HEVC Test Model의 변환 부호화 기술

김재일 · 김문철 (KAIST)

요 약

ISO/IEC와 ITU의 비디오 압축 공동협력팀(Joint Collaborative Team on Video Coding – JCTVC)은 차세대 비디오 압축 표준인 HEVC (High Efficiency Video Coding)를 표준화 중이다. JCTVC에서 제안된 핵심기술들 중 하나인 변환 기술은 크게 핵심 변환 커널 설계 기술과 선택적 변환 기술로 나누어 표준화가 진행되고 있다. 본 논문에서는 변환 방법과 관련된 채택된 기술들을 설명하고, 현재 진행 중인 변환관련 표준화 참여 기술들을 소개한다.

I. 서 론

지난2010년 4월 1차 JCTVC 회의를 시작으로 최근 2011년 7월 6차 JCTVC 회의까지 총 6차에 걸친 HEVC 표준화 회의가 개최되었다. 처음 새로운 표준을 시작하기 위해 92차 MPEG 회의에서 공포된 Cfp (Call for Proposal)에 따라 총 27개 기관에서 테스트모델을 제출하였으며[1] 그 중 동일한 비트 내에서 주관적/객관적 화질이 우수한 제안서의 기술들을 이용하여 1차 회의에서 HEVC TMuC(Test Model under Consideration)가 임시적으로 구성되었다. HEVC TMuC는 블록 기반의 화면간/화면내 예측 방법, 잔차 신호의 블록 기반 변환 방법, 엔트로피 방법 등을 기반으로 확장된 기술과 반복적 블록 구조, 개선된 움직임 벡터 예측 방법, 화면간 예측을 위한 방향성 변환 방법, 루프 내 보간 필터 등의 새롭게 추가된 기술들로 구성되었다[2].

HEVC TMuC의 여러 기술 중 압축 성능과 복잡도를 고려하여 2차 및 3차 JCT-VC 회의에서는 HM을 위한 기본 비디오 압축기술들이 선택되었고, 추가적인 논의를 위하여 다양한 핵심기술들에 대한 CE가 진행되어 최근 6차 JCTVC 회의까지 최적의 기술들이 채택되었다. 〈표 1〉은 5차 JCTVC 회의까지 고효율(High Efficiency)과 저복잡도(Low Complexity) 모드에 따라 HM 3.0에서 최종적으로 채택된 기술들이다.

6차 JCTVC 회의까지 제안된 다양한 핵심 기술 중 변환 기술은 핵심실험(CE: Core Experiment)의 핵심 변환 설계(core transform design) 및 부가적 또는 선택적 변환(additional or alternative transform) 기술들이 표준화가 진행 중이다. 핵심 변환설계 기술의 경우 최적의 변환커널로 채택되기 위해 하드웨어에서의 저복잡도 및 소프트웨어 상의 SIMD(Single Instruction Multiple Data)를 이용한 병렬성에 최적화 되어야

〈표 1〉 HM 3.0까지 채택된 기술[3]

High Efficiency	Low Complexity
Coding Unit 8x8 up to 64x64 in tree structure	
Prediction Units	
Transform unit tree structure (3 level max)	
Transform block size of 4x4 to 32x32 samples Mode-dependent Transform for 4x4 intra prediction	
Spatial Intra Prediction (Max. 34 angular directions and Planar) Adaptive Intra Smoothing Intra Chroma Prediction using Luma samples	
DCT-based interpolation filter for luma samples (1/4-sample, 8-tap)	
DCT-based interpolation filter for Chroma samples (1/8-sample, 4-tap)	
Advanced motion vector prediction	
Coding Unit based Skip & Prediction Unit based merging	
Context adaptive binary arithmetic entropy coding	Context adaptive VLC
Internal bit-depth increase (2 bits)	X
Deblocking filter Sample Adaptive Offset	
Adaptive loop filter	X

하며, CE를 통해 현재 여러 기관에서 제안한 기술들이 부호화 효율과 복잡도 측면에서 기술적인 경쟁을 하고 있다.

선택적 변환 방법은 5차 회의에서 화면내 4x4 방향성 예측 부호화 모드에 따라 DCT(Discrete Cosine Transform)/DST(Discrete Sine Transform)를 다르게 조합하여 적용하는 방법이 채택되었으며, 나머지 화면내/화면간 모드에 대해 적용하기 위한 다양한 방법들이 CE에서 논의 중이다.

