왜곡 공격에 강인한 디지털 워터마크 분할 삽입 기법

송채원[1], 박소현[2], 이일구[3]
[1]성신여자대학교 융합보안공학과 학부생
[2]성신여자대학교 미래융합기술공학과 박사과정
[3]성신여자대학교 융합보안공학과/미래융합기술공학과 교수
20221106@sungshin.ac.kr, 220227022@sungshin.ac.kr, iglee@sungshin.ac.kr

Robust Digital Watermark Segmentation-based Embedding Techniques against Distortion Attacks

Chae-Won Song[1], So-Hyun Park[2], Il-Gu Lee[1,2]
[1]Dept. of Convergence Security Engineering, Sungshin Women's University
[2]Dept. of Future Convergence Technology Engineering, Sungshin Women's University

요 약

최근 디지털 워터마킹 기술은 디지털 콘텐츠의 저작권 보호 및 추적을 위해 활용되고 있다. 그러나 종래의 워터마킹 기술은 이미지에 워터마크 이미지 전체를 삽입하기 때문에 왜곡 공격에 취약하다. 이러한 문제를 해결하기 위해 본 연구에서는 워터마크 분할 삽입 기법을 제안하였다. 워터마크 분할 삽입 기법을 사용하면 종래 방법 대비 20%p의 손실률이 증가하더라도 원본 워터마크를 복구할 수 있어 1.5배 향상된 성능을 보인다.

1. 서론

디지털 워터마킹 기술은 멀티미디어 콘텐츠의 저작권을 보호하기 위해 정보를 삽입하고 삽입한 정보를 검출하는 기술이다 [1]. 디지털 콘텐츠 수요가 증가함에 따라 글로벌 디지털 워터마크 기술 시장은 2031년까지 9,758만 달러에 이를 것으로 예상된다 [2]. 그러나, 단순한 왜곡 공격에도 워터마킹 기능이 무력화될 수 있다.

본 연구에서는 크롭(crop), 블러링(bluring)과 같이 워터마크 왜곡 공격에 강인한 디지털 워터마크 분할 삽입 기반의 워터마크 복원 기법을 제안한다. 또한, 워터마크를 무력화하는 공격에 대해 분할 삽입 기반의 디지털 워터마크 복구 성능을 평가한다.

2. 워터마크 분할 및 중복 삽입 기법

본 연구에서는 디지털 워터마크 분할 삽입 방식을 활용하여 워터마크 왜곡 공격에도 원본 워터마크를 복구할 수 있는 방식을 제안한다.

2.1 워터마크 분할 및 중복 삽입 기법

종래의 방법은 워터마크가 삽입된 이미지는 큰 블록 단위의 반복이므로, 이미지에 일부 공격을 가하면 전체 워터마크가 정상적으로 추출되지 않는다. 본 연구에서는 원본 워터마크를 복수 개의 워터마크 청크로 분할하고, 동일한 워터마크 청크를 중복 삽입하여 워터마크 무력화 공격 상황에도 중복 삽입된 최소 개수의 워터마크 청크를 추출하여 원본 워터마크를 복원하는 기법을 제안한다. 그림 1은 워터마크 삽입과 크롭 공격 및 추출 과정을 나타내는 흐름도이다.

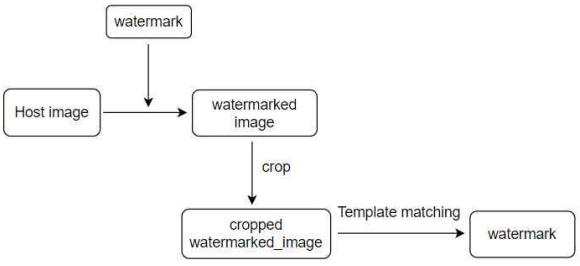

그림 1 Structural diagram of split insertion-based watermarking

원본 워터마크를 여러 조각으로 분할하고 중복률을 설정하여 분할된 워터마크를 호스트 이미지에 중복하여 삽입한다. 중복률이란 워터마크를 중복해서 삽입하는 비율을 의미하고, 본 실험에서는 30으로 설정하여 원본 워터마크를 30번 중복 삽입하였다. 워터마크 삽입하기 위해 서로 다른 이미지를 합성할 때 사용하는 Alpha Blending 알고리즘을 사용하였다. 각 이미지에 워터마크를 임베딩할 때 워터마크

임베딩 강도인 알파 값을 설정하여 호스트 이미지에 분할한 워터마크 청크를 삽입했다. 알파 값이 0에 가까울수록 삽입한 워터마크 이미지의 임베딩 강도가 강해진다. 워터마크 청크가 서로 겹치지 않도록 랜덤한 위치를 설정하여 삽입하였다.

2.2 실험 환경

워터마크 청크를 중복 삽입한 호스트 이미지에 크롭 공격을 시도했을 때 크롭 비율에 따른 워터마크 복원 성능을 평가하였다. 크롭이란 호스트 이미지의 일부를 잘라 내는 것을 의미한다. 워터마크 추출 과정에서는 워터마크 청크를 중복으로 삽입한 호스트 이미지에서 워터마크 청크와 일치하는 이미지의 영역을 찾는 Template Matching 알고리즘을 사용하였다.

