

특집논문 (Special Paper)
방송공학회논문지 제25권 제2호, 2020년 3월 (JBE Vol. 25, No. 2, March 2020)
<https://doi.org/10.5909/JBE.2020.25.2.151>
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

CNN을 이용한 전방위 영상의 워터마크 추출 방법

문원준^{a)}, 서영호^{a)}, 김동욱^{a)†}

Watermark Extraction Method of Omnidirectional Images Using CNN

Won-Jun Moon^{a)}, Young-Ho Seo^{a)}, and Dong-Wook Kim^{a)†}

요약

본 논문에서는 전방위 영상에 대해 알고리즘 기반으로 워터마크를 추출했던 기존 방법의 정확도를 향상시키기 위해 합성곱 신경망 (Convolutional Neural Network, CNN)을 이용한 워터마크 추출 방법을 제안한다. 이 CNN은 전방위 영상의 생성과정에서 발생하는 변형과 악의적인 공격에 대한 보정을 수행하여 워터마크를 추출하는 복원과정과, 추출한 워터마크가 어떤 워터마크인지를 분류하는 분류과정으로 구성된다. 이에 대해 다양한 공격에 대한 실험을 통해 기존 방법보다 추출되는 워터마크의 정확도가 더 높음을 확인한다.

Abstract

In this paper, we propose a watermark extraction method of omnidirectional images using CNN (Convolutional Neural Network) to improve the extracted watermark accuracy of the previous deterministic method that based on algorithm. This CNN consists of a restoration process of extracting watermarks by correcting distortion during omnidirectional image generation and/or malicious attacks, and a classification process of classifying which watermarks are extracted watermarks. Experiments with various attacks confirm that the extracted watermarks are more accurate than the previous methods.

Keyword : virtual reality, watermark, CNN, omnidirectional image

a) 광운대학교 전자재료공학과(Department of Electronic Materials Engineering, Kwangwoon University)

† Corresponding Author : 김동욱(Dong-Wook Kim)
E-mail: dwkim@kw.ac.kr
Tel: +82-2-940-5167

ORCID: <https://orcid.org/0000-0002-4668-743X>

※ 이 논문의 연구 결과중 일부는 한국방송미디어공학회 “2019년 추계학술대회”에서 발표한 바 있음.

※This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1F1A1054552).

· Manuscript received December 19, 2019; Revised February 4, 2020; Accepted February 4, 2020.

1. 서론

VR 산업은 교육, 게임, 간접 체험 등 다양한 분야에서 사용되고 있고, 이에 따라 VR콘텐츠로 사용되는 전방위 영상의 제작과 배포가 활발히 일어나고 있다^[1]. 많은 전방위 영상들이 공유됨에 따라 저작권 보호의 필요성이 높아지는 데, 영상의 저작권 보호에 널리 사용되는 방법이 바로 디지털 워터마크 기술이다.

MPEG-I에서는 등장방형도법(Equi-Rectangular Projection, ERP)과 정방체도법(CubeMap Projection, CMP)을 전방위 영상의 표준으로 정하여 사용하고 있으며, 획득한 영상을 위의 방법으로 투영하는 과정은 획득된 원 영상에 비해 많은 기하학적 왜곡과 화소 값 변화가 발생하기 때문에 기존의 2D 영상과 비디오에 대한 워터마킹 알고리즘을 사용할 때에는 좋은 결과를 낼 수 없다^{2,3)}. 따라서 이렇게 필연적으로 발생하는 왜곡 중에 기하학적인 왜곡을 보정하기 위해 본 연구팀은 전방위 영상에서 특징점을 추출하고 그 위치에 따라 왜곡이 많이 발생하는 상하 부분은 투영방법의 변화를 통해, 왜곡이 비교적 적은 부분은 3DoF(Degree of Freedom) 회전을 통해 영역을 보정하고 워터마크를 추출하는 연구를 진행한 바 있다⁴⁾. 이 방법에서는 전방위 영상이 생성되기 전의 원영상 각각의 소유자가 다른 경우를 감안하여 전방위 영상이 생성되기 전의 원영상에 워터마크를 삽입하고, 전방위 영상을 배포

