
SVC 복호화기에서 Inter Layer 업-샘플링의 효과적인 구조

기대욱* · 김재호**

An Efficient Architecture of Inter Layer Up-sampling in Scalable Video Decoder

Dae-Wook Ki* · Jae-ho Kim**

이 논문은 지식경제부 성장동력개발사업의 연구결과로 수행됨

요 약

본 논문에서는 SVC 복호화기에서 각 계층간 Inter layer 업-샘플링을 효과적으로 구현하기 위한 하드웨어 구조를 제안한다. 제안하는 구조에서 수직, 수평 방향 업-샘플링을 위한 register bank와 보간 모듈이 설계된다. 제안 구조를 사용하여 SRAM 메모리가 감소되고 JSVM과 비교해서 약 41%의 메모리 밴드위스가 감소되었다.

ABSTRACT

This paper proposes an efficient architecture of Inter layer up-sampling in decoder for SVC(scalable video coding). A register bank for horizontal and vertical up-sampling and interpolation units are designed, by introducing the proposed architecture, 41% memory bandwidth is reduced compared to JSVM. For real-time operation for HD 6 layer decoder having CIF, SD, HD resolution for CGS layer, the hardware is designed to operate at 127MHz. The gate count is about 3000.

키워드

스케일러블 비디오 코딩, 레이어간 업 샘플링

Key word

Scalable Video Coding, SVC, Inter layer Up-sampling, JSVM

* 부산대학교 전자전기공학과 박사과정

** 부산대학교 전자전기공학과 교수

접수일자 : 2009. 10. 21

심사완료일자 : 2010. 02. 10

I. 서 론

현재의 네트워크와 단말 환경은 다양한 통신환경과 DMB(Digital Multimedia Broadcasting)와 IPTV 등으로 대표되는 통방융합 등의 비 균일 통신망이 서로 융합하고 있는 추세이며, 다양한 멀티미디어 기기의 보급은 다양한 화면 해상도와 연산 능력을 가진 단말기의 다양성과 이종 디지털 기기간의 디지털 융합(Digital Convergence)로 표현된다. 압축 효율을 최대화시키는 것뿐만 아니라 다양한 단말기 및 변화하는 통신 환경에 적응적으로 대응하여 끊김 없는 품질을 보장하기를 원하는 사용자의 요구를 만족시키기 위해 최근 스케일러블 비디오 부호화(SVC)가 표준화 되었다 [1], [2].

SVC의 핵심 기능은 한 번의 압축된 비트-스트림에서 서로 다른 여러 종류의 프레임 율(Frame rate), 해상도(Resolution), 화질(Quality)을 갖는 영상을 다양한 디바이스와 다양한 네트워크 환경에 적응적으로 복원할 수 있도록 하는 것이다. SVC에서는 H.264/AVC에 이미 정의된 기술들과 추가적인 기술들을 결합함으로써 스케일러빌리티 특징을 유지하면서도 높은 압축 효율을 달성하였다. 공간적 스케일러빌리티(Spatial scalability)를 위해 다중 계층(Multi-layer) 구조를 택하였으며, 계층간 중복도 제거를 위해 계층간 예측(Inter-layer prediction)을 사용하였다.

시간적 스케일러빌리티(Temporal scalability)를 위해 계층적 B-picture(Hierarchical B picture) 구조를 택하여 시간적 분해 레벨을 갖는다. 또한 화질적 스케일러빌리티(SNR scalability)를 위해 FGS(Fine Grain Scalability) 기술 등을 사용하였다[3]. 특히, 공간적 스케일러빌리티의 계층간 업-샘플링은 SVC SD급 6계층을 수직 수평의 화소보간을 하여야 하며, SVC 시스템에서 높은 복잡도와 많은 횡수의 메모리 접근이 필요하며, 디코딩을 하기위한 많은 양의 외부메모리가 필요하다. 따라서 실시간 처리를 위해서는 저전력의 최적화된 하드웨어 구조가 요구된다.

