일반논문-10-15-6-14

H.264/AVC 비트스트림에서 썸네일 추출을 위한 새로운 방법

홍 승 환[a)‡], 김 지 언[a)], 진 영 민[b)], 권 재 철[b)], 오 승 준[a)]

A New Method for Thumbnail Extraction in H.264/AVC Bitstreams

Seung-Hwan Hong[a)‡], Jieon Kim[a)], Youngmin Chin[b)], Jae-Cheol Kwon[b)], and Seoung-Jun Oh[a)]

요 약

IPTV, DMB 등과 같은 방통융합 기반 멀티미디어 서비스에서는 고성능 비디오 인덱싱을 위한 썸네일 기술이 요구되므로 H.264/AVC 비트스트림으로부터 직접 썸네일을 추출하는 방법이 제안되었다. 그러나 화소 공간에서 정의한 일반화 수식을 주파수 공간으로 변환하는 과정에서 H.264/AVC에 정의된 변환 및 양자화, 그리고 인트라 예측의 반올림을 고려하지 않음으로써 썸네일 화질을 저하시키는 문제점이 있다. 본 논문에서는 H.264/AVC 비트스트림에서의 새로운 썸네일 추출 방법을 제안한다. 제안하는 방법은 H.264/AVC 핵심 변환을 기반으로 주파수 공간에서의 썸네일 추출 방법을 이론적으로 정리하고 확률적 방법으로 인트라 예측의 반올림 보상 방법을 제시한다. 제안한 방법을 통해 얻은 썸네일은 참조 복호화기를 통해 얻은 썸네일과 주관적 화질은 크게 차이 나지 않으며 이전에 제안된 방법과 비교하여 주관적 화질 향상과 최대 8.66 dB의 PSNR 이득을 얻었다.

Abstract

Recently, thumbnail techniques are required to index a high-performance video at digital convergence-based multimedia service like IPTV and DMB. Therefore a thumbnail extraction method in H.264/AVC bitstreams has been proposed. However, thumbnail quality deterioration problem at converting the general equation of spatial domain to frequency domain which is generated by not considering about H.264/AVC transform and quantization processing and rounding-off operation in intra prediction. In this paper, we propose a new thumbnail extraction method in H.264/AVC bitstreams. The proposed scheme is based on H.264/AVC core-transform for a thumbnail extraction in frequency domain, and probability theory, intra rounding-off error compensation. Through the implementation and performance evaluation, the subjective quality difference between the output of our scheme and the output of reference decoder is negligible and better than the conventional method, and moreover PSNR gain by up to 8.66 dB.

Keywords : Thumbnail, frequency domain, H.264/AVC core-transform, rounding-off

Ⅰ. 서 론

최근에 디지털 비디오 처리 기술의 발전은 고화질 디지털 방송(HDTV: High Definition Television), 디지털 멀티미디어 방송(DMB: Digital Multimedia Broadcasting), 인터넷 TV(IPTV: Internet Protocol Television)와 같은 다양한 멀티미디어 서비스 발전에 기여하고 있다. 특히 고화질 디

a) 광운대학교 VIA-멀티미디어 센터
VIA-Multimedia Center
b) KT 중앙연구소
Central R&D Lab., KT, Korea
‡ 교신저자 : 홍승환 (shhong@media.kw.ac.kr)
※ 본 연구는 지식경제부와 한국산업기술진흥원의 전략기술인력양성사업, 2010년도 교내학술연구비 지원으로 수행된 결과임
· 접수일(2010년9월13일),수정일(2010년10월19일),게재확정일(2010년11월4일)

지털 방송이 일반화 되면서 다양한 서비스 어플리케이션이 개발되고 있으며, 고화질, 고해상도의 영상을 위한 고속 동영상 처리 기술들이 요구 되고 있다[1-2]. MPEG(Moving Picture Experts Group)과 같은 비디오 압축 표준들은 효율적인 비디오 데이터 저장과 전송을 위해 사용된다. MPEG은 높은 비디오 화질을 유지하면서 많은 데이터의 감소가 가능하다. 그러나 이러한 압축된 데이터는 이미지 처리를 위한 목적으로 사용하기에는 쉽지 않다. 압축된 데이터를 효과적으로 처리하기 위한 고속 검색 및 복원에 대한 많은 연구가 진행되었다[3-4]. 이러한 연구 중 하나가 축소 영상에 관한 연구이다. 축소 영상은 원본 영상에 대한 작은 크기의 영상 혹은 영상의 DC값으로 이루어진 DC 영상을 의미하며, 일반적으로 썸네일(thumbnail)이라 부른다. 이러한 썸네일은 원본 영상의 일반적인 특징을 가지고 있기 때문에 고속 검색이나 색인에 사용된다. 더욱이 썸네일 추출은 영상을 완전한 크기로 복원하는 것보다 적은 메모리가 요구되므로 MPEG기반의 압축 영상의 효과적인 검색에 사용된다[5-8].

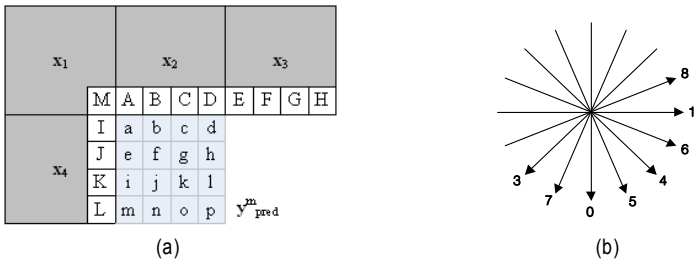
H.264/AVC는 새로운 기술들을 채택함으로써 기존의 비디오 압축 표준보다 높은 압축 효율을 제공한다[9-10]. 이러한 기술 중 인트라 예측은 이전에 복호화 된 주변 블록을 이용하여 현재 블록을 예측하여 공간적 중복성을 제거함으로써 부호화 효율을 높였다. 이러한 기술 특성을 반영하여 H.264/AVC의 인트라 프레임에서 주파수 공간으로부터 썸네일을 추출하는 방법이 이전에 제안되었다[11-14]. 이전에 제안된 H.264/AVC 비트스트림에서의 썸네일 추출 방법은 화소 공간에서 인트라 예측에 대한 일반화 수식을 정의하고, 특수 여과기를 정의하여 변환을 통해 주파수 공간에서 썸네일을 추출 한다[13]. 이 방법은 빠른 썸네일 추출이 가능하지만, 화소 공간에서 정의한 일반화 수식을 주파수 공간으로 변환하는 과정에서 H.264/AVC 표준에 정의된 변환 및 양자화(Transform and Quantization), 그리고 인트라 예측에 사용되는 반올림 연산을 고려하지 않음으로써 썸네일 화질을 저하시키는 문제점이 있다.

본 논문에서는 H.264/AVC 비트스트림에서의 새로운 썸네일 추출 방법을 제안한다. 제안하는 방법은 먼저 화소 공간에서 정의된 일반화 수식을 H.264/AVC 핵심 변환(H.264/AVC Core Transform)을 기반으로 주파수 공간으로 변환하는 과정을 이론적으로 정리하고, H.264/AVC 표준의 역 양자화 과정을 고려한다. 그리고 확률적 방법으로 주파수 공간에서의 인트라 예측의 반올림 오차 보상 방법을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 H.264/AVC의 비트스트림으로부터 썸네일 추출을 위한 이전 연구를 간략히 소개하고 이전 방법의 문제점에 대한 분석을 한다. 3장에서는 2장에서 정리한 문제점을 고려한 새로운 썸네일 추출 방법을 제시한다. 4장에서는 제안한 보상 방법을 적용한 실험과 결과를 보여주며, 5장에서는 결론을 맺는다.

II. H.264/AVC 부호화 기술과 기존의 썸네일 추출 방법

이번 장에서는 H.264/AVC 표준의 부호화 기술 중 인트라 예측과 변환 및 양자화에 대해 간략히 설명한다. 그리고 이를 바탕으로 기존에 제안된 썸네일 추출 방법을 설명하고 문제점에 따른 오차를 수식적으로 분석하도록 한다.



(a) (b)

그림 1. 인트라 4×4 예측 (a) 예측 블록과 참조 화소 (b) 8가지 예측 방향

Fig. 1. Intra 4×4 prediction modes (a) Prediction block and its reference pixels (b) 8 direction of prediction

1. H.264/AVC 인트라 예측

H.264/AVC 인트라 예측 방법은 블록 크기별로 인트라 16×16, 인트라 8×8, 인트라 4×4로 나누어진다. 인트라 16×16 예측은 4가지의 예측 모드를 가지고 인트라 8×8, 인트라 4×4은 9가지의 예측 모드를 가진다. 본 논문에서는 H.264/AVC 고급 프로파일(High Profile)에서 사용되는 인트라 8×8은 고려하지 않고 베이스라인 프로파일(Baseline Profile)에서 사용되는 인트라 16×16, 인트라 4×4만을 고려한다.

