

SIFT 이진화 후-검증 방법을 이용한 근-복사 이미지 검출 시스템 구현

이유진[○] 낭종호

서강대학교 컴퓨터 공학과

yujile09@sogang.ac.kr, jhnang@sogang.ac.kr

Implementation of Near-Duplicate Image Detection using SIFT Binarization Post-Verification

Yu Jin Lee[○] Jongho Nang

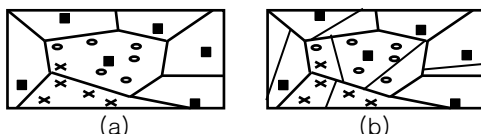
Department of Computer Science and Engineering, Sogang University

요 약

최근 이미지 콘텐츠에 쉽게 접근할 수 있는 인터넷 환경과 이미지 편집 기술들의 보급으로 근-복사 이미지가 폭발적으로 증가하면서 관련 연구가 활발하게 이루어지고 있다. 그러나 근-복사 이미지 검출 방법으로 주로 쓰이는 BoF(Bag of Feature)[1]는 고차원의 지역 특징을 저차원으로 근사화 하는 양자화 과정에서 서로 다른 특징들을 같다고 하거나 같은 특징을 다르다고 하는 문제가 발생할 수 있다. 따라서 양자화의 한계를 극복하기 위한 BoF를 후-검증하는 방법이 필요하다. 이러한 문제점을 해결하기 위하여, 본 논문에서는 BoF의 후-검증 방법으로 SIFT(Scale Invariant Feature Transform)[4]를 128bit의 이진 코드로 변환한 후 코드들간의 거리를 비교하는 시스템을 구현하였다. 구현한 시스템에 의하면 기존의 BoF 방법과 비교하여 근-복사 이미지 검출 정확도가 향상되는 효과가 있다.

1. 서 론

최근 이미지 편집 기술의 보급과 인터넷 환경의 발달로 이미지의 전체 또는 일부가 다양한 변형에 의해 생성된 근-복사 이미지들이 급증하면서 원본 이미지의 저작권 보호를 위한 근-복사 이미지 검출 방법이 활발히 연구되고 있다. 근-복사 이미지 검출의 대표적인 방법으로 이용되는 BoF (Bag of Feature)[1]는 이미지를 지역 특징들의 집합으로 보고 이미지들로부터 지역 특징들을 추출하여 k-means clustering 을 통해 k 개의 코드 워드를 추출한다. 추출된 코드 워드를 이용하여 d 차원의 지역 특징들을 코드 워드들의 인덱스 집합으로 근사화하여 표현하고 이미지들간의 코드 워드 인덱스 집합간 유사도를 측정하여 검출하는 방법이다. 고차원의 지역 특징을 하나의 코드 워드의 인덱스로 사상하기 때문에 검색 공간을 줄일 수 있다는 장점이 있다. 그러나 이미지에 가해진 변형에 의하여 <그림 1>-(a)와 같이 같은 이미지임에도 서로 다른 코드 워드에 사상되거나 (b)와 같이 다른 이미지임에도 불구하고 같은 코드 워드로 사상되어 근-복사 이미지 검출 성능 저하에 영향을 미칠 수 있다.



■=코드 워드, ○=지역 특징, x=지역 특징 잡음

<그림 1> 벡터 양자화 과정으로 인한 한계

이러한 BoF 의 양자화로 인한 문제를 해결하기 위하여 HE (Hamming Embedding)[2]와 같은 후-검증 방법이 제안되었다. HE 는 기존 BoF 방법에서 유사도가 높은 이미지들에 대하여 코드 워드에 해당하는 지역 특징 공간을 임의의 비트로 구성된 이진 코드로 나누어 표현하고 XOR 연산을 통한 이진 공간에서의 거리로 유사도를 판단하여 근-복사 이미지를 검출한다. HE 를 이용한 BoF 방법은 검색을 효율적으로 수행하면서 분별력을 높일 수 있다는 장점을 지니지만 모든 샘플 데이터들을 학습 시키는 과정이 필요하며 매핑 함수가 복잡하여 구현의 어려움이 있다.

