

스케일러블 비디오 코딩을 위한 Open-Loop 프레임 예측 프로세서의 FPGA 설계

정회원 서영호*

FPGA Design of Open-Loop Frame Prediction Processor for Scalable Video Coding

Young-Ho Seo* *Regular Member*

요 약

본 논문에서는 스케일러블 비디오 코딩을 위한 새로운 프레임 예측 필터링 기법과 하드웨어 구조를 제안하였다. MCTF와 hierarchical B-picture는 비디오 프레임간의 상관성을 제거하는 기술의 일종으로 본 논문에서 다루고자 하는 대상이다. 두 기술은 시간에 대해서 비인과성 시스템에 해당하므로 소프트웨어 및 하드웨어 구현 시에 프레임 버퍼링을 위한 대기지연시간이 매우 길고 대용량의 프레임 버퍼를 요구하는 단점이 있다. 이러한 비인과성 시스템을 인과성 시스템으로 재구성하여 효율적으로 구현할 수 있는 구조를 제안하고자 한다. 동일한 연산이 반복으로 수행되는 특성을 이용하여 단위 연산을 수행할 수 있는 프레임 예측 필터링 셀(FPFC : frame prediction filtering cell)을 제안하고 이를 확장하여 전체 연산구조를 재구성하였다. 먼저, 연산의 동작 순서를 분석하고 하드웨어의 구현을 고려한 인과성을 부여한 후 단위 프레임 처리를 위한 셀을 최적화하였다. 제안한 셀의 단순한 확장을 통해서 FPFC 커널을 구성하고, 이를 이용하여 스케일러블 비디오 코딩을 위한 FPFC 프로세서를 구현하였다.

Key Words : MCTF, SVC, FPGA, Lifting, MPEG

ABSTRACT

In this paper, we propose a new frame prediction filtering technique and a hardware(H/W) architecture for scalable video coding. We try to evaluate MCTF(motion compensated temporal filtering) and hierarchical B-picture which are a technique for eliminate correlation between video frames. Since the techniques correspond to non-causal system in time, these have fundamental defects which are long latency time and large size of frame buffer. We propose a new architecture to be efficiently implemented by reconfiguring non-causal system to causal system. We use the property of a repetitive arithmetic and propose a new frame prediction filtering cell(FPFC). By expanding FPFC we reconfigure the whole arithmetic architecture. After the operational sequence of arithmetic is analyzed in detail and the causality is imposed to implement in hardware, the unit cell is optimized. A new FPFC kernel was organized as simple as possible by repeatedly arranging the unit cells and a FPFC processor is realized for scalable video coding.

* 본 연구는 2006년도 한성대학교 교내연구비 지원과제임.

* 한성대학교 정보통신공학과 (www.hansung.ac.kr/design, yhseo@hansung.ac.kr)

논문번호 : KICS2006-03-119, 접수일자 : 2006년 3월 15일, 최종논문접수일자 : 2006년 5월 4일

I. 서론

스케일러블 비디오 코딩은 하나의 비디오 스트림으로 다양한 전송 네트워크와 수신 단말에 적응적 서비스가 가능하게 하기 위한 비디오 부호화 방법으로^{[1][2]}, 기존의 관련 국제 표준 및 활동으로는 MPEG-2 Scalable Profile^[3], MPEG-4 Scalable Profile/FGS^[4], MPEG-4 Part 10 SE(Scalable Extension)^[5] 등이 있다. MPEG-2 Scalable Profile은 기본계층과 향상계층의 구조를 가지면서 공간적, 시간적, 화질에 대한 scalability를 제공하지만 동시에 한 가지 scalability만 제공이 가능하다는 단점을 가지고 있다. MPEG-4 Scalable Profile은 FGS(Fine Granularity Scalability)를 이용하여 화질에 대한 scalability를 제공하고 있으며, MPEG-4 Part 10 SE는 이러한 scalability와 더불어 네트워크 상에서의 강건성, 기본계층 호환성, 저 복잡도 코덱, 단대단 지연, 화질 선택 및 코딩 효율을 고려한다. 현재 표준화 과정이 진행 중인 H.264/AVC AMD1 SE은 다양한 비트스트림을 생성하기 위해 동영상 압축 표준인 H.264/AVC^[6]를 기반으로 한 계층적 접근 방식을 사용함으로써 scalability를 제공한다. 이 중 시간적 scalability를 위해 MCTF와 Hierarchical B pictures를 사용하여 고정된 GOP 구조 내에서 영상들 간의 시간축으로의 재분배 과정을 수행하여 기본계층과 향상계층을 생성한다.

