

CAE를 이용한 Fresnelet 영역의 생성

*이재은 김동욱 서영호

광운대학교 전자재료공학과

*wodms6364@kw.ac.kr

Generation of Fresnelet region using CAE

*Lee, Jae-Eun Kim, Dong-Wook Seo, Young-Ho

Kwangwoon University

요약

본 논문에서는 디지털 홀로그래프 영상을 Fresnelet 변환을 하여 상관도를 확인할 수 있는 데이터로 바꾸고, 컨볼루션 오토 인코더(Convolutional Autoencoder, CAE)를 이용해 압축하고 생성하는 방법을 제안한다. 컨볼루션 계층과 채널 수가 다른 2개의 네트워크로 실험한다. CAE의 인코더를 수행해 영상을 압축하고 디코더를 통해 복원한다. 원본 영상의 Fresnelet 영역과 2개의 네트워크를 진행하여 생성된 Fresnelet 영역을 다시 역 Fresnelet하여 압축률에 따른 PSNR을 비교, 분석한다.

1. 서론

홀로그래프 디스플레이는 실제 공간상에 입체정보를 재현하는 기술로 기존의 3차원 디스플레이 장치에서의 문제점으로 야기되는 시정 시 불편감을 제거할 수 있는 차세대 영상시스템이다. 디지털 홀로그래프를 상용화하기 위한, 시간과 메모리 측면에서의 데이터 크기에 대한 효율성 문제는 중요한 이슈이다.

본 논문에서는 디지털 홀로그래프 영상을 Fresnelet 변환하여 컨볼루션 오토인코더를 이용해 압축하는 방법을 제안한다. 컨볼루션 계층 수와 채널 수를 다르게 설정한 2개의 네트워크를 제안하고, 진행한 결과 영상을 비교, 분석한다.

2. Fresnelet 변환

디지털 홀로그래프는 데이터의 상관도가 적기 때문에 Fresnel 변환을 사용한다. Fresnel 변환은 입력평면으로부터 거리 z 만큼 떨어진 평면에서의 회절현상을 모델링한 수식으로 식 (1)에 나타낸다. $f(x)$ 는 입력, $g(s)$ 는 출력이고 F 는 Fresnelet 변환 필터이다. λ 는 광원의 파장이고 Δx 와 Δs 는 입력 평면과 출력 평면의 샘플링 거리이다[1].

$$\begin{aligned} g(s) &= Ff(x), F = \frac{\Delta s}{\sqrt{\lambda z}} U W V \\ U &= \text{diag}[u_x] u_x = \exp\left[\frac{j\pi}{\lambda z}(x\Delta x)^2\right] \\ V &= \text{diag}[v_x] v_x = \exp\left[\frac{j\pi}{\lambda z}(s\Delta s)^2\right] \\ W &= \text{diag}[w_{xs}] w_{xs} = \exp\left[-\frac{j2\pi}{\lambda z}(x\Delta x)(s\Delta s)\right] \end{aligned} \quad (1)$$

식 (1)의 Fresnel 기저 함수를 이용하여 식 (2)와 같이 각각 저대역 필터(F_0)와 고대역 필터(F_1)를 만들어 Fresnelet 변환을 수행한다. 식 (3)은 역 Fresnelet 변환을 위한 필터이다. L 과 H 는 각각 웨이블릿 변환의 저대역 필터와 고대역 필터로 b-spline 필터를 사용한다.

Fresnelet 변환은 Fresnel 기저함수를 포함하는 식 (2)와 (3)의 필터를 사용하여 1레벨만 수행하고 이 후 웨이블릿 필터를 이용하여 진

행한다[2].

$$F_0 = \frac{\Delta s}{\sqrt{\lambda z}} L U W V, F_1 = \frac{\Delta s}{\sqrt{\lambda z}} H U W V \quad (2)$$