그림 1(a)에서 ympred는 현재 예측 블록을 나타내며 x1, x2, x3, x4는 이전에 부호화되어 복원된 주변 블록이다. 그림 1(b)는 8가지의 예측 모드에 따른 예측 방향을 나타내며 예측 모드 2(DC)경우는 방향이 없기 때문에 표시하지 않고 있다. 최종적으로 인트라 예측 블록은 예측 모드의 방향에 따라 주위 참조 화소간의 산술적인 연산을 이용하여 구해진다. 인트라 16×16도 인트라 4×4와 유사한 계산 방법이 사용된다[9].

2. H.264/AVC의 변환 및 양자화

H.264/AVC에서는 인트라 예측 후에 나오는 예측 잔여 성분들에 대하여 변환 및 양자화 과정을 수행한다. 양자화 된 예측 잔여 블록을 $\mathrm{X_Q}$라고 정의 하면 이는 다음과 같이 정리 된다.

$$\begin{aligned}\mathrm{X_Q} &= \mathrm{HT_Q}(\mathbf{x}) \\ &= ((\mathrm{H}_f\mathbf{x}\mathrm{H}_f^{\mathrm{T}})\otimes\mathrm{MF}+f) \gg qbits \\ &= (\mathrm{W}\otimes\mathrm{MF}+f) \gg qbits\end{aligned} \tag{1}$$

여기서 $\otimes$는 원소의 곱(element multiplication)을 의미하고, f는 양자화 dead zone 조절인자 이며, $qbits = 15 + floor(QP/6)$이고, H_f와 MF(multiplication factor)는 각각 H.264/AVC 표준에서 정의하고 있는 핵심 커널(core kernel)과 양자화 테이블이다. 그리고 W는 입력 블록 $\mathbf{x}$가 핵심 변환 된 블록이다. H.264/AVC 핵심 변환 CT(H.264/AVC Core Transform)은 다음과 같이 정의 된다.

$$\mathrm{CT}(\mathbf{x}) = \mathrm{W} = (\mathrm{H}_f\mathbf{x}\mathrm{H}_f^{\mathrm{T}}) \tag{2}$$

양자화 된 블록 $\mathrm{X_Q}$는 다음과 같은 역 양자화 과정을 거치게 된다. $\hat{\mathbf{x}}$을 화소 공간에서의 역 양자화 된 잔여 블록이라고 정의하면 이는 다음과 같이 정리 된다.

$$\begin{aligned}\hat{\mathbf{x}} &= (\mathrm{H}_i^{\mathrm{T}}((\mathrm{X_Q}\otimes\mathrm{V}) \ll q)\mathrm{H}_i) \gg 6 \\ &= (\mathrm{H}_i^{\mathrm{T}}\mathrm{W}'\mathrm{H}_i) \gg 6\end{aligned} \tag{3}$$

여기서 H_i와 V(scaling factor)는 각각 H.264/AVC 표준에서 정의하고 역-핵심 커널(inverse core kernel)과 역 양자화 테이블이며 $q = floor(QP/6)$이다. H.264/AVC 인트라 프레임 복호화 과정은 위에서와 같이 예측 잔여 블록을 화소 공간으로 복원 후에 주변에 복원이 끝난 화소들을 이용하여 예측 블록을 생성하여 더하여 줌으로써 이뤄지게 된다[9-10].

3. 기존의 썸네일 추출 방법

기존 H.264/AVC 비트스트림에서 썸네일 추출하는 방법은 인트라 예측을 일반화된 수식으로 정리하여 주파수 공간에서 직접 추출한다[13].

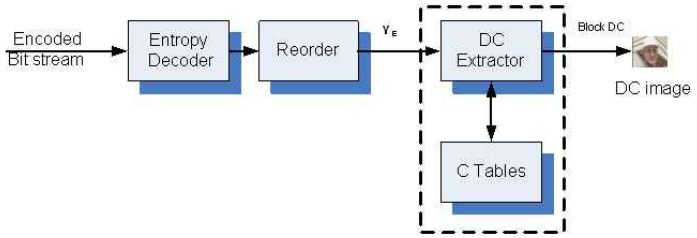


그림 2. 기존의 썸네일 추출기
Fig. 2. Conventional thumbnail extractor

그림 2는 이전에 제안된 썸네일 추출기 구성도이다. 그림 2에서와 같이 H.264/AVC 비트스트림을 완전히 복원하지 않고 DC 추출기(DC Extraction)를 통해 썸네일을 추출할 수 있게 된다. DC 추출기는 식 (4)~(6)을 이용하여 주파수 공간에서 예측 블록의 DC값을 뽑아내고 비트스트림으로부터 예측 잔여 블록의 DC값을 뽑아내어 더해준다.

$$\begin{aligned}\mathrm{Y}^{\mathrm{m}}_{\mathrm{pred},v} &= \mathrm{HT}(\mathbf{x}_{1,v}\mathbf{c}^{\mathrm{m}}_{1,4}+\mathbf{x}_{2,v}\mathbf{c}^{\mathrm{m}}_{2,4}+\mathbf{x}_{3,v}\mathbf{c}^{\mathrm{m}}_{3,4}+(\mathbf{p}^{\mathrm{m}}_{1,4}\mathbf{x}_{4,h})^{\mathrm{T}}) \\ &= \mathrm{X}_{1,v}\mathrm{C}^{\mathrm{m}}_{1,4}+\mathrm{X}_{2,v}\mathrm{C}^{\mathrm{m}}_{2,4}+\mathrm{X}_{3,v}\mathrm{C}^{\mathrm{m}}_{3,4}+(\mathrm{P}^{\mathrm{m}}_{1,4}\mathrm{X}_{4,h})^{\mathrm{T}}\end{aligned} \tag{4}$$

$$\mathrm{Y}^{\mathrm{m}}_{\mathrm{pred},h} = (\sum_{\mathrm{n}=1}^{3}\mathrm{X}_{\mathrm{n},v}\mathrm{Q}^{\mathrm{m}}_{\mathrm{n},h})^{\mathrm{T}}+\mathrm{C}^{\mathrm{m}}_{4,4}\mathrm{X}_{4,h} \tag{5}$$

$$\mathrm{Y}^{\mathrm{m}}_{\mathrm{pred},uni} = \sum_{\mathrm{n}=1}^{3}\mathrm{X}_{\mathrm{n},v}\mathrm{C}'^{\mathrm{m}}_{\mathrm{n}}+(\mathrm{C}'^{\mathrm{m}}_{4})^{\mathrm{T}}\mathrm{X}_{4,h} \tag{6}$$

여기서 m은 인트라 예측 모드를 나타내며 $\mathbf{x}_{\mathrm{n}}$ 은 그림 1(a)에서와 같이 현재 예측될 블록 $\mathbf{y}^{\mathrm{m}}_{\mathrm{pred}}$에 이웃하는 n번째 참조 블록이며 $\mathbf{x}_{\mathrm{n},v}$, $\mathbf{x}_{\mathrm{n},h}$ 는 각각 참조 블록에 수직, 수평 여과기를 통하여 참조 화소로 이루어진 블록이다. $\mathbf{c}^{\mathrm{m}}_{\mathrm{n,i}}$ 는 참조 블록 우측에 곱해지는 상수 행렬을 나타내며 각 모드와 열에 따라 다른 값을 가진다. 화소 공간에서 정의된 $\mathbf{c}^{\mathrm{m}}_{\mathrm{n,i}}, \mathbf{x}_{\mathrm{n},v}, \mathbf{x}_{\mathrm{n},h}$ 는 HT를 이용하여 주파수 공간으로 변환되어 나타내어진다. 이때 HT(H.264/AVC Transform - 근사화된 4×4 DCT)는 식 (7)과 같이 정의된다.

$$\mathrm{HT}(\mathbf{x}) = \mathrm{X} = (\mathrm{H}_f\mathbf{x}\mathrm{H}_f^{\mathrm{T}})\otimes\mathrm{PF}_f \tag{7}$$

여기서 H_f와 PF_f는 각각 H.264/AVC 표준에서 정의하고 있는 정수 변환의 핵심 커널(core kernel)과 후-크기 변환 계수(post scaling factor)이다[9-10]. $\mathrm{C}^{\mathrm{m}}_{\mathrm{n,i}}, \mathrm{X}_{\mathrm{n},v}, \mathrm{X}_{\mathrm{n},h}$ 는 $\mathbf{c}^{\mathrm{m}}_{\mathrm{n,i}}, \mathbf{x}_{\mathrm{n},v}, \mathbf{x}_{\mathrm{n},h}$ 의 HT된 행렬이며 이때 사용되는 Q, C, P 테이블은 인트라 예측 모드에 따른 상수 테이블이므로 미리 계산하여 추가적인 연산을 피한다. DC 추출기에는 다음 블록의 인트라 예측 블록을 복원하기 위해서 $\mathrm{Y}^{\mathrm{m}}_{\mathrm{pred},v}$와 $\mathrm{Y}^{\mathrm{m}}_{\mathrm{pred},h}$를 계산하는 블록이 포함되어 있다.

