

# 효율적이고 확장가능한 3D-WT 압축기법

김성민<sup>o\*</sup> 박시용<sup>\*</sup> 이승원<sup>\*</sup> 이화세<sup>\*\*</sup> 정기동<sup>\*</sup>

부산대학교 전자계산학과

{morethannow<sup>o\*</sup>, sympark<sup>\*</sup>, swlee<sup>\*</sup>, kdchung<sup>\*</sup>}@melon.cs.pusan.ac.kr, hstee<sup>\*\*</sup>@mnu.ac.kr

## An Efficient and Scalable 3D-WT Compression Scheme

Sungmin Kim<sup>o\*</sup> Siyong Park<sup>\*</sup> Seongwon Lee<sup>\*</sup> Hwasee Lee<sup>\*\*</sup> Kidong Chung<sup>\*</sup>

<sup>\*</sup>Dept. of Computer Science, Pusan National University

<sup>\*\*</sup>Dept. of Computer Engineering, Miryang National University

### 요약

기존의 비디오 코딩에서는 연속된 프레임의 시간적인 상관성을 제거하기 위한 방법으로 이전 프레임의 정보를 이용하여 현재 프레임을 예측하는 움직임 예측기법을 많이 사용하고 있다. 정지 화상에 비해서 대용량의 특성을 지니는 비디오 데이터는 이런 움직임 예측을 통해서 대부분의 압축이 일어나게 된다. 하지만 움직임 예측기법은 많은 계산과정을 요구하므로, 전체적인 부호기 복잡도를 높이는 단점을 지닌다. 반면 3D-WT는 움직임 예측을 하지 않으므로, 부호기의 복잡도를 줄일 수 있다.

하지만, 기존의 3D-WT기법들은 부호화를 위한 메모리 요구사항과 복호를 위한 수신측의 지연시간이 가장 큰 단점으로 지적되었다. 따라서, 본 논문에서는 메모리 요구사항과 수신측의 지연시간을 최소화 할 수 있는 효율적이고, 확장가능한 3D-WT 기법을 소개한다.

지연시간을 줄이는 기법을 소개한다.

기존의 3D-WT의 기법들은 연속적인 프레임들에 시간적인 분해를 하기 위해서, GOF를 구성하는 모든 프레임을 입력으로 받아서 처리한 후 전송하게 되지만, 본 논문에서는 GOF의 크기와는 상관없이 4장의 프레임만을 입력으로 받아서 연속적인 프레임에 대한 시간적인 분해를 하게 된다. 복호시에도 기존의 기법들이 GOF의 크기에 해당하는 전 프레임을 다 받던지 일정 개수 이상의 프레임을 받아야만 복호를 시작할 수 있었으나, 본 논문에서 제시하는 기법은 4장의 프레임만을 받으면 복호를 바로 시작할 수 있다. 따라서, 제시하는 기법은 부호화 과정에서 필요한 메모리 요구사항뿐만 아니라, 복호화 과정에서 발생되던 지연시간도 최소화시켰다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구를 살펴보고, 3장에서는 3D-WT 기법을 소개한다. 4장에서는 본 논문에서 제안하는 기법에 대해서 다루며, 5장에서는 실험을 통해 성능을 평가한다. 마지막 6장에서는 연구결과 및 향후과제를 제시한다.

### 2. 관련연구

JPEG, MPEG 표준이 생긴 이후 표준에서 채택되었던 DCT 기법은 이미지 압축을 위해서 가장 훌륭한 대안으로 연구되고 개발되었다. 그러나, 압축율을 높일 때 발생하는 블러킹 현상으로 다양한 네트워크에서 서비스를 제공하기에는 어려움을 가지고 있다[1][6]. 이에 반해, wavelet기법은 압축율에 따른 화질의 저하가 DCT에 비해 거의 발생하지 않으므로, 이미지 코딩 뿐만 아니라 비디오 코딩을 위한 새로운 방법으로도 많은 연구가 진행되었다[2]-[4].

wavelet을 사용하는 비디오 코딩은 기존의 DCT를 사용한 기법과 같이 움직임 예측을 사용하는 경우와 움직임 예측을 사용

### 1. 서론

인터넷의 폭발적인 성장은 오디오, 비디오와 같은 멀티미디어 서비스의 획기적인 발전을 이끌었고, 서비스의 질적인 향상은 사용자들이 더 높은 질의 서비스를 요청할 수 있도록 하였다. 현재 인터넷 상에서 VoD 서비스를 받는다면, 인터넷을 통해 live 방송을 보는 것은 일반적인 서비스 형태로 굳어져 있다. 정지화상과 달리 대용량의 데이터를 가지는 비디오를 인터넷을 통해 전송해서 서비스하기 위해서는 비디오에 대한 여러 가지 코딩 기술들이 필요하다. 연속된 프레임들로 구성된 비디오 영상의 코딩은 시간적으로 프레임들의 정보를 줄일 수 있는 방법이 요구되는데, 기존의 비디오 코딩에서는 이전 프레임의 정보를 현재 프레임에 이용하는 방법인 움직임 예측 기법이 주로 이용되었다.

