# 웨이브릿 변환에서 인지적 가중치를 이용한 SPIHT 비디오 부호기

# SPIHT Video Coder Using Perceptual Weight in Wavelet transform

정용재, 강경원, 문광석

Yong-Jae Jeong, Kyung-Won Kang, Kwang-Seok Moon

## 요약

동영상 부호기에서 화면내 프레임 부호화는 전체 프레임의 화질에 중요한 영향을 미친다. 표준화된 동영상의 부호기는 DCT를 쓰지만, 저 비트율에서의 블록화 현상으로 화질의 열화를 가져올 수 있다. 본 논문에서는 화질의 열화를 감소시키고 인간 시각적인 측면에서의 화질 개선을 위한 비디오 코딩을 제안한다. 제한안 방법에서는 웨이브릿 변환에서 인지적 가중치를 화면내 프레임에 적용하여 SPIHT와 VLC를 이용하여 부호화 하였고, 인간 시각 특성을 고려하여 시각적인 노이즈를 제거하여 주관적인 화질을 향상 시켰다.

## Abstract

The frame coding inside the screen for a video coder has a big influence on the quality of the whole frame. The standardized video coder uses DCT, however it can give rise to a low image quality due to the blocking effect at low bit rate. This paper proposes a video coding method for an image quality improvement in human visual aspects. With the proposed method, the perceptual weight is coded with SPIHT and VLC by applying it into the frame and the visual noises are eliminated.

**Key words** : Video compression, Wavelet transform, Perceptual Weight

## Ⅰ. 서론

최근 멀티미디어에 대한 관심이 높아지면서 관련된 멀티미디어 통신 서비스에 관한 연구가 국내외에서 활발히 진행되고 있다. 멀티미디어 정보 중에서 동영상 정보는 많은 정보량을 가지고 있어서 고압축이 필요하다. 이러한 이유로 효율적으로 동영상을 전송하기 위한 부호화 기법이 필요하게 되었다. 현재 동영상 부호기로 사용되고 있는 것에는 MPEG2, H.263등이 있다[1,2]. 이 중 H.263은 초저속 전송에서 사용되는 규약이다.

위에서 언급한 비디오 영상의 부호화 방법들은 공간영역에서의 잉여데이터를 줄이기 위해서 DCT(discete cosine Transform)를 채택하고 있는데, 공간상에서 8×8 단위의 부 블록으로 나누어 DCT를 적용하기 때문에 시간상의 잉여데이터를 줄이기 위한 움직임 보상 프레임간 예측방법을 이용하는데는 유리하나 한 프레임 내의 블록과 블록간 공간 잉여데이터는 줄이지 못하며 압축률이 높은 경우에 구획 잡음(blocking artifact)이 생기는 단점이 있다. DCT의 이러한 단점 때문에 오늘날 정지 영상 압축 분야에서는 웨이브릿 변환(wavelet transform)을 이용한 영상압축 알고리즘들이 소개되었으며 고압축에서 작은량의 데이터로도 원래 영상을 복원 할 수 있는 장점이 있다. 또한 영상 압축 알고리즘에서 평균 자승 오차는 영상간의 근사 정도, 즉 닮은 정도에 대한 판정의 기준이 되었다. 그러나 영상의 유사성 및 화질에 대한 판단은 인간에 의해서 이루어지므로 단순한 정량적인 척도인 평균 자승 오차를 왜곡 척도로 사용하는 것은 바람직하지 못하다. 따라서 본 논문에서 제안하는 인지 왜곡 척도를 이용한 SPIHT 기반 비디오 부호기는 공간 방향의 잉여-데이터를 줄이기 위해 SPIHT(set partitioning in

hierarchical trees)를 사용하고, 인간 시각 특성을 반영하기 위하여 인지적 가중치를 이용하여 부호화를 하였기 때문에 주관적인 화질이 개선됨을 제시한다.

## II. SPIHT(set partitioning in hierarchical trees)와 SPIHT 비디오 부호기

SPIHT는 웨이브릿 분해된 영상을 효율적으로 부호화하는 방법중의 하나이다. 이 방법은 EZW 부호화시 하나의 심볼에 2비트를 할당하므로 데이터 량이 많아지는 단점을 개선하였을 뿐만 아니라 비트단위로 처리할 수 있어 사용자가 원하는 비트율을 정확하게 맞출 수 있다.