본 논문에서는 스케일러블 비디오 코딩을 위한 새로운 Open-Loop 방식의 프레임 예측 필터링 기법과 하드웨어 구조를 제안하였다.

II. 스케일러블 비디오 코딩을 위한 프레임 예

2.1 H.264/AVC SE

H.264/AVC SE의 공간적, 시간적, 화질에 대한 스케일러빌리티를 제공하기 위한 부호화기의 구조를 그림 1에 나타내었다. H.264/AVC SE의 부호화기는 H.264/AVC Codec을 기반으로 SVC를 접목시키기 위해 추가적인 기본계층과 향상계층의 계층적 접근 방식을 통하여 다양한 비트스트림을 생성 및 추출한다. 공간적 scalability를 위해 spatial decimation 및 interpolation, 시간적 scalability를 위해 MCTF 및 hierarchical B pictures, 화질 scalability를 위해 CGS (Coarse Grain Scalability) 및 FGS (Fine Granular Scalability)에 의해 수행되어진다. 입력 영상의 공간적 spatial decimation을 통한 공간

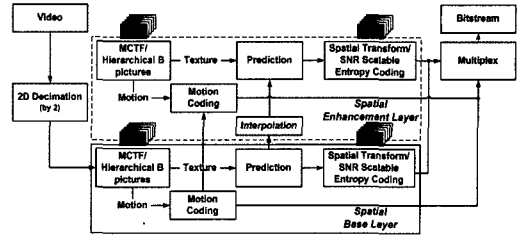


그림 1. H.264/AVC SE 부호화기의 구조
Fig. 1. Encoder architecture of H.264/AVC

적 기본계층과 향상계층들에 대해서 각각 독립적으로 화면간 예측 방법과 함께 MCTF에 의한 시간축으로의 영상 분해 과정 및 화면 내 예측을 수행한다.

낮은 공간적 계층으로부터의 움직임 예측 정보는 높은 공간적 계층의 움직임 예측을 위해 사용되어진다. 공간적 계층들 사이에 존재하는 텍스처 정보들의 중복성을 제거하기 위해 기본 계층으로부터 향상 계층으로의 공간적인 보간을 통해 텍스처 정보들의 차이 값을 구한 후, 압축고정을 거친다. 디코더는 수신단의 전송 통신망과 단말의 상태에 따라 공간적, 시간적, 화질의 기본계층과 향상계층들의 적절한 조합을 통하여 다양한 비트스트림을 생성 및 추출함으로써 scalability를 제공하게 된다.

2.2 MCTF

MCTF는 J. R. Ohm에 의해 처음으로 제안된 알 고리즘으로^[7], 영상들을 시간축으로의 중복성을 제거하기 위한 기술로 그림 2와 같은 피라미드형 분해 구조를 갖는다.

그림 2와 같이, H.264/AVC SE는 고정된 GOP 단위로 Haar와 5/3 필터의 리프팅 기반 MCTF를 적용하고 있으며, 각 레벨(Level)에 대해 예측 및 갱신 과정을 수행하여 각각 고주파 영상과 저주파 영상을 생성한다. 예측 과정은 현재 영상을 기준으로 이전과 이후의 영상을 참조하여 움직임 ME/MC