대용량의 비디오 데이터는 움직임 예측을 통해서 대부분의 정보를 압축한다. 하지만, 이전 프레임을 참조해서 현재 프레임의 정보를 코딩하는 방법이기 때문에, 이전 프레임 픽셀들과의 복잡한 계산과정이 수반되는 단점을 지니고 있다.

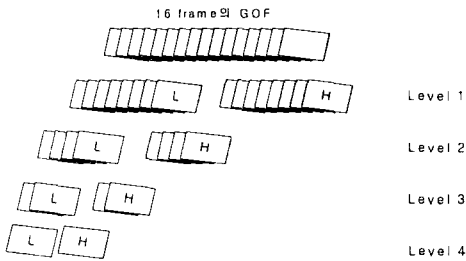
이에 반해 3D-WT기법은 연속적인 프레임들을 하나의 작업공간으로 두고 압축을 하는 기법으로 움직임 예측을 필요로 하지 않는다. 움직임 예측 없이 비디오 데이터를 코딩하므로, 압축을 위해 복잡한 계산과정이 필요 없다는 장점을 지니고 있다 [7]-[9]. 그러나, 움직임 예측은 프레임을 한 장씩 처리하는 반면, 3D-WT는 여러 장의 프레임을 동시에 처리해야 하므로 필요로 하는 메모리의 양이 늘어나는 단점이 있다. 또한, 데이터를 복호하기 위한 수신측에서의 지연시간도 늘어나게 된다. 그러므로, 본 논문에서는 3D-WT가 가지는 메모리 요구사항과

하지 않는 기법으로 나눌 수 있다. 움직임 예측을 사용할 경우에도 DCT를 사용한 방법과 달리 주파수 영역에서 움직임 예측을 하는 기법에 대해서도 연구가 활발히 진행되고 있다.

[10][11]에서는 주파수 영역에서 움직임 예측을 하는 기법들을 소개하고 있다. 또한, 여러 논문에서는 움직임 예측을 사용하지 않고 3차원 영역에서 비디오 데이터를 코딩하는 기법에 대해서도 소개하고 있다[7]-[9].

### 3. 3D-WT

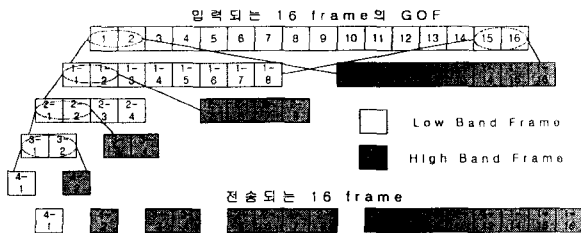
전통적인 3D-WT기법은 전체 비디오를 하나의 시퀀스로 처리한다. 그러나, 전체 시퀀스를 하나의 처리 단위로 보기에는 긴 시간의 비디오 데이터에 적절치 못하다. 따라서, 비디오 시퀀스에 일정 크기의 GOF 개념을 도입하여 처리를 한다.



[그림 1. 16프레임의 시간적인 분해과정]

[그림1]은 16개의 프레임이 GOF로 가지는 3D-WT의 기본 개념을 보인다. 4단계의 분해과정을 가진 원래의 프레임들은 저대역의 1프레임과 15개의 고대역 프레임으로 나누어진다. 시간적인 분해를 통해서 만들어진 고대역 프레임들은 전체적으로 작은 값들로 이루어져 있으므로, 높은 압축 효율을 얻을 수 있다.

[그림1]을 통해 알 수 있듯이 시간적인 분해를 하기 위해서는 우선 16장의 프레임을 입력으로 받아 들여서 처리를 해야 한다. 따라서, 부호화를 위한 메모리 요구량이 많아진다. 또한, [그림2]에서는 수신측에서 첫 프레임을 복호하기 위해서는 최소 9프레임을 수신해야만 재생이 가능함을 설명하고 있다. 즉, 수신측에서 복호를 위한 지연시간이 GOF 크기의 1/2이상으로



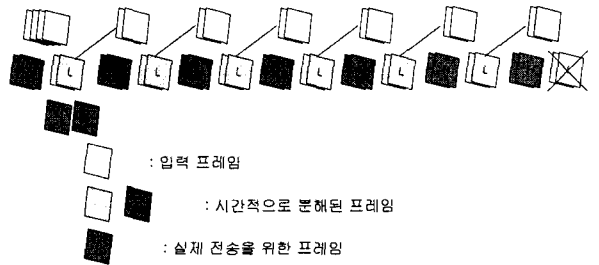
[그림 2. 시간적인 분해 후의 전송 순서]

갈어진다. 지연시간을 줄이기 위해서 프레임의 전송순서를 재조정하는 방법으로 보다 빠른 재생이 가능하지만, 여전히 메모리 요구사항에 대한 단점은 존재한다.

