

사용자에 의한 영상 불변 특징을 이용한 이미지 해쉬 기술

최용수*, 김형중*

*고려대학교 정보경영공학전문대학원

e-mail: {ciechoi, khj-}@korea.ac.kr

Image Hashing Techniques Utilizing User-Defined Image Invariant Features

YongSoo Choi*, HyoungJoong Kim*

*Graduate School of Information Management and Security, Korea University

요 약

이미지 해쉬는 기술자(Descriptor) 또는 구분자(Identifier)로서 영상의 유사성을 측정하는데 사용될 수 있다. 수많은 이미지 해쉬 기술들이 있지만, 그 중에서도 히스토그램 기반의 방법들이 일반적인 영상 처리나 다양한 기하학적 공격들에 강인함을 보여준다. 이 논문에서는 강인한 히스토그램 기반 이미지 해쉬를 생성하기 위하여 영상의 양자화, 사용자 지정 윈도우를 적용하여 영상의 특성화 과정을 적용하며 해쉬 값 결정 알고리즘도 오류에 강하도록 설계하였다. 이러한 기술은 기존의 논문들이 보여주었던 성능을 향상시킨다. 특히, 통계적인 오류측정을 통해 수행결과를 설명함으로써 수행성능의 향상을 객관적으로 평가하였다.

1. 서론

이미지 해쉬 기술에서 비슷한 영상들은 각 영상의 주요 특징들이 비슷하게 유지되어 있다고 할 수 있다. 즉, 영상의 주요 특징들이 같으면 그 영상들은 동일한 영상으로 간주될 수 있다. 컬러변환, 컬러 밸런스 또는 대비(Contrast) 변화, 손실 압축, 크기변환, 회전, 절삭, 필터링 그리고 히스토그램 평활화 같은 일반적인 영상 조작(manipulation) 방법들은 영상의 주요 특징들을 거의 틀리게 만들기도 한다. 동일영상이 위와 같은 영상 조작에 의해 생성되는 변형본을 유사 복제(near-duplicate) 영상이라고 한다. 저작권 위반, 침해와 같은 분야에서 원본과 일부 변형된 유사 변형본을 일치(matching)시키는 기술이 필요로 하게 되었다[2] 유사 복제 영상에 대한 많은 연구들이 수행되어왔다[5].

이와 반대로, 실제로 틀린 영상이지만 비슷하게 보이는 경우도 존재한다. 예를 들어, 동일한 배경에 대해 동일한 촬영 지점에서 촬영된 사진은 유사한 특징들을 가지고 있을 것이다. 유사한 사진들이 하나의 폴더 안에 저장되기 위해서는 동일 영상으로 분류되어야 한다. 사진들은 포토 앨범 같은 것을 만들기 위해 이벤트, 사람, 날짜와 시간, 장소 등으로 정렬되어야 한다. 유사한 영상의 그룹은 한번에 주석이 붙여질 수 있다. 그러므로 영상의 주석을 작성하는 문제는 두 개의 과정으로 구성될 수 있다. 1) 의미적, 시각적으로 비슷한 영상을 찾음 2) 영상의 요약어를 통해 검출된 주요 정보(key phrases)를 검색.

현재 많은 이미지 해쉬 기술이 있으며: scale invariant feature transform (SIFT)은 크기변환과 회전 조작에 강인하다. gradient location and orientation histogram (GLOH)은 SIFT 기술자의 확장된 기술로서 히스토그램을 위해 공간적 특성을 더욱 고려하였다[4]. shape context 방법은 형태를 기술하는 방법의 한가지로 형태 유사성을 측정하고 포인트 연관성을 복원하는 것을 가능하게 한다 [1]. spin image는 양자화된 픽셀의 위치와 밝기 값들의 히스토그램을 이용한다[3]. 대부분의 기술들에서 영상의 히스토그램과 경계선(edge)은 색깔, 문양, 국지 형태, 핵심적 점과 지역과 같은 특징들만큼이나 중요한 특징이다. 조명 처리, 크기변환, 색상 변환에 강인한 Geometric centroid 방법도 개발되었으나 이 방법 또한 많은 처리 시간을 필요로 한다.

