

유사-복제 영상 검색을 위한 영상식별자

전형용*, 방준원*, 김의홍*, 황치정*

*충남대학교 컴퓨터 공학과

e-mail : fantajeon@cnu.ac.kr

isutar84@nate.com

chkim@cs.cnu.ac.kr

cjhwang@cnu.ac.kr

Image Identifier for Nearest-Replica Images Retrieval

Hyeongyong Jeon*, Joonweon Bang*, Euihong Kim*, Chijung Hwang*

* Computer Engineering, Chungnam National University

요 약

미디어와 기술 발달의 따른 매우 원본과 같은 복제 영상인 유사-복제 영상들이 원저자의 동의 없이 사용자간의 교환이 무방비로 유통되고 있다. 이러한 문제점들을 해결하기 위하여, 본 논문에서는 새로운 영상 식별자를 제안한다. 제안된 영상 식별자는 특징점 기반이며, 그것의 주변 밝기 정보의 고유값 분해과정을 거쳐서 지역 기술자를 생성한다. 이 생성된 지역 기술자들을 원본 데이터베이스와 빠르게 검색하여 주어진 질의 영상과 관련된 원본을 찾는 시스템이다. 실험에서는 총 13 종류의 영상을 왜곡을 시행하였으며, 거리의 따른 임계값은 false-positive alarm 을 시행하여 결정하였다. 실험결과 매우 높은 정확도를 가지며, 영상으로부터 영상 식별자 추출 및 데이터베이스 검색속도 또한 매우 우수한 성능을 가진다.

1. 서론

인터넷을 통한 많은 사용자간의 미디어(media) 교환이 많아짐에 따라, 상당수 많은 부분의 복제본에 매우 가까운(유사-복제) 영상들이 교환되는 것이 현실이다. 또한, 원 소유자의 허락도 없이 무방비로 유통되고 있다.

유사 영상검색에서 많이 알려진 방법으로는 블록 기반한 edge-histogram[1,2], color-histogram[3]과 특징점의 기반한 SIFT[4]방법들이 존재를 한다. 그러나, 이러한 방법들은 유사 영상 검색이 기본 개념이므로, 유사-복제를 찾기에는 알고리즘들이 너무 포괄적이라고 할 수 있다. 이러한 문제를 해결하기 위하여, 다른 방법인 Radon Transform[5]이나, Trace Transform[6]들은 수학적인 불변 공간에서 추출하여 빠른 검색을 한다. 하지만, 이러한 공간은 아핀 변환(affine transform)에만 불변하여, 공간이 조금이라도 왜곡되거나, 추가적인 정보가 영상에 삽입된다면, 찾을 수 없다는 치명적인 오류가 있다.

제안한 방법은 이러한 유사-복제 본을 발견하기 위한 가장 기본적인 영상 식별자를 제안한다. 제안한 영상 식별자는 특징점 기반과 고유값으로 구성되어있다.

본 논문에 구성은 2 장에서는 영상 식별자를 제안하며, 3 장에서는 실험과정 및 정확도를 보여준다. 마지막 장에서는 결론으로 마무리를 한다

2. 영상 식별자 (Image Identifier)

제안한 영상 식별자는 특징점 추출 및 특징점의 주변 밝기 정보로 구성된다.

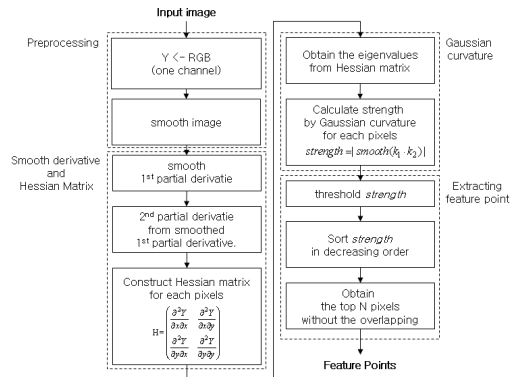


그림 1. 특징점 추출

첫번째 과정은 모든 화소의 gaussian curvature 을 사용하여 특징점을 추출하는 구간이다(그림 1 참조). 그리고 그 추출된 특징점 주위의 밝기 정보를 Singular Value Decomposition (SVD)[7]를 하여, 지역정보를 추출하는 구간이다. 모든 영상들은 gaussian curvature 가 가장 센 점들을 중심으로 고정적인 총 30 개의 특징점과 그 주변 밝기 정보들을 SVD 하여, 상위 18 개의 고유값 (eigen-value)을 지역 기술자로 정의한다. 그러

므로, 총 한 영상당 차지하는 지역 기술자의 정보는 총 540 바이트이다.

