

# H.264/SVC의 계층간 화면내 예측에서 보간법에 따른 부호화 성능 분석

## Performance Analysis of Coding According to the Interpolation filter in Inter layer Intra Prediction of H.264/SVC.

\*길대남, \*\*정차근  
Dae-Nam Gil, Cha-Keon Cheong

**Abstract** - International standard specification, H.264/SVC improved from H.264/AVC, is set up so as to promote free use of huge multimedia data in various channel environments. H.264/AVC is a international standard specification for video compression, adopted and commercialized as standard for DMB broadcasting by JVT of ISO/IEC MPEG and ITU-T VCEG. SVC standard uses 'intra/inter prediction' in AVC as well as 'inter-layer intra prediction', 'inter-layer motion prediction' and 'inter-layer residual prediction' to improve efficiency of encoding. Among prediction technologies, 'inter-layer intra prediction' is to use co-located block of up sampled sublevels as a prediction signal. At this time, application of interpolation is one of the most important factors to determine encoding efficiency. SVC's currently using poly-phase FIR filter of 4-tap and 2-tap respectively to luma components. This paper is written for the purpose of analyzing encoding performance according to the interpolation. For this purpose, we applied poly-phase FIR filter of '2-tap', '4-tap' and '6-tap' respectively to luma components and then measured bit-rate, PNSR and running time of interpolation filter. We're expecting that the analysis results of this paper will be utilized for effective application of interpolation filter. SVC standard uses 'intra /inter prediction' in AVC as well as 'inter-layer intra prediction', 'inter-layer motion prediction' and 'inter-layer residual prediction' to improve efficiency of encoding.

**Key Words** :스케일러블, 계층간 화면내 예측, 상향 표본화, 보간법

### 1. 서론

급속한 기술 발전을 이룩하고 있는 멀티미디어 정보 통신 시대의 전개와 함께, 정보통신기기의 이동성 및 개인화가 빠르게 진전 되고 있다. 이와 함께, 디지털 통신 기술의 획기적인 발전에 힘입어 무선 통신기기 관련 시장이 급속히 확대되고 있다. 특히 PCS와 같은 무선 이동 통신 단말기기 뿐만 아니라, 노트북 PC나 PDA, DMB, PMP와 같은 개인 휴대 정보 단말기 보급의 확대에 유선망의 채널 상태에 구애 되지 않는 동영상 정보의 무선 송수신 공유를 요구 하고 있다.[1]

하나의 영상을 시간(temporal), 공간(spatial), 화질(quality) 적 측면에서 다수의 계층으로 나누어 인코딩하여 다양한 네트워크 및 디바이스 환경에서 적용적으로 복원될 수 있는 스케일리빌리티(scalability)에 관한 연구가 활발히 진행되고 있으며 그 결과로 ISO/IEC MPEG과 ITU-T VCEG의 JVT를 중심으로 H.264/AVC의 확장인 H.264/SVC의 새로운 표준규격이 완료되었다.[2]

하나의 기본 계층(base layer)과 하나 이상의 향상 계층(enhancement layer)으로 구성된 SVC는 '공간', '시간', 'SNR' 그리고 '혼합' 스케일리빌리티를 지원하며 부호화 효율을 높이기 위하여 기존 AVC의 '화면내 예측'과 '화면간 예측' 외에

계층간 상관관계를 이용한 '계층간 화면내 예측', '계층간 움직임 예측' 그리고 '계층간 잔류 예측' 기술을 사용한다. 이때 '공간' 스케일리빌리티는 계층간의 공간 비율(spatial rate)로 상향 표본화한 하위 계층의 정보를 예측 정보로 사용하는 것을 기본으로 한다. 예측 신호의 상향 표본화에 있어서 기본 계층의 움직임 정보를 예측 정보로 사용하는 '계층간 움직임 예측'을 제외하고는 보간법의 적용은 필수적이다. 따라서 SVC는 보간법을 필요로 하는 '계층간 화면내 예측'과 '계층간 잔류 예측'에 대하여 휘도 성분은 4-tap poly-phase FIR 필터를 적용하고 색차 성분과 잔류 성분은 2-tap poly-phase FIR 필터를 적용하고 있다. 이러한 보간법의 적용은 SVC의 부호화 효율을 결정 짓는 가장 중요한 요소 중 하나이다.[3][6]

본 논문은 '계층간 화면내 예측'에서 보간법에 따른 부호화 성능 분석을 그 목적으로 하며 이를 위하여 휘도 성분에 대하여 2-tap, 4-tap, 6-tap의 poly-phase FIR 필터를 각각 적용하여 그에 따른 비트율과 PSNR 및 알고리즘의 복잡도를 측정하는 실험을 수행 하였다. 논문의 구성은 2장에서 '계층간 화면내 예측'의 주요 기법을 기술하며 3장에서 실험 과정의 기술 및 실험 결과를 분석하여 4장에서 결론을 맺는다.

