 알고리즘에서는 LIP (List of Insignificant Pixels), LIS (List of Insignificant Set), LSP (List of Significant Pixels)로 정의되는 3개의 리스트를 이용한다. LIP는 중요하지 않다고 판단된 계수들을 포함하며, LIS는 중요하지 않다고 판단된 계수 집합을 포함한다. 또한 LSP는 중요하다고 판단된 계수들을 포함한

다. 전체적인 특징점 추출 과정은 초기화, 정렬, 양자화 스텝 갱신의 3가지 단계로 구성된다.

2.1. 초기화 단계

첫 번째 초기화 단계에서는 웨이브렛 변환된 영상 X 에서 다음과 같이 초기 양자화 단계 n 을 결정한다.

$$n = \left\lceil \log_2 \left(\max_{(i,j) \in X} |C_{i,j}| \right) \right\rceil \quad (1)$$

그리고, 각 리스트의 초기 엔트리(Entry)를 결정한다. 먼저 가장 낮은 대역의 계수들을 LIS에 넣는다. LSP와 LIP는 엔트리를 가지지 않는다.

2.2. 정렬 단계

정렬 단계에서는 LIS에 있는 트리 T의 중요도를 조사한다. 만약 T가 화소이면, 즉, T가 자식을 가지지 않는다면, 트리 T에 속해 있는 모든 계수들의 중요도에 따라 이 트리는 LIS에서 삭제된다. 이때 만약 트리에 포함되어 있는 계수들이 중요하다고 판단되면, 이 계수는 LSP 리스트에 더해진다. 만약 중요하지 않다고 판단되면, 각 계수들은 LIP 리스트에 더해진다.

만약 현재의 트리 T가 화소가 아니면, 다시 말해, T가 자식을 가지면, 트리의 중요도를 판단한다. 만약 트리가 중요하다고 판단되면, 그림 2와 같이, 트리 T를 분할하고, 분할된 작은 트리들이 LIS에 더해진다. 그런 후에, 현재의 트리 T는 LIS에서 삭제된다. 그림 1은 LIS 정렬 과정을 보여준다.

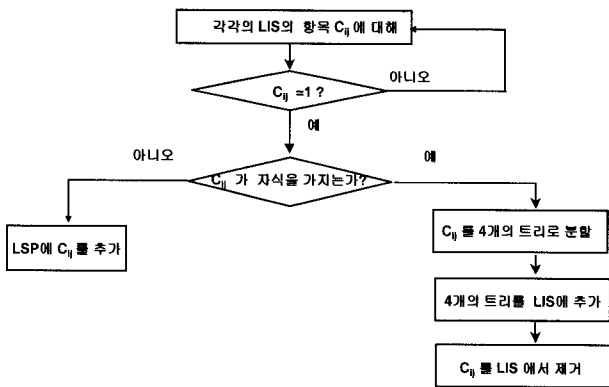


그림 1. LIS의 정렬 과정

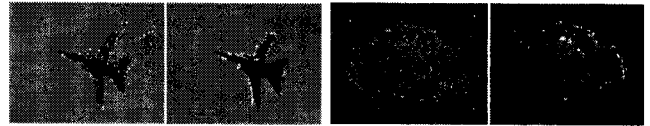
2.3. 양자화 스텝 갱신

이 단계에서는 양자화 단계 n의 값을 1만큼 줄인다. 그리고 검색 과정은 정렬 단계에서부터 다시 시작된다.

그림 2는 영상의 특징점이 100개일 때, Q.Tian[6] 알고리즘의 결과 영상과 본 논문이 제안한 알고리즘의 결과 영상을 비교하여 보여준다. 그림 2(a)와 그림 2(c)는 기존 논문의 결과 영상이고, 그림 2(b)와 그림 2(d)는 본 논문에서 제안한 결과 영상이다.