EZW와의 차이는 웨이브릿 분해된 계수들을 계수의 중요도에 따라 부분 집합으로 분할하여 분할된 계수들을 부호화하는 것이다. 이는 가장 큰 계수가 가장 큰 정보량을 가지고 있다는 점에서 큰 계수를 먼저 전송함으로서 영상의 자기 상관성을 이용하여 효율적인 부호화를 수행하게 된다.

SPIHT알고리듬은 웨이브릿 계수들의 계층적구조를 이용하여 식(1)에서 $2^n$을 임계값으로 하여 분류 과정(sorting pass)과 세분화 과정(refinement pass)에 의해 부호화를 하게된다.

$$n = \lfloor \log_2(\max_{(i,j)}\{|C_{i,j}|\}) \rfloor \qquad (1)$$

여기서, $C_{i,j}$는 좌표 $(i,j)$에서의 웨이블릿 계수이다. 웨이블릿 계수들은 이 임계값보다 크면 유효 계수(significant coefficient)라 하고, 작으면 무효 계수(insignificant coefficient)라 한다. 웨이블릿 계수들은 그림 1과 같이 부밴드 사이의 공간적인 트리구조를 구성할 수 있다. 이러한 구조를 가지는 웨이브릿 계수는 일반적으로 에너지의 대부분이 저주파 부밴드로 갈수록 높아진다.

SPIHT 부호화는 *LSP*(list of significant pixel), *LIP*(list of insignificant pixel), *LIS*(list of insignificant set)라는 세 개의 리스트를 사용하여 수행한다. *LIP*는 무효계수들의 집합이고 *LIS*는 무효계수의 그룹들의 집합이며 *LSP*는 유효계수들의 집합이다. *LIS*는 다시 descendant에 따라 A-type과 B-type으로 나뉘게 된다. A-type은 어떤 계수를 중심으로 하위대역의 모든 descendant들이 무효계수일 경우이고 B-type은 바로 아래 하위대역에 존재하는 계수를 제외한 나머지 하위대역에 존재하는 계수들을 무효계수로서 그룹화하는 것이다.

부호화 알고리즘은 초기화, 분류과정, 세분화 과정과 quantization-step update의 네 단계로 이루어진다.

1) 우선 *LSP*를 비워두고 가장 상위에 있는 계수들을 *LIP*에 넣고 *LIP*의 계수들 중에서 descendant를 *LIS*에 넣는다.
2) 분류화 과정에서는 *LIP*를 순서대로 조사해 가면서 유효계수이면 *LSP*에, 무효계수이면 *LIP*에 각각에 추가시킨다. *LIS*는 descendant에 관한 것으로 A-type과 B-type으로 구분하여 수행 하게된다.
3) 세분화 과정에서는 *LSP*의 정보를 양자화하는 단계이다.
4) 임계값을 2로 낮추어 원하는 비트율에 도달 할 때까지 단계 2)를 반복 수행하게 된다.

위와 같은 단계를 반복적으로 수행하게되어 계수들을 크기 순서로 정렬하게 되고 이렇게 정렬된 계수들을 코딩하게 된다.

또한 비트 단위로 처리되는 과정이므로 부호화 및 복호화에서 사용자가 원하는 비트율을 정확하게 맞출 수 있는 임베디드 특성이 나타나게 된다.

현재 표준화된 동영상 압축 부호화는 기본 알고리즘으로 DCT를 사용하고 있다.

DCT의 단점은 정규화된 블록단위의 연산이 수행되므로 양자화시 발생하는 블록간 이웃되는 픽셀들간의 불일치, 블록단위로 이뤄지는 움짐임 예상과 움직임 보상에서 이웃된 블록들간에 생기는 불일치, 그리고 시스템 버퍼내의 참조 영상에서 일어나는 블록간의 불일치로 인한 블로킹 현상이 발생한다

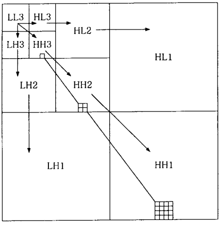
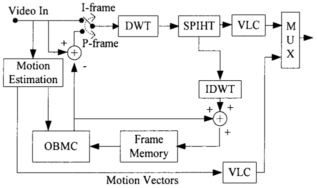



그림 3. 웨이브릿 계수의 대응관계

Fig. 3. The corresponding relationship of wavelet coefficients.

