

정규논문 (Regular Paper)

방송공학회논문지 제19권 제4호, 2014년 7월 (JBE Vol. 19, No. 4, July 2014)

<http://dx.doi.org/10.5909/JBE.2014.19.4.468>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## Foveated Contrast Sensitivity를 이용한 인지품질 기반 비디오 코딩

유 지 우<sup>a)</sup>, 심 동 규<sup>a)†</sup>

### Perceptual Quality-based Video Coding with Foveated Contrast Sensitivity

Jiwoo Ryu<sup>a)</sup> and Donggyu Sim<sup>a)†</sup>

#### 요 약

본 논문은 FCS(foveated contrast sensitivity)를 이용한 인지품질 기반 비디오 코딩 방법을 제안한다. CS(contrast sensitivity)를 이용한 기존의 인지품질 기반 비디오 코딩 방법은 공간주파수에 따라 시각적 인지능력이 달라지는 인간시각체계(HVS, human visual system)의 특징을 이용하여 비디오 압축 시 인지품질의 손상을 최소화하며, FM(foveated masking)을 이용한 방법에서는 HVS의 중심시(central vision)와 주변시(peripheral vision)의 차를 이용한다. 본 연구에서는, 정신물리학 실험을 통하여 기존의 DCT(discrete cosine transform)기반 JND(Just-noticeable difference) 모델과 FM이 서로 의존성을 갖고 동시에 고려된 새로운 FCS 모델을 제안하였고, 이를 HM10.0 부호화기에 적용하여 인지품질기반 부호화를 수행하였다. 제안된 방법으로 부호화된 영상은 인지품질 관점에서 동일한 화질을 유지하면서 평균 10%의 비트율 감소를 보였다.

#### Abstract

This paper proposes a novel perceptual quality-based (PQ-based) video coding method with foveated contrast sensitivity (FCS). Conventional methods on PQ-based video coding with FCS achieve minimum loss on perceptual quality of compressed video by exploiting the property of human visual system (HVS), that is, its sensitivity differs by the spatial frequency of visual stimuli. On the other hand, PQ-based video coding with foveated masking (FM) exploits the difference of the sensitivity of the HVS between the central vision and the peripheral vision. In this study, a novel FCS model is proposed which considers both the conventional DCT-based JND model and the FM model. Psychological study is conducted to construct the proposed FCS model, and the proposed model is applied to PQ-based video coding algorithm implemented on HM10.0 reference software. Experimental results show that the proposed method decreases bitrate by the average of 10% without loss on the perceptual quality.

Keywords : Perceptual quality, Perceptual quality-based video coding, foveated contrast sensitivity, HEVC

---

a) 광운대학교(kwangwoon University)

† Corresponding Author : 심동규(Donggyu Sim)

E-mail: dgsim@kw.ac.kr

Tel: +82-2-940-5470

※ 본 연구는 서울시 산학연 협력사업(SS110004M0229111)의 지원을 받았습니다.

· Manuscript received March 19, 2014 Revised June 23, 2014 Accepted July 23, 2014

## 1. 서론

멀티미디어 서비스에 있어 제한된 대역폭에서 영상의 화질을 향상시키는 것은 가장 중요한 목표 중 하나이다. 무선 네트워크 환경 등 제한된 멀티미디어 전송 환경에서 최소한의 대역폭으로 최대한의 화질을 가진 영상을 전송하기 위하여 다양한 영상 부호화 방법이 연구 및 개발되었는데, 비디오 코딩 분야에서는 2013년 1월에 표준화를 완료한 HEVC(high efficiency video coding) 비디오 코덱 표준이 가장 우수한 성능을 갖는 것으로 평가된다<sup>[1]</sup>. HEVC는 이전 비디오 코딩 표준인 H.264/AVC에 비해 50% 높은 부호화 성능을 갖는 것으로 알려져 있다. Sample adaptive offset(SAO)와 같은 새로운 coding tool의 추가 이외에, HEVC가 큰 폭의 성능 향상을 이룰 수 있었던 주요 원인은 기본 부호화 단위인 coding unit의 크기를 유연하게 변경할 수 있는 coding tree unit 구조로 인해 고해상도 비디오의 부호화 성능이 크게 증가하였기 때문이다. 특히 기존 비디오 코덱 표준에서는 고려하지 않았던 UHD(ultra high definition) 해상도 비디오가 HEVC에서는 주요 타겟이 되었는데, 이것은 비디오 영상 해상도의 꾸준한 증가와 UHD 디스플레이 출시 등의 디스플레이 장치의 대형화와 같은 멀티미디어 시장의 트렌드와 무관하지 않다.

비디오 데이터의 양은 비디오 해상도에 비례하여 기하급수적으로 늘어난다. HEVC가 이전의 비디오 코딩 표준에 비해 50%의 성능 향상을 이루었으나, 해상도의 증가로 인해 배로 늘어난 비디오 데이터의 양과 꾸준히 증가하는 네트워크의 비디오 트래픽으로 인해 멀티미디어 서비스 산업에서는 여전히 더 향상된 비디오 부호화 방법을 필요로 하는 것이 실정이다. 이와 같은 필요성을 충족하기 위해, 최근 인지품질기반 비디오 부호화가 활발히 연구되고 있다.

인지품질 기반 비디오 부호화의 목적은 동일한 인지품질을 유지하면서, 또는 최소한의 인지품질 손상을 통해 높은 압축률로 비디오를 부호화하는 것이다. 인지품질 기반 비디오 부호화의 가장 대표적인 접근방법 중 하나는 인간시각체계(HVS, human visual system)의 contrast sensitivity function(CSF)를 이용하는 것이다<sup>[2][8]</sup>. CSF는 시각적 자극의 공간 주파수에 따라 HVS의 예민한 정도가 변하는 특

성으로, 이를 통하여 인지품질 기반 부호화기에서는 HVS가 덜 민감한 주파수 성분을 영상에서 제거함으로써 인지품질의 저하 없이 높은 압축률을 달성하는 것이 가능하다.

