

CNN을 이용한 전방위 영상의 워터마크 추출

*문원준 서영호 김동욱

광운대학교

*wonjun94@kw.ac.kr

Watermark Extraction of Omnidirectional Images Using CNN

*Moon, Won-Jun Seo, Young-Ho Kim, Dong-Wook

Kwangwoon University

요약

본 논문에서는 CNN을 이용하여 전방위 영상에 대해 워터마크를 추출하는 방법을 제안한다. 네트워크의 입력은 전방위 영상에서 SIFT 특징점을 기준으로 잘라낸 영역들이며, 네트워크를 통해 전방위 영상 생성 과정에서의 왜곡을 보정하고 워터마크를 분류한다. 또한 네트워크의 훈련 집합에는 원본 영상 외에 JPEG 압축, 가우시안 노이즈, 가우시안 블러링, 샤프닝 공격을 가한 영상들도 포함시켜서 학습을 통해 공격에 대한 강인성을 가지도록 한다. 이에 대해 훈련된 네트워크로 추출한 워터마크와 알고리즘으로 추출한 워터마크를 비교하여 제안하는 방법의 유효성을 확인한다.

1. 서론

VR 산업은 다양한 분야에서 사용되고 있고, 이에 따라 VR콘텐츠 중 전방위 영상의 제작과 유통이 활발히 일어나고 있다[1]. 따라서 유통되는 전방위 영상들을 위한 저작권 보호가 필요한데, 영상의 저작권 보호에 널리 사용되는 방법이 바로 디지털 워터마크이다.

본 논문은 카메라에서 획득한 영상에 워터마크를 삽입하고, 삽입된 영상들을 이용하여 만들어진 전방위 영상에서 워터마크를 추출하는 방법을 제안한다. 전방위 영상은 MPEG-I에서 표준으로 사용하고 있는 등장방형도법(Equi-Rectangular Projection, ERP)를 타겟으로 하며 만드는 과정은 많은 기하학적 왜곡과 화소값 변형이 발생하기 때문에 기존의 2D 영상과 비디오에 대한 워터마킹 알고리즘을 사용할 때에는 좋은 결과를 낼 수 없다. 따라서 본 논문은 삽입한 워터마크에 대해 합성곱 신경망(Convolutional Neural Network, CNN)을 이용하여 워터마크를 추출하는 방법을 제안한다.

2. 제안하는 워터마크 알고리즘

2.1. 워터마크 삽입

워터마크의 삽입방법은 그림 1(a)과 같다. 획득한 영상에 SIFT를 이용하여 특징점을 추출한다[2]. 그 중에서 특징점 사이의 거리가 가까워 워터마크 삽입시 다른 점에 영향을 주는 점과 스티칭에서 이용되는 점들은 제외한다. 남은 특징점 주위의 32x32 영역을 3-level DWT를 수행하고 저주파 영역을 제외한 LH3, HL3, HH3 부대역의 계수에 QIM(Quantization Index Modulation) 방법으로 삽입한다.

2.2. 워터마크 추출

워터마크의 추출방법은 그림 1(b)과 같다. 전방위 영상은 y값이 중점에서 멀어질수록 원본 영상과 비교했을 때 왜곡이 심해지므로 이를 보정하기 위해 윗부분과 아랫부분은 CMP(CubeMap Projection)를 이용하여 왜곡을 보정해주고, 중간 영역은 3DoF(Degree of

Freedom)회전을 이용하여 보정해준다[3]. 그렇게 추출된 영역은 삽입 과정과 마찬가지로 3-level DWT를 수행하여 각각의 부대역으로 분리한다. 그 다음 CNN을 이용하는데, 네트워크의 입력은 LH3, HL3, HH3 부대역을 붙인 것(Concatenation)으로 하며, 네트워크의 역할은 QIM에 대한 동작을 수행하여 부대역의 계수를 삽입한 워터마크로 복원하는 것과 복원된 워터마크가 어떤 워터마크인지를 분류하는 것이다.

