

H.264 High-Profile Intra Prediction 설계

이혜윤^{*}, 이용주^{**}, 김호의^{**}, 서기범^{**}

^{*}우송대학교

A design of High-Profile IP For H.264

Hye-Yoon Lee^{**}, Young-Ju Lee^{**}, ho-eui Kim^{**}, Ki-bum Suh^{**}

^{*}Electronic Dept. Graduate School, Woosong University

E-mail : hylee@wsu.ac.kr, yjlee81@wsu.ac.kr, hekim@wsu.ac.kr

요 약

본 논문에서는 AMBA 기반으로 사용될 수 있는 H.264용 High Profile IP를 제안한다. 설계된 모듈은 한 매크로 블록 당 최대 306 cycle내에 동작한다. 제안된 Encoder 구조를 검증하기 위하여 JM 13.2부터 reference C를 개발하였으며, reference C로부터 test vector를 추출하여 설계된 회로를 검증하였다. 우리는 Hardware cost를 줄이기 위하여 plan mode를 제거하였고, SAD 계산 방법을 사용하여 Hardware cost와 cycle을 줄이는 방법을 채택하였다. 제안된 회로는 133MHz clock에서 동작하며, 합성결과 TSMC 0.18 um 공정에 램 포함 25만 gate크기 이다.

ABSTRACT

In this paper, we propose H.264 High Profile Encoder Intra Prediction module. This designed module can be operated in 306 cycle for one-macroblock. To verify the Encoder architecture, we developed the reference C from JM 13.2 and verified the our developed hardware using test vector generated by reference C. We adopt plan removal and SAD calculation to reduce the Hardware cost and cycle. The designed circuit can be operated in 133MHz clock system, and has 250K gate counts using TSMC 0.18 um process including SRAM memory.

키워드

plan removal, SAD calculation

1. 서 론

1999년 MPEG-4 버전2의 표준화가 기술적으로 완료된 이후 MPEG에서는 사실상 동영상 압축 기술에 대한 표준화가 거의 이뤄지지 않고 있었으나, 영상전화 등 주로 통신시스템에 이용되는 동영상 압축 기술에 대한 표준화를 다루는 ITU-T(국제전기통신연합 전기통신표준화부문)에서는 H.263에 이어 H.26L이라는 이름으로 동영상 압축 기술에 대한 표준화를 지속해 왔다. 아울러 2000년 ITU 표준화 조직개편과 함께 MPEG과 유사한 형태의 VCEG이라는 동영상 압

축 기술 전문가 그룹을 구성해 H.26L의 표준화 작업을 더욱 활발하게 추진했다. MPEG-4 버전2 이후 보다 더 좋은 압축 성능을 보이는 기술들이 등장하기 시작하면서 MPEG에서는 새로운 표준화의 요구가 대두됐다. 같은 요구를 만족시키기 위해 MPEG에서 동영상 압축 기술 성능 비교실험을 수행하게 됐다. 그 결과 당시의 H.26L기술이 MPEG-4에 비해 최고 2배 이상의 성능개선 효과를 보이자 이 기술을 바탕으로 하는 동영상 압축 기술을 MPEG-4의 일부인 부분 10으로 새롭게 표준화하기로 결정했다. 이러한 결정은 새로운 분야에 전문가들의 참여를 필요

로 하는 VCEG의 요구와 부합되면서 2001년 12월 두 표준화 기구가 공동으로 표준화 작업을 하기 위한 JVT를 구성하게 된 것이다. 이것은 ISO/IEC 14496-10 라고 하며, 또한 H.264 라는 이름으로 불린다. H.264는 다양한 네트워크, 특히 서로 다른 네트워크 간에 손쉬운 데이터 이동이 가능하도록 하기 위해 동영상 신호를 압축하는 기능을 담당하는 계층과 이 정보를 네트워크상에서 전송될 수 있도록 데이터를 특정 형태로 변환하는 기능을 담당하는 계층으로 분할된다. 이 가운데 압축 기능을 담당하는 계층을 VCL (Video Coding Layer), 네트워크 적응기능을 담당하는 계층을 NAL (Network Abstraction Layer)이라고 부른다.

