

특집논문-07-12-6-06

## 스케일러블 비디오 부호화에서 선택적 계층간 차분 신호 부호화 및 공간적 향상 계층에서의 모드 결정

이범식<sup>a)</sup>, 김문철<sup>a)\*</sup>, 함상진<sup>b)</sup>, 박창섭<sup>b)</sup>, 박근수<sup>b)</sup>

### Selective Inter-layer Residual Prediction Coding and Fast Mode Decision for Spatial Enhancement Layers in Scalable Video Coding

Bumshik Lee<sup>a)</sup>, Munchurl Kim<sup>a)</sup>, Sangjin Hahm<sup>b)</sup>, Changseob Park<sup>b)</sup>, Keunsoo Park<sup>b)</sup>

#### 요 약

본 논문에서는 스케일러블 비디오 부호화의 복잡도 감소를 위해, 공간 계층간 잔차 신호 부호화를 선택적으로 수행하고 그 정보를 바탕으로 향상 계층에서 모드를 고속으로 결정하는 방법에 대해 소개한다. 계층간 잔차 신호 부호화는 스케일러블 비디오 부호화에서 공간 계층간 상관도를 이용하여 계층간의 두 잔차 신호에 대한 차 신호를 부호화하는 방법으로서 부호화 효율을 증대할 수 있는 장점이 있으나 향상 계층의 모든 화면간 모드에 대해 윌-왜곡 비용을 계산하기 때문에 부호화의 복잡도를 크게 증가시키는 주요 요인이 된다. 제안 알고리즘은 하위 계층으로부터 업샘플된 잔차 신호와, 하위 계층의 움직임 벡터, 참조 화면 정보를 이용하여 향상 계층에서 얻은 잔차 신호의 SAD값을 통해 정수 변환 계수의 특징을 미리 판별하여 계층간 잔차 신호 부호화를 선택적으로 수행하는 방법이다. 제안 알고리즘에서는 계층간 잔차 영상의 차 신호에 대한 SAD값에 양자화 계수와 시간 계위에 따라 문턱치 값을 적응적으로 적용시키면서 SAD가 적응적으로 설정된 문턱치 보다 작을때는 잔차 블록에 대한 정수변환 계수가 매우 작다고 판단하여 그 블록에 대해서는 선택적으로 계층간 잔차 부호화를 수행하고 역시 공간적 향상 계층에서 16x16블록에 대해서만 윌-왜곡 최적화를 수행한다. 따라서 계층간 차분 신호에 대한 SAD값 만으로 정수변환계수의 특성을 양자화 계수와 시간 계위에 따라 분류하여 고속의 부호화를 달성함으로써 SVC 부호화기의 복잡도 및 부호화 시간을 크게 감소 시켰다. 제안 알고리즘을 적용하면 다양한 특성을 갖는 영상에 대하여 부호화 시간을 원래의 SVC 참조 소프트웨어 대비 평균 51.5%의 부호화 속도를 향상 하였음에도 이에 따른 PSNR의 감소는 평균 0.03dB, 비트율의 증가는 0.64%로 무시할 수 있을 정도로 작았다.

#### Abstract

In order to reduce the complexity of SVC encoding, we introduce a fast mode decision method in the enhancement layers of spatial scalability by selectively performing the inter-layer residual prediction of SVC. The Inter-layer residual prediction coding in Scalable Video Coding has a large advantage of enhancing the coding efficiency since it utilizes the correlation between two residuals from a lower spatial layer and its next higher spatial layer. However, this entails the dramatical increase in the complexity of SVC encoders. The proposed method is to analyze the characteristics of integer transform coefficients for the subtracted signal for two residuals from lower and upper spatial layers. Then it selectively performs the inter-layer residual prediction coding and rate-distortion optimizations in the upper spatial enhancement layer if the SAD values of residuals exceed adaptive threshold values. Therefore, by classifying the residuals according to the properties of integer-transform coefficients only with SAD of residuals between two layers, the SVC encoder can perform the inter-layer residual coding selectively, thus significantly reducing the total required encoding time. The proposed method results in reduction of the total encoding time with 51.5% in average while maintaining the RD performance with negligible amounts of quality degradation.

Key words : Scalable Video Coding, H.264, MPEG-4 AVC, inter-layer residual prediction, integer transform

## 1. 서 론

스케일러블 비디오 부호화(SVC, Scalable Video Coding)<sup>[1]</sup>는 H.264/AVC의 세가지 영역에서 계위성(Scalability)을 제공하는 확장 표준(Scalable Extension)으로서 시간, 공간 및 화질 계위성을 제공한다. 기존의 비디오 압축 표준인 MPEG-2 Video<sup>[2]</sup>, MPEG-4 Visual (Part 2)<sup>[3]</sup>, H.264/AVC<sup>[4]</sup>등은 하나의 계층으로 부호화하기 때문에 서로 다른 전송용량 및 에러 발생률을 갖는 이중 네트워크 환경에 적합하지 못한 단점이 있다<sup>[5]</sup>. 최근 다양한 멀티미디어 단말과 전송망이 혼재된 이중 멀티미디어 환경에 대응하여 고 압축 효율의 성능을 유지 하면서 압축 비트스트림을 유연하게 표현할 수 있도록 하여 다양한 단말과 가변적인 전송 대역에 적응적으로 데이터 콘텐츠를 전송하기 위해 스케일러블 비디오 부호화 (SVC: Scalable Video Coding)에 대한 표준화 작업이 진행되어 왔다. SVC는 이러한 멀티미디어의 환경에 맞게 네트워크 환경이나 단말 성능에 따라 시간, 공간 및 화질의 관점에서 계층적으로 부호화함으로써 한번의 부호화만으로 서로 다른 환경의 네트워크로의 전송 또는 다양한 단말기에 복호화할 수 있는 최신 동영상 압축 표준이다. SVC는 MPEG (Moving Picture Expert Group)과 VCEG (Video Coding Expert Group)에 의해 구성된JVT (Joint Video Team)에 의해 현재 표준화가 거의 완료된 단계에 있다. SVC는 하나의 기본계층(Base Layer)과 상위에 향상계층(Enhancement Layer)을 구성하여 시간, 공간 및 화질의 계위성 갖도록 부호화 할 수 있다. SVC의 기본계층은 H.264/AVC와 상호 호환적으로 부호화할 수 있으며 향상계층은 기본계층 또는 향상계층의 하위 계층에 대해 계층적으로 부호화할 수 있다<sup>[6]</sup>. 특히 SVC는 공간적 계위성을 지원하고 H.264/AVC 이상의 압축 성능을 제공하기 위하여 공간적인 상관성을 이용한 계층간 예측 기술을 이용한다. 그것을 계층간 텍스처 예측(Inter-layer Texture Prediction), 계층간 움직임 벡터 예측(Inter-layer Motion

Prediction) 그리고 계층간 잔차 신호 예측(Inter-layer Residual Prediction)이라고 한다. 이 세가지 예측 기술은 공간적 계층간의 상관도를 이용하여 SVC의 압축 효율을 높이는 주된 기술 요소 들이다. 또한 SVC는 시간 계위성을 지원하기 위하여 계층적 B-화면(Hierarchical B-picture) 구조를 이용한다. 계층적 B-화면 기술은 다양한 시간 해상도를 지원하기 위한 시간 분해 구조로 주화면(key picture)사이에 B-화면을 삽입함으로써 시간 계위성을 얻을 수 있다<sup>[7]</sup>. SVC에서 화질적 계위성은 FGS(Fine Granular Scalability)와 CGS(Coarse Granular Scalability)를 이용하며, 화질 계위의 기본 계층을 부호화 한 후 이 기본계층을 다시 복원하여 잔차 신호를 구한 후 잔차 신호를 다시 정수 변환하여 기본 계층보다 작은 양자화 파라미터(QP)값을 적용하여 부호화하는 방식이다. 이 세가지 계위성은 모두 계층 구조를 이루고 있으며 NAL(Network Abstraction Layer) 번호로 구별할 수 있도록 설계되어 있어, 이중 환경에 맞게 해당 NAL을 삽입하거나 제거함으로써 적응적으로 NAL 데이터를 처리 할 수 있다.

그러나 이 모든 혼합된 계위성을 지원함으로써 부호화 효율 및 화질은 H.264/AVC의 단일 계층을이용한 부호화에 비하여 두드러 지게 증가하였지만, 그에 따른 부호화의 복잡도 역시 크게 증가하였다. 따라서 SVC를 개발 목적에 맞게 효율적으로 이용하기 위해서는 부호화의 복잡도 감소에 대한 연구가 이루어져야 한다. 본 논문에서는 SVC의 공간적 예측 기술인 계층간 잔차 신호 예측을 시간적 계위성에 사용된 계층적 B-화면의 특성과 비디오의 특성을 이용하여 선택적으로 수행하고 그 결과를 이용하여 공간적 향상 계층에서 블록모드에 대한 윌-왜곡 최적화(Rate-Distortion Optimization, RDO)을 선택적으로 수행함으로써 부호화 시간 및 부호화의 복잡도를 감소시킬 수 있는 방법을 제시한다.