4. 기존의 썸네일 추출 방법에 대한 문제점 분석

앞서서 식 (4)~(6)에서와 같이 인트라 예측에 대한 화소 공간에서 정의된 일반화 수식을 HT를 통하여 주파수 공간으로 정리함으로써 H.264/AVC 비트스트림으로부터 직접 썸네일 추출이 가능함을 설명하였다. 그러나 실질적으로 H.264/AVC에서는 HT가 아닌 식 (1)에서와 같이 양자화기가 결합된 변환 HT_{Q}(H.264/AVC 변환 및 양자화)를 사용하여 부호화 한다. HT_{Q}는 양자화 과정을 포함하고 있으므로 비트스트림에서의 DC값 추출을 위해서는 역 양자화 과정이 반드시 필요하다. H.264/AVC에서 사용하고 있는 양자화기와 DCT가 결합된 정수 변환이 썸네일 추출에 반영되는 문제점을 분석하기 위해서 주파수 공간에서 역 양자화 된 예측 잔여 블록 $\hat{\mathrm{X}}$이 얻어지는 과정을 정리한다.

$$\frac{\mathrm{MF}}{2^{qbits}} \cong \frac{\mathrm{PF}_f}{Qstep} \tag{8}$$

H.264/AVC는 DCT를 근사화 함으로써 구하여진 PF_f와 양자화 간격 크기 $Qstep$을 이용하여 양자화 파라미터 별로 정수화된 테이블 MF를 얻었다. 그러나 MF를 생성하는 과정에서 오차 ε가 존재하게 된다.

$$\mathrm{MF} = \frac{\mathrm{PF}_f}{Qstep}\times 2^{qbits}+\varepsilon \tag{9}$$

여기서 ε은 MF를 정수화 행렬로 정의하기 위한 정수 근사화 오차 행렬이다[16]. 따라서 H.264/AVC의 정수 변환을 통해 생성된 비트스트림에서 역 양자화 과정을 $Qstep$만

$$\begin{aligned}\mathrm{X}_{\mathrm{Q}}\times Qstep &= (((\mathrm{H}_f\mathrm{x}\mathrm{H}_f^{\mathrm{T}})\otimes\mathrm{MF}+f)\gg qbits)\times Qstep \\ &= (((\mathrm{H}_f\mathrm{x}\mathrm{H}_f^{\mathrm{T}})\otimes\left(\frac{\mathrm{PF}}{Qstep}\times 2^{qbits}+\varepsilon\right)+f)\gg qbits)\times Qstep \\ &= (\mathrm{H}_f\mathrm{x}\mathrm{H}_f^{\mathrm{T}})\otimes\mathrm{PF}+\varepsilon_M \\ &\neq \hat{\mathrm{X}}\end{aligned} \tag{10}$$

을 곱하여 수행함으로써 얻어지는 주파수 공간의 값은 H.264/AVC 표준에서 사용하는 역 양자화 테이블을 이용하여 얻어지는 값과 비교하여 오차를 가지게 된다. 그러므로 이전에 제안된 썸네일 추출 방법으로 얻은 $\hat{\mathbf{X}}$을 이용한 썸네일은 오차를 포함하게 된다. 식 (1)과 식 (9)을 이용하여 정리하면 다음과 같다.

여기서 입력 $\mathbf{x}$에 따라 정수 근사화 오차 행렬 $\boldsymbol{\varepsilon}$이 $\boldsymbol{\varepsilon}_M$으로 변하는 것을 알 수 있다. 따라서 주파수 공간에서 예측 잔여 성분 $\mathbf{Y}_e$로부터 발생하는 오차는 예측이 어렵다.

또한 예측 잔여 성분 $\mathbf{Y}_e$을 이용하여 다음 블록의 예측을 위해 저장되는 값들과 그것을 바탕으로 생성된 예측 값은 정확히 HT를 이용한 주파수 공간의 값이라 할 수 없다. 따라서 식 (4)~(6)을 포함한 이전의 썸네일 추출 방법은 일반화 수식을 주파수 공간으로 표현 할 때 직교 변환 문제가 발생하게 된다.

썸네일 추출을 위한 화소 공간에서 정의된 일반화 수식을 HT를 이용하여 주파수 공간으로 정리할 경우 식 (11)~(13)에서 $\boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred},v}$, $\boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred},h}$, $\boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred},uni}$와 같은 오차가 발생하여 직교성이 성립하지 않게 된다.

그리고 H.264/AVC의 인트라 예측은 이웃 블록의 참조 화소들을 이용하여 각 예측 모드에 따른 예측 블록을 생성한다. 이때 주변 참조 화소들 사이에 필터를 취하는데 반올림 연산을 사용한다. 하지만 화소 공간에서 행렬 연산을 이용하여 일반화 수식을 정리하고 이를 주파수 공간으로 변환하여 사용하는 이전의 썸네일 추출 방법에서는 반올림 연산을 고려하지 못한 문제점이 있다.

HT_Q에 의한 변환을 적용하여 이전에 제안된 방법을 통해 얻은 썸네일 추출 값에 존재하는 오차들을 정리하면 다음과 같다.

$$\mathbf{y}_{\mathrm{dc}} = \frac{1}{4}\left(\left(\mathbf{Y}^{\mathrm{m}}_{\mathrm{pred}} + \boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred}} + \boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{r}}\right) + \left(\mathbf{Y}_{\mathrm{e}} + \boldsymbol{\varepsilon}_M\right)\right)_{(0,0)} \tag{14}$$

여기서 HT_Q을 통한 주파수 공간에서 (0,0) 위치 값(DC)은 화소 공간에서의 블록 평균을 나타내며, 화소 공간에서의 블록 평균보다 4배가 크다. 따라서 썸네일 $\mathbf{y}_{\mathrm{dc}}$는 1/4을 곱하여 구할 수 있다. 또한 $\mathbf{Y}^{\mathrm{m}}_{\mathrm{pred}}$은 주파수 공간에서의 현재 예측 행렬이고 이와 함께 $\boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{r}}$은 인트라 예측 모드 별로 반올림 연산이 고려되지 않음으로 인해서 발생하는 오차 행렬, $\boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred}}$은 직교 변환 문제에 따른 오차 행렬이다. 그리고 $\mathbf{Y}_e$은 예측 잔여 행렬, $\boldsymbol{\varepsilon}_M$은 역 양자화 과정에서 발생하는 오차 행렬이다. 따라서 식 (14)에서와 같이 H.264/AVC 구조상 부호화된 비트스트림으로부터 직접 주파수 공간에서 썸네일을 추출하는 이전의 방법은 오차를 포함하게 된다. 그중에서 잔여성분에 대한 오차 $\boldsymbol{\varepsilon}_M$은 예측 잔여 성분에 따라서 오차가 변하므로 H.264/AVC의 변환과 양자화 과정이 분리되지 않는 한 완전한 보상이 어렵다.

앞서 설명된 이러한 오차들은 인트라 프레임 부호화 특성상 추출된 썸네일 화면 우측 하단으로 갈수록 오차가 전파되며 커지는 특성을 가지기 때문에 위에서 설명한 문제점에 대한 고려가 필수적이다. 다음 장에서는 이론적으로 분석된 문제점을 바탕으로 새로운 썸네일 추출 방법에 대해서 설명한다.

$$\begin{aligned}\mathbf{Y}^{\mathrm{m}}_{\mathrm{pred},v} &= \mathbf{X}_{1,v}\mathbf{C}^{\mathrm{m}}_{1,4} + \mathbf{X}_{2,v}\mathbf{C}^{\mathrm{m}}_{2,4} + \mathbf{X}_{3,v}\mathbf{C}^{\mathrm{m}}_{3,4} + \left(\mathbf{P}^{\mathrm{m}}_{1,4}\mathbf{X}_{4,h}\right)^{\mathrm{T}} + \boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred},v} \\ &\neq \mathrm{HT}\left(\mathbf{y}^{\mathrm{m}}_{\mathrm{pred},v}\right)\end{aligned} \tag{11}$$

$$\begin{aligned}\mathbf{Y}^{\mathrm{m}}_{\mathrm{pred},h} &= \left(\sum_{n=1}^{3}\mathbf{X}_{n,v}\mathbf{Q}^{\mathrm{m}}_{n,h}\right)^{\mathrm{T}} + \mathbf{C}^{\mathrm{m}}_{4,4}\mathbf{X}_{4,h} + \boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred},h} \\ &\neq \mathrm{HT}\left(\mathbf{y}^{\mathrm{m}}_{\mathrm{pred},h}\right)\end{aligned} \tag{12}$$

$$\begin{aligned}\mathbf{Y}^{\mathrm{m}}_{\mathrm{pred},uni} &= \sum_{n=1}^{3}\mathbf{X}_{n,v}\mathbf{C}'^{\mathrm{m}}_{n} + \left(\mathbf{C}'^{\mathrm{m}}_{4}\right)^{\mathrm{T}}\mathbf{X}_{4,h} + \boldsymbol{\varepsilon}^{\mathrm{m}}_{\mathrm{pred},uni} \\ &\neq \mathrm{HT}\left(\mathbf{y}^{\mathrm{m}}_{\mathrm{pred},uni}\right)\end{aligned} \tag{13}$$