본 논문에서는 효율성을 개선할 수 있는 SVC 복호화를 위한 계층간 업-샘플링(Inter layer Up-sample) 하드웨어 구조를 제안한다. 메모리 접근 횡수를 감소시키기 위해 블록 데이터(Block Data) 접근 후 수직, 수

평 업-샘플링을 하는 연산구조를 채택하고, 효율적인 연산을 위해 수직, 수평 업-샘플링을 위한 레지스터 뱅크(Horizontal, Vertical Up-sampling Resister Bank, HVURB) 구조를 제안 한다. 이렇게 함으로서 실시간 데이터를 이용하여 수평방향 업-샘플링과 수직방향 업-샘플링을 바로 이어 수행하여 데이터 밴드width를 감소시킨다. 제안된 Inter-layer 업-샘플링 하드웨어 구조를 FPGA(Altera StratixII)를 타겟으로 구현한 결과 127MHz 동작, 그리고 3K 게이트(Gate size)로 합성되었음을 확인하였다. 이는 SVC SD급 SVC 복호화(Decoder)를 위해서 30fps(frame per second)로 실시간 처리가 가능하게 한다.

II. SVC의 계층간 화면내 예측방식

2.1 SVC의 inter-layer 업-샘플링

공간적 스케일러빌리티를 위해 SVC에서는 각 계층이 다른 해상도의 영상으로 부호화되는 다중 계층(Multi-layer) 구조를 채택하였다. 각 계층은 H.264/AVC 방식으로 부호화 되고, 공간 계층간 중복된 정보를 이용하기 위하여 현재 계층의 공간 해상도보다 한 단계 낮은 공간 해상도의 복원된 영상을 예측 신호로 사용하는 계층간 예측(Inter-layer prediction) 기술이 사용된다. 계층간 예측 방식들 중 Base Layer Intra (BL Intra)모드는 복호화하려는 현재 계층의 매크로블록에 대응하는 하위 계층의 복원된 영상을 현재 계층의 매크로블록의 공간 해상도로 업-샘플링(Up-sampling)하여 예측신호로 사용한다. 식 1과 식 2는 상위계층의 위치 X, Y 에 해당하는 하위 계층의 대응하는 위치와 필터 계수를 알 수 있는 중간 값 \hat{x}, \hat{y} 를 계산 하는 수식이다 [3], [4].

$$\hat{x} = \text{Round}\left(\frac{X \times D_x + R_x}{2^{S-4}}\right), \tag{1}$$

$$\text{where } D_x = \text{Round}\left(\frac{2^S \times \text{BaseWidth}}{\text{ScaledBaseWidth}}\right)$$

$$\hat{y} = \text{Round}\left(\frac{Y \times D_y + R_y}{2^{S-4}}\right), \quad (2)$$

$$\text{where } D_y = \text{Round}\left(\frac{2^S \times \text{BaseHeight}}{\text{ScaledBaseHeight}}\right)$$

여기서 *ScaledBaseWidth*, *ScaledBaseHeight*, *S*, *BaseWidth*, *BaseHeight*, 는 상위계층의 넓이, 상위계층의 높이, 스케일 계수, 하위계층의 넓이, 하위계층의 높이를 각각 나타낸다. JSVM에서는 *S*의 값으로 16을 사용한다[5]. 그리고, R_x , R_y 은 오프셋(offset)으로 JSVM에서는 211을 사용한다[5]. 식 1과 식 2의 결과 값인 \hat{x} , \hat{y} 으로부터 먼저, 식 3에서 필터계수를 선택하기 위한 $Phase_x$, $Phase_y$ 를 결정한다.

$$Phase_x = \hat{x} \& 16, \quad Phase_y = \hat{y} \& 16 \quad (3)$$

다음으로, 식 4에서 하위계층의 위치 x , y 를 결정한다.

$$x = \hat{x} \gg 4, \quad y = \hat{y} \gg 4 \quad (4)$$

식 3과 식 4에서 결정된 Phase정보를 이용하여 계층간 업-샘플링 때 사용되는 4개의 필터 계수를 결정한다. 표 1은 16개의 Phase와 그에 대한 필터 계수를 나타낸다.

SVC에서는 복호화하려는 현재 계층의 매크로블록에 대응하는 하위 계층의 복원된 잔여(Residual) 영상을 현재 계층의 매크로블록의 공간 해상도로 업-샘플링(Up-sampling)하여 잔여영상으로 사용한다. 이러한 방법으로 하위 계층의 잔여를 사용함으로써, 현재 계층의 매크로블록의 잔여 신호를 감소되어 부호화 된다. 잔여 신호의 업-샘플링을 위한 보간 필터(Interpolation filter)는 2탭 필터를 사용하며, 대응하는 위치의 소수점 이하 값(fractional value)에 따라 다른 필터 계수가 적용된다. BL Intra 업-샘플링의 식 1에서부터 식 4와 같은 방법으로 대응되는 하위계층의 위치와 필터 계수를 구한다. 계층간 잔여 신호 예측은 매크로블록에 따라 적용

적으로 수행된다. 계층간 잔여 신호 예측을 사용한 결과의 잔여신호가 더 커진 경우 계층간 잔여 신호 예측은 사용하지 않는다. 계층간 잔여 신호 예측의 여부는 플래그로 나타낸다.