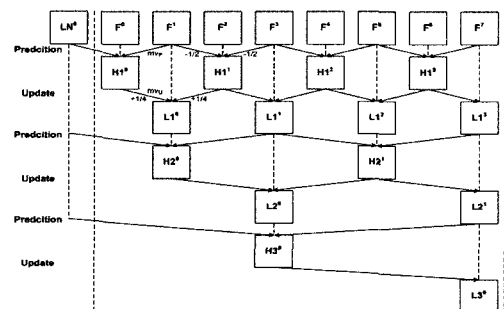


그림 2. H.264/AVC SE의 MCTF 구조
Fig. 2. MCTF structure of H.264/AVC SE

과정을 통해 차영상 즉 고주파 영상을 만들어낸다. 갱신 과정은 이전의 예측 과정에서 구한 움직임 벡터를 이용하여 고주파 영상을 참조하여 움직임 보상을 통해 원 영상의 고주파 성분이 제거된 저주파 영상을 생성한다. 이와 같은 예측 및 갱신 과정은 각 레벨에 대해 수행되어지며 GOP=2N일 경우, N 레벨까지 수행하고 마지막 레벨의 저주파 영상과 각 레벨의 고주파 영상이 부호화기의 다음 단계로 이동한다. 최종적인 저주파 영상은 다음 GOP의 MCTF를 수행할 시 예측 과정을 위해 사용되며, 복호화 시 이러한 영상들의 선택적인 추출로부터 시간적 scalability가 제공된다.

2.3 Hierarchical B pictures

Hierarchical B pictures는 MCTF와 유사하지만 갱신 과정이 없이 예측 과정을 수용하면서 시간적 스케일러빌리티를 제공한다. 이 알고리즘을 사용한 H.264/AVC SE 부호화기의 비트스트림은 기존의 H.264 /AVC Main profile 부호화기와 호환 가능하며, 그림 3에 알고리즘 수행 방법을 보이고 있다. 첫 번째 영상은 IDR(Instantaneous Decoding Refresh) picture로 부호화되며, 이후의 영상들은 GOP 내에서 "B...BP" 혹은 B...BI의 비디오 구조로 부호화된다. 첫 번째 레벨의 B1 영상은 주변 I 영상으로부터 예측되어지며, 다음 레벨의 Bi 영상들은 이전 레벨의 I와 Bj (j<i)로부터 양방향 예측되어진다. 이들 영상은 RPLR(Reference Picture List Reordering)과 MMCO(Memory Management Control Operation) 제어 명령어와 함께 부호화되어 전송된다. 이와 같이 Hierarchical B pictures를 수행한 후 마지막 레벨의 결과인 I, B, P 영상들은 MCTF와 마찬가지로 부호화기의 다음 단계로 이동한다.

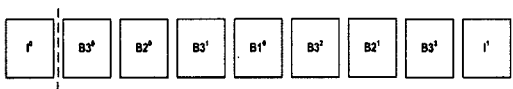


그림 3. H.264/AVC SE의 Hierarchical B pictures 구조
Fig. 3. Hierarchical B pictures structure of H.264/AVC SE

III. 제안한 프레임 예측 필터링의 구조

3.1 제안한 프레임 예측 필터링 방식

프레임 예측 필터링 방식 중의 하나인 MCTF 연산이 수행되는 절차를 도식적으로 나타내면 그림 3과 같다. MCTF는 필터링 연산에 사용되는 계수만 다르고 방식은 동일한 총 4단계의 곱셈과 덧셈과정

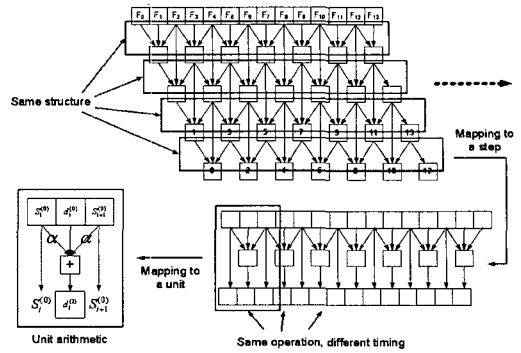


그림 4. MCTF의 구조사상을 통한 단위연산
Fig. 4. Unit arithmetic of 1D lifting by structure mapping

을 거치는데, 이를 식 (1)~(6)에 나타냈다. 식 (4)의 결과는 식 (5)와 (6)의 스케일링 과정을 거친다. 여기서 $x_i^{(j)}$ 는 그림 4에서 하나의 입력 프레임에 해당한다.