SVC의 기본 계층은 H.264/AVC와 상호 호환적으로 부호화 할 수 있도록 설계되어 있다. 2003년에 H.264/AVC의 표준화가 완료된 이후에 H.264/AVC의 부호화기 복잡도 감소에 대한 연구가 많이 진행되어 왔으며 또한 많은 연구 결과들이 발표되었다. 그러나 SVC는 Phase 1[4]에 대한 표준화 작업이 완료되지 않았기 때문에 SVC

a) 한국정보통신대학교(ICU)  
Information and Communications University  
b) 한국방송공사(KBS)  
Korea Broadcasting System  
✉ 교신저자 : 김문철(mkim@icu.ac.kr)

부호화기의 복잡도 개선에 대한 연구가 활발하게 진행되지 않고 있다. 표준화가 진행되면서 부호화 효율을 증가 시키는데 보다 많은 연구가 진행되어 왔기 때문이다. 특히 공간적 계위를 추가 할수록 윗-왜곡 최적화를 수행해야 할 블록 모드의 수가 급격히 많아 지므로 이에 대한 부호화 복잡도 및 시간 감소에 대한 연구가 요구된다. He Li et al은 공간적인 계위성에 대한 복잡도 개선을 위해 향상 계층에서 화면간 모드 결정 시 하위 계층에 대한 모드를 판별하여 향상 계층에서 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4 등 총 7개 모드 중 후보 모드를 줄여 나감으로써 고속 블록 모드 결정을 할 수 있는 알고리즘을 제안하였다<sup>[8]</sup>. 예를 들어, 기본 계층에서 얻어진 블록모드를 공간 계층간 예측을 거쳐 얻어진 모드가 16x8이면 향상 계층에서의 동일한 공간 영역에서는 16x8과 8x8만을 후보 모드로 선택하여 움직임 벡터 예측 및 모드 결정을 수행하여 부호화 시간을 단축하였다. 그러나 He Li의 방법은 기본 계층으로부터 예측된 모드가 16x16 모드가 아닐때는 향상 계층의 윗-왜곡 최적화를 수행해야 하는 후보모드의 수가 급격히 감소하여 고속의 부호화를 수행할 수 있지만 16x16 모드일 경우에는 모든 모드에 대하여 윗-왜곡 최적화를 수행해야 하므로 움직임이 적거나 균일한 영상 영역을 많이 포함하는 영상에 대해서는 효율적이지 못한 단점이 있다. 또한 He Li et al의 방법은 SVC 표준화 초기에 시간적 계위를 지원하기 위해 고려되었던 MCTF(Motion Compensated Temporal Filtering)<sup>[9, 21]</sup> 기법을 사용하였다. 그러나 MCTF를 이용한 시간적 계위는 현재 부호화의 필수사항이 아닌 선택사항으로만 남아있다. 따라서 현재 시간적 계위를 위해 채택된 계층적 B-화면 구조의 특성을 반영하지 못하므로 효율적이지 못하다.

본 논문에서는 공간적 기본 계층 및 향상 계층에서 시간적 계위로 계층적 B-화면 구조를 가지고 있을 경우, 기본 계층에서 얻어진 움직임 벡터와 참조화면 리스트를 이용하여 움직임 보상된 매크로 블록과 향상 계층의 현재 블록간의 잔차 신호를 이용하여 문턱값을 양자화 계수 QP와 계층적 B-화면 구조에서 시간적 계위에 따라 적응적으로 변화시켜 주어진 조건을 만족할 경우 계층간 잔차 부호화를 선택적으로 수행하며, 해당 결과를 이용하여 향상 계층에서도 블록모드를 초기에 결정함으로써 SVC 부호화기의 복잡

도를 감소시키고 부호화 시간을 단축하는 방법을 소개한다.

본 논문은 다음과 같이 구성된다. 2장에서는 SVC의 공간적 향상 계층에서의 모드 결정방법과 계층간 잔차 신호 부호화, 그리고 각 공간 영역에서의 시간적 해상도를 지원하는 계층적 B-화면에 대해서 설명한다. 3장에서는 선택적으로 계층간 잔차 신호 부호화를 위한 결정 조건에 대해서 알아본다. 4장에서는 선택적 잔차 신호 부호화와 공간적 향상 계층에서의 초기 모드 결정을 통한 고속 모드 결정 방법에 대한 제안 알고리즘을 소개한다. 5장에서 본 제안 알고리즘을 이용한 실험결과를 제시하고 6장에서 결론으로 본 논문을 마친다.

## II. SVC 부호화기의 시간적 계위 및 공간적 잔차 신호 부호화

### 1. SVC에서의 계층간 예측 및 잔차 신호 부호화

SVC에서는 공간적 계위성을 지원하기 위하여 세 가지 정보를 하위 계층으로부터 예측하여 차상위 계층에서 예측 신호로 사용한다. 먼저 하위 계층의 매크로 블록이 화면내 예측으로 부호화 되었을 경우, 차 상위 계층의 동일 공간 영역(co-located block)은 하위 계층의 복원 영상 신호를 업샘플링하여 현재 상위 계층의 예측 신호로 사용한다. 이때 이러한 예측 모드를 Intra\_BL 모드<sup>[11]</sup>라고 하고 복원된 텍스처를 업샘플링할 때는 밝기 화소에 대해서는 Cubic Spline 4-tap 필터<sup>[10]</sup>를 적용하고 컬러 화소에 대해서는 2-tap bi-linear 필터를 사용하여 업샘플링에 수반하는 복잡도를 최소화하도록 한다. Intra\_BL 모드는 하위 계층의 동일한 공간 영역이 화면내 예측일때만 적용될 수 있으며 이는 기본 계층이 화면내 예측이 아닐 경우, 움직임 보상 및 복호화를 통해 공간적 텍스처 예측을 수행해야 하기 때문에 부호화기의 복잡도가 크게 증가하게 된다. 이를 Single-loop 복호화<sup>[11]</sup> 텍스처 예측이라고 한다. 두번째로 하위 계층의 대응 블록이 화면간 예측으로 부호화 되었다면 상위 계층의 블록은 하위 계층의 움직임 벡터와 참조화면 리스트를 이용하여 상위 계층에서 움직임 보상을 수행할 수 있다. 만약

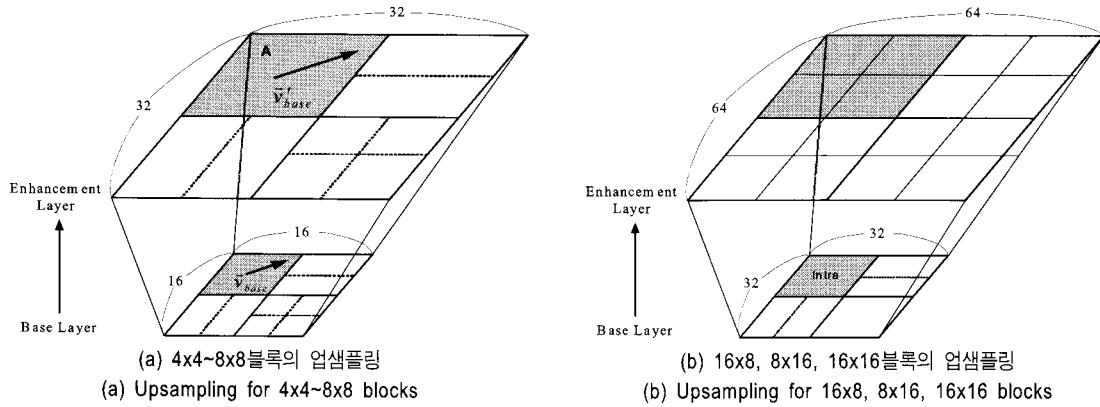


그림 1 업샘플링한 블록과 움직임 벡터  
Fig. 1. The upsampled blocks and motion vectors

이 움직임 벡터와 참조화면 리스트가 최적으로 판단되면 motion\_prediction\_flag\_10 또는 motion\_prediction\_flag\_11 이 매크로 블록 헤더에 부호화되며 상위 계층의 현재 블록은 매크로 블록 헤더에 움직임 벡터, 참조화면 리스트, 매크로 블록의 모드 정보등을 부호화하지 않아도 되기 때문에 부호화 비트를 절약할 수 있는 장점이 있다.

그림 1은 계층간 움직임 정보 예측시 매크로 블록의 분할 정보와 움직임 벡터가 업샘플링되는 것을 나타낸다. (a)는 하위 블록이 4x4~8x8까지 서브 매크로블록으로 분할되었을 때 서브 매크로블록들은 상위 계층에 16x8, 8x16, 16x16 블록으로 각각 매핑되고 만일 하위 계층이 16x8, 8x16, 16x16일때는 모두 16x16 모드로 상위 계층에 매핑된다. 이때 그림 1-(a)의 블록 A처럼 하위 계층에서 상위 계층으로 매핑되는 매크로 블록 모드를  $Mode_{BL\_PRED}$  로 정의한다. 한편  $Mode_{BL\_PRED}$  모드 내의 움직임 벡터는 하위 계층의 움직임 벡터에 두배를 곱하여 사용되며 식 (1)과 같이 표현된다.

$$\vec{v}'_{base} = 2 \cdot \vec{v}_{base} \quad (1)$$

여기서  $\vec{v}_{base}$  는 하위 계층에서 움직임 벡터 예측을 통하여 얻은 움직임 벡터고 1/4 화소 정밀도를 갖는다. 이것을 공간계층간 움직임 벡터 예측을 통하여 2배로 확대한 움직임 벡터  $2\vec{v}_{base}$  를 상위 계층의 대응블록 모드에서 재 사용

이 가능하다. 그러나 1/4 정밀도의 움직임 벡터를 2배로 확대하였기 때문에 벡터의 정밀도는 반화소로 감소한다.

