

# SIFT 기술자 분산을 이용한 근-복사 클립아트 이미지 검출 후-검증 방법

이유진<sup>o</sup> 남종호

서강대학교 컴퓨터 공학과

[yujile09@sogang.ac.kr](mailto:yujile09@sogang.ac.kr), [jhnang@sogang.ac.kr](mailto:jhnang@sogang.ac.kr)

## A Post-Verification Method of Near-Duplicate Clipart Image Detection using SIFT Descriptor Variance

Yu Jin Lee<sup>o</sup> Jongho Nang

Department of Computer Science and Engineering, Sogang University

### 요 약

최근 캐릭터 사업의 확장과 이미지 편집 기술의 발전으로 인하여 근-복사 클립아트 이미지가 급증하면서 저작권을 보호하기 위한 근-복사 이미지 검출 방법 연구에 대한 중요성이 점차 커지고 있다. 그러나 근-복사 이미지 검출을 위하여 주로 이용되는 BoF(Bag-of-Feature)[1]는 특징들의 차원 축소 과정에서 양자화의 한계를 지니며 이를 보완하기 위한 후-검증 방법이 필요하다. 본 논문에서는 BoF에 대한 후-검증 방법으로 SIFT(Scale Invariant Feature Transform)[2]들의 분산을 비교하여 유사도를 판단하는 방법을 제안하고 성능을 분석하였다. 1500장의 원본 클립아트 이미지들에 대하여 실험을 수행한 결과, 기존의 BoF 방법에 비하여 근-복사 이미지 검출 정확도가 6% 향상됨을 보였다.

### 1. 서 론

최근 캐릭터 시장의 확장과 이미지 콘텐츠 편집 기술의 발전으로 근-복사 클립아트 이미지가 급증함에 따라 원본 이미지 콘텐츠에 대한 저작권 침해가 문제가 되고 있다. 그러나 현재까지 근-복사 이미지 검출 연구는 클립아트 이미지보다는 일반 이미지 중심으로 이루어져 왔다. 따라서 근-복사 클립아트 이미지에 대해 강건히 검출할 수 있는 검출 방법에 대한 연구가 요구된다. 일반 이미지들에 대한 근-복사 이미지 검출 연구는 이미지를 작은 부분들의 집합으로 보아 공간적 변형이 가해졌을 때 강건하게 검출할 수 있는 지역 특징 기반의 방법이 대표적으로 이용되고 있다. 하지만 지역 특징을 이용한 경우 근-복사 이미지 검출 시 질의 이미지의 고차원 특징들에 대하여 원본 이미지들의 고차원 특징들과 상호 비교해야 하므로 시간 및 공간적으로 비용이 매우 커지게 된다. 이러한 지역 특징의 비용적인 단점을 해결하기 위하여 고차원의 특징을 저 차원의 인덱스로 사상하는 벡터 양자화 방법에 기반한 BoF(Bag-of-Feature)[1]가 주로 이용되고 있다. 근-복사 일반 이미지의 경우, 원본 이미지의 전체 또는 오브젝트에 대해서 근-복사가 일어나기 때문에 BoF를 적용하여 서로 매칭된 지역 특징의 양을 측정하였을 때 높은 성능을 보인다. 이와 달리 근-복사 클립아트 이미지의 경우, 원본 이미지의 부분 또는 서브 오브젝트에

대해서 복사가 일어나기 때문에 이와 달리 근-복사 클립아트 이미지의 경우, 원본 이미지의 부분 또는 서브 오브젝트에 대해서 복사가 일어나기 때문에 지역 특징의 양으로만 근-복사를 검출하는 데에는 한계가 있다. 한편 클립아트 이미지는 밝기 또는 색 반전과 같은 변환이 근-복사 시에 거의 일어나지 않기 때문에 지역 특징의 변화가 미세하여 질적인 부분에 대한 후-검증이 이루어진다면 클립아트 이미지에 BoF를 적용하였을 때 나타나는 한계를 보완할 수 있다.

본 논문에서는 클립아트 이미지의 특성을 고려하여 BoF를 통해 얻어진 서로 매칭된 128 차원의 SIFT(Scale Invariant Feature Transform)[2]들에 대하여 공간 및 시간 비용을 줄이면서 질적인 연관성을 이용할 수 있도록 각 SIFT의 분산을 구하고 이들간의 차이로 유사도를 측정하는 후-검증 방법을 제안한다. 제안한 방법을 적용한 결과, 기존의 BoF 방법에 비하여 향상된 성능을 확인할 수 있었다.

