

일반논문-09-14-1-01

# GOP 구조 변환을 포함하는 MPEG-2에서 H.264/AVC로의 트랜스코딩

이 강 준<sup>a)†</sup>, 하 창 우<sup>a)</sup>, 정 제 창<sup>a)</sup>

## Transcoding MPEG-2 to H.264/AVC in the GOP Structure Conversion

Kang-Jun Lee<sup>a)†</sup>, Chang-Woo Ha<sup>a)</sup>, and Je-Chang Jeong<sup>a)</sup>

### 요 약

최근에 H.264/AVC 베이스라인 프로파일은 많은 멀티미디어 응용기기에 사용되고 있다. 또한 양방향 예측을 위한 B 픽처를 지원하는 MPEG-2 메인 프로파일은 HDTV, DVD와 같은 많은 멀티미디어 응용기기에 적용되어 왔다. 따라서 MPEG-2 메인 프로파일에서 H.264 베이스라인 프로파일로의 트랜스코딩은 콘텐츠의 범용성을 높이기 위해 필요한 작업이다. 양방향 움직임 추정을 지원하지 않는 H.264/AVC 베이스라인 프로파일로의 전환을 위해 Group of Pictures 구조 변환은 필수적이다. 이러한 Group of Pictures 구조 변환을 포함하는 트랜스코딩 구조에 있어서 제안한 알고리즘은 예측된 움직임 벡터의 선형성 검사를 통한 적응적인 탐색 범위 선택과 참조 영역 복잡성 정보를 이용한 적응적 모드 선택 방법을 통해 화질 열화를 최소화하며 계산상의 복잡성을 획기적으로 줄였다.

### Abstract

Currently, H.264/AVC is used in many multimedia applications. Also, The MPEG-2 main profile which supports B pictures for bi-directional motion prediction is widely used in applications such as HDTV and DVD's. Therefore, transcoding the MPEG-2 main profile to the H.264/AVC baseline is necessary for universal multimedia access. In this transcoding architecture including the GOP structure conversion, the proposed algorithms adopt the adaptive search range selection through the linearity test of a predictive motion vector and adaptive mode selection using the reference region complexity information. The proposed algorithms extremely reduce the computational complexity while maintaining the video quality.

Keywords : Transcoding, MPEG-2, H.264, GOP

## 1. 서 론

높은 부호화 효율을 갖는 H.264/AVC 표준의 등장과 함께 H.264/AVC는 많은 멀티미디어 기기에서 사용되고 있다. 특히 간단한 구현이 가능한 H.264/AVC 베이스라인 프

로파일은 DMB, IPTV, 멀티미디어 저장 장치와 같은 많은 응용에 사용되고 있다. 기존의 비디오 부호화 표준에 있어서는, 양방향 예측을 위한 B 픽처를 지원하는 MPEG-2 메인 프로파일이 HDTV, DVD와 같은 응용에 적용되고 있다. 따라서 MPEG-2 메인 프로파일에서 H.264/AVC 베이스라인으로의 트랜스코딩은 멀티미디어 콘텐츠의 범용성을 높이기 위해 필요하다. H.264/AVC는 다양한 블록 크기의 움직임 추정을 사용한다. 하지만 이것은 높은 계산 복잡도를 유발한다. 따라서 MPEG-2에서 H.264/AVC로의 트랜스코

a) 한양대학교 전자통신컴퓨터공학과

Hanyang Univ. Dept. of Electronics and Computer Engineering

† 교신저자 : 이강준(ee9627@ece.hanyang.ac.kr)

※ 본 연구는 서울시 산학연협력사업으로 구축된 서울 미래형콘텐츠 컨버전스 클러스터 지원으로 수행되었습니다.

딩 성능의 핵심은 화질 열화를 최소화하면서 움직임 추정과 모드 선택 과정에서 발생하는 복잡도를 감소시키는 것이다.

보다 유연한 콘텐츠 변환이 가능한 Cascaded Pixel Domain Transcoder (CPDT) 구조에 있어서 Group of Picture (GOP) 변환을 포함하는 트랜스코딩에서의 빠른 움직임 재추정과 관련된 몇몇 논문이 있다. 시간축을 따라 선형적으로 변화하는 움직임 벡터의 특성을 이용한 방법이 제안되었다<sup>[5]</sup>. 하지만 많은 초기 탐색점을 사용하여 복잡도의 증가를 가져왔다. Xin et al.에 의해 제안한 방법은 단순히 시간축을 따라 변환된 움직임 벡터를 초기 탐색점으로 사용하여 복잡도는 많이 줄였지만 Rate Distortion (RD) 성능의 감소를 가져왔다.

MPEG-2에서 H.264/AVC로의 트랜스코딩에 있어서 빠른 모드 선택을 위해, top-down 분할 구조에 따른 방법이 사용되었고<sup>[7]</sup>, 움직임 보상후에 각 8\*8 블록의 나머지 에너지의 분산에 따른 모드 선택방법도 적용되었다<sup>[8]</sup>. 그러나 위에 언급된 방법은 빠른 skip 모드 결정이 고려되지 않아, 이에 따른 복잡도의 감소와 비트율 감소의 효과를 얻지 못했다. Lu et al.은 주변블록과 임계값을 사용하여 SKIP 모드도 고려한 빠른 모드 선택 방법을 제안하였다. 하지만 모호한 모드 선택은 추가적인 계산 복잡성을 유발하였다.

본 논문에서는 MPEG-2 메인 프로파일에서 H.264/AVC 베이스라인 프로파일로의 트랜스코딩에 초점을 맞춘 GOP 구조 변환을 포함하는 트랜스코딩에서의 빠른 모드 선택과 움직임 재추정 방법에 대해 제안한다. 재추정 방법에 있어서는 시간축을 따라 움직임 벡터의 선형성에 따른 적응적 탐색영역 선택이 적용되었다. 빠른 모드 추정에 있어서는, 참조 영역 복잡도에 따른 적응적 모드 선택 방법이 사용되었다. 이러한 제안한 방법에 의해 화질 열화를 최소화하며 많은 계산 복잡도가 줄어들었다.

본 논문의 2장에서는 GOP 구조 변경을 포함하는 트랜스코딩에 있어서 기본적인 재추정 과정과 참조 영역 정보의 유용성에 대해 설명하고, 3장에서 제안하는 방법에 대해 설명한 뒤, 4장에서 실험 결과를 보여줄 것이다. 이에 따른 결론은 5장에서 설명된다.

