

H.264/AVC SVC에서 DCT 기반의 계층 간 잔여 신호 예측 구조

강진미^o 김성민 정기동

부산대학교 컴퓨터공학과

{wolff98^o, morethannow, kdchung}@pusan.ac.kr

DCT-domain Inter-layer Residual Prediction Architecture in H.264/AVC SVC

Jinmi Kang^o Sungmin Kim Kidong Chung

Department of Computer Science and Engineering, Pusan National University

요 약

H.264/AVC SVC에서는 공간적 확장성을 지원하기 위하여 계층 간 예측 방법을 새롭게 도입하였다. 계층 간 예측은 하위 계층의 움직임 정보, 텍스처 정보, 잔여 신호 정보를 이용하여 계층 간 중복성을 제거하는 방법이다. 따라서 상위 계층의 부호화 효율을 높이는 반면, 복호화 과정에서는 하위 계층의 잔여 신호 정보를 픽셀 단위까지 복원하여 계산 복잡도가 높아지는 문제점이 있다. 본 논문에서는 이러한 H.264/AVC SVC에서 복호화 과정의 계산 복잡도를 줄이기 위하여 DCT 기반의 잔여 신호 예측 구조를 제안하였다. H.264/AVC SVC에서 픽셀 기반의 잔여 신호 예측 구조와 제안하는 구조의 연산 수를 계산하여 계산 복잡도를 비교한 결과 약 33%의 개선이 이루어졌다.

1. 서 론

SVC(Scalable Video Coding)는 한 번의 동영상 압축으로 서로 다른 공간적, 시간적, 화질의 확장성을 갖도록 단계적으로 데이터를 부호화하여, 다양한 네트워크와 단말 환경에 적합한 공간적 해상도, 시간적 프레임율, 프레임의 화질을 갖는 동영상으로 복호화할 수 있는 동영상 압축 방식이다[1,2]. 최근에는 사용자의 미디어 사용 환경이 유무선 네트워크와 단말기 기술의 발달로 매우 다양해짐에 따라 위와 같이 수신자 환경에 적응적인 특징을 가진 SVC의 필요성이 대두되고 있다. ISO/IEC와 ITU-T 공동으로 구성된 JVT (Joint Video Team)는 최신 동영상 압축 표준인 H.264/AVC의 개정안으로 SVC를 표준화하고 있다[3,4].

H.264/AVC SVC는 서로 다른 공간과 화질의 확장성 (scalability)을 용이하도록 하기 위하여 기존의 부호화 방식을 기반으로 데이터를 여러 계층으로 나누어 부호화하는 방식을 추가하였다. 특히 그림 1과 같이 공간적 확장성을 위한 SVC는 기본 계층(base layer)과 여러 개의 상위 계층(enhancement layer)으로 구분된다. 또한 상위 계층의 압축 효율을 향상시키기 위하여 계층 간 예측 방법(inter-layer prediction)을 이용한다[2].

계층 간 예측 방법은 기본 계층의 데이터를 기존의 H.264/AVC 표준과 동일한 방법으로 부호화하고 부호화된 기본 계층의 텍스처 정보, 움직임 정보, 잔여 신호 정보 등을 이용하여 상위 계층의 부호화 시 계층 간 중복되는 정보를 제거하여 압축 효율을 향상시킨다. 따라서 계층 간 예측 방법은 프레임 내 압축 시 계층 간 중복성을 제거하는 텍스처 예측, 기본 계층의 움직임 정보를 이용하는 움직임 예측, 연관된 기본 계층의 잔여 신호의 유사성을 이용하는 잔여 신호 예측으로 총 세 가지 방법으로 분류된다[5].

