

시각적 인지 중복성 제거를 위해 양자화 크기값에 적응적인 최소 인지 왜곡 기반 전처리 방법

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

JND based Video Pre-processing Adaptive to Quantization Step sizes for Perceptual Redundancy Reduction

Sehwan Ki, Munchurl Kim
Korea Advanced Institute of Science and Technology, School of Electronic Engineering

요 약

본 논문에서는 기존의 인지 영상 부호화에 사용되던 Just Noticeable Distortion(JND) 보다 더 압축에 적합한 모델인 Just Noticeable Quantization Distortion(JNQD) 모델을 제시하고, 이를 사용한 인지적 영상 압축 방법을 제안한다. 제안하는 인지적 영상 압축 방식은 영상 코덱 내부의 Rate-Distortion Optimization(RDO)을 수정하지 않고 입력되는 영상의 불필요한 정보들을 미리 제거하는 전처리 과정으로서, JNQD 모델을 사용하여 보다 간단하면서 압축 효율을 크게 증가 시킬 수 있다. 기존 영상 압축의 전처리 방법들은 부호화기의 양자화 값을 전처리 과정에서 고려하지 못하여 부정확한 인지 중복성 제거 결과를 초래하였으나, 제안하는 방법은 영상의 특성뿐만 아니라 양자화 크기 값을 고려하여 적응적으로 인지 왜곡이 발생하지 않는 주관적 인지 중복성 제거를 전처리 과정에서 수행할 수 있다. 거의 유사한 주관적 품질 수준을 유지하면서 HEVC 참조 소프트웨어 대비 약 15%의 압축효율 향상을 보인다.

1. 서론

최근 들어 지상파 방송 서비스 환경하에서 4K-UHD 와 같은 초고화질 영상의 수요가 급격히 증가되는 상황에 있고, 이러한 초고화질 영상을 전송하는 것은 많은 넓은 전송 대역폭이 요구하게 된다. 따라서 주관적으로 동일한 화질을 유지하면서도 최소한의 크기로 초고화질 데이터를 효율적으로 압축하는 인지적 영상 압축 방식에 큰 관심을 보이고 있다 [1].

인지적 영상 압축 방식에서 가장 중요한 점은 사람이 왜곡을 인지하는 경계 지점을 영상의 특성에 따라 추정할 수 있어야 한다. 따라서 기존의 인지적 영상 압축 방식에서는 Just Noticeable Distortion(JND)를 많이 사용하였다 [2], [3]. 이는 영상에 일정한 크기만큼 왜곡을 가하면서 전체 피실험자들 중에서 50% 정도가 왜곡을 인지하기 시작하는 왜곡의 정도를 나타내는 임계값이다. 사람이 인지하지 못하는 최대의 왜곡이라는 개념 그대로 인지적 영상 압축과 동일한 목표이기 때문에 자주 사용되어왔다. 하지만 기존의 JND 값을 찾는 실험방법들은 영상 압축에 잘 부합되지 않는 방법으로 왜곡 임계값을 측정하였기 때문에 영상 압축에 적합하지 않다. 따라서 본 논문에서는 영상 압축에서 발생하는 blocking artifact 와 blur, luminance change 와 같은 왜곡에 해당되는 JND 값을 나타내는 Just Noticeable Quantization Distortion(JNQD) 모델을 사용해서 인지적 영상 압축 방식을 제안한다.

일반적으로 JND 를 영상 압축에 적용한다면 일반적으로 Rate-Distortion Optimization(RDO)을 JND 관점으로

변화해서 압축 효율을 개선하고자 하지만 이와 같은 방식으로는 이미 충분히 작은 값인 영상의 Residual 성분에 대한 양자화 정도를 결정하는 수준에 머무르기 때문에 높은 압축 효과를 기대하기 어렵다. 그대신, 입력되는 영상에서 미리 영상의 인지에 영향을 주지 않는 성분들을 제거하는 인지적 왜곡 경계 관점의 전처리 방식이 더 효과적이라고 볼 수 있다. 하지만 전 처리 과정에서 왜곡 인지 경계만큼 영상의 성분들을 제거하면 영상 압축을 거치면서 양자화 왜곡이 더해진 후에도 여전히 왜곡 인지 경계에 있다라는 것을 보장할 수 없다.