본 논문에서는 JCTVC 6차 회의까지 제안된 변환기술 중 채택된 기술과 핵심논의중인 기술을 중심으로 정리하였다.

Ⅱ. HM의 핵심 변환 설계 기술

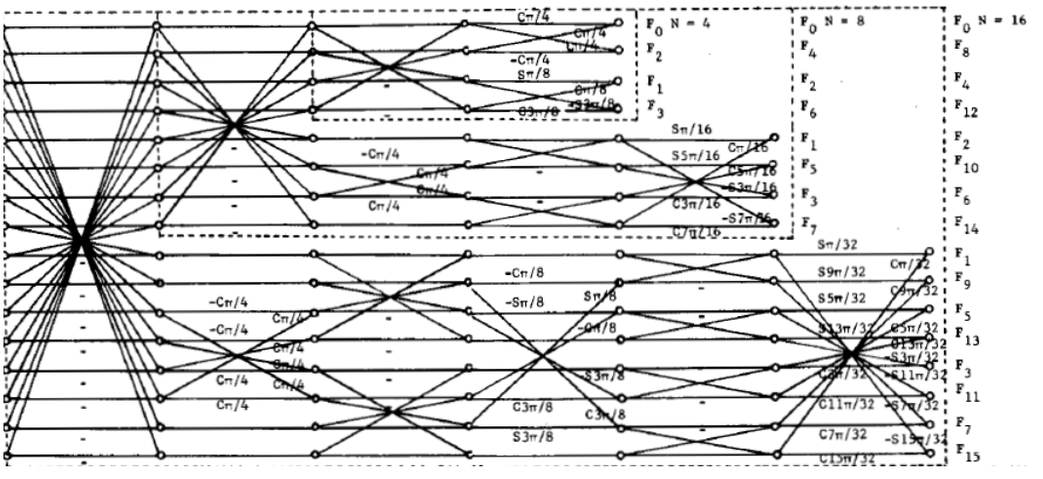
JCTVC 표준화회의에서 핵심 변환 방법은 삼성전자를 비롯한 여러 참여기관에서 CE를 통해 각 제안기술들이 경쟁 중이다. 초기 HM 즉, HM1.0과 2.0에서는 4x4, 8x8 크기 변환커널은 H.264/AVC의 정수 변환(integer transform)을, 16x16, 32x32 크기 변환은 삼성전자에서 제안한 Chen의 고속 DCT 커널을 사용하였다[4]. 하지만, 5차 회의에서 Cisco와 TI에서 제안한 4x4에서 32x32까지 새로운 변환 커널이 HM3.0에 채택되었으나 WD3.0에는 채택을 보류한 상태이며[5] CE를 통해 다양한 제안 변환커널들의 성능을 비교 검증 한 후에 최종 변환커널들을 결정할 것으로 예상된다. 본 장에서는 변환 커널 관련된 현재의 WD 채택 기술과 CE 경합 기술들을 정리하였다.

1. Chen 고속 DCT 방법[4]

삼성전자에서 제안한 Chen 고속 DCT 방법은 HM1.0과 HM2.0과 현재 WD4.0에까지 채택되어 있는 기술로서 16x16, 32x32 크기의 변환 커널에 사용된다. Chen의 DCT는 최적의 고속 DCT 알고리즘은 아니지만 버터플라이 구조로서 구현하기 쉽고, 작은 크기의 변환을 재사용할 수 있기 때문에 처음 HM1.0에서 채택이 되었었다. Chen 고속 DCT 방법의 FDCT (Forward DCT) 구조는 〈그림 1〉과 같으며 DCT 계수인 C_k 는 k값이 1에서 15일 때까지 $\cos(k \cdot \pi / 32)$ 에 64를 곱하여 정수 형태로 계산 가능하며 〈표 2〉와 같다. 미리 정의된 근사화 값을 이용하기 때문에 정수 곱셈과 시프트 연산으로만 변환 과정을 수행할 수 있는 이점을 갖는다.

2. 16 bit 데이터표현을 통한 변환설계[5]

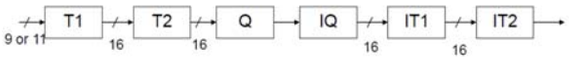
Cisco와 IT에서 제안된 기술은 JCTVC 5차 회의에서 HM3.0에 채택된 변환 커널 설계 기술로서 모든 크기 (4x4, 8x8, 16x16, 32x32)의 변환커널을 제안하였으며, Chen의 고속 DCT 방법과 마찬가지로 작은 크기의 변환이 재사용 가능하며, 커널의 계수도 DCT 매트릭스의 값에 64만큼 값을 곱한 후 반올림하여 정수로 표현하여 정수연한만 수행한다. 4x4와 8x8 변환계수는 아래와 같다.