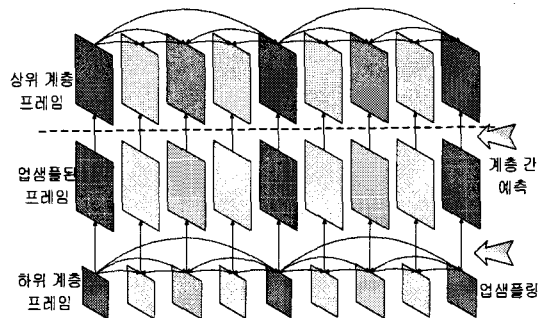


그림 1 SVC의 공간적 확장성

^o이 논문은 2단계 두뇌한국21사업에 의하여 지원되었음

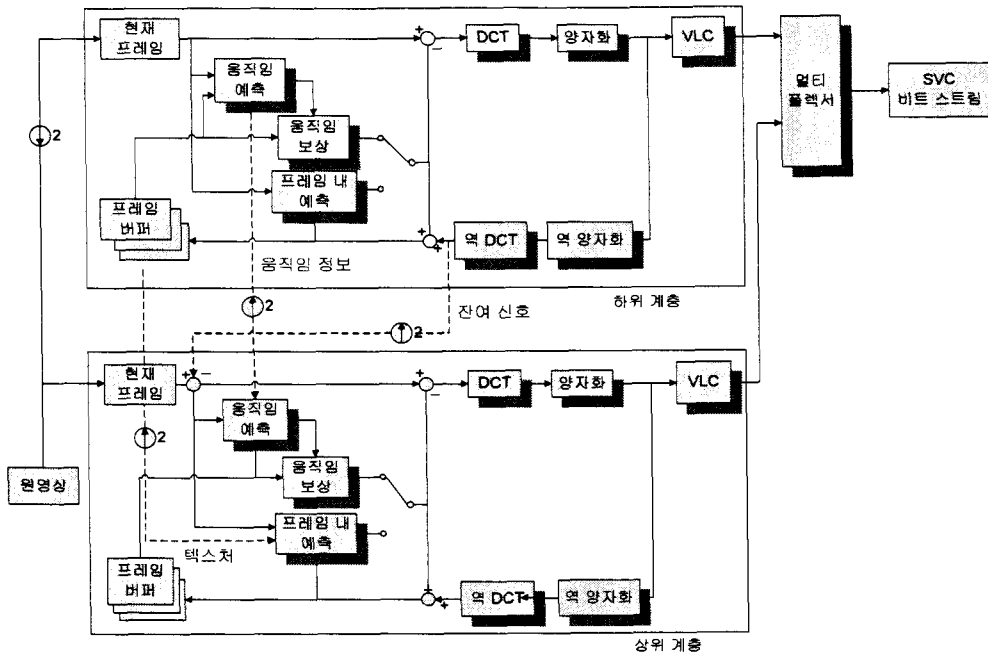


그림 2 공간적 확장성 구조

SVC로 부호화된 데이터는 네트워크와 단말 환경에 적합한 계층만 전송되어 수신자 환경에서 복호화된다. 상위 계층을 복호화하는 최종 과정은 계층 간 예측 방법에 이용된 정보를 추출하기 위해서 하위 계층의 데이터의 복호화가 필요하기 때문에 계산 복잡도가 증가하게 되는 문제점이 있다. H.264/AVC SVC는 프레임 간 예측인 경우 픽셀 단위까지 복호화하지 않고 계층 간 예측에 이용되는 잔여 신호까지만 복호화하여 약간의 계산 복잡도를 감소시켰지만 실시간, 모바일 등의 다양한 수신자 환경에 적용되기 위해서는 복호화 과정의 계산 복잡도의 최소화가 필수적이다[1].

본 논문에서는 공간적 확장성을 위한 SVC의 계층 간 예측 방법을 DCT 기반에서 적용할 수 있는 구조를 제안한다. 픽셀 단위에서 수행되던 계층 간 예측 방법을 DCT 단위에서 적용하여 복호화 과정에서의 계산 복잡도를 효율적으로 감소시킨다. 또한, 반복적인 DCT 변환과 역 DCT 과정을 제거함으로써 DCT 변환으로 인한 화질의 열화를 방지하는 효과도 예상된다.

