

동영상 부호화를 위한 scalable 구조에 관한 연구

신중인*, 한영오, 김형근*, 박상희

연세 대학교 전기공학과

*한국과학기술연구원 응용전자연구실

A Study on the Scalable Structure for Motion Picture Coding

Joong In Shin*, Young Oh Han, Hyung Kon Kim*, Sang Hui Park

Dept. of Electrical Engineering, Yonsei Univ.

*Applied Electronics Laboratory, KIST

ABSTRACT

In this paper, we study the structure of the hierarchical coding method of video signal which can contain the multiresolution video signals. To preserve the compatibility with the conventional coding methods, we accomplished a scalable structure using the subband coding, maintaining enough the international coding structure. The proposed scheme showed the low PSNR, a little, when compared with the conventional scheme, but showed a good image quality perceptually and proved to have a advantage in the H/W implementation in a view of processing speed.

I. 서론

최근 완료되었거나 진행되고 있는 영상신호의 압축에 관한 연구들은 대부분 단일 계층으로 구성된 영상부호화 방식의 성능을 향상시키는데 그 목표를 두고 있다. 그러나 다양한 정보화 사회가 되어감에따라 많은 종류의 영상기기가 출현하게 되었고, 이들을 상호 연결하여 호환성이 있도록 하는 작업이 요구되고 있으며, 가장 최근에 진행되고 있는 이러한 영상신호의 부호화 및 복호화에 관한 국제적인 표준화 작업으로는 MPEG2를 예로 들수 있다. [1] 이러한 일련의 작업에서는 다양한 영상형식간의 상호 호환성을 갖게하는 방법으로서 영상의 scalability와 compatibility를 정의하여 그 부호화 및 복호화 방법을 연구하고 있다.

이 중에서, 영상신호의 scalable한 부호화 방식은 비트열의 일부분만을 복호화했을 때에도 사용가능한 영상들을 출력할 수 있는 비트열로 부호화하는 것으로, non-scalable한 방법에 비하여 약간의 화질열화와 H/W의 복잡도가 증가하는 것을 감수하면서도 다양한 해상도의 영상들을 수용할 수 있도록 하는 것이 최대의 목적이 된다. [2][3] 그러므로 영상신호의 압축구조에 있어서 scalability를 포함한 설계를 행하면, 위와 같은 제한 조건을 갖는 경우에 있어서 많은 양의 신호처리용 H/W나 프레임 메모리를 갖고 있을 필요가 없다는 장점을 갖게 된다.

scalable한 영상신호의 부호화의 방법에는 공간영역(spatial domain)에서 행하는 방법과, 공간 주파수영역(spatial frequency domain)에서 행하는 두가지의 방법이 제안되고 있는데 [1][2][3], 일반적인 구조를 갖고있어서 다중 해상도를 갖는 계층적 부호화 방법이 최근에 많이 제안되고 있다. [8][9]

본 논문에서는 공간 주파수영역(spatial frequency domain)에서 scalable한 영상신호의 부호화를 행하는 방법을 제안하는데 그 목적을 두고 있다. 기존의 제안된 표준 부호화 방식과의

호환성을 유지하기 위하여 기본적으로는 MPEG2 TM3의 구조[1]를 유지하면서 다중해상도(multi-resolution)와 다중주파수(multi-frequency)의 구조를 구현하는데 뛰어난 특성을 갖고있는 것으로 알려진 대역분할(Subband Decomposition)을 이용하고, 또한 복잡한 비트열을 구성함으로써 MPEG2의 기본구조를 유지하면서 공간 주파수영역에서의 scalability를 실현하는 방법을 제시하고자 한다.

II. 대역 분할 부호화

1. 영상신호의 대역분할

대역분할 부호화는 1976년, Crochiere가 처음으로 음성신호의 부호화에 적용하였고, 그 후 영상신호에도 적용되기 시작하여, 1984년 Vetterli가 1차원 신호의 대역분할 및 합성기법을 다차원 특성을 갖는 신호에 확장하여 2차원 직각대칭필터(QMF)를 사용한 2차원 신호의 대역분할 및 합성에 대한 이론을 확립하였다. [4] 1986년, J.W. Woods, S.D.O'Neil등은 정지영상신호에 대하여 4대역 분할 및 합성을 이용하여 대역부호화를 행하여 좋은 결과를 얻었으며[5] 1988년, Westerink등에 의하여 1차원 QMF필터를 이용하여 분리가능한 2차원 필터의 구성이 가능해졌는데[6], 분리가능한 필터를 이용하여 구성된 4대역 대역분할 필터트리는 그림 1과 같고 본 연구에서는 대역분할을 위하여 그림 1의 필터트리를 이용하였다.