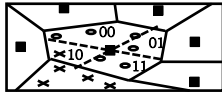
본 논문에서는 이러한 점들을 극복하기 위하여, 간단한 매핑 함수를 이용하여 SIFT(Scale Invariant Feature Transform) [4]를 128bit 로 이진화한 후 이들 간의 이진 거리로 유사도를 측정하는 후-검증 방법을 적용하여 시스템을 구현하였다. 제안한 시스템을 구현 및 실험 결과, 기존의 BoF 보다 향상된 성능을 확인할 수 있었으나 HE 방법과 비교하였을 때 단순한 구현 과정으로 인하여 나타나는 한계점들에 대하여 확인하였다.

2. HE

HE[2]는 BoF의 양자화 과정에서 발생하는 근사화에 따른 문제를 극복하기 위해 세밀하게 양자화할 때의 장점과 희박하게 양자화할 때의 장점을 결합한 방법이다. d 차원의 이진 차원을 d_b 차원이라고 할 때, 지역 특징을 양자화한 코드 워드 인덱스들의 특징 공간을 <그림 2>와 같이 d_b 차원의 이진 코드로 나누어 표현한다. 이미지들의 코드 워드가 일치할 경우 각 지역 특징들이 사상된 이진 코드들의

본 연구는 미래창조과학부 및 한국산업기술평가관리원의 산업융합원천기술개발 사업(정보통신)의 일환으로 수행하였음. [10044615, 클라우드 기반 개방형 소셜 방송미디어 콘텐츠 융합 생성, 편집 및 재생을 위한 미디어 제작 및 전송 시스템 개발]

이진 거리를 구하여 이미지들간의 유사도를 측정하는 방법이다.



■=코드 워드, ○=지역 특징, x=지역 특징 잡음

<그림 2> 지역 특징이 사상되는 코드 워드에 대한 HE 구조

HE 는 질의 이미지가 주어졌을 때 질의 이미지의 지역 특징을 희박한 코드북의 코드 워드 인덱스로 사상한 다음 원본 이미지의 지역 특징 중 같은 코드 워드 인덱스로 사상된 지역 특징들에 대하여 유효하다. 상호의 코드 워드가 일치한 경우에 대하여 d_b 차원으로 다시 한번 나누어진 내부 공간을 한번 더 검색하므로 양자화를 위한 코드 워드의 크기를 어떻게 하느냐에 구애 받지 않고 BoF 보다 한층 분별력을 높일 수 있다. 또한, d_b 차원으로 내부 공간을 나눌 때 공간의 인덱스를 이진 코드로 표현하기 때문에 이미지들 간의 유사도 판단 시 XOR 연산으로 가능하여 빠르게 유사도를 측정할 수 있다. 따라서 HE 는 BoF 를 후-검증하는데 있어서 매우 효과적인 방법이다. 그러나 HE 는 코드 워드 내부 공간을 나누기 위한 추가적인 학습 과정이 필요하고 최대한 지역 특징의 정보가 유지되도록 하기 위한 과정이 복잡한 단계와 연산으로 이루어져 있다. 이로 인해 HE 를 이용한 새로운 시스템을 구축하기에는 시간 및 공간적으로 높은 구현 비용을 요구되며 실제 시스템으로 구축하기에는 구현의 어려움이 존재한다.