〈그림 1〉 Chen 고속 FDCT 구조[4]

〈표 2〉 Chen 고속 DCT의 계수

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}
63/64	62/64	61/64	59/64	56/64	53/64	49/64	45/64	40/64	35/64	30/64	24/64	18/64	12/64	6/64



〈그림 2〉 제안된 16 비트내 연산이 가능한 구조

$$T_{4\times4}=\begin{bmatrix}64 & 64 & 64 & 64\\ 83 & 36 & -36 & -83\\ 64 & -64 & -64 & 64\\ 36 & -83 & 83 & -36\end{bmatrix},$$

$$T_{8\times8}=\begin{bmatrix}64 & 64 & 64 & 64 & 64 & 64 & 64 & 64\\ 89 & 75 & 50 & 18 & -18 & -50 & -75 & -89\\ 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83\\ 75 & -18 & -89 & -50 & 50 & 89 & 18 & -75\\ 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64\\ 50 & -89 & 18 & 75 & -75 & -18 & 89 & -50\\ 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36\\ 18 & -50 & 75 & -89 & 89 & -75 & 50 & -18\end{bmatrix}$$

제안된 방법은 작은 크기의 변환 커널 계수를 따로 저장할 필요가 없이 큰 변환 커널 계수만을 이용해서 작은 커널의 계수를 얻어낼 수 있다. 또한, H.264/AVC와는 다르게 1-norm 즉, 모든 계수의 위치에서 동일만 norm을 사용했기 때문에 모든 위치에 대해 QP에 따라 같은 양자화 스텝만 적용하면 된다. 따라서 양자화 매트릭스를 따로 저장할 필요 없고 양자화 및 역양자화 과정이 간단하다. 또한, 16 비트 내에서 연산이 수행되도록 절삭연산(clipping operation)을 〈그림 2〉와 같이 수행한다.

절삭연산을 수행하는 이유는 2-D 분리변환(2-D separable transform)을 잔차신호에 적용하는 경우 변환계수의 값만큼 결과 값이 커지게 되어 연산 결과 데이터의 동적 범위(dynamic range)가 16 비트를 초과 할 수 있다. 동적 범위가 16 비트를 초과할 경우 추가적인 저장공간이 요구되므로, 이런 경우 추가적인 하드웨어 복잡도를 유발을 피하기 위해 16 비트내에서 변환 연산이 가능하도록 하는 절삭연산 방법을 제안하였다. 절삭 연산을 적용하더라도 부호화 효율 성능에는 큰 차이를 보이지 않아 최종적으로 HM 3.0부터 채택이 되었다.

Ⅲ. HM의 선택적 변환 방법 연구

현재 제안된 HEVC 선택적 변환 방법은 크게 두 가지로 나눌 수 있다. 첫째는 DCT/DST 또는 DCT/KLT의 선택적 변환 방법이고, 두 번째는 DCT 변환이 수행된 이후 저주파 영역에 대해 추가적인 변환을 수행하는 2차 변환 방법이다. 일반적으로 두 방법이 서로 비트-왜곡 성능 면에서 동일한 작용을 하기 때문에 경쟁관계에 있다. 하지만 2차 변환 방법이 선택적 변환 방법을 모든 크기에 적용하는 것보다 적은 연산을 필요로 하기 때문에 작은 크기(4x4)의 화면내 예측 부호화의 경우 선택적 변환 방법이 주로 제안 및 채택 되었으며, 큰 크기(8x8, 16x16, 32x32)의 화면내 예측 부호화의 경우에는 2차 변환 방법이 제안되었다. 나머지 부분에서는 두 종류의 기술들에 대해 자세히 기술한다.

