

내용 기반 이미지 검색에서 사용되는 시각 기술자에 대한 비교 분석*

최효진^o 낭종호
서강대학교 컴퓨터공학과
monsta0622@naver.com

A Comparison of Visual Descriptors for Content-based Image Retrieval

HyoJin Choi^o Jongho Nang
Department of Computer Science and Engineering, Sogang University

요 약

콘텐츠 기반의 이미지 검색에서는 여러 특징들을 이용하여 데이터베이스 내에서 질의 이미지와 유사하다고 판단되는 이미지들을 결과로 보여준다. 다양한 특징들을 이용할 경우 결과 이미지에 대한 오차가 줄어들 것이라는 장점이 있지만, 데이터, 비용 혹은 시간 면에서 소모가 크다고 볼 수 있다. 단일 특징을 이용할 경우 다수의 특징을 이용할 때 발생하는 단점이 보완된다는 이점이 있어, 본 논문에서는 256 RGB, 양자화된 HSV, SIFT를 이용하여 단일 특징으로 유사 이미지 출력에 대한 결과를 확인해보았는데, 단일 특징의 경우 해당 특징에 대해서만 유사도를 판단하여 다른 특징에는 전혀 영향을 받지 않거나 정확한 유사도를 추출하지 못하는 경우가 발생하였다. 또한 실험을 통하여 SIFT 기법은 색상에 전혀 영향을 받지 않음을 알 수 있었는데, 이러한 특징에 착안하여 SIFT와 같이 다른 특징에는 영향을 받지 않는 여러 특징들을 조합하면 정확하고 효율적인 이미지 검색을 할 수 있을 것이 예상할 수 있다.

1. 서 론

인터넷이 현대인의 삶의 주요한 부분을 차지하면서, 이미지 검색 또한 인터넷 정보 사회의 빠질 수 없는 일 부분으로서 사회에 통용화 되어있다. 이미지 검색은 더 새롭고 편리한 방식으로 변화하였는데, 그 중 대표적인 것이 콘텐츠 기반 이미지 검색(Content-Based Image Retrieval, CBIR)이다. 콘텐츠 기반 이미지 검색은 이미지 자체의 특성을 이용해 질의로 준 이미지와 유사한 이미지를 찾는 것으로, 색상, 질감, 모양 등의 특징을 이용하여 이미지를 검색한다[1]. 이 기법은 기존에 키워드를 입력하여 결과를 얻는 텍스트 검색을 기반으로 한 이미지 검색에서 제공할 수 없는 보다 더 직접적인, 혹은 사람의 시각에 맞춘 기법으로서 이미지 검색을 더욱 편리하게 해주며 다양한 사이트에서 해당 서비스를 제공하고 있다. 일반적으로 효과적인 이미지 검색을 위해서는 위에 제시된 특징들을 비롯한 여러 가지의 특징들을 결합하여 유사한 이미지를 찾는데[2], 본 논문에서는 콘텐츠 기반의 이미지 검색에서 사용되는 특징들 중에서도 하나의 특징을 선택하여 사람들이 유사하다고 느낄 수 있는 결과 이미지를 도출해 낼 수 있는지 알아보하고자 하였으며, 단일 특징의 한계점을 보이고자 한다.

2. 검색 알고리즘

*본 연구는 미래창조과학부 및 정보통신기술진흥센터의 SW중심대학지원사업의 연구결과로 수행되었음(R7718-16-1004)

2.1 256 RGB Color 기반 검색

256 Color 기반 검색은[3] 데이터베이스 내에 있는 이미지를 기반으로 RGB 색상들의 정보들을 추출한 후 한 색상의 정보에 대해 256개의 값을 가지도록 양자화 한다. RGB의 값을 그대로 이용할 경우 실제 눈에 보이는 유사도와는 달리 이웃한, 즉 비교적 비슷한 색상 정보를 가졌더라도 그 차이가 클 경우 이웃하지 않은 색상 정보와의 유사도보다도 적게 나오는 경우가 발생한다. 이 경우를 방지하기 위해 양자화의 특징을 이용하여 256개로 색상의 종류를 줄인다.