2. 대역분할 부호화에서의 비트할당

대역분할에 의하여 형성된 각 신호들은 각 대역별로 각각 다른 통계적인 특성을 지니게 되기 때문에 대역별 비트할당이 가능하다. 즉, 랜덤벡터 $X = \{X_1, X_2, \dots, X_k\}$ 가 있을 때, 각 벡터의 요소는 평균 0과 분산,

$$\sigma_k^2 = E[X_k^2] \quad (1)$$

을 갖고, 각 X_k 는 동일한 정규확률밀도함수(Normalized probability density function)를 갖는다고 할 때, 동일한 대역폭을 갖는 대역분할을 가정하고 k번째 대역에 R_k 비트를 할당하면 전체 대역을 부호화하는데 쓰이는 평균 비트율 R 은,

$$R = \frac{1}{M} \sum_{k=1}^M R_k \quad (2)$$

가 된다. 대역분할에 있어서의 비트할당은 각 대역의 분산에 의하여 결정될 수 있는데, 비트할당의 목적은 왜곡이 최소화될 수 있도록 각 요소에 최적의 비트 R_k 를 할당하는 것으로 중첩되지 않는 동일한 대역폭을 갖는 대역분할에 있어서의 최적 비트할당은 다음 식으로 구할 수 있다. [10]

$$R_{k,opt} = R + \frac{1}{2} \log_2 \frac{\sigma_k^2}{\left[\prod_{k=1}^M \sigma_k^2 \right]^{1/M}}$$

$$= R + \frac{1}{2} \log_2 \left[\frac{\sigma_{rk}^2}{\sigma_{gm}^2} \right] \quad (3)$$

여기서, R_k, σ_{rk}^2 은 각 대역의 최적 할당비트, R 은 평균 할당 비트, M 은 분할대역의 수, σ_{rk}^2 은 분할대역의 분산. 그리고 $\sigma_{gm}^2 = \left[\prod_{k=1}^M \sigma_{rk}^2 \right]^{1/M}$ 이며 σ_{rk}^2 의 기하평균이다.

그런데 식 (3)은 높은 비트율에 대해서만 정확한 값을 가지며, 더우기 낮은 비트율에서는 음수의 비트 할당이 발생할 가능성이 매우 높다. 이 문제를 해결하기 위해서는 음수값을 적절한 문턱값인 T 로 절단하고, 원래의 목표 비트율을 초과하는 비트수 만큼을 각 대역의 할당된 비트에서 고르게 차감함으로써 음수의 비트 발생을 억제하는 방법을 사용했다. 즉,

$$R_k = \begin{cases} T & R_k < T \\ R_k & R_k \geq T \end{cases} \quad (4)$$

그리고,

$$R_i = \begin{cases} R_i - (T - R_k) / (BandNum. - 1) & R_k < T \\ R_i & R_k \geq T \end{cases} \quad (5)$$

단, 여기서 $1 \leq i, k \leq BandNumber$, $i \neq k$ 이다.

III. 대역분할을 이용한 scalability

1. 부호화를 위한 전처리 방법

본 논문에서는 서론에서 서술한 바와 같이 기존의 국제 표준화 방식의 골격을 최대한 유지하면서, 간단한 수정만으로도 scalability의 실현이 가능한 그림 2와 같은 대역분할 방법과 그림 3과 같은 부호화를 이용하여 2개의 계층(scale-4,8)을 갖는 scalable구조에 관하여 연구를 하였다.