1. 모드에 따른 선택적 변환 기술(Mode Dependent Alternative Transform)

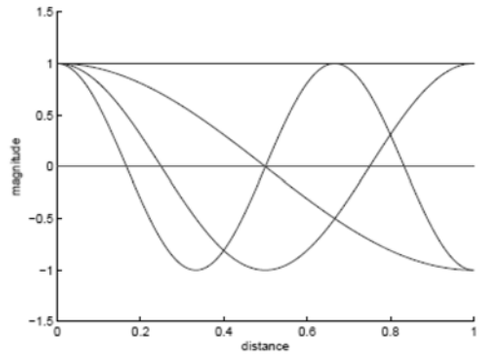
선택적 변환의초기 기술은 퀄컴에서 제안한 화면내 예측 부호화에서 방향성 모드에 따라 선택적 변환을 적용하는 MDDT(mode dependent directional transform)이다[6]. MDDT는 모드에 따라 잔차신호의 각 열과 행의 공분산이 서로 동일하다는 가정하에 2-D separable KLT 변환을 수행한다[7]. MDDT 변환 커널은 최적으로 선택된 방향성 모드에 따른 잔차신호들의 공분산행렬(covariance matrix)에 SVD(singular vector decomposition)를 적용해 구한다. 하지만 이러한 학습된 변환 커널은 학습 방법에 따라 커널의 성능 차이가 생기며, 정확한 해결책이라고 할 수 없다.

따라서 유도된 변환 커널 기술이 현재 논의되고 있다. 변환커널을 유도하기 위해 1차 가우스 마코프 신호(First-order Gauss-Markov)로 가정한 이미지의 상관관계계수(correlation coefficient)를 정의하고, 각 화면내 방향성 예측모드에 따라 상관관계계수를 이용해 유도된 커널을 구한다. 변환커널을 유도하는 방법은 삼성전자[8]와 I2R[9]에서 제안하였으며 결론적으로 상관도가 높고, 방향성이 있는 경우 DST를 적용하는 경우 높은 부호화 효율을 갖는 것으로 보고되었다. 삼성전자가 제안한 기술을 개념적으로 자세히 다음절에서 상세 설명한다.

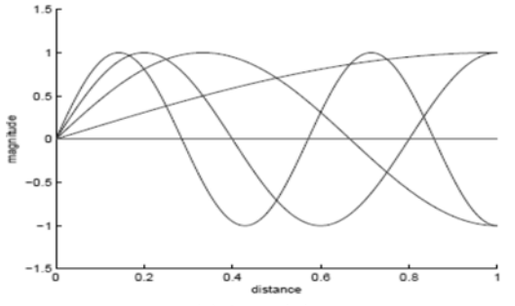
일반적으로 〈그림 3〉과 같이 단방향 화면내 예측이 수행된 후 잔차신호는 참조 신호에서 멀어질수록 커지게 된다. 이러한 경우 잔차신호의 특성은 〈그림 4〉-(a)의 cosine 기저(basis)를 이용해 표현하는 것 보다 〈그림 4〉-(b)의 sine 기저를 이용해 표현하는 것이 신호의 압축 성능 측면에서 유리하다. 왜냐면 방향에 따라 잔차신호가 증가하는 것과 증가하는 기저의 상관도가 높아 에너지가 저주파 영역에 잘 압축되기 때문이다.

예를 들어 잔차신호에 가로/세로 방향으로 아래와 같이 변

〈그림 3〉 단방향예측도식화[8]



(a) Cosine 기저



(b) Sine 기저

〈그림 4〉 cosine/sine 기저[8]

환이 적용된다고 가정하자.

$$Y = CXR$$

이때 X는 잔차신호이고, C는 수직(열 – column) 변환 커널로 잔차신호의 수직방향으로의 신호를 압축하기 위해 적용하며, R은 수평(행 – row) 변환 커널로 잔차신호의 수평방향으로 신호를 압축하기 위해 적용한다. 위의 개념을 적용하여 화면내 예측이 수직방향모드(vertical directional mode)의 경우 수직 방향의 신호를 압축하기 위한 커널은 DST가 적용되어야 하므로 위 식에서 C는 DST가 R은 DCT가 적용되었을 때보다 효과적인 압축효율을 얻게 된다는 것을 직관할 수 있다. 위 방법을 사용해서 〈표 3〉과 같이 화면내각 방향성 예측 모드 별로 DCT/DST를 사용하여 압축하도록 제안 하였으며 TMuC 0.9에 구현하였을 때 복잡도 증가 없이 고효율 모드

〈표 3〉 방향성모드에 따른 DCT/DST 사용

Mode	Vertical (Column) Transform	Horizontal (Row) Transform
VER to VER+8	DST	DCT
HOR to HOR+8	DCT	DST
DC	DCT	DCT
VER–8 to VER–1	DCT	DCT
HOR–8 to HOR–1	DCT	DCT

인트라(High Efficiency Intra)에서 1.2%, 저복잡도 모드 인트라(Low Complexity Intra)에서 1.9%의 압축 효율 성능향상이 있는 것으로 보고되었다. 최종적으로 5차 JCTVC 회의에서는 4x4 화면내 방향성 모드에 따라 DCT/DST 변환 조합에 따른 선택적 변환 기술을 HM3.0에 채택하였다.

