

DCT 기반 영상의 시각 적응적 양자화 방법에 관한 연구

김정현⁰ 박성찬 천승환 이귀상
전남대학교 전산학과
(jkhkim, schpark, shchoen, gslee)@cs.chonnam.ac.kr

Visually-Adaptive Quantization method Based on Block DCT

Jung-Hyun Kim⁰ Sung-Chan Park Seung-Wan Cheon Guee-Sang Lee
Dept. of Computer Science, Chonnam National University

요 약

본 논문에서는 인간의 시각 특성을 고려한 시각 적응적 DCT 영상 압축 방법이 제안되었다. 인간의 시각체계의 특성 중 복잡한 영상이나 대비효과(contrast)가 큰 부분에서는 압축과정에서 발생한 왜곡이 쉽게 눈에 인지되지 않는 특성을 갖는 공간적 마스크링(spatial masking)을 이용하여 영상의 질을 어느 정도 유지하면서 적응적 압축 방법에 의해 압축율을 보다 높이는 방법을 제시하였다. DCT 변환 블록에서 경계영역 정보를 많이 포함하는 부분을 추출한 후 이 영역은 다른 영역에 비해 복잡도가 더 높고 경계성분의 대비효과가 더 크기 때문에 이 영역의 분산을 취해 이 값에 따라 적응적으로 양자화한다. 실험 결과, 제안된 영상 압축 알고리즘은 기존의 시각 적응적 압축 방법보다 좋은 성능을 보이며 특히 고주파 성분을 많이 가진 복잡한 영상에 대하여 높은 압축률을 보였다.

1. 서론

저 비트 영상통신의 요구는 날로 증가함에 따라 효율적인 영상압축을 위한 표준화와 기술개발이 활발히 진행되고 있다[1]. 그 중에서 DCT는 정지영상과 동영상 압축 표준인 JPEG과 MPEG에서 공간적 중복성을 제거하기 위한 알고리즘으로 가장 많이 사용하고 있는 방법이다. DCT가 많이 사용되는 이유는 강력한 에너지 밀집 능력과 고속 알고리즘의 존재라고 할 수 있다. 그러나 이 방식은 인간의 시각적인 면은 전혀 고려하지 않았으며 고변화 영역에서 두드러진 블록화 현상을 초래하게 된다. 예를 들어 입력 영상이 고주파 성분을 많이 포함하고 있는 경우 일반적인 DCT 계수 양자화로 인하여 고주파 성분들을 많이 잃어버리게 된다 [2].

그 동안 영상 압축알고리즘의 성능 및 압축된 영상의 질을 평가하기 위한 척도로써 압축율, 비트율, SNR, PSNR, MSE 등이 사용되어 왔다. 그러나 이러한 척도들에 의하여 계산된 영상의 왜곡량은 사람이 눈으로 보았을 때 직접 느끼는 왜곡량에 비례하여 제대로 나타내지 못한다는 지적들이 있어왔다[3][4]. 이것은 영상의 왜곡을 측정하기 위한 방법들이 인간의 시각체계(human visual system)의 특성에 바탕을 두고 설계된 것이 아니라, 단순히 통계 수학적 방법에만 의존하고 있기 때문이다.

이와 같이 기존의 영상 압축 방법들은 인간의 시각 체계의 특성을 제대로 활용하지 못하고 영상 전체를 일률적인 방법으로 압축하고 있다. 그러므로 인간의 시각 체계의 특성을 토대로 압축 및 복원 알고리즘을 설계한다면 보다 높

은 압축효율을 얻음과 동시에 압축영상의 질을 개선 시키거나 응용목적만족시키는 수준을 유지할 수 있다[5]. 그러나 아직까지 실질적인 시각 적응적 압축방법의 개발에 대한 연구가 미흡한 상태이다.

본 논문에서는 인간의 시각 체계의 특성 중에서 영상이 매우 복잡한 내용을 담고 있거나 대비효과가 큰 부분에서는 압축과정에서 발생한 왜곡이 쉽게 눈에 인지되지 않는 특성인 공간적 마스크링(spatial masking)[3][4][6]을 이용하여 DCT 경계 영역 정보에 따른 시각 적응적 양자화 방법을 제안한다. 2장에서는 관련 연구를 간단히 살펴보고 3장에서는 제안 알고리즘에 관한 설명과 4장에서는 제안한 방법의 실험 결과와 결론을 맺는다.