본 논문의 구성은 다음과 같다. 2장에서는 SVC에서의 공간적 확장성 구조와 계층 간 예측 방법 관련 연구를 살펴보고 3장에서는 본 논문에서 제안하는 DCT 기반의 계층 간 잔여신호 예측 구조를 알아본다. 4장에서는 성능을 평가하고 5장에서는 결론과 향후 연구에 대해 밝히도록 한다.

2. 관련 연구

본 장에서는 H.264/AVC SVC의 공간 확장성을 지원하기 위한 계층 간 예측 방법을 알아보고 SVC와 동일하게 적용되는 H.264/AVC 표준에서 수행되는 DCT 변환 과정을 살펴본다.

2.1 H.264/AVC의 공간적 확장성을 위한 SVC 구조

H.264/AVC SVC는 공간적 확장성을 지원하기 위하여 원 영상을 여러 개의 계층으로 나누어 부호화하는 계층적 부호화 방법을 이용한다. 그림 2와 같이 최하위 계층인 기본 계층은 기존의 프레임 내 예측(intra prediction), 프레임 간 예측(inter prediction)을 수행하기 때문에 H.264/AVC 표준과 호환이 가능하다. 상위 계층에서는 H.264/AVC 구조에 계층 간의 중복성을 제거하기 위하여 계층 간 예측을 수행하는 과정이 추가된다 [6].

계층 간 예측 방법은 상위 계층의 압축 효율을 높이기 위하여 기본 계층을 포함한 하위 계층의 텍스처 정보, 움직임 정보, 잔여 신호 정보를 이용하는 각각 세가지 방법으로 구성된다.

2.1.1 텍스처 정보를 이용한 계층 간 예측

현재 부호화하는 상위 계층의 매크로블록에 대응하는 하위 계층의 블록이 프레임 내 예측이 수행된 경우 적용되는 방법으로 하위 계층의 텍스처 정보를 계층 간 예측에 이용한다. 하위 계층의 공간적 해상도가 현재

계층보다 낮기 때문에 참조 대상이 되는 하위 계층의 블록은 픽셀 단위까지 복호화되어 현재 계층의 공간 해상도와 동일 한 크기로 업샘플된다. 현재 부호화하는 매크로블록은 업샘플한 하위 계층의 텍스처 정보와의 중복성을 제거하는 과정으로 예측이 이루어진다.

2.1.2 움직임 정보를 이용한 계층 간 예측

하위 계층의 대응되는 블록이 프레임 간 예측이 수행된 경우 상위 계층에서는 하위 계층의 움직임 정보를 이용하여 계층 간 예측을 수행한다. 하위 계층의 움직임 정보를 업샘플하여 계층 간 예측 방법에 이용하고, 보다 정확도를 높이기 위하여 업샘플한 움직임 벡터에서 1/4화소 단위로 움직임 예측을 수행하는 1/4화소 정밀 모드도 지원한다. 하위 계층의 움직임 정보가 없거나 현재 계층 내의 프레임 내 예측 효율이 더 높은 경우에는 선택적으로 H.264/AVC 표준의 프레임 내 예측을 수행하여 부호화한다.

2.1.3 잔여 신호 정보를 이용한 계층 간 예측

하위 계층이 프레임 간 예측으로 부호화되거나 텍스처 정보를 이용하지 않고 H.264/AVC 표준의 프레임 내 예측을 수행하여 잔여 신호를 포함한 경우 적용되는 예측 방법이다. 상위 계층의 잔여 신호를 부호화 하는 과정에서 하위 계층의 잔여 신호 정보를 이용하여 계층 간 예측을 수행함으로써 압축 효율을 높인다. 현재 부호화하는 매크로블록과 대응하는 하위 계층 블록의 움직임 정보가 유사한 경우 잔여 신호를 업샘플하여 계층 간 예측으로 중복성을 제거한다. 상·하위 계층간의 움직임 정보의 차이가 크거나 잔여 신호의 중복성이 거의 없는 경우에는 계층 간 예측의 효율이 없기 때문에 움직임 정보의 차이에 따라 적응적으로 수행한다.