## II. GOP 구조 변환에 있어서의 트랜스코딩

### 1. 시간축에서의 MV 조정

MPEG-2 메인 프로파일에서 H.264/AVC 베이스라인 프로파일로의 트랜스코딩에 있어서 트랜스코더에 입력으로 들어오는 비트스트림과 출력으로 나가는 출력 비트스트림의 GOP 구조는 다르다. MPEG-2 메인 프로파일의 B 픽처에서는 세 가지 타입의 움직임 추정 모드가 사용된다. 전방향 예측에 의한 움직임 추정은 과거에 프레임을 참조 프레임으로 사용하고, 후방향 예측에 있어서는 현재 프레임보다 미래의 프레임을 사용한다. 양방향 움직임 추정에는 전방향과 후방향 움직임 추정의 평균값이 사용된다. P 픽처에 있어서는 전방향 움직임 추정만이 사용된다. 세 가지 타입의 움직임 추정에 있어서 참조 프레임은 반드시 I 픽처나 P 픽처이어야 한다.

반면, H.264/AVC 베이스라인 프로파일에서는 전방향 움직임 추정 방법만이 사용된다. 따라서, 복잡도 감소를 위해 움직임 정보의 재사용의 관점에 있어서 MPEG-2 움직임 벡터는 Predictive Motion Vector (PMV)로써 사용되기 위해 조정되어야 한다<sup>[5]</sup>. 간단한 구현을 위해 이전에 부호화된 한 장의 프레임만을 참조 프레임으로 사용하였을 때 움직임 벡터가 시간축을 따라 선형적으로 변한다고 가정하면 PMV는 아래의 식으로 계산 되어진다.

$$\begin{aligned} MV_{pmv} &= MV_{original} / N_f, & \text{전방향 예측} \\ MV_{pmv} &= -MV_{original} / N_b, & \text{후방향 예측} \end{aligned} \quad (1)$$

식 (1)에 있어서  $MV_{pmv}$ 는 H.264/AVC 인코더에서 PMV로써 사용되는 움직임 벡터를 표현하고,  $MV_{original}$ 은 MPEG-2 복호기에서 사용된 움직임 벡터를,  $N_f$ ,  $N_b$ 는 참조프레임부터 현재 프레임 사이에 위치한 B 픽처의 개수이다(현재 프레임 포함). 식 (1)에서 후방향 움직임 벡터의 경우에는 음의 부호를 붙여서 전방향으로 추정하여 계산된다. 식(1)에 의해 계산된 PMV를 그림 1에 표현하였다. 그림 1에서  $PMV_{forward}$ 는 전방향 예측에서 추정된 움직임 벡터를 나타내고,  $PMV_{backward}$ 는 후방향 예측에서 추

정된 PMV를 PMVinter F는 양방향 예측에서 전방향 예측으로부터 추정된 PMV를 PMVinter B는 양방향 예측에서 후방향 움직임 추정으로부터 계산된 PMV를 나타낸다. 식 (1)에 의해 추정된 PMV는 시간축을 따라 선형적으로 움직이는 물체를 포함한 매크로 블록에 있어서 초기 탐색점으로 잘 작동한다. 하지만 시간축을 따라 비선형적으로 움직이는 물체를 포함한 경우에 있어서는 PMV로써 부정확하다 [5].

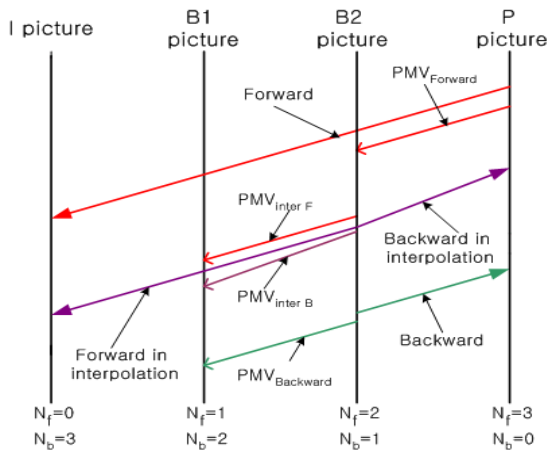


그림 1. 시간축을 따라 조정된 PMV  
Fig. 1. The PMV scaled in temporal direction

2. 매크로 블록의 모드와 참조 영역의 복잡도와의 연관성

부호화 효율을 높이기 위해 H.264/AVC에서는 움직임 추정을 위해 다양한 크기의 블록 크기를 지원한다. H.264/AVC P 슬라이스에서는 다음과 같은 모드들을 지원한다. 아래의 모드들에서 MODE\_는 인터 예측을 나타내고 SKIP은 SKIP 모드를, INTRA는 인트라 예측된 모드를 나타낸다. 각각의 서술자 뒤에 숫자는 예측에 사용된 블록의 크기를 나타낸다.

{SKIP, MODE\_16×16, MODE\_16×8, MODE\_8×16, MODE\_8×8, MODE\_8×4, MODE\_4×8, MODE\_4×4, INTRA\_4, INTRA\_16 }

H.264/AVC 부호화기에 있어서 복잡도의 많은 부분은 RD 최적화를 통한 최적 모드의 선택과정에서 발생한다. 따라서 전체적인 부호화 복잡도의 감소에 있어서 모드 선택 과정의 복잡도 감소는 결정적 역할을 한다. 다양한 모드 중에서 SKIP과 MODE\_16×16은 단순한 화면에 있어서 높은 비율을 차지한다<sup>[10]</sup>. 반면, 복잡한 영역에 있어서는 작은 크기의 블록이 최적의 모드로 선택되는 비율이 높다. 따라서 어떤 영역이 복잡한지 단순한지를 판단할 수 있다면 모드 선택과정에서 특정 모드를 배제하여 계산 복잡성을 줄일 수 있다. 이전 연구에 있어서<sup>[9][10]</sup>, 현재 프레임에 있어서 주변 블록 정보가 모드 선택의 참고로 사용되었다. 하지만 매크로 블록의 부호화 순서 때문에 현재 매크로 블록의 오른쪽 쪽과 아래 부분의 매크로 블록의 정보는 이용되지 못하였고 SKIP과 MODE\_16×16이 혼재되어 있는 부분에 있어서는 모드 예측에 부정확성이 나타났다.

양방향 예측을 지원하는 MPEG-2 메인 프로파일에 있어서, I 픽처와 P 픽처는 B 픽처와 P 픽처의 참조 프레임으로 사용되었다. 그림 2에서 보듯이 I 픽처와 P 픽처에서 참조 영역 복잡도 정보를 알고 있고 현재 모드와의 연관성을 찾을 수 있다면 우리는 현재 매크로 블록의 모드를 예측할 수 있을 것이다.