2. 관련연구

2.1 시각 적응적 영상 압축

시각 적응적 영상 압축은 고손실 또는 저손실 블록을 적응적으로 결정하는 방법으로서, 영상의 각 블록에 대하여 왜곡들의 발생여부를 미리 계산하고 그 결과에 따라 블록에 가해질 손실의 양을 조절하여 압축한다.

2.1.1 휘도마스크링 방법

휘도마스크링(luminance masking)이란 너무 어둡거나 너무 밝은 곳에서는 밝기의 변화를 쉽게 인지할 수 없는 시각체계의 특성을 말한다. DCT 블록에서 DC 성분(D(0))은 블록의 평균적인 밝기를 나타내고 값의 범위는 약 $-1000 \sim 1000$ 정도이다. 이때 평균적인 밝기가 너무 어둡거나 밝은 경우에는 영상의 내용을 판별하기 어렵다고 볼 수 있으며 이러한 블록에서는 왜곡이 많이 발생하더라

도 사람의 눈에 쉽게 인지되지 않는다. 그러므로 임계치의 상한과 하한, L_{min} 과 L_{max} 를 설정하여 이 범위를 벗어나는 블록은 손실을 많이 허용하여 압축하게 된다.

블록의 왜곡을 인지 가능한 평균 휘도값의 범위:

$$L_{min} \leq D(0) \leq L_{max} \quad (1)$$

이 범위에 포함되지 않아서 손실을 많이 허용하여 압축하는 경우에도 DC 성분은 저손실 양자화 테이블을 적용하여 DC 성분에서의 양자화 손실로 인한 왜곡을 최대한 줄였다.

2.1.2 블록효과 발생블록 추출방법

블록효과는 영상에서 블록과 블록 사이의 경계선이 뚜렷하게 드러나는 현상을 말한다. 주어진 블록이 높은 압축률로 압축했을 때 블록효과가 발생할지 판별하기 위해서는 블록이 얼마나 평평한가를 계산함으로써 블록효과를 예측할 수 있다. 이를 계산하기 위해서 S_{low} 와 S_{mid} 를 다음과 같이 정의하였다.

$$S_{mid} = \sum_{i=1}^{30} |D(i)| \quad (1)$$

$$S_{low} = |D(1)| + |D(2)| \quad (2)$$

S_{low} 는 영상의 상대적인 저주파수 성분의 에너지 양을 나타내는데, S_{low} 가 너무 작을 경우에는 블록 효과가 발생되지 않을 정도로 밝기의 변화가 없다는 것을 뜻하고 S_{low} 가 너무 클 경우에는 뚜렷한 경계성분이 존재하게 되어 블록효과가 발생되지 않는다. S_{mid} 는 블록의 복잡도를 나타낸다. S_{mid} 가 너무 크면 블록 전체적으로 복잡한 정보를 많이 포함하게 되어 역시 블록효과가 발생하지 않는다. 그리고 이 두 값이 다음의 범위에 있는 블록을 블록효과가 발생하는 블록이라고 결정한다.

2.2 관심도를 고려한 영상 압축

관심도를 고려한 영상 압축이란 영상 전체를 동일 한 방법으로 압축하는 것이 아니라 영상 중에서 사용자가 관심을 가지고 있는 부분과 나머지 부분을 달리 압축하는 일종의 주관적 압축 방법이다. 이러한 압축 방법은 영상에서 특정영역을 강조하거나 또는 관심영역을 제외한 나머지 부분영역에 대해서는 손실이 비교적 많이 발생하더라도 영상의 활용에 크게 제한을 받지 않는 경우에 매우 유용하게 이용될 수 있다.

압축과정에서 압축영상의 질을 조절할 수 있는 부분이 양자화 과정에 이용되는 양자화 테이블이므로 관심부분과 나머지 부분에 다른 손실 특성을 가지는 양자화 테이블을 적용함으로써 서로 다른 질의 압축 영상을 얻을 수 있다.

3. 제안 방법

DCT 처리에 의해 얻어진 변환 데이터 블록에서 잡음 성분이 적으면서 윤곽선 성분이 많이 분포한 중간 주파수 대역은 그림 1에서 빗금으로 표시된 부분인 B영역이다[7]. 이 영역은 윤곽선 정보가 없는 블록에 비해 더 복잡하고 값의 편차가 심하다.

