

종단간 인공지능망 기반 이미지 압축 기술의 피쳐 공간 상관관계 분석

*이주영, 정세윤, 최진수

한국전자통신연구원

leejy1003@etri.re.kr, jsy@etri.re.kr, jschoi@etri.re.kr

Correlation Analysis of Feature Space Data in End-to-end Image Compression Network

*Jooyoung Lee, Se-Yoon Jeong, Jin Soo Choi

요 약

뉴럴넷 기술이 발전과 함께 다양한 분야에서 획기적인 성능 향상이 이루어지고 있다. 이미지 압축 분야에서도 기존의 전통적인 툴 체인 구조의 압축 방식에서 벗어나 종단간(end-to-end) 뉴럴넷 기반의 이미지 압축 기술에 대한 연구가 활발히 이루어지고 있다. 특히 최근 네트워크를 통해 변환된 피쳐 데이터의 엔트로피를 최소화하는 방식에 대한 연구가 활발히 이루어지고 있으며, 이에 기반한 최근의 연구는 VVC 화면 내 코딩 기술보다 우수한 코딩 효율성을 제공하고 있다. 그러나 변환된 피쳐 데이터에 대한 특성 분석은 부족한 실정이며, 이에 본 논문에서는 엔트로피 최소화 기반 종단간 이미지 압축 네트워크의 피쳐 공간 데이터에 대한 공간적(spatial) 상관관계와 채널 간(inter-channel) 상관관계(correlation)를 분석하고, 나아가 최근 제안된 종단간 이미지 압축 네트워크의 문맥 기반 예측 기능을 통해 잔존하는 데이터 중복성이 효과적으로 제거됨을 보인다.

1. 서론

최근 뉴럴넷 기술이 발전과 함께 다양한 분야에서 획기적인 성능 향상이 이루어지고 있다. 이미지 압축 분야에서도 기존의 전통적인 툴 체인 구조의 압축 방식에서 벗어나 종단간(end-to-end) 뉴럴넷 기반의 이미지 압축 기술에 대한 연구가 활발히 이루어지고 있다. 2016 년 구글 그룹의 연구[1,2]를 시작으로 다양한 연구가 진행되어 왔으며, 특히 최근에는 네트워크를 통해 변환된 피쳐 데이터의 엔트로피를 최소화하는 방식에 대한 연구가 활발히 이루어지고 있다[3-8]. 엔트로피 최소화 기반의 뉴럴넷 이미지 압축 기술은 점진적인 성능향상을 이루어 왔으며, 가장 최근의 연구[8]는 VVC 화면 내 코딩 기술보다 우수한 코딩

효율성을 제공하였다. 그러나 아직 변환된 피쳐 데이터에 대한 특성 분석은 부족한 실정이며, 이에 본 논문에서는 엔트로피 최소화 기반 종단간 이미지 압축 네트워크의 피쳐 공간 데이터에 대한 공간적 상관관계와 채널 간 상관관계를 분석한다.

2. 배경: 엔트로피 최소화 기반 종단간 뉴럴넷 이 이미지 압축기술

초창기 종단간 뉴럴넷 기반 이미지 압축 기술[1,2]은 주로 오토인코더의 특징 중 하나인 차원 축소(dimension reduction) 기능을 기반으로 기술이 개발되었다. 즉 인코더 네트워크는 입력