$$dist = \sqrt{\sum_{i=0}^7 \sum_{j=0}^7 \sum_{k=0}^3 (RGB_{query}[i][j][k] - RGB_{db}[i][j][k])^2} \quad \dots(1)$$

유사도 측정은 식 (1)을 이용하여 데이터베이스 내의 모든 이미지에 대해 그 값을 측정하고 dist 값이 작은 순으로 정렬하여 결과를 도출해 낸다. $RGB[i][j][k]$ 는 이미지의 각 픽셀에서 추출된 RGB를 256 색상으로 양자화 한 값을 나타내는 i, j, k 를 이용하여 히스토그램 배열을 생성한 것이다.

2.2 HSV Color 기반 검색

HSV 색상 기반 검색[4] 또한 RGB 모델과 마찬가지로 그대로 정보를 저장할 경우 엄청난 데이터가 축적되며 색상 히스토그램 데이터 간의 거리를 사람의 시각과는 다르게 도출해내는 경우가 발생하기 때문에 양자화 시켜준다. 얼마나 양자화를 하느냐에 따라 이미지의 유사 순서

가 바뀌어버릴 수도 있기 때문에 RGB 색상 모델에서 검증된 256 색상의 수와 가장 유사하도록 $9 \times 8 \times 8(576)$ 개의 색상이 존재하도록 양자화 한다.

$$dist = \sqrt{\sum_{i=0}^8 \sum_{j=0}^7 \sum_{k=0}^7 (HSV_{query}[i][j][k] - HSV_{db}[i][j][k])^2} \quad \dots(2)$$

식 (2)와 같이 배열을 탐색하여 각각의 이미지 사이의 거리 측정이 가능하다. 여기서 $HSV[i][j][k]$ 는 HSV 값에 해당하는 히스토그램 배열을 나타낸다.

2.3 SIFT (Scale Invariant Feature Transform) 검색

SIFT 기반 검색[5]은 이미지의 특징 점(keypoint)들을 찾는 것으로, 질의 이미지가 주어질 때 데이터베이스 내에 있는 모든 이미지들과의 특징 점 비교를 통해 각각의 특징 점 개수를 구한 후 특징 점 개수가 많을수록 이미지가 유사하다고 판단한다. 기존 SIFT 방법에서는 질의 이미지와 데이터베이스 내의 이미지 특징 점의 가장 가까운 거리 d_1 과 두 번째로 가까운 거리 d_2 를 구하여 $d_1/d_2 < 0.49$ 일 때, 즉 d_1 이 거의 유일하게 존재한다고 보고 특징 점들을 추출한다.

$$dist = \sum_{i=0}^{127} (descr_{Query}[i] - descr_{DB}[i])^2 \quad \dots(3)$$

d_1 과 d_2 는 식 (3)과 같이 해당 기술자에 대한 모든 벡터 값을 비교하여 구할 수 있으며, 식 (3)에서의 $descr$ 가 128 차원의 벡터 정보를 담고 있는 배열이다.

3. 구성 및 설계

3.1 이미지 데이터베이스 구성

이미지 데이터베이스는 결과에 대한 명확한 분석을 위해 물체, 색상을 다양한 범위 내에서 생성하였으며 총 80장의 이미지로 구성하였다. 물체는 [그림 1]과 같이 케익, 꽃, 건물 등으로 나뉘며, 색상은 가장 큰 히스토그램을 가지고 있는 색상을 중심으로 빨간색, 검정색, 회색, 노란색, 흰색 등으로 구성되어 있다.