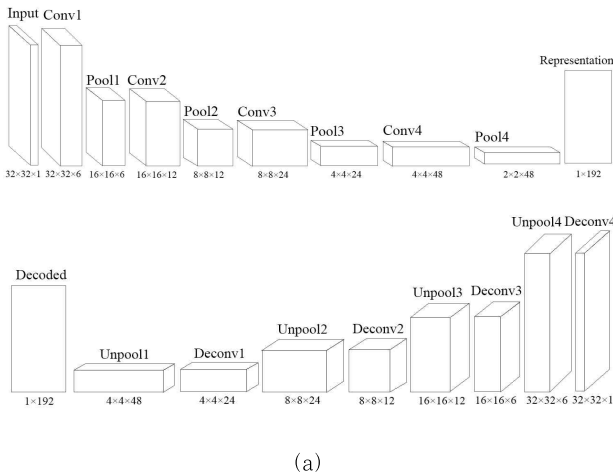
$$F_0^* = \frac{\Delta x}{\sqrt{\lambda z}} V^* W^* U^* L^t, F_1^* = \frac{\Delta x}{\sqrt{\lambda z}} V^* W^* U^* H^t \quad (3)$$

3. 제안하는 CAE(Convolutional Autoencoder) 네트워크 구조

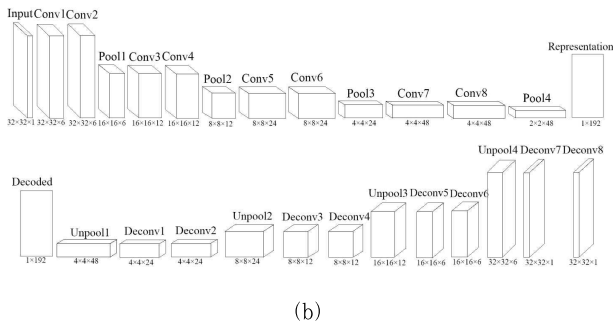
오토인코더는 대표적인 비지도 학습을 하는 신경망이다. hidden 계층의 노드 수가 입력 노드 수보다 적고, 인코더와 디코더가 대칭적인 구조를 가지며, 출력 데이터를 입력 데이터와 유사하게 복원한다. CAE는 오토인코더에서 컨볼루션을 수행하는 신경망이다. CAE는 영상의 특징을 추출하여 학습을 진행하기 때문에 차원 감소와 복원을 할 때 효율적이다.

홀로그래프 데이터는 복소 형태로 구성되어 있기 때문에 실수부와 허수부로 나누어 학습한다. 입력데이터의 대칭성을 유지하기 위해 최댓값과 최솟값의 절댓값을 비교하여 큰 값을 이용해 정규화를 수행한다. 또한 CAE의 컨볼루션 계층에 적용하기 위해 (1,가로,세로,1)의 4차원 형태를 가진 데이터로 변환하여 사용한다.

제안하는 CAE 네트워크는 그림 1의 두 가지 네트워크로 구성한다. 그림 1(a)의 CAE1은 4개의 컨볼루션 계층, 4개의 풀링 계층, 4개의 전치된 컨볼루션 계층, 4개의 언풀링 계층, 1개의 전결합층으로 구성된다. 인코더에서 컨볼루션 계층과 풀링 계층 1개씩 진행시, 채널 수를 6, 12, 24, 48로 증가시키면서 진행한다. 디코더에서는 인코더와 대칭적으로 48, 24, 12, 6으로 감소시키면서 진행한다. (b)의 CAE2은 8개의 컨볼루션 계층, 4개의 풀링 계층, 8개의 전치된 컨볼루션 계층, 4개의 언풀링 계층, 1개의 전결합층으로 구성된다. CAE1에 비해 컨볼루션 계층과 전치된 컨볼루션 계층이 2배로 증가하였고, 채널은 4, 8, 16, 32로, CAE1에 비해 감소한 것을 볼 수 있다. 2개의 네트워크 모두 필터 크기는 7로 설정하여 컨볼루션을 수행하고, 풀링계층은 스트라이드가 (2,2)인 최대풀링으로 설정하여 진행한다.



(a)



(b)

그림 1. CAE 구조; (a) CAE1, (b) CAE2
Figure 1. CAE structure; (a) CAE1, (b) CAE2

4. 실험 결과

표 1. 네트워크에서 사용한 하이퍼파라미터
Table 1. hyperparameter used in network

batch size	lr	optimization	activation function	epoch
16	0.0001	Adam	leaky ReLU, sigmoid	10,000

표 2. CAE1과 CAE2의 결과
Table 1. Result of CAE1 and CAE2

	허수부 손실값	실수부 손실값	PSNR	압축률
CAE1	0.752167	0.710001	25.67dB	5.34:1
CAE2	0.854005	0.803510	25.05dB	8:1

사용한 입력 데이터는 2,048×2,048의 홀로그램 영상 6개를 7레벨 Fresnelet 변환을 수행하여 32×32의 서브밴드를 24,576개로 구성하였다.