하는 형식인 ERP영상에서 워터마크를 추출하는 방법이다. 본 논문은 이전 논문⁴⁾의 워터마크 추출 정확도를 향상시키기 위해 영상처리에서 널리 사용되고 있는 합성곱 신경망(CNN)을 이용하여 비악의적인 공격에 대한 학습뿐만 아니라 악의적 공격에 대해 변형된 데이터도 함께 학습하여 워터마크의 추출 정확도를 향상시키는 방법을 제안한다. 2장에서는 알고리즘적 워터마크 삽입 및 추출방법⁴⁾을 간략히 소개하고 본 논문에서 제안하는 워터마크 추출 CNN을 설명한다. 3장에서는 제안하는 CNN을 학습시키고, 그 결과를 다양한 영상과 공격에 대한 실험결과를 보인다. 이 실험결과를 토대로 4장에서 이 논문의 결론을 맺는다.

II. 제안하는 워터마킹 방법

1. 알고리즘에 의한 워터마킹 방법

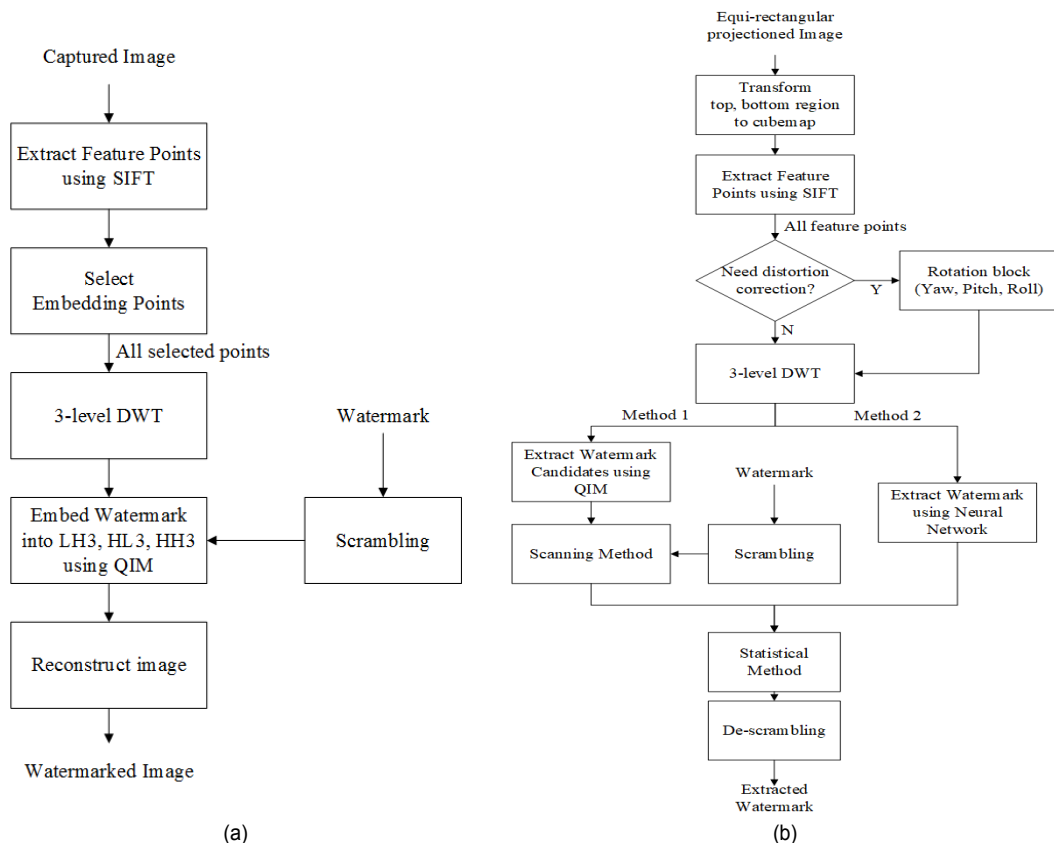


그림 1. 제안하는 워터마킹 방법: (a) 삽입 방법, (b) 추출 방법
 Fig. 1. Proposed watermarking method: (a) Embedding, (b) Extracting