### 2. BoF 및 기존 BoF 기반의 후-검증 방법

BoF는 모든 지역 특징들을  $k$  개의 대표값으로 표현하는 코드북을 생성하고 고차원의 지역 특징들을 1 차원 코드북 인덱스에 사상하여 동일한 인덱스가 많이 나타나는 이미지를 결과 이미지로 검출하는 방법으로 시간 및 공간적 비용을 줄이면서 분별력을 높일 수 있다. 그러나 그림 1과 같이 코드북 대표값의 개수인  $k$ 의 크기에 따라 검출 성능이 크게 달라질 수 있다는 한계가 있다. 예를 들어, 그림 1의 (a)와 같이  $k$ 가 작은 경우에는 같은 특징임에도 다른 인덱스에 사상이 될 수 있으며 (b)와 같이  $k$ 가 큰 경우에는 다른

\*본 연구는 미래창조과학부 및 정보통신기술연구진흥센터의 정보통신-방송 연구개발사업의 일환으로 수행하였음. [R0101-15-0097, 클라우드 기반 개방형 소셜 방송미디어 콘텐츠 융합 생성, 편집 및 재생을 위한 미디어 제작 및 전송 시스템 개발]

특징임에도 같은 인덱스에 사상되어 검출 성능 저하에 영향을 미칠 수 있다.



그림 1 BoF의 벡터 양자화 과정으로 인한 한계

BoF의 한계를 해결하기 위하여 HE[3] 및 SIFT 이진화[4] 등 다양한 후-검증 방법이 제안되고 있다. HE는 인덱스 공간을 이진 코드로 나누어 표현하고 XOR 연산으로 유사도를 판단하여 검색을 효율적으로 수행함과 동시에 분별력을 높일 수 있으나 구현 과정이 복잡하여 실제 시스템 구현에 어려움이 있다. 이와 달리 SIFT 이진화는 128 차원의 SIFT를 각 차원에 대하여 하나의 bit로 표현한 128bit의 이진 코드간의 XOR 연산을 통해 유사도를 판단하여 단순한 구현 과정을 지닌다. 그러나 단순한 구현 과정으로 인하여 BoF의 한계를 충분히 보완하지 못하며 HE보다 분별력이 낮고 공간적 비용이 크기 때문에 실제 시스템에 적용하기에는 한계가 있다.

### 3. SIFT 기술자 분산을 통한 근-복사 클립아트 이미지 검출

본 논문에서는 근-복사 클립아트 이미지에 대하여 시간 및 공간적 비용이 적으면서도 기존의 BoF보다 개선된 성능을 보이는 SIFT 기술자 분산 후-검증 방법을 제안한다. SIFT 기술자는 회전, 크기 및 밝기 변화 등에 강건하여 이미지 검색에서 대표적으로 이용되는 지역 특징으로 128 차원으로 표현된다. 근-복사 클립아트 이미지 검출을 위해서는 SIFT를 직접 비교하는 것이 분별력에 있어 높은 성능을 얻을 수 있지만 모든 원본 이미지와 질의 이미지를 비교하기에는 공간 및 시간적 비용이 많이 소모된다. 본 논문에서는 질적인 연관성을 이용하면서 공간 및 시간적 비용을 줄일 수 있도록 128 차원의 SIFT의 분산 값을 구한다. 그리고 BoF에서 유사도 측정을 통하여 얻어진 짧은 순위 리스트들에 대하여 분산의 차이를 측정하여 다시 한번 유사도를 판단하고 최종적으로 근-복사된 이미지를 검출한다. 다음으로 이어지는 3.1 절에서는 본 논문에서 적용한 SIFT 기술자의 분산의 개념과 연산에 대하여 구체적으로 기술하고 3.2 절에서는 BoF를 통해 구성된 짧은 순위 리스트 이미지들을 대상으로 SIFT 기술자 분산을 이용하여 재 순위화하는 과정에 대하여 자세히 설명한다.

#### 3.1. SIFT 기술자 분산

SIFT 기술자 분산은 128 차원의 SIFT 기술자에 대해서 128 차원 값에 대한 분포의 분산을 구한 것이다. SIFT 기술자  $z_i (i = 1, \dots, 128)$ 에 대한 분산  $\sigma^2$ 은 다음과 같이 정의된다.

$$\sigma^2 = \frac{\sum_{i=1}^{128} (z_i - \mu)^2}{128}, \text{ where } \mu = \frac{\sum_{i=1}^{128} z_i}{128} \quad (1)$$

