

이미지 데이터베이스에서 내용기반 검색을 위한 블록지향 이미지 분리구조의 설계

최현락*, 이병욱*

*경원대학교 전자계산학과

e-mail:lonely98@web.kyungwon.ac.kr

Design of a Block-Oriented Image Decomposition Structure for Content-Based Retrieval in Image Database Systems

Hyun-Lark Choi*, Byung-Wook Lee*

*Dept. of Computer Science, Kyungwon University

요약

이미지 데이터베이스에서의 내용기반 검색 요구가 급증하고 있다. 기존의 이미지 특징벡터는 상세한 이미지를 표현하는데 부정확하므로 시각적 질의를 위한 만족할만한 결과를 제공하지 않는다.

본 연구에서는 내용기반 이미지 데이터 검색을 지원하는 블록지향 이미지 분리 구조를 제시하고, 노나 트리(nona-tree) 데이터 구조를 개선한 블록 지향 이미지 분리 구조를 통해 검색 속도를 개선한다. 이미지 특징을 표현하기 위해서 웨이블릿 변환을 사용하여 특징 벡터를 구성하고 제안한 데이터 모델을 통해 데이터베이스에서 분리된 세그먼트들을 비교하였다. 실험 계산을 통하여 노나 트리 분리 구조와 제안한 이미지 분리 구조를 비교하였고, 계산 결과 재현율과 정확도는 유사하지만, 검색 속도가 기존방법보다 우수하였다.

1. 서론

이미지 데이터베이스에서의 검색은 이미지에서 모양, 색상, 질감, 구도, 다양한 객체의 위치 같은 내용의 특징들을 이용해야 한다. 그러나 이러한 대량의 이미지 정보를 데이터베이스에 저장하고 효율적으로 질의, 검색할 수 있는 실용화된 도구가 있지만 성능 면에서 아직 미흡하다.

기존의 주석기반 검색과 달리 내용 기반 이미지 검색은 질의 이미지의 특징 벡터들과 데이터베이스 이미지들을 탐색, 비교함으로써 제공되는데, 데이터베이스 이미지의 특징 벡터는 그 부 이미지들을 정확하게 표현하지 않는다. 따라서 시각적인 질의에 대해 만족할 만한 결과를 제공하지 않는다. 그러므로 이미지 분리는 이미지 데이터베이스를 탐색하는데 특징 기반 기법의 구현에 필수 불가결하다. 대용량 이미지 데이터베이스에서 이미지의 수동 분리를 피하기 위하여 이미지의 정사각형 트리(quad-tree)에 기반한 블록 지향 접근이 사용되고 있다. 정사각형 트리 분리와 웨이블릿 변환을 사용하여 텍스처 접근에 의한 효과적인

질의가 가능하다.

본 논문에서는 내용기반 이미지 검색에서 사용되는 효과적인 블록 지향 이미지 분리 구조를 제시한다. 이 구조는 정사각형 트리 분리의 확장이다. 그리고 내용기반 이미지 검색을 위한 데이터 모델의 응용을 위해 정사각형 트리(quad-tree)의 변형인 노나 트리(nona-tree)의 데이터 구조를 개선한 블록 지향 이미지 분리 구조를 제안하고, 이미지 검색의 효율을 개선한다. 이미지 특징을 추출하기 위해서 Daubechies 웨이블릿 변환을 사용하고[9], 제안한 데이터 모델을 통해 특징벡터와 데이터베이스에서 분리된 세그먼트들을 비교한다.

성능평가는 노나 트리 분리 구조와 제안한 이미지 분리 구조의 수치 분석을 통해 서로 비교하였고, 이를 통해 제안한 분리가 효율적인 이미지 검색을 위한 새로운 분리 구조임을 증명한다.

2절에서는 노나 트리라고 명명한 블록 지향 데이터 구조와, 계산적 복잡도를 소개한다. 3절에서는 노나 트리 이미지 표현의 응용을 제안하고, 노나 트리

분리 구조와 제안한 이미지 분리 구조를 비교, 평가한다. 4절에서는 결론과 향후 연구방향을 제시한다.