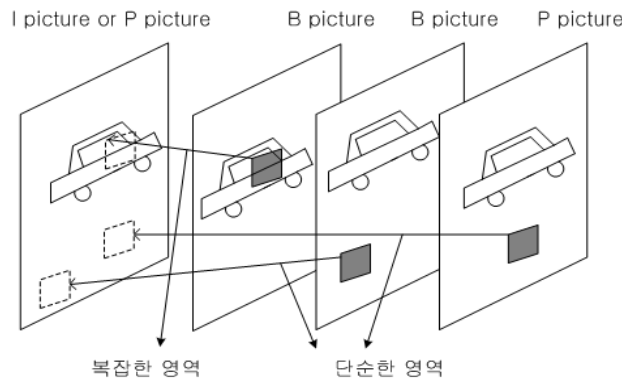


그림 2. MPEG-2 움직임 벡터에 의해 지시되는 참조 영역  
Fig. 2. The reference region indicated by MPEG-2 MV

우리는 참조 영역에 있어서 H.264/AVC에 의해 선택된 모드는 현재의 매크로 블록의 모드와 연관되어 있다는 것을 발견했다. 참조 영역에 있어서 모드가 SKIP과 MODE\_

16×16으로 구성되어 있는 평탄한 영역에 있어서 RD 최적화를 통해 선택된 현재 매크로 블록의 모드의 비율이 표 1에 나타나 있다.

표 1. 평탄한 영역(SKIP or MODE\_16×16)을 참조영역으로 가지는 블록의 현재 부호화 모드의 비율 (%) (입력 bitstream의 GOP 구조는 N=12, M=4이고 부호화단에서 QP = 28에서 실험)

Table 1. The coding mode ratio (%) in the reference region consisting of SKIP mode or 16×16 mode (Input GOP: N=12, m=4, QP=28)

	SKIP	16×16	16×8	8×16	8×8	Intra
Akiyo	89	6	2	2	1	0
Bus	36	32	11	11	9	1
Coastguard	43	30	9	10	8	0
Football	42	29	8	10	10	1
Foreman	55	23	7	8	6	1
Tabletennis	67	16	6	5	6	0

모든 시퀀스에 있어서 참조 영역이 평탄한 영역에 있을 때 현재 모드의 SKIP과 MODE\_16×16의 비율이 매우 높은 것을 확인할 수 있었다. 이것은 현재 매크로 블록의 모드와 참조 영역의 매크로 블록의 모드가 매우 연관되어 있다는 것을 입증한다. 참조 프레임이 뒤따르는 B 픽처와 P 픽처에서 계속적으로 사용됨으로써 GOP 구조 변경을 포함하는 트랜스코딩에 있어서 위의 특성은 매우 효율적으로 사용될 수 있다.

### III. 제안하는 방법

#### 1. 시간축에 따른 선형성 측정을 통한 적응적 움직임 탐색 범위 설정

앞서 설명하였듯이, 식 (1)에 의해 PMV를 유도하였을 경우 시간축에 대해 비선형적으로 움직이는 물체를 포함한 매크로 블록에 있어서 예측 오차가 크게 발생한다. 이러한 문제를 해결하기 위해 본 논문에서는 양방향 예측에 있어서 전방향 예측에서 유도된 PMV와 후방향 예측에서 유도된 PMV를 비교하여 움직임 추정을 위한 탐색 범위를 결정한다. 아래의 그림 3에서 각각의 PMV가 비슷하면 매크로 블록도 선형적으로 움직인다고 가정하고, 만약 각각의 PMV가 크게 차이가 나면 매크로 블록은 비선형적으로 움직인다고 가정한다.

위의 가정을 증명하기 위해 PMVinter F와 PMVinter B의 차이에 따른 PMVinter F와 탐색 범위 32로 전역 탐색 방법에 의해 탐색된 실제 움직임 벡터와의 차이를 그림 4에서 나타내고 있다. x축은 PMVinter F와 PMVinter B의 차이를 화소 단위로 나타내고 y축은 전역 탐색 방법에 의한 움직임 벡터와 PMVinter F의 차이를 나타낸다. 그림 4에서 예측 오차는 PMVinter F와 PMVinter B의 차이가 커짐에 따라 증가하고 있다. 이러한 특성을 이용하여 MPEG-2 비

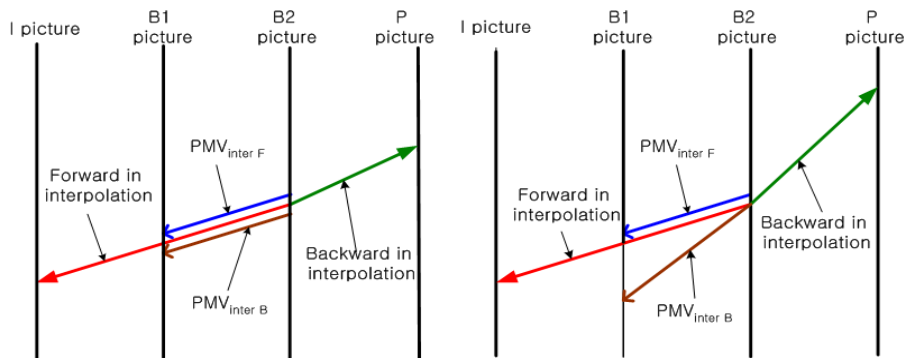


그림 3. 양방향 예측에서 움직임 벡터 선형성 비교  
Fig. 3. The linearity comparison in the interpolative prediction

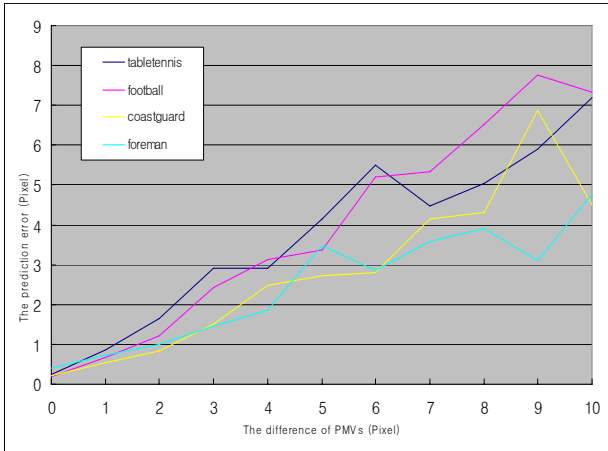


그림 4. 전,후 방향에서 유도된 PMV들의 차이에 따른 실제 움직임 벡터와 PMV의 오차  
Fig. 4. The PMV error in the linearity test

트스트림에서 양방향 예측이 사용된 경우에 H.264/AVC 부호화기에서 적응적인 탐색 범위를 선택할 수 있다. 양방향 예측이 사용되지 않은 전방향, 후방향 예측에 있어서는 복잡도와 RD 성능을 고려한 적절한 탐색 범위를 선택하였다. 다양한 움직임 복잡도를 가진 Akiyo, Bus, Coastguard, Container, Pairs, Football, Foreman, Table-tennis 시퀀스에서 실험을 통해 전방향 예측에 있어서는 탐색 범위를 3으로, 후방향 예측에 있어서는 6으로 고정하여 탐색을 수행하였다.