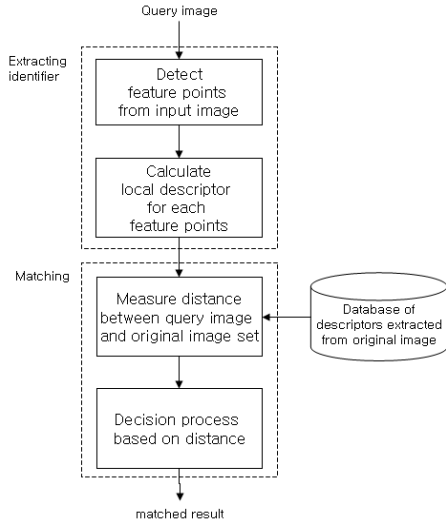


그림 2. 전체 시스템의 블록도

그림 2 는 제안한 시스템의 전체 블록도이다. 첫 과정은 영상 식별자를 추출하는 구간이며, 두번째 과정은 질의 영상으로부터 추출한 식별자만을 가지고, 두 영상간의 거리의 따른, 질의와 원본 데이터베이스에서 어느 원본으로부터 복제된 영상 인지를 판단하는 “true” 또는 “false” 을 결정하는 구간이다.

추출된 영상 식별자를 두 영상간의 거리를 빠르고 효율적으로 계산하기 위하여, penalty algorithm(알고리즘 1)을 제안한다.

```

FUNCTION distance_fp(fpo, fpt, ldd_threshold)
fptnum[1..N] = 0
for j=1..ldoNum,
count = 0;
min_dist = 1000;
min_idx = -1;
for k=1..ldtNum,
cmpdist = localdescriptor_distance( fpo[j], fpt[k] );
if cmpdist < distance_threshold,
if cmpdist < min_dist,
min_dist = cmpdist;
min_idx = k;
end
end
end
if min_idx > -1,
fptnum[min_idx] = fptnum[min_idx] + 1;
end
end
distance = 0;
for i=1..N,
if fptnum[i] > 0,
distance = distance + T / fptnum[i];
end
end
return MAX_DIST-distance;
    
```

알고리즘 1. 거리 계산을 위한 penalty algorithm

여기서, fpo 와 fpt 는 어떠한 두 영상의 image identifier 이고, $ldoNum$ 은 fpo 의 특징점의 개수, $ldtNum$ 은 fpt 의 특징점의 개수이다. $localdescriptor_distance()$ 은 두 local descriptor 간의 거리측정이다.

본 실험에서는 $dist(A, B) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$ 로 두 지역 기술자간의 거리를 측정하였다. $ldd_threshold$ 는 2 개의 지역기술자간의 거리 측정 경계치 값이다. 최종 반환 값은 distance 을 반환하며 이 값이 클수록 원본 영상과 많이 유사하다는 측정을 내릴 수 있으므로, 임의의 정의된 MAX_DIST 로부터 차를 하여, 값이 작을수록 원본의 가깝도록 재 변환을 한다. 또한 T 는 사용자 정의 변수이지만, 반듯이 특징점의 최대 개수보다는 커야 한다. 만약 한 영상당 최대 30 개씩을 특징점으로 뽑았다면, T 는 최소 30 이 되어서 “ $T/fptnum[i]$ ” 가 가분수가 되도록 한다. 실험에서는 $ldd_threshold$ 을 200 으로 하였다.

3. 실험

실험 영상은 총 50,000 장의 영상들로 구성되며, 1 절에서는 임계값(threshold value)을 결정하는 방법과 2 절에서는 임계값과 대응하는 성공률을 설명한다.