2. 이미지 데이터 모델링

이 절에서는 이미지 내용을 표현하는데 사용되는 정사각형 트리와 노나 트리 이미지 분리 구조를 기술하고, 웨이블릿 변환방법을 다룬다.

본 연구의 기반이 되는 정사각형 트리는 이미지 검색을 위해 데이터를 빨리 접근할 수 있는 계층적 이미지 분리구조이다[8]. 이 트리는 이미지를 반복적으로 분리하는 원리에 기반을 둔다. 이미지를 4개의 동등한 크기의 정사각형으로 나누고, 전체 이미지를 표현하는 루트로부터 반복적으로 분리하면서 자식노드를 구성한다.

2.1 노나 트리

노나 트리는 정사각형 트리에서 제안한 반복적 이미지 분리의 변형에 기반한 계층적 이미지 분리 구조이다[8]. 이것은 이미지를 9개의 동등한 크기의 부 이미지로 분리한다. 이미지는 4개의 동등한 크기의 4분면(숫자 1, 2, 3, 4)으로 나뉘고, 동등한 크기의 이미지 중앙 영역(숫자 5)과 중앙 영역의 위, 아래, 왼쪽, 오른쪽 절반을 각각 포함하는 동등한 크기의 4개의 부 이미지(숫자 6, 7, 8, 9)로 분리된다.

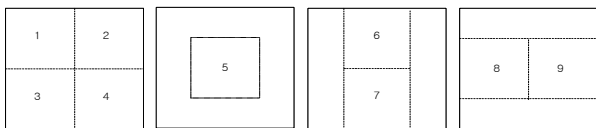


그림 2.1 노나 트리 분리에서 9개의 부 이미지

그림 2.1은 분리된 이미지 세그먼트에서 부 이미지들의 위치를 나타낸다.

노나 트리의 분리 공정은 전체 이미지를 나타내는 루트와 분리된 세그먼트를 나타내는 자식과 함께 반복적으로 표현될 수 있다. 각 내부 노드는 최대 9개의 자식을 가지고 있다. 정사각형 트리 분리[2]의 전략이 노나 트리에서 부 이미지 1, 2, 3, 4의 분리를 위해 사용되고, 분리는 단지 이질적인 텍스처를 가진 세그먼트들에만 적용된다. 4개의 자식노드가 부모노드로부터 생성되기 전에 합병하기 위한 조건을 검사한다. 거리 임계(distance threshold)는 텍스처 특징 추출에 기반한 각 자식노드에 대해 계산된다. 특징 공간에서의 거

리는 부모노드로부터 각각의 자식노드 까지 측정된다. 만약 4개의 자식노드 모두에 대한 거리가 자식노드의 임계값 내에 해당한다면 부모 노드는 단일 텍스처로 간주되고, 더 이상의 분리는 없게 된다. 그렇지 않은 경우에는 자식노드의 배치가 수행될 것이다. 두 개의 이웃한 자식노드가 모두 임계값 내에 있으면 자식노드들은 단일 자식노드로 합병되고, 또한 생성된 1, 2, 3, 4의 부 이미지들의 생성여부에 따라 5, 6, 7, 8, 9의 부 이미지 생성이 결정된다.

2.1.1 부 이미지 적용범위

노나 트리 분리는 분리된 이미지의 일부가 질의 이미지와 매칭 된다면, 적어도 질의 이미지의 $\frac{9}{16}$ 를 포함하는 세그먼트가 존재한다. 그림 2.2는 이 상황을 설명한다. 질의 이미지의 크기가 첫 번째 단계 분리의 이미지와 일치한다고 가정할 때 그림 2.2(a)에서 질의 이미지의 그림자 영역은 $(\frac{3}{4}r)^2$, 즉 $\frac{9}{16}r^2$ 이다. 일반적으로, 그림 2.2(b)에서처럼 질의 이미지의 덮여진 영역 A는 다음의 경우를 따른다.