III. 제안하는 썸네일 추출 방법

1. H.264/AVC 핵심 변환 기반의 썸네일 추출 방법

H.264/AVC 표준에서 사용하고 있는 변환 및 양자화 HT_Q는 식 (1)과 식 (2)를 통해 핵심 변환 CT와 양자화 테이블 MF간의 연산으로 이뤄짐을 알 수 있다. 따라서 핵심 변환을 기반으로 주파수 공간에서 썸네일 추출을 위한 일반화 수식을 정의한다. 핵심 변환을 기반으로 주파수 공간에서 썸네일을 추출하기 위해서는 식 (4)~(6)에서 볼 수 있듯이 일반화 수식을 주파수 공간으로 나타낼 때 직교 변환을 만족한다는 전제 조건이 필요하다. 그러나 H.264/AVC 핵심 변환은 직교성을 가지지 않는다[10]. 따라서 핵심 변환을 이용한 주파수 공간에서 행렬 곱 사이에 직교성을 만족시키기 위해서 다음과 같은 방법을 제안한다.

$$\begin{aligned} CT(\mathbf{xc}) &= H_f \mathbf{x} \mathbf{c} H_f^T \\ &= H_f \mathbf{x} H_f^T O H_f \mathbf{c} H_f^T \\ &= CT(\mathbf{x}) O CT(\mathbf{c}) \end{aligned} \tag{15}$$

여기서 $O = \mathrm{diag}\left\{\frac{1}{4}, \frac{1}{10}, \frac{1}{4}, \frac{1}{10}\right\}$이고 $H_f^T O H_f = I$ 을 만족한다. 따라서 직교성이 성립하게 되며 식 (15)을 이용하여 식 (4)~(6)을 핵심 변환을 기반으로 하는 주파수 공간으로의 표현이 가능해진다.

$$\begin{aligned} Y_{pred,v}^m &= CT(\mathbf{x}_{1,v}\mathbf{c}_{1,4}^m + \mathbf{x}_{2,v}\mathbf{c}_{2,4}^m + \mathbf{x}_{3,v}\mathbf{c}_{3,4}^m + (\mathbf{p}_{1,4}^m \mathbf{x}_{4,h})^T) \\ &= X_{1,v} O C_{1,4}^m + X_{2,v} O C_{2,4}^m + X_{3,v} O C_{3,4}^m + \\ &\quad (P_{1,4}^m O X_{4,h})^T \end{aligned} \tag{16}$$

$$\begin{aligned} Y_{pred,h}^m &= CT(\mathbf{y}_{pred,h}^m) \\ &= \left(\sum_{n=1}^{3} X_{n,v} O Q_{n,h}^m\right)^T + C_{4,4}^m O X_{4,h} \end{aligned} \tag{17}$$

$$\begin{aligned} Y_{pred,uni}^m &= CT(\mathbf{y}_{pred,uni}^m) \\ &= \sum_{n=1}^{3} X_{n,v} O C_n'^m + (C_4'^m)^T O X_{4,h} \end{aligned} \tag{18}$$

따라서 식 (12)~(14)의 직교 변환 문제로 발생된 오차가 제안하는 방법을 통하여 식 (16)~(18)에서와 같이 해결 될 수 있음을 확인 하였다.

그러나 핵심 변환을 기반으로 썸네일을 추출하기 위해서는 예측 블록 생성을 위한 일반화 수식의 주파수 공간으로의 변환뿐만 아니라 비트스트림으로부터 얻을 수 있는 예측 잔여 성분에 대한 역 양자화 과정도 고려되어져야 한다. H.264/AVC 복호화기는 비트스트림으로부터 영상을 복원하기 위해서 엔트로피 복호화 과정을 거친 후에 역 양자화 테이블을 이용하여 역 양자화 과정을 거치게 된다. 즉, 식 (3)에서와 같이 양자화 파라미터에 따른 V가 적용되어 역 양자화를 수행하게 되고 그 후에 얻어지는 W'에 H.264/AVC 핵심 역 변환 과정을 통해 화소공간으로 복원하는 과정을 수행한다. 따라서 이러한 과정을 수식적으로 정리하여 실제 복호화기에 사용되는 화소 공간으로 복원된 예측 잔여 블록 $\hat{\mathbf{x}}$을 핵심 변환 기반의 주파수 공간 값으로 표현하기 위한 방법을 기술한다[15].

식 (3)에서 H.264/AVC에서 화소 공간으로 복원된 예측 잔여 블록 $\hat{\mathbf{x}}$은 다음과 같이 바꾸어 나타 낼 수 있다.

$$\begin{aligned} \hat{\mathbf{x}} &= \frac{1}{64}(H_i^T((X_Q \otimes V) \ll q) H_i) \\ &= \frac{1}{64}(H_i^T W' H_i) \\ &= \frac{1}{64}((H_f)^{-1} P W' P (H_f^T)^{-1}) \end{aligned} \tag{19}$$

여기서 $P = \mathrm{diag}\{4,5,4,5\}$이고 $H_i^T P^{-1} H_f = I$ 을 만족 한다. 식 (19)은 식 (2)의 핵심 변환을 이용하여 주파수 공간으로 나타 낼 수 있다.

$$\begin{aligned} CT(\hat{\mathbf{x}}) &= H_f \hat{\mathbf{x}} H_f^T \\ &= \frac{1}{64}(P W' P) \\ &= \frac{1}{64} W' \otimes P' \end{aligned} \tag{20}$$

여기서 P 행렬은 P'으로 묶어서 나타낼 수 있으며 다음과 같이 정리된다.

$$P' = \begin{bmatrix} 16 & 20 & 16 & 20 \\ 20 & 25 & 20 & 25 \\ 16 & 20 & 16 & 20 \\ 20 & 25 & 20 & 25 \end{bmatrix} \tag{21}$$

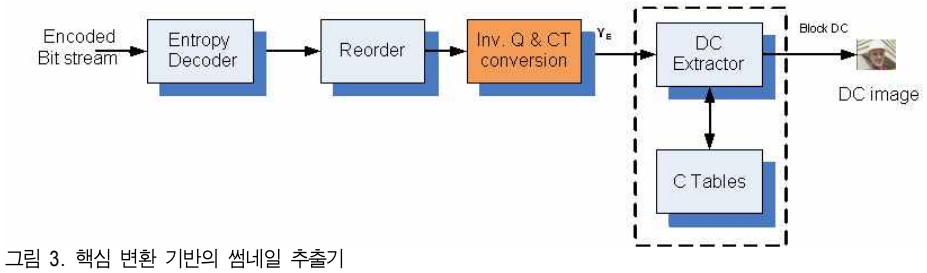


그림 3. 핵심 변환 기반의 썸네일 추출기
Fig. 3. Proposed thumbnail extractor of core transform base

따라서 H.264/AVC 비트스트림으로부터 엔트로피 복호화 과정 후에 얻어지는 예측 잔여 블록 X_Q에 대해서 H.264/AVC 표준과 동일하게 역 양자화 과정을 수행하고 W'에 대하여 식 (20)과 같이 P' 행렬을 이용하여 복호화기에서 얻을 수 있는 화소 공간으로 복원된 예측 잔여 블록 $\hat{x}$을 핵심 변환 기반의 주파수 공간의 값으로 나타낼 수 있다.

그림 3은 핵심 변환 기반의 주파수 공간에서의 새로운 썸네일 추출기 구성도이다.

2. 반올림 예측을 통한 반올림 오차 보상 방법

H.264/AVC의 인트라 예측은 주변 블록의 복호화가 끝난 화소들을 이용하여 다음과 같이 예측 된다.

$$y^3 = \begin{bmatrix} a & b & c & d \\ b & c & d & e \\ c & d & e & f \\ d & e & f & g \end{bmatrix} \tag{22}$$

위 식에서 $a = y^3_{pred,(0,0)} = \lfloor (A+2B+C+2)/4 \rfloor$ 이다. 여기서 A, B, C는 그림 1에서와 같이 인트라 4×4 예측 블록 주변 블록의 참조 화소를 나타낸다. 식 (22)과 같이 모드별 예측 블록 값은 예측 위치 각각에 따른 반올림 연산 후 정수 값으로 나타내게 된다. 그러나 이전에 제안된 썸네일 추출 방법에서 사용되는 일반화 수식에는 반올림에 대한 고려가 없었다. 그로 인하여 반올림 오차가 발생하고 그 오차가 전파되는 현상이 발생하였다. 이러한 오차를 보상하기 위한 방법을 다음과 같이 제안한다. 반올림 오차에 대해서 아래와 같은 가정을 취할 수 있다. 만약 4로 나눌 경우 나머지 값(0, 1, 2, 3)은 모두 1/4의 확률을 갖는다. 따라서 반올림 오차에 대한 예측 값은 다음과 같다[17].

$$\begin{aligned} & y^3_{r,(0,0)} \\ & = (0\times\frac{1}{4}) - (\frac{1}{4}\times\frac{1}{4}) + (\frac{1}{2}\times\frac{1}{4}) + (\frac{1}{4}\times\frac{1}{4}) = \frac{1}{8} \end{aligned} \tag{23}$$

식 (23)에서 $y^3_{r,(0,0)}$은 $y^3_{pred,(0,0)}$에 대한 반올림 오차에 대한 예측 값을 나타낸다. 위의 방법과 같이 나누는 값에 따라 반올림 연산에 대한 예측 값을 구할 수 있다. 인트라 예측은 각 모드에 따라 나누는 값이 결정되어 있다. 그러므로 인트라 예측 모드 정보를 통해 각 모드에 따른 반올림 오차 보상 값을 구할 수 있다. 화소 공간에서 각 모드에 따른 반올림 오차에 대한 예측 행렬을 y^m_r로 정의한다.