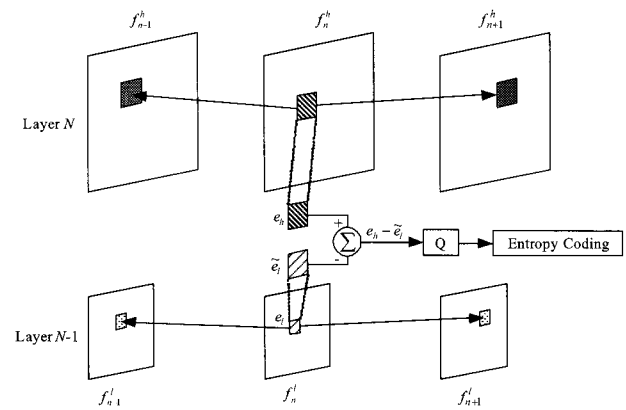


그림 2. 계층간 잔차 신호 예측 및 부호화  
Fig 2. Inter-layer Residual Prediction and Coding

마지막으로 그림 2는 계층간 잔차 신호 예측을 통해 공간 계위 향상 계층의 블록을 부호화하는 과정을 나타낸다. 이를 계층간 잔차 신호 예측(Inter-layer Residual Prediction) 이라고 한다. 현재 계층이 화면간 예측 모드로 부호화 할 때 현재 계층의 매크로 블록의 움직임 정보가 하위 계층의 대응 블록의 움직임 정보와 유사하다면 두 계층간의 잔차 신호에 대한 차분 신호를 부호화하는 것이 부호화 효율 측면에서 유리 할 수 있다. 그림2에서  $e_h$  와  $e_l$  은 각각 향상 계층과 기본 계층에서의 잔차 신호를 의미한다.  $\tilde{e}_l$  는 하위

계층의 블록크기를 향상 계층의 블록크기로 bi-linear 필터를 이용하여 업샘플링한 신호이다. 즉 현재 계층의 블록과 하위 계층의 공간적 대응 블록의 움직임 정보가 유사하다면 두 공간 사이의 잔차 신호인  $e_h - \tilde{e}_l$ 가 작은 값이 되어 부호화 효율 측면에서 작은 비트율로 부호화 할 수 있으므로 매우 유리하다. 한편 계층간 잔차 신호 부호화는 residual\_prediction\_flag를 매크로 블록 헤더에 부호화하여 복호화기에 전송한다.

### 2. SVC에서의 각 공간 계층 내에서의 시간적 계위성 (Temporal Scalability)

그림 3은 각 공간 계위상에서 시간 계위를 구성하기 위한 계층적 B-화면의 구조를 보여준다. 계층적 B-화면 기술은 다양한 시간 해상도를 지원하기 위한 시간 분해 구조이며 시간 계위에 따른 분해 레벨을 갖는다. 그림 3은 GOP의 크기가 8인 계층적 B-화면의 구조이다. 이 그림에서 시간적 거리가 가장 먼 I와 P가 주화면을 구성하며 시간 계위는 0이 된다. 이 주화면이 부호화가 완료되면 이 두화면을 참조

하여  $B_1$ 이 부호화되며 시간 계위는 1이된다. 그 다음으로 시간 계위 2를 구성하는  $B_2$ 가 부호화되며 마지막으로  $B_3$ 가 부호화 된다. 이 구조는 부호화 순서가 계층적인 피라미드 구조로 이루어 지기 때문에 계층적 B-화면 구조라고 말할 수 있다. 여기서 총 구성할 수 있는 시간적 계위는 GOP의 크기로 결정되는데, 만일 GOP의 크기가  $N(N=2^n)$ 이면 총  $(n+1)$ 개의 시간 계위 레벨의 수를 구성할 수 있다. 그림3에서는 GOP의 크기가  $8(=2^3)$ 이므로 총 4가지의 시간 계위를 구성할 수 있다. 한편 계층적 B-화면 구조에서는 주화면을 구성하는 I 및 P화면은 GOP내의 모든 화면의 참조 화면이 될 수 있으므로 가장 좋은 화질을 유지해야 하고 시간적 계위의 레벨이 증가 할 수록 참조될 수 있는 화면의 수가 적어지므로 중요도가 감소하게 된다<sup>[19]</sup>. 이러한 이유로 각 시간 계위별로 서로 다른 QP 값을 적용하여 한 GOP 내의 비트량을 줄일 수 있다. 즉, 주화면을 구성하는 시간 계위 0에 대해서는 작은 QP값을 적용하여 높은 화질을 유지할 수 있도록 하고 시간 계위 레벨이 증가할수록 QP 역시 높은 값을 적용하여 시간 계위에 따른 화질의 중요도로

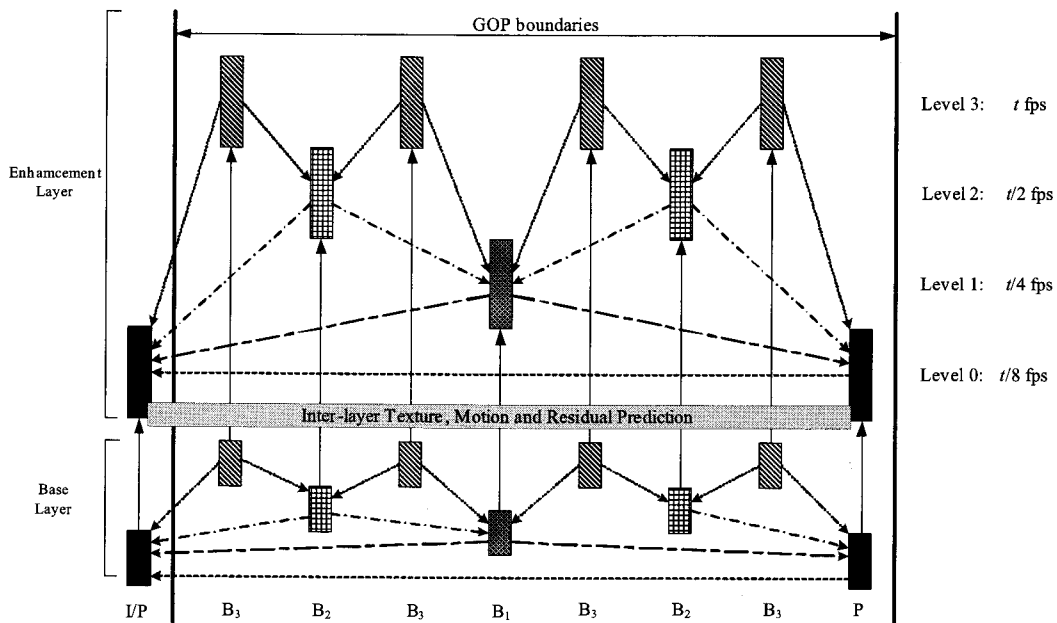
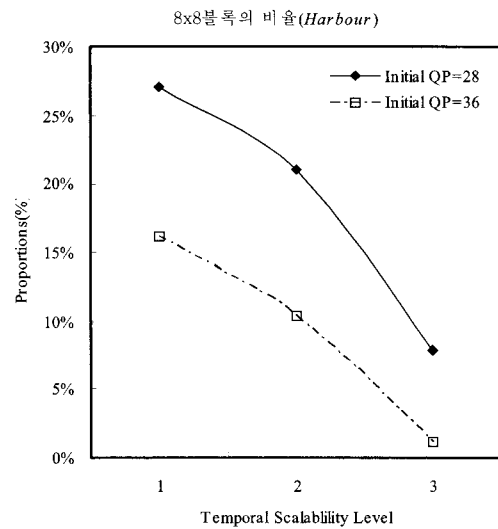
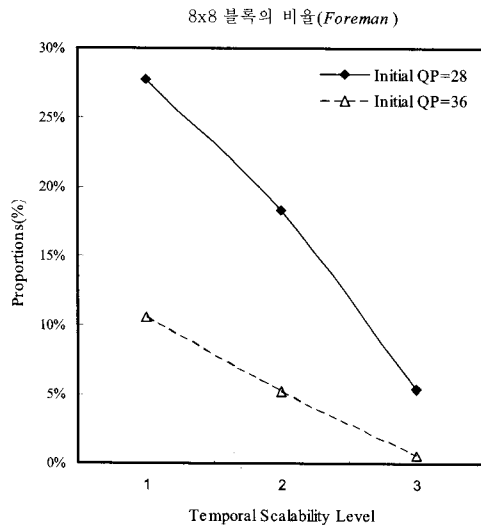


그림 3. 두 공간 계층간에 계층적 B-화면을 통한 시간적 계위  
 Fig. 3. Temporal Scalability via Hierarchical B-picture in two spatial scalabilities



(a) 시간 계위에 따른 8x8 및 그 하위 블록의 비율(Foreman)  
 (a) The proportion of 8x8 and its sub-block according to the temporal scalability level (Foreman)

(b) 시간 계위에 따른 8x8 및 그 하위 블록의 비율(Harbour)  
 (b) The proportion of 8x8 and its sub-block according to the temporal scalability level (Harbour)

그림 4. 시간 계위에 따른 미세한 블록의 부호화 비율

Fig. 4. The occurrence rates of 8x8 and its sub-partitioned block according to the temporal scalability level

따라 비트율을 감소시킬 수 있다<sup>[7, 19]</sup>. 따라서 각 시간 계위 별로 블록 모드의 결정이 다르게 결정된다. 시간 계위에 따른 8x8블록과 같은 미세 블록 모드의 결정 비율은 그림 4에 나타내었다.

