

영상의 복잡도 특징을 기준으로 양자화 왜곡에 대한 최소 인지 왜곡 모델

기세환, 김문철

한국과학기술원 전기 및 전자 공학과
shki@kaist.ac.kr, mkim@ee.kaist.ac.kr

Just noticeable quantization blur model based on the DCT complexity feature of the image

Sehwan Ki, Munchurl Kim

School of Electronic Engineering, Korea Advanced Institute of Science and Technology

요 약

본 논문에서는 기존의 인지적 영상 압축 기법에 사용되었던 Just Noticeable Distortion(JND) 모델이 압축과정에서 생기는 왜곡인 양자화 왜곡에 적합하지 않는다는 것을 보이고, 그 한계점을 해결하기 위하여 Just Noticeable Blur(JNB)의 개념을 적용하여 영상 압축에 적합한 모델을 제시하였다. 주파수 공간에서 영상의 복잡도 특징을 나타내는 Spectral Contrast Index(SCI) 값을 사용해서 영상의 DCT 블록별 JNB 를 추정하고 이를 기반으로 영상의 DCT 계수 값을 감소시켜 최신의 DCT 기반 JND 를 적용한 인지적 압축 영상에 비해 더 낮은 PSNR 을 가지면서 왜곡도 인지되지 않는 영상을 얻을 수 있었다. 새롭게 제시한 모델을 적용하면 인지적 영상압축에서 기존의 방법보다 더 낮은 비트율로 유사한 인지적 화질 성능을 발휘할 것으로 예상된다.

1. 서론

스마트폰의 발달로 이제는 전문가가 아닌 일반인들도 자유롭게 초고화질 영상을 촬영하고 저장하는 것이 일상이 되었다. 스마트폰으로 4K 또는 8K 의 영상을 촬영하다 보면 매우 큰 저장 용량이 필요하고 전송을 위해서도 많은 데이터 사용료가 필요로 되는 문제가 발생하였다. 따라서 영상 압축이 중요한 화두로 떠오르게 되었고 영상 통신의 목적이 사람이 보는 영상을 전송하는 것이므로 사람이 인지하지 못하는 부분을 크게 압축하는 인지적 영상 압축방식에 대한 관심이 높아졌다.

인지적인 압축 방식에는 사람의 영상 인식에 대하여 단순화된 가설을 적용하는 방법과 실제로 다양한 일반인들에게 한 실험을 통해 왜곡이 인지되는 경계를 찾는 방법으로 나눌 수 있다. 먼저, 가설을 적용하는 방법으로는 사람의 시각인지가 구조적인 연관성을 중요하게 인식한다는 것과 객체의 경계의 값에 민감하다는 것과 같이 사람 인지 체계를 단순화하는 모델을 사용하였다 [1]. 이와 같은 방법은 수학적으로 보다 간단하게 구현이 가능하다는 장점이 있으나 실제로 사람의 인지와는 연관성이 떨어지는 부분이 존재하기도 한다. 따라서 일반적인 실험자가 다양한 영상에 대해서 실험을 통해 얻은 결과를 가지고 영상의 특징과 연관성을 찾아 모델을 만드는 것이 더 정확한 방법이라고 할 수 있다.

기존의 인지적 압축 방식에는 Just Noticeable Distortion(JND)를 많이 사용하였다[2],[5]. 이 값은 영상에 일정한 왜곡을 계속 가하면서 전체의 50%의 실험자가 왜곡을 인지하는 그 순간의 왜곡의 크기를 의미하는 값이다. 사람이 왜곡을 인지 못하면서 최대한으로 많이 압축하는 것, 즉 인지적 영상 압축의 목표와 동일하므로 JND 가 많이 적용되어서

