

스케일러블 비디오 코딩에서 방향성 정보를 이용한 모드 결정 고속화 기법

정 현 기^{a)}, 홍 광 수^{a)}, 김 병 규^{a)†}, 김 창 기^{a)}, 유 정 주^{b)}Fast Mode Decision Algorithm for Scalable Video Coding (SVC)
Using Directional Information of Neighboring LayerHyun-Ki Jung^{a)}, Kwang-Soo Hong^{a)}, Byung-Gyu Kim^{a)†}, Chang-Ki Kim^{a)}, and Jeong-Ju Yoo^{b)}

요 약

스케일러블 비디오 코딩은 동영상 압축 표준인 H.264/AVC에서 확장된 표준으로써 시간, 공간, 화질적인 측면에서 확장성을 제공하는 부호화 방법이다. 여러 측면에서 제공하는 확장성으로 압축에 대한 효율 등은 좋아졌지만, 상대적으로 복잡도의 큰 증가로 인해 효율적인 사용을 위해서는 복잡도의 감소가 필수적이다. 본 논문에서는 스케일러블 비디오 코딩의 복잡도를 감소시키기 위해서, B-화면에서의 방향성 정보를 이용하여 계층 간 예측을 효율적으로 수행함으로써 부호화 과정의 복잡도를 크게 감소시킬 수 있는 모드 결정 고속화 기법을 제안한다. 제안하는 방법은 우선 업-샘플링(Up-sampling)된 하위 계층의 최적 모드와 이웃한 매크로블록의 최적 모드 정보를 이용하여 간단하게 SKIP 모드를 검색한다. 그런 후 현재 수행하는 매크로블록의 예측에 이웃한 매크로블록의 최적 모드를 사용하게 되는데, 이 때 업-샘플링된 하위 계층의 최적 모드의 방향성 정보가 순 방향과 역 방향, 양쪽 방향에 따라서 검색 모드 수를 달리하는 방법으로 모드 결정을 고속화하는 방법이다. 이러한 고속화 방법을 수행함으로써 얻은 실험 결과는 SVC의 부호화 과정에서 소요되는 시간을 참조 소프트웨어의 원본 대비 53%까지 감소시키며 이에 따른 비트율의 증가나 화질의 열화는 무시할 수 있을 정도로 적음을 실험을 통해 확인한다.

Abstract

As Scalable Video Coding (SVC) is a video compression standard extended from H.264/AVC, it is a way to provide scalability in terms of temporal, spatial and quality. Although the compression efficiency of SVC is increased due to the scalability in many aspect, it is essential to reduce the complexity in order to efficiently use because the complexity is relatively increased. To reduce the complexity of SVC in the paper, we propose fast mode decision algorithm to reduce the complexity of encoding process using direction information of B-picture by efficiently performing inter-layer prediction. The proposed algorithm is a fast mode decision algorithm that makes different from detection mode number of forward and backward, bi-direction in the way using best mode of base-layer up-sampled after simply SKIP mode detection using the direction information of best mode of base-layer up-sampled. The experimental results show that the proposed algorithm approach can achieve the maximum computational time saving about 53% with almost no loss of rate distortion (RD) performance in the enhancement layer.

Keyword : Scalable video coding, Fast mode decision, H.264/AVC, inter-layer prediction, directional information

a) 선문대학교 컴퓨터공학과

Department of Computer Engineering, SunMoon University, Asan Chungnam

b) 한국전자통신연구원

Smart Television (TV) Research Department, ETRI, Daejeon

† 교신저자 : 김병규 (bg.kim@mpcl.sunmoon.ac.kr)

※ 본 연구는 방송통신위원회의 ETRI 연구개발지원사업의 일환으로 수행되었음 [11921-03001, "Beyond 스마트TV 기술개발"].

· 접수일(2011년10월28일), 수정일(2011년12월15일), 게재확정일(2011년12월18일)

I. 서론

스케일러블 비디오 코딩 (SVC)은 H.264/AVC 표준에서 확장된 표준으로써 시간, 공간 및 화질적인 측면에서의 확장성을 제공하는 부호화 방법이다. 기존의 MPEG-4, H.264/AVC 등의 압축 표준은 단일 계층으로 부호화를 진행하기 때문에 서로 다른 환경에서 효율적으로 동작하는데에 적합하지 않다는 단점이 있다. 이에 SVC는 시간, 공간, 화질의 세 가지 확장성을 제공함으로써 한 번의 부호화만으로 서로 다른 환경 및 기기에 따라 각 환경에 맞는 품질을 제공 할 수 있는 동영상 압축 표준이다. SVC는 하나의 기본 계층 (base layer) 과 1개 혹은 여러 개의 향상 계층 (enhancement layer)로 구성되며 각 계층 별로 시간, 공간 및 화질의 확장성을 갖도록 부호화 할 수 있다. SVC의 기본 계층은 H.264/AVC 표준과 상호 호환 적으로 부호화가 가능하며, 향상 계층은 공간적 확장성을 지원하고 하위 계층의 정보를 이용하여 예측하는 계층 간 예측 (inter-layer prediction) 방법을 사용하여 압축 성능을 올렸다. 또한 시간적인 확장성을 지원하기 위해서 계층적 B-화면 (hierarchical B-picture) 구조를 사용하는데, 이 기술은 다양한 시간적인 확장성을 지원하기 위해 주화면 사이에 B-화면을 사용하는 방법이다.

SVC 에서는 H.264/AVC 표준에서 사용한 모든 예측 방법을 모두 사용하고, 추가적으로 하위 계층에서 수행 된 예측 데이터를 기반으로 수행하는 계층 간 예측이 추가되기 때문에 기존 표준에 비해서 복잡도의 증가는 피할 수 없다. 따라서 SVC 인코더의 효율적인 사용을 위해서는 복잡도의 감소가 필수적이다. 이를 위해서 복잡도 감소를 위해 많은 연구가 진행되고 있는 실정이다.

He Li et al^[3]은 공간적인 확장성에 대한 복잡도 개선을 위해 향상 계층에서 화면 간 모드 결정 시 하위 계층에 대한 모드를 판별하여 향상 계층에서 후보 모드를 줄여 감으로써 모드 결정을 고속화 할 수 있는 기법을 제안하였다. 예를 들어, 기본 계층에서 구한 최적 모드를 업-샘플링을 거쳐 얻어진 모드가 16x8 이라고 하면 향상 계층의 동일한 영역에서는 16x8과 8x8만을 후보 모드로 선택하여 예측하고 모드를 결정하는 방법을 수행하여 부호화의 시간을 단축하였

다.

또 다른 방법으로 공간적으로 이웃한 매크로블록 (macro-block)이 Skip 모드인 블록을 찾거나^[4] 율-왜곡 값을 이용하여^[5] 적응적으로 모드 결정을 수행하는 방법으로 부호화의 시간을 단축하는 방법들을 제시하였다. 또한 H.264/AVC 에서는 B-화면의 화면 간 예측을 진행할 때 16x16 모드의 방향성의 SAD (sum of absolute difference) 값에 따라서 각 방향의 후보 모드를 달리 함으로서 복잡도를 감소하는 방법을 사용하였다^[6].

본 논문에서는 B-화면의 예측 시에 하위 계층의 최적 모드 정보를 이용하여 후보 모드를 줄이고, 하위 계층에서의 최적 모드의 방향성을 이용하여 각 방향성에 후보 모드의 수를 차별화함으로써 압축 효율을 증가시키고 복잡도를 감소시키는 방법을 제시한다.