● $r_1, r_2 \leq \frac{1}{4}r$: 세그먼트 5에 덮여진 영역 :

$$A = (r - r_1) \times (r - r_2) \geq \frac{9}{16}r^2$$

● $r_1 \leq \frac{1}{4}r$ 이고, $r_2 \geq \frac{1}{4}r$: 세그먼트 6에 덮여진 영역 :

$$A = (r - r_1) \times (\frac{1}{2}r + r_2) \geq \frac{9}{16}r^2$$

● $r_1 \geq \frac{1}{4}r$ 이고, $r_2 \leq \frac{1}{4}r$: 세그먼트 8에 덮여진 r 영역 :

$$A = (r - r_2) \times (\frac{1}{2}r + r_1) \geq \frac{9}{16}r^2$$

● $r_1, r_2 \geq \frac{1}{4}r$: 세그먼트 1에 덮여진 영역 :

$$A = (\frac{1}{2}r + r_1) \times (\frac{1}{2}r + r_2) \geq \frac{9}{16}r^2$$

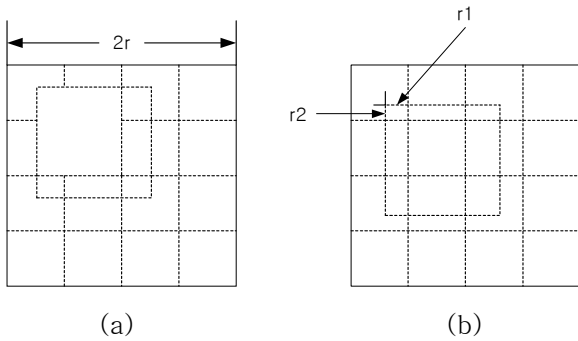


그림 2.2 세그먼트에 의해 덮여진 질의 이미지의 영역

2.2 웨이블릿 변환

웨이블릿 기반 이미지 압축 구조는 폭넓고 다양하게 학문적으로 발전되어 왔다[4,5]. 압축은 이미지 데이터를 표현하기 위하여 웨이블릿 변환을 적용하고, 변환결과 계수를 양자화하고, 양자화된 값을 기호화함으로써 수행된다. 웨이블릿에 의해 변환된 데이터는 내용기반 검색에 이용될 수 있다.

웨이블릿 변환은 특별한 임의 시간에서 신호의 빈도량을 제공하는 신호표현의 형태이다. 처음에는 유사한 속성영역이 분리될 필요가 있다. 이것은 이미지 분리에 의해 성취된다. 압축된 이미지 형식으로부터 속성 특징들을 쫓는 이점을 찾기 위하여 웨이블릿 변환은 대역분리를 이용할 수 있다[5]. 각 대역은 속성정보를 균일하게 하는 정제과정 후에 얻는다. 대역에서 에너지 평가처럼 전형적인 방법은 대역으로부터 특징을 추출하는데 유용하다.

특별한 웨이블릿 변환을 사용하여 노나 트리에서 모든 이미지 조각들에 대한 특징이 추출된다. 대역분리의 3단계가 사용되고[5], 각 대역은 속성정보를 갖게 하는 정제과정 후에 얻는다. 특징 추출은 Daubechies 웨이블릿 변환을 사용하여 수행한다[9]. 각 3단계 변환 정보는 10개의 대역을 생성한다. 특징들은 대역에서 에너지 측정에 의해 추출된다. 두 에너지 특징들, 평균과 분산은 각 대역에서 계산된다. 그러므로 각 특징 벡터는 20개의 원소를 가지고 있다. 이 특징 추출 절차는 각 이미지의 트리 표현에서 모든 이미지 조각들에 대해 적용된다. 노나 트리에서 각 노드는 그러므로 이미지의 조각과 그것의 특징벡터를 표현한다.

3. 분리 구조의 설계

이미지 데이터베이스 시스템에서 내용 기반 이미

지 검색을 위해서 제안한 이미지 데이터 모델을 조사한다. 노나 트리에서의 탐색 절차는 질의 이미지와 데이터베이스 이미지들 사이의 특징벡터들을 비교하기 위해 고안되었다. 웨이블릿 변환은 이미지 조각들의 특징벡터를 생성하는데 사용된다.

