JPEG 압축 및 공모공격에 강인한 디지털 이미지 핑거프린팅 기술

*김광일 **김종원 ***최종욱

상명대학교

{*kikim, **jwkim, ***juchoi}@smu.ac.kr

Digital Image Fingerprinting Technique Against JPEG Compression and Collusion Attack

*Kim, Kwang-Il **Kim, Jong Weon ***Choi, Jong Uk

Sangmyung University

요약

디지털 핑거프린팅(Digital Fingerprinting)은 기밀 정보를 디지털 콘텐츠에 삽입하는 측면에서는 디지털 워터마킹과 동일하다고 볼 수 있으나 저작권자나 판매자의 정보가 아닌 콘텐츠를 구매한 사용자의 정보를 삽입함으로써 콘텐츠 불법 배포자를 추적할 수 있도록 한다는 점에서 워터마킹과 차별화된 기술이다. 이러한 핑거프린팅 기술은 소유권에 대한 인증뿐만 아니라 개인 식별 기능까지 제공해야 하므로 기존의 워터마킹이 갖추어야 할 요구사항인 비가시성, 견고성, 유일성과 더불어 공모 허용, 비대칭성, 익명성, 조건부 추적성 등이 부가적으로 필요하다. 본 논문에서는 행렬의 한 열을 선택 후 쉬프팅 기법을 사용하서 사용자 정보로 조합하여 핑거프린트를 생성하였다. 이렇게 생성된 핑거프린트 정보를 2레벨 웨이블릿 변환 영역 중 LH2, HL2, HH2 부대역에 삽입하였다. 쉬프팅 정보와 도메인 개념을 사용하여 보다 많은 사용자에게 핑거프린트 정보를 삽입할 수 있으며, 공모공격과 JPEG 압축에서도 최소한 1명 이상의 공모자를 검출할 수 있는 핑거프린팅의 기본 조건을 만족 하였다.

1. 서론

최근 개인용 컴퓨터의 보급 확대와 디지털 시대를 맞이하여 다양한 디지털 저작물들이 나오고 있으며, 인터넷의 발달로 멀티미디어 데이터양이 급격히 늘어나고 있다. 그러나 디지털 저작물의 특징상 누구나 손쉽게 무제한으로 복제와 유통이 가능하며, 원본과 복제본과의 구분 또한 불가능하다는 문제점을 가지고 있다. 따라서 불법적인 복제를 막고, 저작물에 대한 저작권, 소유권을 보호하는 것이 필요하다.

디지털 핑거프린팅(Digital Fingerprinting)은 기밀 정보를 디지털 콘텐츠에 삽입하는 측면에서는 디지털 워터마킹과 동일하다고 볼 수 있으나 저작권자나 판매자의 정보가 아닌 콘텐츠를 구매한 사용자의 정보를 삽입함으로써 콘텐츠 불법 배포자를 추적할 수 있도록 한다는 점에서 워터마킹과 차별화된 기술이다. 이러한 핑거프린팅 기술은 소유권에 대한 인증뿐만 아니라 개인 식별 기능까지 제공해야 하므로 기존의 워터마킹이 갖추어야 할 요구사항인 비가시성, 견고성, 유일성과 더불어 공모 허용, 비대칭성, 익명성, 조건부 추적성 등이 부가적으로 필요하다.

핑거프린팅 기술은 워터마킹의 확장된 형태의 기술로써 콘텐츠에 구매자 정보를 삽입하여 불법 배포된 콘텐츠에 대해서는 삽입한 구매자 정보를 추출함으로써 불법 배포가 이루어진 원천지(Source)를 추적할 수 있도록 해 준다.

워터마킹 방법은 한 개인의 단독적인 공격에는 강인하지만 저작권 정보만이 들어가기 때문에 디지털 콘텐츠가 불법 배포되었을 때는 배포자를 추적할 수 없다. 또한 워터마크로써 구매자 정보를 삽입한다고 할지라도 여러 사람이 공모공격(collusion attack)을 가했을 때는 삽입한 워터마크가 모두 손실되기 때문에 삽입한 구매자 정보 및 저작권 정보를 추출할 수 없게 된다.