또한, 최근 6차 회의에서 8x8까지 확장한 방법과 색차(chroma) 신호에 그대로 적용한 방법을 제안하였으며 4x4 색차 블록에 대한 동일한 방법이 CE에 채택되어 다음 회의에서 논의될 예정이다.

2. 2차 변환 방법 (the secondary transform)

2차 변환 방법은 1차 JCTVC 회의에서 삼성전자가 제안한 ROT (Rotational Transform)[4,10]이 대표적인 예로 〈그림 5〉와 같다.

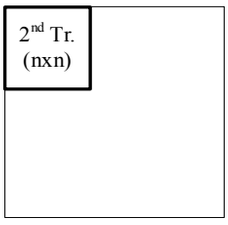
현재 부호화하고자 하는 블록의 화소 값에 대한 예측 과정이 수행된 후 구해진 잔차신호에 대해 DCT를 적용을 한번 저주파 영역에 대해 2차로 변환을 수행하는 방법이 2차 변환 방법이다. 현재 여러 기관에서 비슷한 개념으로 제안을 하고 있으나 아직 채택된 기술은 없다.

2차 변환의 경우 특별히 큰 크기의 변환에 적용하는 것이 유리한데, 일반적으로 방향성에 따라 DCT/DST 또는 DCT/KLT를 적용하는 경우 압축 효율이 높아진다. 하지만 큰 크기의 변환에서는 모든 영역에 대해 2차 변환을 수행하는 것이 복잡도가 높을 뿐만 아니라 실제 큰 크기의 경우 양자화를 고려했을 때 고주파수에서는 신호의 에너지가 작은 관계로 대부분 0의 값으로 양자화 되고, 저주파수 영역에서만 유효하므로 저주파 영역에 대해서만 2차 변환을 수행하는 것이 합당하다. 따라서 대부분 제안된 기고서[10]는 〈그림 6〉과 같이 저주파수 영역에 대해서만 적용을 하며 일반적으로 8x8 크기 또는 4x4크기의 저주파수 영역에 대해 적용한다.

2차 변환 적용하는 방법을 방향성 모드의 고려 여부에 따라 크게 두가지로 나눌 수 있다. 먼저 방향성 모드를 고려하지 않고 2차 변환을 적용한 방법은 NHK에서 제안한 적응적 DCT/DST 선택방법과 삼성전자에서 제안한 ROT이다[10,11]. NHK에서 제안한 적응적 DCT/DST 방법은 화면내/화면간 모두에 적용한 가능한 방법이며 DCT와 DST를 사용할 것인지를 율–왜곡 비용 관점에서 선택하는 방법이다.

삼성전자가 제안한 ROT의 경우 방향성 모드와 상관없이 8x8, 16x16, 32x32 크기의 기존 DCT가 적용된 후 8x8 저주

〈그림 5〉 2차 변환 ROT의 개념도[10]



〈그림 6〉 저주파영역에 적용하는 2차 변환개념도 [10]

파수 영역에 대해 ROT를 적용하며 그때 식은 아래와 같다.

$$Y = R_v^T D^T X' D R_h$$

이때, X'는 잔차 신호 X에서 저주파 영역만을 취한 것이고, D는 기존 DCT 커널, R_v는 수직방향의 ROT 커널, R_h는 수평 방향의 ROT 커널이다. 각각의 경우 해당 각도에 따라 회전 매트릭스를 정의하여 5개의 매트릭스 중에서 율-왜곡 관점에서 가장 성능이 높은 것을 선택한다. 따라서 여러 후보각도에 따른 율-왜곡 값을 계산해야 하기 때문에 복잡도가 20~50% 증가하고, 1% 정도의 성능향상이 있다. 2차 변환에서 가장 높은 성능향상을 가져오나, 복잡도가 높아 채택되지 않았다.