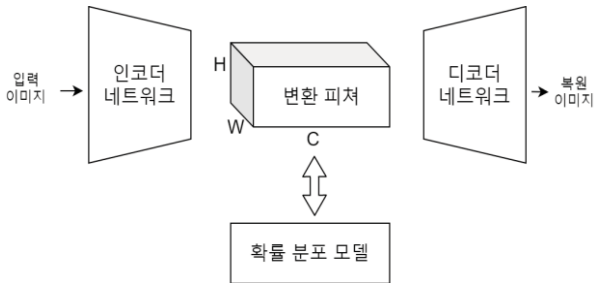


그림 2 엔트로피 최소화 기반 종단간 이미지 압축 네트워크의 기본 구조

데이터의 성분 수보다 작은 수의 피쳐 데이터로 변환하고, 디코더 네트워크는 이를 다시 이미지로 변환하는 구조를 취하였다. 해당 선행 연구들은 왜곡(distortion) 항만을 이용하여 학습을 수행하였고, 이에 인코더 네트워크는 한정된 수의 피쳐 데이터가 최대한 많은 정보를 담을 수 있도록 학습되었다. 결과적으로, 초기 종단간 뉴럴넷 기반 이미지 압축 기술의 인코더 네트워크는 입력 이미지 데이터를 높은 엔트로피의 피쳐 데이터로 변환하는 특성이 있다. 반면 최근의 엔트로피 최소화 기반의 이미지 압축 방법들은 기존 방식과 달리 변환된 피쳐 데이터의 수를 크게 확장하고, 대신 피쳐 데이터가 지니는 엔트로피를 최소화하는 방향으로 네트워크를 학습한다. 기존의 차원 축소 방식이 왜곡 항만 이용하여 학습을 수행한 데 반해, 엔트로피 최소화 기반 방식은 율(rate) 항을 함께 이용하여 율-왜곡 최적화를 수행한다. 여기서 율 항의 계산은 학습 가능한 성분 별 모수적(parametric) 확률 모델을 기반으로 이루어지며, 학습된 확률 모델은 추론 시의 엔트로피 부복호화에도 이용된다. 상기 차원 축소 방식과 엔트로피 최소화 방식을 통해 변환된 피쳐 데이터는 전혀 다른 특성을 보이며, 차원 축소 방식의 경우 데이터의 불확실성이 크게 증가하는 데 반해, 엔트로피 최소화의 경우 일부를 제외한 대부분의 데이터가 0의 값에 근접하는 특성이 있다.

그림 1 은 엔트로피 최소화 기반의 뉴럴넷 이미지 압축 방식의 구성을 간략히 도식화한 것이다. 그림에서 보이듯이 입력 이미지는 인코더 네트워크를 통해 피쳐 형태로 변환되고, 디코더 네트워크는 변환 피쳐를 다시 이미지 형태로 복원한다. 또한 학습 시의 율 항 계산 및 테스트 시의 엔트로피 부복호화를 위한 확률 분포 모델을 이용한다. 최근 제안된 엔트로피 최소화 기반 주요 방식들[3-8]은 변환된 피쳐에 대해 독자적인 확률 분포 추정 모델을 이용하는데, 기술의 발전에 따라 확률 분포의 모수를 직접 학습하는 방식[3,5]부터, 네트워크를 이용하여 모수를 문맥 적응적으로 추정하는 방식[6,7,8]까지 다양하다. 본 논문에서는 최근 더 활발히 연구가 진행되고 있는 엔트로피 최소화 기반 종단간 뉴럴넷 이미지 압축 기술 중 [6]의 방식을 대상으로 데이터 분석을 진행한다.

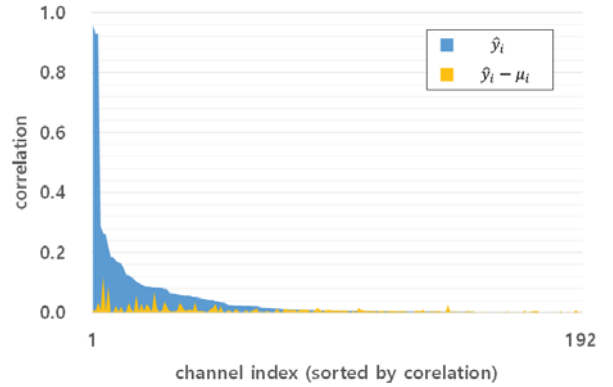


그림 1 종단간 이미지 압축 네트워크로 변환한 피쳐 데이터의 인접 성분 간 상관계수(절대값)

3. 피쳐 공간 상관관계 분석

엔트로피 최소화 기반 종단간 뉴럴넷 이미지 압축기술에서, 변환된 피쳐 데이터 성분 간의 높은 상관관계는 데이터의 높은 중복성을 나타내며 피쳐 데이터 내에 제거해야 할 중복성이 잔존함을 의미한다. 이에 변환 데이터에 대한 피쳐 성분 간 상관관계 분석이 필요하다. 그림 1 에서 보이듯, 변환된 피쳐는 3 차원 형태이며, 피쳐의 폭(W)과 높이(H)로 구성된 공간 축과 채널(C) 축으로 구성된다. 본 논문에서는 실험을 통해 [6] 방식에 따라 변환된 피쳐 데이터의 공간적 (spatial) 상관관계(correlation)와 채널 간 (inter-channel) 상관관계를 측정하였으며, 더불어 [6] 방식에서 데이터의 중복성 문제를 해결하기 위해 도입한 문맥 적응적 피쳐 성분 예측 기술을 통해 예측이 반영된 피쳐 데이터 간의 상관관계가 어떻게 변하는지 측정한다.