그림 2에서 알 수 있듯이, 제안한 알고리즘에서는 기존의 방법에 비해 객체 경계 주위의 특징점들이 영상의 덜 중요한 배경 주위에 추출되는 경우는 줄어들고, 효과적으로 추출되는 것을 알 수 있다. 또한 웨이브렛 계수를 비트 플레인(bit plane)으로 변환하여 set partition-

ing과 linked-list를 이용하기 때문에 반복적인 기존의 방법보다 빠르게 결과를 얻을 수 있다.



(a) 기존방법 (b) 제안방법 (c) 기존방법 (d) 제안방법

그림 2. 영상의 특징점 추출 비교

III. 객체의 중요영역 추출 기법

1. 영상분할 기법

영상분할은 객체단위 영상검색과 영상이해를 위한 필수적인 단계이다. 특히, 영역기반 영상검색의 경우, 영상분할의 성능은 영상검색의 정확성을 크게 좌우한다. 영상분할 과정은 크게 다음의 두 단계로 구성된다.

첫 번째 단계에서는 영상의 질감 정보의 양에 근거한 두 종류의 원형 필터를 사용하여 영상을 분할한다. 원형 필터[8]를 사용하면 다른 모양의 필터에 비해 영역의 모양을 병합이 끝난 후에도 자연스럽게 유지할 수 있다. 또한 원형 필터는 일반적인 영역 뿐만 아니라, 영역병합 방법에 의해 병합되지 않는 줄무늬나 얼룩무늬와 같은 의미없는 작은 영역들을 효과적으로 병합할 수 있다. 두 번째 단계에서는 영역병합과 영역표시 방법을 이용하여 작은 영역들을 근처의 유사한 영역으로 병합한다[8].

2. 중요 객체의 영역 추출

일반적으로 영상의 중요한 객체를 영상의 중심에서 가깝게 위치하도록 구도를 잡고 촬영한다. 그림 3(a)의 원영상은 이러한 경우의 몇 가지 예를 보여주고 있다.

그림 3(a)에 보인 것처럼, 호랑이, 새, 차, 말의 경우에 영상의 객체가 모두 영상의 중심에 가깝게 위치한다. 따라서, 이러한 관점에 따라 구성된 영상 데이터베이스에 포함되어 있는 각 영상에서 중요한 객체들이 위치하는 영역이 영상의 바깥 테두리에 접하는 경우는 매우 드물다. 비록 바깥 테두리의 어느 면에 접한다고 하더라도 되도록 그 영역은 극히 작다고 할 수 있다.

이러한 성질을 이용하여 중요도가 높은 객체가 위치하는 영역에 우선권을 할당하면 효과적인 탐색을 수행할 수 있다. 칼라 영상을 분할한 후, 각 영역에 대해서 영역의 테두리 길이와 넓이를 계산한다. 그 뒤 각 영역이 영상의 테두리에 닿는 화소의 개수를 계산한다. 이 값들을 이용하여 식 (2)의 값을 만족하는 영역을 구한다.

$$\frac{BCL_j}{BL_j} \times 100 \leq ThI \quad (2)$$

여기서 BCL_j는 j번째 영역에서 영상의 테두리 부분과 접하고 있는 화소들의 개수, BL_j는 j번째 영역의 테두리 길이를 나타내며, j는 영상에서 각 영역을 나타내는 인덱스이다. 식 (2)에서 경계값(ThI)은 실험적으로 구했으며, 경계값이 15%일 때 중요 객체영역 추출의 결과가 가장 우수하였다. 그렇지만, 이 경계값은 사용하는 영상 데이터베이스의 특성에 따라서 조정되어야 한다.

