

3차원 웨이블릿 기반 압축 동영상의 적응적 정칙화 복원¹⁾

장윤희, 김태영, 정정훈, 백준기
중앙대학교 첨단영상전문대학원 영상공학과

Adaptive Regularized Restoration Of 3-D Wavelet Coded Video

Younhui Jang, Taeyoung Kim, Junghoon Jung and Joonki Paik
The Graduate School of Advanced Imaging Science, Multimedia and Film
Chung-Ang University

요 약

본 논문에서는 3차원 웨이블릿 부호화 방식으로 압축된 영상 시퀀스를 정칙화 기반 영상복원 방법으로 후처리하는 알고리즘을 제안한다. 우선, 웨이블릿 압축 시스템을 적절한 영상 열화 시스템으로 모델화한다. 그리고, 시간축에 관하여 프레임 간의 같은 위치에 있는 각 픽셀에 대하여 복원을 수행한다. 그 다음으로 2차원 영상 신호에 대하여 복원을 수행하는데, 즉 웨이블릿 변환 계수 정보를 이용하여 영상 및 시간 정보를 여러 스케일의 에지로 분류한 다음, 에지의 방향에 따른 적응적인 제약조건을 사용한다. 이는 각각의 에지 방향에 적합한 고주파 성분을 유지하고, 신호의 각 특성에 적합한 적응적인 정칙화 매개변수를 적용한다. 마지막으로 시간 축에서의 복원과 그것에 이어지는 적응적인 공간 복원에 대한 실험 결과를 보여준다.

1. 서론

웨이블릿 기반 영상 부호화 방식은 높은 압축률과 부호화 및 전송 시의 편리함 때문에 갈수록 많은 응용 분야에서 널리 사용되고 있다. 특히, 3차원 웨이블릿 부호화 기반 동영상 압축 방식은 새로운 동영상 압축 표준 및 응용 분야에서 제안되고 있다.

본 논문에서는 3차원 웨이블릿 부호화 방식으로 압축된 영상 시퀀스를 정칙화 기반 영상복원 방법으로 후처리하는 알고리즘을 제안한다. 3차원 웨이블릿 기반의 영상 시퀀스 압축 시스템은 이미 여러 논문들을 통하여 다양한 알고리즘들이 제안되어 왔다[1]. 여기서 언급되는

3차원 웨이블릿 부호화 기반 영상 시퀀스 압축은 가장 기본적인 방식으로, 프레임의 2차원 영상 정보에 영상 시퀀스의 시간 신호를 더하여 3차원 신호를 만든다. 그런 후, 각 프레임의 공간 변환 뿐만 아니라 이에 해당하는 시간 신호를 시간 축으로 다 해상도를 갖는 3차원 웨이블릿 변환을 적용하는 방법이다.

공간 영역의 압축 과정을 통해 링 현상(ringing artifact) 등이 나타나는데, 공간 영역에서와 마찬가지로 시간 신호의 부호화 시에도 압축 및 전송 시 영상의 고주파 웨이블릿 변환 계수들의 손실에 의한 열화 현상이 발생하며 주로 움직임의 잔상으로 나타난다. 따라서 본 논문에서는 2차원 영상 웨이블릿 계수뿐만 아니라, 시간 축의 웨이블릿 계수들에 대해서도 반복적인 복원 방법을 적용하여 열화 현상을 제거한다.

우선, 웨이블릿 압축 시스템을 적절한 영상 열화 시스템으로 모델화한다. 그리고 압축 시스템에 적합한 제약 조건들을 찾아서 정칙화 기반으로 한 반복적인 영상복원 알고리즘을 도출한다. 일단 시간 축은 프레임 간의 같은 위치에 있는 각 픽셀에 대하여 정칙화 매개변수를 적용해 복원을 수행한다. 그 다음, 2차원 영상 신호에 대해서는 웨이블릿 변환 계수 정보를 이용하여 영상 및 시간 정보를 여러 스케일의 에지로 분류, 해당하는 에지의 방향에 따른 적응적인 제약조건을 사용한다. 이는 각각의 에지 방향에 적합한 고주파 성분을 유지하고, 신호의 각 특성에 적합한 적응적인 정칙화 매개변수를 적용한다.

마지막으로 실험 결과는 제안된 알고리즘이 객관적 및 주관적으로 성능이 뛰어남을 보여준다.