### 2. 계층간 화면내 예측

H.264/SVC는 다양한 공간 해상도를 제공하기 위해 다중

저자 소개

\* 길 대 남 : 호서대학교 메카트로닉스 공학과

\*\* 정 차 근 : 호서대학교 시스템 제어 공학과

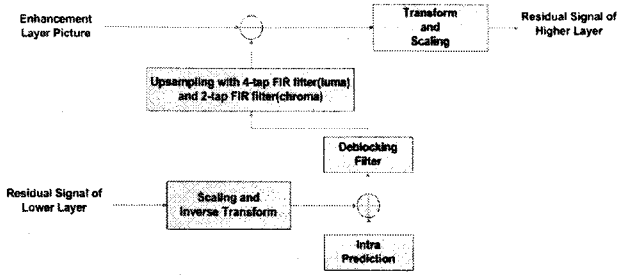


그림 1. '계층간 화면내 예측' 부호화기 구조  
Fig. 1. 'inter-layer intra prediction' encoder structure[6]

해상도(multiresolution)의 계층화 부호화 기법을 기반으로 연속되는 계층간 예측 부호화를 사용한다. 계층간 예측 부호화의 기본은 항상 계층의 부호화 왜곡(rate-distortion) 개선을 위해 하위 계층 정보의 활용을 극대화한 것으로 이와 같은 계층간 예측은 '계층간 화면내 예측', '계층간 움직임 예측', '계층간 잔류 예측' 3 가지가 있다.[2]

'계층간 화면내 예측'은 참조 계층의 동일 위치 매크로 블록을 복호하여 화소 레벨의 예측 신호로 사용하는 것으로 이를 'intra base' 모드라 한다. 'intra base' 모드는 그림 1의 구조를 갖으며 복잡도의 감소를 위해서 참조 계층이 '화면내 예측'으로 부호화된 경우에 한해서 매크로 블록(macro block) 단위로 수행되어 진다. 이와 같은 'intra base' 모드에서 하위 계층의 잔류 신호가 예측 신호로 사용되기 위해서는 화소 단위까지 복호되어 디블록킹(De-Blocking) 필터와 상향 표본화를 과정을 거쳐야 한다.[4]

상향 표본화 과정에서는 보간을 위하여 16단계의 위상을 갖는 poly-phase FIR 필터를 사용한다. 이때 휘도 성분은 4-tap 필터를 색차 성분은 2-tap의 필터를 사용하며 다음의 일반화된 식으로 나타 낼 수 있다.

$$interpolation(E_x, E_y) = \frac{hsum(E_x, E_y) + vsum(E_x, E_y)}{2^{10}}$$

$$hsum(E_x, E_y) = \sum_{n=i}^j (eF(phase, n) \times I(dip(B_{x+n}, Width), B_y))$$

$$vsum(E_x, E_y) = \sum_{m=i}^j (eF(phase, m) \times I(B_x, dip(B_{y+m}, Height)))$$

$$dip(x, y) = \begin{cases} 0, & \text{if } x \leq 0 \\ x, & \text{otherwise} \\ y, & \text{if } y \geq 0 \end{cases}$$

여기서  $E_x, E_y$ 는 보간 하려는 항상 계층의 수평과 수직의 좌표이고,  $B_x, B_y$ 는 참조 계층에서의 대응 좌표이다. 또한  $Width, Height$ 는 참조 계층 영상의 폭과 높이이고  $eF$ 는 가중치 계수이다.[4]

### 3. 보간법에 따른 부호화 성능 분석

본장에서는 휘도 성분의 상향 표본화에 2-tap, 4-tap, 6-tap의 poly-phase 필터를 적용하여 그에 따른 부호화 성능을 분석 한다. 시험을 수행한 플랫폼 환경은 Intel Core(TM)2 Duo 3.00GHz CPU, 2048MB RAM, Windows XP Professional OS 이고 사용한 코덱은 SVC 표준회의 레퍼런스 소프트웨어인 JSVM(Joint Scalable Video Model) Ver15.0 이다. JSVM 부호화

기에 기존의 2-tap, 4-tap poly-phase 필터 이외에 6-tap poly-phase 필터를 추가하였으며 6-tap poly-phase 필터를 위한 계층 테이블의 구성은 JVT-R018-AnnexG.doc[5] 문서를 참고하였다.