인지품질기반 영상 부호화의 또 다른 접근방법은 비디오의 관심영역을 정의한 후 관심영역 이외의 영역을 높은 압축률로 부호화하여 인지품질을 최소화하면서 압축 성능을 높이는 것이다<sup>[9][15]</sup>. 이와 같이 foveated masking(FM)효과를 이용한 방법들은 대부분 시청자의 관심영역을 정의한 후 관심영역 내에서는 화질 중심으로, 이외의 영역에서는 높은 압축률 위주의 부호화를 하여 압축 성능을 향상시킨다. 그러나, FM 효과를 이용한 접근 방식은 관심영역을 정의하는 saliency map을 정확하게 구하는 알고리즘이 없는 한 효율적인 이미지 및 부호화기 적용에 한계가 있다<sup>[3]</sup>.

비디오 영상의 해상도가 점점 증가하고 고해상도의 대형 디스플레이가 시장에 공급되고 있는 현 상황에서는 관심영역을 이용한 인지품질 기반 비디오 부호화가 이전보다 더 유효하게 사용될 수 있다. UHD와 같은 고해상도 영상은 주로 영화관 및 대형 가정용 TV 등 시청자의 시야각의 대부분을 차지하는 대형 디스플레이에 사용된다. 특히 SDTV(standard definition television) 해상도가 사용되던 과거에 비해 디스플레이의 크기가 현저히 증가한 최근에는 시청자의 시야에서 디스플레이 화면이 차지하는 비율이 대폭 증가하였다. 관심 영역을 사용한 인지품질 기반 비디오 코딩은 시청자의 시선이 닿는 지점, 즉 중심시(center vision)에서의 시각적 예민함이 시선이 닿지 않는 주변시(peripheral vision)에 비하여 높은 HVS의 특성을 기본 원리로 하고 있다. 따라서 이와 같은 부호화 방법은 디스플레이 크기에 큰 영향을 받게 되는데, 디스플레이 영역의 크기 증가로 인해 관심영역, 즉 중심시가 되는 영역의 크기가 상대적으로 줄어들고 주변시로 시청하게 되는 영역의 크기가 넓어지기 때문이다.

본 논문에서는, 기존의 CSF기반 및 관심영역 기반의 비디오 부호화 방법의 장점을 동시에 취하기 위해, 중심시에 대해서만 측정되었던 기존의 CSF모형을 주변시로 확장하여 망막 이심률(retinal eccentricity)의 변화에 따른 foveated contrast sensitivity(FCS)의 모델을 실험을 통해 측정한다. 측정된 FCS 모델은 제안하는 인지품질기반 비디오 부호화 방법에 적용되는데, 제안하는 부호화 방법에서 입력 영상

의 관심영역에는 FCS 모델에서 중심시에 대한 CSF를, 비관심 영역에는 주변시에 대한 CSF를 적용하여 부호화한다. 이때, 입력 영상으로부터 관심영역을 계산하기 위하여 Itti의 rapid scene analysis를 통해 계산된 saliency map이 사용된다<sup>[4]</sup>. CSF와 FM을 동시에 고려하는 기존의 알고리즘이 주로 FM효과를 고려한 modulation factor를 계산하여 DCT 계수 just-noticeable difference(JND)를 조정하는 반면, 제안하는 방법은 정신물리학 실험을 통하여 CSF와 FM효과가 동시에 고려된 foveated CSF모델을 적용하여 CSF와 FM 사이의 의존성을 JND 모델에 반영하였다. 변화를 제안하는 방법은 HEVC 참조 소프트웨어인 HM10.0에 구현하였으며, 부호화 실험 결과 제안하는 방법은 비교군인 HM10.0와 동일한 인지품질을 유지하면서 15% 낮은 비트율을 보였다.

본 논문은 다음과 같이 구성되어 있다. II장에서는 공간 주파수와 망막 이심률에 대한 각각의 기존 CSF 모델 및 제안하는 FCS 모델을 설명한다. III장에서는 제안하는 인지 품질 기반 부호화 방법을, IV장에서는 제안한 FCS모델을 이용한 HEVC 부호화 방법이 소개되며, V장에서는 부호화 실험 결과가 HM10.0과 비교된다. 마지막으로, VI장에서는 본 논문의 결론을 맺는다.

## II. Foveated Contrast Sensitivity 모델

본 논문에서 제안하는 FCS 모델은 기존의 공간주파수에 따른 CSF모델과 망막 이심률에 따른 공간주파수별 민감도의 변화를 모델링한 FCS모델과 관련이 있다. 공간주파수에 따른 CSF모델의 경우 HEVC의 DCT(discrete cosine transform) 커널에 대한 CSF에 대한 연구가 많아 실제 비디오 코딩 응용에서 더 정확하고 적용 가능한 연구 데이터가 존재하지만, 망막 이심률의 변화에 따른 공간 주파수별 민감도의 변화를 모델링 하지 않아 주변시에 대한 CSF를 알 수 없다는 문제가 있다. 반면 FCS모델의 경우 망막 이심률에 따른 공간 주파수별 민감도를 파악할 수 있지만 비디오 코딩의 DCT 커널에 적용한 연구가 이루어지지 않았고 실제 비디오 코딩 응용에 적용된 사례도 많지 않다. 따라서, 본 연구에서는 인지실험을 통하여 HEVC의 DCT 커널에 대해 공간

주파수 및 망막 이심률별 민감도를 측정하여 인지품질기반 비디오 코딩을 위한 새로운 FCS모델을 제시한다.

