

인간의 시각특성을 고려한 적응 BTC

윤정필, 김슬기, 손광훈

연세대학교 전파공학과
서울특별시 서대문구 신촌동 134
khsohn@bubble.yonsei.ac.kr

요약

본 논문에서는 중요한 정보의 형태 중 하나인 영상 정보를 빠르고 정확하게 송수신하기 위하여 사용되는 영상 압축 방법 중 간단하면서도 좋은 성능을 가진 기존의 Block Truncation Coding(BTC)을 기반으로 하여 더욱 발전된 성능의 BTC를 구현하기 위한 방법을 제안한다. 인간의 시각적인 특성을 이용하여 각 블록을 명암에 따라 판단하여 차별된 양자화기를 사용함으로써 높은 압축에도 시각적인 측면뿐만 아니라 수치적인 측면에서 더욱 개선된 영상을 얻을 수 있었다. 또한, 모서리 부분의 특성이 잘 보존되도록 한 블록 안의 평균과 로그평균의 차를 이용하여 블록 안의 임계값을 결정하는 새로운 형태의 적응성 BTC 구현을 제안하여 BTC의 상반된 두 가지 특성인 압축률과 오차율을 개선시켰다.

1. 서론

영상 압축 기술에서 높은 해상도와 함께 고려해야 할 것은 고압축률, edge영역의 보호와 잡음에 민감하지 않아야 한다는 것이다. Delp 와 Mitchell[1] 에 의해 소개된 BTC는 쉽게 구현하고 많이 응용될 수 있는 좋은 특성을 가진 영상 압축기술이며, 이미 그래픽과 디지털 비디오, 칼라 디지털 이미지에 응용되어 왔다. BTC의 동작으로는 이미지를 중복되지 않는 $N \times N$ 의 작은 블록으로 나누어 이것을 1bit 양자화기에 통과시킴으로써 개별적으로 코드화시킨다. 양자화 레벨은 여러 양자화 방법에 따라 결정된다. 그러나, 이러한 양자화 방법들은 기본적으로 유사하고, 영상의 부분적인 통계에 적당하지 않은 균일한 양자화기에 의존하고 있다. 그런데 여기에서 이미지의 texture 영역에서의 오류는 균일한 지역에서의 오류보다 눈에 더 잘 보이지 않는다는 것이다. 결과적으로 BTC는 명백한 artifacts를 만들어 내지 않고 texture 영역을 간결하게 코딩한다. 그러나 BTC는 다른 영역에 있는 영상에 대해서 똑같이 잘 수행되지는 않는다. edge를 포함하고 있는 영역에서 잡음을 유발하는 artifacts를 만들어낸다. 또한 BTC에 의해 사용되는 데이터의 형태

는 픽셀의 분산정도가 상대적으로 동일한 영역들에서는 비효율적이다[2]. 본 논문에서는 좋은 압축률을 지닌 BTC의 장점을 보존하면서 인간의 시각적인 특성을 고려하여 임계값의 기준을 결정하였으며 어두운 부분과 밝은 부분에 각각의 차별화된 양자화기를 사용하였다.

2. 이론적 배경

BTC는 전체영상을 $N \times N$ 의 작은 블록으로 중복되지 않게 분할하여 독립적으로 두개의 레벨을 나타내는 1비트 양자화기를 통과시킨다. 이것을 통해 비트맵과 reconstruction level를 발생시켜 수신측에 전달하는 것으로서, 양자화기의 임계값과 두개의 reconstruction level은 각 블록의 통계적인 값에 의해 결정된다. 디코딩은 각 비트맵의 값에 의해서 두개의 reconstruction 값을 적당하게 배열시키는 간단한 과정이다[3]. BTC의 양자화기는 특성에 따라서 Moment preserving 양자화기[4], Absolute moment 양자화기[5], Error minimizing 양자화기[6]로 나눌 수 있으며 Error minimizing 양자화기는 다른 형태의 absolute moment 양자화기임이 증명되

있다.