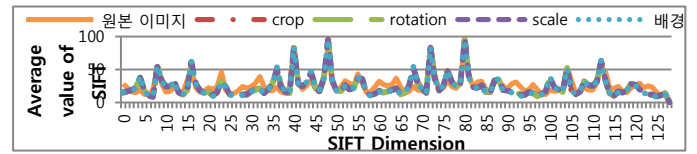
3. SIFT 이진화를 통한 근-복사 이미지 검출

본 논문에서는 실제 시스템 구현이 복잡한 HE 를 대신하여 구현이 간단하면서도 유사한 효과를 얻을 수 있는 SIFT 이진화 후-검증 방법을 제안한다. SIFT 는 특징 점 주변 픽셀 값들의 방향과 크기를 128 차원의 히스토그램으로 표현한 것으로 회전과 크기에 강건하다. 그러므로 근-복사 이미지 검출의 분별력을 높이기 위한 가장 좋은 방법은 이미지에서 추출한 SIFT 를 직접 비교하는 것이지만 모든 원본 이미지들의 SIFT 들에 대해서 질의 이미지의 SIFT 들을 모두 비교하기 위해서는 검색 비용이 매우 커지게 된다. 본 구현 시스템의 방법은 SIFT 가 지닌 픽셀 정보를 충분히 이용하되 검색 비용을 줄이기 위해서 통계적 방법을 이용하여 0~360 의 값을 가지는 128 차원의 SIFT 를 각 차원 별로 0 또는 1 을 부여하여 총 128bit 의 이진 코드로 변환한다. 그리고 BoF 에서 얻어진 짧은 순위 리스트들에 대하여 이진 코드들의 해밍 거리를 측정하여 유사도를 판단하고 근-복사 이미지를 검출한다. 다음으로 이어지는 3.1 절에서는 SIFT 를 128bit 의 이진 코드로 변환하는 통계적 방법을 이용한 이진 매핑 함수에 대해 설명하고 3.2 절에서는 BoF 에서 얻어진 짧은 순위 리스트 이미지들에 대한 재 순위화하는 과정에 대하여 자세히 설명한다.

3.1. 통계적 방법을 이용한 이진 매핑 함수

SIFT 각 차원의 픽셀 정보를 128bit의 이진 코드에 반영한 이진 코드 변환 기준을 구축하기 위하여 원본 이미지, 배경에 원본 이미지가 crop, rotation 그리고 scale 변환된 이미지와 배경 이미지의 각 차원 별 SIFT값을 분석하기 위하여 모든 원본 이미지와 그에 연관된 변환 및 배경 이미지들로부터 차원 별 SIFT 값을 추출한 후 이들의 평균을 구해 <그림 3>의 그래프로 나타내었다. <그림 3>의 그래프를 통해 변환된 이미지들의 그래프와 배경 이미지의 그래프가 유사하게 나타나는 반면 원본 이미지의 그래프가 다른 그래프에 비하여 약 4 정도 값이 차이가 난다는 것을 확인할 수 있다. 이를 통해 원본 이미지에 추가된 배경과 같은 잡음이 SIFT에 영향을 주며 오검출 이미지들과 구분 지어지는 원본 이미지 고유의 값으로 이진 코드를 구성해야 함을 알 수 있다. <그림 3>을 통한 분석을 바탕으로 본 논문에서 구현한 시스템에서는 모든 원본 이미지의 각 차원 별 SIFT의 평균을 이진 코드 변환의 기준 $\tau = (\tau_1, \dots, \tau_{128})$ 로 설정하였다. 원본 이미지의 수가 m 인 모든 원본 이미지 내의 SIFT $x_i, i=1, \dots, m$ 에 대한 변환 기준 τ 는 다음과 같이 계산된다.

$$\tau = \frac{1}{m} \sum_{i=1}^m x_i \quad (1)$$



<그림 3> 모든 원본 및 변환 이미지의 차원 별 SIFT 값

SIFT x 의 128bit 이진 코드를 $b(x) = (b_1(x), \dots, b_{128}(x))$ 라 할 때, 변환 기준 τ 를 이용한 이진 매핑 함수 f_b 는 다음과 같이 정의된다.

$$b(x) = f_b(x), \text{ where } f_b(x) = \begin{cases} 1 & x < \tau \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

