

## 강화학습을 이용한 비활성 영역 패딩

김동신, 오병태

한국항공대학교

dskim20@kau.kr, byungoh@kau.ac.kr

## Inactive region padding by reinforcement learning

Dongsin Kim, Byung Tae Oh

Korea Aerospace University

## 요 약

본 논문에서는 강화학습을 이용하여 비활성 영역 패딩하여 동영상 압축 효율을 향상시키는 기법을 제안한다. 비활성 영역이란 360 영상 혹은 3DOF+ 영상의 예시와 같이 영상 내에서 정보가 존재하지 않는 영역을 의미한다. 하지만 이러한 비활성 영역은 일반적으로 영상의 압축 성능을 제한한다. 기존에는 이를 해결하기 위해 활성 영역과 비활성 영역의 경계부분을 필터링하여 압축 성능을 향상시켰다. 하지만 이러한 방법들은 영상의 특성을 적절하게 반영하지 못하게 된다. 제안하는 기법에서는 영상의 특성과 압축 과정을 고려한 강화학습을 통해서 패딩을 진행하였다. 실험 결과 제안한 기법이 기존 기법에 비해 좋은 성능을 보임을 확인할 수 있다.

In this paper, we propose a new method for inactive region padding using reinforcement learning. Inactive region is an area that has no information, such as 360 or 3DOF+ videos. However, these inactive regions degrade the compression performance in general. To improve the compression performance, simple filtering is applied between active and inactive regions. But it does not fully consider the characteristics of the images. In the proposed method, inactive regions are padded through reinforcement learning that can consider the characteristics of images and the compression process. Experimental results show that the performance is better than the conventional padding method.

## 1. 서론

최근 영상의 하드웨어 기술이 발전함에 따라 고해상도 영상에 더불어 몰입형 영상에 대한 관심 역시 증가하고 있다. 몰입형 영상은 360영상 혹은 3DOF+ 영상과 같이 여러 시점의 영상들을 동시에 제공할 수 있는 영상이다. 따라서 몰입형 영상을 표현하기 위해서는 기존 2D 영상과는 다른 표현법이 필요하다. Joint Video Exploration Team (JVET)에서 제공하는 360

라이브러리에 360 영상을 다양한 포맷으로 형성하는 기법들과 각 포맷들 간의 변환 기법이 있으며[1], Moving Picture Experts Group (MPEG)에서 제공하는 Test Model for Immersive Video (TMIV)를 통해 여러 시점의 360영상을 한 번에 압축할 수 있다[2].

360라이브러리에서는 3D 영상을 구 형태의 영상으로 간주하고 구를 여러 면으로 분리해 투영시켜 2D 영상을 만든다. 이 때 구를 분리하는 방법에 따라서 다양한 포맷을 제공하며

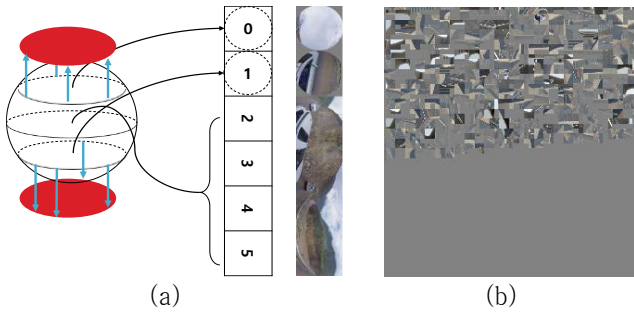


그림 1. 비활성 영역이 존재하는 영상: (a) SSP 영상 (b) 3DOF+ 영상의 frame packing 후 영상

대표적으로 Equirectangular Projection (ERP), Hybrid equiangular cubemap (HEC), Sphere Segmented Projection (SSP)와 같은 투영 방식이 존재한다. 여기서 SSP 방식의 경우 그림 1(a)와 같이 구의 상단과 하단을 원 형태로 투영시키며, 이 때 원의 바깥 부분에 유효하지 않은 화소가 존재하게 된다.

TMIV에서는 여러 시점의 360 영상을 하나로 압축하는 과정에서 각 영상들 간의 중복성을 제거하고 이를 패치단위로 분할하여 저장한다. 이 때 그림 1(b)와 같이 분할된 패치는 임의의 형태를 가지고 있으며, 360영상과 마찬가지로 유효하지 않은 화소가 존재하게 된다.