공모공격은 주로 구매자 정보를 삽입할 때 사용되었던 워터마크와의 상관도 값이 작게 나오도록 하는 방법이 주를 이루고 있고, 다음과 같은 종류의 공모공격이 알려져 있다. 이 공모공격에는 평균(average)공격, 모자이크(mosaic)공격, 최대-최소(Max-Mix)공격[1], 상관계수 음수화(Negative Correlation)공격[1], 상관계수 제로화(Zero Correlation)공격법[2]이 있고, 각각의 공모공격은 하나의 콘텐츠를 서로 다른 ID를 가진 사람들이 내려 받고 각각 소유한 콘텐츠를 공모함으로써 수행된다.

핑거프린팅 기술에 대한 현재까지의 연구를 살펴보면 크게 기존의 워터마킹 기술을 이용하는 기법[3]과 삽입하는 코드 자체가 공모공격에 강인하도록 설계하는 공모보안 코드 방법[4~7]으로 분류할 수 있다.

2. 연구 내용

본 논문에서는 워터마킹 기술을 활용한 핑거프린팅 기술을 개발하기 위하여 2래벨 웨이블릿 변환 영역에 사용자 정보를 갖는 핑거프린트 정보를 삽입하게 된다. 그림 1은 2레벨 웨이블릿 변환을 수행 후 얻어진 주파수 성분의 데이터를 보여주는 것이다. 이 부대역 중에서 비가시성과 강인성을 모두 충족하는 대역을 선택하여 핑거프린트 정보를 각각 삽입하였다.

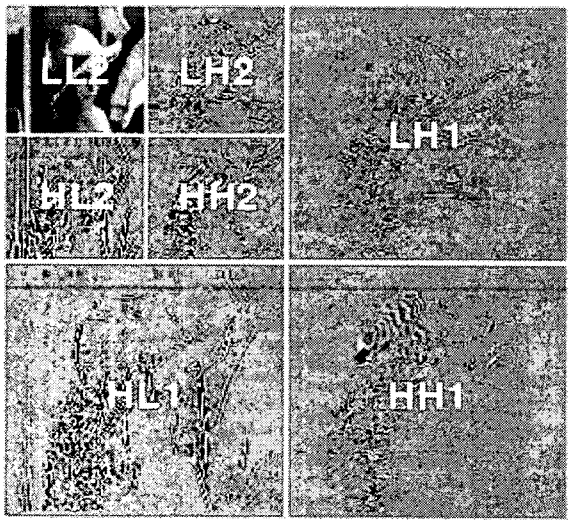

그림 1. 2레벨 웨이블릿 변환 영역

선택된 대역에 삽입할 쉬프트 워터마크는 삽입하고자 하는 사용자 정보를 식(1)에 대입하여 행과 열을 계산하게 된다. 계산된 행과 열의 크기만큼 워터마크를 쉬프트하여 삽입하게 된다. 이렇게 쉬프트된 워터마크는 추출 시 그림 2와 같이 쉬프트된 워터마크 위치 만큼 이동하여 피크가 나타나므로 그 위치 점을 찾으므로 해서 사용자 정보를 얻을 수 있게 된다.

$$S_c = Mod(Userindex, B_L/B_M) \times B_M + B_M/2 \qquad (1)$$
$$S_r = \left\lfloor \frac{Userindex}{(B_L/B)} \right\rfloor \times B_M + B_M/2$$

Sc와 Sr은 각각 이동하게 될 행과 열의 크기를 나타내며, Userindex는 삽입할 사용자 정보를 의미하고, BL은 한 블록의 크기(쉬프팅 하는 워터마크의 크기)이다. BM은 일종의 △성분으로 BM×BM 영역 안에 나타나는 모든 피크를 동일한 정보로 담고 있는 피크로 취급하기 위하여 설정한 것이다. ⌊ ⌋ 는 입력 값보다 작은 가장 가까운 정수 값을 계산한다.