2. 웨이블릿 기반 압축 시스템

기본적인 영상 열화 시스템에 대한 수학적 모델은 다음 식으로 나타낼 수 있다 [2],[3]

$$y = Dx \quad (1)$$

1) 본 연구는 교육부 두뇌 한국 21(BK21) 사업 지원에 의해서 수행되었습니다.

$y, x: N^3 \times 1$ 벡터 $D: N^3 \times N^3$ 행렬
 여기서, y, D 및 x 는 각각 열화된 영상, 열화 시스템
 그리고 원 영상을 나타낸다.

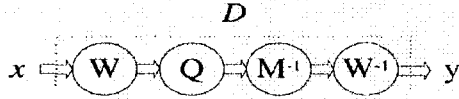


그림 1. 웨이블릿 기반 압축 시스템의 열화 모델

식(1)을 그림 1과 같이 웨이블릿 기반 압축 시스템으로 해석할 수 있다. 여기서, W, Q, M^{-1} 및 W^{-1} 은 각각 웨이블릿 변환, 양자화, 역양자화 그리고 역 웨이블릿 변환을 나타낸다. 블록 기반 이산 여현 변환(BDCT) 압축 시스템과 마찬가지로 웨이블릿 압축 및 재생의 전체 시스템을 간단히 열화 시스템 D 로 표현할 수 있다.

$$D = W^T M Q W \quad (2)$$

이제 3차원 웨이블릿 변환에 대해 알아본다. 먼저 프레임의 2차원 영상 정보에 영상 시퀀스의 시간 신호를 더하여 3차원 신호를 만든다. 그런 후, 시간 신호를 시간 축으로 1차원 웨이블릿 변환을 하고 다시 각 프레임별로 공간 변환을 적용하는 방법이다.[1] 이는 대부분 분리 가능한(separable) 웨이블릿을 사용하기 때문에 시간 축으로의 1차원 변환에 다시 공간의 2차원 변환을 수행하여 단순한 3차원으로 확장할 수 있다.

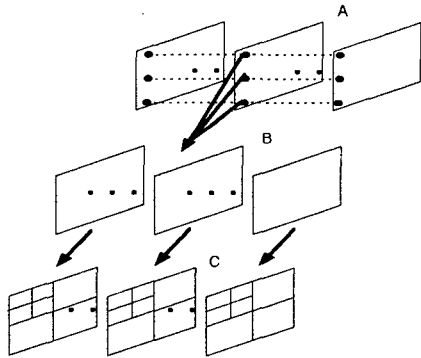


그림 2. 영상 시퀀스의 3차원 웨이블릿 변환

그림 2에서 A는 원 영상시퀀스를, B는 A의 시간정보를 웨이블릿 압축한 시퀀스를 나타내고, C는 이를 다시 각 프레임의 공간 웨이블릿 압축한 시퀀스를 나타낸다. 따라서 본 논문에선 그림 1의 W 는 시간 변환에 공간 변환이 합쳐진 변환이다.

3. 반복적 정칙화 영상 복원

3.1 시간에 대한 반복적 정칙화 복원

영상을 정칙화 기반으로 복원하기 위해선, 식(1)의 해 x 가 범함수(functional)를 최소화 시켜야 한다. 본 논문에서 시간 축에 적용되는 범함수는 기본적인 형태를 이용한다.

$$f(x) = \{\|y - Dx\|^2 + \lambda \|Cx\|^2\} \quad (3)$$

이를 최소화 시키는 해는 반복적 정칙화 방법에 의해 다음과 같이 찾아진다.

$$x^{k+1} = x^k + \beta \{D^T y - (D^T D + \lambda C^T C)x^k\} \quad (4)$$

그러나 이는 웨이블릿 압축 과정에서도 알 수 있듯이 웨이블릿 열화 행렬 D 의 대칭성에 의해

$$D^T D = D D = D \quad (5)$$

식(6)으로 정리된다.

$$x^{k+1} = x^k + \beta \{y - (D + \lambda C^T C)x^k\} \quad (6)$$

여기서 C 는 선형 고 대역통과 필터로 식(3)을 최소화 시키는 과정에서 평활(smoothness) 제약조건으로 작용하며, 본 논문에서는 다음과 같은 간단한 형태를 사용하였다.

$$C = [-0.25 \quad 0.5 \quad -0.25]^T$$

\square 는 정칙화 매개변수(regularization parameter)로써 최소화 과정에서 평활 제약조건을 양을 조절한다.