3.2. BoF의 투표 방식을 통해 생성한 짧은 순위 리스트에 대한 재 순위화

BoF에서는 질의 이미지와 동일한 코드 워드 인덱스가 존재할 경우 원본 이미지에 점수를 누적하여 부여하여 최종 점수가 높은 상위 이미지들에 대하여 짧은 순위 리스트를 구성한다. 본 시스템에서는 BoF에서 얻어진 짧은 순위 리스트 내의 이미지들에 대해서 코드 워드 인덱스가 동일한 경우에 3.1절에서 변환한 128bit로 이진화한 이진 코드에 대한 이진 거리를 계산하고 이진 거리가 임계치 h_t 보다 작은 경우 이미지에 점수를 누적하여 부여하여 재 순위화 하였다. 질의 이미지를 y 라 하고 비교하는 원본 이미지를 x 라 할 때 이진 거리를 이용하여 구한 점수 s 는 다음과 같이 계산된다.

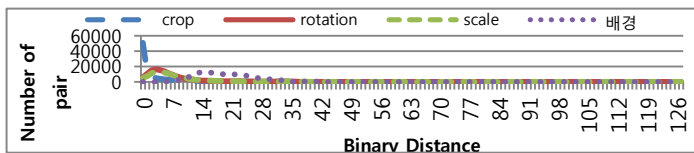
$$s := s + f_{SB}(x, y)$$

이 때, 매칭 함수 f_{SB} 는 다음과 같이 정의된다.

$$f_{SB}(x, y) = \begin{cases} 1 & q(x) = q(y) \text{ and } h(b(x), b(y)) \leq h_t \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{where } h(b(x), b(y)) = \sum_{i=1}^{128} |b_i(x) - b_i(y)| \quad (3)$$

본 구현 시스템에서 적용할 이진 거리 임계치 h_t 를 정하기 위하여 원본 이미지와 올바르게 매칭된 변환 이미지들의 이진 거리와 올바르게 매칭되지 않은 이미지들의 이진 거리를 계산하여 <그림 4>의 분포로 나타내었다. 그래프를 통해 잘못 매칭된 이미지들의 경우 이진 거리가 0에서 먼 것을 확인할 수 있으며 올바르게 매칭 될수록 0에 가깝게 분포하고 crop일 경우 이진 거리가 2일 때, 이외 변형의 경우 이진 거리가 4일 때 분포의 피크를 확인할 수 있다. 본 논문에서는 완벽히 매칭된 값에 대하여 점수를 부여할 수 있도록 모든 변형에 대하여 이진 거리 분포의 피크보다 1씩 작은 값, 즉 1과 3을 h_t 로 하여 시스템을 구현하였다.



<그림 4> SIFT 이진화 코드에 따른 이진 거리 분포

정해진 h_t 를 기반으로 짧은 순위 리스트 내의 이미지 중 질의 이미지와 비교한 점수를 바탕으로 원본 이미지들을 재 순위화하여 가장 많은 점수를 획득한 이미지를 최종 검출된 이미지로 판단한다.

4. 실험 및 분석

본 논문의 실험은 1500장의 배경이 알파인 원본 이미지에 대하여 3가지 변형(crop, rotation, scale)을 가한 후 76장의 배경을 무작위로 선정하여 추가한 질의 이미지 각 1500장으로 실험을 진행하였다. 시스템의 검출 성능 파악을 위해서 정확도를 측정하였으며 정확도 A는 질의 이미지의 수가 Q이고 Q 중 정확하게 검출된 이미지 수를 T라고 할 때 다음과 같이 계산된다.