본 논문은 다음의 구성은 다음과 같다. 2장에서는 SVC의 시간적, 공간적 확장성을 제공하는 기술에 대한 소개를 하고, 3장에서는 본 논문에서 사용하고자 하는 기본 계층의 최적 모드와 그 방향성이 향상 계층과의 연관성을 살펴본다. 4장에서는 살펴본 정보를 토대로 하는 방향성 정보 기반의 모드 결정 고속화 기법을 소개하고, 5장에서는 제안된 기법의 실험 결과에 대해 살펴본다. 마지막으로 6장에서는 결론에 대한 내용을 끝으로 본 논문을 마친다.

II. SVC 에서의 시간적 확장성 (temporal scalability) 및 공간적 확장성 (spatial scalability)

1. SVC 에서의 시간적 확장성 (temporal scalability)

H.264/AVC 표준에서 각 프레임의 종류는 수행에 사용하는 예측 화면의 종류에 따라 I-화면, P-화면, B-화면 총 3가지의 화면으로 나뉘게 된다. 여기서 I-화면은 화면 간 예측을 수행하지 않고 화면 내 예측만을 수행하는 화면이고, P-화면은 이전 화면 즉 L0의 방향으로만 예측을 수행하는 화면이다. 마지막으로 B-화면은 그림 1에서 나타나는 것처럼 이전 화면과 다음 화면 즉 L0과 L1의 방향을 전부

예측에 사용하는 화면이다. 이때의 예측 방향성은 L0과 L1을 각각 한쪽만 사용하는 단방향과 L0, L1 양쪽을 모두 사용하는 양방향(Bi-direction) 이렇게 세 가지의 방향성이 존재한다. 본 논문에서는 L0 방향에 대해 ‘순방향’ L1 방향에 대해 ‘역방향’으로 지칭하여 사용하도록 하였다. 그리고 L0 방향과 L1 방향을 모두 사용하는 방향에 대해서는 ‘양방향’이라고 하였다. 다음 그림 1은 L0 과 L1 의 예측 방향을 나타내고 있는 그림이다.

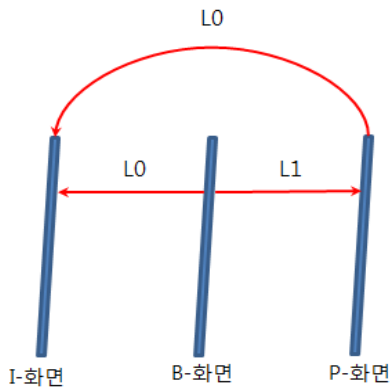


그림 1 L0과 L1의 예측 방향
Fig. 1 Predicted direction of L0 and L1

SVC에서는 시간적인 확장성의 측면에서 기본적으로 계층적 B 구조를 사용하도록 권장한다. SVC가 갖는 구조적인 특징은 다음 그림 2에서 잘 나타내고 있다.

계층적 B 구조는 I 프레임과 P 프레임의 예측을 수행한 후에 GOP (group of picture) 사이즈에 따라 양쪽에서 가장 먼 가운데 프레임을 B 프레임으로 먼저 예측을 수행하고 같은 방법으로 그림 1의 b1 -> b2 -> B3 순서로 예측을 수행하는 방법을 의미한다. 이 구조는 다양한 시간 해상도를 지원하기 위한 구조이며 각각의 시간적 확장성에 따른 분해 레벨을 갖는다. 이 구조의 화면에서는 주 화면을 구성하는 I 및 P 프레임은 GOP 내의 모든 화면의 참조 화면이 될 수 있으므로 가장 좋은 화질을 유지해야 하고 시간적 확장성의 레벨이 증가 할수록 참조 될 수 있는 화면의 수가 적어지므로 중요도가 감소하게 된다. 이러한 이유로 SVC에서는 각 시간적 확장 레벨에 따라 서로 다른 QP 값을 적용하여 한 GOP 내의 비트 량을 줄일 수 있다. 즉, 주 화면을 구성하는 시간적 확장 레벨이 0인 프레임에 대해서는 작은 QP 값을 적용하여 높은 화질을 유지할 수 있도록 하고 시간적 확장 레벨이 증가 할수록 QP 값이 증가하여 시간적 확장 레벨에 따른 화질의 중요도에 따라서 비트 율을 감소

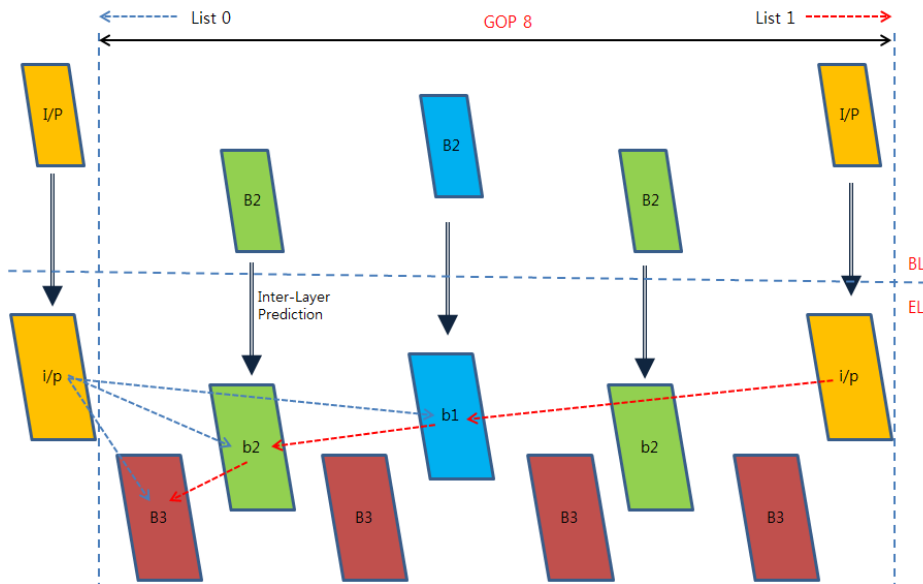


그림 2 SVC의 구조
Fig. 2 Structure of SVC

시킬 수 있다.

2. SVC 에서의 공간적 확장성 (spatial scalability)

SVC에서는 공간적 확장성을 지원하기 위해서 하위 계층으로부터 예측하여 차 상위 계층에서 예측 신호로 사용하게 된다. 이 때 하위 계층의 대응되는 블록이 화면 간 예측으로 수행 되었을 경우 상위 계층의 블록은 하위 계층의 움직임 벡터와 참조 화면 리스트를 이용하여 상위 계층에서 움직임 보상을 수행할 수 있다. 만약 이 움직임 벡터와 참조 화면 리스트가 최적으로 판단되어 진다면 상위 계층의 현재 매크로블록은 헤더에 움직임 벡터, 참조화면 리스트, 모드 정보 등을 부호화 하지 않아도 되기 때문에 비트를 절약 할 수 있다는 장점이 있다.

그림 3은 계층 간 움직임 정보 예측 시 매크로블록의 분할 정보와 움직임 벡터가 업-샘플링 되는 것을 나타낸다.