### 1. Contrast Sensitivity Function

비디오 코딩을 위한 CSF관련 연구는 DCT 커널의 공간 주파수에 따른 HVS의 민감도에 대한 측정 및 이를 이용한 인지품질 기반 비디오 압축 기술에 초점이 맞추어져 왔다. Zhang et al.의 연구에서는 CSF, LA(luminance adaptation) 및 CM(contrast masking)을 이용하여 이미지 코딩 블록의 JND(just-noticeable difference)를 계산하고, 이 값과 Saliency detection을 통해 계산한 해당 블록의 visual attention를 이용해 블록의 JND profile을 산출하여 이미지 부호화에 적용하였다<sup>[6]</sup>. 한편, Bae et al.의 연구에서는 DCT패턴을 이용하여 CSF 및 LA에 대한 JND 모델을 제안하였다<sup>[7]</sup>. 이 연구에서 주목할 일부 저주파 DCT 패턴이 DC 패턴보다 낮은 JND를 보인다는 점이다. 즉 JND는 주파수에 따른 단순 증가함수가 아니며, 특정 공간주파수 영역에서 가장 낮은 JND를 갖는다.

### 2. Foveated Contrast Sensitivity 모델

FCS는 foveation에 따른 HVS의 contrast sensitivity(CS)를 기술하는 모델이다. HVS의 안구 망막의 시세포는 중심와(fovea)와 그 주변에 집중되어 있기 때문에, 중심시, 즉 중심와에 맺히는 이미지가 가장 깨끗하게 지각되며, 상이 중심와에서 멀어질수록, 즉 망막 이심률이 증가할수록 이미지는 흐릿하게 지각된다. 즉 HVS의 sensitivity는 망막 이심률에 반비례한다. Geisler의 연구에서는 망막 이심률에 따른 HVS의 contrast sensitivity의 변화를 실험을 통해 측정하였는데, 그 결과는 다음과 같은 모델로 근사할 수 있다<sup>[3]</sup>.

$$\begin{aligned} CS(f,e) &= 1/CT(f,e) \\ CT(f,e) &= CT_0 e^{(af(e+e_2))/e_2} \end{aligned} \quad (1)$$

이 때,  $f$ 는 공간주파수,  $e$ 는 망막 이심률이며  $CT_0$ 와  $e_2$ 는 각각 spatial contrast decay 와 half resolution eccentricity 상수이다. 수식 (1)로부터 이 FCS모델에서는 HVS의 시각

적 sensitivity는 공간 주파수에 대한 순감소함수인 것을 확인할 수 있는데, sensitivity가 JND와 반비례하는 수치임을 감안 할 때 이것은 II-1에서 소개한 Bae의 CSF모델과 상충됨을 확인할 수 있다.

### 3. 제안하는 FCS모델

망막 이심률에 의한 contrast sensitivity를 고려한 DCT 커널의 CSF를 구하기 위해, 본 연구에서는 인지심리적 실험을 통하여 공간 주파수 및 망막 이심률에 대한 FCS 모델을 구하였다. 실험을 위하여 Samsung UN40D7000LF 40인치 FullHD 디스플레이가 사용되었고, 피실험자는 디스플레이 스크린으로부터 1.9m 떨어진 곳에 위치시켰다. 테스트 이미지는 디스플레이 상의 0.7m×0.41m 크기의 영역에 시험하였다. 시정거리는 DCT 커널의 최대주파수가 약

45cpd가 되도록 설정하였다. 실험 영상은 HEVC의 common test condition class B 영상들이 사용되었다.

FCS모델을 구하기 위한 실험은 다음과 같은 과정으로 진행되었다. 테스트 이미지의 임의의 지점에 조그만 점을 표시하고, 피실험자는 이 점에 초점을 맞추도록 하였다. 그 후, 이 점에서 특정한 거리만큼 떨어져 있는 위치에 있는 64×64이미지 블록에 특정 공간주파수 성분을 갖는 DCT 커널 패턴을 투사시킨다. 이 때, 화면상의 점과 DCT 패턴과의 거리는 망막 이심률이 0°, 1°, 2°, 3°인 지점을 시정거리 및 디스플레이 크기로부터 계산하여 선택한다(이 때, 0°인 지점은 화면의 점과 같은 위치이다). 64×64커널 블록은 실제 화면에서 가로세로 약 2cm정도의 크기를 가지며, 총 64×64개의 커널 중 균일한 수평, 수직 주파수 간격을 두고 25개의 커널이 선택되어 실험에 사용되었다.