사용되었다. 하지만, JND 모델링시 적용된 왜곡의 타입은 영상 압축에서 발생하는 양자화 왜곡과는 다른 특징을 지닌다. 영상 압축과정은 양자화를 통해서 영상의 에너지를 줄이기 위하여 DCT 계수의 크기가 줄어드는 왜곡을 발생시킨다. 하지만 JND 모델을 실험적으로 만들 때에는 무작위로 모든 DCT 계수의 값이 한 개씩 증가되는 왜곡을 발생 시켰을 때, 사람의 왜곡 인지 임계값을 찾았기 때문에 DCT 계수가 동일하게 줄어드는 양자화 왜곡에 대해서는 적합하지 않다. 따라서 새롭게 영상 압축 양자화 왜곡에 적합한 인지적 경계 값을 찾는 실험이 필요하고 이를 이용하여 압축의 양자화를 하는 것이 인지적 영상 압축에 더 효과적인 것으로 판단된다.

양자화 왜곡은 영상의 DCT 계수 값을 줄이는 과정에서 발생되며 영상의 경계가 흐려지는 blur 와 유사한 현상을 보인다. 이와 같은 왜곡에 대한 Just noticeable blur(JNB) 실험은 기존에 연구되지 않았고, 다만 Gaussian blur 에 대한 JNB 연구가 존재하였다[3]. 이는 영상의 선명도 평가 방법의 기준으로 사용되는 것으로 영상의 구체적인 특징으로 모델링을 한 것이 아닌 단순히 최대 밝기와 최소 밝기만의 값을 가지고 결정하는 모델이었다. 따라서 본 논문에서는 주파수 공간에서 영상의 복잡도를 나타내는 Spectral Contrast Index(SCI) 값을 적용하여 영상의 복잡도 기준을 가지고 양자화 정도를 결정할 수 있는 모델을 새롭게 제시하였다. 이를 양자화 왜곡에 따른 JNB 실험을 통해, 구한 경계 값과 제시한 모델이 연관성을 보이는 것을 확인 할 수 있었다.

본 논문의 구성은 2 절에서 새롭게 양자화 왜곡에 맞는 JNB 실험에 대한 설명을 소개하고 3 절에서 설계한 실험의 결과 값과 기존의 JND 모델을 사용하였을 때의 영상과의 비교를 나타낸다. 마지막으로 4 절에서 논문의 결론을 맺는다.

2. 양자화 왜곡 JNB 실험 방법

기존의 주파수 공간에서의 JND 실험에서는, 일정 크기의 영상을 각각의 DCT 계수 한 개씩 일정 값씩 증가시키는 왜곡을 가하고 원본과 차이를 인지하는 왜곡 값을 측정하는 방식으로 진행되었다. 그리고 각각의 계수들을 독립적인 관계라고 보고 확률 식을 만들어서 하나의 DCT block 에서 왜곡이 인지될 확률을 나타내는 형식으로 표현하였다[4],[5].

하지만 양자화 왜곡처럼 DCT 계수를 줄이는 방식으로는 한 개의 계수를 0 으로 줄인다고 왜곡이 눈에 띄지 않을 수도 있고, 또한 독립적으로 왜곡이 생기는 것이 아니라 연관된 계수들과 상충작용이 사라지게 되면서 좁쌀 같은 왜곡이 생기는 것을 볼 수 있었다. 이런 문제들은 JND 모델링에서 적용했던 각각의 계수들이 독립적으로 작용한다는 가정을 어기기 때문에 새로운 실험 방법을 설계하였다.