따라서 시간축을 따라 움직임 벡터의 선형성에 기반한 적응적 움직임 탐색 범위 설정은 다음과 같이 이루어진다.

- Forward prediction :  
search range =3
- Backward prediction :  
search range =6
- Interpolative prediction :  

$$\text{difference} = \text{abs}(PMV_{\text{Inter } F} - PMV_{\text{Inter } B}) \quad (2)$$
  - search range = 2     difference = 0,1
  - search range = 3     difference = 2,3
  - search range = 5     difference = 4,5
  - search range = 7     difference = otherwise

## 2. 참조 영역 복잡도에 기반한 적응적 모드 선택 방법

앞서 언급했던 것처럼, 현재 매크로 블록의 모드는 참조 영역의 모드와 높은 상관성을 가지고 있다. 특히, SKIP과 MODE\_16×16으로 구성된 평탄한 영역에 있어서 현재 매크로 블록 모드가 특정 모드로 부호화 될 비율은 SKIP과 MODE\_16×16의 결합 비율에 따라 달라진다.

그림 5는 참조 영역에 포함되는 블록의 구성 비율에 따른 다양한 참조 영역을 나타내고 있다. ALL\_SKIP은 참조 영역에 모든 블록이 SKIP으로 구성되어 있음을 나타낸다. SKIP\_16은 참조 영역이 SKIP과 MODE\_16×16이 혼합되어 구성되어 있음을 알 수 있다. ALL\_16은 참조 영역이 모두 MODE\_16×16으로 구성되어 있음을 나타낸다. 참조 영역의 모든 블록의 크기가 8×8 이상일 때에는 ABOVE\_8×8로 표현된다. MODE\_8×8보다 작은 크기의 블록이 참조 영역에 포함되어 있을 때는 COMPLEX로 표시한다.

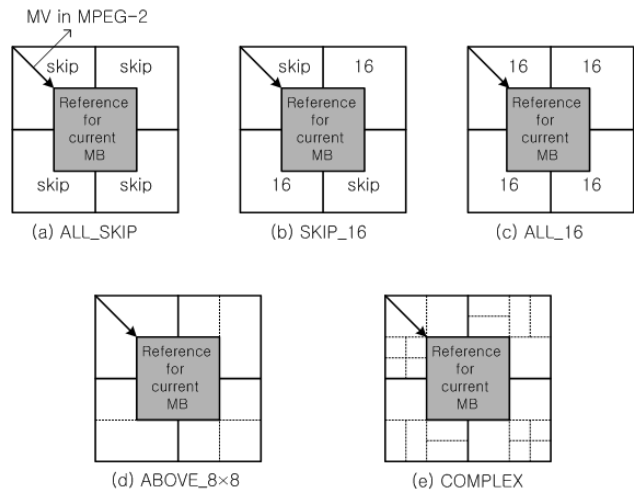


그림 5. 다양한 모드로 구성된 참조 영역들  
Fig. 5. Illustration of the different reference region

표 2는 참조 영역이 ALL\_SKIP과 SKIP\_16, ALL\_16일 때 현재 매크로 블록이 SKIP과 MODE\_16×16으로 부호화 되는 비율을 나타내고 있다. 참조 영역이 ALL\_SKIP 일 때 전체적으로 현재 매크로 블록이 SKIP과 MODE\_16×16으로 부호화 되는 비율이 굉장히 높고 특히 SKIP의 비율

표 2. 참조영역에 따른 현재 블록의 SKIP과 MODE\_16×16의 구성 비율(%)  
Table 2. The coding mode ratio(%) in the different reference region

sequence	ALL_SKIP		SKIP_16		ALL_16	
	SKIP	MODE_16×16	SKIP	MODE_16×16	SKIP	MODE_16×16
Akiyo	96	2	75	13	29	33
Bus	52	26	35	37	17	44
Coastguard	56	26	27	42	7	44
Football	47	32	38	37	26	37
Foreman	73	16	52	28	23	36
Table Tennis	72	13	47	29	27	39

이 높게 나타난다. 참조영역 SKIP\_16에서도 전체적으로 SKIP과 MODE\_16×16의 비율이 높게 나타나지만 SKIP의 비율은 ALL\_SKIP에서 보다는 떨어진다. ALL\_16에서는 MODE\_16×16의 비율이 SKIP 보다 더 높게 나타난다. 이러한 특성을 이용하여 참조 영역의 종류에 따라 적응적 모드 선택 방법을 적용할 수 있다.

추가적으로, 보다 정확한 모드 예측을 위해 세 종류의 경계값이 사용된다. 참조 영역에 따라 다른 경계값을 사용하여 보다 정확한 모드 예측이 가능해진다. 경계값의 계산을 위해 Sum of Absolute Difference (SAD)가 사용된다.

식 (1)은 기본적인 SAD 계산 과정을 나타낸다. 식 (2)는 탐색범위 안에서 최소의 SAD를 나타내고, 식 (3)은 H.264/AVC 부호화기에서 예측된 SKIP 위치에서의 SAD를 나타낸다. SADSKIP, SAD16은 아래의 경계값과 비교된다.

$$SAD(s(x, y), c(x, y, i, j)) = \sum_{x=1}^{16} \sum_{y=1}^{16} |s(x, y) - c(x+i, y+j)| \quad (1)$$

$$SAD_{16} = \min_{\{i, j\} \in SR} (SAD(s(x, y), c(x, y, i, j))) \quad (2)$$

$$SAD_{SKIP} = SAD(s(x, y), c(x, y, i_{SKIP}, j_{SKIP})) \quad (3)$$

식 (4)는 참조영역에서 SKIP 모드인 매크로 블록의 SAD의 평균을 나타내고 식 (5)는 MODE\_16×16 모드인 매크로 블록의 SAD의 평균을 나타낸다. 식 (6)은 이전 프레임에서 SKIP 모드로 부호화된 매크로 블록의 SAD의 평균을 나타