3D-WT 기법은 높은 압축 효율을 위해 많은 프레임을 이용해서 시간적인 분해를 해야만 한다. 하지만, 그에 따른 메모리 요구사항과 지연시간이 증가하게 되는 단점을 가지고 있다.

### 4. 효율적이고 확장가능한 3D-WT

앞에서 언급했듯이 3D-WT의 기법은 여러 가지의 단점을 보인다. 여러 장의 프레임을 받아서 처리하는 프로세싱 절차를 거치므로, 많은 메모리가 필요하고, 부호화 처리 단위가 한 장이 아니므로, 부호 후 전송, 전송 후 복호까지 많은 지연시간을 가지게 된다. 본 장에서는 이러한 단점을 극복할 수 있는 기법을 제시한다.

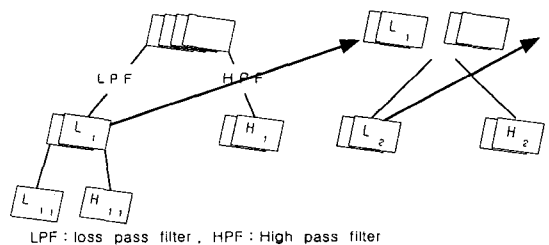


[그림 3. 메모리 요구사항과 지연시간을 줄이는 새로운 알고리즘]

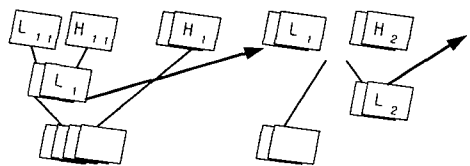
[그림3]과 같이 입력되는 최초의 4프레임을 먼저 2단계 분해한 후 분해된 4개의 프레임을 먼저 전송하고, 이후에 2프레임 연속적으로 입력받아, 계속적으로 시간적인 분해를 하게 된다.

기존의 기법들에서는 16장의 프레임을 모두 받아들이거나, 메모리 요구사항을 줄였다 해도 복호기 측에서의 지연은 어쩔 수 없는 선택이었다. 하지만, 제안하는 기법은 프레임의 입력과 더불어 부호화가 일어날 수 있고, 복호기 측에서는 전송되어 오는 프레임들을 통해 즉시 원래의 프레임을 복구할 수 있는 능력을 제공한다.

제안하는 기법은 4프레임을 작업단위로 가지므로, 메모리 요구사항은 GOF의 크기에 상관없이 4개의 프레임으로 유지되고, 복호기 측에서의 지연시간도 GOF의 크기에 상관없이, 항상 일정하게 유지시킬 수 있다.



(a) 부호화 과정



(b) 복호화 과정

[그림 4. GOF의 처음 6프레임에 대한 부호화와 복호화의 예]

[그림4]에서 부호화와 복호화에 대한 예를 보이고 있다. 최

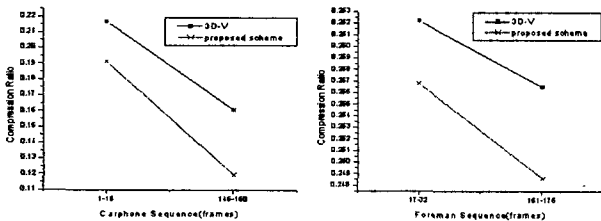
초에 입력되는 4프레임은 2단계의 분해과정을 통해서, 저대역 프레임 1개와 고대역 프레임 3개로 우선 분해한다. 여기서 분해된 4개의 프레임은 바로 전송이 가능하고, 복호측에서는 이 프레임을 받는 즉시 복호가 가능하게 된다.

[그림4-(a)]를 통해서도 알 수 있듯이, 실제 전송되는 데이터와 시간적인 분해를 하기 위한 프레임을 다르게 줌으로써, 연속적인 분해를 위한 메모리 사용을 줄이고, 복호를 위해 긴 지연시간을 요구하지 않고서도 가능하게 하였다. 이러한 과정을 가능하게 하는 1단계 필터와 첫 4프레임에만 적용되는 2단계 필터는 다음과 같다.