JND(x,y)는 수평 및 수직 공간주파수가 각각 x, y인 DCT

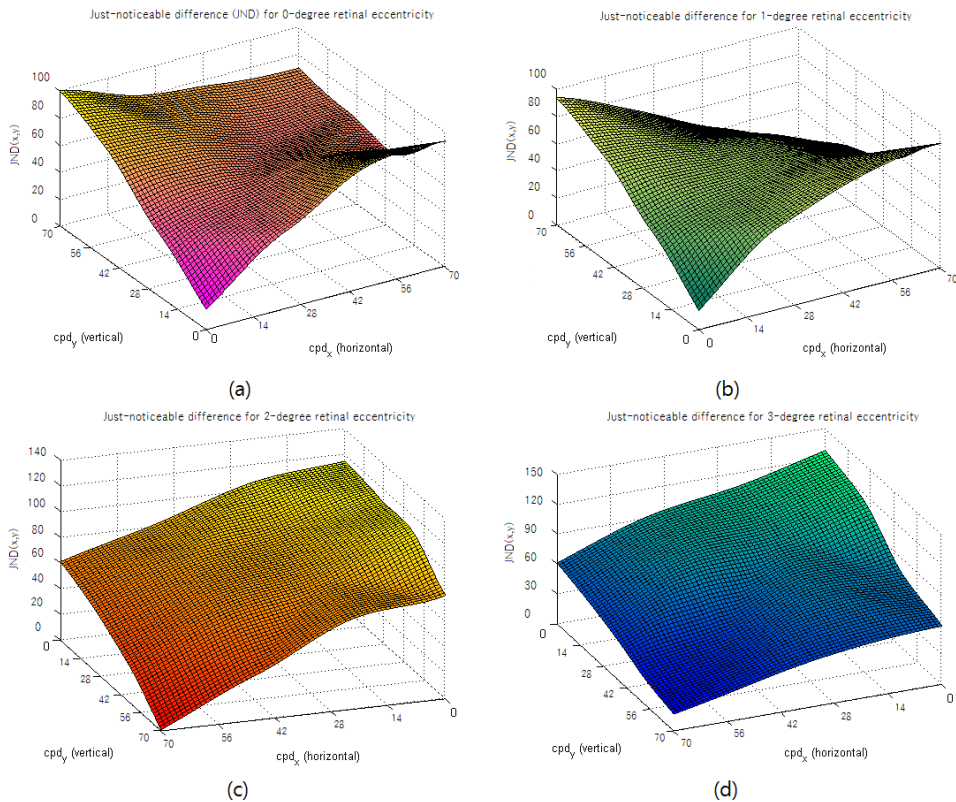


그림 1. 망막 이심률 (a) 0° (b) 1° (c) 2° (d) 3° 의 공간 주파수별 JND 그래프  
 Fig. 1. JND graphs in terms of the spatial frequency for retinal eccentricity of (a) 0° (b) 1° (c) 2° (d) 3°

커널에 대해 위의 실험을 진행하였을 때 50%이상의 피실험자가 변화를 인지할 수 있는 최소한의 가중치 값이다. 실험에 사용된 DCT 커널은 영상 내 64×64 블록에 공간 주파수 및 망막 이심률에 대해 이와 같은 실험을 통해 JND를 구하여 FCS 모델을 완성한다. 실험 데이터는 10명의 피실험자의 실험 결과의 평균치를 통하여 구하였다.

그림 1은 망막 이심률과 공간 주파수에 대한 FCS 모델을 나타낸다. 그림 (a), (b), (c), (d)는 각각 망막 이심률 0°, 1°, 2°, 3°에서의 FCS 모델을 나타내고, 각각의 모델에서  $cpd_x$ ,  $cpd_y$ (cycles per degree)는 각각 수평 및 수직 공간 주파수이며,  $JND(x,y)$ 는 해당 공간 주파수를 갖는 DCT 패턴에 대해 50%이상의 피실험자가 변화를 감지할 수 있는 최소한의 차이값이다.  $JND(x,y)$ 의 값은 해당 DCT 주파수 성분에 대한 피실험자의 시각적 민감도와 반비례한다. 그림 1(a)에서 볼 수 있듯이, 망막 이심률 0°에서 DCT 패턴에 대한 HVS의 민감도는 수직, 수평 단일 고주파 성분에서 가장 높았으며, DC 및 고주파 성분의 민감도는 이에 비해 다소 낮은 것으로 나타났다. 그림 1(b)는 망막 이심률이 1도 증가했을 때의 FCS 모델로, 이 때 FCS는 고주파 성분에 대한 민감도가 크게 감소한 것 이외에 망막 이심률 0°에서의 그것과 큰 차이를 보이지 않는 것을 확인할 수 있다. 이것은 주파수가 높을수록 망막 이심률에 의한 민감도의 감소폭이 커지는 Geisler의 FCS 모델과 상응하는 결과라고 볼 수 있다. 그림 1(c)와 1(d)는 망막 이심률 2° 및 3°에서의 FCS 모델을 보여 준다(x와 y축의 방향이 반대임을 주의하라). 두 모델에서 주목할 점은 DC성분에 대한 민감도는 0°와 1°의 FCS 모델과 큰 차이를 보이지 않는 반면, 고주파 성분에 대한 민감도는 크게 감소하였다는 점이다.

그림 2는 그림 1의 FCS 모델을 수평, 수직 및 대각선 방향 공간주파수 성분에 대해 표현한 그래프이다. 그림 2(a)와 그림 3(b)는 각각 수평, 수직 방향 주파수에 따른 망막 이심률 별 JND의 그래프를 나타내는데, 이 때 각각의 수직, 수평 방향 주파수 성분은 0이다. 그림에서 알 수 있듯이, 수직 방향과 수평방향 주파수에서 망막 이심률에 따른 JND의 변화는 유사하게 나타났다. 다만, 수직 방향에서 망막 이심률의 증가에 따른 고주파에서의 JND 감소율이 수평 방향에서의 그것에 비해 낮은 것으로 나타났는데, 이러한 결과