이 구조에서는, 각 대역의 정보가 서로 다른 특성을 갖기 때문에 DCT 영역에서의 scalability를 실현하는 방법과는 달리, 동일한 정보를 반복하여 부호화할 필요가 없다. 따라서 한정된 대역의 채널을 사용하는 경우 작은 크기의 해상도의 영상만을 추출해도 대역의 낭비를 최소화 할 수 있다는 장점이 있다. 그리고 대역분할에 의하여 일반적인 피라미드형의 구조를 실현하여 복호화기측에서는 원하는 해상도의 영역만을 추출함으로써 보다 간단하게 scalability를 실현할 수 있다.(그림 4)

대역분할을 하기위해서 QMF 필터를 사용하는데, 여기서는 실제 하드웨어 구현시 실시간 처리를 고려하여 본 논문에서는 Johnston의 12A형 필터[7]를 사용하였고, 부호화를 하지 않은 대역분할 및 합성시의 PSNR은 50.85dB였다. 이 때의 실험영상으로는 Popple 30프레임을 이용하여 평균을 취하였다.

대역분할된 영상을 표준안의 부호화에 입력하여 최적의 부호화를 실현하기 위해서는 전처리가 선행되어야 한다. 왜냐하면 표준안의 양자화는 일정한 범위의 값으로 양자화 하도록, 그 값들이 미리 설정되어 있으며, 전송율에 문제가 있는 경우에만 양자화 계단간격을 확장하는 방법을 이용하기 때문에 대역 분할된 영상을 그대로 부호화할 때는 고주파 영역이 손상될 우려가 있기 때문이다. 그리고 대역분할 후의 영상들은 FIR필터를 거친 후의 신호이므로 음수의 계수를 포함할 수 있지만 표준안은 하드웨어의 실용화를 고려하여 정수값만을 처리하도록 하고 있기 때문에 이러한 제한 조건에 맞출 수 있도록 적절한 정규화를 할 필요가 있다. 따라서 본 논문에서는 전처리 방법으로 다음과 같은 방법을 사용하였는데 약간의 추가정보가 발생하였으며, 원래의 대역분할 부호화의 장점인 에너지 집중도의 손실을 가져왔지만 좋은 결과를 얻을 수 있었다. 즉, 대역분할후의 기저대역은 에너지 집중도가 현저히 높은 부분으로서, 이 정보를 기준으로 하여 나머지 대역의 신호들을 정규화하고 부동소수점들은 정수처리한다. 즉,

$$S_{mn}(i, j) = S_{mn}(i, j) - T_{mn} \quad (6)$$

$$NS_{mn}(i, j) = \frac{S_{mn}(i, j)}{D_{mn}} - (D_{11})_1 \quad (7)$$

$$NS_{mn}(i, j) = RoundOff(NS_{mn}(i, j) + (D_{11})_2) \quad (8)$$

여기서, $1 \leq m, n \leq BandNumber$, i, j 는 각 화소의 좌표, $S_{mn}(i, j)$ 은 원래의 신호, $NS_{mn}(i, j)$ 은 정규화된 신호, D_{mn} 은 각 대역의 특성에 기초한 정규화 인자, 그리고 $(D_{11})_1, (D_{11})_2$ 는 기저대역의 특성에 기초한 정규화 인자이다.

정규화하기 전 원래 신호들의 기준값들은 부호화된 데이터를 복호화한 후의 원 영상을 복원하는 과정인 보간과정에서 정확한 계산을 할 수 있도록 부정보로서 각 대역 분할된 영상과 함께 부호화하여 전송한다. 그 값은 부동 소수점으로 나타나지만, 부호화의 효율을 고려하여 정수처리하면 각 대역의 영상에는 0.000032bpp(CCI601-525 4:2:0 규격의 경우)의 추가정보의 발생을 가져온다.

2. 비트열의 구성

본 논문에서 제안한 부호화기(그림 3)에 의하여 발생하는 비트열은 그림 2,5와 같은 대역 분할된 영상들의 복합적인 구조를 갖고 있으며, 각 GOP는 대역분할된 Sub-GOP들의 합으로 구성된다. Sub-GOP의 순서는 그림 3에서 알 수 있는 것처럼, 기저대역만이 모인 Sub-GOP부터 시작해서 고주파 대역순서로 부호화되어 비트열을 구성하게 된다.