영상 내의 어떤 영역이 식 (2)를 만족하면, 그 영역은 중요한 영역이 될 수 있는 후보 영역으로 선택된다. 영상 테두리를 접하지 않으면서 넓이가 큰 영역은 일반적으로 중요한 영역으로 간주할 수 있다. 따라서, 선택된 후보 영역이 경계값 (Th_2)을 넘을 경우, 영역의 넓이를 가지고 정렬한 후, 상위 80% (Th_3)에 해당하는 영역을 중요영역으로 결정한다. 선택된 영역이 경계값(Th_2)을 넘지 못할 경우, 즉, 영상분할의 결과가 좋지 않은 경우에는, 특징점을 이용하여 x, y 좌표에 대해 최대, 최소 좌표를 구하여 ROI를 만든다. 그림 3(b)와 그림 3(c)는 호랑이, 꽃, 차, 말 영상들에 대해 영상분할의 결과 영상과 중요영역의 분할 영상을 보여준다.



(a) 원영상 (b) 분할영상 (c) 중요영역 분할영상

그림 3. 중요영역 분할 결과

3. 특징값 추출

색상정보에 기반한 검색기법은 영상검색에 있어서 아주 중요한 방법이다. 색상정보는 잡음이나 다른 배경의 복잡도에 덜 민감하여 영상 내의 물체의 이동, 회전등의 변화에 영향을 적게 받는 통계 정보를 계산하는데 사용될 수 있기 때문에 유용하게 사용하고 있다. 하지만, 공간적인 정보를 이용하지 않고 영상전체에 대한 색상 특징을 구하기 때문에 서로 다른 영상상에도 같은 영상으로 판단하는 오류를 범할 수 있다.

이러한 단점을 해결하기 위해 본 논문에서는 주요한 영역으로 선택된 객체 영역에 한해서 색상정보를 추출하며 웨이블릿 변환을 이용하여 방향 성분의 질감 정보를 뽑음으로써 색상정보에 기초한 방법들의 단점을 극복하였다.

첫째로 중요영역 안에서 칼라 히스토그램을 구하고 Euclidean 거리를 이용하여 색상 정보의 거리를 계산한다. 두 번째로 추출된 특징점을 중심으로 3x3 화소에 해당하는 영역의 웨이블릿 계수를 이용하여 수평, 수직, 대각선 방향의 정보를 얻어낸다. 이렇게 얻어진 값을 이용하여 아래와 같은 방법으로 수평과 수직의 질감 정보의 비, 수평과 대각선의 질감 정보의 비를 각각 구하여 질의영상 Q와 데이터베이스 영상 T사이의 질감정보의 거리를 구한다. 식 (3)은 데이터베이스 영상과 질의 영상간의 질감정보의 거리를 계산하는데 이용된다.

$$d_{Q,T}^r = \left| \frac{Yd_Q}{Xd_Q} - \frac{Yd_T}{Xd_T} \right| + \left| \frac{XYd_Q}{XYd_Q} - \frac{XYd_T}{XYd_T} \right| \quad (3)$$

식 (4)을 이용하여 검색을 위한 최종적인 거리를 계산한다. 여기서 w_1 와 w_2 는 가중치로서, 실험적으로 우수한 검색 결과를 내는 값인 0.75, 0.25를 각각 사용하였다. H_d 는 색상정보 거리, D 는 질감정보 거리를 의미하며,

식 (4)에서 계산된 점수(Score)값을 내림차순으로 정렬하였다.

$$Score = w_1 \cdot H_d + w_2 \cdot D \quad (4)$$

그림 4는 본 논문에서 제안한 영역기반 영상검색 시스템의 구조를 보여주고 있다. 이 시스템은 데이터베이스 안의 영상을 처리하는 부분과 질의 영상을 검색하는 과정으로 나뉘며, 각 부분은 유사한 과정으로 진행된다.

첫 번째 단계에서 제안한 특징점 추출 방법을 이용하여 입력영상에서 특징점을 추출한다. 추출된 특징점들은 질감 정보를 얻어내거나 ROI를 형성하는데 이용된다. 두 번째에서는 영상분할 과정이 수행된다. 영상분할 후, 제안한 기법을 사용하여 중요한 영역들이나 ROI가 결정된다. 선택된 중요영역 안에서 얻어진 칼라 정보와 특징점을 사용하여 계산된 질감정보를 특징값들로 사용한다.