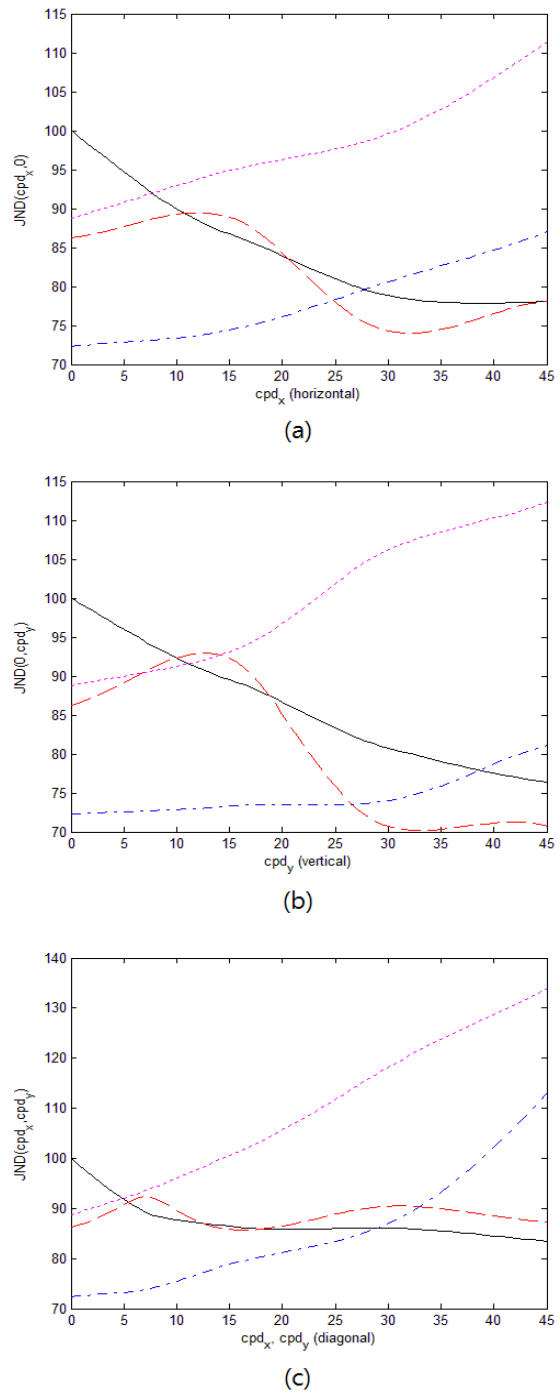


그림 2. 망막 이심률 0° (black), 1° (red), 2° (blue) 3° (magenta)에서 (a) 수평 (b)수직 (c) 대각 공간 주파수에 따른 JND 그래프  
 Fig. 2. JND graphs in terms of (a) horizontal (b) vertical (c) diagonal frequency component for retinal eccentricity of 0°, 1°, 2°, 3°.

로부터 HVS는 고주파 성분에 대해 수평 방향보다 수직 방향에서 망막 이심률 증가에 따른 민감도 감소가 적은 것으로 미루어 볼 수 있다. 그림 2(c)는  $cpd_x = cpd_y$  일 때 대각 주파수 성분에 대한 망막 이심률 별 JND를 나타낸 그래프이다. 이 그래프는 고주파 부분에서 그림 2(a), 2(b)의 그래프와 확연히 차이가 나는 것을 확인할 수 있는데, 이것으로부터 대각 방향 고주파 성분에 대한 HVS의 민감도는 망막 이심률이 1°이상만 되어도 큰 폭으로 떨어진다는 것을 확인할 수 있다.

제안하는 FCS 모델은 망막 이심률이 0°인 경우 기존의 CSF와 유사한 모델을, 망막 이심률이 증가하였을 경우 0°의 CSF에서 망막 이심률의 증가에 따라 주파수가 높은 성분부터 민감도가 급격히 떨어지는 것을 관찰할 수 있었다. 이렇게 구한 FCS모델은 인지품질 기반 비디오 부호화기를 설계하는데 있어 변환코딩 직후 변환 계수들에 대한 억제를 위한 JND로 사용할 수 있다.

### III. 제안하는 인지품질 기반 비디오 부호화 방법

제안하는 인지품질 기반 비디오 부호화는 2장에서 소개한 제안하는 FCS모델을 기반으로 구현하였다. 2장에서 구한 FCS모델의 망막 이심률 및 DCT 커널의 주파수 별 변환

계수에 대한 JND는 HEVC의 변환계수 코딩 과정에 그대로 적용 가능하기 때문에, 제안하는 부호화 방법에서는 HEVC 부호화기의 변환 코딩 후 양자화 직전 JND 모델을 통하여 변환계수를 억제하는 coefficient suppression 방법을 적용하였다. 제안하는 방법은 부호화기에만 적용되기 때문에, HEVC표준에 부합하는 방법이다.

그림 3은 제안하는 인지품질 기반 부호화기의 블록도를 나타낸다. 제안하는 방법은 먼저 부호화기의 입력인 비디오 프레임으로부터 Ittt et. al의 rapid scene analysis 방법을 이용하여 saliency map을 계산한다<sup>[14]</sup>. 계산된 saliency map은 이후 해당 프레임에 대한 변환계수를 억제(coefficient suppression)하는데 사용되는데, 이 때 saliency map에서 변환 계수에 해당하는 블록의 saliency 값을 이용하여 해당 블록에서의 망막 이심률의 추정치를 구하고, 이 추정치와 변환 계수의 주파수를 통해 해당 변환 계수에 어떤 JND모델을 적용할지 계산하고, 선택된 JND모델을 이용해 계수를 억제한다.

$$JND = \begin{cases} JND_3 & 0 \leq saliency < 64 \\ JND_2 & 64 \leq saliency < 128 \\ JND_1 & 128 \leq saliency < 196 \\ JND_0 & 197 \leq saliency \leq 255 \end{cases} \quad (2)$$