$$x_{2i+1}^{(1)} = x_{2i+1}^{(0)} + \alpha(x_{2i}^{(0)} + x_{2i+2}^{(0)}) \quad (1)$$

$$x_{2i}^{(1)} = x_{2i}^{(0)} + \beta(x_{2i+1}^{(1)} + x_{2i-1}^{(1)}) \quad (2)$$

$$x_{2i+1}^{(2)} = x_{2i+1}^{(1)} + \gamma(x_{2i}^{(1)} + x_{2i+2}^{(1)}) \quad (3)$$

$$x_{2i}^{(2)} = x_{2i}^{(1)} + \delta(x_{2i+1}^{(2)} + x_{2i-1}^{(2)}) \quad (4)$$

$$x_{2i} = \zeta x_{2i}^{(2)} \quad (5)$$

$$x_{2i+1} = x_{2i+1}^{(2)} / \zeta \quad (6)$$

MCTF는 동일한 연산구조를 가지므로 그림 4의 오른쪽 아래와 같은 하나의 단계로 대체시킬 수 있다. 마지막으로 하나의 단계로 사상된 MCTF 연산은 동일한 구조를 가지고 동일한 연산을 수행하면서 시간적으로 중복되지 않으므로 그림 4의 왼쪽 아래쪽과 같은 하나의 단위 연산으로 사상될 수 있다. 이 경우 연속적으로 입력되는 데이터를 처리할 수 있어야 하는 기본사항을 만족시켜야 하는데, 그림 5에서 단위 연산을 시간에 따라 겹치지 않으면서 입력되는 데이터에 대해 MCTF 연산을 수행할 수 있도록 재구성하는 과정을 그림 5에 나타냈다. 그림 5에 나타낸 MCTF 연산과정을 그림 6에 자세히 나타내었다. 그림 6의 연산과정은 H/W 구현을 대상으로 하고 있고, 시간적으로 연속된 비디오 신호를 처리하는 프레임 단위의 실시간 MCTF 연산 방식을 제안하고 있다.

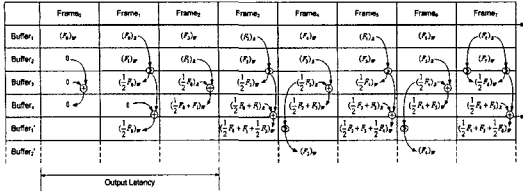


그림 5. 제안한 MCTF의 재구성
Fig. 5. The proposed reconfiguration of MCTF

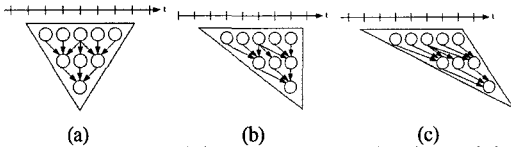


그림 6. MCTF의 프레임 재구성 (a) 원래 시스템 (b) 인과시스템 (c) 비디오 처리를 위한 제안한 구조
Fig. 6. Frame re-configuration of MCTF (a) original system (b) causal system (c) the proposed system for processing video

3.2 Hierarchical B-picture로의 확장

또다른 프레임 예측 필터링 방식인 Hierarchical B-picture의 연산은 MCTF와 동일한 구조를 가지고 수행된다. 하지만 MCTF와는 달리 갱신과정이 없는 것이 특징이다. 또한 MCTF에서는 고속 웨이블릿 연산인 리프팅을 사용하기 때문에 리프팅 계수를 각 프레임의 예측에 적용하지만 hierarchical B-picture에서는 사용하지 않는다. 따라서 MCTF 연산을 위한 알고리즘에서 갱신과정은 생략하고 필터링 연산에 사용되는 계수를 1로 선택할 수 있으면 동일한 알고리즘으로 hierarchical B-picture까지 수행할 수 있다.

3.3 제안한 FPFC의 H/W 구조

MCTF와 hierarchical B-picture는 비디오 프레임 간의 상관성을 제거하는 기술의 일종으로 시간적인 순서를 고려할 때 다수의 비디오 프레임에 대한 비인과성 시스템에 해당하므로 소프트웨어 및 하드웨어 구현 시에 프레임 버퍼링을 위한 대기지연시간이 매우 길고 대용량의 프레임 버퍼를 요구하는 단점이 있다. 따라서 이러한 비인과성 시스템을 인과성 시스템으로 재구성하고 구현이 용이한 구조가 필요하다.