알고리즘을 이용한 이전 논문^[4]의 워터마크 삽입 방법은 그림 1(a)와 같다. 먼저 삽입하는 위치를 선정하기 위해 획득한 영상에 SIFT^[5] 특징점을 추출하고, 그 중 특징점 사이의 거리가 가까워 워터마크 삽입 시 다른 특징점에 영향을 주는 점과 스티칭에서 이용되는 점들은 제외한다. 남은 특징점 주위의 32×32 영역을 3-level DWT를 수행하고 최저 주파 영역을 제외한 레벨 3의 LH3, HL3, HH3 부대역의 계수에 QIM(Quantization Index Modulation) 방법으로 워터마크를 삽입한다.

추출 방법은 그림 1(b)의 방법 1(Method 1)을 따르며, 추출의 대상은 ERP 영상이다. ERP 영상은 y값이 중점에서 멀어질수록 한 점이 투영되는 면적이 넓어져 원본 영상과 비교했을 때 왜곡이 더욱 심해지므로, 윗부분과 아랫부분은 CMP를 이용하여 투영에서 생기는 왜곡을 보정하고, SIFT 특징점을 추출한다. 추출한 점들 중에 보정이 필요한 점은 3DoF 회전을 통해 보정해준 후 특징점 주위의 영역을 선정하고, 삽입과정과 동일하게 3-level DWT를 수행하여 삽입 과정에서 사용한 부대역에 QIM 방식에 따라 워터마크 정보를 추출하여 통계적인 방법으로 최종 워터마크를 선정한다.

2. 네트워크에 의한 워터마킹 방법

워터마크 삽입방법은 이전 논문^[4]과 동일하며, 본 논문에서 제안하는 추출방법은 그림 1(b)의 방법 2(Method 2)로 나타내었다. 투영 변환 및 3DoF 회전을 통해 보정한 특징점 주위 영역에 3-level DWT를 수행하여 분리한 LH3, HL3, HH3 부대역을 순차적으로 네트워크의 입력으로 삽입한다. 또한 악의적인 공격에도 강인한 특성을 가지기 위해 워터마크가 삽입된 ERP영상에 공격을 가하여 위의 방법으로 추출된 부대역 역시 입력영상으로 포함하여 학습한다. 네트워크의 역할은 부대역의 계수를 QIM을 수행하여 삽입한 워터마크로 복원하는 것과 복원된 워터마크가 어떤 워터마크인지를 분류하는 것이다.

제안하는 네트워크는 그림 2에서 나타내었으며, 복원과 분류과정으로 나뉜다. 복원은 3개의 합성곱 계층에 각각 32, 64, 128개의 필터로 구성되어 합성곱을 수행한 다음 배치정규화(Batch Normalization, BN)와 ReLU 활성화 함수를 거치도록 구성하였다. 그 다음 복원된 워터마크를 추

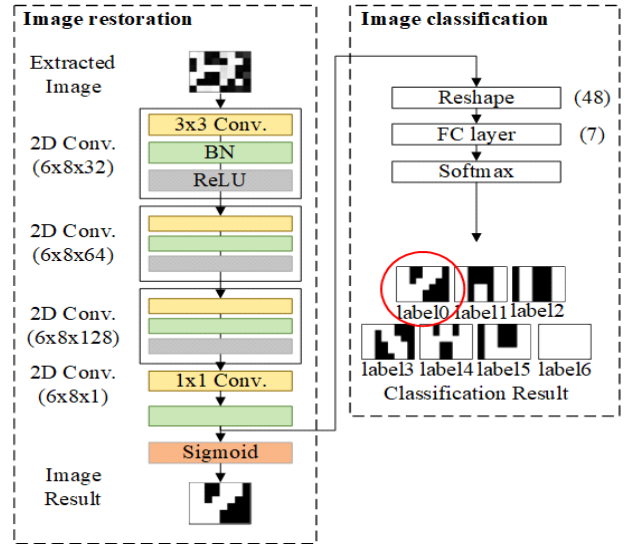


그림 2. 제안하는 워터마크 추출 네트워크
 Fig. 2. Proposed watermark extraction network

출하기 위해 1×1 합성곱을 통해 한 채널로 감소시켰고, 마찬가지로 배치정규화 이후 시그모이드(sigmoid) 활성화 함수를 이용하여 워터마크를 추출하였다. 또한 BN된 값을 분류과정으로 넘겨주어 전결합(Fully-Connected, FC) 계층을 거쳐서 소프트맥스(softmax) 함수를 통해 각 워터마크에 대한 확률을 추출하였다. 이 때 레이블의 갯수는 삽입한 워터마크의 종류가 6개이고 삽입을 하지 않은 것도 판단해야 하므로 총 7개로 구성하였다.