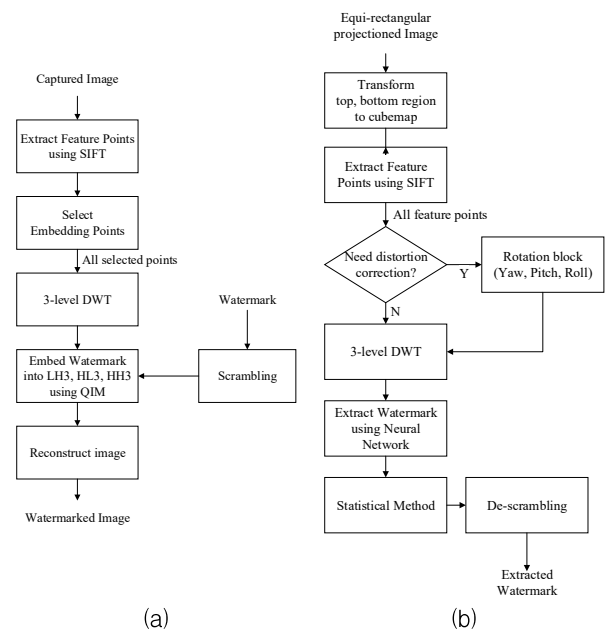


그림 1. 제안하는 워터마킹 방법: (a)삽입 방법, (b)추출 방법

자세한 네트워크는 그림 2에서 나타내었는데, 복원 과정에서는 3개의 합성곱 계층(Convolution layer)에 각각 32, 64, 128개의 필터를 수행하였다. 그 다음 복원된 워터마크를 뽑기 위해 1x1 합성곱을 통한 채널로 감소시켰고, 배치정규화(Batch normalization) 이후에 Sigmoid 활성화 함수를 이용하여 워터마크를 추출하였다. 또한 배치정규화된 값을 분류과정으로 넘겨주어 전결합 계층(Fully-connected layer)을 거쳐서 소프트맥스 함수를 통해 각 워터마크에 대한 확률을 추출하였다.

여기서 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)로 표현할 수 있다.

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

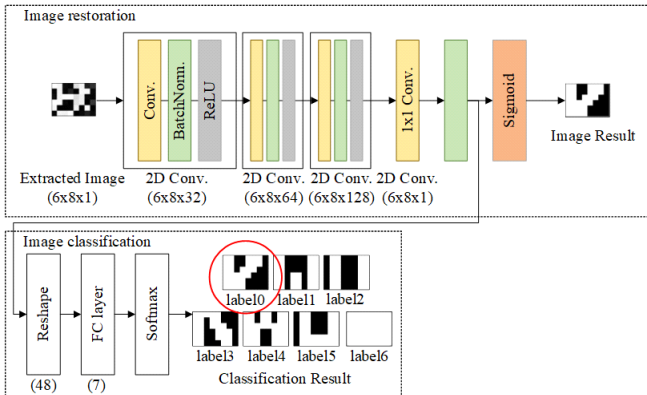


그림 2. 워터마크 추출 네트워크

3. 실험 및 결과

3.1. 실험환경 및 파라미터

본 실험에 사용된 파라미터는 표 1과 같다. 훈련 집합(Training set)은 6개의 ERP에 대해 원본 영상과 JPEG압축(Quality 60/100), 가우시안 노이즈(3%), 가우시안 블러링(5x5), 샤프닝(3pixel)을 수행한 영상에서 워터마크를 추출한 것이고, 검증 집합(Validation set)은 2개의 ERP에 대해 원본 영상과 위의 공격들에 대한 워터마크를 추출한 것이다.

배치크기와 학습율(learning rate)은 실험적으로 구하였고, 에폭(epochs)은 검증 집합에 대한 손실함수가 수렴할 때까지 수행하였다.

표 1. 네트워크에 사용된 파라미터와 데이터 세트

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

3.2. 실험결과

3.1절의 환경으로 학습한 결과를 그림 3에서 나타내었는데, 약 280 에폭 전까지는 모든 손실 값이 감소하는 경향을 보이지만, 280에폭 이후로는 검증집합에 대한 레이블의 손실 값이 오히려 증가한다. 따라서 과적합(Overfitting)되기 전 값인 280에폭을 기준으로 테스트를 수행하였다.

각각 원본영상과 JPEG압축, 가우시안 노이즈, 가우시안 블러링, 샤프닝에 대해 공격된 영상으로 실험하였고, 실험 결과를 알고리즘으로 추출한 이전 논문[3]과 비교하여 그림 4에 나타내었다.

공격의 강도가 강해질수록 워터마크의 추출률은 떨어지는 것이 일반적인데, 이전 논문에 비해 모든 공격과 모든 강도에서 떨어지는 폭이 작은 것을 볼 수 있다. 이는 네트워크의 학습에서 훈련 집합에 해당 공격에 대해 변형된 데이터들도 포함이 되고, 이를 보정하는 것을 네트워크가 학습하는 것이라 사료된다.

