

# H.264 스케일러블 확장을 위한 시각적 가중치 기반 비트 할당 알고리즘

전선국<sup>†</sup>, 하호진<sup>\*\*</sup>

## 요 약

본 논문은 매크로 블록 레벨에서 인간의 시각적인 특성을 고려한 새로운 H.264 스케일러블 확장을 위한 비트 할당 알고리즘을 제안한다. 제안된 알고리즘은 2 단계로 구성된다. 우선, 인간의 시각적인 특성에 기초한 관심영역을 중심으로, 각 매크로 블록에 대한 공간 주파수(spatial frequency)의 분석을 위한 시각적 가중치 모델을 제안한다. 두 번째 단계에서는, 제안된 모델을 기초로 각각의 화질 계층에 존재하는 관심영역의 크기를 조절하여 주어진 비트 량에서, 최적의 비디오 화질을 얻기 위한 비트 할당 알고리즘을 제안한다. 실험을 통하여 제안된 알고리즘은 주관적 및 객관적 화질 평가에서 기존의 알고리즘보다 향상된 결과를 보여주었다.

## A Visual Weighting-Based Bit Allocation Algorithm for H.264 Scalable Extension(SE)

Shan Guo Quan<sup>†</sup>, Hojin Ha<sup>\*\*</sup>

## ABSTRACT

This paper proposes a novel bit allocation algorithm for H.264 scalable extension(SE) based on a human visual system (HVS) to improve the coding efficiency. The proposed algorithm is consist of two stages: visual weighting model and visual weighting-based bit allocation algorithm. In the first stage, the visual weighting for each macroblock (MB) is analyzed according to the region of interests. Then the adaptation of the visual weighting into the bit allocation routine for each quality layer is performed for improving the visual quality. In the simulation results, it is observed that the proposed scheme can improve the subjective and objective video quality in the same bit rate, compared to the previous scalable video coding in H.264.

**Key words:** Bit Allocation(비트 할당), Scalable Video Coding(스케일러블 비디오 코딩), Visual Weighting(시각적 가중치), H.264

## 1. 서 론

최근에 멀티미디어 통신 기술의 발전과 함께 다양한 통신환경에서 많은 양의 비디오 데이터를 효과적으로 전송하기 위한 연구들이 활발히 진행되고 있다.

특히 H.261, H.264, MPEG-1, MPEG-2 그리고 MPEG-4 같은 표준화된 압축 기술에 더불어, 다양한 전송채널의 대역폭에 적응하면서 다양한 화질을 제공할 수 있는 스케일러블 비디오 코딩 등에 대한 표준화 연구가 활발히 진행되고 있다[1-4].

※ 교신저자(Corresponding Author): 하호진, 주소: 경기도 수원시 영통구 매탄 3동 416 삼성전자 DMC 연구소(443-742), 전화: 031)279-7970, E-mail: hojini@samsung.com

접수일: 2010년 11월 3일, 수정일: 2011년 3월 3일  
완료일: 2011년 4월 11일

<sup>†</sup> 종신회원, 남서울대학교 전자공학과  
(E-mail: sgquan@nsu.ac.kr)

<sup>\*\*</sup> 정회원, 삼성전자 DMC 연구소

※ 본 연구는 남서울대학교 교내공모(자유)연구(2010-자유-9)지원으로 수행되었음.

비디오 코딩 기술의 궁극적인 목적은 낮은 비트 량으로 신뢰성 있는 비디오 화질을 어떻게 제공할 것인가에 초점을 맞추고 있다. 일반적으로 기존의 비디오 코딩 기술은 원본 영상과 복원된 영상의 mean square error (MSE)를 최소화하도록 설계된다[1-2]. 하지만, MSE에 기초한 객관적인 평가는 주관적인 화질평가와 비교했을 때, 실제로 화질과 연관성이 적음을 보여주고 있다[5-6].

따라서, 사용자들은 비디오 화질평가를 객관적인 판단보다는 주관적인 판단에 많이 의존하게 되었고, 이러한 인간의 시각적인 특성을 비디오 코딩 기술에 최대한 반영하려는 연구들이 기존에 많이 진행되어 왔다. 특히 인간의 시각적인 특성은 관심영역과의 거리에 의해서 민감도의 차이를 갖기 때문에, eye-tracker, 색깔, edge 등의 정보를 이용하여 관심영역을 찾는 연구들이 진행되어 왔다[7-9]. 모션의 크기 및 블록에 대한 정보들을 이용하여 시각적으로 민감도가 높은 영역을 찾기 위한 연구도 진행되었다[10]. 그리고 홍채에 존재하는 시신경의 비정형적인 분포를 이용하여, 시각적인 중요도의 레벨을 결정하는 포비에이션 (foveation)에 관한 연구가 진행되어 왔다 [12-13].