각 대역별로 움직임보상 DPCM루트를 거쳐지게 되는데, 이 때의 움직임 벡터는 각각의 움직임보상 DPCM루트에서 발생한 것들을 각 대역별로 따로 이용하였다. 움직임 벡터를 찾기위한 움직임 추정의 탐색영역은 표준안의 탐색범위의 1/2인 범위에 대하여 실험을 하였으며, 대역분할된 영상에 대한 실험결과 전체 탐색 영역을 적용하는 것보다 우수하거나 비슷한 실험결과를 얻을 수 있었다. 따라서, 화질의 열화가 없기 때문에 표준안의 탐색범위의 1/2인 범위탐색영역을 적용하는 것이 부호화의 수행속도가 2배이상 됨으로써 훨씬 유리함을 알 수 있었고 여기에 하드웨어 구성시에 각 대역별로 병렬처리를 하는 경우에는 원래의 표준화안의 구조보다 더욱 빨라지리라는 것을 예측할 수 있다.

대역분할에 의하여 생기는 추가정보는 그림 5와 같은 비트열을 구성하여 처리하였다. 즉, 각 대역별로 GOP내에 Sub-GOP를 구성하였기 때문에 각 대역군을 구별하기 위한 Sub-GOP 식별용 부호가 4개 필요하게 되며, 부호의 길이는 표준안의 Start Code와 동일한 32비트의 부호를 사용했고 그 각각은 각 대역군의 제일 선단에 위치하게 되는데 이로 인한 추가정보의 발생은 0.000063bpp이다. 전송율조정은 전송한 바와 같은 비트할당 방법을 이용하여, 각 대역군별로 평균분산을 계산하여 영상열(Video Sequence)의 시작부분에서 Sub-GOP에 대한 비트할당을 한 후, 표준안 TM3의 전송율 조정방법[1]을 이용하여 각 대역별 부호화를 하였다.(그림 2)

IV. 실험 및 결과 고찰

실험용 영상으로 사용한 동영상은 국제표준 영상인 CCI601 4:2:0-525규격의 각 화소당 8비트 계조도를 갖는 Popple (720x480) 60프레임을 사용하여 9Mbps의 전송율로 실험을 하였고 각 대역에 대하여 움직임벡터를 찾기위한 움직임추정의 탐색영역을 전송한 바와 같이 표준안의 1/2로 설정하여 실험하였다. 기존의 제안된 방법[1][3]은 그림 6과 같은 DCT의 부집합을 취하는 방법을 이용하여 실험을 하였다.

복원된 영상의 화질을 평가하기 위한 지표로는 객관적 평가에 가장 널리 이용되는 철두치 신호대 잡음비(PSNR)를 이용하였는데, PSNR은 다음식으로 정의된다.

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [s(i,j) - \hat{s}(i,j)]^2} \quad [dB]$$

$s(i,j)$: original signal, $\hat{s}(i,j)$ = reconstructed signal

사진 1은 4대역 분할, 부호화, 복호화를 거쳐서 다시 대역 합성하여 얻은 하위계층(고해상도를 갖는 계층)인 scale-8의 영상인대(Popple의 24번째 프레임), 제안된 방법은 기존의 방법 보다는 다소 낮은 PSNR을 보이고 있지만(표1), 사진 1에서 알 수 있듯이 제안된 방법은 시각적으로 대체로 충실하게 복원되었음을 알 수 있다. 상위계층(저 해상도를 갖는 계층)의 성능은 대역분할된 영상인 SB11을 부호화, 복호화했을 때의 PSNR로 평가하였다. (표2)

V. 결론

본 연구에서는 동영상의 부호화에 있어서 최근에 실용화를 목표로 연구되고 있는 다중해상도를 지닌 scalable구조의 적용 사례에 대하여 연구하였다. 기존의 DCT영역의 분할방법이 아닌 대역분할 방법을 이용하였고, 4대역으로 분할하여 전처리과정을 거친뒤에 각 대역에 대한 부호화를 하여 복합적인 비트열을 구성하는 방식을 취하였다.