3.1 노나 트리 구조의 설계

(1) 탐색기법의 설계

데이터베이스 이미지들이 노나 트리에 의해 표현된다고 가정한다. 특징 생성절차는 노나 트리에서 각 노드에 적용된다. 그러므로, 각 노드는 원래 픽셀 데이터보다 오히려 이미지 조각들의 특징 벡터를 포함한다.

모든 이미지들이 노나 트리에서 그들의 특징 벡터에 의해 표현되는 이미지 데이터베이스를 고려하자. 질의 이미지를 Q라 할 때 질의 이미지의 특징벡터와 데이터베이스 이미지들과 일치하는 절차를 생각해 보자. 일치하는 먼저 데이터베이스 이미지 내용과 특징 벡터 사이에서 이루어지고, 검색은 그 다음에 이 특징벡터들과 질의 이미지 Q의 특징벡터 사이의 관계에 기반해서 수행된다. 우리는 이 이미지들을 질의 이미지를 포함하거나 유사한 대용량 이미지 데이터베이스 안에 위치시킬 수 있다.

주어진 질의 이미지 Q와 데이터베이스 내의 어떤 이미지 M 사이의 관계는 Q의 특징벡터와 M의 다른 부분을 비교함으로써 결정된다. 이것은 루트로부터 잎노드까지 노나 트리를 통한 탐색에 의해 가능하다. 질의 이미지의 특징벡터와 데이터베이스 이미지 조각들의 특징벡터 사이의 거리를 비교하기 위하여 평균의 평방미터를 사용하였다. 이것은 주어진 특징 벡터 $v_1(a_1, \dots, a_n)$ 과 $v_2(b_1, \dots, b_n)$ 사이의 거리를 다음과 같이 계산한다.

$$dist(v_1, v_2) = \sqrt{\frac{\sum_{i=1}^n (a_i - b_i)^2}{n}}$$

주어진 임계값 t에 대하여 질의 이미지와 데이터베이스 이미지의 세그먼트 사이의 거리가 t보다 작다면, 이 데이터베이스 이미지는 매칭된 이미지로 선택된다.

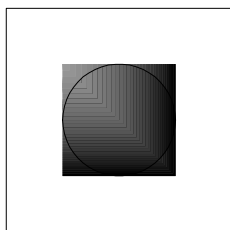
각각의 노나 트리의 탐색은 더 많은 시간의 소비 동안에 수행된다. 이것은 분리된 이미지 세그먼트의 수가 증가함에 따라 발생된다. 그러므로 주어진 질의

이미지에 대한 효율적인 검색을 위해서는 모든 이미지 세그먼트의 생성된 특징벡터들을 포함하는 특징 공간에서 특징 벡터들을 분류하기 위해 색인 기법을 적용해야 한다.