$$A = \frac{T}{Q} \quad (4)$$

본 논문의 실험에서는 BoF 에서의 짧은 순위 리스트 크기를 30으로 설정하였으며 이에 따른 실험 결과는 [표 1]과 같다. [표 1]에 대한 분석을 통해 SIFT 이진화 후-검증 방법이 기존 BoF의 검출 정확도에 비하여 일반 코드북의 경우 약 0.03 향상되었으며 속도를 향상시키기 위하여 적용한 계층적 코드북[5]의 경우 약 0.04 향상된 것을 확인할 수 있다. 그러나 [표 1]의 결과로 보았을 때 본 구현 시스템의 정확도가 소폭 향상되긴 하였으나 HE와 비교하였을 때 SIFT가 가진 픽셀 정보를 최대한 내포하기 위한 128bit라는 이진 코드의 크기로 인하여 비교적 큰 오버헤드를 유발할 수 있다. 예를 들어, 질의 이미지 한 장에 대해 약 17byte 이상의 storage 오버헤드가 발생할 수 있으며, 1500장의 이미지라고 하였을 때 25,500byte의 storage 오버헤드가 발생할 수 있다. 또한, 속도 면에서는 SIFT의 이진화를 하는 과정에서 약 20ms의 오버헤드가 발생 가능한 한계점이 있다.

[표 1] 각각의 질의 이미지에 대한 정확도

| 코드북 | BoF | | | | BoF + SIFT 이진화(Ours) | | | |
|---------------|------|----------|-------|---------------|----------------------|----------|-------|---------------|
| | crop | rotation | scale | 평균 검출 속도 (ms) | crop | rotation | scale | 평균 검출 속도 (ms) |
| Original 128K | 0.94 | 0.81 | 0.82 | 4768 | 0.97 | 0.93 | 0.94 | 4781 |
| 1K | 0.91 | 0.78 | 0.76 | 494 | 0.96 | 0.82 | 0.83 | 518 |
| 2K | 0.91 | 0.78 | 0.76 | 500 | 0.95 | 0.82 | 0.83 | 525 |

5. 결론

본 논문에서는 BoF의 양자화로 인한 한계점을 보완하기 위해서 구현이 복잡한 HE가 아닌 SIFT의 이진화를 이용한 후-검증 방법을 제안하였다. 제안한 방법은 BoF에서 질의 이미지와 원본 이미지의 코드 워드가 매칭될 경우 원본 이미지에 점수를 부여하여 짧은 순위 리스트를 구성하고 리스트 내의 후보 이미지와 질의 이미지로부터 추출한 128차원의 SIFT를 각 차원 별 평균으로 이루어진 간단한 이진 매핑 함수를 통해 128bit의 이진 코드를 생성하여 이진 거리로 유사도를 측정하였다. 그러나 본 논문에서 제시한 방법은 HE에 비하여 매우 단순한 이진 매핑 함수와 이를 통한 구현으로 인하여 작은 비트로 양자화가 지닌 문제점을 충분히 보완하지 못하는 한계점이 있다. 비트 수가 큰 이진 코드를 사용하므로 storage와 속도에 있어서 오버헤드가 발생 가능하다. 따라서, 향후 HE 구현을 통한 HE에 대한 명확한 프로세스 이해와 이를 기반으로 발전시킨 간단하면서 추가적인 오버헤드를 줄인 BoF에 대한 새로운 후-검증 방법 연구가 필요하다.

참고 문헌

[1] S. Lazebnik, C.Schmid, and J.Ponce, "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, vol.2, pp.2169-2178, 2006.

[2] H. Jegou, M. Douze, and C. Schmid, "Improving Bag-of-Feature for Large Scale Image Search," *International Journal of Computer Vision*, vol.87, no.3, pp.316-336, 2010.

[3] K. He, F. Wen, and J. Sun, "K-means Hashing: an Affinity-Preserving Quantization Method for Learning Binary Compact Codes," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp.2938-2945, 2013.

[4] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol.60, no.2, pp.91-110, 2004.

[5] 이제민, 이유진, 남종호, "Bag of Feature를 이용한 근-복사 이미지 검출 시스템 구현," *한국정보과학회 2014 한국컴퓨터종합학술대회 논문집*, pp.1427-1429, 2014.