그리고 앞에서 설명한 관심영역에 따른 공간 주파수의 민감도를 비디오 압축 기술들에 적용하는 많은 연구가 진행되고 있다[10-12]. 시각적인 마스킹 레벨에 따라 양자화 파라미터를 조절하는 알고리즘이 [11]에서 제안되었고, 웨이블릿에 기초한 임베디드 압축 기술을 인간의 시신경 분포를 이용한 시각적인 마스킹 레벨에 적용하여, 낮은 비트 량에서도 효율적으로 화질을 개선하는 연구들도 진행되었다[12-13].

본 논문에서는 H.264의 스케일러블 확장을 위해서, 인간의 시각적인 특성을 이용하는 새로운 비트 할당 알고리즘을 제안한다. 제안된 알고리즘은 두 단계로 구성된다. 인간의 시각적인 특성을 MB 레벨에서 분석하기 위한 시각적 가중치 모델을 제안하는 단계와 분석된 가중치 모델을 각각의 화질계층에 적용하여 비트를 할당하는 단계로 구성된다.

먼저, 시각적 가중치 모델은 관심영역을 중심으로 각 MB의 위치에 따른 공간 주파수(spatial frequency)의 민감도를 계산하는 단계이다. 관심영역에서 가까울수록 높은 민감도로 갖기 때문에, 화질 왜곡에 대한 민감도가 증가한다. 반면에, 관심영역에서

멀어질수록 공간 주파수에 낮은 민감도를 갖는다 [7-13]. 관심영역에 기초한 공간 주파수에 대한 계산은 포비에이션 알고리즘을 이용한다.

두 번째 단계에서는, 각각의 화질 계층에 존재하는 관심 영역의 크기를 제안된 시각적 가중치 모델을 이용하여 조절함으로써, 주어진 인코딩 비트 량에 최적의 비디오 화질을 얻기 위한 비트 할당 알고리즘이 제안된다. 낮은 화질 계층은 최소한의 관심영역을 유지함으로써, 적은 비트를 효율적으로 사용하여 화질을 향상시킨다. 반면에 높은 화질계층은 많은 인코딩 비트를 가지고 있기 때문에, 관심영역을 최대한 확장함으로써, 전체적인 화질을 향상시킨다. 관심영역에 의한 시각적 특성을 비트할당 알고리즘에 반영하여 위해서, 각 MB내의 DCT 계수를 필터링하는 방법을 이용한다.

결국, 낮은 공간 주파수를 갖는 MB는 작은 시각적 민감도를 갖기 때문에 적은 수의 DCT 변환 계수를 유지한다. 반면, 높은 공간 주파수를 갖는 MB는 큰 시각적 민감도를 갖기 때문에, 최대한 많은 수의 DCT 변환 계수를 갖도록 함으로써, 주파수의 민감도를 반영한다. 이러한 시각적 가중치 모델을 이용한 제안된 비트 할당 알고리즘은 상대적으로 비트 량이 적은 기초 계층에서 많은 화질 개선의 효과가 있음을 알 수 있었다.

본 논문의 구성은 다음과 같다. II장에서는 시스템 모델에 대해 간단히 설명하고, III장에서는 제안된 시각적 가중치 분석 모델을 설명한다. 제안된 모델을 기초로, H.264 SE의 시각적 가중치를 이용한 비트 할당 알고리즘이 IV장에서 제안된다. 실험 결과가 V장에서 보이고, VI장에서 결론을 제시한다.