인트라 예측 모드에 따라 y^m_r이 다르게 정의 되므로 인트라 예측 모드 2에 대해서 예를 들어 설명하도록 한다. 인트라 예측 모드 2는 왼쪽과 위쪽의 참조 화소들의 평균을 통해 예측 블록을 생성한다. 모드 2의 예측 블록은 다음과 같은 값으로 채워진다.

$$\begin{aligned} & y^2_{pred,(i,j)} = (A+B+C+D+I+J+K+L+4)/8, \\ & i,j = 0,1,2,3 \end{aligned} \tag{24}$$

여기서 A~D, I~L은 인트라 예측을 위한 주변 참조 화소를 나타낸다. 식 (23)과 같은 방법으로 반올림 오차에 대한 예측 값을 구하면 다음과 같다.

$$\begin{aligned} y^2_{r,(i,j)} &= (0\times\frac{1}{8})-(\frac{1}{8}\times\frac{1}{8})-(\frac{2}{8}\times\frac{1}{8})-(\frac{3}{8}\times\frac{1}{8}) \\ &\quad +(\frac{4}{8}\times\frac{1}{8})+(\frac{3}{8}\times\frac{1}{8})+(\frac{2}{8}\times\frac{1}{8})+(\frac{1}{8}\times\frac{1}{8}) \\ &= \frac{1}{16} \end{aligned} \tag{25}$$

식 (25)을 이용하여 예측 모드 2에 대한 오차 보상 행렬을 다음과 같이 정리 할 수 있다.

$$\mathbf{y}^2_r = \frac{1}{16}\mathbf{u} \tag{26}$$

인트라 예측 블록을 생성하기 위해서는 전체 주변 블록의 값이 필요하지 않고 $\mathbf{x}_{n,v}$, $\mathbf{x}_{n,h}$ 값만 필요하다. 또한 썸네일 추출 방법은 주파수 공간에서 적용된다. 따라서 반올림 오차 보상 역시 인트라 예측 모드 별로 핵심 변환을 이용하여 주파수 공간에서의 LUT 형태로 저장하게 된다.

$$\begin{aligned} &\mathbf{Y}^m_{r,v} = \mathrm{CT}(\mathbf{y}^m_{r,v}) = \mathbf{Y}^m_{r,v} \\ &\text{여기에서 } \mathbf{y}^m_{r,v} = \mathrm{V}(\mathbf{y}^m_r) \end{aligned} \tag{27}$$

$$\begin{aligned} &\mathbf{Y}^m_{r,h} = \mathrm{CT}(\mathbf{y}^m_{r,h}) = \mathbf{Y}^m_{r,h} \\ &\text{여기에서 } \mathbf{y}^m_{r,h} = \mathrm{H}(\mathbf{y}^m_r) \end{aligned} \tag{28}$$

식 (27), (28)을 통해 $\mathbf{Y}^m_{r,v}$, $\mathbf{Y}^m_{r,h}$과 같이 각 모드에 따른 반올림 오차 보상 행렬을 LUT로 저장하여 사용하게 된다. 다음 블록에서의 썸네일 추출을 위해서는 인트라 예측 블록 생성에 필요한 $\mathbf{Y}_V$, $\mathbf{Y}_H$ 성분을 저장하게 된다. 따라서 인트라 예측에 대한 반올림 오차 보상을 위해서는 식 (27), (28)에서 구한 값을 $\mathbf{Y}_V$, $\mathbf{Y}_H$ 성분에 포함하여 적용해야만 다음 인트라 블록을 위해 저장되는 값에 반올림 오차에 대한 예측 값을 보상함으로써 오차에 대한 전파를 막을 수 있다. 또한 화면에 출력되어져야 하는 썸네일에도 오차 보상 DC값이 포함되어져야 한다.

3. 제안하는 썸네일 추출 방법의 구조

위에서 설명한 것과 같이 본 논문에서는 H.264/AVC 핵심 변환을 기반으로 주파수 공간에서의 썸네일 추출 방법과 확률적 방법을 통한 인트라 반올림 오차 보상을 제안하

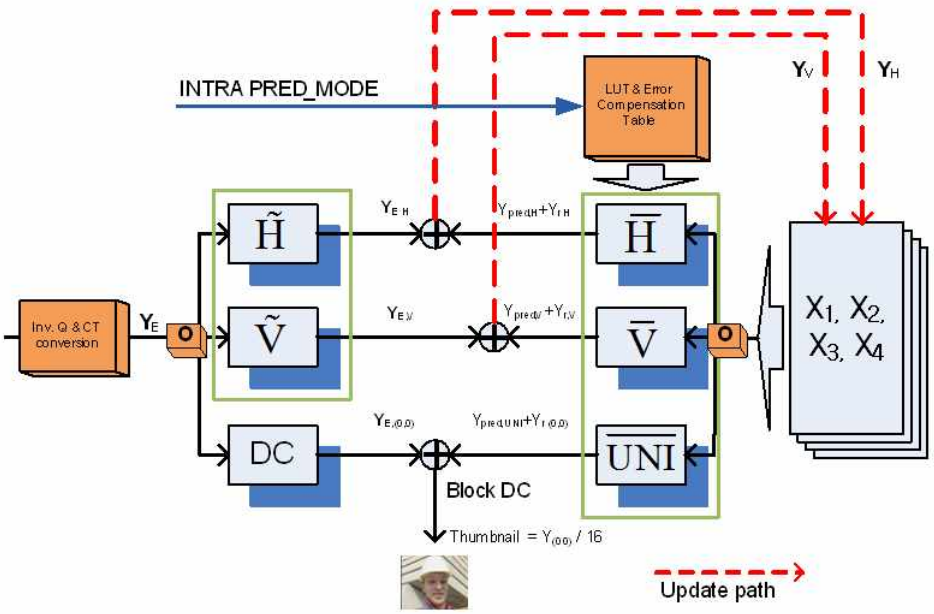


그림 4. 제안하는 DC 추출기의 구조
Fig. 4. The structure of the proposed DC extractor

였다.

그림 4는 제안하는 방법에 대한 DC 추출기의 구조이다. 제안하는 추출기는 핵심 변환을 기반으로 하는 주파수 공간에서 모든 동작이 이뤄지며 순서는 다음과 같다.

Step 1) 입력된 비트스트림으로부터 제안하는 역 양자화 과정을 거쳐 생성된 $\mathbf{Y}_{e}$에 대하여 각각 직교성이 성립하기 위하여 O행렬을 이용하여 수직과 수평 여과 과정을 수행하여 $\mathbf{Y}_{e,v}$와 $\mathbf{Y}_{e,h}$를 생성한다.

Step 2) 직교성이 성립하기 위하여 O행렬을 이용하여 LUT와 이전에 계산된 $\mathbf{x}_{n,v}$, $\mathbf{x}_{n,h}$을 통해 마찬가지로 $\mathbf{Y}_{pred,uni}$, $\mathbf{Y}_{pred,v}$, $\mathbf{Y}_{pred,h}$을 구하고 오차 보상 테이블을 이용하여 $\mathbf{Y}_{r,(0,0)}$, $\mathbf{Y}_{r,v}$, $\mathbf{Y}_{r,h}$을 구하여 각각을 더한다.

Step 3) $(\mathbf{Y}_{pred,uni}+\mathbf{Y}_{r,(0,0)})$와 $\mathbf{Y}_{e,(0,0)}$을 더한 결과는 핵심 변환의 주파수 공간에서의 DC값이므로 1/16을 곱하여 썸네일 값을 구한다.

Step 4) 다음의 인트라 예측 블록 생성을 위하여 $\mathbf{Y}_{V}$, $\mathbf{Y}_{H}$을 저장한다.

Ⅳ. 실험 및 결과

이 장에서는 본 논문에서 제안하는 썸네일 추출 방법과 이전에 제안된 썸네일 추출 방법에 대해 객관적인 화질 및 주관적 화질을 비교한다. 참조 영상 REF(Reference)은 원 영상을 4x4 블록 당 16개 화소의 평균을 취하는 방법으로 가로, 세로 각각 1/4의 크기로 축소한 영상을 사용한다. 제안하는 방법의 성능 실험을 위하여 H.264/AVC 비트스트림으로부터의 썸네일 추출 방법을 나열하면 다음과 같다. SDM(Spatial Domain Method)은 H.264/AVC 레퍼런스 복호화기를 통해 화소 공간으로 완전히 복호화 된 영상을 가로, 세로 각각 1/4의 크기로 축소한 영상을 추출하는 방법이다. HTM(H.264/AVC Transform Method)은 기존에 제안된 방법으로써 HT(H.264/AVC 변환)를 이용한 주파수 공간에서의 썸네일 추출 방법이다. CTM(Core Transform Method)은 본 논문에서 제안하는 방법으로써 CT(H.264/AVC 핵심 변환)을 이용한 주파수 공간에서의 썸네일 추출 방법이고, CTMEC(Core Transform Method with Error Compensation)은 제안한 CTM에 반올림 오차 보상을 추가한 방법이다. 실험에 적용된 양자화 파라미터는 일반적으로 많이 사용되는 양자화 대역을 선택하였다. 표 1은 실험 환경을 나타낸다.