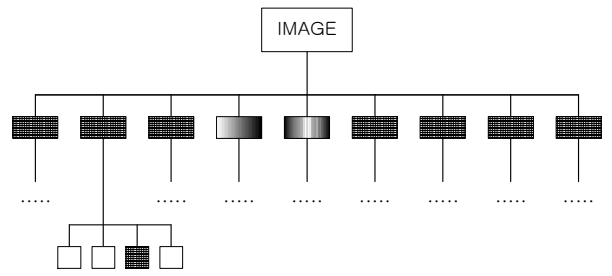
(2) 자료구조의 설계

노나 트리에서 생성된 이미지 세그먼트의 수는 노나 트리 구조의 탐색 시 더 많은 시간을 소비한다. 그러나, 이 문제는 이미지 세그먼트로부터 생성된 특징 벡터를 포함하는 색인 기법에 의해 해결될 수 있다. 이미지에서 세그먼트들의 위치에 기반해서 카테고리들을 분류한다. 즉, 데이터베이스 이미지들은 각각의 9개의 다른 위치 중에서 하나의 특별한 위치를 표현하는 9개의 부 공간으로 구성된다. 데이터베이스 이미지 분리의 상위 단계 분리에서 세그먼트들의 위치에 대한 이 분류는 데이터베이스 이미지 내에 있는 부 이미지의 적절한 위치에 기반을 둔 질의들에 대한 효율적인 검색이 가능하다. 그 다음에 두 번째 단계부터는 정사각형 트리(quad-tree)의 분리구조를 따른다. 즉, 상위 단계에서는 노나 트리 분리구조를 적용하고, 반복되는 부 이미지 세그먼트들은 4개의 부 공간으로 구성된다. 이것은 노나 트리의 탐색시간을 감소시킬 수 있다. 검색의 재현율(recall)과 정확도(precision)를 향상시키기 위해서 노나 트리 구조를 첫 번째 분리 단계에 적용하고, 세그먼트 수의 감소를 위해서 두 번째 분리 단계부터는 정사각형 트리 구조를 적용한다.

그림 3.1은 부 이미지가 점진적으로 순서화 되는 노나 트리 데이터 구조의 개선된 이미지의 표현을 나타낸다. 예를 들면, 첫 번째 분리 단계에서 주어진 이미지는 9개의 부 이미지로 분리된다. 모든 부 이미지들이 이질적이므로 계속적인 분리가 필요하고, 두 번째 분리 단계부터는 정사각형 트리 분리 구조를 적용하여 부 이미지가 4개의 부 이미지로 분리된다. 다른 부 이미지들도 유사하게 분리된다.



(a)



(b)

그림 3.1 (a) 이미지; (b) 제안한 노나 트리

이 방법은 특히 첫 번째 이미지 분리 후 이미지 세그먼트내의 이미지 포함영역이 적을 경우에 매우 효율적이다.

3.2 블록 지향 트리 구조의 비교

표 1과 같이 잎 블록(leaf block)의 크기는 최소 16 X 16으로 제한한다. 그러므로, 256 X 256 픽셀의 데이터베이스 이미지에 대한 트리의 깊이는 4단계까지 진행된다.

표 1 트리의 깊이와 이미지 크기

단계	0	1	2	3	4
이미지 크기	256*256	128*128	64*64	32*32	16*16

정사각형 트리 구조에서 세그먼트의 분리는 표 2와 같이 각 레벨 i 에서 4^i 개의 최대 세그먼트를 포함한다.

표 2 정사각형-트리에서 최대 세그먼트의 수

단계	i	0	1	2	3	4
정사각형 트리의 세그먼트 수	4^i	1	4	16	64	256

노나 트리에서의 최대 세그먼트의 수는 표 3에 있다. 레벨 i 에서의 세그먼트 수를 S_i 라고 했을 때, $S_0 = 1, S_1 = 9$ 이고, 레벨 $i+1$ 에서는 9^i 대신에 $S_{i+1} = (2^{i+1}-1)^2$ 이다.

반면에 제안한 노나 트리 모델에서의 최대 세그먼트 수는 노나 트리에 비해 하위 단계로 갈수록 줄어든다. 제안한 모델은 정사각형 트리와 노나 트리 구조의 장점을 서로 보완하여 적용함으로써 검색의 정확성을 유지하면서 검색시간을 단축하였다.

표 3 최대 세그먼트 수의 비교

단계	i	0	1	2	3	4
노나 트리의 세그먼트 수	$(2^{i+1}-1)^2$	1	9	49	225	961
제안한 노나 트리의 세그먼트 수	$4^{(i-1)} * 9$ 단, $i \geq 1$	1	9	36	144	576
세그먼트 수 감소 비율	0 %	0%	0%	26%	36%	40%

표 3과 같이 제안한 노나 트리의 최대 세그먼트 수는 첫 번째 단계에서는 노나 트리의 최대 세그먼트 수와 차이가 없지만 점차 단계가 높아질수록 노나 트리의 세그먼트 수보다 감소한다.