(2)는 이미지의 각 변환계수 블록에 대해 적용할 JND 모

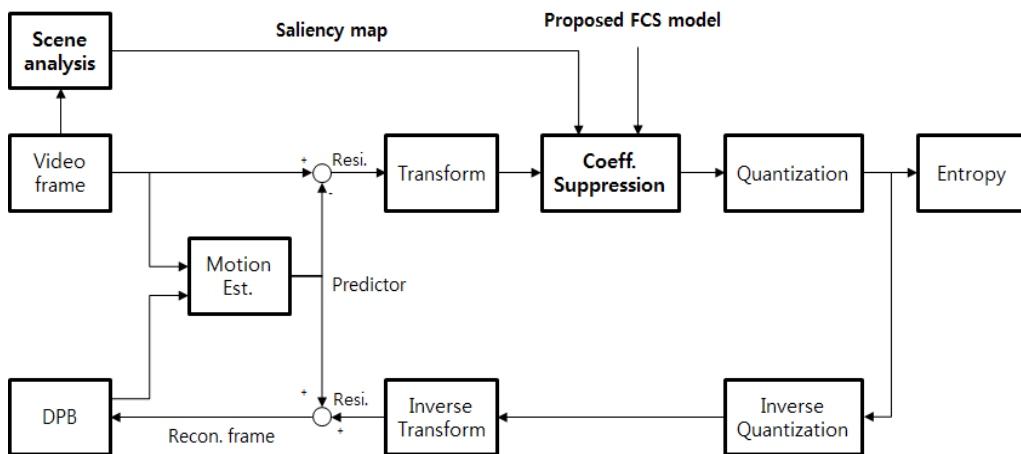


그림 3. 제안하는 인지품질 기반 비디오 부호화기  
 Fig. 3. The proposed perceptual quality-based video encoder

델을 구하는 식이다. 이 때, JND는 그림 1의  $i^o$  망막 이심률에 대한 JND 모델을 나타내며, saliency는 해당 변환 블록의 saliency값의 평균이다. 변환 블록의 saliency값이 높을수록 시청자의 시점이 해당 블록에 머물 확률이 높기 때문에, 제안하는 방법에서는 더 높은 saliency를 갖는 블록에 대하여 더 낮은 망막 이심률에 대한 JND 모델을 적용하였다.

$$F(x,y) = \begin{cases} F(x,y) & F(x,y) > JND(x,y) \\ 0 & otherwise \end{cases} \quad (3)$$

식 (3)은 식 (2)를 통해 선택된 JND 모델을 통하여 해당 블록의 변환 계수를 억제하는 알고리즘을 나타낸다. 이 때  $F(x,y)$ 는 원본 변환 계수이며  $F(x,y)$ 는 JND 모델을 통하여 억제된 변환 계수이다.

제안하는 인지품질 기반 부호화는 FM효과가 반영된 JND모델을 사용하고 있으며, FM효과는 영상의 saliency에 크게 의존적인 특징을 지니고 있다. 따라서, 제안하는 인지품질 기반 부호화기는 saliency 검출 알고리즘의 성능에 의해 그 성능이 크게 좌우될 수 있다.

#### IV. 인지품질 기반 비디오 부호화 실험 및 결과

제안하는 방법은 HEVC 참조 소프트웨어인 HM10.0상에 구현되었다. 실험에서 HM10.0과 제안하는 방법이 비교되었으며, 실험 결과에 대한 성능 평가를 위해 객관적 척도로 RD (rate distortion) 그래프 및 BDbitrate가 비교되었고, 주관적 성능의 척도로 MOS(mean opinion score)를 사용하였다. 실험에서는 각각 HEVC Class A, B, C, D, E영상을 CTC(common test condition) All-Intra, QP 22, 27, 32, 37로 부호화하였고, 그림 3의 인지품질 기반 부호화 방법과 같이 프레임의 모든 TU(transform unit)에 대하여 양자화 전 DCT 계수에 주파수 및 망막 이심률에 따라 그림 1의 JND 모델을 적용하여 계수를 억제하였다. 이 때, 현재 부호화 TU의 망막 이심률은 III장에서 언급하였듯이 Itti의 방법을 사용하여 구한 현재 부호화 프레임의 saliency map을 통하여 계산된다. 변환계수는 제안하는 JND모델에 의해 억

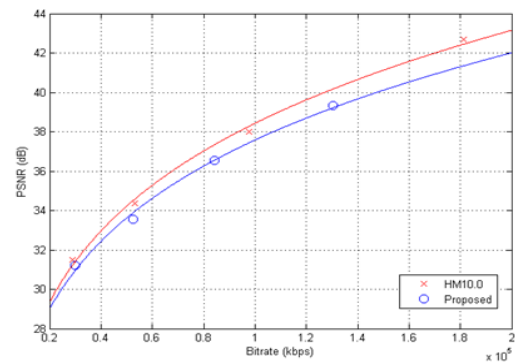
제되므로, 제안하는 방법을 통해 부호화된 영상의 비트율은 기존 HM10.0의 그것에 비해 항상 낮다.