2. 히스토그램 기반 이미지 해쉬

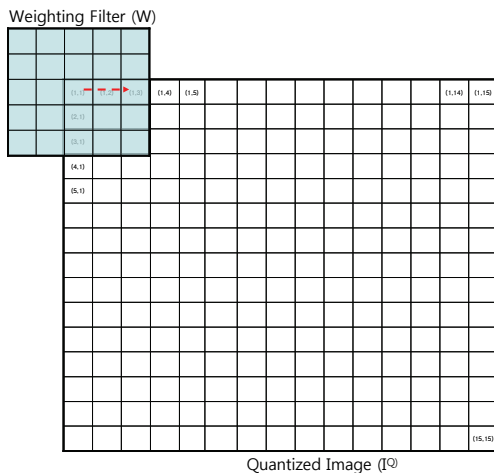
이미지 해쉬 기술이 일반적인 영상 처리 및 기하학적 변형 등에 강인하다할지라도 완벽한 기술은 없다고 가정한다면, 이미지 해쉬 알고리즘의 성능을 보다 향상시키기 위해 다른 영상들 간의 구분과 비슷한 영상들을 식별하는 두 가지 요구사항의 절충점을 찾는 것도 중요하다. 제안된 방법에서는 크게 두 가지의 주요 전처리 과정을 제시한다. 첫째로 입력영상I(실험에 사용된 그레이 영상 데이터베이스)은 N진 표현으로 양자화 된다. 이렇게 양자화된 영상(I^Q)의 밝기(Intensity)는 원래 영상의 기하학적 형태를

유지한다. 둘째로 제공되는 대칭형 가중치 윈도우가 N진 양자화된 영상에 적용(I^W)된다. 이 수직, 수평, 대각으로 대칭적인 윈도우는 기하학적 변화에 강인하기 위한 목적으로 사용된다.

V_6	V_5	V_4	V_5	V_6
V_5	V_3	V_2	V_3	V_5
V_4	V_2	V_1	V_2	V_4
V_5	V_3	V_2	V_3	V_5
V_6	V_5	V_4	V_5	V_6

(그림 1) 대칭형 가중치 윈도우(W)

이 대칭 가중치 윈도우는 영상의 왼쪽 최상위 픽셀부터 오른쪽 최하위 픽셀의 순서로 중심을 이동하며 필터링 연산이 되어 진다. 이 과정을 거치며 영상의 히스토그램들은 각각의 특징적인 형태들을 가지게 된다.

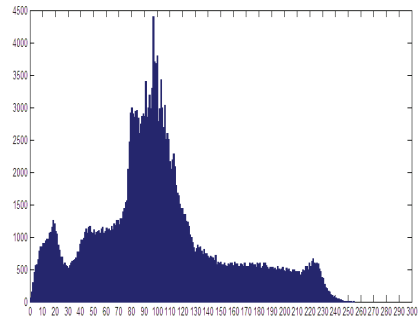


(그림 2) 양자화 영상에 가중치 필터의 적용 순서

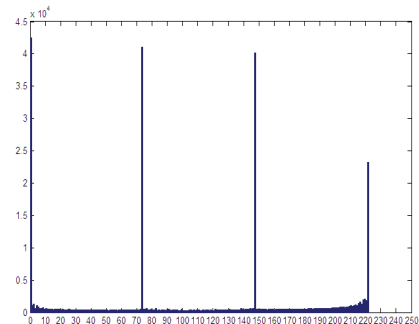
그림 3은 원본 영상의 히스토그램이다. 히스토그램은 특유의 특징이나 형태를 가지지 않는다. 하지만, 논문에서 제안하는 두 가지 과정을 동일한 영상에 대해 적용한 후 히스토그램을 구한다면 영상에 대한 특징으로서 N개의 Peak와 나머지 히스토그램 값들을 얻을 수 있다[그림 4]. I^Q 는 0~3의 숫자로 이루어져 있고 W의 합은 $\sum W$ 이므로, I^W 의 픽셀은 최소 0, 최대 $\sum W*(N-1)$ 를 가진다. 아래의 과정들은 데이터베이스의 모든 영상을 읽을 때까지 반복된다.