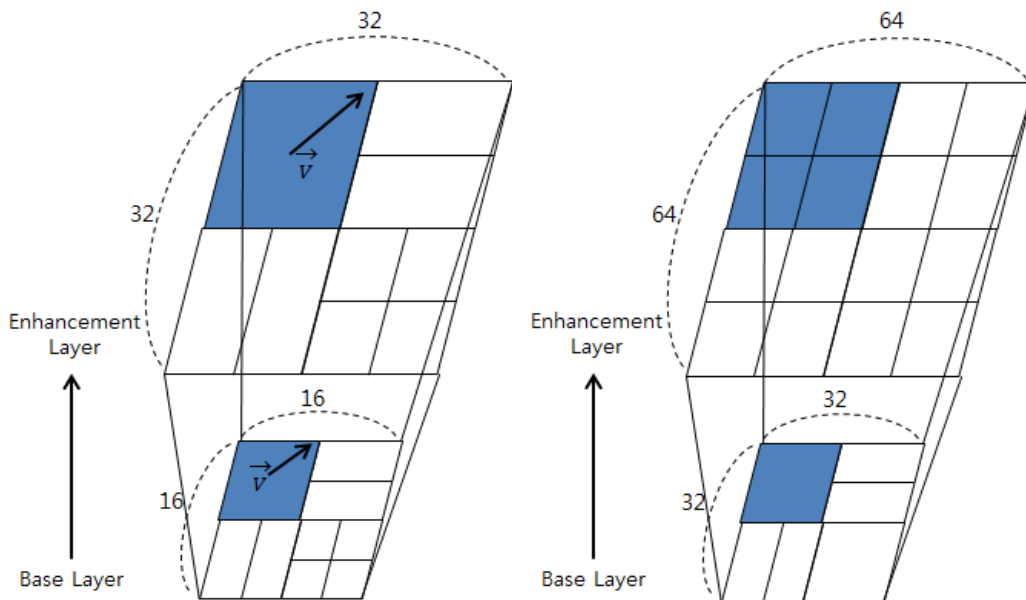
(a)는 하위 블록이 4x4 및 8x8까지의 서브 매크로블록 모드로 분할되었을 때 서브 매크로블록들은 상위 계층에

16x8, 8x16, 16x16 모드로 각각 매핑 되어 지고 하위 계층의 모드가 16x8, 8x16, 16x16 일 경우는 모드 16x16 모드로 상위 계층에 매핑 되게 된다. 이 때의 움직임 벡터는 하위 계층의 움직임 벡터에 각각 두 배를 곱하여 표현되어 진다.

3. SVC 에서의 율-왜곡 최적화 (Rate-Distortion Optimization) 기반 모드 결정

SVC는 H.264/AVC 표준에서 확장된 부호화 방법이기 때문에 기본 계층은 H.264/AVC 기술과 상호 호환성을 유지한다. 따라서 SVC의 기본 계층에서의 블록 모드는 화면 간 예측에서 H.264/AVC 표준과 동일하게 수행된다. 다음 그림 4는 화면 간 예측 시 7가지의 서로 다른 블록 모드를 나타낸 것이다.

또한 화면 내 예측의 경우도 H.264/AVC 기술과 동일하게 4x4 블록에 대하여 9가지의 예측 신호를 갖고, 16x16에 대해서는 4가지의 예측 신호를 갖고, 8x8 역시 4가지의 예측 신호를 갖는다.



(a) 4x4~8x8블록의 업-샘플링
(a) Up-sampling for 4x4~8x8 blocks

(b) 16x16,16x8,8x16 블록의 업-샘플링
(b) Up-sampling for 16x16, 16x8, 8x16 blocks

그림 3. 업-샘플링 한 블록과 움직임 벡터
Fig. 3. The up-sampled blocks and motion vectors

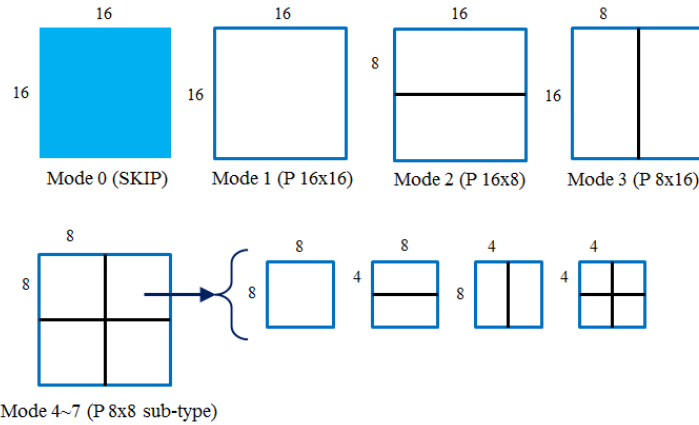


그림 4 화면 간 예측 시 7가지 서로 다른 블록 모드
 Fig. 4 7 Different block modes for inter-frame prediction

최적의 블록 모드의 결정은 모든 블록에 대하여 율-왜곡 (rate-distortion) 값을 계산하여 최소의 값을 갖는 것을 최적의 모드로 결정한다. 다음 식 (1)은 최적 모드를 결정하기 위한 율-왜곡 함수를 나타낸다.

$$J(s,c,Mode|QP,\lambda_{Mode}) = SSD(s,c,Mode) + \lambda_{Mode} \cdot R(s,c,Mode|QP) \quad (1)$$

여기서 s는 원 영상, c는 복원된 영상을 의미하며, QP는 모드 결정 시 사용되는 양자화 계수를 뜻한다. λ_{Mode} 는 라그랑지 계수로서 $0.85 \times 2(QP-12)/3$ 와 같이 계산된다. SSD (sum of squared differences)는 원 영상과 복원된 영상 간의 왜곡 도를 나타내며, 값이 클수록 왜곡이 심하다.

향상 계층 (enhancement layer)에서는 위에서 설명한 블록들 이외에 하위 계층으로부터 계층 간 예측 모드를 수행하여 후보 모드가 추가되기 때문에 율-왜곡 값을 이용해서 모드를 결정할 때 복잡도가 급격하게 증가 하게 된다.

III. 기본 계층과 향상 계층 간 방향성 및 모드 정보의 연관성

1. 기본 계층과 향상 계층 간 모드 정보의 연관성

SVC의 기본 계층과 향상 계층은 시간적, 공간적, 화질적

으로 제공되는 확장성의 차이만 둔 동일한 영상을 부호화

표 1. 기본 계층과 향상 계층 간 모드 정보의 연관성 (%)
 Table 1. The relevance of mode information between the base layer and enhancement layer (%)

Bus		업-샘플링 된 기본 계층의 최적 모드			
		16x16	16x8	8x16	8x8
향상 계층의 최적 모드	SKIP	26.30	0.84	0.00	0.00
	16x16	44.08	2.80	4.44	5.68
	16x8	10.16	37.82	4.81	3.41
	8x16	8.74	3.64	37.04	6.82
	8x8sub	10.72	54.90	53.70	84.09
Foreman		업-샘플링 된 기본 계층의 최적 모드			
		16x16	16x8	8x16	8x8
향상 계층의 최적 모드	SKIP	49.20	0.00	0.00	0.00
	16x16	35.07	1.98	0.00	0.00
	16x8	4.92	40.59	2.63	0.00
	8x16	5.93	1.98	71.05	0.00
	8x8sub	4.88	55.45	26.32	100.00
Football		업-샘플링 된 기본 계층의 최적 모드			
		16x16	16x8	8x16	8x8
향상 계층의 최적 모드	SKIP	16.58	0.00	0.11	0.00
	16x16	42.00	3.07	4.21	2.58
	16x8	12.42	44.33	4.32	3.15
	8x16	14.56	5.79	50.11	8.02
	8x8sub	14.43	46.81	41.25	86.25
CITY		업-샘플링 된 기본 계층의 최적 모드			
		16x16	16x8	8x16	8x8
향상 계층의 최적 모드	SKIP	51.48	0.00	1.64	0.00
	16x16	35.56	9.09	1.64	0.00
	16x8	2.68	18.18	1.64	0.00
	8x16	5.10	9.09	60.66	0.00
	8x8sub	5.18	63.64	34.43	100.00

하는 것이기 때문에 계층 간에 상당한 연관성이 있다. 앞에서 설명 했듯이 기본 계층이 화면 간 예측을 하게 되면 업-샘플링을 통해서 향상 계층에 적용되는데 표 1은 기본 계층의 최적 모드를 업-샘플링 한 후 해당 모드에 대해 향상 계층의 최적 모드의 분포를 나타낸 것이다.