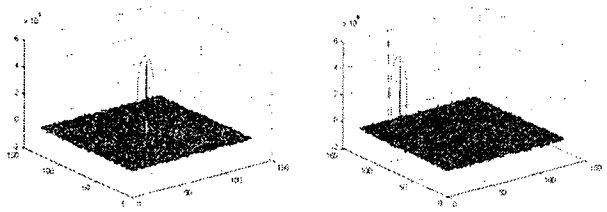

그림 2. 워터마크 쉬프트의 원리

위와 같은 방법으로 워터마크를 쉬프팅 한 후 선택된 웨이블릿 영역에 쉬프트 워터마크를 삽입하게 된다.

가. 핑거프린트 생성 및 삽입

본 논문에서는 핑거프린트를 생성하기 위해 다음과 같은 절차를 수행하게 된다.

1. 저작권 정보를 부호화한다.
2. 부호화된 저작권 정보를 기초로 16384*128 크기의 Domain 행렬을 생성한다.
3. 구매자의 도메인(1~128)을 결정하고 결정된 도메인 번호에 해당하는 Domain 행렬의 열을 선택한다.(예, 도메인을 3으로 선택했다면 Domain 행렬의 3번째 열을 선택)
4. 결정된 도메인의 열을 이용하여 128×128 도메인 인덱스 행렬을 생성한다. (예, 3번째 열의 1×16382 배열을 128×128 행렬로 변환 생성)
5. 도메인 내에서 사용자 정보를 인덱스화하여 식(1)에서와 같이 쉬프팅 할 행과 열을 지정한다. 지정된 행과 열만큼 쉬프팅하여 사용자 정보를 나타낸다.

Domain 행렬은 Malvar에 의해 제안된 듀얼 워터마킹/핑거프린팅(dual WM/FP)[3]에서 제안된 워터마크를 숨기는 반송 신호와 같은 방법으로 생성된 M×N의 2차원 행렬로, 표준편차가 B인 정규분포를 따른다.

이와 같이 생성한 핑거프린트는 웨이블릿을 통해 얻어진 주파수 영역 중 특정 부대역을 선택하여 삽입된다. 그림 3은 제안된 핑거프린트 삽입 알고리즘 구조를 나타낸 것이다.

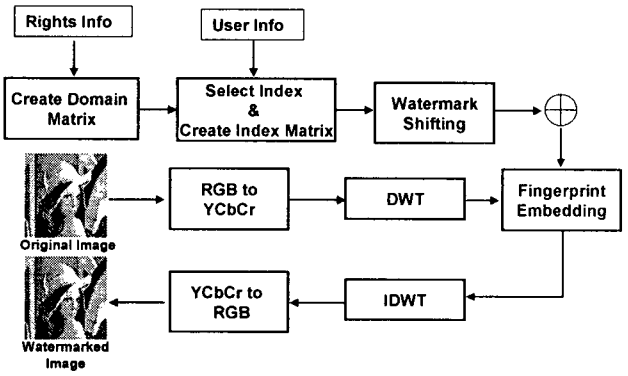

그림 3. 핑거프린트 삽입 알고리즘

입력된 원 영상을 휘도와 색차 성분으로 분해 후, 휘도 성분을 추출하여 휘도 성분에 2레벨 웨이블릿 분해를 실시한다. 앞 절에서 설명한 바와 같이 핑거프린트를 생성하여 쉬프트 시킨 후 LH2, HL2, HH2 영역에 삽입하게 된다.

나. 핑거프린트 추출

그림 4는 핑거프린트 추출 알고리즘을 나타내는 것으로 Preprocess 부분과 Detect precess 두 부분으로 나누어져 있다.