낸다.

$$AVER_{SKIP} = Aver \left\{ \sum_{\{MB_i | MODE_i = skip\}} SAD_{skip}(MB_i) \right\} \quad (4)$$

$$AVER_{16} = Aver \left\{ \sum_{\{MB_i | MODE_i = 16\}} SAD_{16}(MB_i) \right\} \quad (5)$$

$$AVER_{PRE\_SKIP} = Aver \left\{ \sum_{\{MB_i | MODE_i = skip_{previous}\}} SAD_{skip}(MB_i) \right\} \quad (6)$$

전체적인 모드 선택 과정에 있어서, 참조 영역이 ABOVE\_8×8일 때는 움직임 보상 후 남은 각각의 8×8블록의 DCT 계수의 절대값의 합의 분산을 이용하여 모드를 판단한다. 왜냐하면 이러한 방법은 매크로 블록이 MODE\_16×16, MODE\_16×8, MODE\_8×16, MODE\_8×8으로 부호화되는 비교적 덜 복잡한 영역에 있어서 좋은 성능을 나타내기 때문이다<sup>18)</sup>.

MPEG-2에서 B 픽처로 부호화 된 프레임의 모드 선택에 있어서, 표 2에서 살펴보았듯이 참조 영역이 ALL\_SKIP 일 경우 SKIP과 MODE\_16×16으로 부호화되는 비율이 거의 대부분을 차지한다. 따라서 본 논문에서는 AVERSKIP과 SADSKIP의 비교에 의해 SKIP과 MODE\_16×16 두 가지 모드 중에서 하나를 선택한다. AVERSKIP에 상수 1.2를 곱한 값을 경계값으로 사용하여 ALL\_SKIP 영역에서 SKIP 모드를 예측하는데 이것은 예측의 정확도는 다소 감소하여

표 3. 가중치에 따른 모든 선택의 정확도와 선택과정에 의해 결정되는 매크로 블록의 비율  
 Table 3. The accuracy and the portion of modes selected by the weighting method

sequence	1.0		1.1		1.2	
	정확도(%)	참여도(%)	정확도(%)	참여도(%)	정확도(%)	참여도(%)
Akiyo	98	28	96	49	95	52
Foreman	72	3	71	5	70	5
Paris	96	23	94	35	93	36
Container	91	42	91	40	91	42
Table Tennis	92	9	89	15	86	16

도 이러한 예측에 참여하는 매크로 블록의 비율을 높여 복잡도 감소 효과를 가져 온다. 각각의 가중치에 따른 예측 정확도와 이에 따라 감소효과를 가져오는 매크로 블록의 비율을 표 3에 나타냈다.

SKIP\_16 영역에 있어서는 표 2에서 볼 수 있듯이 SKIP과 MODE\_16×16의 비율이 시퀀스에 따라 다르게 나타난다. 따라서 이 영역에 있어서는 SAD 비교에 의한 모드 선택의 방법은 정확성이 떨어진다. 우리는 두 개의 SAD 값을 사용하여 정확성을 높였다. 만약 SADSKIP이 AVERSKIP보다 작고 SAD16이 AVER16보다 작다고 하면 모드를 SKIP으로 판단한다. 그렇지 않고 SAD16만이 AVER16보다 작다면 RD 비용(cost)이 SKIP과 MODE\_16×16의 선택을 위해 계산된다. 만약 두 개의 SAD 값이 모두 경계값보다 크다면 각각의 8×8블록의 움직임 보상후의 에너지의 분산이 모드 선택을 위해 사용된다.

참조 영역이 ALL\_16 영역에 있는 경우에는, 표 2에서 보듯이 MODE\_16×16의 비율이 SKIP과 비교하여 높다. 또한, ALL\_16 영역에 있어서는 AVERSKIP을 사용할 수 없으므로 AVERpre\_sKIP을 이용하여 SADSKIP이 AVERpre\_sKIP보다 작은 경우에는 SKIP으로 그렇지 않고 SAD16만이 AVER16보다 작다면 MODE\_16×16으로 부호화되고 나머지 경우에 있어서는 각각의 8×8블록의 움직임 보상후의 에너지의 분산이 모드 선택을 위해 사용된다.

제안한 방법에서 뒤따르는 B 픽처와 P 픽처에 있어서 참조 영역 복잡도를 판단하기 위한 참조 프레임으로 사용되는 P 픽처에서의 정확한 모드 선택은 이 프레임을 참조 프레임으로 사용하여 참조 영역 복잡도를 판단하는 뒤따르는 B 픽처와 P 픽처의 RD 성능에 많은 영향을 미친다. 따

라서, P 픽처에 있어서 참조 영역이 ALL\_SKIP, SKIP\_16, ALL\_16의 경우에는 SKIP, MODE\_16×16, MODE\_16×8, MODE\_8×16, MODE\_8×8 들의 RD 비용을 계산하여 최소의 RD 비용을 가진 값을 최적의 모드로 결정한다. P 픽처의 다른 참조 영역에 있어서는 B 픽처의 방법과 동일하게 모드 선택이 이루어진다.

COMPLEX 영역에 있어서는 모든 모드들의 RD 비용이 계산되고 최소의 RD 비용을 갖는 모드가 선택되어 진다. 비록 이러한 과정에서 높은 계산 복잡도가 요구되지만 기존의 모드 예측에 있어서 많은 예측 오차를 발생시키는 복잡한 영역에서의 모드 선택의 오차를 줄일 수 있고 복잡한 영역에서만 이러한 과정을 수행하여 계산 복잡도를 최대한 줄였다. 제안한 방법의 전체적인 수행과정은 아래의 의사코드와 같다.

< MPEG-2 P 픽처에서의 움직임 재추정과 모드 선택 방법 >

```

The PMV for the ME is adjusted with equation (1).
The search range is determined by the equation (2).

if (reference region = ALL_SKIP or ALL_16 or SKIP_16)
{ coded in RDO with enabling SKIP, MODE_16×16,
  MODE_16×8, MODE_8×16, MODE_8×8.
}
else if(the reference region = ABOVE_8×8)
{ the energy variance is compared. }
else
{ the current MB is coded in RDO concept with enabling
    
```

the all inter prediction mode.

}

#### < MPEG-2 B 픽처에서의 움직임 재추정과 모드 선택 방법 >

The PMV for the ME is adjusted with equation (1).