[표 1. 각 단계에서 사용되는 부호화 필터]

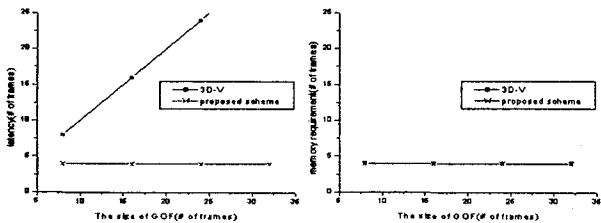
	level 1	level 2
low pass filter	[1/2, 0, 1/2, 0]	[1/2, 1/2]
high pass filter	[1/2, 0, -1/2, 0]	[1/2, -1/2]

5. 실험 및 결과



(a) carphone.qcif

(b) foreman.qcif



(c) 수신측 지연시간

(d) 메모리 요구사항

[그림 5. 3D-V와의 압축율, 지연시간, 메모리 요구사항 비교]

실험은 foreman과 carphone 시퀀스를 통해 [8]에서 제시된 3D-V 기법과의 성능을 비교하였다. 메모리 요구사항을 제외한 수신측에서의 재생을 위한 지연시간, 부호화 과정을 통한 압축율에서 3D-V 기법보다 나아짐을 알 수 있다. 전체적인 화면의 움직임이 적은 carphone에서 foreman보다 훨씬 향상된 압축 성능을 보인다. 또한, 3D-V와는 달리 GOF의 크기에 상관없이 수신측에서의 지연시간은 일정하게 유지되는 것을 확인할 수 있었다.

6. 결론 및 향후과제

\* 압축율 = 시간분해 후의 데이터의 양 / 입력되는 데이터의 양

비디오 코딩 기법에서 복잡한 계산과정의 움직임 예측을 피하기 위한 3D-WT 기법은 계산과정을 단순화시킨 장점에 비해, 메모리 요구사항과 수신측의 지연시간에 있어서 단점을 지니고 있다.

본 논문에서는 부호화할 때의 메모리 요구사항과, 수신측에서 재생을 위해 필요한 지연시간을 최소화할 수 있는 기법에 대해서 논의하였고, 기존의 기법들이 가지던 단점을 보완하여 좀 더 향상된 능력을 보여주고 있다. 하지만, 실제적인 서비스를 위한 압축율과 에러에 대한 강인성 측면에서 검증되지 않은 많은 부분이 남아 있다.

향후에는 움직임 예측을 통한 비디오 코딩 기법과의 데이터 압축율을 비교할 것이고, 시간적인 분해를 통해 압축된 프레임 정보가 에러에 대해 어느 정도의 강인성을 지니고 있는지 조사할 것이다.

참고문헌

- [1] "MPEG-2 video", ITU-T Recommendation H.262-ISO/IEC 13818-2, Jan. 1995.
- [2] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3463, Dec. 1993.
- [3] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243-250, June 1996.
- [4] Zixiang Xiong, Ramchandran K., Orchard M.T. and Ya-Qin Zhang, "A comparative study of DCT- and wavelet-based image coding", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 692-695, Aug 1999
- [5] Aksay, A., Bozdagi, G., Akhan, M.B., Temizel, A. and Duhamel, P.B.D., "Motion wavelet compression", *Time-scale and Time-Frequency Analysis and Applications, IEE Seminar on*, 2000
- [6] Wing-Kuen Ling and Bing Zeng "A novel method for blocking effect reduction in DCT-coded images", *Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium on*, vol. 4, pp. 46-48, Jul 1999
- [7] Salvarlo J. and Sousa L., "Video Coding by using the 3D Zero-Tree approach in the Wavelet transform domain", *Digital Signal Processing, 14th International Conference on*, vol. 2, pp. 683-687, 2002
- [8] E. Moyano, F.J. Quiles, A. Garrido, L. Orozco-Barbosa and J. Duato, "Efficient 3D Wavelet Transform Decomposition For Video Compression", *Digital and Computational Video, 2001. Proceedings. Second International Workshop on*, pp. 118-125, 2001
- [9] C. I. Podilchuk, N. S. Jayant, and N. Farvardin, "Three-dimensional subband coding of video", *IEEE Trans. Image Processing*, vol.4, pp.125-139, Feb. 1995
- [10] Ming Fai Fu, Au, O.C. and Wing Cheong Chan, "Novel motion compensation for wavelet video coding using overlapping", *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, vol.4 pp.3421-3424, 2002
- [11] H.W. Park and H.S. Kim, "Motion estimation using low-band-shift method for wavelet-based moving-picture coding", *Image Proc, IEEE Trans.*, vol.9, pp.577-587, Apr 2000,