그림 4는 계층적 B-화면 구조 내에서 시간적 계위에 따른 8x8블록 및 8x8하위 블록의 비율을 나타낸 것이다. GOP의 크기는 8이고 실험 영상은 Foreman(QCIF) 및 Harbour(QCIF)를 사용하였다. 그림에서 보는 것처럼 시간 계위가 증가할수록 8x8 및 그 하위 블록이 최적의 블록으로 결정되는 비율이 급격히 낮아진다. 이것은 시간 계위별로 QP를 다르게 적용하였기 때문에 균일 영상 보다는 미세한 영상을 표현하기에 적합한 8x8 및 그 하위 블록이 QP의 영향으로 시간 계위가 큰 영상에서 제거가 되기 때문이다.

### 3. SVC에서의RDO 기반 모드 결정

SVC 부호화기의 기본 계층은 H.264/AVC와 상호 호환적으로 부호화 한다. 따라서 SVC의 기본 계층에서의 블록 모드는 화면간 모드일 경우 그림 5와 같이 7가지의 서로

다른 모드를 가진다.

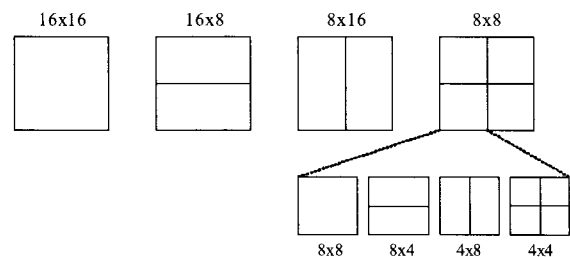


그림 5. 화면간 예측시 7가지 서로 다른 블록 모드

Fig. 5. 7 Different block modes for inter-frame prediction

또한 화면내 예측일 경우 4x4 서브 블록에 대하여 9가지의 방향 예측신호, 16x16에 대하여 4가지 신호를 이 블록들 중에서 최적의 블록 크기의 결정은 모든 블록에 대하여 RD 비용을 계산하여 최소의 RD 비용을 갖는 것을 최적의 블록으로 결정하는 원칙을 따른다. 따라서 다음 식 (2)를 최소화 하는 블록을 최적으로 블록으로 판단하는 것이다.

$$J(s, c, Mode|QP, \lambda_{Mode}) = SSD(s, c, Mode) + \lambda_{Mode} \cdot R(s, c, Mode|QP) \quad (2)$$

여기서  $s$ 는 원영상,  $c$ 는 복원된 영상을 의미한다.  $QP$ 는 모드 결정 시 사용되는 양자화 계수를 의미한다.  $\lambda_{Mode}$ 는 라그랑지 계수로서  $0.85 \times 2^{(QP-12)/3}$ 와 같이 계산된다<sup>[12]</sup>. SSD(sum of Squared Differences)는 원영상과 복원된 영상간의 왜곡도를 나타내며, 값이 클수록 왜곡이 심하며 모드 결정이 잘못된 것을 의미한다<sup>[12]</sup>. 그러나 향상 계층에서는 위에서 설명한 블록 이외에 하위 계층으로부터 계층간 예측된 모드인  $Mode_{BL\_PRED}$ 까지 후보 모드로 포함해야 하므로 RDO를 이용한 모드 결정 시 RDO의 비용계산으로 인한 복잡도가 급격히 증가하게 된다.

### III. 선택적 계층간 잔차 신호 부호화를 위한 조건의 결정

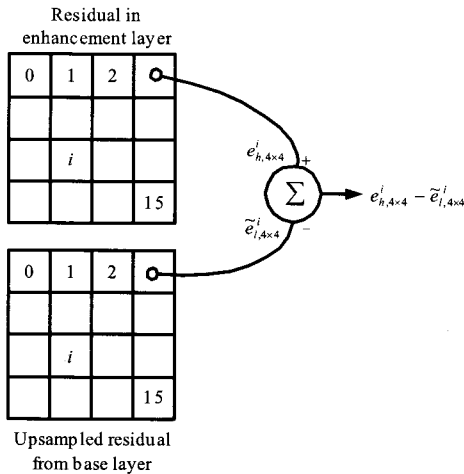


그림 6. 계층간 잔차 신호 부호화  
Fig. 6. Inter-layer Residual Coding

본 장에서는 계층간 잔차 신호에 대한 부호화를 선택적으로 수행하기 위한 결정 조건에 대하여 알아본다. 그림 6은 현재 향상 계층이 화면간 부호화를 하는 매크로 블록과 대응하는 하위 계층의 블록을 bi-linear 필터를 이용하여 업샘플링한 블록에 대하여 계층간 잔차 신호 부호화를 나타낸 그림이다. 현재 매크로 블록을 16개의 4x4 블록으로 분할하여 대응하는 블록끼리의 계층간 잔차 신호를 식 (3)같

이 나타낼 수 있다.

$$e_{R,4 \times 4}^i(x, y) = e_{h,4 \times 4}^i(x, y) - \tilde{e}_{i,4 \times 4}^i(x, y) \quad (3)$$

여기서  $i$ 는 그림 6에서 총 16개의 4x4블록 중 하나를 나타내며  $e_{R,4 \times 4}^i(x, y)$ 는 그 블록의 잔차신호 값을 나타낸다. 그리고  $e_{h,4 \times 4}^i(x, y)$ 는 향상 계층에서의 매크로 블록내의  $i$ 번째 4x4블록의 잔차 신호를 의미하며  $\tilde{e}_{i,4 \times 4}^i(x, y)$ 는 대응하는 업샘플링된 하위 4x4 블록의 잔차 신호를 의미한다. 여기서  $e_{h,4 \times 4}^i(x, y)$ 가 작다는 것은 물리적으로 상위 계층과 하위 계층의 동일 공간영역(co-located block)이 움직임의 형태가 비슷하여 두 계층간의 잔차 에너지가 적음을 의미하고 RDO 측면에서도 부호화 효율을 높일 수 있음을 의미한다. 여기서  $e_{h,4 \times 4}^i(x, y)$ 는 다음 두 경우를 고려할 수 있다.

$$e_{h,4 \times 4}^i(x, y) = \begin{cases} f_{nh}(x, y) - \begin{pmatrix} \omega_{L0} \cdot \hat{f}_{L0h}(x-v_{hxL0}, y-v_{hyL0} | L0_h) \\ \omega_{L1} \cdot \hat{f}_{L1h}(x-v_{hxL1}, y-v_{hyL1} | L1_h) \end{pmatrix} & (a) \\ f_{nh}(x, y) - \begin{pmatrix} \omega_{L0} \cdot \hat{f}_{L0b}(x-2v_{bxL0}, y-2v_{byL0} | L0_b) \\ \omega_{L1} \cdot \hat{f}_{L1b}(x-2v_{bxL1}, y-2v_{byL1} | L1_b) \end{pmatrix} & (b) \end{cases} \quad (4)$$

식 (4)-(a)는 현재 블록과 현재 향상계층에서 독립적으로 움직임벡터를 찾아 움직임 보상을 통하여 구한 잔차 신호를 나타낸다. 여기서  $f_{nh}(x, y)$ 는 현재 부호화 하고자 하는 향상 계층에서의 원 신호를 나타낸다. 식 (4)에서 첨자  $h$ 는 상위 계층을 나타낸다. 움직임벡터  $(v_{hx}, v_{hy})$ 는 1/4픽셀 정밀도를 가지는 움직임벡터로서 향상 계층에서 H.264/AVC와 동일한 구조로 예측된 움직임 벡터이고, 양방향 참조화면 리스트  $L0_h, L1_h$ 을 갖는다. 반면 식 (4)-(b)는 II장에서 설명한 계층간 움직임 벡터 예측(Inter-layer motion prediction)을 통하여 얻어진 움직임 벡터  $(2v_{bx}, 2v_{by})$ 와 참조화면리스트  $L0_b, L1_b$ 를 이용하여 움직임 보상을 한 화면과 현재 화면과의 차이를 나타낸 것이다. 식 (4)-(a)와 (4)-(b) 모두 부호화 모드로서 유효하나 식 (4)-(a)를 구할때는 현재 향상 계층에서 각 그림 5의 모든 블록 모드에 대하여 움직