3.2 공간 반복적 정칙화 영상 복원

본 논문에서 공간에 적용되는 범함수(functional)는 다음과 같다[2].

$$f(x) = P\{\|y - Dx\|^2 + \lambda_1 \|Cx\|^2 + \lambda_2 \|a - x\|^2\} \quad (7)$$

이 식에서 두 번째 항의 C 는 선형 고 대역통과 필터로 위 식의 최소화 과정에서는 평활(smoothness) 제약조건으로 작용한다. 세 번째 항에서 a 는 미디언(Median) 필터된 압축영상으로 에지 부근의 링 현상을 빨리 감소시켜 주는데 도움을 준다. λ_1 과 λ_2 는 정칙화 매개변수(regularization parameter)로써 두 제약조건을 양을 조절하는 라그랑주 승수(Lagrange multiplier)이다.

P 는 정칙화 과정 밖에서 사용되는 hard 제약조건으로 본 논문에서는 두 가지 방법이 사용된다. 첫 번째는 양자화 제약조건으로서, 정칙화 과정 동안 변경될 수 있는 각 화소의 밝기 값들이 반드시 해당 양자화 레벨 범위 안에 속하도록 한다. 두 번째는 gradient descent 방법 대신 비선형 함수를 사용하여 근사화 된 해를 찾는 방식으로[4], 에지 주변은 물론 넓게 퍼진 링 현상을 줄이는데 도움이 된다.

시각적으로 중요한 영상의 고주파 성분을 살리기 위한 방법으로 적응적인 고 대역통과 필터를 사용한다[5]. 이를 위해서는 우선 에지 정보가 필요하나, 압축된 영상들은 많은 고주파 성분들이 손실되어 있다. 이로 인해 에지 주변에는 비정상적인 고주파 성분이 포함되어 있어 정확한 에지의 검출은 어렵다. 따라서 멀티스케일 에지 탐색기(multiscale edge detector)[5]를 이용하여 보다 정확한 에지를 찾아 낸다. 이렇게 검출된 에지의 방향에 따른 서로 다른 고 대역통과 필터($C_m, m=1..M$) 사용한다. M 은 에지 방향의 수를 나타낸다. 즉, M 을 5로 결정한다면, 에지의 방향은 1은 평활 영역, 2는 수평 에지, 3은 수직 에지, 4와 5는 각 대각 성분의 에지에 해당한다. 위와 같은 방법들을 이용하여 식(3)을 최소화 시키는 해를 찾으면 식(4)과 같다.

$$x^{k+1} = x^k + \beta(b - \sum_{m=1}^M I_m T_m x^k) \quad (4)$$

$$T_m = D + \lambda_1 C_m^T C_m + \lambda_2 I \quad (5)$$

여기서 M 은 에지 방향 갯수를 의미한다. I_m 은 대각선 성분이 1이나 0으로 된 대각 행렬으로써 각 에지에 적합한 C_m 를 선택하도록 한다.

4. 실험결과

1차원, 시간 변환에 이어 공간 변환을 통해 압축된 영상 시퀀스에 대해 실험하였다. 이 때 시간 축에 대한 웨이블릿은 Haar 웨이블릿을 사용하였다. 그리고, 복원시에는, 시간 축에 대해서 복원을 수행하고, 다시 이 결과값들을 공간에 대해서 반복적 복원을 적용하였다.

본 논문의 실험에 사용된 프레임들의 개수는 총 16개이다. 시간과 공간에 모두 3번의 대역분할($L=3$)을 수행하고 단순히 양자화만을 하는 간단한 압축 방식을 택하였다. 그리고, 복원과정은 10번에 걸쳐 반복하였다.

이 실험에 사용된 매개변수의 값은 다음과 같다. 먼저 1차원 복원에서는 $\beta=0.3$, $\lambda=0.3$ 을, 공간 복원에서는 $\beta=0.001$, $\lambda_1=2.0$, $\lambda_2=0.5$ 을 사용하였으며, 여기에 사용된 변수의 값들은 실험적으로 얻어진 것들이다.

다음은 16개의 프레임들 중 임의의 한 프레임의 PSNR 변화를 보여준다.

