

지역 특징 매칭의 공간적 일관성을 이용한 근-복사 클립아트 이미지 검출

이유진^o 이제민 낭중호

서강대학교 컴퓨터 공학과

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

Near-Duplicate Clipart Image Detection using Spatial Consistency of local-feature-matching

Yu Jin Lee^o Jemin Lee Jongho Nang

Department of Computer Science and Engineering, Sogang University

요 약

멀티미디어 콘텐츠에 대한 저작권 관리의 중요성이 높아지면서 이미지 프로세싱 기반의 변환에 대한 근-복사 이미지 검출 방법이 주로 연구되고 있다. 그러나 클립아트와 같은 벡터 그래픽 형태의 콘텐츠의 경우 지역 특징에 대한 변화가 매우 적고 본질적으로 컴포지션 과정을 통해서 생성되는 근-복사 이미지가 많다. 따라서 지역 특징들의 매칭되는 양보다는 복사 영역에 대한 질적인 일관성을 이용하여 근-복사 이미지를 검출하는 것이 필요하다. 이러한 문제점을 해결하기 위하여, 본 논문에서는 질의 이미지와 원본 이미지 간의 공간적 일관성을 기반으로 한 근-복사 이미지 검출 방법을 제안한다. 제안한 방법에 의하면 기존의 방법과 비교하여 컴포지션에 민감하게 반응하지 않고 정확도가 향상되는 효과가 있다.

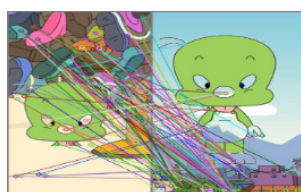
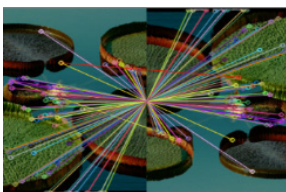
1. 서 론

최근 멀티미디어 콘텐츠의 폭발적인 증가와 함께 저작권 관리가 중요해지면서 근-복사 이미지 검출에 대한 관심이 높아지고 있다. 근-복사 이미지에 대한 여러 변환에 대하여 검출 방법들이 연구되고 있지만 대부분 비트맵 이미지를 대상으로 하고 있다. 벡터 그래픽에 대한 저작권 관리도 중요해 지고 있기 때문에 본 논문에서는 클립아트에 대한 효과적인 근-복사 검출 방법을 연구하였다.

근-복사 이미지 검출 방법들 중 영상을 지역 특징들의 집합으로 표현하여 비교 영상들의 집합간 유사도로 검출하는 BoF(Bag-of-Feature)[1] 기반의 방법이 주로 사용되고 있다. BoF 기반의 근-복사 검출 방법은 두 비교 영상의 서로 매칭되는 지역 특징들의 양을 측정하여 일치 여부를 판단한다.

그러나 클립 아트의 경우 본질적으로 근-복사가 컴포지션으로 이루어지기 때문에 주로 서브 오브젝트에 대하여 발생할 뿐만 아니라 단순화하여 표현되기 때문에 지역 특징들의 양으로 원본인가를 검출하는 데에는 한계가 있다. 한편 클립아트에 대한 근-복사 시 밝기나 색상 변화와 같은 이미지 프로세싱 기반의 변환은 잘 일어나지 않으므로 지역 특징에 대한 변화가 매우 적어 복사 영역에 대한 질적인 일관성을 이용하는 것이 적합하다고 볼 수 있다.

본 논문에서는 이와 같은 클립아트 근-복사 이미지 검출 문제의 특성을 고려하여, 영상을 BoF 로 표현하고 원본과의 공간적 일관성을 기반으로 검출하는 방법을 연구하였다. 제안한 방법은 먼저 크기나 회전에 강건하면서 컬러 공간 상에 주어진 유용한 특징들이 활용 가능한 SIFT[2]를 지역 특징으로 하여 매칭되는 특징들의 양으로 후보 이미지들을 선별하고 이들에 대하여 지역 특징들의 방향성과 상대적 좌표 정보를 이용한 공간적 일관성을 기반으로 유사성을 판단한다. 방향성에 대해서는 매칭된 지역 특징들의 방향 차이 히스토그램을 계산한 후 그 중 피크를 이용하여 두 영상 사이에 일관성이 있는지를 먼저 판단하는 방법을 적용하였다. 그 다음 방향적 일관성이 있는 특징들에 대하여 상대적 거리 차이의 변화량을 이용하여 상대 좌표를 통한 공간적 일관성을 판단하였다. 실험에 의하면 클립아트 근-복사 이미지 검출에 있어서 기존의 BoF 방법보다 본 논문에서 제시하는 방법이 정확도가 향상됨을 보였다.