2.2 H.264/AVC에서의 DCT 변환

H.264/AVC에서는 공간적 중복성을 제거하기 위하여 4x4 블록 단위로 DCT 변환을 수행하며 DCT 종류는 기존의 4x4 DCT 변환을 변형시킨 형태인 4x4 정수형 DCT 변환을 이용한다[7].

식 1은 4x4 정수형 DCT 변환을 행렬 연산 형태로 나타낸 것이다. 4x4 정수형 DCT 변환은 DCT 변환의 부호화 후 다시 복호화되었을 때 손실이 없으므로 DCT 변환 과정으로 인한 오류를 발생시키지 않는다. 또한 계산을 빠르게 하기 위하여 나눗셈을 양자화 과정으로 이동하여 수행하는 특징이 있으며 나눗셈을 제외한 DCT 변환을 Core 변환이라고도 한다.

$$Y = CXC^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} [X] \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \quad \text{식 1}$$

DCT 변환이 행렬 연산으로 이루어지기 때문에 행렬 연산을 이용하여 해상도 조절, 트랜스코딩 등을 수행하는 DCT 기반의 구조에 관련된 연구가 활발히 진행되어 왔다[8]. DCT 기반 구조는 픽셀 기반 구조에 비해 계산 복잡도를 줄이면서 DCT 변환 후 역 DCT 변환 과정에서 발생하는 오류를 제거하여 화질 개선의 효과도 기대할 수 있다.

3. 제안하는 DCT 기반 계층 간 잔여 신호 예측 구조

본 논문에서는 H.264/AVC SVC의 공간적 확장성을 위한 계층 간 예측 구조 중 DCT 기반의 잔여 신호 예측 구조를 제안한다. 앞 장에서 살펴본 H.264/AVC 표준의 DCT 변환에서의 행렬 연산 특징을 이용하여 계산 복잡도를 감소시킨다.

3.1 픽셀 기반의 계층 간 잔여 신호 예측 구조

공간적 확장성을 위하여 추가된 SVC 구조 중 잔여 신호 예측은 현재 H.264/AVC에서는 픽셀 기반에서 수행된다. 그림 3은 픽셀 기반의 잔여 신호 예측의 부호화 과정과 상위 계층의 복호화 과정을 나타낸다. 계층 간 잔여 신호 예측의 부호화 과정은 하위 계층의 잔여 신호의 픽셀 값을 업샘플하여 상위 계층의 잔여 신호의 픽셀 값과 차분하여 계층 간 예측을 수행한다. 계층 간 예측 후의 픽셀 값을 DCT 변환하는 과정으로 부호화가 진행된다. 상위 계층에서는 이러한 계층 간 예측 과정이 추가됨으로써 하위 계층과의 중복성을 제거하기 때문에 부호화 과정의 압축 효율을 높일 수 있다.