- [1] 영상 데이터베이스에서 원본 영상을 읽음
- [2] 영상크기를 지정된 크기로 변환
 - 가. 지정된 왜곡 공격(왜곡의 종류만큼 왜곡영상을 생성)을 수행

- [3] 영상(원본과 왜곡 영상들)을 양자화시켜 N진 표현 형식으로 저장
- [4] 원본 및 왜곡 영상들에 대칭 가중치 윈도우를 한 픽셀씩 이동하며 적용



(그림 3) 원본 영상의 히스토그램



(그림 4) 전처리 후 영상의 히스토그램

- [5] 모든 영상들의 히스토그램을 연산
 - 가. 지정된 만큼의 히스토그램 값들을 합하여 히스토그램 Bin의 빈도를 연산
- [6] N개의 Peak값을 가지는 Bin을 제거(본 논문에서는 N진 표현을 사용하므로 Peak는 N개임)
- [7] 해쉬 생성 알고리즘에 따라 해쉬 열을 생성
- [8] 원본 영상과 각 왜곡 영상과의 해밍 거리를 측정한다. 왜곡의 종류가 N_d 라면 N_d 개의 해밍거리가 생성

모든 영상들에 대해 위에서 제시한 8개의 단계를 거치면, 각 왜곡 공격별로 해밍 거리의 히스토그램을 계산하는 것이 가능하다. 과정 5에서 얻어진 히스토그램 값들은 M개의 독립된 히스토그램 bin(bin)으로 묶여진다. 히스토그램 bin이 M개인 경우, 하나의 영상에서 생성되는 해쉬 열의 길이는 $\frac{n_{bin} - N \times (n_{bin} - N - 1)}{2}$ (단, N진 양자화를 적용한 경우)이다. 해쉬 생성을 위해 Bin Population을 이용하는 것은 이전의 논문(Xiang et. al)을 따른다[8]. 하지만 더 세밀한 해쉬 값 결정 알고리즘[그림 5]을 개발함으로써 해쉬 생성과정에서 발생할 어려들을 감소시킨다.

원본영상과 왜곡 영상의 해쉬 값의 차이에 대한 히스토그램이 모두 얻어지면 임계값(Threshold)을 정함으로서

제 1종 오류(Type I error)를 측정할 수 있다. 제 2종오류 (Type II error)를 측정하기 위하여 서로 다른 원본영상 각각들 간의 해밍거리를 측정할 필요가 있다. 여기서 통계적 가설검정을 적용하며 귀무가설(H_0)은 차이를 구하는 두 영상이 같다고 가정 할 경우, 대립가설(H_1)은 두 영상이 틀리다고 가정 할 경우를 전제로 한다.

```

for i=1:m
  for j=(i+1):n
    pp = freq(i) / (freq(i)+freq(j));
    stat = 2*sqrt(nn) * (pp-0.5);
    if (abs(stat) > th)
      hash(k) = sign(stat);
    else
      hash(k) = 0;
    end;
    k = k+1;
  end;
end;

```

(그림 5) 이미지 해쉬 값 결정 루틴

영상 원본영상의 개수가 n개라면, 하나의 영상은 n-1개의 영상들과 상호배타적인 영상의 쌍이 되어 해밍거리를 구해가 된다. 즉, 영상 데이터베이스를 통틀어 $\frac{n \times (n-1)}{2}$ 개의 해밍거리가 구해진다. 최종적으로 해밍거리 집합의 히스토그램을 계산한 후, 제 1종 오류에서 지정한 임계치를 적용하여 제 2종 오류를 구한다. 제 1,2종 오류의 임계치는 오류의 범위를 정하는 것이 우선이나 본 논문에서는 임의적인 해밍거리 임계치를 설정한 후 두 종류의 오류를 측정하였다.