## 2. 시스템 모델

기존의 H.264 SE는 각 화질 계층에 존재하는 MB들의 시각적 주파수에 대한 민감도의 차이를 고려하지 않는다. 그러나 한 프레임내의 영상은 제한된 비트 량에서, 다양한 복잡도를 갖고, 관심영역에 기초하여 다양한 주파수의 민감도를 가지기 때문에, 일정한 시각적 중요도를 각 MB에 적용하는 것은 비디오 코딩의 화질 저하를 가져오게 된다[10-13]. 본 논문에서는 제안된 시각적 가중치 모델을 이용하여, MB내의 시각적 중요도를 선택적으로 할당함으로써, 제

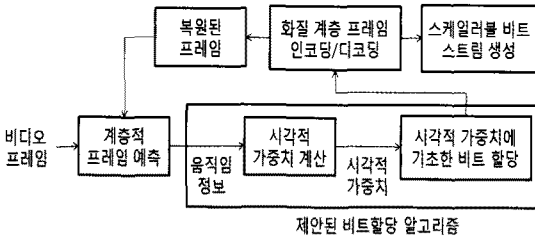


그림 1. 기존 H.264 SE 프레임상의 비트 할당 알고리즘의 전체 구조

한된 인코딩 비트에서 화질을 향상시키는 비트 할당 알고리즘을 제안한다. 그림 1은 제안된 시각적 가중치 모델을 이용한 스케일러블 비디오 비트 할당 알고리즘의 전체 구조를 나타내고 있다. 시각적 가중치 모델은 계층적 예측 모델을 이용하여 모션의 크기를 이용하여 관심영역을 예측한다[3-4]. 예측된 영역과 인간의 시각적 모델에 근거한 포비에이션 알고리즘을 이용해서, 시각적 공간 주파수를 계산한다. 예측된 관심영역을 중심으로 시각적 공간 주파수는 지수승으로 감소하는 형태를 갖게 된다.

각 MB의 DCT 계수와 계산된 공간 주파수를 이용하여, 제안된 시각적 스케일러블 알고리즘은 각 화질 계층의 관심영역의 크기를 조절함으로써, 관심영역을 중심으로 화질이 점차적으로 향상되는 결과를 갖게 된다. 화질 계층이 높아지면, 관심영역을 확대하고, 화질 계층이 낮아지면, 관심영역을 최소한으로 가져감으로써, 제한된 인코딩 비트를 효율적인 사용하게 된다.

### 3. 시각적 가중치 모델

제안된 시각적 가중치 모델은 2단계로 구성된다.

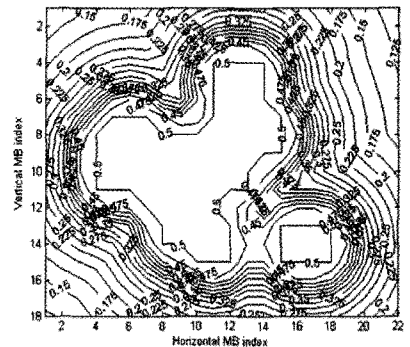


그림 2. 시각적 가중치 모델의 결과 (a) 원본 영상 (b)  $I_k(i,j)$ 의결과 (c) 0.5로 정규화된 각 MB의 공간 주파수

우선, 모션정보를 이용하여 관심영역을 선정한 후, 선정된 관심영역과 포비에이션 알고리즘을 바탕으로 각 MB에 대한 공간 주파수를 계산한다.

첫 번째로, 관심영역에 대한 예측은 다음과 같은 가정을 이용한다. 인간의 시각적 특성은 정지되어 있는 배경보다 움직이는 물체에 더 많은 주의를 기울인다[11-12]. 결국, 인간은 움직이고 있는 물체에 더 많은 관심을 갖고, 이 영역의 화질 감소에 더 민감하다. 반면에 정지되어 있는 배경에 대한 주파수의 민감도는 감소하게 된다. 본 논문에서는 각 프레임의 관심영역을 설정하기 위해서, 각 프레임 내의 MB 모션정보를 이용한다. k번째 프레임의 (i, j)번째 MB에 대한 모션의 크기는 다음과 같이 정의 된다.

$$I_k(i,j) = \sqrt{MVX_k(i,j)^2 + MVY_k(i,j)^2} \quad (1)$$

여기에서  $MVX_k(i,j)$ 와  $MVY_k(i,j)$ 는 (i, j)번째 MB의 수평과 수직의 모션 크기를 나타낸다. 만약  $I_k(i,j)$ 가 0보다 크고  $I_{MAX}$ 보다 작으면, 우리는 현재의 MB를 관심영역으로 간주한다. 0보다 크다는 의미는 물체가 움직인다는 의미이고, 인간의 시각적 특성은 움직임이 있는 물체에 더 민감하게 된다. 또한  $I_{MAX}$ 를 사용한 이유는 움직임이 너무 큰 물체들은 오히려 주파수에 대한 민감도를 감소시키기 때문이다[7-11]. 각 MB들은 모션의 크기를 이용하여, 관심영역과 비 관심영역으로 구분된다.