The search range is determined by the equation (2).

if (the reference region = ALL\_SKIP)

{

if((AVERSKIP × 1.2) > SADSKIP){ SKIP }

else { MODE\_16×16 }

}

else if (the reference region = SKIP\_16)

{

if(AVERSKIP > SADSKIP && AVER16 > SAD16)

{ SKIP }

else if(AVER16 > SAD16) {determined among SKIP and MODE\_16×16 in RDO}

else{ the energy variance is compared. }

}

else if (the reference region = ALL\_16)

{

if(AVERpre\_sKIP > SADSKIP && AVER16 > SAD16) { SKIP }

else if(Aver16 > SAD16){ MODE\_16×16 }

else { the energy variance is compared. }

}

else if (the reference region = ABOVE\_8×8)

{ the energy variance is compared. }

else

{ the current MB is coded in enabling the all inter prediction mode with RDO. }

## IV. 실험 결과

트랜스코더를 구현함에 있어서, 복호기 부분의 MPEG-2 메인프로파일을 위해 TM5<sup>[11]</sup>를 사용하였고 부호기 부분의 H.264/AVC 베이스라인 프로파일을 위해 J.M 8.6<sup>[12]</sup>이 사용되었다. 입력 비트스트림을 위해 CIF(352\*288) 크기의 시퀀스를 초당 30프레임씩 4Mbps로 부호화하였고 GOP 구조는 I 픽처간의 거리는 12 (N=12)로, P 픽처간의 거리는 4 (M=4)를 사용하였다. 출력 비트스트림을 위해 H.264/AVC 베이스라인 부호화기는 이전에 부호화된 한 장의 프레임을 참조 프레임으로 사용하였다. 따라서 출력 비트스트림의 GOP 구조는 N=∞, M=1이다. 모든 실험은 1Gb의 메모리 크기와 함께 Intel Pentium Core2 1.86GHz에서 수행되었다. H.264/AVC 부호화기의 부호화 파라미터 설정은 표 4와 같다.

표 4. H.264/AVC 부호화 파라미터

Table 4. H.264/AVC Encoding parameters

Number of reference frames	1
RDO	on
Search range	16
Rate control	off
FME	off
Use Hadamard	on
Output GOP structure	N=∞, M=1 (IPPP)

제안한 방법의 성능을 비교하기 위해, MPEG-2 복호기에 서 복호된 영상을 JM 부호화기 지원하는 RDO 방법을 이용하여 최적의 모드를 선택한 방법과 비교하였다. Adaptive Search Range Selection (ASRS)에 있어서는, 탐색영역은 식 (2)에 의해 결정된 탐색영역에서 모든 모드의 RD 비용을 구해서 최소의 RD 비용을 가지는 모드가 최적의 모드로 선택된다. ASRS+Adaptive Mode Decision Method(AMDM)에 있어서는, 탐색영역은 식 (2)의해 결정되고 모드 선택은 앞서 설명한 참조 영역 복잡성에 의한 적응적 모드 선택 방법에 따라 결정된다.



표 5는 QP=18에서 제안한 트랜스코더의 성능을 나타낸다. Reference와 비교하여 평균적으로 ASRS는 0.008dB의 PSNR 감소와 0.72%의 bitrate 증가를 보이고 ASRS+AMDM은 0.07dB의 PSNR 감소와 3.89%의 bitrate 증가율을 보인다. 반면 수행시간에서는 각각 18%와 72%의 감소를 가져왔다. 표 6은 QP=24에서의 제안한 트랜스코더의 성능을 보여준다. 평균적으로 ASRS는 0.003dB의 PSNR 감

소와 1.48%의 bitrate 증가를 보이고 ASRS+AMDM은 0.05dB의 PSNR 감소와 7.29%의 bitrate 증가율을 보인다. 수행시간에서는 각각 20%와 73% 이상의 감소를 보인다. 표 7은 QP=30에서의 RD 성능과 수행시간 감소의 성능을 나타낸다. Reference와 비교하여 평균적으로 ASRS는 0.01dB의 PSNR 감소와 2.11%의 bitrate 증가를 보이고 ASRS+AMDM은 0.05dB의 PSNR 감소와 12.24%의 bi-

표 5. 제안한 방법의 성능 비교 (QP = 18)

Table 5. The performance comparison of the proposed algorithm (QP=18)

sequence	ASRS			ASRS+AMDM		
	ΔPSNR (dB)	Δ Bitrate (%)	Δtime (%)	ΔPSNR (dB)	Δ Bitrate (%)	Δtime (%)
Akiyo	0.01	0.04	-22.13	0	3.30	79.23
Bus	0.02	2.38	-16.17	-0.07	4.80	69.53
Container	0	0.07	-19.35	-0.04	2.32	73.83
Paris	0	0.37	-18.12	-0.05	2.29	75.66
Football	0.02	0.87	-15.88	-0.21	6.90	68.98
Foreman	0	0.64	-18.27	-0.06	3.74	68.54

표 6. 제안한 방법의 성능 비교 (QP = 24)

Table 6. The performance comparison of the proposed algorithm (QP=24)

sequence	ASRS			ASRS+AMDM		
	ΔPSNR (dB)	Δ Bitrate (%)	Δtime (%)	ΔPSNR (dB)	Δ Bitrate (%)	Δtime (%)
Akiyo	-0.01	0.23	24.36	0.03	4.16	80.45
Bus	0	4.39	18.60	-0.09	7.96	69.76
Container	-0.02	0.47	22.19	-0.02	7.54	77.02
Paris	0.02	0.64	20.24	-0.02	4.10	77.31
Football	-0.01	1.52	18.18	-0.19	11.17	68.79
Foreman	0	1.64	21.89	-0.05	8.82	70.18

표 7. 제안한 방법의 성능 비교 (QP = 30)

Table 7. The performance comparison of the proposed algorithm (QP=30)

sequence	ASRS			ASRS+AMDM		
	ΔPSNR (dB)	Δ Bitrate (%)	Δtime (%)	ΔPSNR (dB)	Δ Bitrate (%)	Δtime (%)
Akiyo	0.02	0.05	25.82	0.09	7.23	81.98
Bus	0.03	6.71	20.57	-0.13	14.18	70.83
Container	0.03	0.47	24.64	-0.01	14.31	80.17
Paris	0	1.11	21.92	0.03	6.86	78.07
Football	0.01	2.29	20.55	-0.19	16.38	69.72
Foreman	0.01	2.04	24	-0.1	14.53	72.85

표 8. 제안한 방법의 성능 비교 (QP = 36)