(a) 비트맵 이미지

(b) 클립아트 이미지

<그림 1> 기존 BoF 방법으로 매칭한 비트맵 이미지와 클립아트 이미지

본 연구는 미래창조과학부 및 한국산업기술평가관리원의 산업융합원천기술개발사업(정보통신)의 일환으로 수행하였음.

[10044615, 클라우드 기반 개방형 소셜 방송미디어 콘텐츠 융합 생성, 편집 및 재생을 위한 미디어 제작 및 전송 시스템 개발]

2. 기존의 BoF 기반의 근-복사 이미지 검출 방법

기존 BoF 기반의 근-복사 이미지 검출 방법은 질의 영상과 대상 영상에서 추출한 지역 특징들의 집합을 구성하고 집합들 간의 유사도 측정을 하는 방법을 주로 이용하고 있다. 미리 대상 이미지의 지역 특징들을 집합으로 표현하고 대상 이미지들의 집합에 대한 역 테이블을 구성한다. 질의 이미지의 특징 집합을 역 테이블 내의 집합들과 비교하여 가장 유사한 집합을 찾은 뒤 점수를 부여하고 최종 점수가 최대값이 되는 이미지를 검출된 이미지라고 판단한다.

그러나 모든 대상 이미지들의 지역 특징들에 대하여 검색을 하기 때문에 BoF 는 대상 이미지들이 많은 경우 질의 이미지와의 검색을 위한 공간이 매우 커지게 되므로 사용하기에 적합하지 않다.

이에 기존의 근-복사 이미지 검출 방법들은 검색 공간의 크기로 인한 문제를 해결하기 위하여 근사가 적용되는 LSH[3], BoW(Bag of Words)[4]와 같은 기법들이 사용되고 있다. 그 중에서 메모리 사용량이 적을 뿐만 아니라 비교적 성능이 좋아 주로 사용되는 BoW 는 k-means clustering 을 이용하여 k 개의 코드 워드를 추출한 후 이 워드들을 바탕으로 양자화한다. d 차원의 지역 특징은 양자화를 통해 하나의 코드 워드 인덱스에 사상되며 두 이미지가 유사하다면 각각을 대표하는 지역 특징은 양자화를 통하여 동일한 코드 워드 인덱스를 가리킬 확률이 높다. 두 지역 특징이 같은 코드 워드 인덱스로 양자화 되었을 경우 두 이미지가 매칭되었다고 판단하며 벡터 양자화를 이용한 매칭에서는 코드 워드의 수를 나타내는 k 의 값이 검색 성능에 영향을 주게 된다.

하지만 클립아트 이미지의 경우 주로 서브 이미지에 대한 복사가 발생하기 때문에 복사된 영역과 배경이 근접한 부근의 지역 특징들은 변형이 심하게 일어난다. 이를 벡터 양자화에 적용하였을 시 같은 객체임에도 불구하고 완전히 다른 코드 워드로 사상되게 된다. 반면 복사된 객체 내에 존재하는 지역 특징들은 비교적 그 값이 잘 보존되기 때문에 사상되는 코드 워드의 양이 적더라도 질적으로 정확하게 매칭된다.

따라서 근-복사 클립아트 검출을 위해서는 점수를 계산하는 투표 방식과 같은 지역 특징의 출현 빈도수로 유사도를 측정하는 방법은 알맞지 않다. 이에 따라 질적 유사도를 고려하는 검출 방법을 이용하여 정확하게 매칭되는 부분을 찾아 검출하는 방법이 적합하다.