그림 2의 (b)는 SOCCER 테스트 영상에서 24번째 프레임의 모션 크기( $I_k(i,j)$ )를 이용해서 관심영역으로 선정된 MB를 표시한 결과이다.

두 번째 단계에서는, 설정된 관심영역을 중심으로, 각 MB들에 시각적 특성을 고려한 가중치를 할당한다. 각 매크로 블록의 모션 정보를 이용하여, 선정

된 포비에이션 포인트를 중심으로, 주위에 존재하는 매크로 블록의 시각적 가중치가 계산된다[13]. 여기에서 포비에이션 포인트는 인간의 눈이 가장 높은 공간 주파수를 갖는 영역으로 정의된다[12-13]. 포비에이션 포인트에서 가장 높은 주파수의 민감도를 갖고, 그 점에서 멀어질수록 민감도는 지수 승으로 감소하게 된다. 본 논문에서는 관심영역 내의 MB를 포비에이션 포인트로 간주하고, 가장 큰 시각적 가중치를 할당한다. 포비에이션 포인트의 거리에 비례해서, 각 MB에 공간 주파수를 할당하는 방법을 사용하여 가중치를 할당하게 된다[10-13].

이미지 프레임에는 관심영역에 존재하는 포비에이션 포인트,  $\vec{p}$ 와 비 관심영역에 존재하는 어느 한 점,  $\vec{p}'$ 가 존재한다.  $\vec{p}'$ 와 눈과의 거리 그리고  $\vec{p}$ 와  $\vec{p}'$ 의 거리를 각각  $\alpha$ 와  $\beta$ 로 정의 하면, 시각적 이심율(eccentricity)을 아래와 같이 계산할 수 있다[13].

$$e(\alpha, \vec{p}') = \tan^{-1} \left( \frac{\beta}{\alpha} \right) \quad (2)$$

여기에서  $\alpha$ 는  $N$ (영상의 가로 픽셀 값)과  $v$ 의 곱으로 정의되고,  $v$ 는 3으로 설정한다. 또한 비 관심영역  $\vec{p}$ 와 눈과의 거리는  $\sqrt{\alpha^2 + \beta^2}$ 를 이용해서 계산 될 수 있다. 시각적 이심율로부터, 최소한의 시각적 명암을 분별할 수 있는 공간적 컷 오프 주파수,  $f_c(e(\alpha, \vec{p}'))$ 는 아래와 같이 정의 한다.

$$f_c(e(\alpha, \vec{p}')) = \frac{e_2 \ln \left( \frac{1}{CT_0} \right)}{\alpha(e(\alpha, \vec{p}') + e_2)} \quad (3)$$

여기에서,  $CT_0$ 을 최소 명암 임계치 (minimal contrast threshold),  $e_2$ 는 반 해상도 이심율 상수 (half resolution eccentricity constant),  $\alpha$ 는 공간 주파수 감쇠 상수 (spatial frequency decay constant), 최소 대비 임계치(contrast threshold)이다. 또한, 실제적인 디지털 영상에서 시각적 해상도 (visual resolution),  $r$ 은 가시거리  $v$ 와 영상의 해상도  $N$ 으로부터 다음과 같이 유도되어 진다[13].

$$r \approx Nv \frac{\pi}{180} \quad (4)$$

Nyquist 주파수로부터, 가장 높은 시각적 해상도는 다음과 같이 유도되어 진다[13].

$$f_d = \frac{r}{2} = Nv \frac{\pi}{360} \quad (5)$$

식(3)과 (5)로부터, (i, j)번째 MB의 공간 주파수 (spatial frequency),  $f_s^{i,j}(v)$ 은 다음과 같이 유도되어 진다.

$$f_s^{i,j}(v) = \min_{\vec{p} \in (i,j)^{\#}MB} (avg(f_c(e, \vec{p})), f_d(v)) \quad (6)$$

여기에서,  $avg(f_c(e, \vec{p}))$ 은 (i, j)번째에 포함되는 픽셀,  $\vec{p}$ 에 대한 공간적 컷 오프 주파수의 평균값이고  $v$ 는 눈과 영상사이의 거리를 나타낸다. 그림 2. (c)는 0.5로 정규화된 각 MB들의 공간 주파수를 보여주고 있고, 관심영역을 중심으로 점차적으로 감소되는 공간적 주파수의 값을 볼 수 있다.  $f_s^{i,j}(v)$ 는  $v$ 를 일정하게 고정시킴으로서,  $f_s^{i,j}$ 으로 나타낼 수 있다.