Table 8. The performance comparison of the proposed algorithm (QP=36)

sequence	ASRS			ASRS+AMDM		
	$\Delta$ PSNR (dB)	$\Delta$ Bitrate (%)	$\Delta$ time (%)	$\Delta$ PSNR (dB)	$\Delta$ Bitrate (%)	$\Delta$ time (%)
Akiyo	0.01	0.48	27.79	0.12	12.34	83.70
Bus	-0.06	8.88	23.57	-0.21	25.76	72.72
Container	-0.03	-0.80	26.74	0.07	15.90	82.97
Paris	-0.01	1.69	24.19	-0.05	10.79	79.30
Football	0.02	3.17	23.24	-0.25	21.64	71.25
Foreman	-0.05	1.78	25.67	-0.2	21.66	75.83

표 9. 제안한 방법의 성능 비교 (QP = 42)

Table 9. The performance comparison of the proposed algorithm (QP=42)

sequence	ASRS			ASRS+AMDM		
	$\Delta$ PSNR (dB)	$\Delta$ Bitrate (%)	$\Delta$ time (%)	$\Delta$ PSNR (dB)	$\Delta$ Bitrate (%)	$\Delta$ time (%)
Akiyo	0.02	-0.77	29.42	0.09	16.25	85.19
Bus	0.06	8.72	25.90	-0.37	41.00	74.95
Container	0.03	-0.12	28.11	-0.03	14.40	84.54
Paris	0.02	1.16	25.95	-0.07	16.90	81.12
Football	0.01	3.89	25.28	-0.39	25.32	73.22
Foreman	0.06	1.51	26.90	-0.42	26.78	78.90

trate 증가율을 보인다. 수행시간에서는 각각 22%와 75% 이상의 계산 복잡도 감소를 가져왔다. 표 8에서 보이는 QP=36에서는 평균적으로 ASRS는 0.02dB의 PSNR 감소와 2.53%의 bitrate 증가를 보이고 ASRS+AMDM은 0.08dB의 PSNR 감소와 18.01%의 bitrate 증가율을 보인다. 수행시간에서는 각각 25%와 77% 이상의 감소를 보인다. 표 9는 QP=42에서 제안한 트랜스코더의 성능을 나타낸다. Reference와 비교하여 평균적으로 ASRS는 0.03dB의 PSNR 감소와 2.39%의 bitrate 증가를 보이고 ASRS+AMDM은

0.19dB의 PSNR 감소와 23.44%의 bitrate 증가율을 보인다. 반면 수행시간에서는 각각 26%와 79%의 감소를 가져왔다.

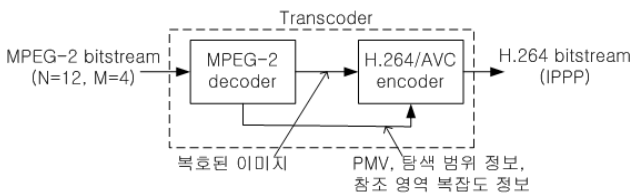


그림 6. 제안된 트랜스코더 구조  
Fig. 6. The proposed structure for transcoding MPEG-2 to H.264/AVC

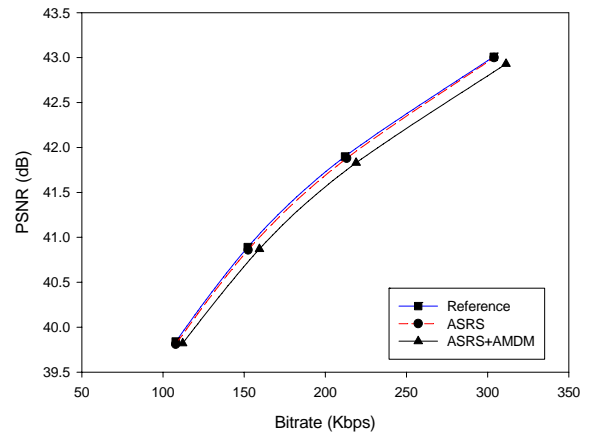


그림 7. Akiyo 시퀀스의 R-D 성능  
Fig. 7. The R-D performance in the Akiyo sequence

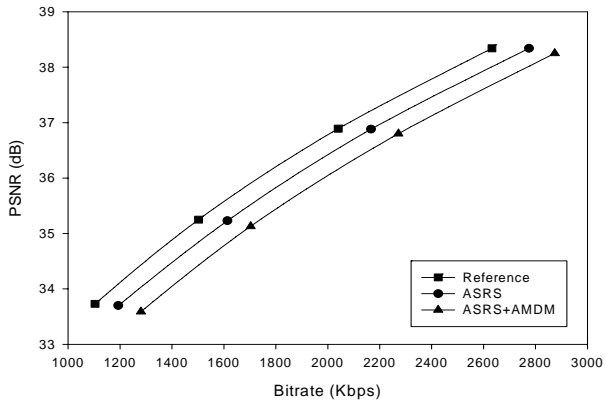


그림 8. Bus 시퀀스의 R-D 성능  
Fig. 8. The R-D performance in the Bus sequence

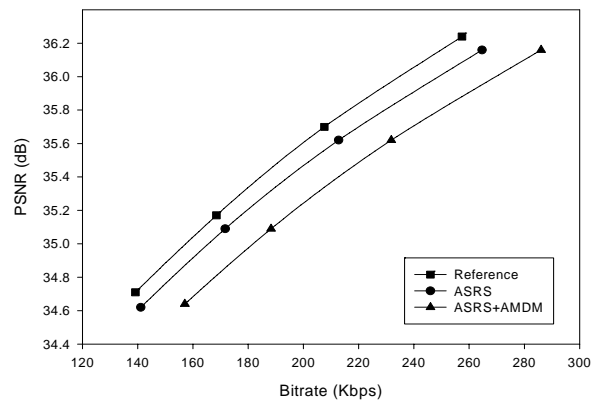


그림 9. Container 시퀀스의 R-D 성능  
Fig. 9. The RD performance in the Container sequence