네트워크의 학습에 사용된 손실함수는 복원된 결과 영상 ($I_{(x,y)'}$)과 정답 영상($I_{(x,y)}$)의 차이를 교차 엔트로피 오차로 계산한 것과 분류과정에서의 각 레이블에 대한 확률 값(l'_n)과 원-핫 인코딩(one-hot encoding)된 정답 레이블(l_n)의 차이를 교차 엔트로피 오차로 계산한 것의 합으로 식(1)과 같다.

$$L_{total} = - \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} I_{(x,y)} \log f(I_{(x,y)'}) - \sum_{n=0}^{L-1} l_n \log f(l'_n) \quad (1)$$

III. 실험 및 결과

1. 실험환경 및 파라미터

사용한 훈련 및 검증집합을 위한 영상 데이터는 GoPro

Hero4 카메라를 리그에 장착하여 직접 촬영하고 그 결과를 Autopano Giga 4.4로 전방위 영상을 생성하여 사용하였다^[4]. 본 실험에 사용된 파라미터는 표 1과 같다. 비악의적인 공격에도 강인하기 위해 훈련 집합(training set)은 6개 ERP의 원본 영상에서 뿐만 아니라 각 영상에 JPEG압축(quality 60/100), 가우시안 노이즈(3%), 가우시안 블러링(5x5), 샤프닝(3 pixels)을 수행한 영상도 포함하였고, 검증 집합(validation set)은 훈련집합에 포함되지 않은 2개의 ERP에 대해 원본 영상과 위의 공격들에 대한 워터마크를 추출하였다.

표 1. 네트워크에 사용된 파라미터와 데이터 세트
Table 1. Parameters and datasets used in the network

Parameter	batch size	512
	epoch	500
	learning rate	0.00005
Dataset (Train) (8×6 image before QIM)	Training set	25,100
	Validation set	7,100

훈련집합과 검증집합에는 각 영상에서 구한 각각의 SIFT 특징점 주변 32×32 영상과 워터마크 삽입 후 공격한 영상을 모두 합하여, 훈련집합에는 총 25,100개, 검증집합에는 총 7,100개의 데이터를 포함하고 있다. 배치크기와 학습율(learning rate)은 실험적으로 구하였고, 에폭(epoch)은 검증 집합에 대한 손실함수가 수렴할 때까지 수행하였다.

2. 실험결과

3.1절의 환경으로 학습한 결과를 그림 3에서 나타내었고, 훈련집합의 영상과 레이블(Train_Image, Train_Label), 검증집합의 영상과 레이블(Valid_Image, Valid_Label)의 손실 값(a)과 정확도(b)를 구분하여 표기하였다.

약 280에폭 전까지는 모든 손실값이 감소하는 경향을 보이지만, 280에폭 이후로는 검증집합에 대한 레이블의 손실값이 오히려 증가한다. 이러한 경향은 과적합(overfitting)이며, 과적합 이후에는 일반적인 데이터에 모두 적합하도록 학습하는 것이 아니라 훈련에 사용된 데이터에만 최적화 되는 학습이기 때문에 유효한 결과라고 볼 수 없다. 따라서 과적합 되기 전 값인 280에폭의 결과를 기준으로 테스트를 수행하였다.

각각 원본영상과 JPEG압축, 가우시안 노이즈, 가우시안 블러링, 샤프닝에 대해 공격된 영상으로 실험하였고, 실험 결과를 알고리즘으로 추출한 이전 논문^[4]과 비교하여 그림 4에 나타내었다. 그림 4에서 이전 논문의 결과는 ‘Previous’로, 제안한 네트워크의 결과는 ‘Proposed’로 표기하였으며, 강인성은 NCC(Normalized Cross Correlation)과 BER(Bit Error Rate)를 모두 사용하였다.