입 벡터를 계산해야하기 때문에 연산량이 급격히 증가하는 반면 식 (4)-(b)는 계층간 움직임 벡터 예측을 통하여 하위 계층에서 얻어진 최적의 벡터를 별도의 움직임 벡터 예측 없이 사용하므로 적은 연산량을 통하여 계산 할 수 있다. 본 제안 알고리즘에서는 계층간 움직임 벡터를 통하여 얻어진 움직임 벡터와 참조화면리스트를 이용하는 식 (4)-(b)를 이용하여 얻어진  $SAD_{4 \times 4}^i$ 의 값을 바탕으로 현재 부호화하고자 하는 향상계층의 매크로 블록의 정수변환계수의 특성을 계층간 잔차 신호의 SAD값 만으로 미리 판단하여 선택적으로 계층간 잔차 부호화를 수행하는 알고리즘을 제안한다. 먼저  $i$  번째  $4 \times 4$  블록에 대한 정수변환<sup>[1, 14]</sup>은 다음 식 (5)와 같이 표현할 수 있다.

$$C_I^i(u, v) = \sum_{x=0}^3 \sum_{y=0}^3 e_{R,4 \times 4}^i(x, y) \cdot |K(x, u) \cdot K(y, v)| \quad (5)$$

여기서  $K(\alpha, \beta)$ 는 다음 식 (6)과 같이 정의 된다.

$$K(\alpha, \beta) = \left\lfloor \frac{2.5C(\beta)}{\sqrt{2}} \cos \frac{(2\alpha+1)\beta\pi}{8} \right\rfloor \quad (6)$$

$C(\beta)$ 는 정수변환의 정의에 의해 식 (7)과 같이 나타낼 수 있다.

$$C(\beta) = \begin{cases} 1/\sqrt{2}, & \text{if } \beta = 0 \\ 1, & \text{otherwise} \end{cases} \quad (7)$$

또한  $\lfloor \cdot \rfloor$ 는 가장 가까운 정수값을 갖게 하는 라운드오프(round-off) 연산자이다.  $C_I^i(u, v)$ 는 매크로 블록 내  $i$  번째  $4 \times 4$ 블록의 정수변환 계수이고 정수변환에 대한 양자화 계수를 다음과 같이 적용하여 양자화 할 수 있다.

여기서 %는 모듈러 연산자이고 >>는 이진 쉬프트 연산자

$$C_{Q,4 \times 4}^i(u, v, r, QP) = \text{sign}\{C_I^i(u, v)\} \times \left( |C_I^i(u, v)| \cdot MF(QP\%6, r) + f \cdot 2^{qbits} \right) \gg qbits \quad (8)$$

(binary shift right)이다.  $qbits$ 은 정수 변환에서 나누기연산을 피하기 위하여 도입된 쉬프트 비트로서  $qbits = 2^{15+QP\%6}$ 이 되고,  $f$ 는 변환계수값 0주위의 dead zone<sup>[14]</sup>을 결정하는 인자로서  $f \in [0, 0.5]$ 의 값을 갖는다. SVC 참조소프트웨어 JSVM에서는 화면내 예측일 경우 1/3, 화면간 예측인 경우 1/6을 갖는다<sup>[13]</sup>. 그리고  $MF(QP\%6, r)$ 는 곱셈 계수(Multiplication Factor)라고 하며 각 주파수 영역별로 이미 지정된 값이고 (9)와 같이 정의한다<sup>[13]</sup>.

$$MF(QP\%6, r) = \begin{bmatrix} 5243 & 8066 & 13107 \\ 4660 & 7490 & 11916 \\ 4194 & 6554 & 10082 \\ 3647 & 5825 & 9362 \\ 3355 & 5243 & 8192 \\ 2893 & 4559 & 7282 \end{bmatrix} \quad (9)$$

식 (8)와 (9)에서  $r = 2 - (u\%2) - (v\%2)$ 이다.  $4 \times 4$ 블록내의 정수변환계수의 위치를 그림 7에 나타낸다. 즉 그림 7과 같이 MF의 첫번째 열( $r = 0$ )은  $4 \times 4$  정수변환 계수의 (1,1), (1,3), (3,1), (3,3) 위치를 세번째 열( $r = 2$ )은 (0, 0), (0, 2), (2, 0), (2, 2) 위치를 마지막으로 두번째 열( $r = 1$ )은 그 이외의 위치를 나타낸다.

(0,0)	(0,1)	(0,2)	(0,3)
(1,0)	(1,1)	(1,2)	(1,3)
(2,0)	(2,1)	(2,2)	(2,3)
(3,0)	(3,1)	(3,2)	(3,3)

그림 7.  $4 \times 4$  정수변환영역에의 주파수 요소

Fig. 7. Frequencies Components for integer transformed coefficients in the  $4 \times 4$  block

[15, 16]에서 설명한 것처럼  $e_{R,4 \times 4}^i(x, y)$ 의 정수변환된 모든 계수가 0이 되는 조건이 정수변환 이전 단계에서 조기에 검출되려면 다음과 같은 충분 조건이 필요하다.



$$C_I^j(u, v) \leq \sum_{x=0}^3 \sum_{y=0}^3 |e_{R,4 \times 4}^i(x, y)| \cdot |K(x, u) \cdot K(y, v)| \quad (10)$$

식 (5)와 식 (10)로부터 든 정수변환계수가 0이 되는 충분 조건은 다음과 같다.

$$\sum_{x=0}^3 \sum_{y=0}^3 |e_{R,4 \times 4}^i(x, y)| \leq \frac{2^{qbits} - f \cdot 2^{qbits}}{4 \cdot MF(QP\%6, 0)} \quad (11)$$

식 (11)은  $r=0$ 일 때 문턱치를 최대로 하는 충분 조건이다. [16]에서 식 (11)을 각 주파수 성분 별로 일반화 하여 나타내기 위하여 다음과 같은 관계식을 얻는다.

$$\sum_{x=0}^3 \sum_{y=0}^3 |e_{R,4 \times 4}^i(x, y)| \leq \frac{2^{qbits} - f \cdot 2^{qbits}}{K_{\max}(r) \cdot MF(QP\%6, r)} = Th(r) \quad (12)$$

여기서  $K_{\max}(r)$ 은 주파수 영역에 대하여  $|K(x, u) \cdot K(y, v)|$ 이 최대가 되는 값을 의미한다.  $K_{\max}(r)$ 은 각 주파수 영역에 대하여 다음과 같이 표현할 수 있다.

$$K_{\max}(u, v) = 2^{2-r}, \quad r = 2 - (u\%2) - (v\%2) \quad (13)$$

식 (13)은 각 정수변환 계수의 주파수 위치에 해당하는 영변환계수를 나타낸다. 다음 그림 8은  $Th(r)$ 에 따

라 정수변환 계수가 0이 되는 주파수의 위치를 나타낸다. 여기서 짙은 색이 변환계수가 0이 되는 변환계수의 위치이다.

제안 알고리즘에서는 이런 주파수 계수별 영변환계수를 가지는 조건을 계층적 B-화면의 특성을 이용하여 다음 식 (14)와 같이 다시 정리한다.

$$\sum_{x=0}^3 \sum_{y=0}^3 |e_{R,4 \times 4}^i(x, y)| \leq \frac{2^{15+QP_i\%6} - f \cdot 2^{15+QP_i\%6}}{2^{2-\alpha(t_i)} \cdot MF(QP_i\%6, \alpha(t_i))} \quad (14)$$

$$= Th(\alpha(t_i), QP_{t_i})$$

여기서  $t_i$ 는 한 GOP내의 시간 계위를 나타내고  $QP_{t_i}$ 는 각 시간계위의 화면에 적용되는 양자화 계수이다. 따라서 2장에서 설명했던 것처럼 각 시간계위별로 화면의 중요도가 다르므로 주화면 또는 시간계위가 낮은 화면에 대해서는  $Th(\alpha(t_i), QP_{t_i})$ 를 작게하여 모든 주파수 변환 계수가 0일 때를 All Zero Coefficients Block으로 검출하고 시간 계위가 커질수록 그 조건을 완화하여  $Th(\alpha(t_i), QP_{t_i})$ 를 큰 값을 갖도록 한다. 한편  $\alpha(t_i)$ 는 각 시간 계위와 변환 계수의 위치와의 관계를 나타내며 다음 식 (15)와 같이 정의 한다.

$$\alpha(t_i) = \begin{cases} 0, & 0 \leq t_i \leq 1 \\ 1, & 2 \leq t_i < 3 \\ 2, & 3 \leq t_i \end{cases}, \text{ for GOP size} \leq 8 \quad (15)$$