영상의 질을 측정하는데 큰 영향을 주는 공간적 마스크링 (spatial masking)은 영상이 매우 복잡한 내용을 담고 있거나 대비효과가 큰 부분에서는 압축과정에서 발생한 왜곡이 쉽게 눈에 인지되지 않는 특성을 가진다. 이러한 인간 시각체계의 특성을 이용하기 위해 각 블록에서 윤곽선 정보가 포함된 B영역에 속한 블록들의 분산을 구한다. 분산

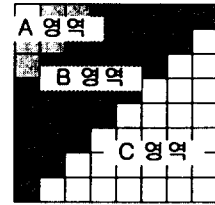


그림 1. 8x8 블록의 영역 구분

값이 크면 영상이 더 복잡하고 경계의 차이가 뚜렷하여 대비효과도 커지게 되어 사람의 눈에 왜곡이 쉽게 인지되지 않기 때문에 더 큰 값으로 양자화를 하고 분산 값이 작으면 영상이 덜 복잡하고 차이가 심한 경계도 많지 않다는 의미이므로 더 작은 값으로 양자화 해 줌으로써 적용적으로 양자화를 하게 된다. 제안된 알고리즘의 큰 흐름은 그림 2와 같다.

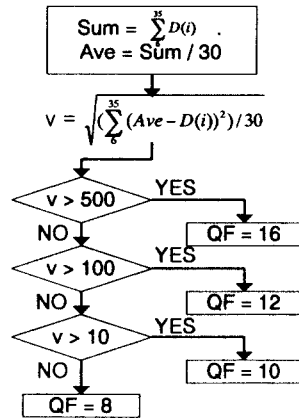


그림 2. 제안 알고리즘

8x8 블록 단위로 DCT 변환된 결과에서 그림 1의 B영역에 포함된 값들의 분산을 구한 후 이 값에 따라 적절한 양자화 값(QF)을 결정한다. 이때 구분 기준이 되는 값들은 영상의 복잡도와 경계의 대비에 따른 분산의 분포를 분석하여 실험 후 결정하였다. 다음 표는 각각 다른 복잡도의 영상과 그에 따른 분산 값이다.

| | | | | | |
|------|------|-----|------|------|------|
| | | | | | |
| 1024 | 1770 | 356 | 30.3 | 10.3 | 12.8 |

표 1. 영상의 복잡도에 따른 분산 값 비교

분산 값에 따라 각 블록에 대한 양자화 값이 결정되면 그 블록 내에서 다시 더 세분화 되어 양자화가 수행 되는데 A영역에 속한 성분들은 블록 전체의 주요 정보와 에너지가 집중되어 있기 때문에 이 영역의 손실을 더 작게 하기 위해 양자화 값을 조정하고 B영역은 처음에 결정된 값을 그대로 사용하고 C영역에 속한 성분들은 A, B영역에 비해 정보의 중요도가 덜 하므로 손실을 좀 더 많이 하여 양자화 한다. 결국 두 번의 양자화 사이드 결정 과정을 거치면서 시각적으로 덜 민감한 부분은 손실을 많게 하고 중요 정보가 포함된 부분은 손실을 적도록 양자화 함으로써 기준

의 다른 시각 적응적 압축 방법들 보다 향상된 화질과 압축률을 얻을 수 있었다.

그리고 본 논문에서는 칼라 성분인 Cb, Cr 성분들은 수평 및 수직 방향에 대해서 각각 2:1로 subsampling 하여, 전체적으로는 1/4 정도의 공간 해상도로 낮추어 압축하였다. 따라서 이미 subsampling 되어 있는 칼라 성분에 대해서 다시 고손실의 양자화 과정을 적용할 경우에는 영상에 대한 왜곡이 너무 심하게 되므로 본 논문에서는 칼라 성분에 대해서는 표준안 그대로 양자화 하여 압축하고 대상을 Y 성분만으로 한정하였다.

4. 실험 결과 및 결론

4.1 실험 결과

실험은 Chen DCT 를 이용하였으며 비교 방법들은 기존의 표준방식과 관련연구에서 설명했던 휘도 마스크 방법 그리고 블록 추출 방법을 함께 실험하여 다음 세 가지 척도로 비교하여 보았다. 실험 영상은 352x240 사이즈의 tennis, bike, flower 영상이고 각각 50 프레임씩 압축하였다.