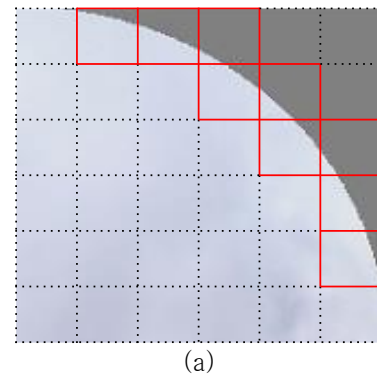
이러한 유효하지 않은 화소들의 영역을 비활성 영역이라고 한다. 비활성 영역의 화소 값들은 활성 영역의 화소 값들과 급격한 차이를 보이며, 이는 압축 성능을 크게 저하시키는 원인이다. 본 논문에서는 강화학습을 통해서 비활성 영역 패딩을 진행하여 압축 성능을 향상시킬 수 있는 기법을 제안한다.

## 2. 제안 알고리즘

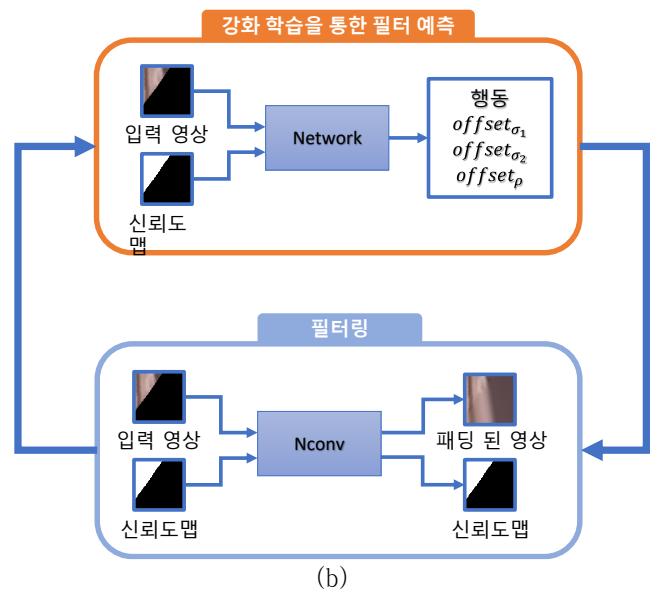
일반적으로 비활성 영역과 활성 영역은 그 경계에서 계단 함수 형태의 화소 값을 가지고 있다. 화소 값의 급격한 변화는 영상의 압축과정을 고려했을 때 비트율과 화질 모두 성능 저하를 발생시킨다. 이를 방지하기 위해서 기존 360라이브리에서는 비활성 영역의 값을 128로 채워 그 차이를 최소화했으며 활성 영역의 화소 값을 방향에 따라 일정 거리만큼 비활성 영역의 화소 값과 선형적으로 보간하여 압축 성능을 향상시켰다[3]. 이외에도 영상을 16x16 블록 단위로 나눈 후 활성 영역과 비활성 영역의 경계 근처에서 블록 단위로 패딩을 진행하는 기법[4], 주변 화소 값 혹은 구 형태의 면에 대응하는 사각형 투영면에서 화소 값을 가져오는 기법 등이 존재한다[5].

위 기법들은 비활성 영역과 활성 영역의 화소 값 간의

차이를 줄여 압축 성능을 높이는 방법을 사용했다. 하지만 모델 기반의 기법들의 한계로 인해 각 영상의 특징을 고려하지 않았으며 이에 따라 최적의 방식이 아닐 가능성이 높다. 따라서 제안하는 기법에서는 활성 영역의 값을 비활성 영역으로 전파하기 위해 필터링을 사용하며 활성 영역과 비활성 영역의 화소 값 간의 차이를 줄였으며, 화소 값을 전파할 때는 보다 효율적으로 전파할 수 있는 normalized convolution 기법을 사용한다[6]. 또한 영상의 필터링을 통해 활성 영역의 화소 값을 전파할 때 영상의 특성을 고려해 압축 성능을 높이기 위해 다변량 정규분포를 따르는 필터를 구성한다. Normalized convolution은 필터의 형태에 따라 값을 전파하는 방향이 달라지며, 신뢰도 맵을 통해서 각 위치에서 값을 얼마나 가져올지 결정한다. 따라서 필터의 값이 다변량 정규분포를 따른다면 필터의 형태를 쉽게 조절해 값을 전파하는 방향을 조절할 수 있으며, 다변량 정규분포의 형태를 결정하는 변수들은 강화학습을 통해서 예측할 수 있다. 또한 활성영역인지 아닌지를 나타내는