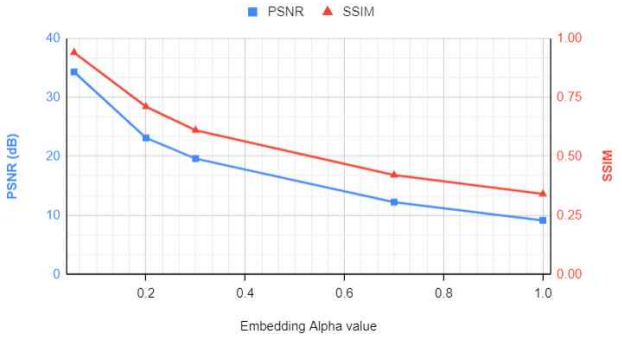

그림 2 Image similarity according to embedding ratio

이 실험을 위해 100픽셀 크기의 워터마크 이미지를 4X4, 즉 16조각으로 나누고 중복률을 30으로 설정하여 총 480개의 워터마크 청크를 호스트 이미지에 랜덤하게 삽입하였다. 임베딩 강도 설정 값인 알파를 0.055로 설정하고 호스트 이미지와 워터마크 분할 삽입을 적용한 호스트 이미지 간의 이미지 유사도를 측정하였다. 그림 2는 워터마크 임베딩 강도에 따른 호스트 이미지와 워터마크 임베딩된 이미지 간의 이미지 유사도를 보여준다. 이미지 유사도는 PSNR(peak signal-to-noise ratio), SSIM(structural similarity index measure)로 평가하였다. 임베딩 강도를 0.055로 하여 480개의 워터마크 조각을 삽입했을 때, PSNR이 34.3dB로 측정되고, SSIM은 0.94로 측정되어 비가시성과 이미지 간의 유사성을 모두 보장할 수 있음을 알 수 있었다.

2.3 실험 결과

분할 워터마크가 삽입된 호스트 이미지의 30%에서 80%까지 10% 단위로 크롭 비율을 늘리며 공격을 시도하였다. 표 2에서는 이미지 크롭 비율(%)에 따른 종래 방식과 제안 방식의 워터마크 복원 능력을 평가하였다. 복원 가능 능력을 평가할 때 20번의 실험을 진행하고 워터마크 평균 추출 값을 도출하였다. 종래 방식은 워터마크를 삽입한 호스트 이미지의 손실률이 40%를 초과하면 원본 워터마크를 추출할 수 없다. 그러나 본 논문에서 제안하는 방식은 워터마크를 분할하고 중복하여 호스트 이미지에 삽입함으로써 공격이 이루어져 최대 60%의 손실이 발생하여도 원본 워터마크를 복원할 수 있다. 또한, 70%, 80%의 크롭이 이루어진 경우, 16개의 워터마크 청크를 모두 추출할 수 없지만 평균적으로 각각 14개, 13개의 워터마크를 추출할 수 있었다. 워터마크의 일부가 손실되면 워터마크 전체를 복구할 수 없는 종래의 방식과 비교하면 워터마크의 87%, 81%를 추출할 수 있다.

category	Crop ratio of images [%]					
	30	40	50	60	70	80
Conventional	O	O	X	X	X	X
Proposed	O	O	O	O	X	X

표 2 Comparison of watermark restoration possibilities between conventional and proposed methods according to image cropping ratio (%)

3. 결론

본 연구에서는 워터마크 왜곡 공격에 강인한 워터마크 분할 삽입 기반의 오류 정정 및 워터마크 추출 기법을 제안하였다. 호스트 이미지의 일부를 잘라내는 크롭 공격을 시도했을 때, 제안하는 방법은 종래 방법 대비 20%p 손실률이 증가하더라도 워터마크 추출이 가능하여 1.5배 향상된 성능을 보였다. 향후 연구로는 노이즈, 픽셀 값 변경 등 왜곡 공격을 고도화시켜 분할 삽입 기법의 성능을 평가하고 효율적인 워터마크 손실 대응을 위하여 오류 정정 코드 기반의 워터마크 복원기법을 연구하고자 한다.

ACKNOWLEDGMENT

본 논문은 2024년도 산업통상자원부 및 한국산업기술진흥원의 산업혁신인재성장지원사업 (RS-2024-00415520)과 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재4.0 사업의 연구결과로 수행되었음 (No. IITP-2022-RS-2022-00156310)

참고문헌

[1] 저작권기술 동향, 한국저작권위원회, 2019년-02호, p11
[2] Digital Watermark Technology market Size, Share, Growth, and Industry Analysis, By Type (Invisible Digital Watermark), By Application (Broadcasting and Television Industry), Regional Insights and Forecast to 2031, Business Research INSIGHTS, 2023.12