3.2 색상 기반 모델 결과 추출

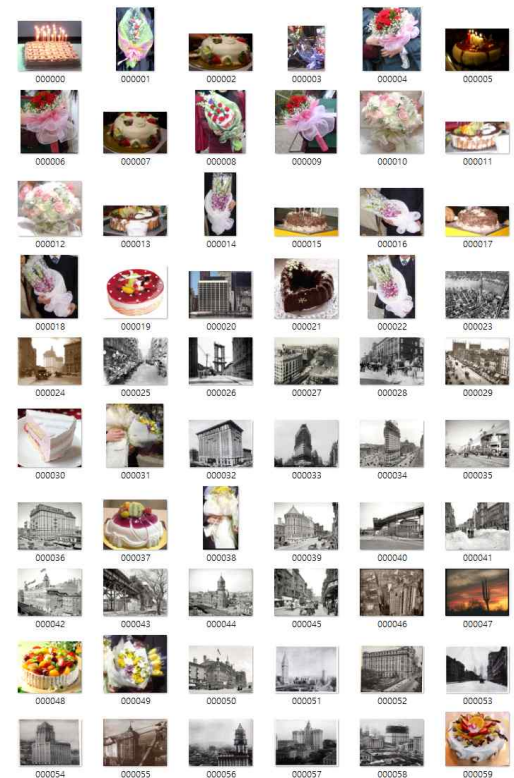
[3]과 양자화 된 HSV에서는 이미지의 유사도를 판단하기 위해서 각 이미지들의 히스토그램 데이터를 저장하는 공간인 인덱스 파일에서 비교하려는 이미지를 찾아 그 차이 값을 L2 거리를 이용하여 구한다. 해당 과정은 다음과 같다.

1단계 : 질의 이미지 입력

2단계 : 데이터베이스 내의 이미지와 L2 Distance 계산

3단계 : 해당 결과 값을 결과 배열에 정렬

위의 2, 3단계를 모든 이미지에 대해 반복하면 최종적으



[그림 1] 이미지 데이터베이스

로 질의 이미지와 유사도가 높은 이미지 10개가 결과 이미지에 출력되는 구조이다.

3.3 SIFT 결과 추출

[5]에서는 동일한 이미지를 찾는 것이 목적이거나, 본 논문에서는 유사도가 높은 이미지를 찾는 것이 목적이므로 유일성이 아닌 유사성에 초점을 두고 $\frac{d_1}{d_2}$ 의 임계값을 높여 특징 점 개수를 추출하여 결과 이미지 배열에 정렬한다. 임계값은 0.49에서부터 증가시켜가며 특징 점의 개수들의 차이가 100을 넘지 않을 때의 분포에 이를 때를 찾아 해당 지점인 0.9를 임계값으로 택하였다.

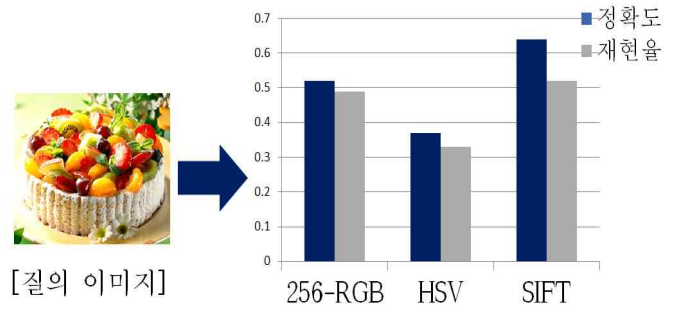
4. 구현 및 결과 분석

결과 분석을 위해 [그림 2]와 같이 질의 이미지에 대한 정답 이미지를 사람의 시각에 맞추어 지정한 후, [그림 3]의 결과와 같이 각각의 특징에 따른 결과에서 정답에 해당하는 이미지를 확인하고 이를 이용하여 정확도(Precision)와 재현율(Recall)을 구할 수 있다.