#### 4. H.264 SE를 위한 시각적 가중치 기반 비트 할당 알고리즘

시각적 스케일러블 비디오 코딩 알고리즘은 2 단계로 구성되어 있다. 첫 번째 단계는 화질 계층으로부터 관심영역의 경계를 설정하고, 공간 주파수의 가중치를 각 MB마다 할당한다. 두 번째는 결정된 공간 주파수의 가중치를 이용하여, 각 MB내의 DCT 계수를 필터링하는 과정으로 구성된다. 제안된 스케일러블을 위한 비트 할당 알고리즘을 각 화질 계층의 4x4 DCT 도메인에서 제안되기 때문에, 각 MB에 대해서 연속적인 공간 주파수의 가중치,  $f_s^{i,j}$ 은 이산적인 4x4 DCT 도메인으로 양자화 된다.  $f_s^{i,j}$ 로부터 양자화된 공간 주파수를  $f_q^{i,j}$ 라고 하면,  $f_q^{i,j}$ 은 다음과 같이 계산된다.

$$\begin{aligned} \tilde{x}, \tilde{y} &= \operatorname{argmin}_{x,y} |f_N(x,y) - f_s^{i,j}| \\ f_q^{i,j} &= f_N(\tilde{x}, \tilde{y}) \end{aligned} \quad (7)$$

여기에서  $f_N(x,y)$ 는  $N \times N$  DCT 도메인에서, 공간 주파수를 나타내며[10], 표 1은 4x4 DCT 도메인에서

표 1.  $N \times N$  DCT 도메인에서, x와 y위치에서의 공간 주파수 ( $f_N(x,y)$ )

DCT Coefficient Index	0	1	2	3
0	0.00	0.13	0.25	0.38
1	0.13	0.18	0.28	0.40
2	0.25	0.28	0.35	0.45
3	0.38	0.40	0.45	0.50

$f_N(x,y)$ 의 값을 보여준다. 여기에서  $x$ 와  $y$ 는  $4 \times 4$  DCT 도메인에서 수평과 수직의 좌표를 각각 나타낸다.

첫 번째 단계에서, 제안된 알고리즘은 관심영역의 크기에 따라 시각적 가중치를 각 MB에 할당한다. 관심영역의 크기는 인코딩 비트 량에 비례해서 조절된다. 만약 현재 관심영역의 MSE가 목표 MSE보다 작으면, 그 다음 레벨의 확장된 영역으로 관심영역을 확장한다. 따라서 적은 인코딩 비트 량을 갖고 있는 낮은 화질 계층에서는 가장 작은 크기로 관심영역을 제한함으로써, 시각적 중요도가 큰 영역을 중심으로 비트를 할당한다. 반면에, 높은 화질 계층에서는 낮은 화질계층보다 인코딩 비트 량이 충분하기 때문에, 관심영역의 크기를 비 관심영역으로 확장시킴으로써, 결과적으로 전체적인 화질의 향상을 얻게 된다.

두 번째 단계에서는 MB내의 DCT 계수를 할당된 시각적 가중치에 따라서 필터링하는 과정이다. 관심영역으로 선택된 MB는 모든 DCT 계수를 인코딩 할당 한다. 이에 반하여 비 관심영역에 속하는 MB는 가장 높은 공간 주파수대역의 계수만을 전송하고 나머지 영역의 DCT 계수는 0으로 설정함으로써 필터링 과정을 수행하게 된다.

그림 3은 할당된 공간 주파수에 따른 DCT 계수의 필터링의 범위를 SOCCER 테스트 영상의 24번째 프레임을 보여준다. 가장 큰 값의  $f_q^{i,j}$ 은 9로 그리고 가장 작은 값의  $f_q^{i,j}$ 은 0으로 각각 맵핑된다. 화질 계층 0은 가장 작은 관심영역을 유지하기 위해서,  $f_q^{i,j}$  값이 9인 영역만이 관심영역으로, 나머지 영역은 비 관심영역으로 분리되고, 이를 중심으로 DCT 계수가 필

터링 된다. 결과는 그림 3 (a)에서 보여주고 있다. 그림 3 (b)은 화질계층 4에서,  $f_q^{i,j}$  값이 6인 부분까지 관심영역으로 분리되고, 나머지 영역은 비관심영역으로 분리되면서 그림 3 (b)과 같이 필터링 된다. 결국 화질계층이 증가함에 따라서 점점 관심영역이 점점 증가하게 되고, 전체적인 화질이 향상되는 결과를 볼 수 있다.