1) Moment preserving 양자화기

Moment-preserving 양자화기는 블록 내 표본들의 모멘트가 양자화기를 거친 후에도 보존 되도록 설계되었다. 영상은 $N \times N$ 행렬의 블록으로 나누어지고, X_1, X_2, \dots, X_m 은 원 영상의 주어진 블록에서의 픽셀 값이다. ($m = n^2$)

$$\hat{X}_i = \begin{cases} a, & \text{if } X_i < X_{th} \\ b, & \text{if } X_i > X_{th} \end{cases} \quad \text{for } i=1,2,3,\dots,m \quad (2)$$

$$X_{th} = (\bar{X}^2)^{1/2} \quad (3)$$

$$m\bar{X} = (m-q)a + qb \quad (4)$$

$$m\bar{X}^2 = (m-q)a^2 + qb^2$$

q 를 X_{th} 보다 큰 X_i 의 갯수라고 정의한다면 다음 두 식(4)에 의해 a 와 b 를 구할 수 있다.

$$a = \bar{X} - \sigma \sqrt{\frac{q}{m-q}}, \quad b = \bar{X} + \sigma \sqrt{\frac{m-q}{q}} \quad (5)$$

moment preserving 양자화기의 단점은 제곱(squaring)과 제곱근(square root)이 사용되어 실제로 하드웨어로 구현함에 있어서 비교적 계산이 복잡하다는 단점이 있다.

2) Absolute-moment quantizer

Moment-Preserving 양자화기의 단점을 피하기 위해 Lema와 Mitchell[5]은 AMBTC라는 새로운 양자화 방법을 개발해 내었다. 이것은 $N \times N$ 픽셀의 표본 평균뿐만 아니라 1차 절대 중심 모멘트도 보존이 되도록 설계되었다. $N \times N$ 픽셀의 평균과 1차 절대 모멘트(a)는 다음과 같다.

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad a = \frac{1}{m} \sum_{i=0}^m |X_i - \bar{X}| \quad (6)$$

양자화기의 임계값을 \bar{X} 라고 하면, 재생 레벨은 다음과 같다.

$$a = \bar{X} - \frac{ma}{2(m-q)}, \quad b = \bar{X} + \frac{ma}{2q} \quad (7)$$

$$\hat{X}_i = \begin{cases} a, & \text{if } X_i < X_{th} \\ b, & \text{if } X_i > X_{th} \end{cases}, \quad X_{th} = m \quad (8)$$

여기서, q 는 $N \times N$ 블록 내의 평균보다 큰 픽셀의 수이다. 이 방식은 다른 두 가지 방식에 비해 실제로 구현하였을 때 상당히 높은 성능을 보이는 양자화 방법이다.

3 기존의 적용 BTC[7]

전체 영상을 4×4 블록의 크기로 분할하였을 때 이 4×4 블록은 3가지 영역 중 한 영역에 속하게 된다. 첫 번째 영역은 픽셀값들이 거의 일정한 smooth 영역이고, 두 번째 영역은 texture를 포함하고 있는 texture 영역이며, 세 번째 영역은 edge를 포함하고 있는 edge 영역이다. 영상을 간결하게 코딩하기 위해서는 각 경우에 대해서 다른 방법으로 코딩해야 한다. smooth 영역은 그 영역의 평균값을 전송하는 1-레벨 양자화기를 사용하고, texture 영역은 2-레벨인 absolute moment 양자화기를 사용하였으며 edge영역은 4개의 작은 블록(2×2)으로 나뉘어 임계값 보다 작은 영역은 다시 smooth 영역으로, 임계값 큰 부분은 edge 부분을 보호하기 위해 픽셀값이 그대로 코딩된다. 영역을 결정하기 위해서는 각 영역에 대한 임계값을 결정해야 하며, 이 값은 한 블록 내에서 픽셀값의 최대값과 최소값의 차이를 이용한다. 즉, 임계값이 13-40인 것을 smoothness threshold라 하고 이 값보다 작은 영역은 smooth 영역에 속하게 된다. 임계값이 120인 것을 edge threshold라 하고, 이 값보다 큰 블록은 edge 영역에 속하게 되어 2×2 블록으로 다시 분할된다. 분할된 블록 중 edge threshold 값의 반인 60을 임계값으로 선택하여 2×2 블록의 smooth 영역과 edge 영역으로 나뉘게 되며, edge 영역에 속하는 2×2 블록은 원래의 픽셀값 그대로 코딩된다. 세가지 영역 중 smooth영역과 edge영역 중간에 있는 texture 영역은 AMBTC의 방식으로 코딩된다.