(a)



(b)

그림 2. 제안하는 패딩 방식: (a) 64x64 블록 기반 패딩 (b) 패딩 시스템 구성도

이진 마스크를 초기 신뢰도 맵으로 사용하면 활성영역의 값 만을 사용해 비활성 영역 패딩을 진행할 수 있다. 마지막으로 블록단위 영상 압축을 고려해 영상을 64x64 블록 단위로 나눈 후 활성 영역과 비활성 영역이 동시에 존재하는 블록에 대해서 비활성 영역 패딩을 적용한다.

제안하는 기법에서 사용하는 패딩 방식은 그림 2와 같다. 실험을 위해서 비활성 영역이 존재하는 영상 중 하나인 SSP 영상을 사용했으며 해당 영상을 normalized convolution 기법을 통해 필터링한다. 한 번 필터링 된 영상은 사용된 필터의 크기에 따라 일정 거리만큼 활성 영역의 값이 비활성 영역으로 전파되었으며, 비활성 영역의 화소 값들을 활성 영역에서 전파된 값으로 모두 채우기 위해 해당 필터링 과정을 여러 번 반복한다. 각 반복에서는 원본 활성 영역의 값이 변하지 않도록 하기 위해 필터링을 통해 변화된 활성 영역의 화소 값을 활성영역인지 아닌지 알려주는 마스크를 통해서 원본 화소 값으로 대체하는 과정이 존재한다. 또한 값을 전파하는 관점에서 마스크를 통한 화소 값 대체 기법은 기존 화소 값이 가지는 에너지를 보존해 더욱 효율적으로 전파시킬 수 있다.

강화 학습을 통해서 다변량 정규분포를 결정하기 위해서는 평균과 공분산 행렬이 필요하다. 제안 기법에서는 행동의 복잡도를 고려하여 2차원 가우시안 필터를 사용하였으며, 이에 따라 공분산 행렬을 결정하기 위한 변수로 2개의 표준편차와 1개의 상관계수가 필요하다. 또한 표준편차와 상관계수는 연속적인 값을 가지기 때문에 일정 범위 내의 불연속적인 보상 값을 행동으로 해 그 값을 조정하였다. 또한 필터링 중인 영상과 그에 대응하는 신뢰도맵은 영상의 특징과 필터링 된 정도 등 영상의 현재 상태를 효율적으로 나타낼 수 있기 때문에 강화학습의 상태로 사용하였다. 마지막으로 행동이 적용되기 전과 후의 영상에 대해서 각각 압축을 진행한 후 Bjøntegaard delta bit-rate (BDBR)[7]의 차이를 계산해 보상으로 사용하였다. 강화학습을 위한 모델으로는 DQN[8]을 사용하였으며 네 개의 컨볼루션 층과 두 개의 FC 층을 사용하였다.

### 3. 실험

시스템을 학습하기 위해 Youtube에서 제공하는 4K 영상을20개 사용하였으며 각 영상에서 10프레임당 한 장 단위로 총 20장을 추출하여 학습 데이터셋을 구성하였다. 검증의 경우 MPEG에서 제공하는 4개의 4K 영상을 사용하였으며 각 영상에서 100장을 추출하여 검증 데이터셋을 구성하였다.

제안 기법을 평가하기 위해 JVET에서 제공하는 360 라이브러리의 패딩 기법이 적용된 영상들과 BDBR을 측정해 평가

표 1. BDBR 성능 비교

실험영상	BDBR
DrivingInCity	-8.79%
DrivingInCountry	-6.97%
PoleVault	-6.67%
AerialCity	-7.88%
Average	<b>-7.58%</b>

지표로 사용하였다. 또한 블록기반 처리과정을 고려하여 강화학습이 적용된 CTU에 대해서 성능을 측정하였다. 표 1에서 그 결과를 볼 수 있으며 제안하는 기법을 통해 높은 성능 향상을 얻을 수 있음을 알 수 있다.