시간 계위가 낮은 화면에서는  $\alpha(t_i)$ 이 작은 값이 입력되

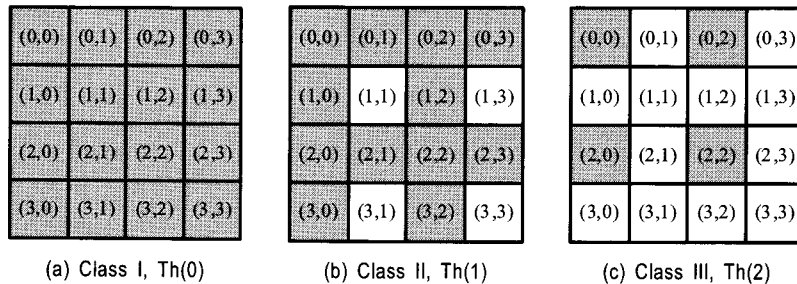


그림 8. 각 문턱치 값에 따라 0이 되는 주파수 계수

Fig. 8. Zero frequencies components according to the threshold values

므로 문턱값,  $Th(\alpha(t_l), QP_{t_l})$ 가 작아져 선택적으로 계층간 잔차 부호화 할 수 있는 기회가 줄어들는다. 4x4 블록의 잔차 신호에 대한 SAD의 문턱값은 계층적 B-화면 구조와 그 화면에 적용되는 QP값에 대한 관계식으로 나타낼 수 있다. 또한 한 GOP내의 QP도 시간 계위에 따라 다음 식 (16)과 같은 크기로 적용된다<sup>[6]</sup>.

$$QP_{key} < QP_{t_1} < QP_{t_2} < QP_{t_3} \quad (16)$$

이것은 2장에서 설명한 것처럼 시간계위에 따라 화면의 중요도에 따라 다른 QP를 적용함으로써 총 부호화 비트를 절약할 수 있기 때문이다. 시간 계위와 QP를 적용했을 때 정수 변환 계수의 특성을 결정짓는 문턱치 값은 다음 부등식 (17)과 같은 순서의 크기로 결정됨을 알 수 있다.

$$Th(\alpha(t_0), QP_{t_0}) < Th(\alpha(t_1), QP_{t_1}) < Th(\alpha(t_2), QP_{t_2}) \quad (17)$$

그림 8은 시간 계위와 QP에 따라 SAD값에 따른 4x4 서브 블록을 3가지 클래스로 분류한 것이다. 즉 시간 계위가 아주 낮은 화면은 모든 정수 변환 계수가 0인 조건만을 선택하도록 하고 중간 수준의 시간 계위에서는 그림 8-(b)와 같이  $r=1$  인 주파수 요소를 포함하여 정수 변환 계수가 12개 이상일 때를 Class II로 한다. 마지막으로 시간 계위가 아주 낮은 화면은 조건을 완화하여  $r=2$  인 주파수 요소를 포함(DC 포함), 4개 이상의 주파수 요소가 0일 때 클래스 III로 한다. 이와 같이 시간 계위와 QP 그리고 계층간 잔차 신호의 SAD에 따라서 4x4 블록을 세 가지 특성으로 분류할 수 있으며 16개의 4x4블록에 대하여 다음 식 (18)과 같은 조건을 만족하면 선택적 계층간 잔차 신호 부호화를 수행할 수 있도록 한다.

$$\sum_{i=0}^{15} f(i) < \Gamma_S \quad (18)$$

제안 알고리즘에서  $f(i)$ 는 다음과 같이 계산한다

$$f(i) = \begin{cases} 1, & \text{if } \sum_{x=0}^3 \sum_{y=0}^3 |e_{R,4 \times 4}^i(x, y)| < Th(\alpha(t_l), QP_{t_l}) \\ 0, & \text{else} \end{cases} \quad (19)$$

식 (19)는 한 매크로 블록이 선택적 잔차 부호화로 결정될 수 있는 지 여부를 결정하는 조건식이다. 여기서 16개의 4x4 블록에 대하여 식 (19)의 조건이 만족되면  $f(i)$ 를 1로 설정하고 그렇지 않으면 0으로 한다. 총 16개의  $f(i)$ 의 합이  $\Gamma_S$ 보다 작으면 선택적 잔차신호 부호화를 할 수 있고 그렇지 않으면 정상적인 부호화 과정을 따른다.  $\Gamma_S$ 은 공간 계층 내의 시간 계위에 따라 결정되며 식 (20)과 같다.

$$\Gamma_S = 1 + \chi_{inc}(t_l), \quad \chi_{inc}(t_l) = t_l + 1, \quad 0 \leq t_l \leq n+1 \quad (20)$$

여기서  $\chi_{inc}(t_l)$  시간 계위에 따른 증분값(increment)이며 시간 계위가 증가할수록 그 값이 커진다. 식 (17)과 (20)를 고려했을 때 시간 계위가 작을수록 선택적으로 계층간 잔차 부호화하는 기회가 줄어들고 커질수록 선택적 계층간 잔차 부호화의 비율이 커짐을 알 수 있다. 이것은 시간 계위가 작을수록 화면의 높은 화질을 유지하여 한 GOP내의 화질을 전체적으로 유지하기 위함이고 시간계위가 커질수록 영향을 GOP내의 다른 화면에 영향을 미칠 수 있는 경우가 작기 때문에 보다 문턱치가 큰 완화된 조건으로 선택적 잔차 부호화를 수행할 수 있다. 이것은 그림 4에서 보는 것처럼 시간 계위에 따라 세밀한 블록모드의 결정 비율이 매우 달라 지는 성질을 이용한 것이다.

#### IV. 선택적 잔차 신호 부호화와 공간적 향상 계층에서의 조기 모드 결정

3장에서 선택적 잔차 신호 부호화를 할 수 있는 조건에

대하여 알아 보았다. 본 장에서는 선택적 잔차 신호 부호화 정보를 이용하여 공간적 향상 계층에서 부호화 고속화 방법과 고속 모드 결정 방법에 대하여 알아본다. 3장에서  $e_{R,4 \times 4}^i(x, y)$ 가 매우 작다면 두 계층간의 움직임 요소가 매우 유사하여 계층간의 잔차 신호가 매우 작음을 의미한다. 식 (4)-(b)를 이용하여 식 (18)을 만족하였다면 향상 계층에서의 별도의 움직임 벡터 예측없이 계층간 움직임 벡터 예측을 통해 하위 계층으로부터 얻어진 움직임 벡터를 이용하여 계층간 잔차 신호 부호화를 수행하므로 향상 계층에서의 연산량은 급격히 감소한다. 그러나, 하위 계층으로부터 예측된 움직임 벡터는 식 (1)처럼 하위 계층의 움직임 벡터에 2배를 하여 얻어진 벡터이므로 1/4 화소 정밀도의 절반 수준인 1/2 화소 정밀도로 감소하므로 벡터 정확도 측면에서 효율이 감소할 수 있다. 따라서 식 (18)을 만족하는 블록의 경우에는 향상 계층에서는 그림 9의 블록 다이어그램과 같이 세가지 모드만을 후보 모드로 선택하여 RDO를 계산한 후에 부호화를 완료한다. 선택적 계층간 잔차 부호화 결정 조건을 만족했을 경우에 대한 부호화 과정을 그림 9에서 나타낸다. 그림 9에서 보는 것처럼 조건을 만족한 경우 공간적 향상 계층에서는 RD 비용 계산을 3회 밖에 실행하지 않는다. 블록 다이어그램의 (a)에서는 식(4)-(b)와 같이 현재 업샘플된 하위 계층의 잔차 신호와, 현재 계층의 원신호와 하위 계층으로부터 얻은 움직임 벡터와 참조화면 리스트를 이용하여 구한 움직임 보상 화면과의 잔차 신호와의 계층간 잔차 신호를 이용한 RD 비용을 구한다. (b)에서는 그림 4에서의 하위 계층으로부터 예측된 모드 분할 정보와 움직임 벡터를 이용하여 움직임 보상된 화면을 이용한 RD 비용이다. 이 경우는 H.264/AVC와 같은 형태로 계층 독립적으로 RD 비용을 계산하되, 움직임 벡터와 참조 화면 정보 등은 하위 계층으로부터 얻는 차이가 있다. (c)는 H.264/AVC와 동일한 방식으로 RD비용을 구하는 것으로 조건을 만족하였을 경우 16x16 매크로 블록 모드에서만 움직임 벡터 등을 예측하여 RD 비용을 구한다. 계층간 잔차 신호 부호화 조건을 만족하였을 경우는 계층간 잔차신호가 작다는 것을 의미하고 물리적으로 움직임이 적거나 균일한 영상에서 가능

하므로 이 경우에는 향상 계층에서 16x16 매크로 블록에 대해서만 RD 비용을 구함으로써 연산량을 크게 줄일 수 있다. 따라서 이 경우에는 실제로 움직임 벡터 측정을 1회, 향상계층의 16x16블록 모드에 대해서만 실행하므로 매우 효율적으로 부호화를 완료할 수가 있다.