4. 제안된 적용 BTC

제안된 적용 BTC는 두 개의 차별화된 양자화기를 가지며 각 양자화기는 edge 영역과 texture 영역, 그리고 smooth 영역으로 나누어준다. 전체 영상을 4×4 크기의 블록으로 나누어 각 블록의 평균을 구하고, 그 값은 어두운 부분과 밝은 부분을 구별해 주는 임계값과 비교를 한다. 임계값은 여러 번 실험을 통해 시각적인 면과 수치적인 면에서 최적의 값(143)을 선택하였다. 블록의 평균이 143보다 큰 값을 가질 경우, 이 블록은 밝은 영역으로 구분되기 때문에 압축을 하는 과정에서 압축률보다는 화질에 중점을 둔 압축방법을 택하게 된다. 따라서 영역을 결정하는 양자화 레벨이 비교적 낮은 값을 가진다. 블록의 평균이 143보다 작아 어두운 영역으로 구분되어지는 블록의 경우 사람의 눈이 민감하게 반응하지 않는 부분을 많이 포함하고 있으므로 화질보다는 압축률을 높이는 방향의 양자화가 이루어진다. 임계값 143에 의해 나누어진 블록들은 두 개의 차별된 양자화기에 의해 각각 세 영역으로 분리되며 그 영역을 결정하는 기준

은 식(9)와 같다. 한 블록 내에서 평균과 밑이 2인 로그 평균을 이용하여 임계값을 구하면 픽셀값의 변화가 없는 smooth 영역에서는 두 평균의 차이가 작고 edge가 포함된 영역에서는 차이가 크기 때문에 edge 부분을 보호할 수 있다.

$$\begin{aligned}
 mn_1 &= \frac{1}{m} \sum_{i=1}^m X_i \\
 mn_2' &= \frac{1}{m} \sum_{i=1}^m \log_2 X_i \\
 mn_2 &= 2^{mn_2'} \\
 \text{Threshold} &= |mn_1 - mn_2|
 \end{aligned} \tag{9}$$

m : 한 블록 내에서의 픽셀의 수
 X_i : 한 블록 내에서의 픽셀값

어두운 부분에서의 변화보다는 밝은 부분의 변화에 민감한 인간의 시각 특성을 이용하여 설계된 차별화된 양자화기를 통과한 영역 중 smooth 영역에서는 블록의 평균값이 코딩되고, texture 영역에서는 두 개의 재생레벨과 비트맵이 AMBTC 형태로 코딩된다. edge 영역에서는 다시 4개의 작은 블록(2×2)으로 분할되며 이 블록들은 임계값에 의해 다시 smooth 영역과 edge 영역으로 나뉘어져서 smooth 영역은 블록의 평균값이, edge 영역은 픽셀값 그대로 코딩되어 edge 부분을 보호해 준다. 그림1은 제안된 BTC의 전체적인 알고리즘을 블록다이어그램으로 나타내었다

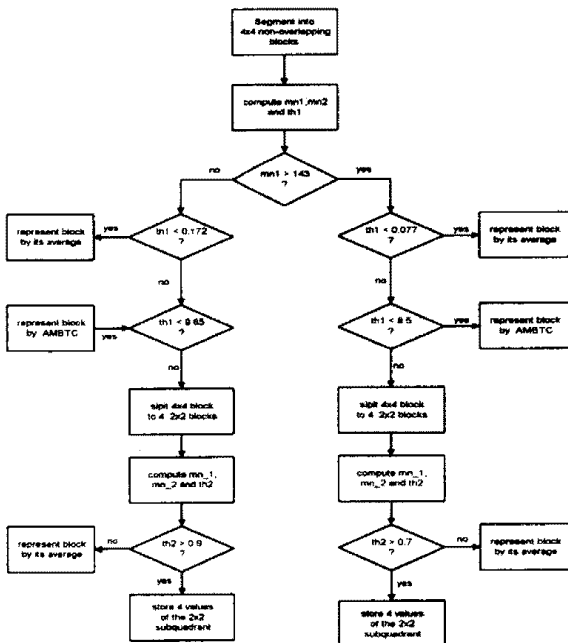


그림 1 제안된 BTC의 전체 알고리즘

5. 시뮬레이션

기존 BTC와 제안된 BTC의 실험영상은 그림2와 같다

며, 여기서 기존의 적용 BTC는 smooth 영역과 texture 영역 사이의 임계값을 18, texture 영역과 edge 영역 사이의 임계값을 120으로 하였다. 이 논문에서 제시한 BTC의 각 임계값을 변화시켜서 반복적으로 실험을 하여 최적의 값을 선택하고 RMSE와 압축률뿐만 아니라 시각적인 면에서 기존의 BTC와 비교를 하였다.