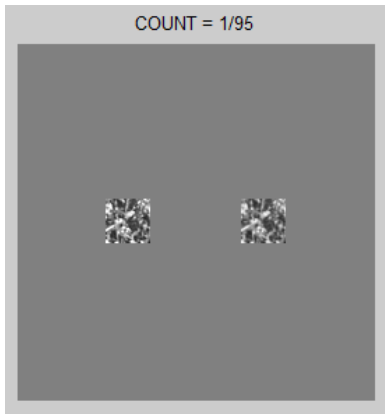


그림 1. 양자화 왜곡 실험 화면

일반적으로 사람의 왜곡 인지 수준이 영상의 복잡도와 연관성이 크기 때문에 양자화 왜곡에도 영상의 복잡도가 중요한 특성으로 영향을 미치기 때문에, JND 모델링에서 사용되었던 주파수 공간의 SCI 를 사용하여 실험 영상을 추출하였다[4]. SCI 는 DCT 영역에서 구해지는 특징 값으로서, 영상의 contrast 와 kurtosis 값의 비로 나타낼 수 있다[4]. SCI 는 사람이 영상의 화질을 인지할 때 전체적인 밝고 어두움이 흐릿한 영상에 비해 뚜렷한 영상에서 더 복잡함을 느끼고, DCT 계수들의 에너지가 저 주파수 대역에 몰려서 뾰족한 분포를 가지는 것보다 고르게 퍼져 완만한 분포를 가지는 경우가 사람이 인지하기에 더 복잡하다는 시각 인지 특성을 반영하고 있다.

$$\tau_{SCI} = \frac{C_t^\alpha}{K_t^\beta} \quad (1)$$

τ_{SCI} : Spectral Contrast Index

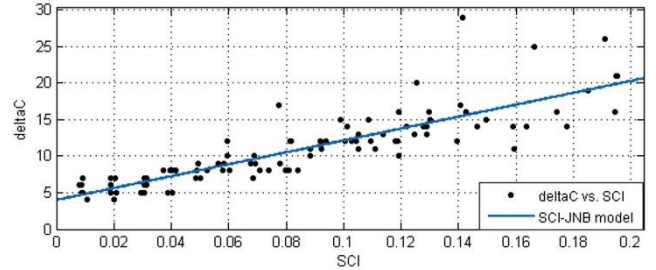
C_t : Contrast intensity

K_t : Kurtosis

TID2013 영상 화질평가 데이터를 사용하여 SCI 값이 0.01 부터 0.2 까지 총 95 장의 32x32 크기의 영상을 추출하여 일반적인 시각을 가진 피실험자 10 명에게 왜곡 인지 실험을 하였다. 영상에 가해진 왜곡은 양자화로 영상의 DCT 계수가 줄어드는 것을 가정하여 8x8 DCT 계수가 모두 동일하게 1 만큼 감소 되고 계수가 0 보다 작아지지 않도록 설계하였다. 실험 환경은 일반적인 연구실 조도 환경이고, 삼성 55 인치

Full HD TV 에서 시청거리를 화면높이의 1.9 배(133 cm)로 고정하고 실험을 진행 하였다. 실험은 원본 영상과 왜곡영상을 같은 화면에 출력을 하고 일정 왜곡을 계속 가하면서 실험자가 원본과 차이를 느끼는 지점을 측정한다. 왜곡이 인지 되었을 때의 감소된 DCT 계수 값 ΔC 가 그 영상의 JNB 값이 된다