[그림 4]의 측정값은 각각 케익, 꽃, 오토바이, 자동차, 건물을 물체로 가지는 이미지들을 질의 이미지로 가져와 나온 정확도와 재현율을 평균하여 나온 값이다. HSV는 실제 사람의 시각과 유사한 시스템을 가져 유사도가 높



[정답 이미지]
[그림 2] 질의 이미지에 따른 정답 이미지



[그림 4] 이미지 유사도 비교 그래프

것의 경우 물체의 구성에 초점을 두어 유사도를 판단하므로 동일한 이미지에서 색상만이 다른 경우가 아닌 이상 유사함을 느끼기 어렵다. 그러므로 단일한 특징을 이용해서가 아닌 복수의 특징들과 조합하여 유사한 결과를 내야하며, [5]와 같이 다른 특징에 영향을 받지 않는 특징들을 이용하면 중복되지 않는 효율적인 이미지 검색 기법을 구현할 수 있다. 그러므로 그러한 특징들이 어떠한 것이 있는지 찾아보고, 어떠한 조합으로 특징들을 최소화 한 최적의 기법을 구현할 수 있는지 탐색한다.

6. 참고문헌

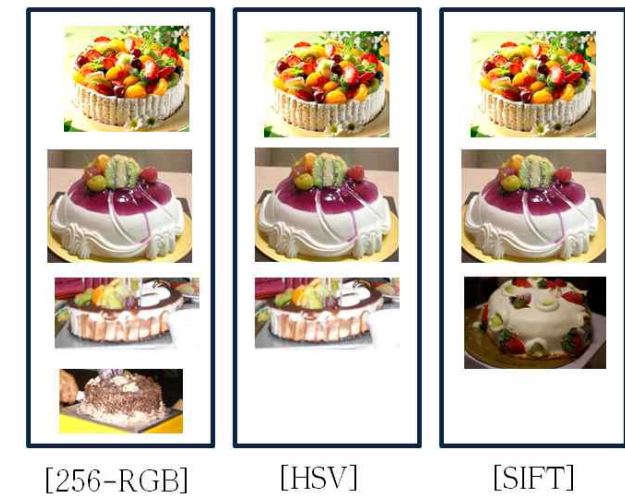
[1] S. Shirazi, N. Khan, A. Umar, M. R. Naz, B. A. Haqbani, "Content-Based Image Retrieval Using Texture Color Shape and Region," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 1, pp. 418-426, 2016.

[2] P. S. Hiremath, J. Pujari, "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement," *Proceedings of the 15th International Conference on Advance Computing and Communications*, pp. 780-784, 2007.

[3] S. Jeong, C. SunWon, R. M. Gray, "Image Retrieval Using Color Histograms Generated by Gauss Mixture Vector Quantization," *Computer Vision and Image Understanding*, vol. 94, Issues 1-3, pp. 44-66, 2004.

[4] J. Q. Ma, "Content-Based Image Retrieval with HSV Color Space and Texture Features," *Proceedings of the 2009 International Conference on Web Information Systems and Mining*, pp. 61-63, 2009.

[5] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.



[그림 3] 정답이미지와 매칭되는 결과이미지 목록 예시

게 나오는 것이 보편적이거나, 본 논문에서는 HSV를 양자화 하였으므로 [2] 보다도 낮은 유사도 결과를 낸다.

[5]의 경우 색상 기반에 비하여 물체 기반 이미지 유사도 측정이 비교적 높은 정확도를 나타내나, 본 논문의 실험 모두 그 값이 0.8 이하로 유사도 측정이 정확히 이루어지지 않으며, 동일한 이미지를 색 반전 시킨 후 두 이미지에 대해 질의 이미지와의 유사도를 관찰한 경우, [3]과 양자화 된 HSV 모두 색상이 달라져 유사도에 영향을 받으나 SIFT의 경우 두 경우 모두 질의 이미지에 대한 유사도에 변함이 없어 색상에 영향을 받지 않는다.

5. 결론 및 향후 연구

단일한 특징을 이용할 경우 해당 특징이 아닌 다른 특징에 대해서는 유사도를 전혀 측정할 수 없으므로 모든 특징에 부합하는 결과를 추출해 낼 수 없다. 색상의 특징을 이용한 결과가 더 정확하게 나오기 위해서는 색상의 배치 또한 고려 대상이 되어야 하며 특징 점을 찾는