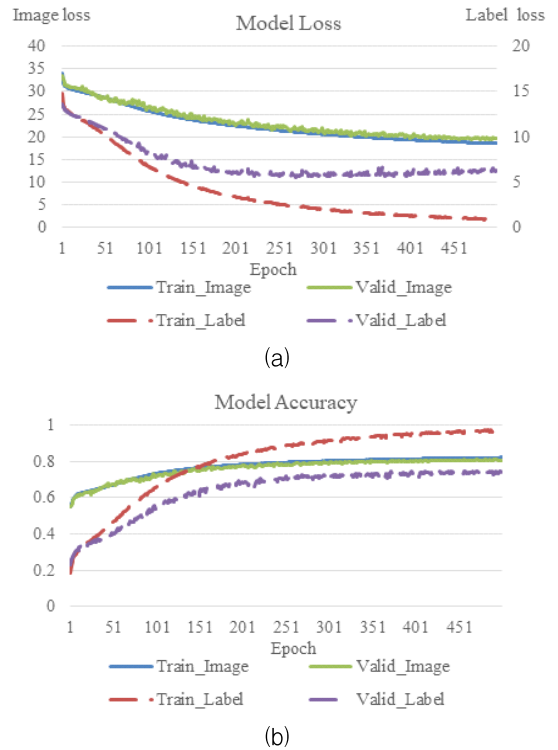


그림 3. 에폭에 따른 훈련 결과: (a) 훈련 및 검증 집합에 대한 복원영상과 레이블의 손실 값 (b) 정확도

감사의 글

이 논문은 2016 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2016R1D1A1B03930691)

참고문헌

- [1] Youtube Virtual-Reality Channel.. Available : <https://www.youtube.com/channel/UCzuqhhs6NWbgTzMuM09WKDQ>
- [2] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, Jan. 2004.
- [3] 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

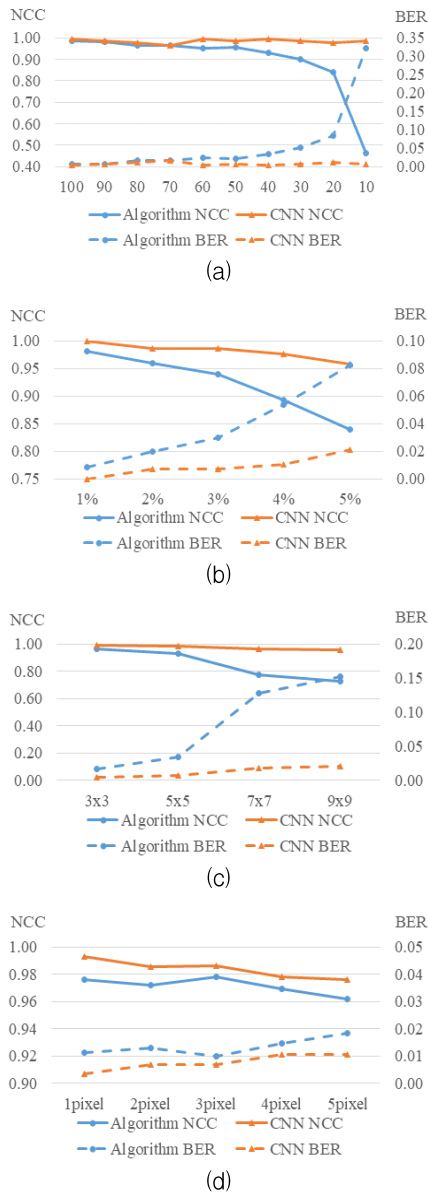


그림 4. 이전 논문과 제안하는 방법으로 추출한 워터마크 비교: (a) JPEG 압축, (b) 가우시안 노이즈, (c) 가우시안 블러링, (d) 샤프닝

4. 결론

본 논문은 원본 영상에 워터마크를 삽입하고, 전방위 영상에 대해 CNN을 통한 워터마크 추출 방법을 제안하였다. CNN의 역할은 전방위 영상으로 만들면서 생기는 왜곡과 악의적인 공격에 데이터가 변하는 것을 보정하여 삽입한 워터마크 그대로 추출해내는 것이다. 이에 대해 알고리즘을 통한 추출 결과와 비교하여 워터마크의 정확도가 더 높은 것을 검증하였다. 따라서 이 기술은 전방위 영상의 워터마크 추출에 유용하게 사용될 것이라 예상된다.