본 논문에서는 이와 같은 영상 압축 전처리 과정의 문제점을 해결하기 위해서 압축 영상의 왜곡 인지 영상을 생성하여 그 영상과 전처리 후 압축된 영상을 비교하여 전 처리로 제거되는 인지 중복 데이터의 양을 조절한다. 이와 같이 영상의 압축을 고려하여 전처리 과정에서 제거되는 인지 중복 데이터의 양을 적응적으로 조절할 수 있는 방법을 제안하여 기존의 영상 압축 과정에서의 양자화 효과를 고려하지 못한 전처리 방법을 개선한다. 본 논문의 구성은 2 절에서 영상 압축에 적합한 왜곡 인지 경계 모델을 제시하고, 3 절에서 양자화 효과를 고려한 인지 영상 압축의 전처리 방법과 실험 결과를 제시하고, 마지막으로 4 절에서 논문의 결론을 맺는다.

2. 양자화 효과를 고려한 왜곡 인지 경계 모델

영상 압축에서 발생하는 왜곡은 영상의 에너지를 감소시키는 방향으로 진행되면서 발생된다고 볼 수 있다. 즉, 주파수 성분으로 보면 성분들의 크기가 줄어드는 방향이라고 할 수 있다. 하지만 기존의 주파수 영역에서의 JND 모델의

경우 주파수 성분을 독립적으로 증가시켜가는 왜곡 실험을 통해 왜곡의 인지경계를 찾았다 [4]. 이와 같이 에너지가 증가되면서 발생하는 왜곡은 영상 압축에서 발생하는 것과 상관성이 낮기 때문에 기존의 왜곡 인지 경계 모델을 영상 압축에 적용할 경우 압축 후 왜곡을 쉽게 인지하는 경우가 많다. 따라서 영상의 주파수 성분의 절대값을 감소시키는 방향으로 왜곡 인지 경계 값을 측정하는 JNQD 모델을 제안하고, 이를 적용하여 영상 압축에서 더 효과적으로 왜곡 인지 경계를 찾을 수 있도록 하였다. 이를 위해, 입력 영상을 8x8 block DCT 변환 후 각 block 별로 모든 주파수 성분을 동일한 크기만큼 줄여가면서 왜곡의 인지 경계를 찾았다. 왜곡 인지 경계값(JNQD)은 영상의 특성에 따라 가변적이므로 블록 내 영상 복잡도 특성을 나타내는 Spectral Contrast Index(SCI) 값에 따라 지수함수의 형태로 결정하였다 [4].

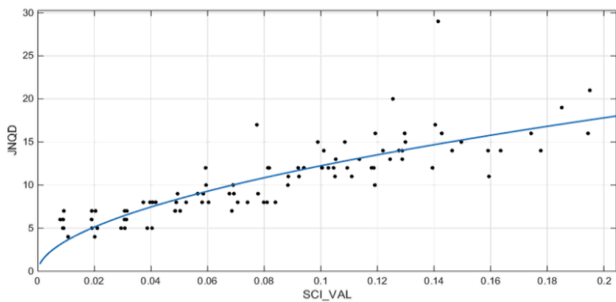


그림 1. 회귀법을 이용한 JNQD 모델링

$$JNQD = 42.5 \times \tau_{SCI}^{0.54} \quad (1)$$

$$\tau_{SCI} = \frac{C_t^\alpha}{K_t^\beta} \quad (2)$$

식 (2)에서 C_t 는 Contrast intensity이고 K_t 는 Kurtosis 이다. 또한 τ_{SCI} 는 SCI 인데, 주파수 영역에서 계산되는 영상 복잡도를 나타내는 특성 값으로서 영상의 contrast 와 kurtosis 값의 비로 나타낼 수 있다. 사람이 영상을 인지할 때 contrast 가 낮은 영상에 비해 contrast 가 높은 영상을 더 복잡한 것으로 인지한다. kurtosis 값이 낮은 영상들은 DCT 주파수 성분별 에너지가 저주파수 대역에 집중되어 뾰족한 형태를 보이고, kurtosis 값이 높은 영상들은 고주파 대역쪽으로 완만한 분포를 보이는 경향이 있다. 주관적 실험을 통해 알 수 있듯이, 단순한 텍스처 영역에서는 주파수 계수 크기를 조금만 줄여도 왜곡을 쉽게 인지하며, 복잡한 영역에서는 상대적으로 큰 왜곡에 대해서도 둔감하게 인지하는 경향을 확인할 수 있다.