3.1 임계값 결정

두 영상이 같다는 판단을 내리기 위한, 임의의 임계값을 결정하기 위하여, 총 10,000 장의 서로 다른 영상들의 거리를 계산을 하였다. 계산된 총 영상 쌍의 개수는 $\frac{10,000 \times (10,000 - 1)}{2} = 49,995,000$ 이다.

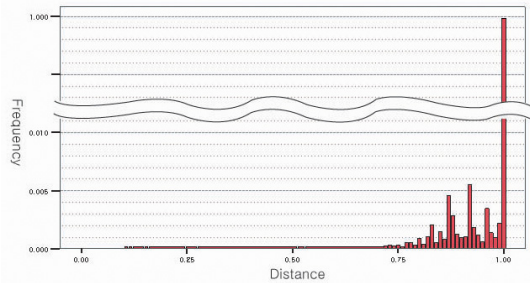


그림 3. 정규화된 거리 히스토그램

그림 3 은 이 거리 히스토그램이며, 거리값을 정규화 하였다. 실험결과 백만 장당 1 장을 오판할 확률인 1ppm(pair per million) false positive alarm 과 대응되는 거리는 0.45 이다.

3.2 정확률

본 실험에서는 2,000 장의 원본과 13 개의 왜곡 항목을 생성하였다. 실험에 사용된 총 왜곡 영상은 $13 \times 2000 = 26,000$ 이다. 이 모든 26,000 개의 영상은 본 시스템의 질의로 사용하였으며, 원본 데이터베이스와 비교하여 거리차가 임계값보다 작으면서, 가장 가까운 것을 원본과 같다고 판단하였다. 표 1 은 실험결과

이다. 실험에서 영상 식별자 추출 시간과 10,000 의 데이터베이스 원본을 찾는 시간은 Pentium 4 GHz 기준 평균 1.2 초였다.

Hammarling, A. McKenney, and D. Sorensen, "LAPACK User's Guide", Third Edition, SIAM, Philadelphia, 1999.

비교적 공통 영역이 많이 있는 왜곡들에서는 매우 높은 성공률을 보여주지만, 공통 면적이 작은 trans_10 과 같은 왜곡에서는 매우 낮은 실패율을 나타냈다.

표 1. 정확률

modification_items	success ratio
bright+5	95.10
bright-5	95.30
cropping_90	98.30
gauss_noise_2.5	98.50
jpg_95	98.60
monochrome	98.70
perspective_-2	98.20
perspective_2	98.20
rotation_10	88.10
scale_90	90.35
skew_-2	98.50
skew_2	98.40
trans_10	44.60
Total	92.37

4. 결론 및 향후 연구

본 논문에서는 제안한 영상 식별자는 매우 높은 정확률과 빠른 검색을 할 수 있다는 것을 실험적으로 증명하였다. 향후, 제안한 영상 식별자를 조금 더 강인한 방법이 필요하며, 추가적인 왜곡 종류를 더 추가할 것이다.

참고문헌

- [1] B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan and Akio Yamada, "Color and Texture Descriptors", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 11, No. 6, pp. 703-715 (2001)
- [2] Seong-O Shim and Tae-Sun Choi, "Edge Color Histogram For Image Retrieval", IEEE ICIP 2002, Vol. 3, pp. 957-960 (2002)
- [3] Michael J. Swian and Dana H. Ballard, "Color Indexing", International Journal of Computer Vision, Vol. 1, No. 7, pp. 11-32 (1991)
- [4] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 20, pp. 91-110 (2003)
- [5] Jin S. Seo, Jaap Haitsma, Ton Kalker and Chang D. Yoo, "A robust image fingerprinting system using the Radon transform", Signal Processing: Image Communication, Vol. 19, No. 14, pp. 325-339 (2004)
- [6] Alexander Kadyrov and Maria Petrou, "The Trace Transform and Its Applications", IEEE Trans. PAMI, Vol. 23, No. 8, pp. 811-828 (2001)
- [7] Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S.