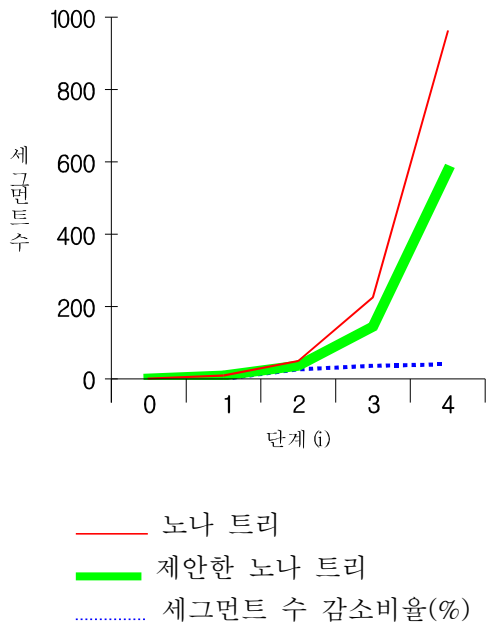


그림 3.2 노나 트리와 제안한 노나 트리의 세그먼트 수의 비교

정확성을 위해 첫 번째 분리단계에서는 노나 트리 구조를 사용하였고, 두 번째 단계 이후에는 정사각형 트리 구조를 적용하여 0% ~ 40%의 검색속도를 향상한다.

4. 결론

본 논문에서는 내용기반 이미지 데이터의 효과적인

이고, 효율적인 검색을 지원하는 이미지 데이터 표현과 검색에 관한 방법을 제시하였다. 제안한 분리 구조는 정사각형 트리 분리와 노나 트리 보다 더 좋은 블록지향 트리 분리 접근법을 제공한다. 질의 이미지가 데이터베이스 이미지 보다 같거나 작다는 가정아래 노나 트리 분리 구조와 제안한 이미지 분리 구조의 단계별 세그먼트 수와 부 이미지의 적용범위를 서로 비교 분석하였다. 노나 트리의 단점인 많은 세그먼트 수를 개선하여 검색 효율을 수치 분석함으로써 제안한 방법이 효율적인 이미지 검색을 위한 새로운 분리 구조를 제공함을 증명하였다. 제안한 분리 구조는 특히 분리된 이미지 세그먼트 내에 이미지가 차지하는 비율이 적을 경우 효율적이다.

향후 연구과제로 블록 이미지 세그먼트들의 공간 탐색을 위한 색인에 관한 연구가 필요하다.

참고 문헌

- [1] Edward Remias, Gholamhosein Sheikholeslami, Aidong Zhang and Tanveer Fathima Syeda-Mahmood, "Supporting Content-Based Retrieval in Large Image Database Systems," Multimedia Tools & Applications, V.4 153-170(1997).
- [2] John R. Smith and Shih-Fu Chang, "Quad-Tree Segmentation for Texture-Based Image Query," in Proc. of ACM Multimedia 94, San Francisco, California, October 1994, pp. 279-286.
- [3] J. Guo, A. Zhang, E. Remias and G. Sheikholeslami, "Image decomposition and representation in large image database systems", The Journal of Visual Communication and Image Representation, 8(2); 167-181, June 1997.
- [4] 김영로, 홍원기, 고성제, "웨이블릿 변환 계수의 비트 평면을 이용한 영상 부호화", 한국통신 학회 논문지 97-4 Vol. 22 No.4.
- [5] S. Mallai, "A theory for multiresolution signal decomposition: the wavelet representation." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, pp. 674-693, 1989.
- [6] Dantong Yu, Lei Zhu, Aidong Zhang, Ling Bian, "A Multi-Resolution Geographical Data Retrieval System", Proceedings of the 11th International Conference on Scientific and Statistical Database Management, pp. 277-277, 1999.

- [7] K.-I. Lin, H.V.Jagadish, and C. Faloutsos, "The TV-Tree:an Index Structure for High-Dimentional Data," The VLDB Journal, Vol, 3, pp. 517-542, 1994.
- [8] S.K. Chang, Principles of Pictorial Information Systems Design, Prentice Hall, Englewood Cliffs, NJ 1989.
- [9] G. Strang and Truong Nguyen, Wavelet and Filter Banks, Wellesley-Cambridge Press, 1996.