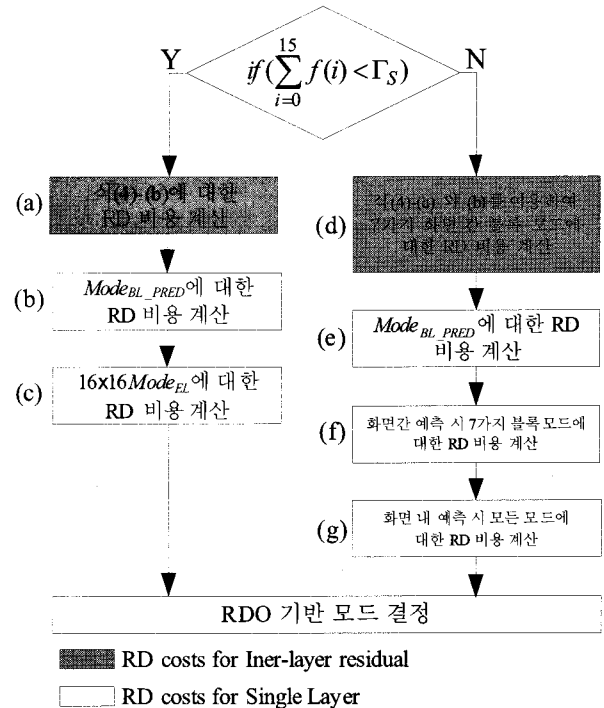


그림 9. 선택적 잔차 신호 부호화에 대한 제안 알고리즘

Fig.9. The proposed algorithm for selective inter-layer residual coding

(d), (e), (f), (g)는 모두 일반적인 부호화 과정을 이며 모든 블록 모드에 대하여 RD 비용을 측정하는 과정을 거친다. 블록다이어그램에서 검은색 블록은 계층간 잔차 신호 부호화를 수행하는 부분이다.

## V. 실험 결과

실험을 위해 SVC의 참조 소프트웨어인 JSVM 9.0<sup>[17]</sup>을 사용하였다. 참조 프레임의 개수는 1개, 각공간 계층에서 시간적 계위성을 위해 GOP의 크기를 8로 한 계층적 B-화

면 구조를 사용하였다. 움직임벡터의 정밀도는 1/4로 설정하였다. 표 1은 제안알고리즘에 대한 부호화 실험환경을 나타낸다.

표1. 부호화 파라미터 선택

Table1. Encoding Parameters Options

참조화면의 수	1
GOP 크기	8(IBBBBBBBP)
움직임벡터정밀도	1/4 화소
움직임벡터탐색범위	32화소
총 화면 수	100
비용함수	정수화소 탐색시: SAD 부화소 탐색시: SATD
공간 계층의 수	2(기본계층(QCIF) + 향상계층(CIF))
화질 계위	사용안함
컴퓨터 환경	Pentium Dual Core 2 2.4 GHz cpu , 2GB 주메모리

다양한 특성을 가지는 시험 영상을 이용하기 위하여 움직임이 적은 시험 영상부터 움직임이 많은 것까지 다양한 특성을 가지는 영상, 즉 *Container*, *Mother&Daughter*, *Silent*, *Foreman*, *Harbour*, *Ice*, *Tempete*을 사용하였다. 실험결과는 복잡도 감소로 인한 부호화 시간의 감소와 이에 따른 PSNR과 비트율의 변화량을 측정하여 성능을 비교하였다. 본 알고리즘의 목적은 부호화 과정의 복잡도를 감소시켜 전체 부호화 시간을 감소시키는 것으로 한다. 따라서 부호화 시간을 측정하여 원 참조 소프트웨어보다 얼마만큼의 부호화 시간의 감소를 가져왔는지 측정한다. 그리고 본 알고리즘 적용 전 후의 PSNR 변동량과 비트율의 변화량을 dB와 %로 각각 측정하였다. 알고리즘의 적용 전후의 RD 성능의 변화량을 측정함으로써 알고리즘의 적용 후 복잡도 감소에도 RD 성능을 유지하는지 보기 위함이다. 시험 영상으로는 움직임이 적은 순서부터 *Container*(C), *Mother&Daughter*(M-D), *Silent*(S), *Foreman*(F) 영상을 사용하였고 다음으로 움직임은 적으나 텍스처가 미세한 특징을 보여주는 *Harbour*(H)영상 그리고 스포츠 영상인 *Ice*(I)영상을 사용하였다. 시험 영상은 기본 계층에는

30fps의 QCIF 해상도를 향상계층에는 30fps의 CIF 해상도를 100프레임을 사용하였다. 양자화 계수는 기본계층과 향상계층에 동일한 값을 사용되되 28, 32, 36, 40을 사용하여 실험하였다. 전체 비트율 구간의 PSNR과 비트율의 변화를 알아보기 위하여 [18]에서 제시한 BDPSNR(dB)과 BDBR(%)를 계산하였다.

표 2. 제안 알고리즘의 BDBR과 BDPSNR

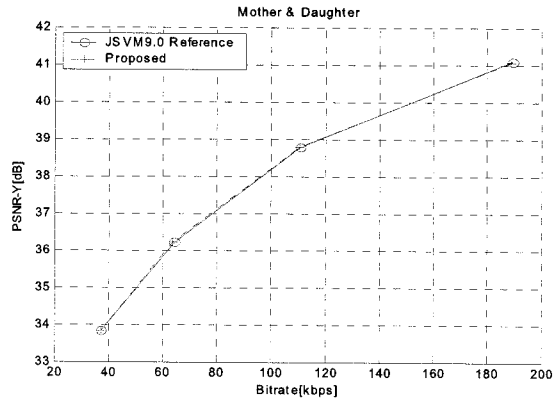
Table 2. BDBR and BDPSNR for the proposed algorithm

Seq.	BDBR (%)	BDPSNR (dB)	Avg. Time Saving (%)
C	0.01	0.00	61.4
M-D	0.00	-0.04	62.8
S	0.76	0.00	55.6
F	1.49	-0.06	54.6
T	0.15	0.00	42.0
H	0.40	-0.02	31.1
I	2.15	-0.12	55.6
CT	0.18	-0.01	48.9
Avg	0.64	-0.03	51.5

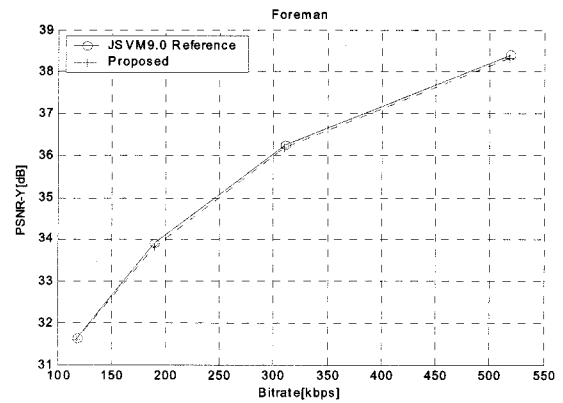
그림 9에서 보는 것처럼 움직임이 작은 *Mother&Daughter* 영상부터 텍스처가 복잡하고 카메라 움직임이 큰 *City*영상까지 RD 성능의 저하가 없는 것을 그래프를 통해 알 수 있다. 반면 부호화의 시간은 모든 영상에 걸쳐 평균적으로 50%이상의 부호화 시간을 단축할 수 있었다. 이것은 불필요한 계층간 잔차 부호화에 대한 윌-왜곡 최적화 수행을 선택적으로 수행함으로써 얻어지거나 이에 해당하는 블록에 대하여 큰 블록크기에 대해서만 움직임 벡터를 예측 및 보상을 수행함으로써 부호화 시간을 크게 단축 시켰기 때문이다.

## VI. 결론

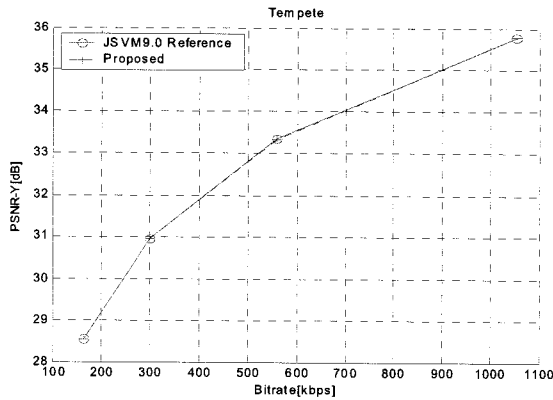
본 논문에서는 공간 계위성을 지원하기 위해 도입된 계층간 잔차 신호 부호화로 인해 급격히 증가한 스케일러블



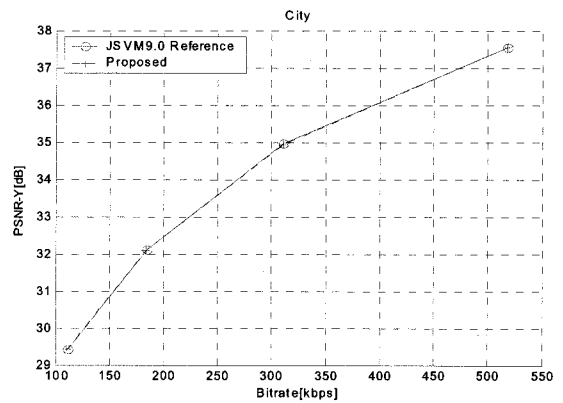
(a) RD 성능 그래프 (Mother&Daughter)  
(a) RD performance graph (Mother&Daughter)



(b) RD 성능 그래프 (Foreman)  
(b) RD performance graph (Foreman)



(c) RD 성능 그래프 (Tempete)  
(c) RD performance graph (Tempete)



(d) RD 성능 그래프 (City)  
(d) RD performance graph (City)