업-샘플링 시 8x8 보다 큰 모든 블록이 16x16 으로 매핑 되기 때문에 기본적으로 업-샘플링 된 기본 계층의 최적 모드는 16x16 이 가장 많이 존재한다. 또한 표 1에서 살펴보면 향상 계층의 최적 모드는 업-샘플링 된 기본 계층의 최적 모드와 같은 경우 많이 나타나는 것을 볼 수 있다. foreman 영상의 업-샘플링 된 기본 계층의 최적 모드가 16x8인 경우와 같이 향상 계층의 최적 모드가 업-샘플링 된 기본 계층의 최적 모드의 확률과 다른 모드에서 가장 높은 확률로 나타나는 것을 볼 수 있는데 이것의 이유는 8x8sub 모드가 1가지의 모드가 아니고, 향상 계층은 기본 계층보다 화질이 좋기 때문에 좀 더 세밀한 부분이 많이 나타날 수 있기 때문이다. 이러한 연관성을 이용하여 알고리즘에 적용한다면 보다 효율적인 부호화가 가능 할 것이다.

2. 기본 계층과 향상 계층 간 방향성 정보의 연관성

앞 절에서 설명한 것과 마찬가지로 모드 정보 이외에 방향성 정보에도 상당한 연관성을 갖는다. 표 2는 업-샘플링 된 기본 계층의 최적 모드에서 방향성이, 동일한 위치의 향상 계층 최적모드의 방향성과 일치 할 확률을 나타낸 표이다. 실험에 사용된 영상의 양자화 계수는 기본 계층이 32 향상 계층이 26 이며 해상도는 기본 계층이 176x144의

표 2. 업-샘플링 된 기본 계층의 최적 모드의 방향성이 향상 계층의 방향성과 일치할 확률 (%)

Table 2. Probability that the direction of the best mode of the up-sampled base layer corresponds to that of enhancement layer (%)

영상이름	업-샘플링 된 기본 계층의 최적 모드			
	16x16	16x8	8x16	8x8sub
CITY	64.044	64.286	81.818	40.000
football	60.620	74.684	76.687	83.274
foreman	67.483	77.895	79.167	82.353
bus	61.074	67.456	83.012	82.105

QCIF 영상, 향상 계층이 352x288의 CIF 영상을 사용하였다.

표 2를 통해 살펴보면 기본 계층의 최적 모드에 대해 상당히 높은 확률로 방향성이 일치하는 것을 확인 할 수 있다. 그러나 업-샘플링 된 기본 계층의 최적 모드는 기본적으로 기본 계층의 최적 모드가 8x8 모드보다 큰 크기를 가질 때 모두 16x16으로 매치되기 때문에 16x16 모드에 대부분의 빈도수가 집중되고 나머지 모드에 대해서는 그 경우의 수가 상당히 적다. 때문에 이를 효과적으로 활용하기 위해서는 추가적인 정보가 더 필요하다. 본 논문에서는 현재 매크로블록의 주변 블록들의 정보들을 추가적으로 이용하여 위의 데이터를 좀 더 효과적으로 사용하도록 하였다.

표 3은 현재 매크로블록의 주변 블록들의 방향성이 업-샘플링 된 기본 계층의 최적 모드의 방향성과 일치할 확률을 나타낸 것이다.

표 3. 업-샘플링 된 기본 계층의 최적 모드의 방향성이 현재 매크로블록의 주변 매크로블록들과 방향성과 일치할 확률 (%)

Table 3. Probability that the direction of the best mode of the up-sampled base layer corresponds to that of neighboring macroblock of current macroblock (%)

		업-샘플링 된 기본 계층의 최적 모드			
		16x16	16x8	8x16	8x8sub
City	Left	65.9	45.91	39.67	46.18
	Above	66.63	50.35	52.57	53.62
football	Left	57.93	46.61	43.73	38.11
	Above	56.68	48.62	54.02	40.59
foreman	Left	64.46	43.67	43.33	48.34
	Above	63.14	39.22	43.23	45.31
bus	Left	63.61	51.86	45.62	42.27
	Above	61.47	53.67	55.22	46.72

표 3의 데이터만을 살펴보면 39.67% ~ 66.63% 까지 그리 크지 않은 확률이 나타나는 것을 볼 수 있다. 그러나 해당 방향성이 일치하는 경우에 대해서 향상계층의 방향성 또한 일치하는 경우는 표 4를 통해 확인 할 수 있다. 다음 표 4는 표 3에서 나타난 결과의 경우에 현재 매크로블록의 방향성이 일치할 확률을 나타내고 있다. 표 4를 통해 살펴보면 대체적으로 70%가 넘는 확률이 나타나는 것을 알 수

있다. 이런 경우 방향성에 따라 예측모드를 달리 하는데 효율적이라고 하겠다.

표 4. 업-샘플링 된 기본 계층의 최적 모드의 방향성이 현재 매크로블록의 주변 매크로블록들의 방향성과 일치할 때 현재 매크로블록의 방향성이 일치할 확률(%)

Table 4. Probability that the direction of current macroblock corresponds when the direction of the best mode of the up-sampled base layer corresponds to that of neighboring macroblock of current macroblock (%)

		16x16	16x8	8x16	8x8
CITY	Left	74.33	77.78	80.39	75.00
	Above	75.41	66.67	84.31	75.00
Football	Left	73.23	76.51	72.34	61.86
	Above	72.03	73.63	66.10	66.67
Foreman	Left	59.83	63.22	75.00	71.43
	Above	61.45	67.82	64.29	64.29
Bus	Left	76.31	66.80	72.69	77.92
	Above	75.22	73.05	76.21	66.23

표 5는 향상 계층의 최적 모드들의 분포를 나타낸 것이다. 이 표에서는 SKIP 및 16x16모드가 상당히 높은 비중을 차지하는 것을 알 수 있으며, 영상에 따라 그 차이는 있지만 작은 크기의 모드로 갈수록 낮은 비중을 차지하는 것을 확인할 수 있다.

이러한 정보를 기반으로 본 연구에서는 화면 간 예측에서 기본 계층의 방향성 정보 기반의 모드 결정 고속화 기법을 제안한다.

