

시각적 동영상 부호화 효율을 위한 부 대역 부호화 기반 전처리 알고리즘

최광연, *송병철
 인하대학교
 *bcsong@inha.ac.kr

Sub-Band Coding based Pre-Processing Algorithm for Perceptual Video Coding Efficiency

Kwang Yeon Choi, *Byung Cheol Song
 Inha University

요 약

본 논문에서는 시각적 동영상 부호화 기법으로서 부 대역 부호화 기반의 시각적 민감도가 낮은 고주파 영역을 억제하는 기법을 동영상 코덱 전처리로 사용하는 것을 제안한다. 먼저 억제하기 위한 고주파 대역 설정을 위해 입력 시퀀스의 프레임들을 특정 주파수 대역으로 나눈다. 나누어진 주파수 대역 중 고주파 대역에 한해 적절한 이득 값 조절을 통하여 고주파 영역을 억제한다. 이와 같은 전처리 이후 동영상 압축 표준에 적용하여 인간의 시각적 구조의 인지 특성을 고려한 동영상 압축 효율을 얻는다. 모의 실험을 통해 제안 기법을 H.264 의 전처리로 적용한 결과와 전처리 없이 적용한 결과를 비교하여 시각적 차이 없이 평균 9.56%의 데이터 감소 효과를 얻는다.

1. 서론

최근 들어 다양한 디스플레이 기기들의 발전에 따라 특정한 시청 환경에서 영상을 시청하는 경우가 증가하고 있다. 이러한 경향으로 인해 다양한 시각적 자극에 따른 인간의 시각적 구조 특성을 고려하여 시각적 동영상 부호화 기법들이 개발되고 있다[1]. 시각적 민감도에 영향을 주는 요소들은 주로 공간적, 시간적 주파수와 밝기, 움직임 등의 정보들과 같은 것들이 있다. 시각적 동영상 부호화 기법들은 이러한 요소들의 정보를 제거하여 시각적 동영상 압축 효율을 얻는다.

기존 기법들은 시각적으로 두드러진 영역에 대한 뱀을 도출하거나 JND(Just Noticeable Difference) 모델링을 통해 시각적 동영상 부호화 기법을 적용하는 방법을 제안하였다[2, 3]. 그러나 기존 기법들은 시각적 민감도가 낮은 영역에 대하여 화질의 열화가 인지되고, 동영상 부호화 구조 내부에서 적용되어 코덱 표준을 변경되기 때문에 복잡도가 증가한다.

본 논문에서는 특정 영역의 화질 열화와 코덱 표준의 변경 없이 전처리로서 부 대역 부호화 기반의 전처리 기법을 제안한다. 부 대역 부호화로 인하여 나누어진 여러 주파수 대역들 중에서 시각적 민감도가 낮은 고주파 대역에 한해 이득 값을 조절하여 해당 고주파 대역을 억제한다. 이후 주파수 대역들을 다시 합성하여 영상을 생성한 뒤 일반적인 동영상 코덱의 입력으로 사용한다.

본 논문의 구성은 다음과 같다. 2 절에서는 부 대역 부호화 기반의 주파수 대역 분해에 대해 살펴본 후, 3 절에서는 나누어진 주파수 대역에 적용되는 이득 값 조절에 대해 설명한다. 4 절에서는 제안한 기법의 성능을 실험을 통해서

확인한다. 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

2. 부 대역 분해

입력 시퀀스의 각 프레임에 대하여 식 (1)과 같이 저역 통과 필터를 적용한 결과를 원본에서 빼주어 서로 다른 주파수 대역을 얻는다.

$$\begin{cases} R'_n(x, y) = I(x, y) - F_n(x, y) * I(x, y), & 1 \leq n \leq N-1, \\ R'_n(x, y) = I(x, y), & n = N, \end{cases} \quad (1)$$

$R'_n(x, y)$ 는 n 번째 주파수 대역, N 은 대역의 수, $I(x, y)$ 는 입력 프레임, $*$ 은 컨볼루션 연산, $F_n(x, y)$ 은 가우시안 필터이다. 주파수 영역이 겹치지 않도록 식(2)와 같이 분해를 적용한다.

$$\begin{cases} \bar{R}_n = R'_1, & n = 1, \\ \bar{R}_n = R'_n - R'_{n-1}, & 2 \leq n \leq N, \end{cases} \quad (2)$$

\bar{R}_n 은 부 대역 분해로 얻어지는 주파수 대역을 나타낸다. F_n 에서 가우시안 분포의 표준 편차를 나타내는 σ_n 은 식 (3)과 같은 조건을 만족한다.

$$\sigma_{n+1} > \sigma_n, \quad n=1,2,\dots,N-1 \quad (3)$$

3. 이득 값 조절

이득 값은 1 보다 작은 값으로 식 (4)와 같이 설정하여 해당 고주파 대역을 억제한다.

$$G_n(x, y) = \exp\left(-1 \times \left(\frac{NR_n(x, y)}{\lambda}\right)^{\bar{\sigma}_n}\right) \quad (4)$$

λ 는 이득 값의 최소값 설정을 위한 상수이고 NR_n 과 $\bar{\sigma}_n$ 은 아래 식 (5)와 같이 나타낸다.

$$NR_n(x, y) = \frac{|\bar{R}_n|}{\max(|\bar{R}_n|)}, \quad \bar{\sigma}_n = \frac{\sigma_n}{\max(\sigma_n)} \quad (5)$$

σ_n 과 λ 값 조절을 통해 이득 값을 조절한다. 이득 값은 그림 1 의 (a)에서와 같이 고주파 대역에 해당하는 작은 n 에서 더 작은 값이 되어 고주파 대역이 더욱 강하게 억제된다. 또한 그림 1 의 (b)에서 λ 값에 대한 이득 값 조절을 통해 전체적인 억제 강도를 조절한다.