동일한 구조의 반복적인 연산을 통해서 수행되는 Open-Loop 방식의 프레임 예측 필터링의 특성을 이용하여 단위 연산을 수행할 수 있는 프레임 예측 필터링 셀을 제안하고 이를 확장하여 전체 프레임 예측 필터링을 위한 프레임 버퍼링 구조를 재구성하였다. 먼저, 앞 절에서 설명한 것과 같이 연산의 동작 순서를 분석하고 하드웨어의 구현을 고려한

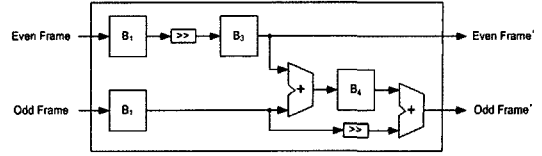


그림 7. FPFC의 구조
Fig. 7. FPFC architecture

인과성을 부여한 후 단위 프레임 처리를 위한 셀을 최적화하였다. 제안한 셀의 단순한 확장을 통해서 리프팅 커널을 구성하고, 이를 이용하여 그림 7과 같이 스케일러블 비디오 코딩을 위한 프레임 예측 필터링 셀(FPFC : frame prediction filtering cell)의 구조를 제안하였다.

IV. 하드웨어 구현결과

제안한 FPFC의 연산방식과 cell의 구조를 이용하여 H/W를 구현하였다. 구현한 하드웨어의 구조를 특징에 따라서 그림 7, 8, 그리고 9에 나타내었다. 그리고 그림 10에는 MCTF cell 기반의 MCTF Kernel로 구성된 MCTF processor의 구조를 나타내고 있다. 제안된 H/W 구조는 셀 단위로 독립적인 동작이 가능하여 단순한 셀의 추가만으로 동작 및 성능의 확장이 가능하다. 본 논문에서 내부 메모리 혹은 버퍼는 SRAM을 가정하였고, 외부 메모리의 경우에는 고속의 SDRAM을 대상으로 하였다.

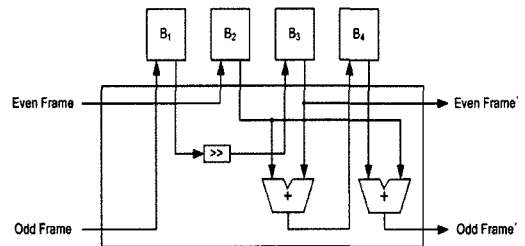


그림 8. 내부 메모리(버퍼)를 이용한 FPFC의 구조
Fig. 8. H/W architecture of FPFC using on-chip memory

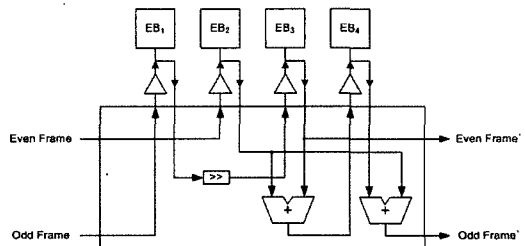


그림 9. 외부 메모리(SDRAM)를 이용한 FPFC의 H/W구조
Fig. 9. H/W architecture of FPFC using off-chip memory (SDRAM)

제안한 프레임 예측 필터링 연산방식과 FPFC 기반의 FPFC Kernel로 구성된 FPFC processor의 구조를 그림 10에 나타냈다. 제안된 H/W 구조는 셀 단위로 독립적인 동작이 가능하여 단순한 셀의 추가만으로 동작 및 성능의 확장이 가능하다. 본 논문에서 내부 메모리 혹은 버퍼는 SRAM을 가정하였고, 외부 메모리의 경우에 고속의 SDRAM을 대상으로 하였다.

구현된 H/W는 Altera사의 Stratix FPGA에서 약 13만개의 LE와 60개의 ESB를 점유하여 구현되었고, 약 100MHz의 클럭에서 동작이 가능하여 일반적인 비디오 처리 시스템과의 동작이 충분히 가능함을 보였다.