표 1은 CAE에서 사용한 하이퍼파라미터를 보여준다. 하이퍼파라미터는 CAE1과 2에서 동일하게 설정하였다. 배치 사이즈(batch size)를 16, 학습률(learning rate, lr)은 0.0001로 하여 Adam 최적화(optimization)를 진행한다. 활성화함수(activation function)는, 마지막에 위치한, 전치된 컨볼루션 계층을 진행할 때 sigmoid 함수를 사용하고 그 외에는 모두 leaky ReLU 함수를 사용한다. 학습을 진행할 때 모든 훈련 데이터를 학습한 반복횟수를 나타내는 에폭(epoch)은 허수부와 실수부 네트워크에서 각각 10,000번씩 진행한다.

표 2는 네트워크 2개의 손실값, 압축률과 결과를 역 Fresnelet하여 구한 PSNR을 비교한다. CAE1의 허수부와 실수부 네트워크의 손실값은 0.752167, 0.710001이고 CAE2의 경우, 0.854005과 0.803510를 가진

다. CAE1의 경우, 인코더 출력의 크기는 192(2×2×48)이고 1024(32×32)인 입력 데이터와 비교했을 때 압축률은 5.34:1이 된다. 네트워크의 출력 데이터를 역 Fresnelet 변환한 결과의 PSNR은 25.67dB를 가진다. CAE2의 인코더 출력 크기는 128(2×2×32)이 된다. 그 결과, 압축률은 8:1, 결과의 PSNR은 25.05dB를 가진다.

그림 2은 역 Fresnelet 변환한 결과 영상이다. (a)는 입력 데이터를, (b)는 CAE1 네트워크의 출력 데이터를, (c)는 CAE2 네트워크의 출력 데이터를 적용한 영상이다. 네트워크의 손실 값으로도 예측이 가능하지만 PSNR은 CAE1의 결과 영상이 CAE2의 결과 영상보다 높게 측정됐다. 하지만 두 영상의 압축률을 생각해보았을 때, 5.34:1과 8:1로 거의 2배정도 차이가 나는 것에 비해 PSNR 차이는 약 0.62dB 정도 가지는 것을 볼 수 있다.

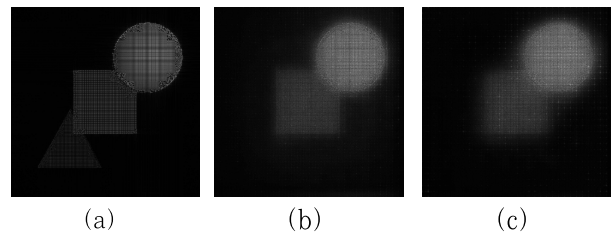


그림 2. 역 Fresnelet 변환 결과; (a) 원본 영상, (b) CAE1을 수행한 영상, (c) CAE2를 수행한 영상

Fig. 2. Results of inverse Fresnelet transform; (a) original image, (b) image using CAE1, (c) image using CAE2

5. 결론

제안하는 CAE 네트워크는 필터 크기가 7인 컨볼루션 계층, 스트라이드를 2로 설정한 최대풀링 계층, 전치된 컨볼루션 계층, 언풀링 계층으로 구성되어있다. 2개의 네트워크를 제안하는데, CAE2는 CAE1에 비해 컨볼루션 계층을 2배로 가지고 채널 수를 다르게 설정했다.

구현결과, CAE1이 5.34:1인 압축률에서 PSNR 25.57dB를 갖고 CAE2는 8:1인 압축률에서 PSNR 25.05dB를 갖는다. 2개의 네트워크를 비교해보면, CAE2는 PSNR 대비 높은 압축 효율을 가지고 있음을 알 수 있다. 따라서 더 깊은 컨볼루션 오토인코더 네트워크를 구성한다면, 더 높은 압축률이 나올 것이라 기대된다.

감사의 글

이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2018RID1A1B07043220).

참고문헌

[1] U. Schnar and W. Jueptner, "Digital Holography", Springer, Berlin, Germany, 2005.
[2] M. Nazeer and D.-G. Kim, "An Efficient Data Hiding Technique in Frequency domain by using Fresnelet basis", WCE 2012, London, U.K, 2012.