공격의 강도가 강해질수록 워터마크의 추출률은 떨어지는 것이 일반적인데, 이전 논문에 비해 실험한 모든 공격에서 강도가 강해짐에 따라 감소하는 폭이 더 적은 것을 볼 수 있다. 이는 네트워크의 학습 과정에서 훈련 집합에 해당 공격에 의해 변형된 데이터를 포함하고, 이를 보정하는 것

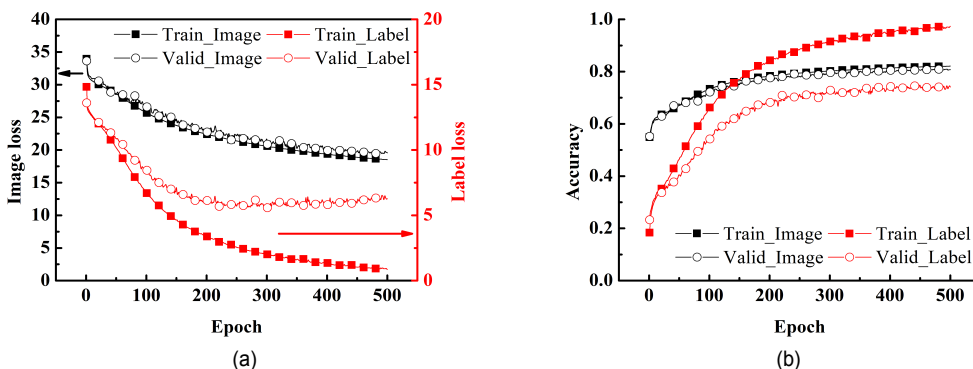


그림 3. 에폭에 따른 네트워크의 학습 및 검증 결과: (a) 복원 영상과 레이블의 손실 값, (b) 정확도
Fig. 3. Network training and validation result according to epoch: (a) restoration image and label loss, (b) accuracy