특히 고압축일 경우 블록화 현상이 심하게 나타나는 단점이 있다. 이러한 단점을 극복할 뿐만 아니라 스케이러블한 전송 특성을 얻기 위해 웨이브릿 변환 영역에서 SPIHT를 이용한 동영상 부호화 방법이 있다.[5,6] SPIHT비디오 부호기는 화면내(intra) "I" 픽처와 화면간(inter) "P" 픽처로 구성되어 부호화가 되며, 변환된 픽처는 SPIHT와 VLC(variable length coding)에 의해 부호화되고, 나머지 남아있는 프레임인 화면간 픽처는 H.263의

OBMC(overlapped block motion compensation)에 의해 생성된 보상영상을 SPIHT에 의해 부호화한다. SPIHT는 목표로하는 비트율에서 정지영상 부호화시 가장 좋은 성능을 보이는 부호기중 하나로 인식되어져 있다. 알고리듬의 임베디드 특성은 목표로하는 비트율에서 정확히 프래임을 부호화하는 것이 가능하도록한다. 그림 2는 SPIHT를 이용한 동영상 부호화의 블록도를 나타낸다.



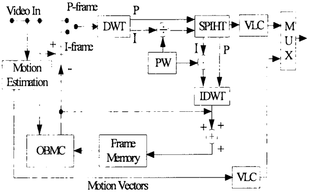

| | |
|---|---|
| DWT | Discrete Wavelet Transform |
| IDWT | InverseDiscrete Wavelet Transform |
| SPIHT | Set partitioning in hierachical trees |
| MC | Motion Compensation |
| VLC | Variable Length Coder |
| Out | Out bits |
| MV | Motion vector |

그림 4. SPIHT 비디오 부호기
Fig. 4. SPIHT Video Coder

## III. 제안한 웨이브릿 변환에서 인지적 가중치를 이용한 SPIHT 비디오 부호기

SPIHT 비디오 부호기는 H.263 부호기와 같이 처음 한번 화면내(Intra)부호화를 시작하고 전체 동영상의 부호화가 끝날 때까지의 나머지 프레임은 화면간(Inter)부호화로 진행된다.

제안한 웨이브릿 변환에서 인지적 가중인자를 이용한 SPIHT 비디오 부호기는 시작 프레임이 부호기로 입력이 되면 "I" 픽쳐로 인식하고 웨이브릿 변환을 하게 된다. 웨이브릿 변환된 프레임은 인지적 가중치를 통하여 각각의 웨이브릿 부대역에 인지적 가중치가 적용된다[7]. 다음 단계에서 SPIHT를 수행하게 된다. SPIHT를 수행하여 양자화 과정을 거친 프레임은 VLC과정을 통하여 비트스트림화 되어 압축형식으로 출력되고 또한 SPIHT를 통하여 양자화된 프레임은 두 번째 들어올 원영상과의 비교를 통한 차영상을 만들기 위해서 역 SPIHT 과정을 진행한 다음 역 인지적 가중치를 적용한 후 역 웨이브릿 변환을 통하여 영상을 복원하여 메모리에 저장하게 된다. 메모리에 저장된 영상과 다음에 들어오게되는 영상을 ME(motion estimation)에서 MV(moving vector)를 만들고 MC(motion compensation)를 이용하여 복원된 영상을 만들게 된다. 복원된 영상은 다시 두 번째 영상과 차를 구해서 차 영상으로 만들어 MV는 VLC하여 복호기로 보내고 차영상은 "P" 프레임인 화면간 프레임으로 DWT를 통하여 변환된다. 영상의 정보가 거의 없는 차영상이므로 인지적 가중치는 적용되지 않고 단지 SPIHT를 통하여 양자화되고 VLC과정을 통하여 비트스트림화되어 압축된 형식으로 출력된다. 이후, 화면간 프레임의 부호화 과정은 비디오 부호가 끝날 때까지 계속 반복되어진다. 제안한 웨이브릿 변환에서 인지적 가중인자를 이용한 SPIHT 비디오 부호기는 그림 5에 나타내었다.




| | |
|---|---|
| DWT | Discrete Wavelet Transform |
| IDWT | Inverse Discrete Wavelet Transform |
| SPIHT | Set partitioning in hierachical trees |
| MC | Motion Compensation |
| VLC | Variable Length Coder |
| Out | Out bits |
| MV | Motion vector |
| PW | Perceptual-weights |