BoF를 이용하여 질의 이미지와 원본 이미지 내의 SIFT 기술자의 코드북 인덱스가 일치하는 경우, 올바르게 코드북

인덱스가 매칭된 경우도 있지만 BoF 양자화 과정의 한계로 인하여 올바르게 사상되어 매칭된 코드북 인덱스에 사상된 기술자가 존재할 수 있다. 올바르게 사상되어 매칭된 코드북 인덱스들이라면 SIFT의 각 차원 별 값이 유사하게 나타날 것이다. 즉, SIFT 기술자의 분산을 적용한다면 유사한 SIFT 기술자들의 경우 SIFT 기술자들의 분산 차이가 거의 나타나지 않을 것이다. 본 논문에서는 이러한 점을 이용하여 BoF 후 짧은 순위 리스트 내의 이미지들에 대하여 기술자 분산을 적용하였다. SIFT 기술자 분산을 이용할 경우 128 차원의 SIFT 기술자를 직접적으로 비교하지 않고 분산을 이용하여 유사도를 측정할 수 있어 공간적으로는 원본 이미지의 128 차원 SIFT 기술자가 아닌 분산 값만을 저장하면 되므로 비용이 최소화되며 시간적으로는 차원 축소로 인하여 유사도 측정에 필요한 비용을 최소화 가능하다는 장점이 있다.

#### 3.2. SIFT 기술자 분산을 이용한 재 순위화

본 논문에서는 BoF 후의 짧은 순위 리스트  $C$  내의 이미지들  $[I_1, I_2, \dots, I_m]$ 에 대해서 코드북 인덱스가 동일한 경우 SIFT 기술자 분산 차이를 계산하고 차이가 임계치  $\gamma$ 보다 작은 경우 이미지에 점수를 부여하며 모든 이미지에 대한 점수 부여가 종료된 후 짧은 순위 리스트 내의 이미지들에 대해서 순위를 재 측정한다. 질의 이미지  $k$ 의 SIFT 기술자들을  $k_i (i = 1, \dots, x)$ ,  $x$ 을 질의 이미지의 SIFT 기술자 수라 하고 원본 내 비교 이미지  $o$ 의 SIFT 기술자들을  $o_{i,j} (i = 1, \dots, y_j)$ ,  $y_j$ 를  $j$ 번째 이미지의 SIFT 기술자 수라 할 때 비교 이미지  $j$ 에 대한 점수  $s_j$ 는 다음과 같이 정의된다.

$$s_j := s_j + f(k_i, o_{i,j}) \quad (2)$$

이 때, 재 순위화 함수  $f$ 는 다음과 같이 정의된다.

$$f(k, o) = \begin{cases} 1 & q(k) = q(o) \text{ and } |\sigma^2(k) - \sigma^2(o)| \leq \gamma \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

본 논문의 실험에서 적용할 분산 값들 차이의 임계치  $\gamma$ 를 정하기 위하여 원본 이미지와 원본 이미지를 변형한 이미지들간의 코드워드가 일치한 경우의 분산 차이와 원본 이미지와 관련이 없는 이미지들간의 코드워드가 일치한 경우의 분산 차이를 계산하여 그림 2의 분포로 나타내었다. 그래프를 통해 원본 이미지와 관련이 없는 배경 이미지들의 경우 코드워드가 일치하였다 하더라도 분산 차이가 0에 거의 분포하고 있지 않음을 확인할 수 있으며 원본 이미지와 관련도가 높을 수록 코드워드가 일치하는 경우에 대하여 분산의 차이가 0에 가깝게 분포한다는 것을 확인할 수 있다. 특히, 크롭의 경우 차이가 0일 때, 회전과 크기의 경우 차이가 10일 때 분산 차이의 분포에서 피크를 나타냄을 확인할 수 있다. 본 논문에서는 SIFT 기술자들의 매칭이 올바르게 이루어진 값에 대하여 점수가 부여될 수 있도록 각 변형에 대하여 분산 차이 분포에서 보인 피크와 동일한 0과 10을 임계치  $\gamma$ 로 하여 실험을 수행하였다.

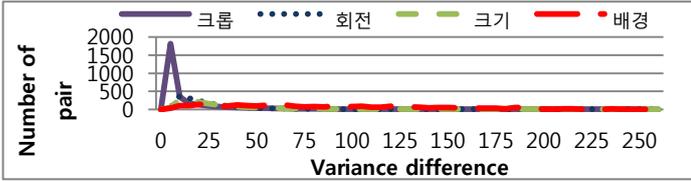


그림 2 원본 및 질의 이미지의 SIFT 기술자 분산 차이 분포  
 설정된  $\gamma$ 를 바탕으로 짧은 순위 리스트 내의 이미지와 질의 이미지를 재 비교한 점수를 이용하여 짧은 순위 리스트 내의 원본 이미지들을 재 순위화하고 그 중 가장 높은 점수의 이미지를 최종 결과 이미지로 검출한다. 그림 3 은 3 절에서 설명한 SIFT 기술자 분산 후-검증 방법에 대한 알고리즘이다.

```

function MATCHANDRERANK(  $C = [I_1, I_2, \dots, I_m]$  )
    Let  $S = [s_1, s_2, \dots, s_m] \in \mathbb{N}^k$ 
     $i \leftarrow 1, S_{p \in [1, m]} \leftarrow 0$ 
    while  $i \leq m$ 
        Let  $k$  be the number of  $x_i$  in  $I_i$ 
        Let  $l$  be the number of  $x_y$  in  $I_y$ 
        for  $\forall_j \in [1, k], \forall_t \in [1, l]$ 
            if  $Q(x_i^j) = Q(x_y^t)$  and  $\text{abs}(\sigma^2(x_i), \sigma^2(x_y)) \leq \gamma$ 
                 $S_i \leftarrow S_i + 1$ 
            end if
        end for
    end while
     $p \leftarrow \text{argmax}_p S_p$ 
    return  $I_p$ 
end function
    
```