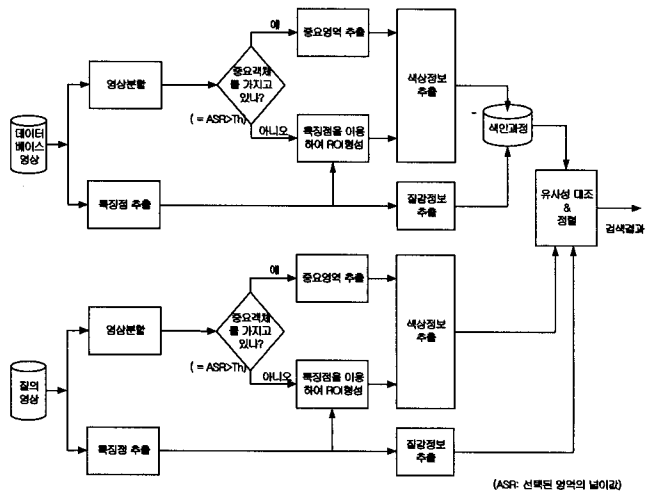


그림 4. 제안한 영역기반 영상검색 시스템의 구조도

IV. 실험결과 및 평가

본 논문에서는 영상검색 결과의 성능을 평가하기 위해 다음과 같은 정확도(Precision)를 사용하였다.

$$정확도 = \frac{검색된 영상 - 잘못 찾아진 영상}{검색된 영상} \quad (5)$$

본 실험에서는 Corel Photo CD (<http://corel.digitriver.com/>)와 Web에서 수집한 자연 영상, 그래픽 영상, 그림 영상 등을 포함하는 3000개의 영상 데이터베이스를 사용하였으며, 8개의 그룹(호랑이, 새, 자동차, 해, 꽃, 사자, 말, 표지판)에 대해서 실험하였다.

그림 5는 새 그룹에 있는 독수리 영상의 검색 결과를 보여준다. 정합의 순위는 위에서 아래로 왼쪽에서 오른쪽으로 갈수록 낮아진다. 제안한 알고리즘은 배경이 서로 다르고 객체의 위치가 바뀌었다 해도 영상의 객체에 해당하는 주요한 영역을 자동적으로 정확하게 찾았다. 찾아진 중요 영역과 특징점을 사용하여 색상과 질감 정보를 뽑아 내어 검색하였기 때문에 영상 전체에 대한 색상 정보나 질감 정보를 이용한 경우보다 훨씬 좋은 검색 결과를 얻을 수 있었다.



그림 5. 제안한 알고리즘의 검색결과

그림 6은 각 그룹에 본 논문에서 제안된 알고리즘을 적용한 경우, 전역적인 색상을 이용한 경우, 전역적인 질감 정보를 이용한 경우, 기존의 특징점 추출 방법을 적용하여 지역적 칼라 정보를 사용한 Q'tian [6]의 알고리즘, 각각에 대해서 계산된 정확도를 보여주고 있다.