### 5. 실험 결과

시각적 가중치를 이용한 제안된 비트할당 알고리즘(visual weighting SVC, VW-SVC)은 성능 평가를 위해서, 기존의 H.264 SE에 구현하였다[4]. 모션의 검색 범위는 32이고,  $4 \times 4$  DCT가 공간주파수의 양자화를 위해서 사용되었다. 순방향 및 역방향의 계층적 예측 구조가 압축율의 향상을 위해서 사용되었다.  $I_{MAX}$ 의 값은 실험적으로 50을 설정하였다. 테스트 영상은 CIF의 해상도에 초당 30프레임을 갖고 있는 Soccer, City, Stefan, 그리고 Silent 가 사용되었다. 화질계층의 구성은 한 개의 기저 계층과 2개의 향상계층으로 구성하였다. GOP의 크기 및 최대 시간계층의 크기는 각각 16과 5로 구성하였다.

성능평가를 위해서, 기존의 스케일러블 알고리즘(Non visual weighting SVC, NVW-SVC)을 사용하였다. 또한 HVS의 특성을 반영한 성능평가를 위해서, foveal-PSNR (FPSNR)를 사용하였다[12]. 먼저, 포비에이션의 시각적 가중치를 이용한 foveal-MSE (FMSE)가 다음과 같이 계산된다.

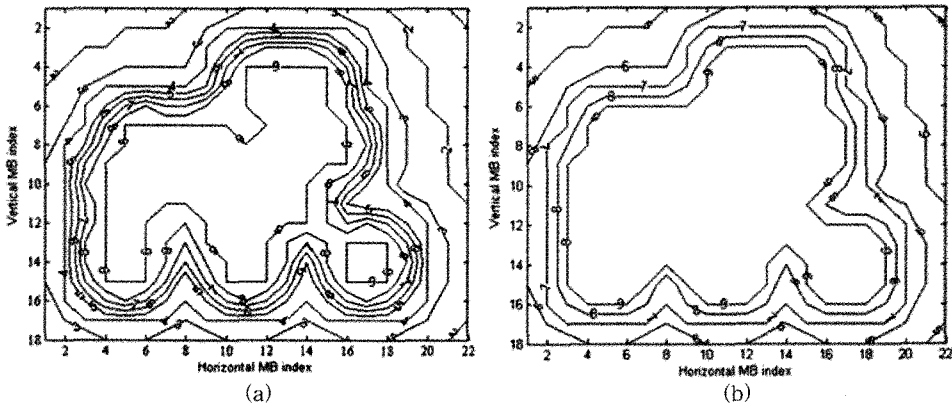


그림 3. 각 화질 계층에서 관심영역의 크기에 따른 공간 주파수의 변화. (a) 화질 계층 0 (b) 화질 계층 4

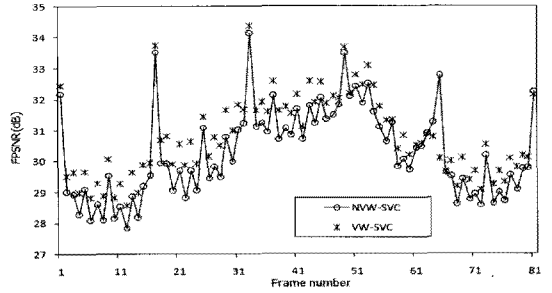
$$FMSE = \frac{1}{\sum_{i=1}^I \sum_{j=1}^J (f_q^{i,j})^2} \cdot \sum_{i=1}^I \sum_{j=1}^J (o_{i,j} - r_{i,j})^2 (f_q^{i,j})^2 \quad (8)$$

여기에서  $I \times J$  비디오 프레임에서  $i$ 와  $j$ 는 수평과 수직을 나타낸다.  $o_{i,j}$ 는 원본 비디오 프레임이고  $r_{i,j}$ 는 복원된 프레임을 나타낸다. FMSE를 이용해서 foveal-PSNR(FPSNR)을 다음과 같이 계산할 수 있다.