3. 공간적 일관성을 이용한 근-복사 검출 방법

본 논문에서는 공간적 일관성을 이용한 근-복사 검출 방법을 3 가지 단계로 나누었다. 첫 번째 단계는 대상 이미지와 질의 이미지의 양자화된 지역 특징들을 매칭한 후 2 에서 소개한 기존의 BoF 기반의 방법과 같이

매칭되는 지역 특징들이 일정량 이상 존재하는 이미지에 대하여 점수를 부여하고 계산된 점수에 따라 후보 이미지들의 짧은 순위 리스트(short ranked list)를 구성한다. 기존 BoF 방법이 짧은 순위 리스트에서 가장 높은 점수를 가지는 이미지를 검출하였다면 본 논문에서 제시하는 방법은 공간적 일관성을 고려하는 검증 단계를 수행한다. 검증을 위한 두 번째 단계에서는 짧은 순위 리스트에 해당하는 대상 이미지들에 대하여 모든 매칭된 지역 특징의 쌍들 중 대략적으로 공간적 일관성이 있는 특징들의 쌍을 찾기 위하여 대상 이미지와 질의 이미지의 방향 차이 히스토그램을 계산한다. 마지막 검증으로 두 번째 단계에서 선택한 특징 쌍들의 좌표 값을 활용하여 질의 이미지가 대상 이미지와 일치하는 지 확인한다. 다음으로는 두 번째 단계와 마지막 단계에 대한 좀 더 자세히 설명한다.

A. 방향 차이 히스토그램 생성

방향 차이 히스토그램은 매칭된 SIFT 특징 쌍들 사이의 오리엔트 차이의 분포를 나타낸다. 두 이미지 r 과 q 의 방향 차이 히스토그램은 다음의 식에 따라 계산될 수 있다.

$$h_i := h_i + f_i(l_i^r, l_i^q), \text{ for } q(l_i^r) = q(l_i^q)$$

$$\text{where } f_i(l, l') = \begin{cases} 1 & |ori(l) - ori(l')| < \frac{2\pi}{n} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

l_i^r 은 r 이미지의 i 번째 특징을 나타내며 $ori(l)$ 은 특징 l 의 방향을 나타낸다. 이 때, $ori(l)$ 은 SIFT 의 특성 상 $0 \sim 360^\circ$ 를 가질 수 있으므로 매칭되는 특징의 방향 차이인 $|ori(l) - ori(l')|$ 는 $-360 \sim 360^\circ$ 의 값을 지닌다. 본 논문에서는 실험적 결과에 의하여 10° 씩 73 차원으로 나누어 방향 차이 히스토그램을 구현하였다. 비교 영상들 간에 공간적 일관성이 있을 경우 이들의 방향 차이 히스토그램에는 방향적 일관성을 보여주는 피크가 있을 것이다. 이러한 피크에 해당하는 특징들을 반영하여 또 다른 검증 단계를 거친다면 보다 정확하게 매칭된 지역 특징들을 찾을 수 있다.

B. 좌표 값 차이 매트릭스를 이용한 개선

방향 차이 히스토그램을 사용할 경우, 공간적 일관성 영역을 보다 효율적으로 확인할 수 있다. 그러나 잘못 매칭된 특징들이 실제 매칭된 쌍들의 차이와 일치할 수 있기 때문에 몇몇의 잘못된 결과를 가져올 가능성이 있다. 그러므로 두 이미지 간의 좌표 값의 차이를 이용하여 부가적인 일관성 계산 방법을 적용하였다. 방향 차이 히스토그램에서 정점을 나타내는 집합의 모든 지역 특징들에 대하여 각각 이미지 r 또는 q 내에서 다른 지역 특징들과의 좌표 값들 사이의 거리를 계산한 좌표 매트릭스 D^r 과 D^q 가 있다. 이 매트릭스 D 는 다음과 같이 계산될 수 있다.

$$D \in \mathbb{R}^{m \times m}, D_{ij} = |pt(l) - pt(l')| \quad (2)$$

위의 식에서 m 은 방향 차이 히스토그램의 정점이 지니는 값과 동일한 값을 가진다. l 과 l' 은 동일한 이미지 내에

존재하면서 방향 차이 히스토그램에서 피크에 해당되는 각각의 지역 특징을 의미하며 $pt(l)$ 은 각 이미지 내에 존재하는 특징 l 의 좌표 값이다.