3. 실험 결과 및 분석



$$\widehat{\Delta C} = 81.48 \times SCI + 3.96 \quad (2)$$

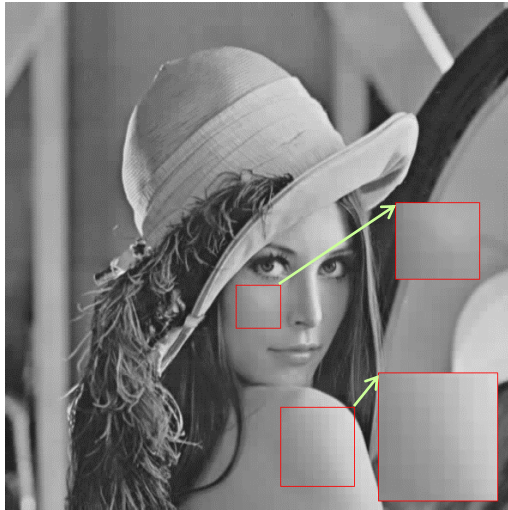
ΔC : 실험적으로 측정된 왜곡이 인지 될 때의 DCT 계수 감소량

그림 2. 영상의 복잡도와 실험으로 측정된 ΔC 와의 관계

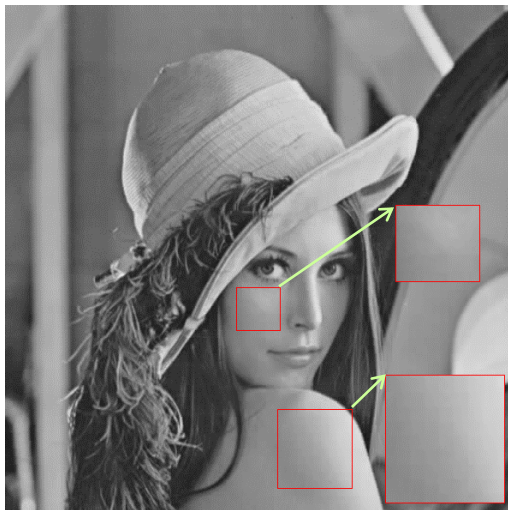
실험을 통해 얻은 ΔC 값과 영상의 복잡도 즉, SCI 간의 연관성을 그림 2 에 나타내었다. 그림 2 에서 볼 수 있듯이 두 변수가 선형관계를 보이는 것을 알 수 있다. 작은 SCI 값을 가지는 영상은 작은 ΔC 값을 가지고 큰 SCI 을 가진 영상은 큰 ΔC 값을 나타내는 것을 볼 수 있다. 일반적으로 blur 왜곡은 복잡한 영상에서 쉽게 인지되는 것으로 알려져 있지만 실제로 DCT 계수의 값의 변화량으로 보면 단조로운 영상에 비해서 복잡한 영상의 경우 더 많이 변해도 왜곡 인지가 덜 되는 경향이 있다. 이는 비율로 보는 것과 값의 변화의 차이로 볼 수 있다. Gaussian blur 처럼 DCT 계수들을 비율적으로 감소시키는 경우, 영상이 복잡할 때 작은 비율로 변화되어도 DCT 계수의 크기 변화로 보면 단순한 영상에서 큰 비율로 변하는 것보다 많이 변하는 것이다.

단순한 일차함수의 선형회귀 방법을 사용해서 얻은 (2)의 식인 SCI-JNB 모델을 사용해서 일반적인 영상에 DCT 계수 감소 왜곡을 JNB 만큼 왜곡을 준 영상을 만들고 주관적으로 평가를 해보았다. TID2013 영상 화질평가 데이터에서 9 장의 원본 영상과 JNB 만큼 왜곡한 영상을 비교했을 때 시각적으로 차이가 나지 않는 것을 확인할 수 있었다. 기존의 Bae's DCT-based JND 모델[5]을 적용해서 동일한 왜곡을 JND 경계까지 가했을 경우는 왜곡이 인지가 되었다. 특히 그림 3 에서 볼 수 있듯이 단조로운 영역에서 blocking 왜곡이 두드러지게 드러났다. 이는 JND 모델을 만들 때 DCT 계수를 모두 감소시키는 형태의 왜곡에 맞춰서 실험을 한 것이 아니기 때문에 단조로운 영역에서 정확한 결과를 얻지 못하였다.

하지만 새롭게 제시하는 SCI-JNB 모델에서는 실험 영상의 크기가 32x32 크기로 16 개의 8x8 DCT 블록에 모두 동일한 왜곡을 가하는 형식으로 수행했기 때문에 피실험자가 blocking 왜곡도 같이 고려할 수 있었다. 따라서 기존의 블록 JND 에서는 고려하지 못했던 영상의 단조로운 부분의 blocking 왜곡에 대해 더 정확하게 왜곡인지 경계를 예측할 수 있다. 그리고 복잡한 영역에서는 DCT 계수를 더 줄여도 왜곡이 인지가 되지 않는 것을 적용하여 Bae's JND 모델을 적용한 영상에 비해 왜곡 값을 더 증가시킬 수 있었다.