표 5. 향상 계층의 최적 모드의 분포(%)
Table 5. Distribution of Best Mode in Enhancement Layer(%)

	SKIP	16x16	16x8	8x16	8x8sub
bus	27.03	40.03	11.10	9.01	12.82
football	21.74	36.57	13.06	12.51	16.12
foreman	43.16	38.36	6.89	6.64	4.94
city	47.58	37.65	3.97	6.22	4.58
crew	22.10	43.19	10.27	14.87	9.56

IV. 제안하는 방향성 정보 기반의 모드 결정 고속화 기법

앞 장을 통해서 기본 계층과 향상 계층 간의 모드 정보가 상당한 연관성이 있음을 확인하였다. 본 장에서는 기본 계층의 최적 모드 및 방향성 정보를 이용하여 모드 결정을 고속화하는 기법을 제안한다. 제안하는 기법의 전체 순서도는 다음 그림 5와 같다.

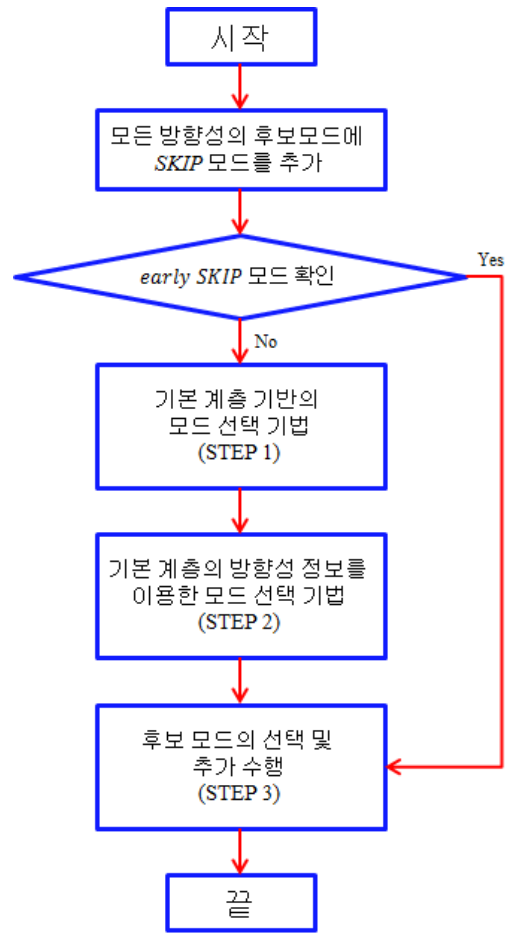


그림 5. 제안된 기법의 순서도
Fig. 5. Flowchart of the proposed algorithm

1. 기본 계층 기반의 모드 선택 기법

본 절에서는 기본 계층의 모드를 기반으로 후보 모드를

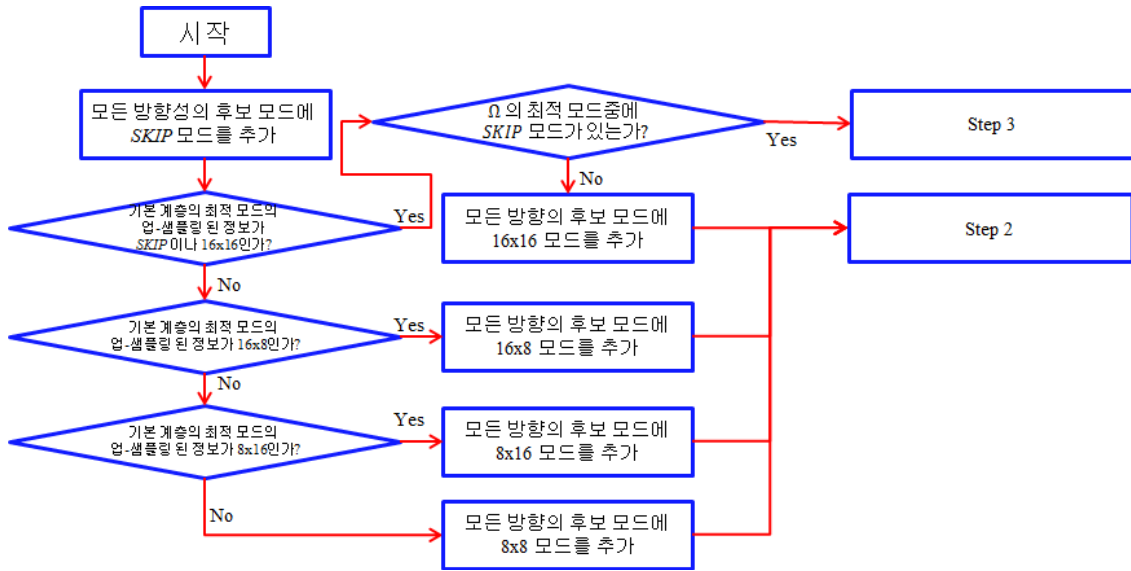


그림 6. 기본 계층의 모드를 기반으로 후보모드를 선택하는 기법의 순서도 (STEP 1)
 Fig. 6 Flowchart of selection of candidate modes based on the mode of the base layer (STEP 1)

선택 하는 기법을 소개한다. 그림 6은 기본 계층의 모드를 기반으로 후보 모드를 선택하는 기법의 순서도이다.

먼저 순, 역, 양쪽 의 모든 방향성의 후보 모드에 SKIP 모드를 추가한다. SKIP 모드를 추가하는 이유는 SKIP 모드는 기본적으로 계산 복잡도가 상당히 적고, 최적 모드로 평가된 많은 수가 선택 되어 지기 때문에 기본적으로 수행 되어도 속도 저하에 영향을 주지 않는다. 때문에 기본적으로 모든 후보 모드에 SKIP 모드를 추가한다.

그 후 같은 위치에 있는 기본 계층의 최적 모드가 업-샘플링 된 정보를 이용하여 후보 모드에 모드를 추가 하게 되는데 업-샘플링 된 데이터가 SKIP이나 16x16인 경우는 Ω의 최적 모드를 확인 한 후 그 안에 SKIP 모드가 있으면 그대로 SKIP 모드만을 후보 모드로 갖는다. 반대로 SKIP 모드가 없다면 모든 방향의 후보 모드에 16x16 모드를 추가하게 된다.

다음으로 업-샘플링 된 데이터가 SKIP이나 16x16이 아닌 경우에는 해당하는 모드를 모든 방향의 후보 모드에 추가한다. 이 과정에서 업-샘플링 된 데이터는 8x8 보다 작은 모드가 존재하지 않으므로 업-샘플링 된 데이터가 8x8이 된다면 8x8 이하 모든 서브 모드를 후보 모드에 추가한다.

2. 기본 계층의 방향성 정보를 이용한 모드 선택 기법

앞 절에서 결정 된 후보 모드에 기본 계층의 방향성 정보를 이용해서 각 방향성에 후보 모드를 차별적으로 추가함으로써 속도 향상과 화질의 열화를 최소화 하는 과정을 추가한다. 모드를 추가하기 위해 다음 식 2와 같이 집합 Ω을 정의 한다.

$$\Omega = \{ R_{(x,y,p)} \mid q, p = 0,1 (q+p=1) \}, \quad (2)$$

여기서 $R(x,y)$ 는 현재 매크로블록을 의미 하는 것으로 x, y 값은 각각 한 화면 내 에서 매크로블록의 위치를 좌표로 나타낸 것을 뜻한다. 즉 위의 식 2 은 현재 매크로블록의 이웃한 블록을 나타내는 것으로 q 가 1이면 현재 매크로블록의 왼쪽에 위치한 매크로블록을 의미하고, p 가 1이면 현재 매크로블록의 위쪽에 위치한 매크로블록을 의미한다. 앞서 3장에서 살펴봤듯이 기본 계층과 향상 계층의 방향성은 상당한 연관성을 갖기 때문에 이 정보를 이용하여 좀 더 많이 예측 되는 방향성 부분에만 모드를 후보 모드에

추가한다. 방법은 다음과 같은 순서를 따른다.