실험은 표준 영상인 Soccer, City, Crew 3가지 영상을 대상으로 수행하였으며 12에서 40범위의 QP를 2간격으로 증가시키면서 3단계의 공간 계층을 가지도록 부호화 하였다. 복잡도는 보간법 수행 과정에서 소비되는 시간을 uS 단위로 측정하였고, 비트율과 PSNR을 이용하여 효율성을 판단하였다. 표 1은 실험에 사용한 JSVM의 설정으로 모든 항상 계층의 영상을 'intra base'모드로 부호화 하기위해 GOP, IDRPeriod 그리고 InterLayerPeriod를 1로 설정 하였다.

#### 3.1 알고리즘 복잡도

JSVM은 한 프레임의 BL 영상을 EL과의 공간 비율로 보간하기 위해서  $((W \text{ of } EL * H \text{ of } EL) + (W \text{ of } EL * H \text{ of } BL)) * tap$  만큼의 MAC(multiply and accumulate) 연산을 수행한다. 여기서  $W$ 는 영상의 너비,  $H$ 는 영상의 높이이다. 표 2는 보간법의 복잡도에 대한 실험결과로서 QCIF와 CIF 해상도의 영상을 각각CIF와 4CIF의 해상도로 상향 표본화 할 때 보간법에 따른 MAC연산의 연산 횟수와 수행 시간이다. 측정된 수행 시간으로부터  $T_0(N) = 1.265 * N + 2.85, T_1 = 4.928 * N + 11.68$ 의 입력 탭수에 대한 실행 시간 함수를 구할 수 있으며 실행 시간 함수에서 영상의 크기에 따라 좌우되는 상수 성분을 제외 하면 보간법의 복잡도는 입력 탭수의 비율을 가지며 선형적으로 증가함을 확인할 수 있다

#### 3.2 부호화 효율 분석

부호화 효율 분석을 위한 비트율과 PSNR을 표 3과 그림 2 그림 3에 정리 하였다. 그림 2은 실험 영상 중 가장 큰 비트율 변화를 보인 표준 영상 Soccer의 항상 계층 0에서의 비트율 왜곡 곡선으로 2-tap 필터는 약간의 비트율 변화를 보이나 4-tap 필터와 6-tap 필터간은 비트율 변화를 거의 보이지 않고 있다. 때문에 6-tap 필터를 제거하고 가장 큰 비트율 변화 폭을 보이는 40~140kbit/s 구간에서의 비트율 왜곡

표 1. JSVM의 실험 환경

Table 1. condition of experiment in JSVM

Frame Rate	30	GOP	1
FramesToEncoded	90	IDRPeriod	1
NumLayers	3	QP min	12
SymbolMode	1	QP max	40
InterLayerPeriod	1	QP increment	4

표 2. 보간법의 복잡도에 대한 실험 결과

Table 2. Experimental results in complexity of interpolation

complexity		Computation	running time
		Quantities	(mS)
EL0-CIF	2-tap	304128	5.38
	4-tap	608256	7.93
	6-tap	912384	10.44
EL1-4CIF	2-tap	1216512	21.54
	4-tap	2433024	31.42
	6-tap	3649536	41.25

표 3. 보간법에 따른 부호화 효율

Table 3. Performance according to the interpolation filter

PSNR & Bitrates Suquence	2-tap		Bit saving (%)	4-tap		Bit saving (%)	6-tap	
	bit-rate (kbit/s)	Y-PSNR (db)		bit-rate (kbit/s)	Y-PSNR (db)		bit-rate (kbit/s)	Y-PSNR (db)
Soccer_CIF(EL0)	4496.0027	39.1001	-1.83	4413.6107	39.0942	-0.39	4396.5840	39.0940
City_CIF(EL0)	5483.0187	38.9465	-1.82	5383.2960	38.9304	-0.42	5360.5387	38.9267
Crew_CIF(EL0)	3130.2187	40.8349	-1.54	3082.1360	40.8222	-0.32	3072.1867	40.8307
Soccer_4CIF(EL1)	16329.1227	39.5060	-1.52	16080.3013	40.0506	-0.38	16019.3360	40.0505
City_4CIF(EL1)	20286.6293	40.0521	-1.34	20015.0667	39.5019	-0.39	19937.1627	39.4988
Crew_4CIF(EL1)	9885.3653	40.7726	-1.16	9770.6987	40.7795	-0.23	9748.5333	40.7785