그림 12, 13, 그리고 14에 핵심 블록인 MCTF Cell, SDRAM 제어기, 그리고 ME에 대한 Synplify TM를 이용한 FPGA 합성 결과를 보였다. 전체 FPFC 프로세서의 FPGA 합성결과는 시각적으로 확인하기 어렵기 때문에 본 논문에서 생략하였다. 아

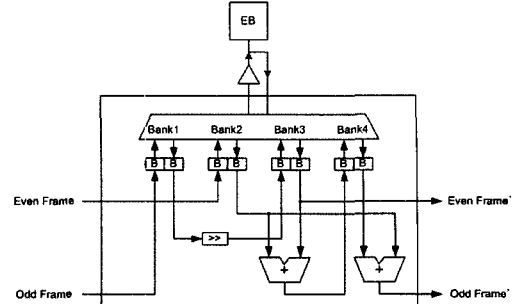


그림 10. 외부메모리(SDRAM)를 이용한 FPFC의 H/W구조2
Fig. 10. H/W architecture2 of FPFC using off-chip memory (SDRAM)

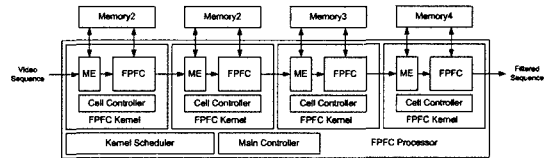


그림 11. FPFC 프로세서의 H/W 구조
Fig. 11. H/W architecture

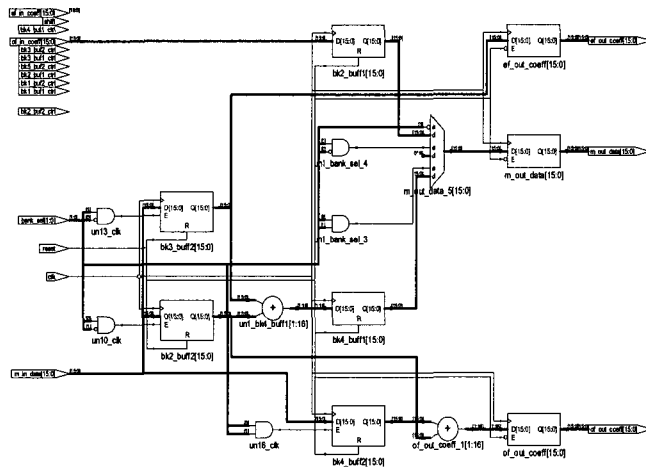


그림 12. FPFC의 FPGA 구현결과
Fig. 12. FPGA synthesis result of FPFC

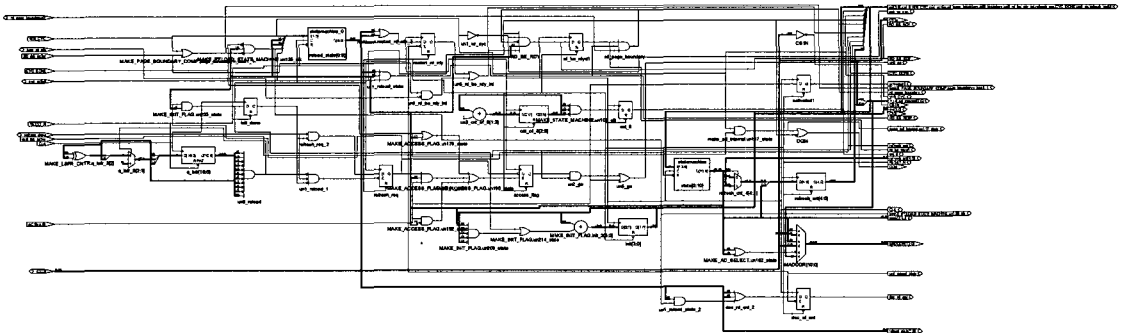


그림 13. SDRAM 제어기의 FPGA 구현결과
Fig. 13. FPGA synthesis result of SDRAM controller

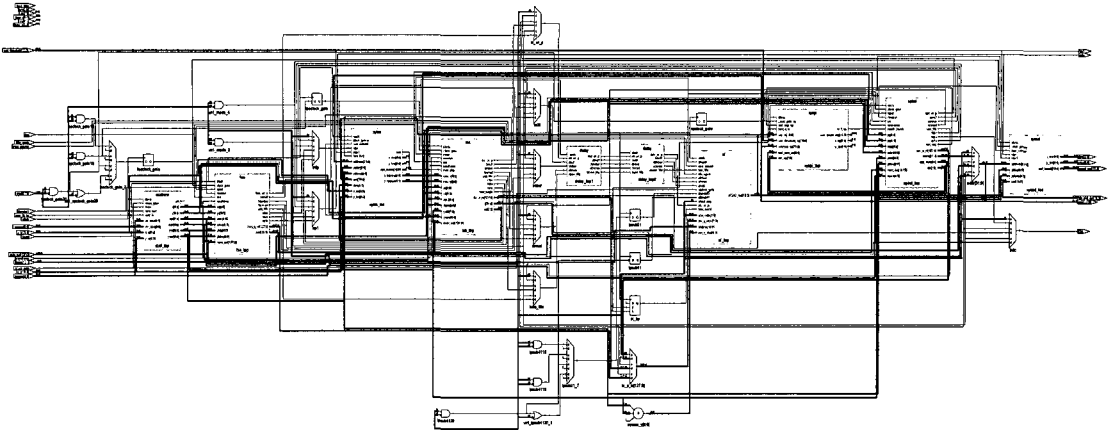


그림 14. ME의 FPGA 구현결과
Fig. 14. FPGA synthesis result of ME

직까지 MCTF와 hierarchical B-picture를 H/W로 구현한 구체적인 연구사례가 없기 때문에 연구 결과를 직접으로 비교하기는 어렵다.

V. 결론

본 논문에서는 스케일러블 비디오 코딩을 위한 새로운 시간적인 프레임 예측을 위한 필터링 연산 방식과 그에 따른 H/W 구조를 제안하였다. 프레임 단위 리프팅 연산의 특성을 분석하여 실시간으로 처리가 가능한 연산 구조를 보였다. 제안된 구조는 H/W의 구현이 용이하고 최적화된 메모리 사용율을 가져 최소화된 비용이 소모되므로 앞으로 MPEG-2, MPEG-4, H.264/AVC 그리고 JPEG2000과의 통합을 통해 효율적인 스케일러블 비디오 코딩을 구현할 수 있는 솔루션이 될 수 있을 것으로 사료된다.

참고 문헌

[1] J. Bormans, J. Gelissen, and A. Perkis, "MPEG-21: The 21st Century Multimedia Framework", *IEEE Trans. on Signal Processing Magazine*, vol. 20, pp. 53-62, 2003.