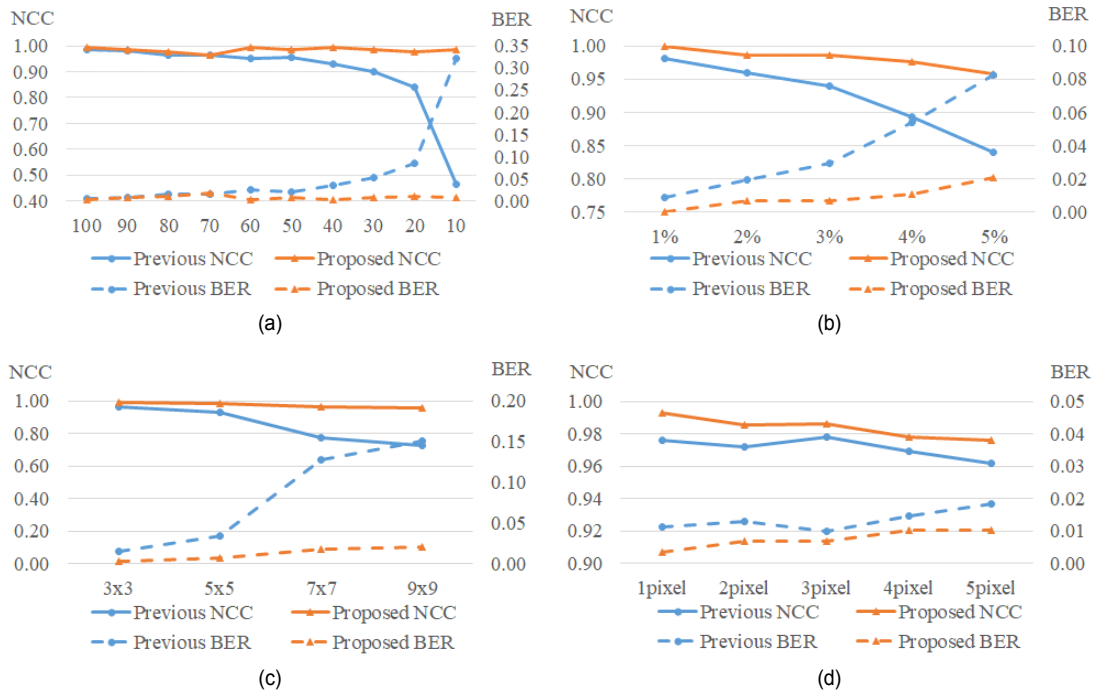


그림 4. 이전 논문과 제안하는 방법으로 추출한 워터마크 비교: (a) JPEG 압축, (b) 가우시안 노이즈, (c) 가우시안 블러링, (d) 샤프닝
 Fig. 4. Comparison of extracted watermark between previous and proposed: (a) JPEG compression, (b) Gaussian noise, (c) Gaussian blur, (d) sharpen

을 네트워크가 학습하는 것이라고 판단할 수 있다.

IV. 결론

본 논문은 원본 영상에 워터마크를 삽입하고, 전방위 영상에 대해 CNN을 통해 워터마크를 추출하는 방법을 제안하였다. CNN의 역할은 전방위 영상 제작과정에서 필연적으로 생기는 왜곡과 악의적인 공격에 데이터가 변하는 것을 보정하여 삽입한 워터마크 그대로 추출해내는 것이다. 이에 대해 알고리즘을 통해 추출한 이전논문과 비교하여 워터마크의 정확도가 더 높은 것을 검증하였다. 따라서 이 기술은 전방위 영상의 워터마크 추출에 유용하게 사용될 것이라 예상된다. 또한 본 연구에서의 방법은 다른 워터마크 기술개발에도 충분히 응용가능하리라 사료된다.

참고 문헌 (References)

- [1] Youtube Virtual-Reality Channel.. Available : <https://www.youtube.com/channel/UCzuqhhs6NWbgTzMuM09WKDQ>
- [2] MPEG-I, Information Technology - Coded Representation of Immersive Media (MPEG-I) - Part 2: Omnidirectional Media Format, document N17563, SO/IEC FDIS 23090-2:201x, ISO/IEC JTC1 SC29/WG11 MPEG, 122nd Meeting, San Diego, USA, Apr., 2018
- [3] I. S. Kang, W. J. Moon, Y. H. Seo, and D. W. Kim, "Distortion in VR 360 degree panoramic image," Summer conference on the KIBME, Jeju, Korea, pp.194-196, June, 2017
- [4] W. J. Moon, Y. H. Seo, and D. W. Kim, "SIFT Feature Based Digital Watermarking Method for VR Image," Journal of Broadcast Engineering, Vol.24, No. 6, Nov., 2019
- [5] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, Jan., 2004

저 자 소 개



문 원 준

- 2018년 2월 : 광운대학교 전자재료공학과 졸업(공학사)
- 2018년 3월 ~ 현재 : 광운대학교 전자재료공학과 (공학석사)
- ORCID : <https://orcid.org/0000-0002-9620-9524>
- 주관심분야 : Virtual Reality, 워터마킹, 2D 영상 신호처리, 딥 러닝



서 영 호

- 1999년 2월 : 광운대학교 전자재료공학과 졸업(공학사)
- 2001년 2월 : 광운대학교 일반대학원 졸업(공학석사)
- 2004년 8월 : 광운대학교 일반대학원 졸업(공학박사)
- 2005년 9월 ~ 2008년 2월 : 한성대학교 조교수
- 2008년 3월 ~ 현재 : 광운대학교 전자재료공학과 정교수
- ORCID : <http://orcid.org/0000-0003-1046-395X>
- 주관심분야 : 실감미디어, 2D/3D 영상 신호처리, 디지털 홀로그램, SoC 설계



김 동 욱

- 1983년 2월 : 한양대학교 전자공학과 졸업(공학사)
- 1985년 2월 : 한양대학교 공학석사
- 1991년 9월 : Georgia 공과대학 전기공학과(공학박사)
- 1992년 3월 ~ 현재 : 광운대학교 전자재료공학과 정교수
- ORCID : <http://orcid.org/0000-0002-4668-743X>
- 주관심분야 : 3D 영상처리, 디지털 홀로그램, 디지털 VLSI Testability, VLSI CAD, DSP설계, Wireless Communication