3. 영상 데이터베이스를 이용한 실험 및 결과

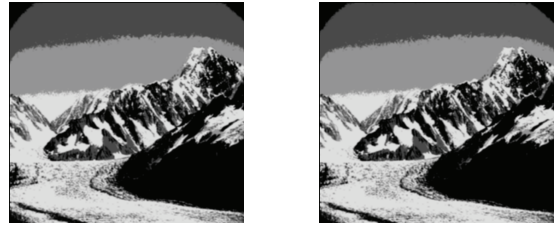
그림에서는 이미지 해쉬 왜곡 처리 중에서도 강력한 공격으로 꼽히는 히스토그램 평활화를 왜곡 공격 종류로



(그림 6) 원본 영상과 왜곡 영상(히스토그램 평활화)



(그림 7) 원본영상과 왜곡 영상의 차



(그림 8) 전처리 과정을 거친 원본 영상과 왜곡 영상



(그림 9) 전처리 과정 거친 원본영상과 왜곡 영상의 차

선택하여 예를 보였다. 원본 영상과 왜곡 영상의 차 영상이 영상의 형태를 모두 보존하고 있을 정도로 왜곡이 큰 것을 그림 7을 통해 볼 수 있다. 이 논문에서 제안한 양자화 방법과 대칭형 가중치 윈도우를 적용한 후 차영상을 구하면 상당부분의 차이 값이 0에 가깝게 표현됨을 볼 수 있다. 즉, 일반적인 영상을 이용하여 이미지 해쉬를 생성하게 되면 동일 영상일지라도 작은 왜곡을 받은 경우 틀리게 판단한 원인이 커지게 된다. 하지만, 그림 9에서 보인바와 같이 제안한 사용자 정의 전처리 과정을 이용하여 변환된 형태의 영상을 얻으면 왜곡 후에도 거의 동일한 영상이 얻어진다. 시각적으로 비슷한 영상은 히스토그램을 통한 해쉬 생성에서도 동일한 해쉬를 생성해 줄 것이라는 가능성을 의미한다.

표1과 표2를 통해 실제 실험 결과를 본다. SVD 실험의 경우 참조논문에서 설정한 임계치 값을 사용하였으며 [9] 제안된 논문의 경우 30, 35의 임계치를 각각 적용하였다. 표 1에서는 왜곡 방법들에 따른 오류를 측정하였으며 잡음 첨가 및 중간값 필터링 왜곡에서 SVD에 비해 성능 저하가 많이 일어났다. 하지만 아래의 표 2에서는 영상 데이터베이스의 모든 영상들에 대해 모든 가능한 영상의 쌍을 구성하여 오류를 측정하였으며 SVD 알고리즘에 비해 뛰어난 2종 오류의 감소가 나타난다. 즉, SVD 알고리즘은 영상의 왜곡 발생 시에도 원본 영상과 비슷한 해쉬 열을 생성해 내지만 완전히 다른(비슷한 특징들을 가진)영상을 구분해 내는 능력은 떨어짐을 알 수 있다.

4. 결론 및 향후 연구

본 논문에서는 히스토그램 기반의 이미지 해쉬 생성 기술에 대한 정확도 향상을 목표로 하였다. 사용자 지정 양자화 수준과 대칭형 가중치 윈도우를 적용하는 전처리 과정을 통하여 원본의 시각적 특징은 보존한 채 영상의 픽셀 값을 새로운 영역으로 변환하였다. 변환된 영상에서