피쳐 성분 간 공간적 상관관계 측정을 위해 본 논문에서는 CLIC 학습용 이미지셋[9]에서 임의로 추출한 5,000 장의 512 x 512 사이즈의 패치를 이용하였으며, 이 경우 변환 피쳐의 공간적 크기는 32 x 32 이다. 각 패치 별로 공간적으로 인접한 200 개의 피쳐 쌍을 이용하여 인접 성분 간 피쳐스 상관계수의 절대값을 측정하였다. 상기 측정은 피쳐 데이터를 구성하는 192 개 채널 각각에 대해 개별적으로 수행되었다. 그림 1 에서 파란색 바 그래프는 변환 피쳐의 인접 성분 간 상관계수의 절대값을 나타낸 것으로, 상관관계가 큰 순서대로 도식화한 것이다. 그림에서 보이듯 대부분의 채널에서 비교적 낮은 상관계수의 절대값을 보이나, 일부 채널에서는 공간적으로 인접한 성분 간의 상관관계가 유지됨을 알 수 있다. 이는 데이터의 모든 공간 영역에 반복적으로 동일한 필터를 적용하는 컨볼루션(convolutional) 레이어의 특성에 기인한 것으로,

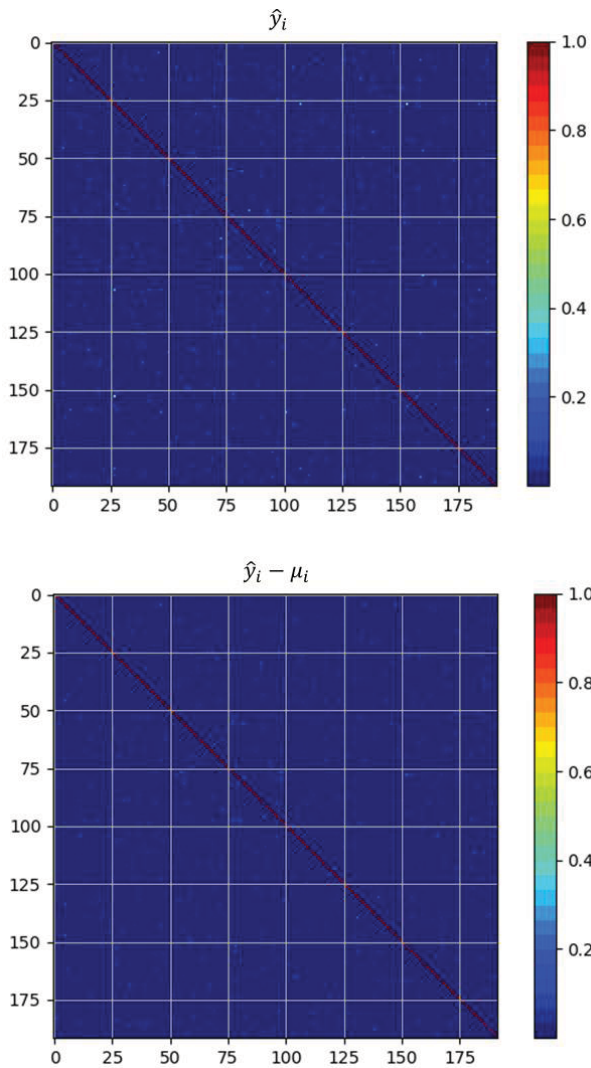


그림 3 종단간 이미지 압축 네트워크로 변환한 피쳐 데이터의 채널 간 상관계수(절대값)

[6]에서는 상기 공간적인 중복성을 제거하기 위해 피쳐의 인접 성분을 문맥으로 활용하여 인코딩(디코딩)할 성분의 평균과 표준편차 값을 예측하는 네트워크를 도입한다. 그림 2 의 노란색 바 그래프는 피쳐 성분 값과 문맥을 이용해 예측한 성분값의 차분 데이터에 대해 인접 성분간 상관계수의 절대값을 측정했 것으로, 그림에서 보이듯이 [6]은 문맥 기반 예측 기능을 이용하여 대부분의 인접 성분 간 공간적 중복성을 제거하고 있다.