3. 양자화 크기에 적응적인 인지적 압축 방법

전처리 과정을 통해 입력영상의 인지 왜곡 경계 영상을 영상 압축 과정에 입력하게 되면 압축 과정에서 발생하는 양자화 왜곡으로 인해 최종 결과 영상에서는 왜곡이 눈에 띄는 경우가 발생될 수도 있다. 따라서 본 논문에서는 입력하기 전의 원본 영상의 왜곡 인지 경계만큼 주파수 계수 값을 감소시키는 것이 아니라 출력될 때 원본의 압축영상과 전처리 후 압축영상의 인지 왜곡 경계를 비교하는 방식을 제시한다.

$$JNQD_{QP}(r, c) = JNQD(r, c) \times \alpha(r, c), \alpha \in [0, 1] \quad (3)$$

식(3)과 같이 기존의 JNQD 모델에서 얻은 왜곡 인지 경계 값에서 α 변수를 곱함으로써 주파수 계수 값을 감소시킬 수 있도록 하였다. 이때 α 는 DCT 블록 마다 다른 값을 가지도록

하여, 식(4)와 같이 영상의 특성과 영상의 압축 양자화 크기 모두에 적절하게 결정될 수 있도록 하였다.

$$\hat{\alpha} = \underset{\alpha \in [0, 1]}{\operatorname{argmin}} \frac{1}{8^2} \sum_{(i, j) \in 8 \times 8 \text{ block}} \left(C_{HM, (i, j)}^{JNQD} - C_{HM, (i, j)}^{pre}(\alpha) \right)^2 \quad (4)$$

$C_{HM, (i, j)}^{JNQD}$: 압축영상의 JNQD 영상의 8x8 block 내 (i, j) 위치의 DCT 계수

$C_{HM, (i, j)}^{pre}(\alpha)$: 전처리 후압축영상의 8x8 block 내 (i, j) 위치의 DCT 계수

따라서 최적의 α 을 구하는 방법은 식(4)와 같이 압축영상의 JNQD 영상과 임의의 α 를 JNQD 에 곱하여 전처리를 한 후 전처리를 수행하지 않고 압축한 후 복원한 영상과 전처리를 수행한 후 압축을 수행하고 이를 복원한 영상 간의 MSE 가 최소가 되는 값으로 결정하면 된다. 이 식은 NP(Non-deterministic Polynomial time) 문제이므로 그림 2 와 같이 반복적으로 α 을 변화하면서 압축을 진행하고 얻어진 MSE 를 직접 비교 후 최소가 되는 값을 결정한다.

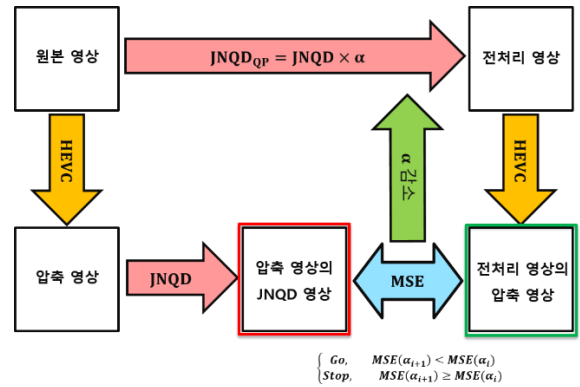


그림 2. 양자화 크기에 적응적인 인지적 압축 방법

이와 같은 방법으로 압축을 진행 하였을 때 QP가 22 일 때에는 대부분의 DCT 블록의 α 값이 1 로 선택되어짐을 그림 3 에서 볼 수 있다. 이는 영상 압축으로 인한 왜곡이 거의 발생되지 않는 작은 QP 에서는 원본에서 왜곡 인지 경계에 맞게 감소시킨 것이 출력 후에도 왜곡 인지 경계에서 크게 벗어나지 않았음을 알 수 있다. 하지만 QP 37 에서와 같이 압축 왜곡이 크게 발생하는 경우에는 α 값이 0 에 가까운 값들도 충분히 많이 선택된다. 이는 전처리 과정에서 왜곡 인지 경계 만큼 감소시킬 때 양자화 왜곡이 발생되어도 인지 범위 안에 들도록 충분한 여유를 두고 감소 시켜야 함을 보여준다.