그림 5. 제안한 SPIHT 비디오 부호기
Fig. 5 Proposed SPIHT VIdeo Coder

화면내 부호화는 전체 동영상 부호화에서 단 한번만 수행하고 또한 화면간 부호화은 화면내 프레임의 복원 프레임의 차 성분으로 비디오 부호화가 되기 때문에 화면내 프레임의 부호화는 동영상 전체에서 아주 중요하다. 기존의 방법에서의 인지적으로 중요한 계수 값들의 coarse한 양자화과정으로 인해 인지적 왜곡을 제거하지 못하여 주관적 화질의 떨어짐이 있었다. 본 논문에서 제안하는 인지적 웨이트의 적용은 화면내 프레임의 웨이브릿 변환영역에서 부대역 마다 인지적 가중치를 적용하여 인지적 잉여성분을 제거하고, 영상 전체에서 인지적 왜곡

을 제거하여 비트를 보존할 수 있다.

## V. 실험 및 고찰

본 논문에서 제안한 웨이브릿 변환에서 인지적 가중치를 이용한 SPIHT 비디오 부호기의 성능 평가를위해 H.263표준 영상 포맷인 cif(352 × 288)크기의 flower 와 foreman영상을 이용하여 실험하였다. 웨이브릿 필터는 쌍직교 9/7 필터를 사용하여 영상을 3레벨 분할하였다 . 화면내 프레임은 인지적 가중치를 적용하였고, 화면간 프레임은 인지적 가중치를 포함하지 않고 부호화 하였다.

기존의 방법과의 성능 평가를 위해 flower, foreman 영상의 각 10 프레임을 대상으로 부호화 하여 화면내 프레임의 PSNR과 화면간 프레임의 평균 PSNR을 비교한다. 제안한 방법에서 foreman 영상은 동일한 비트량을 가지는 화면내 프레임에서 foreman 영상의 제안한 방법은 H.263보다는 1.27dB가 향상되었고, SPIHT보다는 1.93dB가 떨어졌다. 하지만 SPHIT에서 양자화 과정은 coarse하기 때문에 인지적 잉여 성분들이 제거하지 못하게 된다. 화면 전체의 인지적 잉여들을 제거하기 위해서 전처리과정 없이 인지적 가중치를 이용하여 영상 전체의 인지적 왜곡을 제거하여 주관적인 화질의 개선됨을 알 수 있다. 화면간 프레임은 H.263보다는 프레임 평균 PSNR이 1.61dB 향상 되었고, SPIHT보다는 2.03dB떨어졌다.

표1. 화면내 프레임의 PSNR 결과

Table1. PSNR result of Intra frame

| Input Sequence | Size | H.263 (dB) | SPIHT (dB) | 제안한 방법 (dB) |
|---|---|---|---|---|
| foreman | 8kbyte | 34.68 | 37.88 | 35.95 |
| flower | 25kbyte | 31.45 | 33.92 | 32.56 |

표2. 화면간 프레임의 평균 PSNR 결과

Table2. Average PSNR result of Inter frames

| Input Sequence | bit rate | H.263 (dB) | SPIHT (dB) | 제안한 방법 (dB) |
|---|---|---|---|---|
| foreman | 56kbps | 33.82 | 37.46 | 35.43 |
| flower | 56kbps | 29.23 | 33.63 | 32.12 |

그림 7은 10프레임 동안의 PSNR을 그래프로 나타낸 것이다. 그림에서 보는바와 같이 화면내 프레임이 전체 프레임에 큰 비중을 차지함을 알 수 있다.

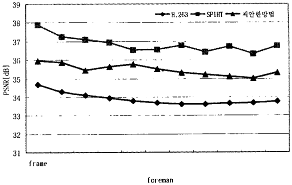



그림 7. 동영상 부호기들의 PSNR비교(전체 10프레임)

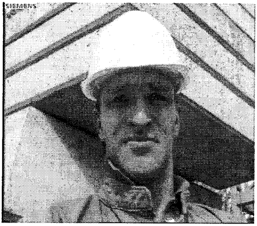
Fig 7. Compare of PSNR of Video coder(Total 10 frame)

그림6 부호기들의 화면내 프레임을 나타낸 것이다. foreman 영상의 (a)는 원영상이고, (b)는 H.263에서의 화면내 프레임, (c)는 SPIHT 비디오 부호기에서의 화면내 프레임 그리고, (d)는 제안한 방법에서의 화면내 프레임이다. 파일의 크기는 8Kbyte이고 보는것과 같이 H.263에서는 원영상과 비교했을 때 화질의 열화가 있음을 알 수 있다. SPIHT 비디오 부호기에서의 화면내 프레임은 원영상과 거의 같고 제안한 방법또한 비록 PSNR은 떨어졌지만 주관적인 화질에서는 원영상과 거의 비슷함을 알 수 있다.