두 번째 방법들은 파나소닉과 삼성에서 제안하였으며 방향성 모드를 고려해서 2차 변환을 적용한다[12,13]. 파나소닉에서 제안한 모드에 따른 방법은 JCTVC-B024와 JCTVC-C108에 소개된 이론적 유도 과정을 통해 8x8, 16x16, 32x32 크기의 변환에서 1차 DCT 를 적용한 이후 저주파 4x4 크기의 변환 계수에 대해 2차 KLT 변환을 수행하는 방법이다. 반면, 삼성전자의 제안 기술은 8x8, 16x16, 32x32 크기의 변환 중 4x4 저주파 영역에 대해 2차 변환을 적용하는 것은 동일하지만, KLT 대신 DST를 사용하도록 설계하여 기존의 모드에 따른 4x4 DCT/DST 방법에 구조를 그대로 이용할 수 있도록 제안하였다.

Ⅳ. 결 론

본 논문에서는 HM에서 제안 및 채택된 변환부호화 방법으로서 크게 핵심 변환 커널 설계 방법과 추가적/선택적 변환 방법 2가지로 나누어 설명하였다. 변환 커널이 부호화/복호화 연산을 수행할 때 가장 중요한 과정중의 하나이므로 변환커널을 설계 할 때 소프트웨어와 하드웨어의 복잡도를 모두 고려하고, 작은 크기의 커널 계수가 큰 크기의 커널에 재사용 가능하도록 논의되고 있다.

추가적 선택적 변환 방법의 경우 4x4 크기의 변환에서는 DCT/DST를 모드에 따라 선택적으로 사용하도록 하는 방법이 채택되었으며, 8x8 이상 크기 변환에서는 2차 변환을 모드에 따라 적용하는 방법이 논의되고 있다.

참고문헌

[1] http://ftp3.itu.int/av-arch/jctvc-site/
[2] JCT-VC, “Test Model under Consideration (TMuC),” JCTVC-B205_draft007, Geneva, Switzerland, Jul 2010.
[3] JCT-VC, “HM3: High Efficiency Video Coding (HEVC) Test Model 3 Encoder Description” JCTVC-E602, Geneva, CH, 16-23 March, 2011.
[4] JCT-VC, “Samsung’s Response to the Call for Proposals on Video Compression Technology” JCTVC-A124, Dresden, DE, 15-23 April, 2010.
[5] JCT-VC, “Transform design for HEVC with 16 bit intermediate data representation” JCTVC-E243, Geneva, CH, 16-23 March, 2011.
[6] Y. Ye and M. Karczewicz, “Improved H.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning,” IEEE Int.Conf. Image Process 08 (ICIP08), San Diego, U.S.A., Oct., 2008.
[7] Gerbrands, J. J. "On the relationships between SVD, KLT and PCA," Pattern Recognition 14(1-6): 375-381.
[8] JCT-VC, “Jointly optimal intra prediction and adaptive primary transform” JCTVC-C108, Guangzhou, CN, 7-15 October, 2010.
[9] JCT-VC, “Mode-Dependent Fast Separable KLT for Block-based Intra Coding” JCTVC-B024, Geneva, CH, 21-28 July, 2010.
[10] JCT-VC, “CE 7: Experimental Results for the Rotational Transform” JCTVC-F294, Torino, IT, 14-22 July, 2011.
[11] JCT-VC, “CE7.5: Performance analysis of adaptive DCT/DST selection” JCTVC-F229, Torino, IT, 14-22 July, 2011.
[12] JCT-VC, “CE7: Mode Dependent 2-step Transform for Intra Coding” JCTVC-F224, Torino, IT, 14-22 July, 2011.
[13] JCT-VC, “On secondary transforms for intra prediction residual” JCTVC-F554, Torino, IT, 14-22 July, 2011.

김 재 일

2001년 한국기술교육대학교 정보통신공학과 전자공학전공 학사.
2001년~현재 한국과학기술원 정보통신공학과 석박사 통합과정.
〈관심분야〉 비디오부호화, 패턴인식, 변환기술, 방송 시스템

김 문 철

1989년 강북대학교 전자공학과 학사.
1992년 Univerisity of Florida, Electrical & Computer Engineering 석사.
1996년 Univerisity of Florida, Electrical & Computer Engineering 박사.
1997년~2001년 ETRI, 방송미디어연구부, 선임연구원.
2001년~2009년 한국정보통신대학교 공학부, 조교수/부교수.
2009년~현재 한국과학기술원 전기및전자공학과 부교수.
〈관심분야〉 비디오부호화, 패턴인식, 통계적기계학습, 멀티미디어방송