

CNN 기반 HEVC 압축된 동영상의 삭제 검출 기법

홍진형, 양윤모, 오병태
한국항공대학교

hong_1993@kau.kr, yym064@kau.kr, byungoh@kau.ac.kr

Detection of Frame Deletion for HEVC-coded Video Using CNN

Jin Hyung Hong, Yoonmo Yang, Byung Tae Oh
Korea Aerospace University

요 약

최근 딥 러닝 기술의 발전이 가속화됨에 따라, 기존의 알고리즘과 융합하여 뛰어난 성능 향상을 보이는 연구가 급격히 증가하고 있다. 본 논문에서는 딥 러닝을 이용하여 HEVC 로 압축된 동영상의 일부 프레임의 삭제 여부를 검출하는 알고리즘을 제안한다. 영상의 삭제 정보가 포함되어 있는 HEVC 의 부호화 파라미터를 추출하여 간단한 전 처리 과정을 통해 데이터의 크기를 효과적으로 압축한 뒤, 동영상의 시간적 특성을 고려할 수 있도록 CNN 네트워크를 구성한다. 실험 결과, 효과적으로 다양한 압축 환경에 강인한 영상 삭제 검출 성능을 보이는 것을 확인하였다.

1. 서론

최근 들어, 급격히 발전하고 있는 딥 러닝 기술에 힘입어, 영상처리의 다양한 분야에 접목되어 효과적으로 성능을 강화하고 있다. 이와 함께, 과학 기술의 발전하며 시민들의 디지털 영상 획득 경로가 다양해지고 있다. 디지털 문화가 발전함에 따라, 다양한 방식으로 디지털 영상의 조작과 수정이 이루어지고 있다. 법정에서는 이러한 사회적 현상이 여러가지 어려움을 야기시킬 수 있다. 따라서 과학의 급격한 발전에 맞추어 디지털 포렌식 기술의 발전이 필수불가결하다. 디지털 포렌식의 한 분야인 비디오 포렌식 기술에는 크게 동영상의 시/공간적 상관관계를 이용한 조작 검출 방식과, 물리적 불일치성을 이용한 조작 검출 방식, 동영상을 여러 번 압축할 때 발생하는 조작 위치에서의 흔적(footprint)을 검출하는 방식 등이 있다. 이 때, 동영상을 조작하며 필수적으로 다수 발생하는 동영상 압축은 동영상의 진실성을 떨어뜨린다. 따라서 동영상의 압축 횟수가 동영상의 조작 여부를 판단하는 기준이 되기도 한다 [1-4]. 본 논문에서는 최근 유행하고 있는 딥 러닝 기술을 이용하여 동영상의 일부 영상 삭제 조작 이후 이중 압축(double compression) 여부를 검출하는 기법에 대해 소개하고자 한다.

대부분의 비디오 포렌식 연구는 지금까지 CIF(352×288) 및 QCIF(176×144)와 같은 저해상도 동영상을 기준으로 진행되어 왔다. 하지만, 과학 발전의 흐름에 맞추어 이제는 HD(1920×1080) 또는 UHD(3840×2160)와 같은 고해상도 동영상을 포함하는 포렌식 연구가 진행되어야 한다. 또한, MPEG-2 또는 H.264/AVC 등의 동영상 압축 표준 뿐만 아니라 고성능 동영상 압축 코덱(HEVC)을 기반으로 하는 동영상의 조작 검출 연구가 진행될 필요가 있다. 따라서 본

논문에서는 HEVC 동영상 압축 코덱을 이용하여 압축된 고해상도의 동영상의 조작 검출을 소개한다. 이 알고리즘은 고해상도의 동영상을 HEVC 동영상 압축 파라미터를 딥 러닝에 접목하여 빠르고 강인하게 삭제 범위를 탐색할 수 있다.

본 논문의 구성은 다음과 같다. 2 장에서는 본 논문에서 소개하는 주요 알고리즘에 대해 구체적으로 기술하고, 3 장에서는 제안하는 알고리즘의 성능을 실험을 통해 확인한다. 4 장에서는 마지막으로, 본 논문에 대한 결론을 맺는다.