(a) 표준 BTC

(b) AMBTC



(c) Adaptive BTC

(d) 제안된 적용 BTC

그림2. 기존의 BTC와 제안된 BTC의 영상

6. 성능 비교

표2는 잡음이 발생되지 않았을 때 BTC, AMBTC, 기존의 적용 BTC, 제안된 BTC의 성능을 RMSE(Root Mean Square Error)와 BPP(bit per pixel)를 이용하여 비교하였으며 제안된 BTC가 성능이 우수함을 알 수 있었다.

$$RMSE = \left[\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2 \right]^{\frac{1}{2}}$$

$\hat{f}(x, y)$: Reconstruction 영상 ($M \times N$)
 $f(x, y)$: 원 영상 ($M \times N$)

표2. 여러 종류의 BTC 성능 비교

	BTC	AMBTC	기존의 ADBTC	제안된 BTC
RMSE	4.99	4.89	5.00	5.00
BPP	2	2	1.526	1.507

표3. Smoothness Threshold에 의한 기존의 BTC 성능 비교

S.T	기존의 BTC		
	RMSE	BPP	PSNR
13	4.85	1.691	34.4169
15	4.91	1.608	34.3032
18	5.00	1.526	34.1406
30	5.39	1.307	33.4940
40	5.60	1.220	33.1710

표4. 제안된 BTC의 어두운 부분과 밝은 부분을 결정하는 임계값의 변화에 따른 성능비교

th _i	RMSE	BPP	PSNR
123	4.97347	1.53438	34.1976
133	4.99155	1.52130	34.1661
143	5.00662	1.50775	34.1399
153	5.02640	1.49493	34.1057
163	5.05232	1.48083	34.0610
193	5.11444	1.44540	33.9548

표3.은 smooth 영역의 임계값의 변화에 따른 기존의 적용 BTC의 성능을 나타내고 있으며, 표4.는 제안된 BTC의 밝은 부분과 어두운 부분의 임계값의 변화에 따른 성능을 나타내고 있다. 그 외 임계값도 같은 방식으로 실험하여 최적의 값을 얻었다.



(a) BTC

(b) AMBTC



(c) ADBTC

(d) 제안된 BTC

그림 3. 각 BTC 에 따른 edge 영역의 보존 비교

그림3.는 영상을 확대하여 각 BTC에 대해서 edge영역의 보존 여부를 서로 비교하여 보여주고 있으며 이것을 통해 적용 BTC가 표준 BTC나 AMBTC 보다 edge영역이 잘 보존됨을 알 수 있으며, 제안된 BTC는 기존의 적용

BTC보다 blocking artifact가 적다는 것을 알 수 있다.

7. 결론

본 논문에서는 인간의 눈의 감도를 고려하여 영상을 차별적으로 양자화 함으로써 화질에 비중을 두면서도 압축률을 손상시키지 않는 양자화기를 구현하여 기존에 있는 적용 BTC 등과 성능을 서로 비교, 분석하였다. 그 결과 edge영역의 정보를 그대로 보호하면서 기존의 BTC와 비교하여 좋은 압축률과 낮은 RMSE를 가지고 있으면서, 인간의 시각적인 측면에서도 그림3. 과 같이 texture 영역에서 blocking artifact가 많이 생기지 않아 시각적으로 좋은 성능을 보이고 있다. 또한, 기존의 적응성 BTC가 작은 변화의 임계값에도 성능에 있어서 커다란 차이를 보인 것과는 달리 제안된 BTC는 임계값을 크게 변화시켜도 성능에 있어서의 급격한 변화가 나타나지 않았다. 이는 설정한 임계값들이 어느 정도 최적화된 값에 근접해 있다는 것을 보여주고 있다.

8. 참고 문헌

- [1] E.J.Delp and R.Mitchell, "Image compression using block truncation coding", *IEEE Trans. on comm*, vol.27, pp.1335-1342, 1979.
- [2] N. C. Griswold, D. R. Halverson and G. L. Wise, "A Note on Adaptive Block Truncation Coding for Image Processing", *IEEE Trans. on Accoustics, Speech, and Signal Processing*, vol. Assp-35, no. 8, 1987.
- [3] Majid Rabbani and Paul W. Jones, "Digital Image Compression Techniques", Georgia Institute of Technology, 1991.
- [4] D. R. Halverson and N. C. Griswold and G. L. Wise, "A Generalized Block Truncation Coding Algorithm for Image Compression", *IEEE Trans. on Accoustics, Speech, and Signal Processing*, vol. Assp-32, no. 8, 1984.
- [5] M.D. Lema and O.R. Mitchell, "Absolute moment block truncation coding and its application to color image", *IEEE Trans. Comm*, vol. 32, pp.1148-1157, 1984
- [6] Vishwas R. Udpikar and Jewan P. Raina, "BTC Image coding Using vector Quantization", *IEEE Trans. on Comm.* vol. com-35. no.3, pp. 352-356, 1987.
- [7] Panos Nasiopoulos, Rabab K. Ward, and Daryl J.Morse, "Adaptive Compression Coding", *IEEE trans. on comm*, vol. 39, No.8, pp. 1245-1254, 1991.