그림 9. 제안 알고리즘을 이용하였을 때의 RD 성능 그래프

Fig. 9. RD performance graphs for the proposed algorithm

비디오 부호화에 대한 선택적 잔차 신호 부호화를 수행함으로써 SVC 부호화기의 복잡도를 크게 개선한 알고리즘을 제안하였다. 제안된 방법은 현재 부호화하고자 하는 매크로 블록이 화면간 예측일 때 계층간 잔차 신호의 SAD값이 시간적 계위와 양자화 파라미터에 의해 적응적으로 변화하는 문턱치 값보다 작으면 정수 변환된 계수의 특성에 따라서 선택적으로 계층간 잔차 신호 부호화를 수행하는 알고리즘이다. 또한 그 조건이 만족되면 공간적 향상 계층에서는 움직임 벡터 예측 및 RDO 수행을 16x16 블록에서만 수행하여 연산량을 크게 줄일 수 있다. 제안된 알고리즘은

SVC의 공간 및 시간 계위의 특성을 이용하여 고안되었으며, 움직임이 적은 영상부터 많은 영상까지, 단순한 텍스처 영상뿐만 아니라 복잡한 영상에 이르기 까지 다양한 실험 영상을 통하여 전체 부호화 시간 대비 평균 51.5%의 부호화 시간의 감소 효과를 얻을 수 있었다. 그러나 PSNR 감소는 0.03dB, 비트율 증가는 0.64%에 불과해 무시할 수 있을 정도로 작은 값이었다. 따라서 제안된 알고리즘은 실시간 SVC 부호화기 구현 등에 아주 유효하게 사용될 수 있는 방법이라 할 수 있다.

참 고 문 헌

[1] T.Wiegand, G.Sullivan, J.Reichel, H.Schwarz and M.Wien, ISO/IEC JTC 1/SC 29/WG 11 and ITU-T SG16 Q.6: JVT-W201 'Joint Draft 10 of SVC Amendment,' 23th Meeting, San Jose, California, April 2007.

[2] ITU-T and ISO/IEC JTC 1, "Generic Coding of Moving Pictures and Associated Audio Information Part 2:Video,"ITU-T Recommendation H.262 ISO/IEC 13818-2, 1994

[3] ISO/IEC 14496-2,"Information technology-Coding of audio-visual objects -part 2", International Standard, second edition, December 2001

[4] ITU-T and ISO/IEC JVT 1, "Advanced video coding for generic audiovisual services," ITU-T Recommendation H.264 and ISO/IEC 14496-1-(MPEG-4 AVC), Version 1:May 2003, Version 2: Jan 2004, Version 3:Sep.2004, Version 4:July 2005

[5] 최해철, 강정원, 배성준, 유정주, "스케일러블 비디오 부호화 Phase 1 기술", 전자공학회지, 제 34권, 8호, pp.18-26, 2007년 8월

[6] 이범식, 김문철, 함상진, 김병선, 이근식, 박근수, "스케일러블 비디오 부호화에서 통계적 가설검증을 이용한 고속 화면간 모드 결정", 제 12 권 3호 pp.250-265, 2007년 5월

[7] H.Schwarz, D.Marpe and T.Wiegand, "Hierarchical B pictures," Joint Video Team, Doc.JVT-P014, Poznan, Poland, July 2005

[8] H.Li, Z.Li and C.Wen, " Fast Mode Decision Algorithm for Inter-Frame Coding in Fully Scalable Video Coding", IEEE Transaction on Circuits and Systems for Video Technology, vol 16, no 7, pp.889-895, July 2006.

[9] D.S.Turga, M.van der Schaar, Y.Andreopoulos, A.Munteanu, P.Schelkens, "Unconstrained motion compensated temporal filtering(UMCTF) for efficient and flexible interframe wavelet video coding," Signal Processing :Image Communication, 20(2005) 1~19

[10] A.Segall and J.Zhao, ISO/IEC JTC 1/SC 29/WG 11 and ITU-T SG16 Q.6: JVT-U42 'CE:Evaluation of Texture Upsampling with 4-tap Cubic-Spline Filter,' 21st Meeting, Hangzhou, China, October 2006

[11] H.Schwarz, T.Hinz, D.Marpe and T.Wiegand, "Constrained Inter-Layer Prediction for Single-Loop Decoding in Spatial Scalability," IEEE International Conference on Image Processing, vol. 2, pp.870-873, September 2005

[12] D.Wu, F.Pan, K.P.Lim, S.Wu, Z.G.Li, X.Lin, S.Rahardja and C.C.Ko, " Fast Intermode Decision in H.264/AVC Video Coding," IEEE Transaction on Circuits and Systems for Video Technology, vol.15, no.6, pp.953-958, July 2005

[13] H.Wang, S.Kwong, "Hybrid Model to Detect Zero Quantized DCT Coefficients in H.264," IEEE Transaction on Multimedia, vol.9, no.4, pp.728-735, June 2007

[14] H.Hallapuro, M.Karczewicz and L.Kerofsky, "Low Complexity Transform and Quantization in H.264/AVC," IEEE Transaction on Circuits and Systems for Video Technology, vol.13, no 7, pp.598-603, July, 2003

[15] L. A. Sousa, "General method for eliminating redundant computations in video coding," Electronics Letters, vol.36, no.4, pp.306-307, Feb. 2000.

[16] Y.H.Moon, G.Y.Kim and J.H.Kim, "An Improved Early Detection Algorithm for All-Zero Blocks in H.264 Video Encoding," IEEE Transaction on Circuits and Systems for Video Technology, vol.15, no.8, pp.1053-1057, August, 2005

[17] J.Vieron, M.Wien and H.Schwarz, ISO/IEC JTC 1/SC 29/WG 11 and ITU-T SG16 Q.6: JVT-V203 'JSVM 9 Software' 22th Meeting, Marrakech, Morocco, January 2007.

[18] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," Doc. VCEG-M33, April 2001

[19] Heiko Schwarz, Detlev Marpe and Thomas Wiegand, "Analysis of Hierarchical B Pictures and MCTF," IEEE International Conference on Multimedia & Expo (ICME 2006), Canada, July 9-12, 2006

[20] H. Schwarz, D. Marpe, and T. Wiegand, "MCTF and Scalability Extension of H.264/AVC," Proc. PCS 2004, San Francisco, CA, USA, December 2004.

저 자 소 개



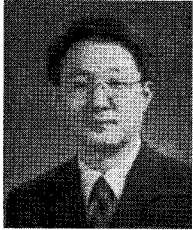
이 범 식

- 2000년 8월 : 고려대학교 전기전자전공공학부 학사
- 2006년 8월 : 한국정보통신대학교 공학부 석사
- 2006년 8월 ~ 현재 : 한국정보통신대학교 공학부 박사 과정
- 2000년 9월 ~ 2003년 9월 : (주)기산텔레콤
- 주관심분야 : Scalable Video Coding, H.264/AVC, 차세대 비디오 부호화, 모바일 영상 통신, 내용기반 영상 검색, 패턴인식

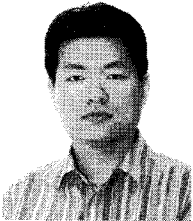
---

 저 자 소 개
 

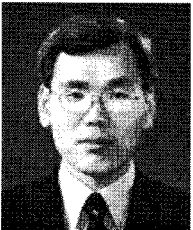
---

**김 문 철**

- 1989년 : 경북대학교 학사
- 1992년 : University of Florida, Electrical and Computer Engineering, 석사
- 1996년 : University of Florida, Electrical and Computer Engineering, 박사
- 1997년 1월 ~ 2001년 2월 : 한국전자통신연구원 실감영상연구팀/영상미디어 연구팀 팀장
- 2001년 2월 ~ 현재 : 한국정보통신대학교 부교수
- 2001년 1월 ~ 현재 : MPEG 포럼 운영위원, SC29-Korea 전문위원
- 주관심분야 : 차세대 비디오 부호화, 모바일 영상 통신, 영상 및 신호처리, 패턴인식, SuperHDTV/UDTV/, IPTV personalization

**함 상 진**

- 1996년 2월 : 연세대학교 전자공학과 학사
- 1998년 8월 : 연세대학교 전자공학과 석사
- 2001년 7월~현재 : 한국방송 방송기술연구팀, 현 U콘텐츠처리보호 선임연구원
- 주관심 분야 : 고효율 영상압축, 콘텐츠 저작권 보호 기술, 영상 신호 처리

**박 창 섭**

- 1980년 : 한양대학교 전자공학과 (학사)
- 1982년 : 한양대학교 산업대학원 (석사)
- 2001년 : 한국과학기술원 정보 및 통신공학과 (박사)
- 1981년 - 현재 : 한국방송 방송기술연구팀
- 주관심분야 : 영상신호압축, 컴퓨터 비전, 입체TV

**박 근 수**

- 1982년 : 서울대학교 제어계측공학과, 학사
- 1984년 : KAIST 전기전자공학과, 석사
- 1984년 3월 ~ 현재 : 한국방송 방송기술연구소 근무, 현 U콘텐츠처리보호 수석연구원
- 주관심 분야 : HD/SD 영상호과 및 처리 방송제작기술, DTV 방송콘텐츠 보호기술, 유비쿼터스 콘텐츠 처리기술