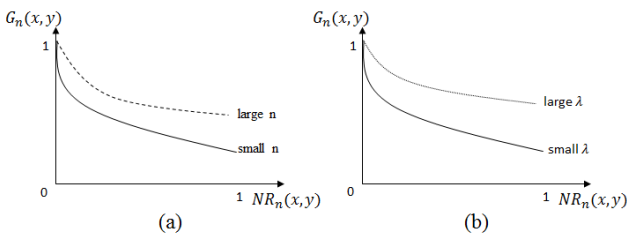


그림 1. (a) σ_n 에 대한 이득 값 (b) λ 에 대한 이득 값

최종적으로 각 주파수 대역에 대하여 식 (6)과 같이 곱하여 결과를 얻는다.

$$I' = \sum_{n=1}^N \bar{R}_n \cdot G_n \quad (6)$$

4. 모의실험 결과

실험을 위해 1920×1080 해상도, 4:2:0 색 포맷, 30 프레임의 테스트 영상인 Cactus, Toys and calendar, Traffic, Flag shoot 시퀀스와 15 프레임의 촬영한 4 개의 시퀀스를 사용하였다. 촬영 영상은 삼성 카메라 NX1 의 연사기능(15Hz)을 활용하여 획득하였다. 동영상 코덱 표준은 H.264 JM 9.0[4]을 사용하였다. 인코더 설정은 IPPP 포맷으로 양자화 파라미터를 24, 28, 32, 36 으로 하고 프레임 비율은 30, 움직임 추정은 EPZS(Enhanced Predictive Zonal Search), 참조 프레임은 2 개, 비율 왜곡 최적화와 비율조절은 적용하지 않았다.

제안 기법의 주관적 관점의 동영상 압축 효율을 확인하기 위해 주관적 시각 테스트 중 하나인 ACJ (Adjectival Categorical Judgment) 기반의 테스트를 진행하였다[5]. 전처리 미 적용 영상과 적용 후 영상에 대해 동일 양자화 파라미터 값에서 동영상 코덱을 적용한 영상을 비교하여 두 영상의 차이가 인지 되는지를 확인하였다. 시각적 차이 유무는 그림 2 에서 확인할 수 있다. 시각적 시청환경은 표 1 과 같이 설정하였다.

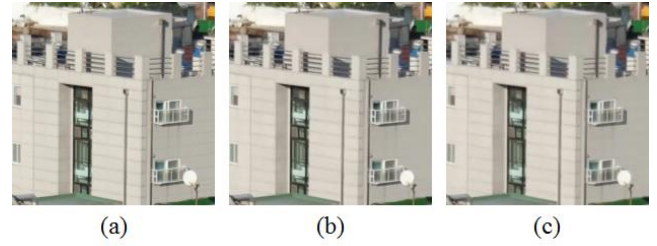


그림 2. (a) 원본 영상 (b) 전처리 미 적용 후 코덱 적용 (c) 제안 기법 적용 후 코덱 적용

표 1. 시청환경

모니터	삼성 UN46F8000AF
형태, 크기	LED, 46 인치
해상도	1920×1080
실험 인원, 시청 거리	10, 2H

정량적인 화질 비교를 위해 MS-SSIM(multiple scale-structural similarity)를 사용하였다[6]. 전처리로 제안 기법 적용 이후에 코덱 적용 결과에 대하여 모든 실험 영상과 양자화 파라미터에 대한 평균값으로 확인하였다. MS-SSIM 는 전처리 없이 0.9795, 전처리 적용 후 0.9783 으로 화질은 비슷하지만 9.56%의 데이터 감소효과를 얻었다.

5. 결론

본 논문은 시각적 민감도가 낮은 고주파 영역을 억제하기 위한 부 대역 부호화 기반의 코덱 전처리로 적용되는 시각적 동영상 부호화 기법을 제안한다. 대역 분해 과정에서 나누어 지는 주파수 대역 중 고주파에 한해 이득 값 조절을 통해 고주파 영역을 억제한다. 전처리를 적용하지 않고 동영상 코덱을 적용한 결과와 비교하여 시각적 차이 없이 평균 5.56%의 데이터 감소 효과를 얻는다. 기존 기법들과 비교하여 시각적 민감도가 낮은 배경 영역의 화질의 열화와 코덱 표준 변경 없이 적용 가능하다.

참고문헌

[1] H. R. Wu and K. R. Rao, *Digital Video Image Quality and Perceptual Coding*. Boca Raton, FL: CRC Press, Nov. 2005.

[2] Z. Li, S. Qin, and L. Itti, "Visual attention guided bit allocation in video compression," *Image Vis. Comput.*, vol. 29, no. 1, pp. 1- 14, Jan. 2011.

[3] Jaeil Kim, Sung-Ho Bae, and Munchurl 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, Nov. 2015.

[4] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG, Joint Model Reference Software, version 9.0.

[5] ITU, "Methodology for the subjective assessment of the quality of television pictures," Geneva, Switzerland, ITU-R BT.500-11, 2002.

[6] Z. Wang, E. P. Simoncelli and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *Proc. IEEE Asilomar Conf. Signals, Systems, Comput.*, vol. 2, pp1398-1402, Nov. 2003.