표 1. 실험 환경
Table 1. Experimental Environment

구분		실험환경
참조 소프트웨어		JM 12.2[18] BP, Level 4.0
프레임 부호화 모드		인트라 모드
적응형 프레임/필드 모드		프레임 모드
양자화 계수		22, 27, 32, 37
실험 영상	704×576	City, Crew, Harbour, Soccer
	1920×1080	Crowd run, Rush hour, Sunflower, Toys and Calendar
썸네일 추출 크기		176×144 / 480×270

제안하는 CTM의 역 양자화 방법의 필요성을 증명하기 위해서 각각의 방법에 대한 예측 잔여 블록의 오차를 그림 5에서 확인한다.

그림 5는 영상의 휘도 성분에 대해서 각각의 방법으로 구한 예측 잔여 블록과 실제 복호화기에서 얻을 수 있는 인트라 예측 잔여 블록 사이의 오차를 실험 영상의 프레임 당 평균 SSE(Sum of Squared Error)를 통하여 비교한 것이다. 그림 5의 결과는 앞서 2장에서 수식으로 증명 하였듯이 HTM에서와 같이 양자화 간격 크기 $Qstep$만을 곱하여 역 양자화를 수행하게 될 경우 H.264/AVC의 역 양자화 과정과 비교하여 예측 잔여 블록에 오차가 있음을 나타내고 있다. 이렇게 생성되는 HTM에서의 예측 잔여 블록은 썸네일 추출에 필요한 수직/수평 성분에 대해서도 오차를 포함하게 되는데 이러한 오차가 다음 블록에 전파되어 썸네일 화질이 저하되는 문제가 있다. 특히 양자화 파라미터가 37일 때 HTM의 오차가 급격히 줄어드는 것을 볼 수 있는데 이것은 양자화 파라미터가 큰 값 일수록 예측 잔여 블록의

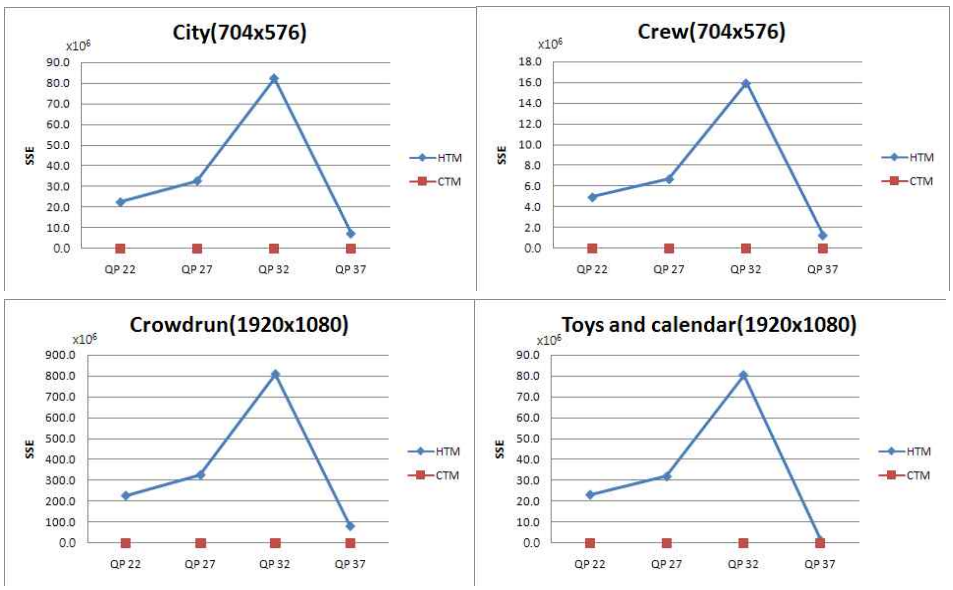


그림 5. 역 양자화 후의 예측 잔여 블록에 대한 오차
Fig. 5. The predicted residual block error after inverse quantization

성분들이 0값을 많이 가지게 됨으로 양자화 오차도 적어지게 되는 것이다. 또한 일반적으로 인트라 예측에서 복잡한 영상은 평탄한 영상에 비해 예측이 잘 되지 않음으로 예측 잔여 블록 성분의 절대 값 들이 커지고 주파수 공간에서 AC 성분들이 많이 분포하게 된다. 이것은 2장에서 식 (10)을 통해 설명한 것과 같이 HTM의 역 양자화 오차는 예측 잔여 블록에 의존적이게 되는데 특히 AC 성분에 오차가 많이 포함되게 된다. 그에 따라 City(704×576)와 Crowdrun(1920×1080) 같은 복잡한 영상에서 역 양자화 오차가 크게 발생함을 확인 할 수 있다. 이에 반해 CTM은 양자화 파라미터와 예측 잔여 블록에 관계없이 언제나 역 양자화 오차가 없음을 보인다. 따라서 CTM에서와 같은 역 양자화 방법을 사용함으로써 예측 잔여 블록에서 오차가 발생하는 문제점을 해결하였음을 보였다.

표 2는 객관적 화질을 비교하기 위하여 각각 704×576 크기와 1920×1080 크기의 실험 영상에 양자화 파라미터 별로 썸네일 추출 방법을 달리하여 PSNR을 측정 한 것이다. PSNR은 참조 썸네일 REF와 각각 HTM, CTM,

표 2. 양자화 파라미터 별 썸네일 추출 방법에 따른 PSNR 비교
Table 2. Performance comparison in PSNR for thumbnail extraction methods at each QP

Sequence	HTM				CTM				CTMEC			
(704x576)_100frame	QP22	QP27	QP32	QP37	QP22	QP27	QP32	QP37	QP22	QP27	QP32	QP37
City	21.21	18.95	17.63	30.73	21.58	18.73	17.97	31.95	26.78	22.11	20.68	35.24
Crew	21.01	20.41	22.34	29.99	21.19	20.35	22.96	33.22	26.34	23.84	26.47	37.45
Harbour	20.69	17.93	17.14	32.67	20.53	17.98	17.01	33.97	25.75	21.59	19.98	30.56
Soccer	20.54	18.80	20.03	29.63	20.35	18.60	19.99	30.82	24.60	21.50	22.37	34.75

Sequence	HTM				CTM				CTMEC			
(1920x1080)_60frame	QP22	QP27	QP32	QP37	QP22	QP27	QP32	QP37	QP22	QP27	QP32	QP37
Crowd run	15.31	16.43	14.43	26.19	15.82	16.07	14.76	26.32	23.41	19.83	16.96	33.56
Rush hour	16.00	17.10	20.44	25.31	16.15	17.48	21.29	29.65	21.05	20.40	24.39	33.96
Sunflower	13.58	12.71	14.75	23.27	13.65	12.85	15.00	25.21	18.36	15.32	17.90	30.95
Toys and calendar	15.42	15.01	16.05	26.09	15.50	15.02	16.13	28.04	19.26	17.45	18.61	31.08

CTMEC 방법들로 추출된 썸네일과 비교하여 측정하였다.

표 2의 결과에서 이전에 제안된 방법인 HTM과 비교하여 본 논문에서 새롭게 제안하는 방법인 CTM과 CTMEC은 각각 최대 4.34 dB와 최대 8.66 dB의 PSNR 이득을 얻었다. 그리고 표 2의 결과에서 양자화 파라미터가 증가할 때 PSNR이 증가하는 현상을 보인다. 이것은 인트라 예측 모드의 선택 비율과 연관되어 설명된다.

그림 6은 양자화 파라미터 별로 인트라 예측 모드의 선택 비율을 나누어 놓은 것으로써 각각 Crew 영상을 100프레임, Crowd run 영상을 60프레임 부호화 하였을 때 4×4 블록 마다 예측 모드의 선택 비율을 나타낸 것이다. 그림 6에서 보듯이 양자화 파라미터가 증가할수록 반올림 오차가 없는 인트라 16×16 예측 수직, 수평 모드가 많이 선택됨을 알 수 있다. 이와 함께 표 2에서 양자화 파라미터가 증가할수록 PSNR이 높아지는 현상이 발생하게 된다. 이를 통해 반올림 오차를 발생시키는 예측 모드가 많이 선택되는 양자화 파라미터에서의 PSNR을 보게 되면 CTMEC를 통하여 화질을 개선하였지만 반올림 오차에 대한 보상이 완전하지 않았다는 것을 보인다. 이것으로 화소 공간에서 수행된 연산을 주파수 공간에서 예측하여 보상하는 과정으로는 완전한 반올림 오차 보상이 수행 되지 않는다는 것을 알 수 있다.

그림 7은 기존에 썸네일 추출 방법인 HTM에서 역 양자화 과정과 일반화 수식의 직교 변환 문제를 해결한 제안하는 썸네일 추출 방법인 CTM을 통해 주관적 화질을 비교하여 보여준다.