2. 제안 알고리즘

제안하는 알고리즘에서는 인위적으로 조작된 동영상의 특징을 추출하기 위해 이중 압축 검출에 중점을 두었다. 조작된 동영상의 경우, 최초 획득된 동영상에 대해 수정이 이루어지고, 저장하는 과정에서 한 번 이상의 압축이 추가적으로 수행된다. 따라서 동영상의 원본 여부를 확인하기 위해서, 이중 압축 여부를 조사하는 것은 타당하다고 할 수 있다. 또한 동영상에서의 일부 영상을 삭제하는 과정이 수행된다면, 동일한 압축 코덱과 그 파라미터를 이용하지 않는다면, 그 이후의 압축 과정에서 영상의 종류 변화는 필연적이다. 따라서 영상 타입 변화를 특징으로서 이용할 수 있다면 이중 압축을 검출할 수 있고, 나아가 원본 여부를 검출할 수 있다. 제안하는 알고리즘에서는 영상의 종류 변화를 검출하기 위해 HEVC 의 부호화 정보를 이용하여 특징을 추출하고, 적절한 전 처리 과정을 수행하여 미리 구성된 CNN 네트워크의 입력으로 제공한다. 아래 그림 1 은 전반적인 알고리즘의 순서도이다.

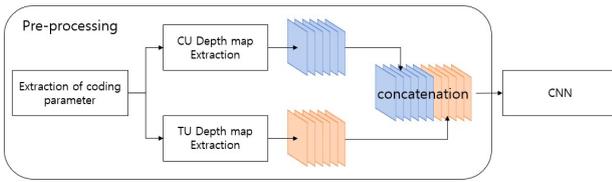


그림 1. 제안 알고리즘의 순서도

A. HEVC의 기본 단위

앞서 언급한 것처럼, HEVC에서의 동영상의 일부 영상 삭제를 검출하기 위해 HEVC의 기본 구조를 이용할 필요가 있다. HEVC 부호화기에서는 MPEG-2, H.264/AVC와 달리 매크로블록(Macro Block)을 사용하지 않고, 더 큰 단위인 Coding Tree Unit (CTU)와 그 세부 단위인 Coding Unit (CU), Prediction Unit (PU), Transform Unit (TU)을 이용하여 동영상을 부호화 한다. CTU는 모든 압축과정에서 CU로 분할된다. CU는 압축 효율에 따라 크게 64x64에서 작게는 8x8까지 다양한 크기로 분할된다. 각각의 CU는 화면 내 예측 또는 화면 간 예측 모드 중 하나로 압축되며, 화면 내 예측의 경우 영상 내에서만 참조하고, 화면 간 예측의 경우 이웃의 영상을 참조한다. 각각의 CU는 필요한 경우에 한해 변환 및 양자화 과정을 거치는데, 이 때 TU를 기준으로 해당 과정이 수행된다. 이 두 가지 CU와 TU는 뒤에서 제안하는 알고리즘의 특징 추출에서 사용하게 된다.

B. 특징 추출

CU와 TU는 동영상의 일부 영상이 삭제되었을 때, 해당 영상의 이웃 영상으로부터 참조하는 과정에서 원본 동영상과의 오차가 발생하게 된다. 동영상의 압축 과정에서 일부 영상 삭제로 인해 참조 영상과의 거리가 멀어지게 되면, 동영상 압축의 기본 단위인 CU와 변환과 양자화를 위한 단위인 TU의 분할 정보가 원본과 차이를 가질 수 있다. 이러한 특징을 이용하기 위해 CU와 TU의 분할 정보를 압축된 동영상의 영상 단위로 추출한다. 또한 화면 간 예측 과정은 Group of Picture (GOP)단위와 상관관계가 높으므로, 해당 데이터를 그림 1에서와 같이 GOP단위로 구성한다. CU와 TU의 깊이 맵을 동시에 고려하기 위해 시간의 차원으로 두 특징을 연결하여 하나의 데이터로 구성한다. 이와 같이 GOP단위로 데이터를 구성할 경우, 일부 영상의 삭제가 포함된 GOP와 삭제 이후의 GOP는 영상의 삭제가 이뤄지지 않은, 삭제 이전의 GOP와 다른 부호화 패턴을 보이게 되며, 이러한 부호화 패턴을 구분하기 위해 합성곱 신경망(CNN)을 구성한다.