본 연구에 의하여 얻은 결과는 다음과 같다. 첫째, 표준안의 구조에 대역분할 부호화 방법을 적용하여 scalable구조를 실현할 수 있음을 알 수 있었으며 둘째, 대역분할 방법을 적용해도 부정보가 전체적인 영향을 거의 미치지 않음을 알 수 있었고 셋째, 대역분할된 영상에서의 움직임 벡터의 탐색영역을 반차소 탐색영역의 반으로 줄임으로서 화질의 열화없이도 처리속도를 향상시켰으며, 영상들을 각 대역별로 병렬처리할 수 있도록 하드웨어를 구성하는 경우에는 그 처리 속도가 더욱 향상될 수 있음을 추정할 수 있었다.

그러나, 대역분할에 의한 2계층의 scalable구조를 실현하는 경우 상위 계층의 화질은 매우 뛰어나지만, 각 대역을 모두 합성하여 하위계층을 구현하는 경우에는 시각적으로는 판단하기 힘들어도 불구하고 PSNR은 단일계층의 복원영상보다 약간의 열화가 있음을 알 수 있었다. 그러므로 대역분할 및 합성단의 최적의 필터를 구성하여 하위계층의 해상도를 향상시킬 수 있는 방안과 더불어, 기존의 부호화 방법과의 호환성을 유지하면서도 3계층 이상의 다중 해상도를 갖는 구조에 관한 연구가 뒤따라야 할 것이다.

<참고문헌>

- [1]ISO-IEC/JTC1/SC29/WG11, MPEG92/328, "MPEG-2 Test Model 3", Nov.1992
- [2]M.R.Civanlar, A.Puri, "Scalable Video Coding in Frequency Domain", SPIE Visual Communication & Image Processing, Boston, MA, pp.1124-1134, Nov.1992
- [3]J.Arnold, E.O'Brien & R.Brandis, "A Simple Frequency Scalable Coder for Interlaced Video Sources", IEEE Visual Signal Processing & Communications, Workshop Proceedings, Melbourne Australia, Sep. 1993
- [4]M.Vetterli, "Multi-dimensional Subband Coding : Some Theory and Algorithm", Signal Processing, vol.6, NO.2, pp. 97-112, April 1984.
- [5]J.W.Woods & S.D.O'Neil, "Subband Coding of Images" IEEE Trans. on Acoust., Speech, Signal Processing, vol. ASSP-34, No. 5, pp.1278-1288, October, 1986.
- [6]P.H.Westerink, D.E.Boekee, J.Biemond & J.W.Woods, "Subband Coding of Images Using Vector Quantization", IEEE Trans. on Communications, vol.36, NO.6, pp.713-719, June, 1988

[7]J.D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," ICASSP, pp.291-294, June, 1980

[8]Iole Moccagatta, M.Kunt, "Motion Compensated Generic Coding of Video based on a Multiresolution Data Structure", Optical Engineering, vol.32, NO.7, pp.1559-1570, July, 1993

[9]H.Paek, R.C.Kim & S.U.Lee, "On the Motion Compensated Transform Coding Technique Employing Subband Decomposition", SPIE Visual Communication & Image Processing, Boston, MA, pp.253-264, Nov. 1992

[10]N.S.Jayant & P.Noll, "Digital Coding of Waveforms" Englewood Cliffs NJ : Prentice-Hall, 1984

	대역분할			DCT		
	Y	Cb	Cr	Y	Cb	Cr
I Picture	30.68	36.72	37.73	35.62	42.22	41.97
B Picture	29.58	35.89	35.09	33.19	40.44	41.07
P Picture	30.55	37.08	37.79	34.37	39.93	39.95

표1. 하위계층의 성능, Popple, 9Mbps
대역분할방법 : 0.592bps

	Y	Cb	Cr
I Picture	40.65	42.23	42.32
B Picture	39.15	41.67	41.85
P Picture	43.03	44.09	43.74

표2. 대역분할시의
상위계층의 성능,
Popple.