그림 7 (a)에서 볼 수 있듯이 HTM을 통한 썸네일은 역 양자화 및 직교 변환 문제로 발생한 오차가 H.264/AVC 인트라 프레임 부호화 특성상 우측 하단 방향으로 계속하여 전파되어져 감을 확인 할 수 있다. 이에 반해 그림 7 (b)에

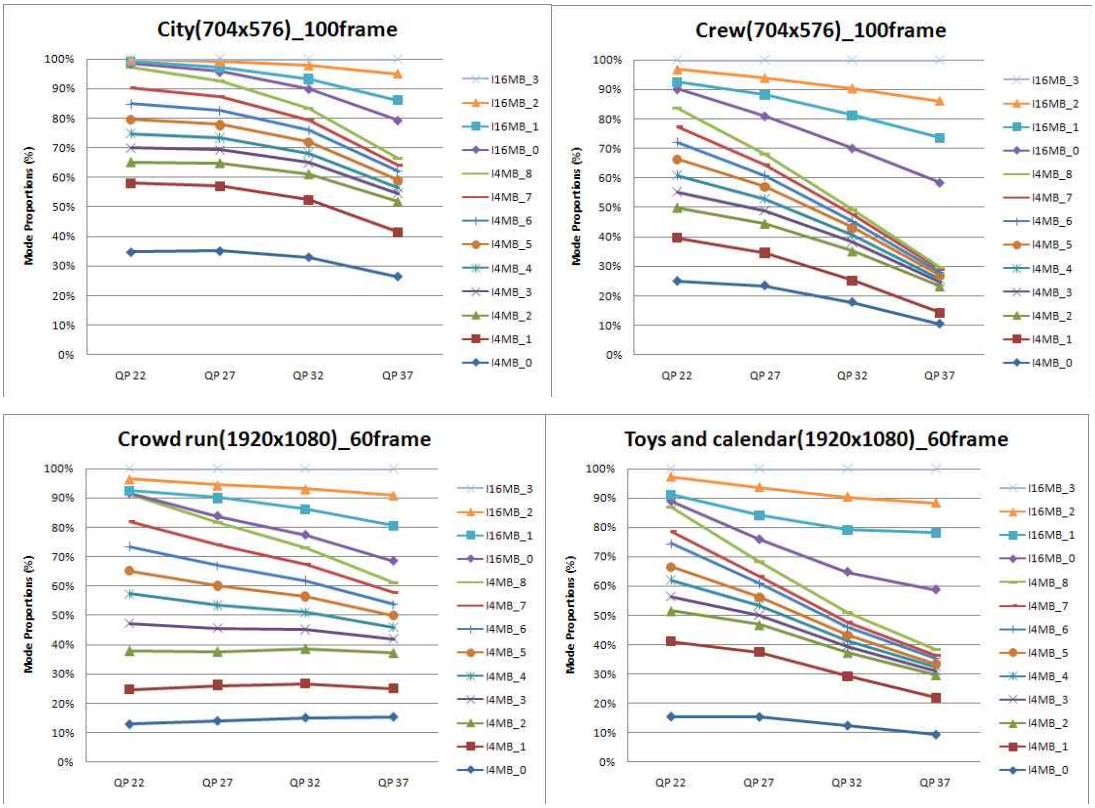


그림 6. 양자화 파라미터 별 인트라 예측 모드 선택 비율
Fig. 6. Intra prediction mode selection ratio per QP

(a) (b)

그림 7. Rush hour 영상에서의 화질비교 (480×270, QP=37) (a) HTM, (b) CTM

Fig. 7. Subjective quality performance of Rush hour sequence (480×270, QP=37) (a) HTM, (b) CTM

서 볼 수 있듯이 역 양자화 및 직교 변환 문제를 해결한 CTM을 통한 썸네일은 주관적 화질이 개선되었음을 보인다.

그림 8~11는 각각의 방법을 통해 실험 영상의 첫 번째 프레임에서 추출된 썸네일의 주관적 화질 비교를 보여준다.

그림 8과 그림 9는 724×576 크기의 실험 영상으로부터 추출한 크기가 176×144인 썸네일이 된다. 그리고 그림 10과 그림 11는 1920×1080 크기의 실험 영상으로부터 추출한 크기가 480×270인 썸네일이 된다. 그림 8~11 (b)에서와 같이 HTM을 통한 썸네일이 보이는 문제점들이 있다. 이러한 문제점은 그림 7 (a)에서와 같이 썸네일의 우측 하단부로 갈수록 화면이 어두워지거나 화면이 고르지 못하게

(a) (b) (c) (d)

그림 8. City 영상에서의 화질비교 (176×144, QP=22) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC 썸네일

Fig. 8. Subjective quality performance of City sequence (176×144, QP=22) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC thumbnail

(a) (b) (c) (d)

그림 9. Soccer 영상에서의 화질비교 (176×144, QP=27) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC 썸네일

Fig. 9. Subjective quality performance of Soccer sequence (176×144, QP=27) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC thumbnail

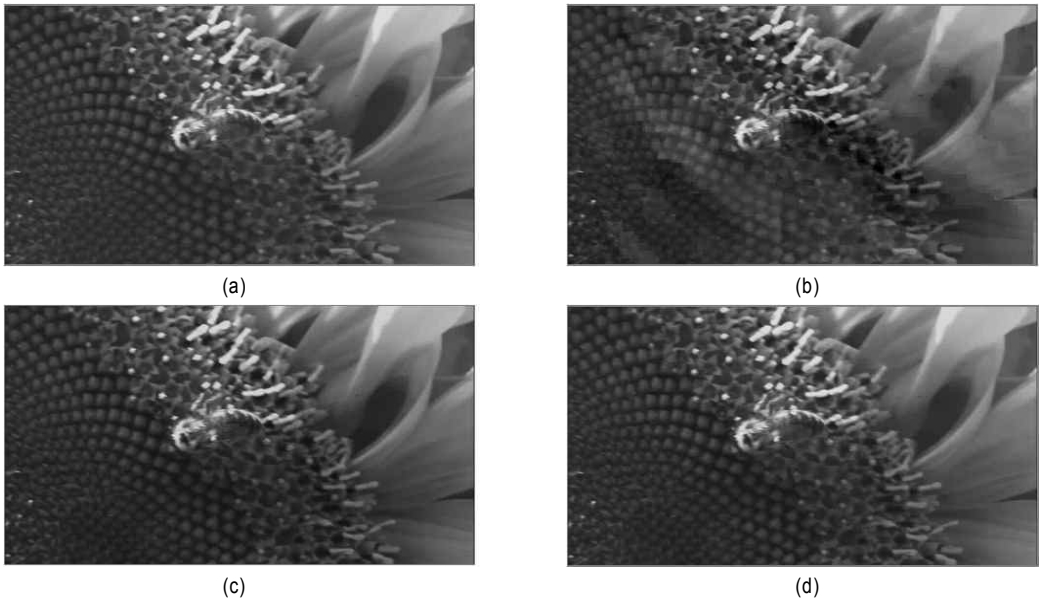


그림 10. Sunflower 영상에서의 화질비교 (480×270, QP=37) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC 썸네일
Fig. 10. Subjective quality performance of Sunflower sequence (704x576, QP=37) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC thumbnail

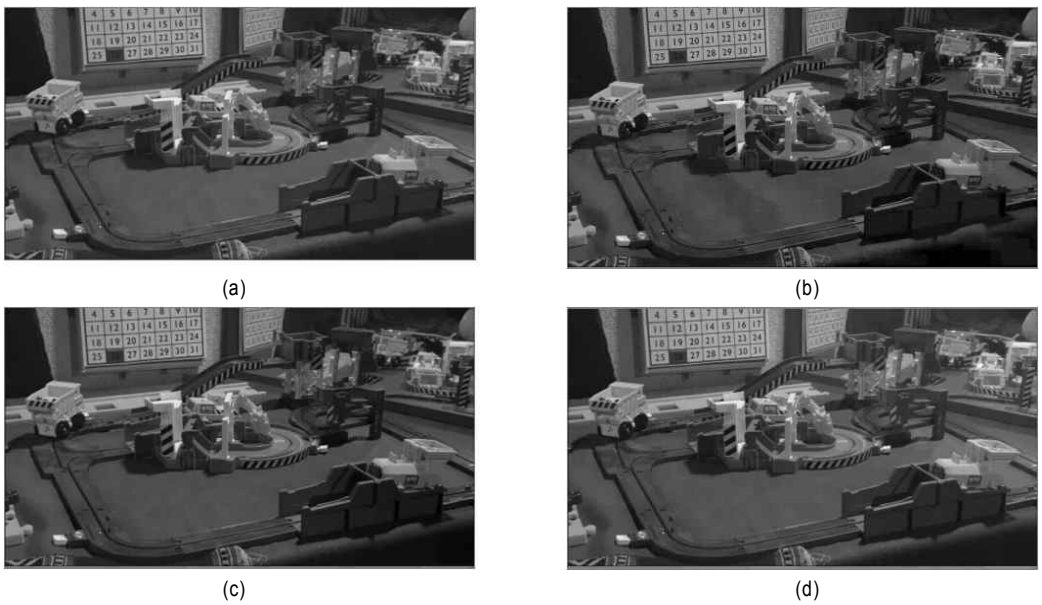


그림 11. Toys and calendar 영상에서의 화질비교 (480×270, QP=37) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC 썸네일
Fig. 11. Subjective quality performance of Toys and calendar sequence (704x576, QP=37) (a) SDM, (b) HTM, (c) CTM, (d) CTMEC thumbnail

되는데 이는 썸네일 추출 시 발생하는 역 양자화 오차와 반올림 오차가 H.264/AVC의 인트라 프레임 부호화 특성상 계속 전파되어 나가는 것이다. 그림 8~11을 통해서 이러한 오차들 중에서 역 양자화 오차와 일반화 수식의 직교 변환 문제를 해결한 CTM을 통한 (c)에서 썸네일 화질이 약간 개선되었으며 반올림 오차 보상을 포함한 CTMEC를

통한 (d)에서의 썸네일은 H.264/AVC 복호화기를 통해 화소 공간으로 복원한 영상을 다운-샘플하는 방법인 SDM을 통한 (a)에서의 썸네일과 유사한 주관적 화질을 보임을 확인하였다.