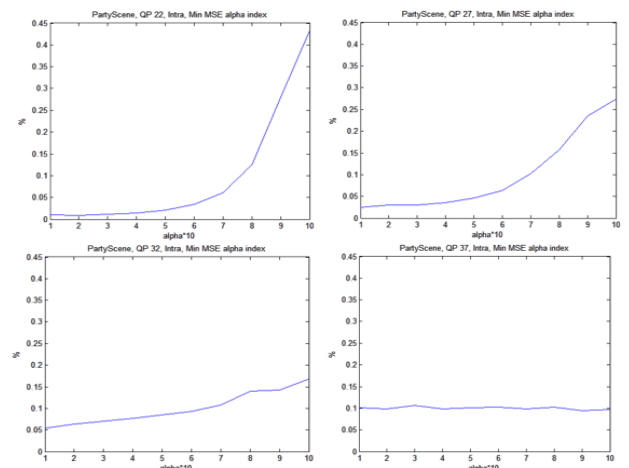


그림 3. 'PartyScene' 에서 QP 별로 선택된 α 의 확률 분포

제안하는 인지적 압축 방법의 압축 수행 결과는 표 1 에서 볼 수 있듯이 기존 HM 대비 평균적으로 15%의 압축 효율 증가를 보인다. 그리고 일반적인 시각을 가진 5 명의 피실험자가 일반 연구실 조도환경에서 55 인치 Full HD TV 을 통해 재생되는 영상을 화면 높이의 3 배(210cm)의 시청거리를 가지고 주관적 영상 품질 평가를 진행하여 주관적으로 품질의 변화가 없음을 확인하였다. 제안된 방법은 원본과 압축 영상의 왜곡 비교를 해서 전처리 값을 결정하는 것이 아니라 압축된 영상이 기준이 되기 때문에 기존의 JND 관점의 인지적 영상 압축 방법에 비해 큰 QP 값에서 더 많은 압축률 향상을 보인다. 이는 기존의 방법들은 왜곡 비교 대상이 원본 영상이기 때문에 큰 QP 값에서는 원본과의 왜곡이 이미 발생되었기 때문에 더 이상 감소시키지 못하게 된다 [3]. 하지만 본 논문에서 제안하는 방법은 비교 대상이 압축영상이기 때문에 QP 값이 크더라도 이미 왜곡이 발생한 영상의 JNPD 값을 구하여서 그 영상과의 인지 왜곡 경계 값을 찾는 방식이므로 모든 QP 에 대해서 약 8% 이상의 높은 압축률 향상을 보인다.

Sequence (all intra coding)	QP	HM	Prop. method		DMOS ($MOS_{prop.} - MOS_{HM}$)
		bitrate	bitrate	$\Delta rate(\%)$	
BasketballDrill	22	20143	16843	16.38	0.0
	27	10847	9304	14.22	0.0
	32	5816	5108	12.18	0.0
	37	3200	2943	8.01	0.0
BQMall	22	28810	22926	20.42	0.0
	27	17833	14645	17.88	0.0
	32	10688	9145	14.44	-0.3
	37	5999	5339	11.00	0.0
PartyScene	22	50999	39592	22.37	-0.3
	27	32368	25810	20.26	0.0
	32	19538	16049	17.86	0.3
	37	10817	9211	14.85	0.0
Average			15.82	-0.03	

표 1. 제안된 인지적 영상 압축 방법을 적용한 결과

4. 결론

이 논문에서는 영상 압축에서 발생하는 왜곡을 반영하는 JNPD 모델을 사용해서 영상 압축의 전처리 과정에 적용하는 방법을 제시하였다. 전처리 과정으로 영상의 불필요한 성분들을 RDO 관점에서보다 더 많이 감소시킬 수 있었고, 압축 후의 결과 영상이 왜곡 인지 경계 내에 존재하는가에 대한 불확실성도 양자화 크기에 따라 적응적으로 바뀌는 α 값으로 보완할 수 있었다. 그리고 α 값을 결정할 때 원본이 기준이 되어서 결정하는 것이 아니라 압축 영상이 기준이 되어 압축 영상의 인지 왜곡 영상과의 MSE 가 최소가 되도록 결정하기 때문에 QP 의 크기에 압축률 향상 폭이 적게 영향을 받는다. 이와 같이 기존의 방법들이 가지고 있는 문제들을 해결하여 모든 QP 에 대해서 HM 대비 8% 이상의 압축률 향상을 보이고 평균적으로 15%의 압축률 향상을 보인다. 추후 진행할 연구 방향은 최적의 α 을 찾기 위해서 반복적인 압축과정이 필요하기 때문에 발생하는 계산량을 줄이기 위해, 영상의 특징과 최적의 α 간의 연관성을 가지고 반복적인 과정 없이 α 을 추정할 수 있는 방법을 찾는 것이다.

Acknowledgement

이 논문은 2016 년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (B0117-16-1006, 초고실감 미디어 서비스 실현을 위해 HEVC/3DA 대비 2 배 압축을 제공하는 5 세대 비디오/오디오 표준 핵심 기술 개발 및 표준화)

참고문헌

- [1] Rehman, Abdul, and Zhou Wang. "SSIM-inspired perceptual video coding for HEVC," *IEEE International Conference on. Multimedia and Expo (ICME)*. 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., in press*, Jan. 2014.
- [3] S.-H Bae, Jaeil 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.
- [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, Aug. 2014.