- i) 같은 위치의 기본 계층의 최적 모드 방향성을 확인한다.
- ii) Ω 의 최적 모드를 확인한다.
- iii) 기본 계층의 최적 모드의 방향성이 순 혹은 역방향이면 해당 방향성에만 ii)에서 확인한 최적 모드들을 후보 모드에 추가한다.
- iv) 기본 계층의 최적 모드의 방향성이 양방향이면 모든 방향의 후보 모드에 ii)에서 확인된 최적 모드를 추가한다.

그림 7은 방향성 정보를 이용한 모드 선택 기법의 순서도이다.

3. 후보 모드의 선택 및 추가 수행

1, 2절을 통해서 후보 모드들을 추가한 후에 추가된 모든 후보 모드들의 율-왜곡 값을 계산한다. 또한, Ω 의 율-왜

곡 값들의 평균값을 구하고 후보 모드들 중에서 계산된 가장 최소값을 갖는 율-왜곡 값과 비교하여 선택된 모드가 적절한지를 확인하는 과정은 다음과 같은 식 3에 의해 결정된다.

$$Avg. RD-cost_N < Best RD-cost_C \tag{3}$$

위 식 3에서 Avg. RD-costN 은 Ω 의 율-왜곡 값의 평균을 의미하며, Best RD-costC는 후보 모드들의 율-왜곡 값 중 최적의 값을 의미한다. 위 식의 조건이 맞으면 후보 모드에 따라 후보 모드에 SKIP 모드만 있다면, 추가적으로 큰 크기의 모드들만의 율-왜곡 값을 계산하고 그 중 최적 모드를 결정하고, 그 이외의 경우에 대해서는 추가적으로 모든 모드 검색 (full search)을 통해 최적 모드를 결정하게 된다. 반대로 위 식의 조건이 맞지 않다면 현재 결정되어있는 최소 율-왜곡 값을 갖는 모드를 최적 모드로 결정하게 된다.

다음 그림 8은 후보 모드들 중 최적 모드의 선택 및 추가 수행을 위한 기법의 순서도이다.

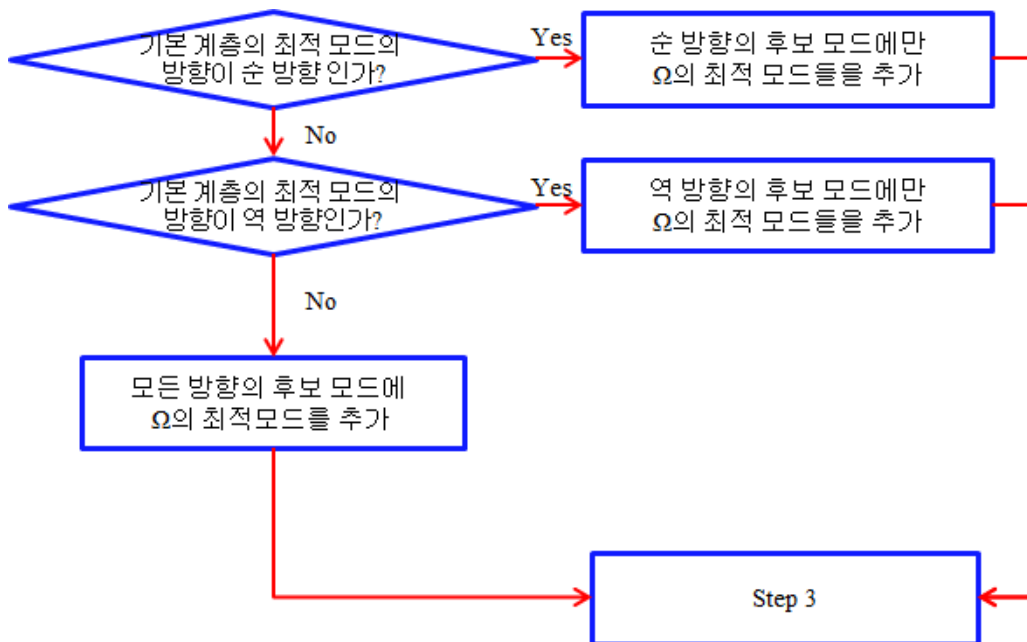


그림 7. 방향성 정보를 기반으로 후보 모드를 선택하는 기법의 순서도 (STEP 2)
 Fig. 7. Flowchart of selection of candidate modes based on the directional information of the base layer (STEP 2)

V. 실험 결과

다음 표 6은 제안된 알고리즘의 실험을 위해 JSVM 9.12^[7]에서 설정한 환경을 나타낸 것이다.

표 6. 부호화 파라미터
Table 6. Encoding Parameter

참조 화면의 수	1
GOP의 크기	8 (IBBBBBBP)
움직임 벡터 정밀도	1/4 화소
움직임 벡터 탐색 범위	32화소
사용 동영상	Foreman, Football, Bus, CREW
공간 계층의 수	2 (기본 계층 QCIF + 향상 계층 CIF)
컴퓨터 환경	Intel Core i5-2500 3.30GHz CPU 4GB 주 메모리 Windows7 64bit 운영체제

제안된 기법은 부호화 과정에서 연산량을 감소시켜 전체 부호화 시간을 줄이고, 그에 따른 비트의 증가나 화질의 열화는 최소로 하는 것을 목표로 한다. 따라서 제안된 기법의 부호화 시 수행 시간을 측정하고, 이것을 원 참조 소프트웨어 (JSVM) 보다 얼마만큼 감소했는지를 확인하였다.

또한 본 기법의 적용 전 후의 PSNR과 비트율의 변화량을 각각 dB와 %로 측정하였다. 시험 영상은 기본 계층에는

30fps의 QCIF 영상을 사용하였고, 향상 계층에는 30fps의 CIF 영상을 사용하였다. 또한, 양자화 계수 (quantization parameter)의 변화에 따른 성능을 확인하기 위해서 기본 계층의 양자화 계수가 32, 36, 40 일 때 향상 계층의 양자화 계수가 각각 26, 30, 34에 대한 성능을 테스트하였다. BD-rate 및 BD-PSNR을 구하기 위해서는 총 4가지의 양자화 계수 값들이 필요한데 이를 위해서 기본계층 28, 향상계층 22의 양자화 계수일 때의 수치를 추가하여 각각의 값을 구했다. 또한 실험에 사용된 영상은 모두 제공되는 영상의 마지막 화면까지 실험하였다.

다음 표 7은 제안된 알고리즘과 비교 논문인 Li의 기법^[3]의 BD-rate(%)와 BD-PSNR (dB)을 계산한 것이며, 표 8은 Li의 기법과 Δ Bit-rate, Δ PSNR, 부호화 감소 시간을 각각 계산한 것이다.