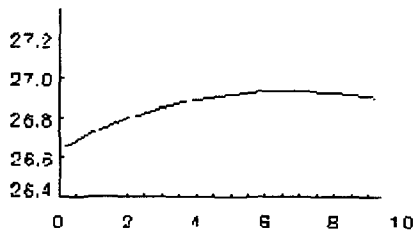


그림 3. 시간 축으로의 복원 결과

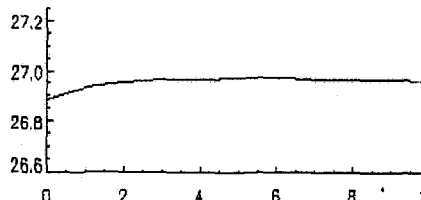


그림 4. 공간의 반복적 정칙화 복원 결과

그림 3에서 가로축은 반복 횟수를, 세로축은 반복 횟수에 따른 PSNR값을 나타낸다. 이 그림에서 알 수 있듯이 초기 PSNR 값인 26.7에서 반복적 적응 방식으로 복원된 값은 26.884로, 0.16만큼 증가했다. 또한 PSNR의 수렴성을 알 수 있다.

그림 4도 공간에 대한 복원 결과를 반복 횟수에 따른 PSNR 변화값을 나타낸다. PSNR이 26.9369으로 수렴함을 볼 수 있다. 즉, 시간-공간 복원에 의해 PSNR이 26.7에서 26.937로 0.293228만큼 증가했다.

그림 5는 3차원 웨이블릿 방식으로 압축된 프레임들을 보여준다. 그림 6은 그림 3에 해당하는 결과 영상으로, 시간 축으로 반복적 복원을 적용한 것이다. 그림 7은 공간 계수에 적응적인 방법을 반복적으로 적용하여 얻어진 결과 영상이다.



그림 5. 3차원 웨이블릿 기반으로 압축된 영상



그림 6. 그림 5의 시간축에 대한 반복적 복원을 한 결과



그림 7. 그림 6의 공간 적응적 반복 복원을 한 결과

5. 앞으로의 연구 계획

처리속도와 성능을 향상시키기 위해 계층적인 구조로 확장 시켜야 하며, 실시간 처리를 위한 방법도 간구해야 한다. 그리고 시간 신호에 대한 보다 적응적인 복원 방법을 이용하여야 할 것이며, 특히 움직임이 보강된 동영상을 처리할 수 있는 알고리즘에 대한 연구가 더 필요하다.

참고 문헌

- [1] Christine I. Podilchuk, "Three-dimensional subband coding of video," *IEEE Transactions on image processing*, vol 4, no.2, February 1995.
- [2] A. K. Katsaggelos, "Iterative image restoration algorithms," *Optical Engineering*, vol. 28, no. 7, pp. 735-748, July 1989.
- [3] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.
- [4] Meiyin Shen, "Postprocessing for removing wavelet coding artifacts," ISO/IEC JTC1/SC29/WG1N935, June 1998.
- [5] Stephane Mallat, Sifen Zhong, "Characterization of signals from multiscale edges," *IEEE Trans on pattern analysis and machine intelligence*, vol. 14, no.7, July 1992.
- [6] J. H. Jung, J. H. Shin, and J. K. Paik, "Spatio-temporally adaptive image sequence interpolation," *Proc. 1998 Int. Tech. Conf. Circuits, Systems, Computers, Communications*, vol. 1, pp. 43-46, July 1998.
- [7] Seung-Jong Choi and John W. Woods, "Motion-Compensated 3-D Subband Coding of video," *IEEE Transactions on image processing*, vol. 8, no. 2, February 1999.
- [8] Ian K. Levy and Roland Wilson, "Three Dimensional Wavelet Transform Video Compression," *Proceedings of the IEEE Multimedia Systems '99*, vol 2, pp. 924-928, June 1999.
- [9] J. H. Jung, S. C. Joung and J. K. Paik, "Regularized constrained restoration of wavelet compressed image," *Proc, SPIE Visual Comm, Image Proc.*, Australia, June 2000.
- [10] J. H. Jung, S. C. Joung, Y. H. Jang and J. K. Paik, "Enhancement of Region-of-Interest Coded Images by Using Adaptive Regularization," *Digest of Technical Papers, 2000 Int. Conf. Consumer Electronics*, to appear, LA, June 2000.
- [11] J. H. Jung, and J. K. Paik, "Restoration of wavelet-compressed images using adaptively masking smoothness constraints," *Proc. 2000 Int. Conf. Image Processing*, to appear, Canada, October 2000.