표 1은 제안하는 방법의 HM10.0 대비 BD-rate를 나타낸다. 제안하는 방법은 HM대비 평균 8%의 BD-rate 손실을

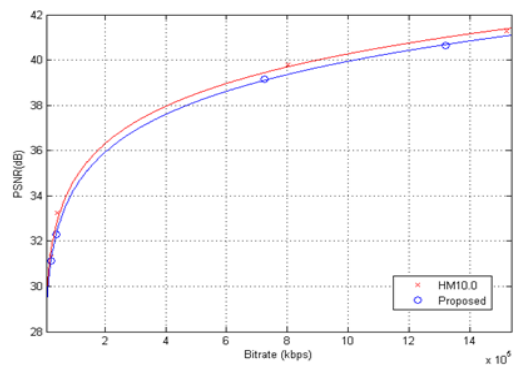
표 4. 제안한 인지품질 기반 부호화 방법의 HM10.0 대비 BD-rate 및 부호화 시간

Table 4. BD-rate and encoding time of the proposed PQ-based encoding algorithm compared to HM10.0 anchor

	Y	U	V
Class A	4.5%	6.2%	6.1%
Class B	6.1%	8.1%	8.1%
Class C	10.2%	16.7%	16.6%
Class D	11.2%	19.3%	19.1%
Class E	7.8%	10.1%	10.1%
Average	8%	12.1%	12%
Encoding Time	102%		



(a)



(b)

그림 4. (a) BasketballDrill (b) BQTerrace 영상의 부호화 실험 결과 RD 그래프

Fig. 4. RD graphs of encoding of (a) BasketballDrill (b) BQTerrace



갖는 것으로 나타났으며, Class A 영상에서 4.5%로 손실률이 가장 낮았으며 Class D에서 11.2%으로 가장 높은 손실을 보였다. 그림 4는 Class C 일부 영상의 RD 나타낸다. 그림 4(a)와 그림 4(b)는 각각 BasketballDrill과 BQTerrace 영상에 대한 RD그래프이다. 제안하는 방법의 비트율이 HM에 비하여 평균 10%가량 낮은 반면, 계수 억제로 인한 영상 왜곡의 증가로 인해 BD-rate는 제안하는 방법이 오히려 10% 정도 낮은 것으로 나타났다. 그러나, 인지품질 기반 영상 부호화에서 객관적 품질 평가 방법인 PSNR(peak-signal to noise ratio)는 성능 평가에 있어 중요한 척도가 아니다.

제안하는 방법으로 부호화 한 영상에 대한 주관적 화질 평가는 그림 4의 Class C 두 영상을 이용하여 수행하였다.

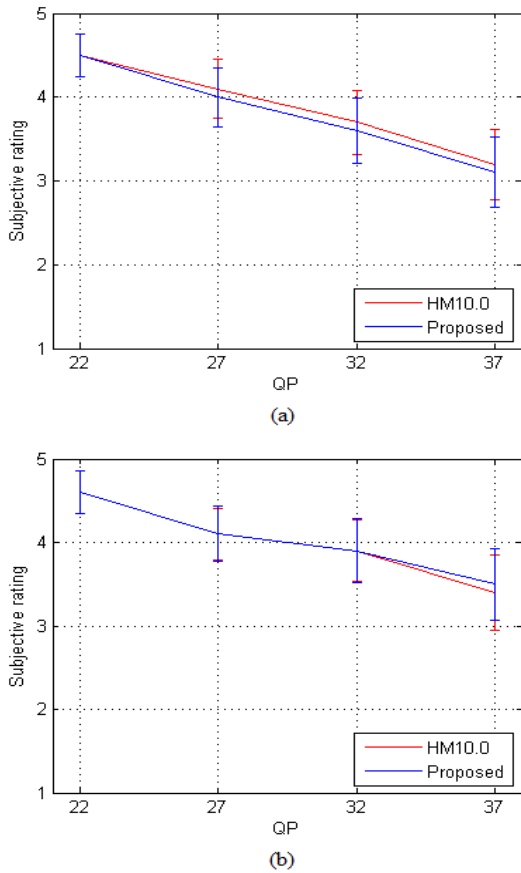


그림 5. HM10.0 및 제안하는 방법으로 부호화된 (a) BasketballDrill (b) BQTerrace 영상의 주관적 화질 평가 결과  
 Fig. 5. Subjective quality test result of (a) BasketballDrill (b) BQTerrace encoded with HM10.0 and the proposed method

그림 5는 제안하는 알고리즘을 통해 부호화한 영상에 대한 주관적 화질 평가 결과를 나타낸다. 주관적 화질 평가는 10명의 피 실험자들로부터 ITU-R BT.500의 SS(single-stimulus)방법을 통해 MOS를 측정하는 식으로 진행되었다. 그림 5(a)와 그림 5(b)는 각각 BasketballDrill과 BQTerrace 영상에 대한 MOS그래프를 나타내며, 그래프의 에러 바는 MOS의 1 표준편차 범위를 나타낸다. 그림 5의 두 그래프의 HM10.0과 제안하는 방법의 MOS를 비교하였을 때, BasketballDrill의 경우 제안하는 방법의 MOS가 HM10.0보다 최대 0.1 낮으며, BQTerrace의 경우 QP37일 때 HM10.0보다 0.1 높은 것을 제외하고 동일한 것으로 나타났다.

그림 6은 각 실험 영상에 대한 제안하는 방법의 비트율의 감소비율을 나타낸다. 비트율의 감소율은 제안하는 방법으로 부호화한 영상의 비트율을 동일 영상 및 QP에 대해 HM10.0으로 부호화한 영상의 비트율로 나누어 계산하였다. 그림에서 확인할 수 있듯 비트율의 감소는 Basketball Drill영상이 BQTerrace영상보다 높게 나타났으며, 높은 QP에서보다는 낮은 QP에서 비트 감소율이 더 높은 것으로 나타났다. 흥미로운 것은 BasketballDrill과 BQTerrace영상이 낮은 QP에서 비트 감소율이 각각 27%, 15%로 비교적 높은 반면 이에 따른 주관적 화질, 즉 MOS의 감소율은 그림 5에서 볼 수 있듯 두 영상 모두 0인 것으로 나타났다. QP27에서도 이와 비슷한 현상을 확인할 수 있는데, 반면 상대적으로