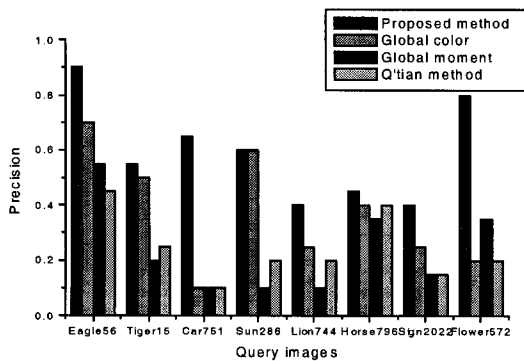


그림 6. 각 그룹에 대한 정확도

사자와 말과 같이 배경이 복잡하지 않은 영상의 경우, 제안한 알고리즘의 정확도는 기존의 전역적인 색상과 질감 검색 방법의 검색 결과에 있어서 큰 차이를 나타내지 않는다. 그러나 배경 영상이 단조롭지 않은 영상의 경우, 기존의 방법에 비해 배경의 영향에 덜 민감한 영역기반 검색 방법의 검색 결과가 뛰어난 것을 알 수 있다. 특히 자동차, 꽃과 같이 복잡한 배경 영상으로 이루어져 있는 경우에는 영상 전체 영역에 걸쳐 색상 혹은 질감 정보를 뽑아 검색했을 때 검색의 성능이 좋지 못했다. 그러나 본 논문에서 제안한 알고리즘을 사용한 경우 복잡한 배경에도 중요한 객체를 자동적으로 정확히 추출해 내어 그 객체 안의 색상과 질감 정보를 검색에 이용하였기 때문에 우수한 검색 결과를 나타냈다.

V. 결론

본 논문에서는 영상의 특징점과 영상분할 결과를 이용하는 영역기반의 영상검색 시스템을 제안하였다. 본 논문에서 제안된 특징점 추출 알고리즘은 set partitioning과 linked-list 방법을 사용하여 기존의 방법보다 빠르고 정확하게 특징점을 추출해 낸다. 중요영역은 영상분할 결과와 특징점을 이용하여 찾아낸다. 선택된 영역의 넓이가 경계값을 넘는 경우에 선택된 영역을 객체 주변의 중요 영역으로 인식하고 그 영역에서 칼라 정보

를 추출하였다. 그렇지 않은 경우, 특징점을 이용하여 객체 주위에 ROI를 형성하여 그 영역에서 칼라 정보를 추출하였다. 한편, 웨이블릿 변환을 이용하여 특징점의 주위 영역에 대해 방향 질감 정보를 추출하여 검색에 이용하였다. 중요한 영역에 대한 색상, 질감 정보를 사용하였기 때문에 전역적 칼라 또는 질감을 이용한 방법보다 검색 결과의 성능이 상당히 향상되었다.

감사의 글

본 연구는 광주과학기술원(K-JIST) 초고속 광네트워크 연구센터(UFON)를 통한 한국과학재단 우수연구센터(ERC)와 교육부 두뇌한국21(BK21) 정보기술사업단의 지원에 의한 것입니다.

참고문헌

- [1] S. Chang, J. Smith, M. Beigi and A. Benitez, "Visual Information Retrieval from Large Distributed Online Repositories," *Communications of ACM*, vol. 12, pp. 12-20, 1997.
- [2] M. Flicker, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC system," *IEEE Computer special issue on Content Based Picture Retrieval System*, vol.28, no. 9, pp. 23-32, 1995.
- [3] W. Y. Ma and B. S. Manjunath, "Netra: A Toolbox for Navigating Large Image Database," *IEEE Conf. on Image Processing*, vol. 1, pp. 568-571, 1997.
- [4] J. R. Smith and S. F. Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System," *ACM Multimedia*, pp. 87-98, Nov. 1996.
- [5] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, "Blobworld: A system for Region-based Image Indexing and Retrieval," *Proc. Int. Conf. Visual Inf. Sys*, vol.3, pp. 509-516, June 1999.
- [6] Q. Tian, N. Sebe, M. S. Lew, E. Loupias, and T. S. Huang, "Image Retrieval using Wavelet-based Salient Points," *Journal of Electronic Imaging, Special Issue on Storage, Processing and Retrieval of Digital Media*, vol.10, no.4, pp. 3-11, Oct. 2001.
- [7] A. Said and W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Trans. on Circuit and Systems for Video Tech.*, vol. 6, no. 3, pp. 243-250, June 1996.
- [8] B. C. Ko, H. S. Lee and H. R. Byun, "Region-based Image Retrieval System using Efficient Feature Description," *International Conference on Pattern Recognition*, vol.4, pp. 284-286, Sept. 2000.