최근 JVT의 활동은 Geneva meeting(Jan. 29 - Feb. 1, 2002)와 Fairfax meeting (2002 May 6~10)을 통해 Committee Draft로 승인되고, Klagenfurt Meeting(2002 July 22-26)에서 Final Committee Draft [1] 로 확정되었다. 또한 Geneva Meeting (2002 9-17 October), Awaji meeting(5-13 December, 2002)을 통해서 DIS (Draft International Standard: JVT-F100)의 Document를 발간하였다. 따라서 현재 JVT의 표준은 거의 완성을 이룬 것으로 보인다. 이에 따라 JVT 영상 코덱을 지원하는 하드웨어를 필요로 하는 application에 하드웨어 코덱의 코어의 확보는 시급한 문제가 된다. 따라서 본 논문에서는 JVT Video Encoder에서 핵심적인 코어인 화면 내(Intra) 예측 모듈을 개발하는 것을 목적으로 한다.

본 논문에서는 hardware cost를 줄이고, 부호화 수행 시간을 줄이기 위해 Luma 16x16과 Chroma8x8 모드 중 선택 되는 확률이 다른 모드에 비해 적고 Hardware cost가 높은 plane Prediction mode를 제거하는 방법[2]과 SAD 계산법 그리고 8 pixel 병렬처리 등을 제안한다.

II. 제안된 High Profile IP 구조

그림 1은 High-Profile Intra Prediction 의 하드웨어 구조를 보여준다. 간략한 내부 흐름은 다음과 같다.

- 1 : original data 저장
- 2 : Inter prediction data 저장
- 3 : Intra prediction 수행
- 4 : 수행결과를 Intra prediction ram에 저장
- 5 : post 프로세싱 수행 (DCT/Q/IQ/IDCT)
- 6 : REC 후 DB 모듈 내의 ram에 저장
- 7 : CAVLC 모듈에 필요한 coefficient data 연산

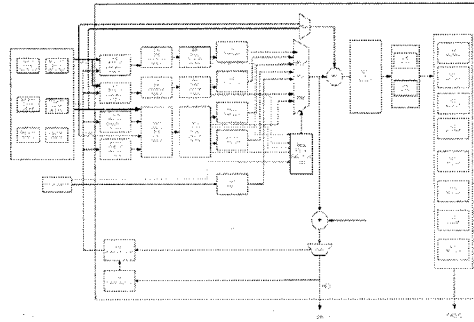


그림 1. Intra Prediction hardware 구조

2.1 dual buffering 기법 사용

그림 2는 High Profile IP 내부의 메모리를 보여준다. Luma의 org data를 저장하는 메모리가 4개 존재하는데, 8x8 Prediction과 4X4 Prediction, 16X16 Prediction 모듈을 병렬적으로 처리하기 위해 2개의 Luma data 메모리를 사용하고, 2개씩 짝을 이루어 듀얼 버퍼링을 수행한다. Chroma 메모리 역시 2개로 듀얼 버퍼링을 사용한다.

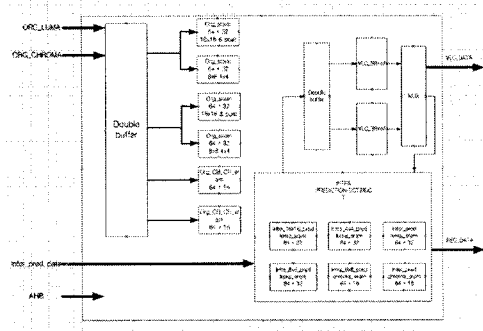


그림 2. 메모리 구조

2.2 Plane Prediction mode 제거

High Profile에서는 실제 Baseline 보다 선택될 수 있는 모드의 종류가 더 많아 졌다. 표 1을 보면 Baseline 은 14개의 모드를 지원하고, High profile는 24개의 모드를 지원 한다. 우리가 제안하는 구조(WSU)는 Plane Prediction mode를 제거한 23개의 모드를 지원하는 것이다. H/W cost가 높은 Plane 모드의 제거로 인한 화질저하는 Baseline 보다 그 영향이 적다. 하지만 H/W cost는 현저하게 떨어진다. Plane Prediction mode 제거로 인한 PSNR의 변화는 V장에서 설명 하겠다.

profile	Base	High	WSU
4x4 mode	9 개	9 개	9 개
8x8 mode	0 개	9 개	9 개
16x16 mode	4 개	4 개	3 개
inter_P(16x16 만 적용)	1 개	1 개	1 개
inter_B(16x16 만 적용)	0 개	1 개	1 개
총 mode 수	14 개 모드	24개 모드	23개 모드

표 1. profile mode

2.3 SAD Cost 선택

SAD Cost 선택은 H/W cost를 낮추고 더 빠르게 동작 시킬 수 있도록 하기 위해 선택 하였다. 실제 인코더를 하드웨어로 설계 하였을 경우 inter cost는 SAD로 계산 될 수밖에 없기 때문에 인트라 모듈에서 Hadamard 보다 SAD로 cost를 계산하는 것이 더 유리해지고 inter MB와 intra MB의 판별로 정확해 진다. 하지만 intra 안에서의 모드 결정은 Hadamard 보다는 떨어지게 된다. High Profile에서는 Baseline 보다 선택의 폭이 B슬라이스와 P슬라이스에 더 많이 집중되어 있으므로 화질저하는 Baseline에서 SAD를 선택 하는 것 보다 적어진다.