표 1. BL Intra 업-샘플링의 필터 계수
Table. 1 Filter coefficients of BL Intra Up-Sampling

Phase	C_0	C_1	C_2	C_3
0	0	32	0	0
1	-1	32	2	-1
2	-2	31	4	-1
3	-3	30	6	-1
4	-3	28	8	-1
5	-4	26	11	-1
6	-4	24	14	-2
7	-3	22	16	-3
8	-3	19	19	-3
9	-3	16	22	-3
10	-2	14	24	-4
11	-1	11	26	-4
12	-1	8	28	-3
13	-1	6	30	-3
14	-1	4	31	-2
15	-1	2	32	-1

2.2 기존의 업-샘플링 구현 방식

기존 H.264/AVC에서는 1/4 화소 정밀도 업-샘플링을 Motion Compensation (MC)를 위해 수행한다. 이를 위해 H.264/AVC에서는 6탭과 bi-linear 보간(Interpolation) 방법을 사용하였다 [7]. H.264/AVC의 엔코더, 디코더 시스템에서 1/4 화소 보간 방법으로 인해 복잡도가 증가되었고, 또한 많은 양의 데이터를 메모리로부터 읽기/쓰기(Read/Write)를 하기 때문에 메모리의 사용량이 증가 되었다. 따라서 Dajing Zhou[8]는 메모리 접근을 감소시키며 Bandwidth를 줄이기 위해 2D array 메모리를 제안하여 수직 수평방향의 데이터 접근을 효율적으로 하였고, Nai-Ran Zhang[9]는 수직방향의 4개의 FIR 필터와 수평방향의 9개 FIR 필터로 구성된 병렬(Parallel)구조를 제안 함으로서 메모리 접근과

Bandwidth을 절약 하였다.

SVC의 화소 보간(Interpolation) 알고리즘은 H.264/AVC의 6탭 필터의 계수가 한가지인 것과 달리, 16종류의 4탭 필터 중 한가지를 적응적으로 선택하는 구조이다. 그리고 식 1에서 식 4와 같이 대응되는 하위계층의 위치 값과 Phase를 각각 상위계층의 위치에 대해서 계산을 하여야 한다. 따라서 SVC의 화소 보간 방법은 2007년 7월에 표준안 된 레퍼런스 소프트웨어(JSVM)이 완성 되었으며, 기존의 H.264/AVC의 화소 보간 구현방식과 다른 새로운 하드웨어 구현 필요 하다. 본 논문을 위해 조사해본 결과 현재까지 SVC 하드웨어 구조를 제안한 사례를 찾을 수 없었다.

III. 제안방식

본 장에서는 BL_Intra 모드의 업-샘플링을 위한 하드웨어 구조를 제안한다. 하드웨어 구조는 그림 1과 같다. 그림 1은 크게 레지스터인 HVURB, 수평/수직 방향 필터 계수 컨트롤러인 Hor_Filter/Ver_Filter, 그리고 수평/수직방향업-샘플링 모듈인 Horizontal Up-sampling/Vertical Up-sampling 모듈로 구성된다. 먼저, 외부 메모리에서 읽은 Base Layer의 데이터를 HVURB의 내부 메모리인 4x16 Register에 저장 한다. 내부 메모리 구조는 입력으로 클락(clock)과 수평방향, 수직방향 시프트 신호인 H_SFT와 V_SFT가 입력되며, H_SFT와 V_SFT에 따라 내부 메모리의 데이터가 수평방향 또는 수직방향으로 시프트 된다. 한번의 수직방향 시프트가 수행 되면 Hor_Filter에서 선택된 계수들을 이용하여 한번의 수평방향 업-샘플링이 수행되고, 업-샘플링된 하나의 데이터와 이전에 업-샘플링 되었던 3개의 데이터들을 Ver_Filter에서 선택된 계수들을 이용하여 수직방향으로 2번 업-샘플링을 수행한다.