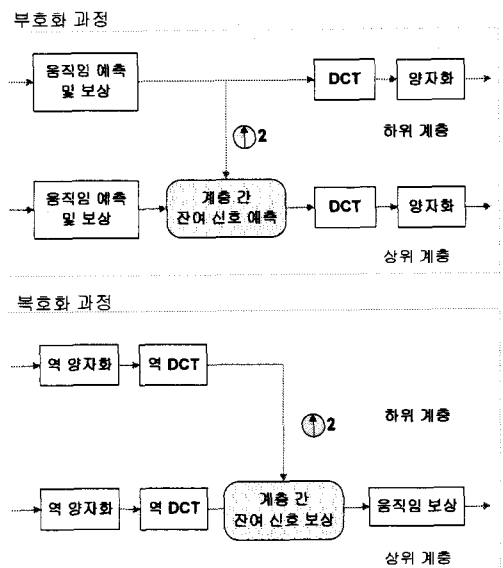


그림 3 픽셀 기반의 계층 간 잔여 신호 예측 구조

픽셀 기반의 계층 간 잔여 신호 예측 구조에서는 상위 계층의 복호화 과정에서 계산 복잡도가 높아지는 문제점을 가진다. 상위 계층의 잔여 신호를 복호화하기 위해서는 부호화 과정의 계층 간 예측에 이용된 하위 계층의 잔여 신호를 참조하는 과정이 수행되기 때문이다. 따라서 하위 계층의 잔여 신호를 픽셀 단위까지 복호화해야 하기 때문에 픽셀 기반 구조에서는 계산 복잡도가 높아지게 된다.

그러므로 본 논문에서는 DCT 변환 행렬의 특징을 이용하여 잔여 신호 정보의 변화없이 복호화 과정의 계산 복잡도를 감소시키는 DCT 기반의 계층 간 잔여 신호 예측 구조를 제안한다.

3.2 DCT 기반의 계층 간 잔여 신호 예측 구조

상위 계층의 압축 효율을 위하여 하위 계층의 잔여 신호 정보를 이용하여 계층 간 예측을 수행한다. 앞서 살펴본 바와 같이 픽셀 기반에서 이루어지는 경우 상위 계층과 하위 계층 간의 잔여 신호의 픽셀 값의 차분을 구하고 부호화 과정에 따라 DCT 변환 과정을 거치게 된다. 본 논문에서는 각 계층에서 잔여 신호를 픽셀 단위로 차분하지 않고 DCT 변환을 수행한 후 변환된 잔여 신호 정보를 차분하는 DCT 기반의 계층 간 잔여 신호 예측 구조를 제안한다.

그림 4는 제안하는 DCT 기반의 계층 간 잔여 신호 예측 구조의 부호화 과정과 복호화 과정을 나타낸다.

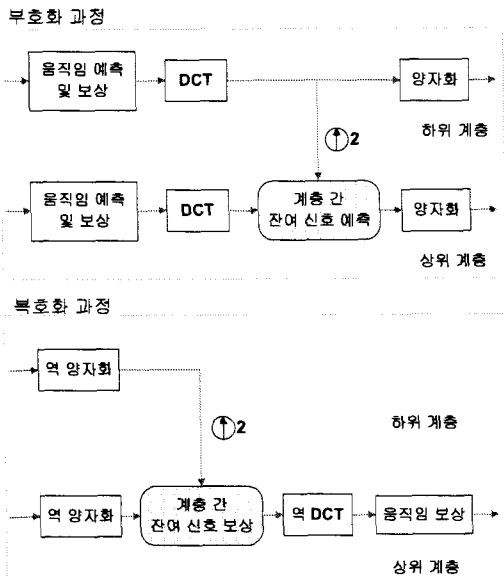


그림 4 DCT 기반의 계층 간 잔여 신호 예측 구조

우선, 부호화 과정을 살펴보면 하위 계층에서 잔여 신호의 정보를 DCT 변환 과정을 수행한 후 변환된 잔여 신호를 상위 계층의 해상도와 동일한 크기로 DCT

기반 압샘플 방법을 적용한다. 상위 계층에서도 프레임 내 예측 및 프레임 간 예측 과정을 수행한 잔여 신호를 계층 간 예측을 수행하지 않고 DCT 변환을 적용한다. 각각 DCT 변환된 잔여 신호를 이용하여 계층 간 예측을 수행한다. 픽셀 기반의 구조와 비교하면 잔여 신호를 이용한 계층 간 예측의 순서와 각 계층의 DCT 변환 과정의 순서가 변경되었음을 알 수 있다.