그림 3과 4는 각각 PSNR 과 평균 MSE를 비교한 그래프이다. 제안 방법이 표준 방식과 약간의 차이가 있지만 휘도 마스크 방법이나 블록 추출방법보다는 좋은 결과임을 알 수 있다. 그림 5는 file size를 비교한 것인데 블록 추출



그림 3. 기존 방법과의 PSNR 비교

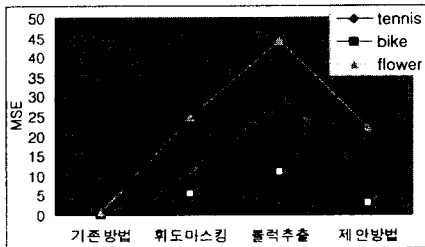


그림 4. 기존 방법과의 평균 MSE 비교

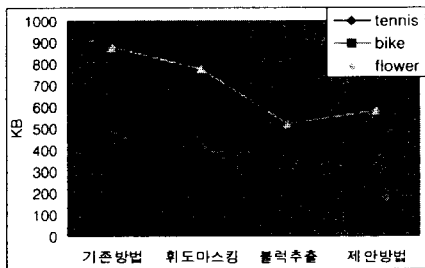


그림 5. 기존 방법과 File Size 비교

방식보다 약간 압축율이 떨어지지만 화질에 비하면 제안방법의 압축율이 가장 좋다고 볼 수 있다.

제안된 방법으로 압축된 이미지는 육안으로 보아서 원래의 표준 방식으로 압축된 것과 거의 차이가 없었고 압축 효율은 더 높았다. 특히 이 방법을 적용하여 압축율에서의 이득을 많이 볼 수 있는 영상은 전체적으로 영상의 내용에 고주파 성분이 많이 포함되어 있는 영상이라는 것을 알 수 있었다. 고주파 성분이 골고루 많이 포함된 flower 영상에서 압축 효율이 상대적으로 많이 향상되었다.

실험 결과에서 알 수 있듯이 PSNR이나 MSE의 수치의 변화에도 불구하고 실제 영상을 육안으로 확인했을 때는 거의 차이가 없었다. 기존의 왜곡량 측정 방법은 사람이 느끼는 정도의 영상의 질을 정확하게 반영하지 못한다는 것을 알 수 있었다.

4.2 결론 및 향후 과제

본 논문에서는 블록 단위의 변환 압축 방법으로 영상을 압축하는데 있어서 영상 전체를 일괄적인 방법으로 압축하지 않고 압축 과정에서 손실이 더 발생하더라도 시각적으로 크게 영향을 받지 않은 부분과 그렇지 않은 부분으로 구분하여 적응적으로 양자화하는 방법에 대해서 제안하였다. 이 시각 적응적 압축 방법을 다양한 영상에 대해서 실험한 결과 영상의 질과 압축율을 동시에 높일 수 있는 방법임을 확인 하였고, 이미지 전체를 일괄적으로 압축하는 방식보다 개선된 압축 방법이라고 할 수 있다.

앞으로의 연구과제로는 이미지의 질을 세분화 시켜 조절하기 위한 양자화 테이블 설계 및 자동 영역 구분에 관한 연구가 많이 이루어져야 할 것이다.

5. 참고 문헌

- [1]. S. C. Joung, H. B. Woo, J. K. Paik, and Y. W. Choi, "Enhancement of inter-frame coded images based on convex projections," Proc. 1999 IEEE Region 10 Conf., vol. II, pp.1275-1278, September 1999.
- [2]. 김일구, 조남익, "주파수 적응 변환을 이용한 영상 압축 방법", 신호처리합동학술대회, pp. 109-111, 10.1999
- [3]. D.R.Fuhrmann, J.A.Baaro and Jerome R.Cox, "Experimental evaluation of psychophysical distortion metrics for JPEG-encoded images", Proc. Of SPIE. Vol. 1993, Pp.179-190. San Jose, CA, USA, 1993
- [4]. S.A.Karunasekera and N.G.Kingsbury, "A Distortion Measure for Image Artifacts Based on Human Visual Sensitivity", ICASSP-94, Vol. V pp.405-408, Mineapolis, April, 1993
- [5]. 허봉식, 김민환, "블록 DCT기반의 시각적응적 이미지 압축에 관한 연구", 정보과학회 논문지(A) 제 22권 제 10호 10. 1995
- [6]. S.A.Karunasekera and N.G.Kingsbury, "A Distortion measure for monochrome image compression", Proc. Of SPIE, Visual Communications and Image Processing '93. Vol. 2094, pp.474-486, Cambridge, MA, USA
- [7]. 정화자, 정기현 외, "DCT를 이용한 윤곽선 추출", 정보과학회 논문지(C) 제 3권 제 1호, pp. 90-97, 1997