16번의 수직방향 시프트가 수행되면서 첫 번째 데이터가 본래의 자리로 돌아오면 다시 반복하여 1번의 수평방향 시프트, 16번의 수직방향 시프트를 수행한다. Horizontal Up-sampling과 Vertical Up-Sampling 모듈에서 계산된 상위 계층의 데이터는 출력되어 SDRAM에 저장된다. 그림 2는 HVURB의 인터페이스이다.

제안 방식에서는 수평방향 업-샘플링 수행 후 바로 수직방향 업-샘플링을 수행하기 때문에 수평방향 업-샘플링된 모든 값을 메모리에 저장할 필요가 없다. 단지 HVURB의 레지스터에서 수직 방향으로 13, 14, 그리고 15 번 위치 데이터들의 수평 방향 업-샘플링 수행 후 계산된 수평방향업-샘플링된 화소값들만 아래 매크로 블록의 업-샘플링을 위해서 SRAM에 저장 된다. 제안한 하드웨어 구조는 수평방향의 화소보간을 할 때 수직방향의 화소보간을 바로 수행할 수 있는 효율적인 메모리 구조를 사용하여, 모든 수직방향의 업-샘플링된 데이터를 읽기/쓰기를 할 필요가 없다. 또한, 수평방향의 화소보간을 결과 값을 외부메모리(SDRAM)가 아닌 내부 메모리(SRAM)에 쓰고 읽기 가능한 구조 제안하여, 요구되는 외부 메모리의 크기를 감소 하였으며, 메모리의 접근 횟수를 줄이므로 JSVM에서의 요구되는 메모리와 Bandwith를 167MB/ Sec(SDRAM)에서 99MB/Sec(SDRAM) + 8.44KB(SRAM) 으로 절약 한다. 그림 3은 이를 나타낸다.

잔여신호 업-샘플링(Residual up-sample) 알고리즘은 16개의 Phase를 가지는 2탭 필터(Bilinear interpolation)를 사용한다. Phase는 BL Intra 식 1에서 식4와 동일하게 결정된다. 잔여신호 수평방향 업-샘플링 수식은 식 5와 같다.

$$E(X, Y) = \begin{cases} 16 \times B(x, y), & \text{if } X \% 4\alpha = 0, \\ 16 \times B(x-1, y), & \text{else if } X \% 4\alpha = 7, \\ (16 - Phase) \times B(x-1, y) + Phase \times B(x, y), & \text{else,} \end{cases}$$

$$\text{where, } x = \begin{cases} \text{Int}(X / \alpha) & \text{when, } X = \text{even} \\ \text{Int}(X / \alpha + 1) & \text{otherwise} \end{cases}$$

(5)

식 6은 수직 방향의 잔여 신호 업-샘플링 수식이다.

$$E(X, Y) = \begin{cases} 16 \times B(x, y), & \text{if } Y \% 4\alpha = 0, \\ 16 \times B(x, y-1), & \text{else if } Y \% 4\alpha = 7, \\ (16 - Phase) \times B(x, y-1) + Phase \times B(x, y), & \text{else,} \end{cases}$$

$$\text{where, } y = \begin{cases} \text{Int}(Y / \alpha) & \text{when, } Y = \text{even} \\ \text{Int}(Y / \alpha + 1) & \text{otherwise} \end{cases}$$

(6)

제안하는 잔여신호 업-샘플링 구조에서 첫 라인(line)의 수직방향의 Bilinear interpolation를 하지 않고, 다음 라인(line)의 Bilinear interpolation의 결과값과 첫 라인에서 수평방향 화소보간한 화소 값과 수직방향 화소보간(VERTICAL interpolation)을 한다. 잔여신호 업-샘플링(Residual up-sample)은 SVC 인코더에서 적응적으로 On/Off 를 할 수 있다. 하위 계층의 잔여신호(Base Layer의 Residual Data)를 외부 메모리로부터 읽어(Read) Register1(R1)에 저장하고 이웃하는 하위계층의 잔여신호(Base Layer의 Residual Data)를 R0에 저장하여 수평방향의 시프트(Horizontal shift)를 하여 R1는 R0으로 이동시켜서 Bilinear interpolation을 수행한다. 이때 하위계층의 위치에 따라서 필터계수가 적응적으로 선택된다. 그림 4는 제안한 잔여신호 업-샘플링의 하드웨어 구조이다. 그림 4에서 R0, R1은 하위계층 잔여신호 화소 값이며, R0은 $B(x-1,y)$, R1은 $B(x,y)$ 의 화소 값이다. BI(Block Index)은 블록 인덱스이며, $BI0=Int(x-1)/4$, $BI1=Int(x)/4$ 이다. $B(x-1,y)$ 의 블록 인덱스 BI0과 $B(x,y)$ 의 블록 인덱스 BI1이 같으면 2탭 업-샘플링을 하며, 다를 경우에서 하위계층의 위치에 따른 Phase에의 화소 값이 결정된다. 수직 방향의 하드웨어 구조는 동일하다.