2.4 Q 와 IQ 모듈의 공유

제한한 High profile IP는 4x4 prediction 과 8x8 prediction을 동시에 수행하게 되는데 DCT 모듈은 서로 다른 모듈을 사용하지만 H/W cost를 줄이기 위해 8x8 과 4x4의 동작 cycle를 조절하여 1개의 Q, IQ 모듈을 사용 하였다.

2.5 luma DC 와 Chroma DC의 pre processing

post 과정에서 I16MB 일 경우 LUMA DC값을 처리 해야 하는데 DC값을 추출하기 위해서는 DCT를 수행 하여야 한다. chroma DC 역시 DCT 수행 후 DC 값을 얻을 수 있다. 그렇게 된다면 POST 과정에서 Cycle 소모가 커지게 될 것이다. 그래서 제안한 IP 구조에서는 16x16 prediction 하는 동안 DC값을 미리 추출하고 하다마드를 수행 하게 된다. chroma DC 역시 chroma prediction 수행 시 DC 값을 추출하여 하다마드를 수행해 놓는다. 그로인하여 POST 과정에서 DC값을 얻기 위한 Cycle 소모를 없앨 수 있다

2.6 8픽셀 병렬 처리

High profile의 고해상도 영상을 처리하기 위해서는 많은 cycle를 요구 하게 된다. 그래서 제안한 IP 구조에서는 빠르게 수행하기 위해 8픽셀

로 병렬 처리 하여 Intra Prediction을 수행한다. IP 내부의 데이터 이동은 64bit 인터페이스를 사용 하였다.

2.7 High profile CABAC 과 CAVLC 지원

Baseline에서는 CAVLC만 지원 하였지만 High profile 에서는 CABAC를 지원하므로 제안한 IP 구조역시 CABAC를 지원 하도록 설계 하였다. CAVLC와 CABAC의 큰 차이점은 8X8 prediction에서 zig-zag 스캔 부분이다. CABAC에서는 8X8 블록 단위의 zig-zag 스캔을 하고 CAVLC는 8X8 블록 에서 4x4 블록크기의 네 부분으로 나누어 zig-zag 스캔을 수행한다. 제안된 IP 모듈은 CAVLC와 CABAC에서 바로 사용할 수 있도록 coefficient data를 생성 한다.

III. 실험 결과

3.1 전체 cycle

그림 3은 제안된 IP의 전체 cycle이다. 그림에서 보는 바와 같이 크게 PRE processing 과정과 POST processing 과정으로 나누어진다. PRE processing 과정에서는 Intra prediction을 수행하는 구간이고, POST processing 과정은 DCT, Q, IQ, IDCT를 수행하는 구간이다. Pre 과정은 210 cycle 동안 수행되고, POST 과정은 98 cycle로 수행되어 총 308 cycle로 처리 가능 하다.

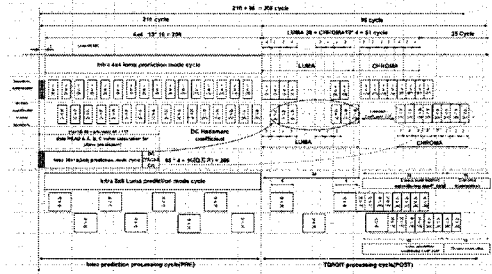


그림 3. IP 전체 cycle

3.2 PSNR

기존의 구조와 제안하는 구조의 비교 실험을 위한 QCIF 영상을 사용 하였으며 각 시퀀스 당 300프레임 IBBPBBP(IDR15)로 코딩 하였다. QP 값은 22,24,26,28,30,32,34를 사용하여 각 비트율에 대한 휘도성분과 색차 성분의 PSNR을 비교 하였으며 참조 소프트웨어는 JM 13.2를 사용하여 실험을 수행하였다.