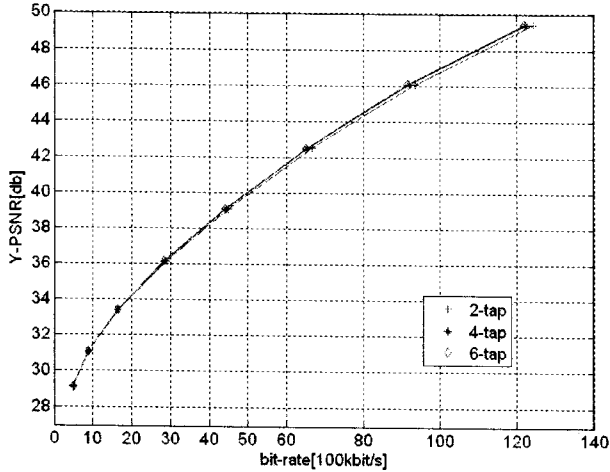


그림 2. Soccer\_CIF 비트율 왜곡 곡선

Fig.2. R-D curve for Soccer\_CIF

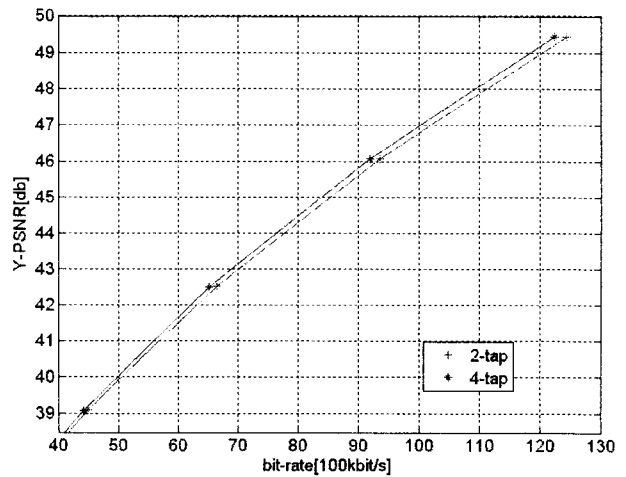


그림 3. Soccer\_CIF 비트율 왜곡 곡선(비트율 40~130 구간)

Fig.3. R-D curve for Soccer\_CIF(bit-rate 40 ~ 130)

곡선을 그림 3에 도시한다.

표3은 QP 24로 부호화했을 때의 부호화 효율로서 2-tap 필터와 4-tap 필터간의 비트율은 1.16%에서 1.83% 사이의 감소율을 가지어 평균적으로 1.53%의 비트율 절감 효과를 얻을 수 있다. 또한 6-tap 필터는 4-tap 필터에 비해 최대 0.42% 이하의 비교적 작은 비트율 절감 효과를 가지고 있다. 이러한 6-tap 필터의 과중한 연산량에 비해 얻을 수 있는 미미한 비트율 절감 효과는 6-tap 필터가 SVC 표준화 과정에서 제외된 원인으로 분석 된다.

#### 4. 결론

본 논문에서는 최근 국제 표준화가 완료된 H.264/SVC 부호화기의 '계층간 화면내 예측'에 있어서 적용되는 보간법에 따른 부호화 성능을 분석하였다. 현재 SVC 표준은 상향 표본화에 있어서 휘도 성분은 4-tap poly-phase FIR 필터, 색차 성분은 2-tap poly-phase FIR 필터를 사용해 보간하고 있으며 6-tap 필터가 제외된 원인을 분석 결과를 통해 찾을 수 있다. 또한 4-tap 필터가 2-tap 필터에 비해 얻을 수 있는 부호화 효율도 입력 탭수의 비율을 가지며 선형적으로 증가하는 연산량에 비하면 그리 크다고 판단되지는 않는다.

제한적인 하드웨어 성능을 가지는 임베디드(embedded) 시스템에서 4-tap 필터의 이러한 연산량은 SVC의 도입을 방해하는 요인으로 작용할 수도 있을 것이다. 따라서 프로파일과 레벨에 따른 2-tap 필터와 4-tap의 필터의 선택적 적용이 고

려되어야 할 것이다.

#### 참고 문헌

- [1] Thomas Schierl, Thomas Stockhammer, Thomas Wiegand, "Mibile Video Transmission Using Scalable Video Coding", IEEE, vol. 17, no. 9, pp 1204-1217, 2007
- [2] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Trans. on Circuits and Video Tech., vol.17, no.9, pp.1103-1120, 2007.
- [3] D. W. Song and C. W. Chen, "Scalable H.264/AVC Video Transmission Over MIMO Wireless System With Adaptive Channel Selection Based on Partial Channel Information," IEEE Trans. on Circuits and Video Tech., vol.17, no .9, pp.1218-1226, 2007.
- [4] T.Wiegand, G. Sullivan, J. Reichel, H. Schwarz, and M.Wien, "Joint Draft ITU-T Rec. H.264 | ISO/IEC 14496-10 / Amd.3 Scalable video coding," Joint Video Team, Doc. JVT-X201, July 2007
- [5] J. Reichel, H. Schwarz, and M. Wien, "Scalable Video Coding-Working Draft", JVT-R018-AnnexG.doc, 2006
- [6] 정차근, 김대남, "C-모델 시뮬레이터 기반 H.264/SVC 복호기 시스템 구현", 한국콘텐츠학회 논문지, vol.9, no.2, 2009.