표 7의 가장 윗줄의 BQP는 기본 계층의 양자화 계수를 의미하고, EQP는 향상 계층의 양자화 계수를 의미하며 두 번째 줄의 TS는 원본 대비 시간의 감소율(Time Saving)을 나타낸다. 한다. 우선 표 7을 통해 살펴보면 제안된 기법의 부호화 속도는 원본 대비 최대 53% 감소하는데 반해 화질의 열화는 PSNR 대비 평균 0.06dB~0.11dB 정도로 낮고, 비트의 증가량 또한 평균 -1.18%~2.16% 정도로 무시할 만한 수준임을 알 수 있다. 또한 표 8을 통해 살펴보면, 모든 영상에 대해서 Li의 기법 보다 PSNR의 감소는 적고

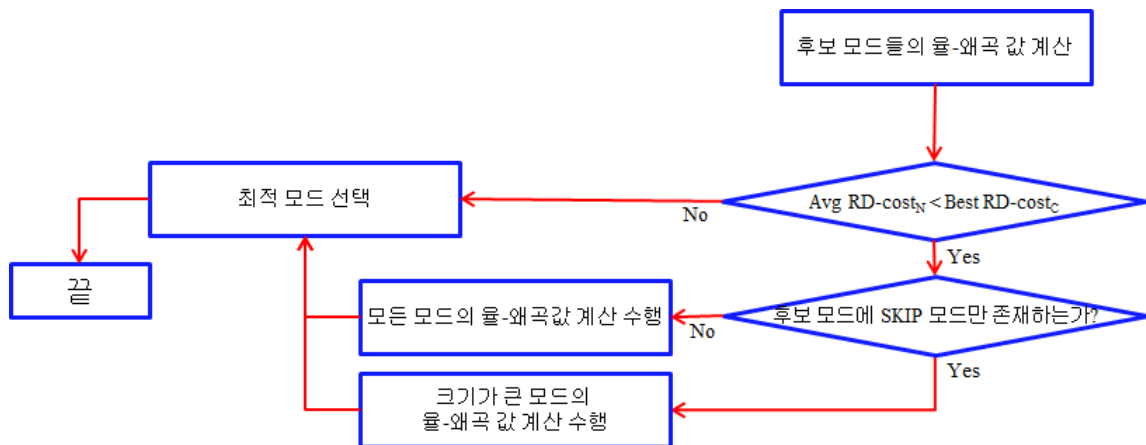


그림 8. 최적 모드의 선택 및 추가 수행을 위한 기법의 순서도 (STEP 3)
Fig. 8. Flowchart of algorithm for selection of best mode and additional perform (STEP 3)

표 7. Δ Bit-rate, Δ PSNR과 부호화 시간 감소율
Table 7. Δ Bit-rate, Δ PSNR and Encoding Time

		BQP 32, EQP 26			BQP 36, EQP 30			BQP 40, EQP 34		
		Δ PSNR	Δ Bit(%)	TS(%)	Δ PSNR	Δ Bit(%)	TS(%)	Δ PSNR	Δ Bit(%)	TS(%)
bus	Li's	-0.04	3.02	-50.31	-0.09	2.93	-52.26	-0.12	2.57	-52.94
	Proposed	-0.05	2.22	-50.93	-0.09	1.04	-49.68	-0.10	-0.36	-49.02
foreman	Li's	-0.08	3.13	-58.05	-0.13	1.62	-58.84	-0.16	0.32	-58.97
	Proposed	-0.08	2.43	-52.68	-0.11	-0.09	-53.06	-0.13	-1.79	-52.41
CREW	Li's	-0.11	3.41	-47.59	-0.19	2.06	-50.14	-0.26	0.18	-51.29
	Proposed	-0.07	1.76	-47.86	-0.12	0.00	-51.26	-0.15	-1.88	-52.44
football	Li's	-0.07	1.98	-40.58	-0.11	2.08	-44.58	-0.18	2.02	-47.30
	Proposed	-0.05	1.63	-52.46	-0.05	0.72	-52.41	-0.10	-0.54	-50.48
average	Li's	-0.07	2.88	-49.13	-0.13	2.17	-51.45	-0.18	1.27	-52.62
	Proposed	-0.06	2.01	-50.99	-0.09	0.42	-51.60	-0.12	-1.14	-51.09

Bit-rate의 증가 역시 적은 것을 알 수 있다. Crew 영상과 같은 경우는 BD-PSNR은 0.09dB의 차이가 나고, BD-rate은 2.6% 정도 차이가 나는 것을 확인 할 수 있다.

또한 Li의 기법과 비교해보면, 부호화의 속도는 제안된 기법 보다 Li의 기법이 대체적으로 빠르지만 표 7에서 살펴보면 football 영상과 같이 움직임이 많은 영상에 대해서는 제안된 기법의속도가 더 빠른 것을 확인 할 수 있다. 또한 양자화 계수를 달리하여 한 실험 결과를 살펴보면 평균적으로 양자화 계수가 작을수록 Li의 기법보다 좋은 속도가 나타나는 것을 알 수 있다. 한편, 화질적인 측면에서 살펴보면 표 7에서 Δ PSNR은 Li의 기법과 평균 0.01dB ~ 0.06dB

까지 차이가 나타나며, 특히 Δ Bit-rate는 Li의 기법 보다 양자화 계수가 작을 때 에는 평균 0.8% 정도로 큰 차이가 없지만 양자화 계수가 커지게 되면 2%가 넘는 차이가 나타나는 것을 볼 수 있다.

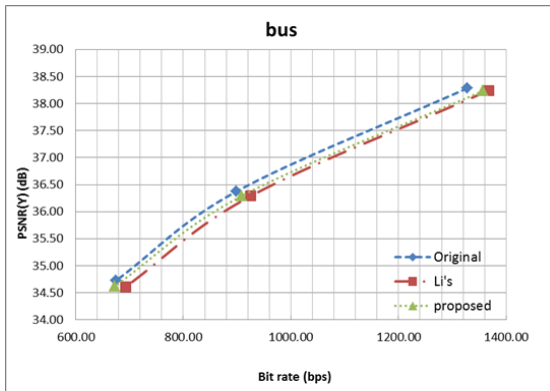
표 7의 결과에서 기본계층의 양자화 계수가 40이고 항상 계층의 양자화 계수가 34인 경우에 bit-rate가 감소하는 것을 볼 수 있다. 이것은 블록을 선택할 때 율-왜곡 최적화를 하는 과정에서 bit 부분만을 고려하여 선택하지 않기 때문에 나타날 수 있는 현상이다. 예를 들어 원본 실험에서 모든 모드를 확인하여 최적 모드로 결정된 모드가 화면 내 예측을 통한 모드가 결정 되었을 경우 빠른 모드 결정시 화면 내 예측을 수행하지 않게 된다면 화면 간 예측에서 수행되는 모드로 결정이 되는데 일반적으로 화면 내 예측 모드가 화면 간 예측 모드보다 bit의 크기가 크기 때문에 이와 같은 경우에는 bit가 감소 할 수 있다.