Preprocess 단계에서 좀 더 높은 상관도를 얻기 위해 위너필터를 사용하였다. 이 방법은 위너필터를 이용해 워터마크를 삽입하기 전의 영상을 예측하는 방법으로 위너 필터는 주로 잡음을 제거하기 위한 방법으로 많이 활용되었다. 워터마크된 이미지를 위너 필터링하고, 워터마크된 이미지와 이 필터링된 이미지와의 차 영상을 구하는 Pre-precess 절차를 수행 후 다음에서 설명할 Detect precess 절차를 실시하게 된다.

Detect precess 단계에서는 Pre-precess 절차에서 구한 차 영상을 2레벨 웨이블릿 분해하여 분해된 부대역중 LH2, HL2, HH2 영역의 계수들을 합한다. 이 계수들의 합과 삽입과정에서 생성한 Domain 행렬의 1열부터 128열까지 각각의 모든 열들을 삽입과정 3 에서와 같이

128×128 행렬로 변환하여 상관도를 구하게 된다. 일정 임계치 이상의 값이 추출되면, 추출됐을 시 사용되었던 행렬을 만든 Domain의 열 번호를 먼저 구하고, 이 피크의 쉬프트된 값을 계산한다. 이 두 값, 즉 도메인을 나타내는 행렬의 열 번호와 쉬프트된 값을 계산해서 얻을 정보를 조합하여 사용자 정보를 얻게 된다.

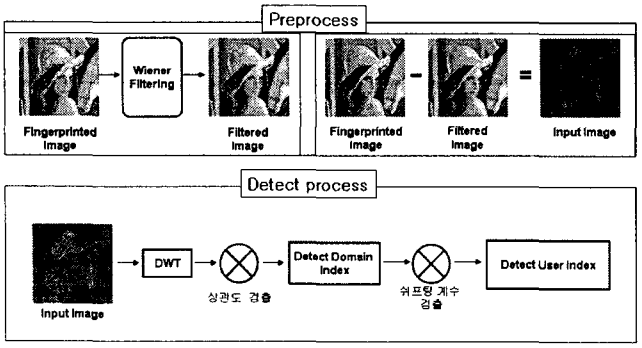

그림 4. 핑거프린트 추출 알고리즘

3. 실험 결과

본 논문에서는 제안한 알고리즘의 성능을 측정을 위하여 512×512 영상 5개(*lena, airplane, girl, milk, pepper*)를 실험하였고, 공모 공격에 대한 강인성을 확인하기 위하여 평균 공모공격, 모자이크 공모공격을 실시하였다. 공모자의 수는 최대 20명으로 설정하였으며, 각 공모자 수에 따른 JPEG압축 강인성 성능측정도 함께 실시하였다.

가. 평균 공모공격

표 1은 평균 공모공격에 대해 검출된 공모자의 수를 공격한 공모자의 수와 JPEG Quality Factor에 변화를 주면서 실시한 실험의 결과이다. 공모자의 수는 다섯 명에서 스무 명까지 2~3 명씩 증가시키면서 실험을 하였고, JPEG Quality Factor는 90에서 30까지 10~20 %씩 감소시키면서 실험을 하였다.

JPEG을 적용하지 않았을 때의 공격은 최대 20명의 공모자가 공격했을 시에도 10명 이상의 공모자를 검출하였다. 또한 Quality Factor 60%까지는 17명의 공모자 중 핑거프린팅 기술의 요구 조건중의 하나인 최소한 한명 이상의 공모자를 검출할 수 있었다.