표 2는 Foreman 영상의 제안된 구조인 Plane

mode 제거와 SAD 계산 방법을 적용 시킨 것과 SAD 계산 방법만 적용시킨 것 그리고 원본 영상을 QP 값에 따른 비트율과 PSNR을 나타내고, 그림 4에서는 이에 대한 R-D 곡선을 보여주고 있다. 그림 5번(parkrun HD영상)과 같이 다른 영상을 사용한 실험에서도 이와 비슷한 결과를 얻었다. 그림 5를 보면 QP가 낮은 구간은 참조 영상과 거의 동일한 PSNR을 갖고, QP가 30 이상으로 높아지면 제안된 구조의 IP가 0.5db 정도 떨어진다.

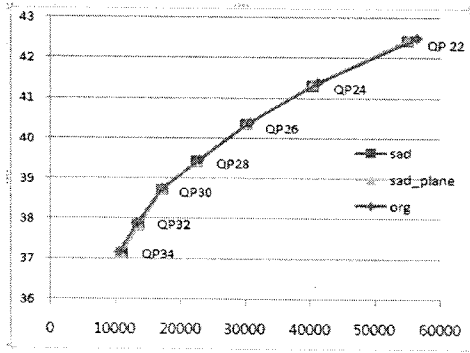


그림 4. foreman qcif R-D 곡선

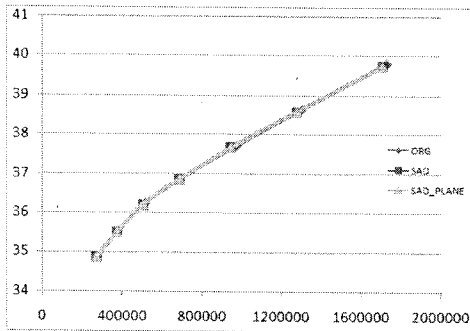


그림 5. parkrun 1280x720p R-D 곡선

qp	org		sad		sad plane	
	PSNR	bit rate	PSNR	bit rate	PSNR	bit rate
22	42.4937651	1691025	42.4411946	1650660	42.4466711	1651038
24	41.3973859	1234370	41.3331879	1211134	41.3465503	1211667
26	40.3958221	918289	40.352245	903708	40.3408289	903626
28	39.472047	683702	39.4286779	675544	39.4267215	674231
30	38.7652383	512173	38.7227919	515011	38.7130503	513562
32	37.8939195	395871	37.8540034	405349	37.8437282	405175
34	37.1679868	315521	37.1402718	327299	37.1205302	325547

표 2. foreman 영상의 PSNR과 비트율

3.3 Modelsim 시뮬레이션

그림 6은 Modelsim 시뮬레이션 결과 파형이다. 제안된 하드웨어 구조는 Verilog HDL를 이용하여 설계 되었으며 JM 13.2를 이용하여 검증 되었다.

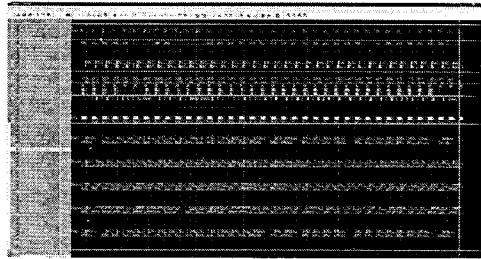


그림 6. simulation 결과 파형

IV. 결론

본 논문에서는 AMBA 기반으로 사용될 수 있는 H.264/AVC를 위한 효율적인 High Profile IP의 하드웨어 설계 구조를 제안하였다. 제안된 plane mode 제거 방법과 SAD 계산법 그리고 8 pixel 병렬처리 등을 통해 hardware cost 및 부호화 수행 시간을 줄였다. 제안된 구조를 적용하여 실험한 결과는 그림 4와 그림 5에서 보는 바와 같이 QP가 낮은 경우에 화질저하는 거의 없고, Qp가 30이상일 때 0.5db정도의 화질 저하를 보였다. 제안된 구조는 0.5db의 화질 저하 대신 낮은 H/W cost와 Full HD 영상을 충분히 인코딩 할 수 있는 빠른 cycle를 얻을 수 있었다. 제안된 구조는 25만 gate 와 Full HD 1080@30fps에서 133MHz clock에서 동작 시킬 수 있음을 확인하였다.

본 연구에 사용된 CAD Tool은 IDEC으로부터 지원 받았음.

참고문헌

- [1] Draft ITU-T Recommendation and Final Draft International Standard of joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), Sep. 2005.
- [2] C. Ku, C. Cheng, G. Yu M. Tasi, T. Chang "A High-Definition H.264/AVC Intra-Frame Codec IP for Digital Video and Still Camera Applications", IEEE, vol.16, pp. 920-922, Aug. 2006.