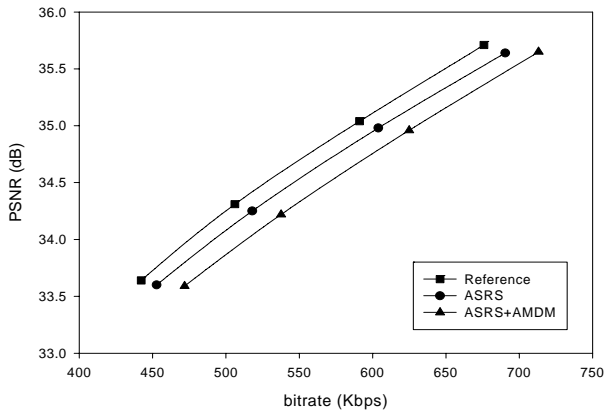


그림 10. Paris 시퀀스의 R-D 성능  
Fig. 10. The R-D performance in the Paris sequence

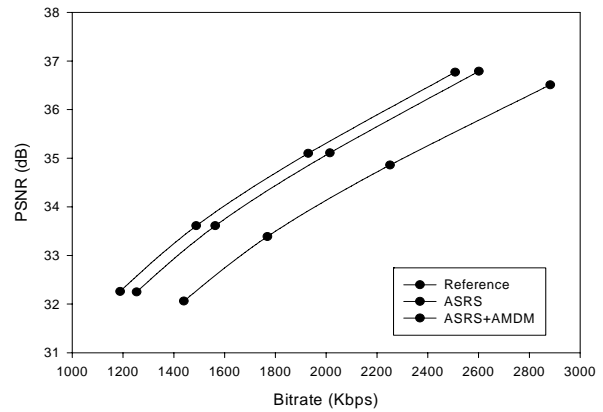


그림 11. Football 시퀀스의 R-D 성능  
Fig. 11. The R-D performance in the Football sequence

그림 6-11은 제안한 트랜스코더의 RD 성능을 나타낸다.

## V. 결론

제안한 방법은 MPEG-2 메인 프로파일에서 H.264/AVC 베이스라인 프로파일로의 변환과 같은 GOP 구조 변경을 포함하는 트랜스코딩에 있어서 효과적이다. 제안한 방법은 움직임벡터의 선형성 측정에 의한 적응적 탐색영역 선택과 참조 영역 복잡도에 기반한 적응적 모드 선택 방법에 의해 계산 복잡도를 획기적으로 줄일 수 있었다. 저비트율에서 RD 성능의 감소는 더 보완하여야 하는 과제이다.

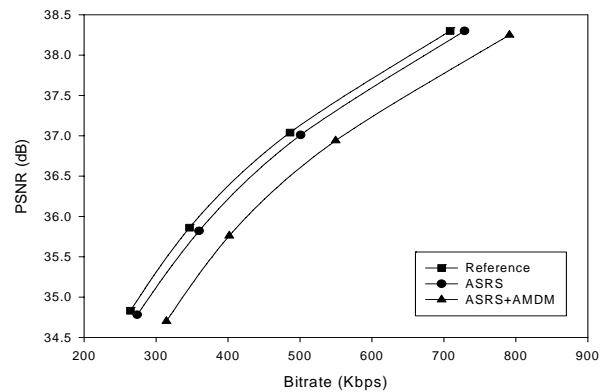


그림 12. Foreman 시퀀스의 R-D 성능  
Fig. 12. The R-D performance in the Foreman sequence

## 참고 문헌

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, ITU-T Rec. H.264 | ISO/IEC 14496 - 10 AVC, 2003.
- [2] Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Video, ISO/IEC 13818-2, 1995.
- [3] R. Mohan, J. R. Smith, and C. S. Li “Adapting multimedia Internet content for universal access,” IEEE Trans. Multimedia, vol. 1, no. 1, pp. 104 -114, Mar, 1999.
- [4] H. Sun, W. Kwok, and J. W. Zdepski, “Architectures for MPEG compressed bitstream scaling,” IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 2, pp. 191-199, Apr. 1996.
- [5] T. Shanableh, M. Ghanbari, “The importance of the bi-directionally predicted pictures in video streaming,” IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 3, pp. 402 - 414, Mar, 2001.
- [6] J. Xin, A. Vetro, S. Sekiguchi, and K. Sugimoto, “Motion and mode mapping for MPEG-2 to H.264/AVC transcoding,” in Proc. IEEE Int. Conf. Multimedia and Expo, 2006, pp. 313-316.
- [7] Z. Zhou, S. Sun, S. Lei, and M. T. Sun, “Motion information and coding mode reuse for MPEG-2 to H.264 transcoding,” in Proc. IEEE Int. Symp. Circuits Syst. 2005, vol. 2, pp.1230-1233.
- [8] G. Chen, Y.Zhang, S. Lin, and F. Dai, “Efficient block size selection for MPEG-2 to H.264 transcoding,” in Proc. ACM Int. Conf. Multimedia, 2004, pp.300-303.
- [9] X. Lu, A. M. Tourapis, P. Yin, and J. Boyce “Fast mode decision and motion estimation for H.264 with a focus on MPEG-2/H.264 transcoding,” in Proc. IEEE Int. Symp. Circuits Syst. 2005, vol.2, pp.1246- 1249.
- [10] C. Grecos, and M. Y. Yang, “Fast inter mode prediction for p slices in the H.264 video coding standard,” IEEE Trans. Broadcasting, vol.51, no. 2, pp. 256-263, June, 2005.
- [11] Test Model 5, ISO/IEC JTC1/SC29/WG11, N0400, Apr. 1993.
- [12] H.264/AVC Reference Software JM 8.6 [Online] Available: <http://bs.hhi.de/~suehring/tml/>

## 저 자 소 개



## 이 강 준

- 2003년 : 강릉대학교 전자공학과 학사
- 2006년 : 한양대학교 전자통신컴퓨터공학과 석사
- 2006년 ~ 현재 : 한양대학교 전자통신컴퓨터공학과 박사과정
- 주관심분야 : 비디오 트랜스코딩, H.264/AVC 및 차세대 영상 압축 기술



## 하 창 우

- 2005년 2월 : 강원대학교 전자공학과 졸업
- 2007년 8월 : 한양대학교 전자통신컴퓨터공학과 석사
- 2007년 9월 ~ 현재 : 한양대학교 전자통신컴퓨터공학과 박사과정
- 주관심분야 : 영상 화질 개선, H.264/AVC 및 차세대 영상 압축 기술



## 정 제 창

- 1980년 2월 : 서울대학교 전자 공학과 졸업
- 1982년 2월 : KAIST 전기전자 공학과 석사
- 1990년 : 미국 미시간대학 전기 공학과 공학박사
- 1980년 ~ 1986년 : KBS 기술연구소 연구원(디지털 TV 및 뉴미디어 연구)
- 1991년 ~ 1995년 : 삼성전자 멀티미디어 연구소 (MPEG, HDTV, 멀티미디어 연구)
- 1995년 ~ 현재 : 한양대학교 전자통신컴퓨터공학과 교수 (영상통신 및 신호처리 연구실)
- 1998년 11월 27일 : 과학기술자상 수상
- 1998년 12월 31일 : 정보통신부장관상 표창
- 주관심분야 : 영상처리 및 영상압축