표 1. 평균 공격에 대한 공모자 검출 실험

lena, airplane, girl, milk, pepper (PSNR:38.46db)						
JPEG / 공모자수	90	80	60	50	40	30
5	5	5	5	5	5	4.6
7	7	7	7	7	7	5.6
10	10	10	8.6	5	2.4	0
13	12.6	9.2	2.8	1	0.2	0
15	14.6	11	3	0.8	0.4	0
17	16.2	12.2	2.4	1.2	0.4	0
20	11.2	3.6	0.2	0.4	0	0

나. 모자이크 공격

표 2는 모자이크 공격에 대해서 검출된 공모자의 수를 공격한 공모자의 수와 JPEG Quality Factor에 변화를 주면서 실시한 실험의 결과이다. 압축을 하지 않았을 때는 최대 16명의 공모자중 평균 8.6명을 검출하여 50% 이상의 검출률을 기록하였고, 50%이상의 JPEG Quality Factor에서도 최대 16명의 공모자중 최소한 1명 이상의 공모자를 검출할 수 있어 핑거프린팅 기술의 요구 조건을 만족하였다.

표 2. 모자이크 공격에 대한 공모자 검출 실험

lena, airplane, girl, milk, pepper (PSNR:38.46db)						
JPEG / 공모자수	90	80	60	50	40	30
2	2	2	2	2	2	2
4	4	4	4	4	4	4
6	6	6	6	6	6	6
8	8	8	7.8	7.6	6.2	3
16	8.6	4.2	1.6	1.2	0.2	0.2

4. 결론

본 논문에서는 저작권 정보를 부호화한 Domain 행렬을 생성한 다음, 이 행렬의 한 열을 선택하여 하나의 새로운 행렬을 만들어 쉬프팅 기법을 활용, 사용자 정보로 조합하여 핑거프린트를 생성하였다. 이렇게 생성된 핑거프린트 정보를 2레벨 웨이블릿 변환 영역 중 LH2, HL2, HH2 부대역에 삽입하였다. 추출과정에서는 Domain 행렬의 열에 해당하는 인덱스를 검출하고, 쉬프팅된 정보를 조합하여 사용자 정보를 얻게 된다. 쉬프팅 정보와 도메인 개념을 사용하여 보다 많은 사용자에게 핑거프린트 정보를 삽입할 수 있었다. 평균화 공격의 경우 최대 17명까지의 공모 공격에 대해 JPEG Quality Factor 60% 이상까지는 최소한 1명 이상의 공모자를 검출할 수 있는 핑거프린팅의 기본 조건을 만족 하였다. 또한 모자이크 공격의 경우 최대 16명의 공모자중 JPEG Quality Factor 50% 이상까지는 최소한 1명 이상의 공모자를 검출할 수 있는 핑거프린팅의 기본 조건을 만족 하였다.

[참고문헌]

[1] H. S. Stone, "Analysis of Attacks On Image Watermarks with Randomized Coefficients, " NEC Res. Inst., Tech. Rep. 96-045, 1996

[2] V. Wahadaniah, Y. L. Guan, and H. C. Chua, "A New Collusion Attack and Its Performance Evaluation, "Proceedings of IWDW, 2002, pp. 88-103. 2002.

[3] D. Kirovski, H. S. Malvar, and Y. Yacobi, "Multimedia Content Screening Using a Dual Watermarking and Fingerprinting System, " ACM Multimedia, 2002.

[4] D. Boneh and J. Shaw, "Collusion-Secure Fingerprinting for Digital Data," IEEE Trans. Inf. Theory, vol. 44, no. 5, pp. 1897-1905, Sept., 1998.

[5] J. Dittmann, "Combining Digital Watermarks and Collusion Secure Fingerprints for customer Copy Monitoring," Proc. IEE Seminar Sec. Image & Image Auth, pp. 128-132, Mar., 2000.

[6] J. Domingo-Ferrer and J. Herrera-Joancomart, "Simple Collusion-secure Fingerprinting Schemes for Images," in IEEE International Conference on Information Technology: Coding and Computing, ISBN -7695-0540-6, pp. 128-132. 2000.

[7] W. Trappe, M. Wu, and K. J. R. Liu, "Collusion Resistant Fingerprinting for Multimedia," Proc, of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing(ICASSP. 02), vol. IV, pp. 3309-3312. M