그림 3 SIFT 기술자 분산 후-검증 방법 알고리즘

4. 실험 및 분석

본 실험에서는 SIFT 기술자 분산을 이용한 근-복사 클립아트 이미지 검출 후-검증 방법의 성능을 실험 및 분석하기 위하여 배경이 알파인 원본 이미지 1500장에 대하여 실험을 수행하였다. 질의 이미지는 원본 이미지에 76장 중 하나의 배경을 무작위로 추가하고 3가지 변형(크롭, 회전, 크기)을 가한 4500장으로 실험을 진행하였다. 검출 성능을 파악하기 위하여 정확도를 측정하였으며 정확도 A는 질의 이미지의 수가 Q일 때 Q 중 정확하게 검출된 이미지 수가 T인 경우 다음과 같이 정의된다.

$$A = \frac{T}{Q} \tag{4}$$

본 논문에서 제안하는 후-검증 방법의 성능 확인을 위하여 BoF 에서 구성되는 짧은 순위 리스트의 크기를 30 으로 설정하여 실험을 수행하였으며 이에 대한 결과는 표 1 과 같다. 표 1 에 따르면 기존 BoF 방법에 비하여 제안 방법의 검출 정확도가 크롭 변형 이미지에 일반 코드북을 적용한 실험보다 약 0.03 향상되었다. 또한 검출 속도 개선을 위하여 크롭 변형 이미지에 계층적 코드북[5]을 적용한 실험의 경우 검출 정확도가 약 0.06 향상되었으며 SIFT 이진화[4]를 적용한 후-검증 방법에 비하여서도 약 0.01 향상된 결과를 확인할 수 있었다. 이러한 결과를 통해 BoF 의 양자화로 인하여 서로 다른 코드북 인덱스임에도 동일한 인덱스에

사상되어 있던 특징들에 대한 후-검증이 수행되었으며 이로 인해 성능이 향상된 것을 확인할 수 있다. 마지막으로 SIFT 이진화 방법이 질의 이미지 한 장당 17 byte 이상의 storage 오버헤드를 발생 시키는 반면, 제안 방법의 경우 질의 이미지 한 장당 4 byte 이상의 storage 오버헤드를 발생시켜 질의 이미지 한 장당 약 13 byte 의 공간적 비용을 줄일 수 있다.

표 1 각각의 질의 이미지에 대한 정확도

코드북 크기	BoF				BoF + SIFT 기술자 분산(Ours)			
	크롭	회전	크기	평균 검출 속도 (ms)	크롭	회전	크기	평균 검출 속도 (ms)
Original 128K	0.94	0.81	0.82	4768	0.97	0.93	0.94	4521
1K	0.91	0.78	0.76	494	0.97	0.83	0.84	481
2K	0.91	0.78	0.76	500	0.95	0.82	0.83	493

5. 결 론

본 논문에서는 BoF 가 지닌 양자화의 한계를 보완하기 위하여 SIFT 기술자들 간의 질적인 연관성을 이용하면서 시간 및 공간적인 비용을 최소화하는 SIFT 기술자 분산을 이용한 후-검증 방법을 제안하였다. 제안 방법은 BoF 에서 질의 이미지와 원본 이미지 간의 코드북 인덱스가 일치할 경우 부여한 점수를 바탕으로 상위 원본 이미지들에 대해서 짧은 순위 리스트를 구성한 후, 짧은 순위 리스트 내의 이미지들 중 코드북 인덱스가 일치하면서 SIFT 기술자의 분산 차이가 임계치보다 작을 경우 해당 이미지에 점수를 부여하여 가장 높은 점수를 얻은 이미지를 결과 이미지로 검출하였다. 그 결과 기존의 BoF 및 후-검증 방법들에 비하여 향상된 성능을 확인할 수 있었다.

참 고 문 헌

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

[3] 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.

[4] 이유진, 남종호, "SIFT 이진화 후-검증 방법을 이용한 근-복사 이미지 검출 시스템 구현," *2014 한국정보과학회 제 41 회 정기총회 및 동계학술발표논문집*, pp.1059-1061, 2014.

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