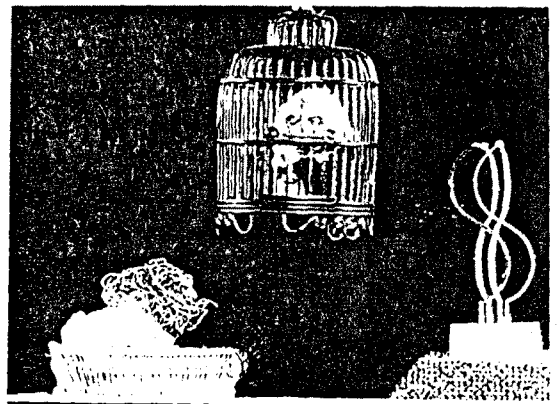


사진 1. 복원된 Popple의 하위계층의 영상

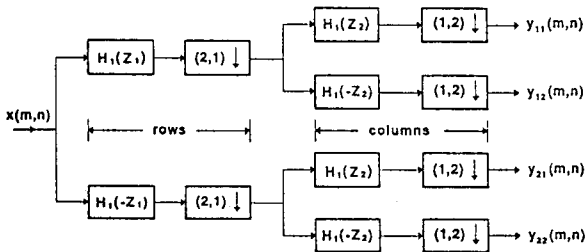
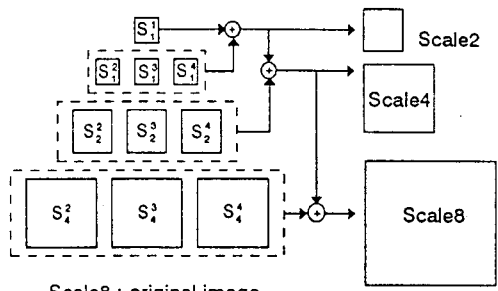


그림 1. 분리가능한 4대역 대역분할 필터트리



Scale8 : original image
Scale4 : half size
Scale2 : quarter size

그림 4. 대역분할의 피라미드형 구조

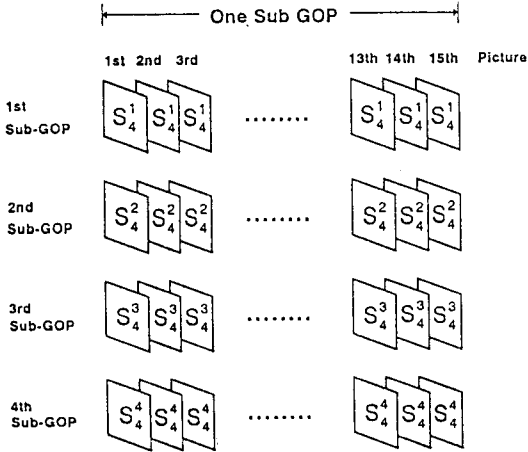


그림 2. 대역분할 GOP의 구조

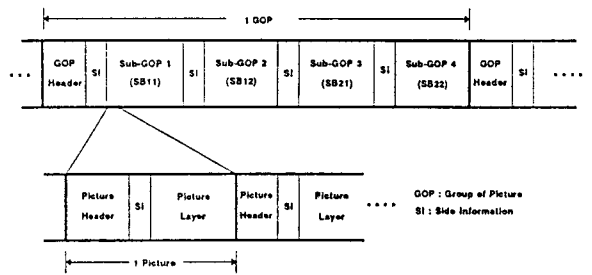


그림 5. 비트열의 구조

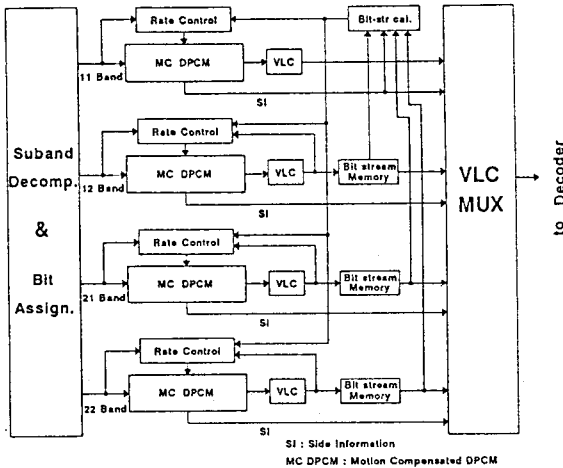


그림 3. 제한된 부호화기의 구조

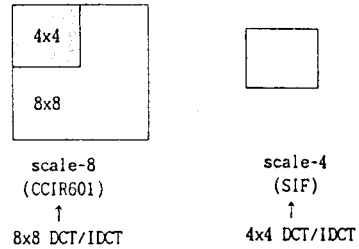


그림 6. DCT의 부집합을 취하는 방법