 [2] ISO/IEC TR 21000-1:2001(E), "Part 1: Vision, Technologies and Strategy", 2000.
 [3] J. Reichel, H. Schwarz, and M. Wien, "Working Draft 3 of ISO/IEC 14496-10: 2005/AMD1 Scalable Video Coding", *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6 Document N7310*, 2005.

[4] ISO/IEC 13818-2, "Generic Coding of moving pictures and associated audio information-Part 2: Video", also *ITU-T Recommendation H.262*, 1995.
 [5] ISO/IEC 14496-2, "Coding of audio-visual objects-Part 2: Visual", 1998.
 [6] ISO/IEC 14496-10 "Coding of Audio-visual Objects-Part 10: Advance Video Coding", 2003, also *ITU-T Recommendation H.264 "Advanced Video Coding for Generic Audiovisual Services"*, 2003.
 [7] J. R. Ohm, "Three Dimensional Subband Coding with Motion Compensation", *IEEE Trans. on Image Processing*, vol. 3, pp. 559-571, 1994.

서영호 (Young-Ho Seo)

정회원



1999년 2월 광운대학교 전자재료공학과 졸업(공학사)
 2001년 2월 광운대학교 일반대학원 졸업(공학석사)
 2000년 3월~2001년 12월 인티스닷컴(주) 연구원
 2004년 8월 광운대학교 일반대학원 졸업(공학박사)

2003년 6월~2004년 6월 한국전기연구원 연구원
 2004년 12월~2005년 8월 유한대학 연구교수
 2005년 9월~현재 한성대학교 전임강사
 <관심분야> 2D/3D 영상 및 비디오 처리, 디지털 홀로그램, SoC 설계, 워터마킹/암호화