## VI. 결론

본 논문에서는 기존의 H.263과 SPIHT 비디오 부호기의 주관적 화질을 향상시키기 위해서  인지적 가중치를 이용한 SPIHT비디오 부호기를 제안하였다. 제안한 방법은 화면내 프레임의 웨이브릿 변환영역에서 인지적 가중치를 적용하여 화면내 프레임에서의 인지 잉여를 제거함으로써 주관적 화질을 기존의 두 비디오 부호기(H.263, SPIHT)보다 높여서 화면간 프레임의 주관적 화질 또한 높아지게 만들어 동일한 비트율에서 전체 동영상의 주관적 화질을 개선하였다. 향후 화면간 프레임에서의 인지적 가중치에 관한 연구가 된다면 좀더 확실하게 전체 영상의 주관적 화질을 개선시킬 수 있을 것으로 사료된다.

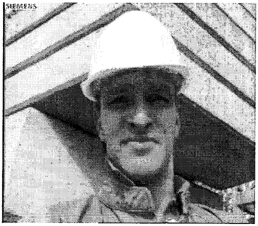
접수일자 : 2001. 12. 4　　수정완료 : 2002. 1. 23



(a) Original

(b) H.263

(c) SPIHT



(d) 제안한 방법

그림 6. 비디오 부호기에서의 화면내 프레임 : (a)원영상, (b)H.263, (c)SPIHT, (d)제안한 방법

Fig 6. Intra frame at video coder : (a) Original image, (b)H.263, (c)SPIHT, (d)Proposed methode

## [참고문헌]

[1] Draft ITU-T Recommandation H.263, "Video coding for low bitrate Communication," Dec, 1995.

[2] ISO-IEC/JTC1/SC29/WG11, MPEG93/457, Coded Representation of Picture and Audio Information, Test Model 5, April 1993

[3] Shapiro, J.M. "Embedded image coding using zerotrees of wavelet coefficients," Signal Processing, IEEE Transactions on , vol. 41, No.12 , pp. 3445 -3462, Dec. 1993

[4] Said, A. Pearlman, W.A. "A new, fast, and efficient image codec based on set partitioning in hierarchical trees." Circuits and Systems for Video Technology, IEEE Transactions on , vol. 6 No. 3 , pp. 243 -250, June 1996

[5] Karlekar, J. Desai, U.B. "SPIHT video coder." TENCON '98. 1998 IEEE Region 10 International Conference on Global Connectivity in Energy, Computer, Communication and Control , vol. 1 , pp. 45 -48, 1998

[6] Martucci, S.A Sodagar, I. Chiang, T. Ya-Qin Zhang "A zerotree wavelet video coder." Circuits and Systems for Video Technology, IEEE Transactions on , vol. 7 No. 1 , pp. 109 -118, Feb. 1997

[7] Hontsch, I. Karam, L.J. Safranek, R.J. "A perceptually tuned embedded zerotree image coder" Image Processing, 1997. Proceedings., International Conference on , vol. 1 , pp. 41 -44 , 1997

**정용재**(Yong Jae Jeong)
準會員
1999년 부경대학교 전자공학과 (공학사)
2001년-현재 부경대학교 대학원 전자공학과 석사 과정 재학중
관심분야 : 영상신호처리, 동영상압축

**강경원**(Kyung Won Kang)
準會員
1996년 부경대학교 전자공학과 (공학사)
1998년 부경대학교 대학원 전자공학과 (공학석사)
1998년~현재 부경대학교 대학원 박사 과정 재학중
1998년 3월~1999년 8월, 미국 Jackson State University 연구원
관심분야 : 영상신호처리, 멀티미디어 통신

**문광석**(Kwang Seok Moon)
正會員
1979년 2월 경북대학교 전자공학과 (공학사)
1981년 2월 경북대학교 대학원 전자공학과(공학석사)
1989년 2월 경북대학교 대학원 전자공학과(공학박사)
1988년 1월~12월 일본 동경대학교 공학부 연구원
1997년 8월~1998년 7월 미국 Jackson State University 객원교수
1990년~현재 부경대학교 전자공학과 교수
관심분야 : 영상신호처리, 적응신호처리, 멀티미디어 통신 등.