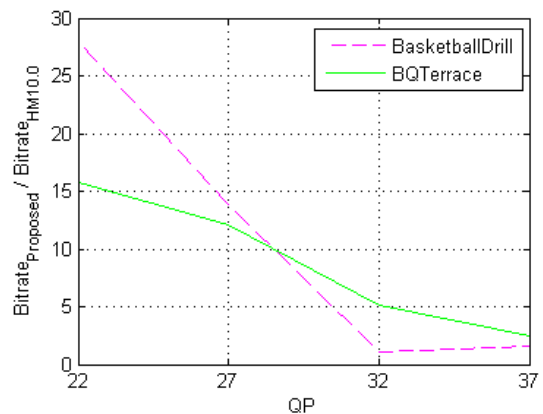


그림 6. 제안하는 방법의 QP별 비트율의 감소비율 그래프  
 Fig. 6. Bitsavings of the proposed method with different QPs



로 높은 QP인 32, 37에서는 비트율의 감소가 상대적으로 낮으면서 주관적 화질의 감소도 더 높은 것으로 나타났다. 따라서, 제안하는 방법은 낮은 QP에서 더 높은 성능을 보이는 것으로 해석할 수 있다.

## V. 결론

본 논문에서는 인지심리적 실험을 통하여 망막 이심률과 공간주파수에 따른 DCT 커널의 새로운 FCS 모델을 구하고, 이 모델에 기반한 인지품질 기반 HEVC 부호화 방법을 제안하였다. 제안하는 FCS 모델은 0°, 1°, 2°, 3°의 망막 이심률에서 수평, 수직의 공간주파수에 대하여, DCT 변환계수에 곱해졌을 경우 시청자가 변화를 인지하지 못하는 최소값인 just noticeable 변환 계수 JND를 구한다. 제안하는 FCS 모델은 HM10.0 부호화기에 적용되었으며, 이 모델이 적용된 제안하는 부호화기는 변환 계수에 대해 FCS모델 기반 계수 억제를 통하여 인지적 무손실 압축을 수행한다. 실험 결과, 제안하는 부호화기는 약 10%의 BD-rate 손실을 가지나 주관적 화질 측면에서 볼 때 거의 동일한 주관적 화질로 비트율을 평균 15% 낮추는 것으로 나타났다.

## 참고 문헌 (References)

[1] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," Vol. 22, No. 12, pp. 1649-1668, Dec. 2012

[2] H. Oh, W. Kim, "Video processing for human perceptual visual quality-oriented video coding," IEEE Transactions on Image processing, Vol. 22, No. 4, pp. 1526-1535, Dec. 2012

[3] Yang Xia, RuiMin Hu, Zhongyuan Wang, "Perceptual video compression based on DCTdomain foveated JND model," 2011 4th International Congress on Image and Signal Processing, Shanghai,

China, 2011 Oct.

[4] H. Oh, W. Kim, "Video Processing for Human Perceptual Visual Quality-Oriented Video Coding," IEEE Transactions on Image Processing, Vol. 22, Issue 4, pp. 1526-1535, April 2013

[5] H. H. Chen, Y.-H. Huang, P.-Y. Su, T.-S. Ou, "Improving video coding quality by perceptual rate-distortion optimization," 2010 IEEE International Conference on Multimedia and Expo (ICME), Suntec City, Singapore, pp. 1287-1292, July 2010

[6] D. Zhang, L. Gao, D. Zang, Y. Sun, "A DCT-domain JND model based on visual attention for image," 2013 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Melaka, Malaysia, pp. 1-4, Oct. 2013

[7] S. H. Bae and M. Kim, "A Novel DCT-based JND Model for Luminance Adaptation Effect in DCT Frequency," IEEE Signal Processing Letters, vol. 20, no. 9, pp. 893-896, Sept. 2013

[8] Z. Wei, Ngan, K.N., "A temporal just-noticeable distortion profile for video in DCT domain," 15th IEEE International Conference on Image Processing, San Diego, CA, pp. 1336-1339, Oct. 2008

[9] M. J. Nadenau, J. Reichel, M. Kunt, "Wavelet-based color image compression: exploiting the contrast sensitivity function," IEEE Transactions on Image Processing, Vol. 12, No. 1, pp. 58-70, Jan. 2003

[10] M. G. Albanesi, M. Ferretti, F. Guerrini, "Adaptive image compression based on regions of interest and a modified contrast sensitivity function," 15th International Conference on Pattern Recognition, Barcelona, Spain, Vol. 3, pp. 215-218, Sep. 2000

[11] H. Hadizadeh, I.V., Bajic, "Saliency-preserving video compression," IEEE International Conference on Multimedia and Expo (ICME), Barcelona, Spain, pp. 1-6, July 2011

[12] D. Grois, E. Kaminsky, O. Hadar, "Dynamically adjustable and scalable ROI video coding," IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Shanghai, China, pp. 1-5, March 2010

[13] W. S. Geisler, and J. S. Perry, "A real-time foveated multiresolution system for low-bandwidth video communication," Proceedings of SPIE, vol. 3299, 1998.

[14] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp. 1254-1259, Nov. 1998

[15] Grois, D., Kaminsky, E., Hadar, O., "Dynamically adjustable and scalable ROI video coding," 2010 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Shanghai, China, pp. 1-5, March 2010

---

저 자 소 개

---



유 지 우

- 2011년 : 광운대학교 컴퓨터공학과 (학사)
- 2013년 ~ 현재 : 광운대학교 컴퓨터공학과 (석사)
- 주관심분야 : 영상처리, 비디오 코덱



심 동 규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 바로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 부교수
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전