표 8. BD-rate(%)과 BD-PSNR (BQP : 28, 32, 36, 40; EQP : 22, 26, 30, 34)

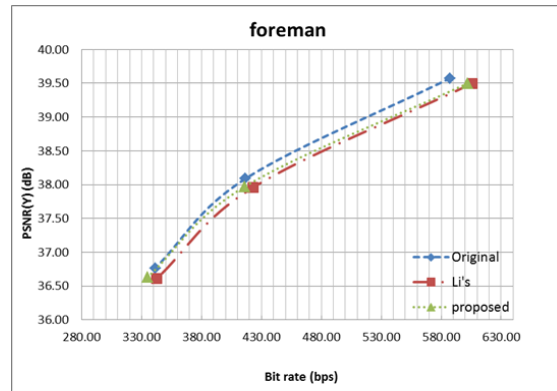
Table 8. BD-rate(%) and BD-PSNR (BQP : 28, 32, 36, 40; EQP : 22, 26, 30, 34)

	BD-PSNR		BD-rate(%)	
	Li's	Proposed	Li's	Proposed
Bus	-0.19	-0.14	3.98	2.81
Crew	-0.24	-0.13	5.68	3.04
Foreman	-0.18	-0.14	4.47	3.46
Football	-0.19	-0.13	3.57	2.41

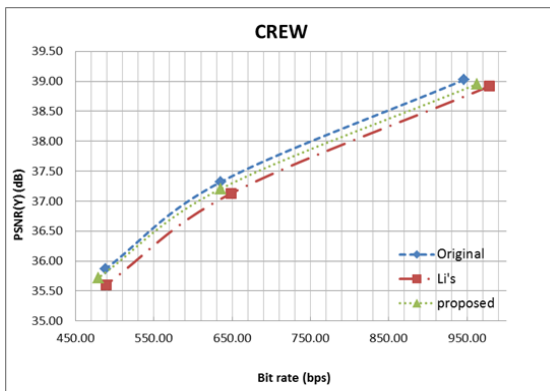
그림 9는 R-D 성능 그래프로서 화질적인 측면에서는 제안된 기법이 더 좋음을 확인 할 수 있다. Li의 기법이 원본에 비해서 그래프가 상당히 아래쪽에 형성되는 것으로 화질의 하락이 있는데 제안되는 기법은 영상의 종류에 관계없이 Li의 기법과 원본 영상의 가운데에 위치함으로 Li의 기법보다 화질적인 측면에서는 더 좋은 것을 확인 할 수 있다.



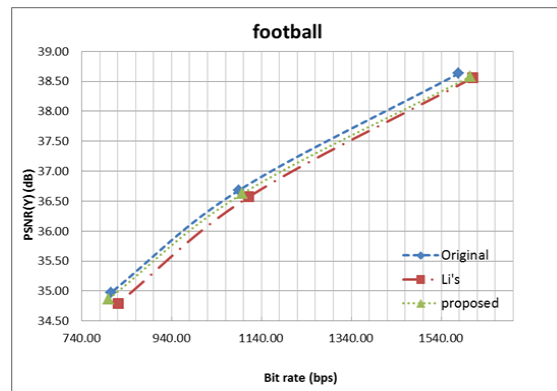
(a) R-D 성능 (bus)
(a) R-D performance (bus)



(b) R-D 성능 (foreman)
(b) R-D performance (foreman)



(c) R-D 성능 (CREW)
(c) R-D performance (CREW)



(d) R-D 성능 (football)
(d) R-D performance (football)

그림 9. 제안된 알고리즘을 이용했을 때의 R-D 성능
Fig. 9. R-D performance for the proposed algorithm

VI. 결론

본 논문에서는 스케일러블 비디오 코딩의 공간적 확장성의 측면에서 기본 계층과 향상 계층 간의 연관성을 이용하여 모드 결정을 고속화 하는 방법을 수행함으로써 스케일러블 비디오 코딩 부호화기의 복잡도를 크게 개선한 기법을 제안하였다. 제안된 방법은 현재 매크로블록과 같은 위치의 업-샘플링 된 기본 계층의 최적 모드와 방향성 정보를 이용하여 후보 모드를 추가하는 방식으로 진행되며 여러 가지 영상의 테스트를 통해 전체 부호화 시간 대비 최대

53%의 부호화 시간의 감소 효과를 얻을 수 있었다. 그러나 상대적 PSNR 감소나 비트 율 증가는 최소 -0.05dB , -2.01% 에서 최대 -0.15dB , 3.74% 에 불과해 무시 할 수 있을 정도로 작은 값이었다. 따라서 제안된 기법은 실시간 부호화기 구현 등에 유용하게 사용 될 수 있을 것이다.

참고 문헌

- [1] H.Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," IEEE Trans. Circuits Syst. Video Technol., vol.17, no. 9, pp. 1103-1120, Sep. 2007

- [2] Information technology-Coding of audio-visual objects Part 10: Advanced Video Coding, document ISO/IEC 14496-10, May. 2009.
- [3] H. Li, Z. Li, and C. Wen, "Fast mode decision algorithm for inter-frame coding in fully scalable video coding," IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 7, pp. 889-895, Jul. 2006.
- [4] S. W. Jung, S. J. Baek, C. S. Park, and S. J. Ko, "Fast mode decision using all-zero block detection for fidelity and spatial scalable video coding", IEEE Trans. Circuits Syst. Video Technol., Vol. 20, No. 1, Jan. 2010
- [5] J. F. Ren, N.Kehtarnavaz, "Fast adaptive early termination for mode decision in H.264 scalable video coding", International Conference on Image Processing (ICIP) 2008, pp. 2464-2467, Oct. 2008
- [6] Jong-Ho Kim, Hyo-Sung Kim, Byung-Gyu Kim, Hui-Yong Kim, Se-Yoon Jeong, Jin Soo Choi, "Fast Block Mode Decision Scheme for B-picture Coding in H.264/AVC", IEEE International Conference on Image Processing, pp. 2873-2876, Sep. 2010
- [7] Draft Reference software for SVC, document JVT-V203, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Jan. 2008

— 자 소 개 —



정 현 기

- 2010년 2월 : 선문대학교 컴퓨터정보학부 졸업
- 2010년 3월 ~ 2011년 현재 : 선문대학교 컴퓨터공학과 석사과정
- 주관심분야 : 영상비디오 신호처리, 스케일러블 비디오 코딩 (SVC)



홍 광 수

- 2011년 2월 : 선문대학교 컴퓨터공학부 졸업
- 2011년 3월 ~ 2011년 현재 : 선문대학교 컴퓨터공학과 석사과정
- 주관심분야 : 영상비디오 신호처리, 스테레오 비전



김 병 규

- 1995년 2월 : 부산대학교 공학사
- 1998년 2월 : 한국과학기술원 (KAIST) 공학석사
- 2004년 2월 : 한국과학기술원 (KAIST) 공학박사
- 2009년 2월 ~ 현재 : 선문대학교 컴퓨터공학과 교수
- 주관심분야 : 비디오 신호처리, 비디오 압축 알고리즘, 비디오 기반 감성인식 기술



김 창 기

- 1995년 2월 : 부산대학교 공학사
- 1997년 2월 : 부산대학교 공학석사
- 2005년 2월 : Carnegie-Mellon University 공학석사
- 1997년 3월 ~ 2000년 : 삼성전자 연구원
- 2000년 11월 ~ 현재 : 한국전자통신연구원 선임연구원
- 주관심분야 : 비디오통신 신호처리, 비디오 압축 알고리즘, Smart TV 기술

저 자 소 개



유 정 주

- 1982년 2월 : 광운대학교 공학사
- 1984년 2월 : 광운대학교 공학석사
- 2001년 7월 : Lancaster University 공학박사
- 1984년 ~ 현재 : 한국전자통신연구원 책임연구원
- 주관심분야 : QoS, 비디오 압축 알고리즘, Smart TV 기술