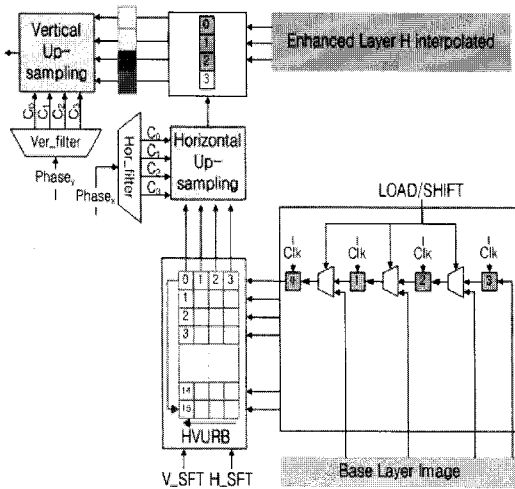


그림 1. BL Intra의 하드웨어 구조
Fig. 1 The hardware architecture of BL Intra

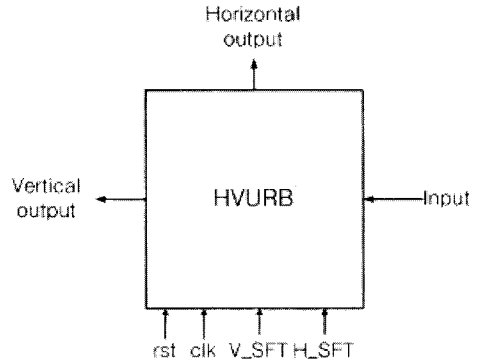


그림 2. 수평 수직 업-샘플링 레지스터 뱅크의(HVURB) 인터페이스
Fig. 2 The Interface of HVURB

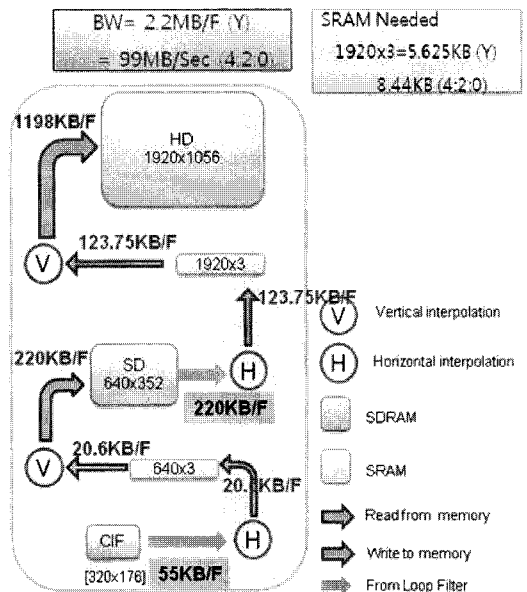


그림 3. 제안한 하드웨어의 Bandwidth와 메모리
Fig. 3 The memory and bandwidth of proposed Hardware

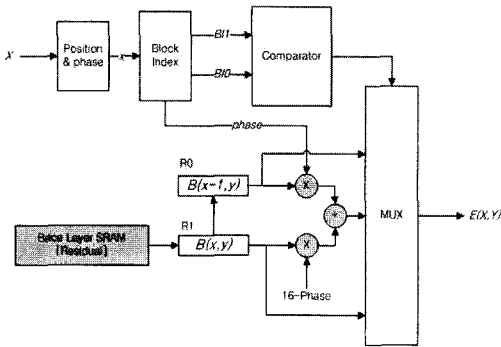


그림 4. 잔여신호 업-샘플링의 하드웨어 구조
Fig. 4 The hardware architecture of residual up-sampling