(a) JND 왜곡 영상



(b) SCI-JNB 왜곡 영상

그림 3. Lena 영상 양자화 인지 경계 왜곡 영상 비교

아래의 표 1 에서 보면 주관적 화질 측면에서는 SCI-JNB 모델을 적용하여 왜곡한 영상이 인지적으로 왜곡이 보이지 않았고 Bae's JND 모델을 적용한 경우 60% 이상의 실험 영상에서 왜곡이 인지된 것을 확인할 수 있었다. 그리고 얼마나 많은 왜곡을 가했는지에 대한 값인 PSNR 값은 평균 2.4 dB 더 낮으므로 새롭게 제시하는 SCI-JNB 모델이 실제 사람의 DCT 계수의 감소 왜곡 즉 양자화 왜곡에 JND 모델에 비해 더 적합하다.

표 1. 양자화 왜곡 영상의 PSNR 과 왜곡 인지 여부 비교

Test image (TID 2013)	PSNR (dB)		왜곡의 인지 여부	
	Bae's JND	SCI-JNB	Bae's JND	SCI-JNB
stream	31.46	30.02	X	X
paintedhouse	36.64	32.56	X	X
flowersonih35	36.07	28.95	0	X
lena	37.39	36.87	0	X
monarch	37.70	35.13	0	X
sailing1	36.77	33.87	0	X
bikes	33.92	31.71	X	X
womanhat	37.18	36.93	0	X
parrots	38.14	38.29	0	X
Average	36.17	33.82	66.7%	0%

4. 결론

본 논문에서는 기존의 인지 비디오 부호화 방식에서 사용되던 JND 모델이 실제적으로 압축 과정에서 발생하는 왜곡, 즉 양자화 왜곡에 적합하지 않다는 것을 제시하였다. 따라서 이와 같은 문제를 해결하기 위해서 양자화 왜곡과 유사한 DCT 계수를 감소시키는 왜곡을 다양한 복잡도에 영상에 대하여 실험을 하였고, 실험의 결과로 복잡도를 나타내는 SCI 라는 DCT 기반 복잡도 측정법과 왜곡 인지 경계 값의 선형 관계를 찾을 수 있었다. 단순한 선형 함수 관계로 영상의 특성에 따라 얼마만큼의 DCT 계수를 감소시켜도 인지적으로 차이가 없는지에 대한 왜곡 임계치를 예측할 수 있었다. DCT 계수를 사용하는 방법이므로 다소 계산 량이 증가될 수 있지만 픽셀 값만 사용하는 방법에 비해 매우 높은 정확도를 보이는 것을 확인할 수 있었다. 그리고 이와 같은 왜곡이 많이 나타나는 영상 압축에서는 실제로 전송되는 데이터가 DCT 변환된 계수이므로 손쉽게 적용 될 수 있다. 향후 연구로, 실제 최신 영상 압축 표준인 High Efficiency Video Codec (HEVC)에 적용하여 다른 인지적 영상 압축 방식 기법들과 성능 비교를 진행할 예정이다.

Acknowledgement

본 논문 연구는 연구재단 중견연구자사업 핵심연구(개인) 과제 (과제번호: 2014R1A2A2A01006642)로 수행되었습니다.

참고문헌

[1] A. Rehman and Z. Wang, "SSIM-inspired perceptual video coding for HEVC," *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 497-502, July 2012.

[2] J. Kim, S.-H. Bae, and M. Kim, "An HEVC-Compliant Perceptual Video Coding Scheme based on JND Models for Variable Block-sized Transform Kernels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 11, pp. 1786-1800, Jan. 2014.

[3] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. on Image Process.*, vol. 18, no. 4, pp. 717-728, Apr. 2009.

[4] S.-H Bae and M. Kim, "A Novel Generalized DCTbased JND Profile Based on An Elaborate CM-JND Model for Variable Block-Sized Transforms in Monochrome Images," *IEEE Trans. on Image Process.*, vol. 23, no. 8, pp. 3227-3240, Aug. 2014.

[5] S.-H Bae, J. Kim, and M. Kim, "HEVC-Based Perceptually Adaptive Video Coding using a DCT-based Local Distortion Detection Probability Model," *IEEE Trans. on Image Process.*, vol. PP, no. 99, May. 2016.