채널 간 상관관계의 경우에도 동일한 이미지 패치 셋을 대상으로 측정하였으며, 동일한 공간적 위치 상에 속하는 서로 다른 채널의 피쳐 성분 간의 상관계수를 측정하였다. 그림 3 은 측정된 채널 간 상관계수를 그래프 행렬(gram matrix) 형태로 도식화한 것이다. 그림 3 의 상단에서 보이듯이 채널 간 상관관계의 경우 공간적 인접 성분 간의 상관관계와 달리 인코더 네트워크에 의한 변환 단계에서 대부분 제거되고 있음을 알 수

있다. 이는 학습된 인코더 네트워크의 비상관화(decorrelation) 성능이 매우 우수하고, 변환된 피쳐 데이터 내 채널 간 데이터 중복성이 매우 낮음을 보인다. 간헐적으로 채널 간 상관관계가 잔존하는 경우가 있으나, 이는 그림 2 의 하단과 같이 [6]의 문맥 기반 예측 기능을 통해 대부분 제거된다.

4. 결론

본 논문에서는 엔트로피 최소화 기반 종단간 이미지 압축 네트워크의 피쳐 공간 데이터에 대한 공간적 상관관계와 채널 간 상관관계를 상관계수 측정을 통해 보였다. 채널 간 상관관계는 대부분 인코더 네트워크를 통한 변환 단계에서 제거됨을 확인하였고, 반면 피쳐 인접 성분 간의 공간적 상관관계의 경우 컨볼루션 네트워크의 특성 상 변환 피쳐의 일부 채널에 잔존함을 확인하였다. 또한 일부 잔존하는 공간적 또는 채널 간 데이터 중복성은 최근 제안된 종단간 이미지 압축 네트워크의 문맥 기반 예측 기능을 통해 효과적으로 제거됨을 보였다

감사의 글

이 논문은 2018 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (2017-0-00072, 초실감 테라미디어를 위한 AV 부호화 및 LF 미디어 원천기술 개발)

참고문헌

[1] G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell, "Full resolution image compression with recurrent neural networks," in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2017. [Online]. Available: <http://arxiv.org/abs/1608.05148>

[2] N. Johnston, D. Vincent, D. Minnen, M. Covell, S. Singh, T. Chinen, S. Jin Hwang, J. Shor, and G. Toderici, "Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

[3] J. Ballé, V. Laparra, E. P. Simoncelli, "End-to-end optimized image compression", Proc. Int. Conf. Learn. Representat, pp. 1-27, 2017.

[4] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," 6th Int. Conf. on Learning Representations, 2018, accepted. [Online]. Available: <https://openreview.net/forum?id=rkcQFMZRb>.

[5] L. Theis, W. Shi, A. Cunningham, and F. Huszar, "Lossy image compression with compressive autoencoders," International Conference on Learning Representations, 2017.

- [6] J. Lee, S. Cho, and S.-K. Beack, "Context-adaptive entropy model for end-to-end optimized image compression," in the 7th Int. Conf. on Learning Representations, May 2019.
- [7] D. Minnen, J. Ballé, and G. Toderici, "Joint autoregressive and hierarchical priors for learned image compression," in Advances in Neural Information Processing Systems, May 2018.
- [8] J. Lee, S. Cho, M. Kim, "An End-to-End Joint-learning Scheme of Image Compression and Quality Enhancement with Improved Entropy Minimization," arXiv preprint arXiv:1912.12817, 2019. [Online]. Available: <https://arxiv.org/abs/1912.12817>
- [9] "Workshop and challenge on learned image compression," 2019. [Online]. Available: <https://www.compression.cc/>