IV. 실험결과

본 논문에서 제안된 하드웨어의 구조를 VerilogHDL로 구현하고 Altera Quartus II v8.1에서 합성(Synthesis)하였다. 타겟 FPGA는 Altera Stratix II이다. 테스트 환경은 표 2와 같다. 제안 방식의 모듈은 127Mhz의 동작 주파수와 3K의 게이트로 합성되었다. 합성 후 FPGA에서 사용된 LE(logic Element)의 수는 Altera 사 Stratix II 제품의 데이터 시트(Data sheet)를 참고하여 Altera Quartus II에서 측정된 ALUTs(Adaptive Look-Up Table)의 수를 1.25배로 환산하여 계산하였다[6]. 총 4x16블록 당 최대 652 사이클이 소요되므로, 한 HVURB 당 SD급 404,800 (880 * 460) cycles 이 소요되고, HS급은 2,284,480 (3520 * 649) cycles 이 소요 된다. 구현된 모듈을 127MHz에서 동작 하여, 3.37ms, 19.04ms이고, SVC HD 6 layer(CIF, CIF(CGS), SD, SD(CGS), HD(CGS))에서 한 HVURB를 처리하는 데 요구되는 시간은 22.41ms 이다.

표 2. 테스트 환경
Table. 2 Environments for test

RTL Description Language	Verilog HDL
RTL Simulation Tool	Modelsim SE 6.0d
Synthesis	Altera Quartus II v8.1
Target FPGA	Altera Stratix II EP2S180F1020C3

V. 결론

본 논문은 SVC 복호기에서 inter-layer 업-샘플링을 위한 Intra 업-샘플링과 잔여 신호의 업-샘플링의 하드웨어 구조를 제안하였다. 효율적인 수평, 수직 업-샘플링을 위한 HVURB를 구현하여 SRAM 크기를 감소시키고 메모리 밴드width를 감소시켰다. 제안한 하드웨어 구조를 Altera Stratix II FPGA를 타겟으로 하여 127 동작 주파수로 HD(1920x1052)급 6 layer를 복호화 한다. 합성된 결과를 바탕으로 제안한 하드웨어 구조는 HD 6 layer에서 초당 30 프레임으로 실시간 복호화 될 수 있다.

참고문헌

- [1] T. Wiegand, G. J. Sullivan, J. Reichel, H. Schwarz, and M. Wien (Eds.), Amendment 3 to ITU-T Recommendation H.264 (2005) | ISO/IEC 14496-10:2005, Scalable video coding, July 2007.JTC1/WG11, Doc. m11244, Palma de Mallorca, Spain, Oct. 2004
- [2] Standardization Trend of Next Generation Video Coding after H.264, 전자통신동향분석 제23권 제1호, ETRI, 2008년 2월.
- [3] C.Andrew Segall and Gary J. Sullivan, "Spatial Scalability within the H.264/AVC Scalable Video Coding Extension", IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 9, pp. 1121-1135, Sept. 2007
- [4] Y.Taur and T.H.Ning, *Fundamentals of Modern VLSI Devices*, Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [5] ITU-T and ISO/IEC JTC1, "JSVM 12.1 Software"
- [6] Altera Corporation, "Stratix II Device Family Data Sheet", 2005. http://www.altera.com/literature/hb/stx2/stx2_sii5v1_01.pdf
- [7] Joint Video Team H.264/AVC Reference, Software Version. JM. 15.1 <http://iphome.hhi.de/suehring/ttml/download/>

- [8] Dajiang Zhou, Peilin Liu, "A Hardware-Efficient Dual-Standard VLSI Architecture for MC Interpolation in AVS and H.264," Proceedings of 2007 IEEE International Symposium on Circuits and Systems, 2007
- [9] Nai-Ran Zhang, Mo Li, Wu-Chen Wu, "High Performance and Efficient Bandwidth Motion Compensation VLSI Design for H.264/AVC Decoder," 8th International Conference on Solid-State and Integrated Circuit Technology, 2006.

저자소개



기대욱(Dae-wook Ki)

2003년 2월: 동서대학교 학사
2005년 2월: 부산대학교 석사
2006년 2월~현재: 부산대학교
전자전기공학과 박사과정

※ 관심분야: 영상 압축, 영상 처리, 로직설계



김재호(Jae-ho Kim)

1980년 2월: 부산대학교 학사
1982년 2월: 한국과학기술원 석사
1990년 2월: 한국과학기술원 박사
1992년 2월~현재: 부산대학교 교수

※ 관심분야: 영상통신 시스템, 영상 압축, ASIC