각 계층에서 DCT 변환 과정 후 변환된 잔여 신호를 차분하였기 때문에 제안하는 DCT 기반 구조의 복호화 과정에서는 그림 4와 같이 역 DCT 변환 과정의 이전 단계에서 계층 간 잔여 신호 보상 과정이 수행된다. 따라서 상위 계층의 복호화 과정에서 하위 계층의 잔여 신호를 픽셀 단위까지 복원하기 위하여 수행되었던 역 DCT 변환 과정을 제거하여 복호화 과정에서의 계산 복잡도를 감소시킬 수 있다.

4. 성능 평가

본 장에서는 제안한 DCT 기반의 계층 간 잔여 신호 예측 구조의 계산 복잡도 성능을 평가한다. H.264/AVC SVC 구조인 픽셀 기반의 계층 간 잔여 신호 예측 구조를 비교 대상으로 하고 계산 복잡도는 수행되는 연산의 수를 계산하여 비교한다.

4.1 픽셀 기반 구조의 계산 복잡도

픽셀 기반 구조에서는 상위 계층의 부호화에서 하위 계층의 픽셀 단위의 잔여 신호를 차분하고 DCT 변환을 수행한다. 이 때 수행되는 연산은 다음과 같다.

$$R = R_E - R_B \quad \text{식 2}$$

$$\tilde{R} = C \cdot R \cdot C^T \quad \text{식 3}$$

$$\tilde{R}_B = C \cdot R_B \cdot C^T \quad \text{식 4}$$

R_B 는 하위 계층의 잔여 신호, R_E 는 상위 계층의 잔여 신호를 나타내고 픽셀 기반의 계층 간 예측은 식 2와 같이 표현된다. 식 3과 식 4는 상위 계층과 하위 계층에서 수행되는 DCT 변환 연산을 각각 나타낸다.

복호화 과정은 식 5와 같이 하위 계층에서 역 DCT 변환 과정으로 잔여 신호를 복원하고 식 6과 같이 상위 계층에서 역 DCT 변환 과정을 수행하여 잔여 신호를 각각 복원한다. 최종으로 상위 계층의 잔여 신호는 식 7의 계층 간 잔여 신호 보상 과정으로 복원된다.

$$R_B = C^T \cdot \tilde{R}_B \cdot C \quad \text{식 5}$$

$$R = C^T \cdot \tilde{R} \cdot C \quad \text{식 6}$$

$$R_E = R + R_B \quad \text{식 7}$$

4.2 DCT 기반 구조의 계산 복잡도

제안한 DCT 기반 구조에서는 하위 계층에서 DCT 변환 된 잔여 신호와 상위 계층에서 DCT 변환 된 잔여 신호의 차분을 수행하므로 부호화 연산은 다음과 같다.

$$\tilde{R}_E = C \cdot R_E \cdot C^T \quad \text{식 8}$$

$$\tilde{R}_B = C \cdot R_B \cdot C^T \quad \text{식 9}$$

$$\tilde{R} = \tilde{R}_E - \tilde{R}_B \quad \text{식 10}$$

식 8과 식 9는 상위 계층과 하위 계층에서의 DCT 변환 연산을 각각 나타내고 식 10는 잔여 신호 예측을 나타낸다.

복호화 과정은 역 DCT 변환을 수행하기 전의 단계에서 식 11의 계층 간 보상을 통해 DCT 변환 된 상위 계층의 잔여 신호를 추출한다. 식 12의 역 DCT 변환 과정을 수행하여 최종 잔여 신호를 복원하게 된다.

$$\tilde{R}_E = \tilde{R} + \tilde{R}_B \quad \text{식 11}$$

$$R_E = C^T \cdot \tilde{R}_E \cdot C \quad \text{식 12}$$

4.3 계산 복잡도 성능 비교

계산 복잡도의 성능은 각 구조의 상위 계층과 하위 계층에서 수행되는 연산의 수로 측정한다. 계층 간 잔여 신호 예측이 16x16의 매크로블록 단위로 수행되므로 매크로블록을 기준으로 계산하였고 상위 계층의 매크로블록에 대응되는 하위 계층은 1/2 해상도로 가정하여 8x8블록으로 계산하였다. 업샘플 과정은 공통으로 수행되므로 제외하였다.