3. 실험결과

시간적 특성을 이용하기 위해 GOP단위로 데이터셋을 구성함에 따라 데이터 셋을 구성하기 위해 동영상의 수가 많이 필요하다. 본 논문에서는 총 58개의 동영상을 사용하였다. 훈련에 사용된 동영상은 무작위로 50개이며, 평가에 사용된 동영상은 나머지 8개로 진행되었다.

일부 영상이 삭제되면 두 번째 압축에서 해당 위치의 영상의 예측 모드 타입이 변한다. 따라서 GOP내 삭제 위치를 바꿔가며 해당 GOP 끝까지 삭제하는 방식으로 삭제 길이를 조절하였다. 예를 들어 GOP 크기가 8인 동영상의 경우, GOP내 5번째 위치에서 영상의 삭제가 발생한다면, 총 4장의 영상이 삭제되는 방식이다. 이와 같이 영상 삭제 경우를 다양하게 준비하여, 여러가지 상황에서 강인하게 영상 조작 여부를 검출할 수 있도록 실험 환경을 조성하였다.

표 1. 동영상 조작 검출 결과

	LowDelayP	RandomAccess
Precision	0.9842	0.9888
Recall	0.9873	0.9866
F1-score	0.9857	0.9877

표 1에서 보면, 제안하는 알고리즘에서 LowDelayP의 재현율(Recall)은 0.9873, 정확률(Precision)은 0.9842, F1-score는 0.9857이다. 또한 RandomAccess의 재현율은 0.9888, 정확률은 0.9866, F1-score는 0.9877이다. 즉, 제안하는 알고리즘이 HEVC 압축 환경에서 고해상도의 동영상 데이터셋에 대해 시간적인 특성을 고려하여 일부 영상의 조작 여부를 동영상 압축 모드에 강인하게 효과적으로 검출하고 있음을 잘 보여주고 있다.

4. 결론

본 논문에서는 HEVC로 압축된 고해상도의 동영상의 삭제 여부를 딥 러닝 네트워크를 통해 판단하기 위해, 압축된 동영상의 여러 부호화 패턴을 분석하여 일부 영상의 삭제로 인해 CU와 TU에서 발생하는 특별한 분할 정보를 추출하였으며, 딥 러닝 네트워크에 사용하기 알맞게 전 처리 과정을 수행하고, 시간적인 특성을 고려하여 CNN을 이용하여 영상 삭제 여부를 검출하는 알고리즘을 제안하였다. 제안하는 알고리즘은 공정한 상황에서의 실험 성능 도출을 위해 객관적이고, 다양한 실험환경에서 데이터 셋을 구성하였으며, 실험 결과 높은 성능으로 일부 영상 삭제 여부를 검출하는 것을 확인할 수 있었다.

감사의 글

본 연구는 2016년도 정부(교육부)의 재원으로 한국연구재단 기초연구사업(NRF-2016R1D1A1B03930917)과 경기도 지역협력 연구센터 사업(GRRC) (2017-B02, 이동형 미디어 디바이스를 위한 영상처리 및 UI 플랫폼 기술연구)의 지원을 받아 수행되었음.

참조문헌

[1] Milani, S., Fontani, M., Bestagini, P., Barni, M., Piva, A., Tagliasacchi, M., and Tubaro, S.. An overview on video

forensics. *APSIPA Transactions on Signal and Information Processing*, 2012, 1: e2.

- [2] Wang, W., and FARID, H.. Exposing digital forgeries in video by detecting double MPEG compression. In: Proceedings of the 8th workshop on Multimedia and security. ACM, 2006. p. 37-47.
- [3] Wang, W., and FARID, H.. Exposing digital forgeries in video by detecting double quantization. In: Proceedings of the 11th ACM workshop on Multimedia and security. ACM, 2009. p. 39-48.
- [4] Shanableh, T.. Detection of frame deletion for digital video forensics. *Digital Investigation* 2013, 10.4, pp. 350-360.