Ⅵ. 결 론

본 논문에서는 이전에 제안된 썸네일 추출 방법이 화소 공간에서 정의한 일반화 수식을 주파수 공간으로 변환하는 과정에서 H.264/AVC에 정의된 변환 및 양자화, 그리고 인트라 예측의 반올림을 고려하지 않음으로써 썸네일 화질을 저하시키는 문제점을 지적하고 H.264/AVC 비트스트림에서의 새로운 썸네일 추출 방법을 제안하였다. 제안한 방법은 H.264/AVC 핵심 변환을 기반으로 주파수 공간에서의 썸네일 추출과 확률적 방법으로 인트라 반올림 오차를 보상한다. 제안한 방법은 , H.264/AVC 비트스트림 주파수 공간으로부터 썸네일을 추출하기 위하여 표준의 역 양자화 과정을 고려였고, 화소 공간에서 정의된 일반화 수식을 H.264/AVC 핵심 변환(H.264/AVC Core Transform)을 기반으로 주파수 공간으로 변환하는 과정을 이론적으로 정리하였다. 또한 확률적인 방법으로 인트라 반올림 오차를 예측하여 주파수 공간에서 보상한다. 제안한 방법을 통해 얻은 썸네일은 참조 복호화기를 통해 얻은 썸네일과 주관적 화질은 크게 차이 나지 않으며 이전에 제안된 방법과 비교하여 주관적 화질 향상과 최대 8.66 dB의 PSNR 이득을 얻었다. 하지만 제안하는 방법은 인트라 반올림 오차를 확률적으로 예측하여 보상하였기 때문에 참조 복호화기를 통해 추출한 썸네일과 비교하여 화질 차이가 존재 한다. 이를 보완하기 위해서는 향후에 정확한 주파수 공간에서의 인트라 반올림 연산 방법을 위한 연구가 수행되어야 할 것이다.

참 고 문 헌

[1] B. Yeo and B. Liu, "Rapid scene analysis on compressed video", IEEE Trans. Circuit and System for video Technology, vol. 5, no. 6, pp.533-540, 1995.

[2] Wing-San Chau, Oscar C. Au, Tak-Song Chong, Tai-Wai Chan, Chi-Shun Cheung, "Efficient Scene Change Detection in MPEG Compressed Video Using Composite Block Averaged Luminance Image Sequence", IEEE ICICS, pp. 688-691, 2005.

[3] 윤자천, 설상훈, “압축영역에서 빠른 축소 영상 추출을 위한 다중부호 룩업테이블 설계”, 한국방송공학회, vol. 10, no. 3, pp. 413-421, 2005.

[4] N. Merhav and V. Bhaskaran, "Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation", IEEE Trans. Circuit and System for video Technology, vol. 7, pp. 468-476, 1997.

[5] Taskiran, C. Jau-Yuen Chen, Bouman, C.A. Delp, E.J. "A compressed video database structured for active browsing and search," IEEE image Processing, vol.3, pp.133-137, Oct. 1998.

[6] Ja-Cheon Yoon, Hyeokman Kim, Seong Soo Chun, Jung-Rim Kim, Sanghoon Sull, "Real-Time Video Indexing System for Live Digital Broadcast TV Programs," Lecture Notes in Computer Science, CVIR2004, vol. 3115, pp. 261-269, Jul. 2004.

[7] G.C. Feng and J. Jiang, "Image extraction in DCT domain," IEE Proc.-Vis. Image Signal Process., Vol. 150, No. 1, pp.20-27, Feb. 2003.

[8] J. Song and B.L. Yeo, "Fast Extraction of Spatially Reduced Image Sequences from MPEG-2 Compressed Video," IEEE Trans. Circuits and Systems for Video Technology, vol. 9, no.7, pp. 1100-1114. Oct. 1999.

[9] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), 2003.

[10] Henrique S. Malvar, Antti Hallapuro, Marta Karczewicz, and Louis Kerofsky, "Low-Complexity Transform and Quantization in H.264/AVC", IEEE Trans. Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 598-603, July 2003.

[11] Chen Chen, Ping-Hao Wu and Chen H, "Transform-Domain Intra Prediction for H.264," IEEE ISCAS, pp.1497-1500, May 2005.

[12] 유상준, 윤명근, 김은석, 손채봉, 심동규, 오승준, “H.264/AVC 비트스트림에서 효율적으로 축소 영상을 추출하는 방법,” 한국방송공학회, 제13권, 제2호, pp,222-235, Mar, 2008.

[13] Eun-Seok Kim, Tae-Woong Um, and Seoung-Jun Oh, "A Fast Thumbnail Extraction Method in H.264/AVC Video Streams," IEEE Trans. Consumer Electronics, vol. 55, no. 3, pp. 1424-1430, Aug. 2009.

[14] 홍승환, 이여송, 조혜정, 안창범, 오승준, “H.264/AVC 비트스트림에 대한 공간주파수 영역에서의 썸네일 추출 방법,” 한국방송공학회 학술대회, pp,277-280, Nov, 2009

[15] Yu-Kuang Tu, Jar-Ferr Yang, and Ming-Ting Sun, "Efficient Rate-Distortion Estimation for H.264/AVC Coders", IEEE Trans. Circuits and Systems for Video Technology, vol. 16, no. 5, pp. 600-611, May 2006.

[16] Jeong Mee Moon, and Jae Ho Kim, "A New Low-Complexity Integer Distortion Estimation Method for H.264/AVC Encoder", IEEE Trans. Circuits and Systems for Video Technology, vol. 20, no. 2, pp. 207-212, Feb. 2010.

[17] Yuichiro Nakaya, Yoshinori Suzuki, and Satoshi Misaka, "Avoidance of error accumulation in motion compensation," Doc. Q15-A-35R1, June 1997

[18] H.264/AVC Reference Software Model (JM 12.2) : http://iphome.hhi.de/suehring/tml/download/old jm

저 자 소 개

홍 승 환

- 2009년 2월 : 광운대학교 전자공학과 졸업 공학사
- 2009년 3월 ~ 현재 : 광운대학교 대학원 전자공학과 석사과정
- 주관심분야 : H.264/AVC, 비디오 압축, Thumbnail

김 지 언

- 2009년 2월 : 광운대학교 전자공학과 졸업 공학사
- 2009년 3월 ~ 현재 : 광운대학교 대학원 전자공학과 석사과정
- 주관심분야 : H.264/AVC, 비디오 압축

진 영 민

- 1983년 : 연세대학교 전기공학과 학사
- 1985년 : 연세대학교 전기공학과 석사
- 1989년 : 연세대학교 전기공학과 박사
- 1989년 ~ 현재 : KT 중앙연구소 상무보
- 주관심분야 : IPTV, 플랫폼/STB, CAS/DRM, 양방향서비스 기술, UI/UX 등

권 재 철

- 1986년 : 한양대학교 전자공학과 졸업
- 1988년 : KAIST 전기및전자공학과 석사
- 2003년 : KAIST 전기및전자공학과 박사
- 1988년 ~ 현재 : KT 중앙연구소 연구전문 수석연구원
- 주관심분야 : 멀티미디어통신, 영상코덱, 패킷비디오, 영상품질측정

오 승 준

- 1980년 2월 : 서울대학교 전자공학과 졸업(학사)
- 1982년 2월 : 서울대학교 전자공학과 대학원 졸업(석사)
- 1988년 5월 : 미국 Syracuse University 전기/컴퓨터공학과 졸업(박사)
- 1982년 3월 ~ 1992년 8월 : 한국전자통신연구원 근무(멀티미디어연구실 실장)
- 1986년 7월 ~ 1986년 8월 : NSF Supercomputer Center 초청 학생연구원
- 1987년 5월 ~ 1988년 5월 : Northeast Parallel Architecture Center 학생연구원
- 1992년 3월 ~ 1992년 8월 : 충남대학교 컴퓨터공학부 겸임교수
- 1992년 9월 ~ 현재 : 광운대학교 전자공학부 교수 (VIA-Multimedia Center 멀티미디어연구실)
- 2002년 3월 ~ 현재 : SC29-Korea 의장 및 MPEG Forum 부의장
- 주관심분야 : 비디오 데이터 처리, 비디오 코덱 시스템, 비디오 및 영상압축, 멀티미디어시스템