표 1 계산 복잡도 비교

	픽셀 기반 구조	제안된 DCT 기반 구조
부호화 과정	8 DCT+16 DCT +1A	8 DCT+16 DCT +1A
복호화 과정	8 DCT+16 DCT +1A	16 DCT +1A

DCT = 1024M + 768A

DCT 변환의 행렬 연산은 4x4 블록 단위로 이루어 지므로 한 블록당 1024번의 곱셈(M)과 768번의 덧셈(A) 연산이 필요하다. 16개의 블록으로 이루어진 한 개의 매크로블록 당 필요한 총 연산 수는 16384번의 곱셈과 12288번의 덧셈이 된다.

표 1은 총 연산 과정에 따른 계산 복잡도를 비교한 결과이다. 부호화 과정에서는 제안한 DCT 기반의 계층 간 잔여 신호 예측 구조는 비교 대상인 픽셀 기반의 구조의 연산 수와 동일하므로 추가적인 계산이 필요하

지 않음을 알 수 있다. 복호화 과정에서는 매크로블록 당 8 DCT = 1024M+768A 연산의 수를 줄임으로써 복호화 과정 중 약 33%의 계산 복잡도를 감소하는 결과를 보였다.

5. 결론 및 향후 과제

본 논문에서는 H.264/AVC SVC에서 복호화 과정의 계산 복잡도를 줄이기 위하여 DCT 기반의 잔여 신호 예측 구조를 제안하였다. 픽셀 기반에서 수행되던 잔여 신호 예측 과정을 DCT 기반 영역에서 수행하는 구조로 제안하여 H.264/AVC SVC의 문제점인 복호화의 계산 복잡도를 개선하도록 하였다. H.264/AVC SVC에서 픽셀 기반의 잔여 신호 예측 구조와 복호화 과정의 계산 복잡도를 연산 수 비교 방법으로 계산한 결과 약 33% 정도 개선됨을 알 수 있었다. 향후에는 움직임 정보와 텍스처 정보의 계층 간 예측 방법에 적용할 수 있는 연구를 수행할 것이다.

참고 문헌

- [1] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Trans. Circuits Syst. Video Technol., Sep., 2007.
- [2] H. Huang, W. Peng, T. Chiang, H. Hang, "Advances in the scalable amendment of H.264/AVC," IEEE Communications Magazine, vol.45, pp.68-76, Jan., 2007.
- [3] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Joint Draft 10 of SVC Amendment," JVT-W201, April, 2007.
- [4] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Proposed modifications for Joint Scalable Video Model," JVT-W202, April, 2007.
- [5] 박성호, 김원하, 한우진, "H.264/AVC-Scalable Extension의 표준화 연구동향과 알고리즘 분석," 방송공학회논문지, 제10권 제4호, pp. 515-527, Dec., 2005.
- [6] K. Wolf, D. Schrijver, W. Neve, S. Zutter, P. Lambert, R. Walle, "Analysis of Prediction Mode Decision in Spatial Enhancement Layers in H.264/AVC SVC," CAIP 2007, LNCS 4673, pp.848-855, Aug., 2007.
- [7] Iain E. G. Richardson, "H.264 and MPEG-4 Video Compression," Wiley, England, 2003.
- [8] J. Xin, A. Vetro, H. Sun, "Converting DCT Coefficients to H.264/AVC Transform Coefficients," PCM 2004, pp.939-946, Nov., 2004.