언어진 히스토그램에 이미지 해쉬 생성 알고리즘을 적용하였고 여기서 언어진 이미지 해쉬들을 실험에 적용하였다. 원본과 왜곡 영상의 유사성 판별, 다른 영상들 간의 구분성 판별이라는 두 가지 환경에 대해 실험을 수행하였다. 제안한 알고리즘은 통계학에서 이용하는 제 1종 오류에서는 기존의 고성능 알고리즘인 SVD에 비해 일부 낮은 성능을 보이며 특히, 잡음과 같은 왜곡에서 90%가 넘는 오류를 보이지만 잡음 제거 필터 등을 이용한 알고리즘 개선을 통해 상당부분 개선이 가능하다. 제 2종 오류에 대해서는 탁월한 성능을 보임을 확인하였으므로 논문의 목적이었던 제 1종, 2종 오류의 절충점을 성공적으로 찾을 수 있음을 보였다.

후향 가설검증 단계에서 제 1종, 2종 오류를 측정함에 있어 일반적인 통계검증과 같이 사용하여 일반화된 오류를 측정할 필요가 있으며 잡음 제거 필터링 등의 저주파 통과 필터를 추가 적용할 경우 알고리즘 성능의 개선이 기대된다.

<표 1> 두 가지 임계치를 적용한 경우 제 1종 오류의 측정 비교

Type of Distortion	Proposed		SVD
	Th_Proposed		Th_SVD
	30	35	0.02
JPEG 70	14.32 %	10.74 %	0 %
JPEG 80	8.87 %	6.39 %	0 %
JPEG 90	2.39 %	1.62 %	0 %
JPEG 100	0 %	0 %	0 %
Noise Addition 0.001	94.97 %	92.16 %	3.24 %
Noise Addition 0.005	94.71 %	92.24 %	3.50 %
Noise Addition 0.01	94.54 %	91.73 %	3.75 %
Rotation 1	0.5 %	0.17 %	0.43 %
Rotation 2	1.36 %	0.68 %	1.53 %
Rotation 5	12.79 %	6.65 %	8.27 %
Rotation 10	44.25 %	31.29 %	26.34 %
Histogram Equalization	0.6 %	0.17 %	10.66 %
Cropping 10	20.12 %	12.70 %	12.45 %
Cropping 25	9.21 %	4.35 %	1.96 %
Median 3X3	39.81 %	26.68 %	0 %
Median 5X5	73.74 %	58.91 %	0.26 %

<표 2> 두 가지 임계치를 적용한 경우 제 2종 오류의 측정 비교

	Proposed		SVD
	Th_Proposed		0.02
	30	35	
1173 Images	2.09 %	4.13 %	25.73 %

감사의 글

본 연구는 지식경제부 및 한국산업기술평가관리원의 IT 산업원천기술개발사업의 일환으로 수행하였음.(과제번호: KI001917, 익명성 기반의 u-지식 정보보호기술개발)

참고문헌

- [1] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.4, pp.509-522, 2002.
- [2] J. Fridrich, D. Soukal, and J. Lukas, "Detection of copy move forgery in digital images," Proceedings of Digital Forensic Research Workshop, 2003.
- [3] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.21, no.5, pp.433-449, 1999.
- [4] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," Proceedings of the Computer Vision and Pattern Recognition, pp.257-263, 2003.
- [5] N. Shivakumar and H. Garcia-Molina, "Finding near-replicas of documents and servers on the web," Proceedings of Workshop on Web Databases, pp.204-212, 1998.
- [6] A. W. M. Smeulders, M. Worring, S.Santini,A.Gupta, and R.Jain,"Content based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, pp.1349-1380, 2000.
- [7] M. J. Swain and B. H. Ballard, "Color indexing," International Journal on Computer Vision, vol.7, no.1, pp.11-32, 1991.
- [8] S. Xiang, H. J. Kim, and J. Huang, "Histogram-based image hashing scheme robust against geometric deformations," Proceedings of the ACM Multimedia and Security Workshop, pp.121-128, 2007.
- [9] Suleyman S. Kozat, and Ramarathnam Venkatesan, M.Kivanc Mihcak,"Robust Perceptual Image Hashing via Matrix Invariants," Proceedings of International Conference on Image Processing, pp.3443-3446, 2004