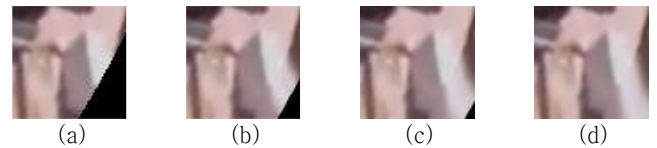


그림 3. 필터링 반복 횟수에 따른 비활성 영역의 변화: (a) 원본 (b) 5 번 반복 (c) 10 번 반복 (d) 15 번 반복

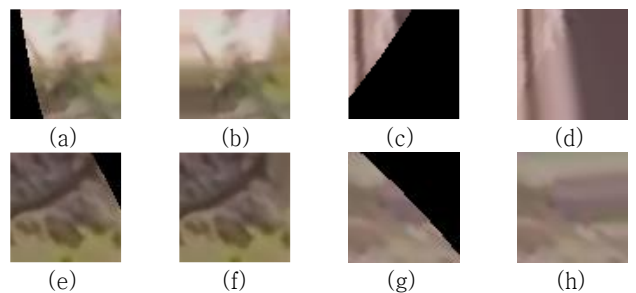


그림 4. 제안 기술이 적용된 CTU의 몇 가지 예: (a), (c), (e), (g) 원본 영상 (b), (d), (f), (h) 패딩 후 영상

그림 3은 활성 영역의 값을 전파하기 위해 필터링 과정을 여러 번 반복할 때 각 반복 단계에서의 필터링 후 영상을 보여준다. 필터링이 반복될수록 값이 전파됨을 확인할 수 있다. 그림 4는 패딩이 끝난 CTU 영상이다. 해당 영상의 비활성 영역에 채워진 값들을 보면 기존 활성 영역의 패턴을 비슷하게 유지하는 것을 볼 수 있다. 이는 전파 방향을 고려한 필터링에 의한 결과이다. 또한 단순히 가우시안 필터링을 취한 결과와 달리

활성 영역의 값들을 그대로 보존하며 패딩이 진행됨을 확인할 수 있다.

#### 4. 결론

본 논문에서는 강화학습을 통해 비활성 영역을 패딩 하는 기법에 대해서 제안하였다. 제안하는 방법은 블록 단위로 패딩을 적용하였으며 각각의 블록에서 여러 번 필터링을 반복하여 활성 영역의 값을 비활성 영역으로 전파하였다. 또한 강화학습을 통해서 분산과 상관계수를 예측해 필터를 구성하기 위한 2차 가우시안 분포를 특정함으로써 전파 방향을 효율적으로 결정할 수 있었다. normalized convolution을 이용한 필터링 기법을 통해 활성 영역의 값 만을 사용해 전파에 사용하였다. 실험 결과를 통해 제안하는 기법이 기존 기법에 비해 좋은 성능을 보임을 확인할 수 있다. 추후 연구에서는 필터의 자유도를 높이거나 필터링에 사용될 화소 값들을 학습을 통해 결정하는 등 제안 기법의 한계를 극복하는 방향으로 연구를 진행할 예정이다.

#### 감사의 글

본 연구는 2020 년도 정부(교육부)의 재원으로 한국연구재단 기초연구사업(NRF- 2019R1F1A1063229)과 경기도 지역협력 연구센터 사업 (GRRC) (2020-B02, 3 차원 공간 데이터 처리 및 응용기술 연구)의 지원을 받아 수행되었음.

#### 참조문헌

- [1] Y. Ye, E. Alshina, and J. Boyce, "Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 5)," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-H1004, Oct. 2017.
- [2] B. Salahieh, B. Kroon, J. Jung, M. Domański (Eds.), "Test model 2 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N18577, July 2019.
- [3] Y.-H. Lee, H.-C. Lin, J.-L. Lin, S.-K. Chang, and C.-C. Ju, "EE4: ERP/EAP-based segmented sphere projection with different padding sizes," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-G0097, July 2017.
- [4] A. Abbas, "AHG8: An Update on RSP Projection," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-H0056, Oct. 2017.

- [5] Y.-U. Yoon, D.-H. Park, J.-G. Kim, Y. Ahn, and D. Sim, "Methods of padding inactive regions for segmented sphere projection," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-I0023, Jan. 2018.
- [6] Eldesokey, Abdelrahman, Michael Felsberg, and Fahad Shahbaz Khan. "Confidence propagation through cnns for guided sparse depth regression." IEEE transactions on pattern analysis and machine intelligence 42.10 (2019): 2423-2436.
- [7] Bjøntegaard, G. (2001). Calculation of average PSNR differences between RD-curves. VCEG-M33.
- [8] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602, 2013.