계산된 두 매트릭스 D^r 과 D^q 를 이용하여 비교 이미지들 사이에 공간적 일관성이 있는지 알아보기 위한 상대 거리 비율들의 집합을 구한다. 이 집합의 분산이 임계 치보다 작은 경우 공간적 일관성이 있는 특징들이라 판단할 수 있다. 이러한 상대 거리 비율들의 집합 R 과 집합 R 에 대한 분산 V 는 다음과 같이 계산될 수 있다.

$$V = \text{Var}(R) < T$$

$$\text{where } R = \{D_{ij}^r/D_{ij}^q\}_{i,j \in [1,m]} \quad (3)$$

D^r 를 D^q 의 값으로 나누어 상대 거리 비율을 구한 뒤 이들의 분산이 임계 치 T 보다 작은 경우 다른 지역 특징과의 거리 비율이 일정하다는 것으로 공간적 일관성이 높은 매칭되는 특징들 즉, 정확하게 매칭된 특징들이라고 판단할 수 있다. 짧은 순위 리스트에 포함된 후보 이미지들에 대하여 질의 이미지와의 분산 V 를 구했을 때 임계 치보다 분산이 작을 뿐 아니라 후보 이미지들 중에서 최소값을 가지는 분산을 가지는 후보 이미지에 대하여 질의 이미지에 대하여 근-복사된 클립아트 이미지라고 판단한다.

4. 실험적 결과

본 논문의 실험은 10000장의 원본 클립아트 이미지에 대하여 4가지 변형이 가해진 각 100장의 근-복사 클립아트 이미지로 실험하였다. 4가지 변형은 원본 클립아트 이미지에 대하여 그대로 크롭핑한 Crop1, Crop1에서 배경을 제거한 Crop2, Crop2에 단순한 배경을 더한 BGS, Crop2에 복잡한 배경을 더한 BGC로 구성하였으며 크롭핑된 이미지는 해상도를 달리하였다. 검색 성능을 평가하기 위해서 정확도를 측정하였으며 질의 이미지 수를 Q 라 할 때, Q 중 정확하게 검색된 이미지 수를 T 라 하면 정확도 A 는 다음과 같이 계산된다.

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

본 논문의 실험에서는 후보군 이미지를 10개로 설정하였으며 이에 따른 시험 결과는 표 1과 같다.

[표 1] 각각의 질의 이미지에 대한 정확도

	Crop1	Crop2	BGS	BGC	평균
기존 방법	0.91	0.51	0.47	0.28	0.54
본 논문 방법	0.90	0.75	0.70	0.35	0.67

표 1 은 제안한 방법이 기존의 BoF 방법보다 근-복사 이미지 검출 정확도에서 성능이 더 좋아짐을 보여주고 있으며 기존의 방법보다 복사된 영역과 배경이 근접한

부근의 지역 특징들의 변화에 민감하게 반응하지 않음을 알 수 있다.

5. 결 론

본 논문에서는 클립아트 이미지를 대상으로 하여 기존의 이미지 프로세싱 변환에 기반한 근-복사 이미지 검출 방법이 아닌 공간적 일관성을 고려한 근-복사 이미지 검출 방법을 제안하였다. 제안한 방법은 먼저 양자화된 대상 이미지와 질의 이미지의 지역 특징들이 매칭 될 경우 해당 대상 이미지에 점수를 부여 하여 후보 이미지를 선별하였다. 후보 이미지와 질의 이미지들 사이의 SIFT 특징의 방향 차이 히스토그램으로 방향적 일관성이 있다 판단되는 정점에 해당하는 특징들만 남겼다. 마지막으로 방향적 일관성이 있는 지역 특징들에 대하여 각 특징간의 상대적 거리 차이의 변화량을 이용하여 공간적 일관성을 판단하였다.

그러나 본 논문에서 제시한 방법은 근-복사된 전체 또는 부분 클립아트 이미지 내에 SIFT 지역 특징이 1 개 이상 추출 가능한 것들을 대상으로 구현된 것으로 지역 특징이 충분히 추출되지 않는 이미지에 대해서는 제안된 방법을 적용하기 어려운 문제점이 있다. 따라서, 본 논문에서 제시한 지역 특징 이외에 다른 특징들을 더하여 유사도를 측정한다면 더욱 좋은 결과를 얻을 수 있을 것으로 판단되며 향후 이를 이용한 유사도 측정 방법에 대한 연구가 필요하다.

참 고 문 헌

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

[2] D. Lowe, "Object Recognition from Local Scale-Invariant Features," *Proc. 7th IEEE International Conference on Computer vision*, pp.1150-1157, 1999.

[3] P. Indyk and R. Motwani, "Approximate Nearest Neighbor: Towards Removing the Curse of Dimensionality," *Proc. 30th Annual ACM Symposium on Theory of Computing*, pp.604-613, 1998.

[4] J. Sivic and A. Zisserman, "Efficient Visual Search of Videos Cast as Text Retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol.31, no.4, pp.591-606, 2009.

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