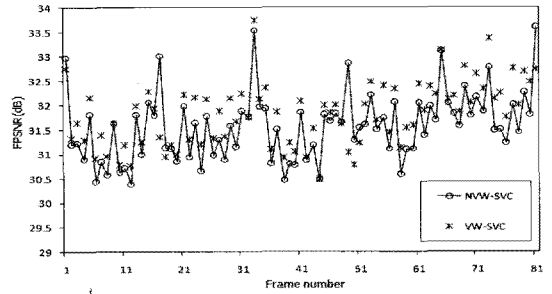
$$FPSNR = 10 \cdot \log_{10} \frac{255^2}{FMSE} \quad (9)$$

그림 4는 기초 계층에서의 각 프레임 당 FPSNR 결과를 각 테스트 시퀀스에 대해서 보여주고 있다. 또한 표 2는 각 테스트 시퀀스의 평균 FPSNR을 보여주고 있다. NVW-SVC와 비교하여, 제안된 알고리즘은 Soccer, City, Stefan, 그리고 Silent에 대해서 각각 0.5, 0.3, 0.3, 그리고 0.2 dB의 평균 FPSNR 성능 향상을 보여 주고 있다. 이러한 화질 개선은 비 관심 영역과 관심영역의 존재하는 각 매크로 블록의 DCT의 계수를 조절함으로써, 제한된 인코딩 비트를 효율적으로 사용하여 관심영역의 화질을 개선했기 때문이다.

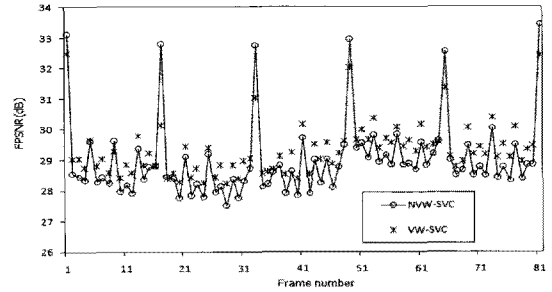
그림 5와 6에서는, SOCCER의 28번째와 CITY의 76번째 비디오 클립을 사용한 주관적 화질 평가의 결과를 나타내었다. SOCCER 영상은 운동선수의 움직임에 따라, 관심영역이 움직이는 특징을 갖는 반면, CITY는 중간에 존재하는 타워를 중심으로 관심영역이 일정하게 유지되는 것을 특징으로 하는 테스트 영상이다. NVW-SVC는 관심 영역을 전체 영역으로 생각하기 때문에, 낮은 비트 량에서, 주관적 화질이 많이 떨어짐을 알 수 있었다. 이에 반하여, 제안된 알고리즘이 관심영역을 중심으로 비트를 할당하기 때문에, 낮은 화질 계층에서도 NVW-SVC보다 주관적 화질의 향상을 볼 수 있었다. 제안된 관심영역을 중심으로 한 스케일러블 비디오 코딩의 비트 할당 알고리즘은 낮은 비트율에서, 효과적으로 주관적 화질을 향상시킬 수 있을 뿐만 아니라, 높은 화질 계층에서도, 선택된 관심영역을 중심으로 비트를 할당



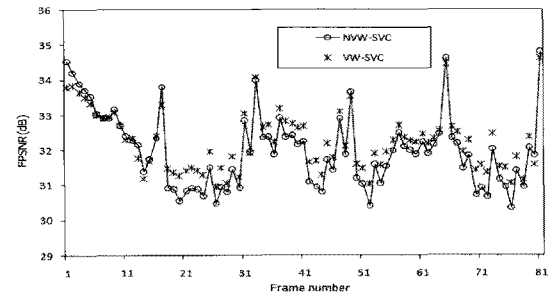
(a)



(b)



(c)



(d)

그림 4. 기초 계층에서의 FPSNR 결과. (a) Soccer (b) City (c) Stefan (d) Silent

표 2. 각 테스트 시퀀스의 평균 FPSNR

테스트시퀀스 알고리즘	Soccer	City	Stefan	Silent
NVW-SVC	30.26	31.55	28.99	31.97
VW-SVC	30.79	31.81	29.32	32.22

함으로써, 스케일러블 비디오 코딩의 화질을 효과적으로 향상시킬 수 있었다.

제안된 비트 할당 알고리즘은 매크로 블록내의 DCT 계수를 공간 주파수의 민감도에 따라, 필터링하는 방법을 제안하였다. 향후에는 시각적 가중치를

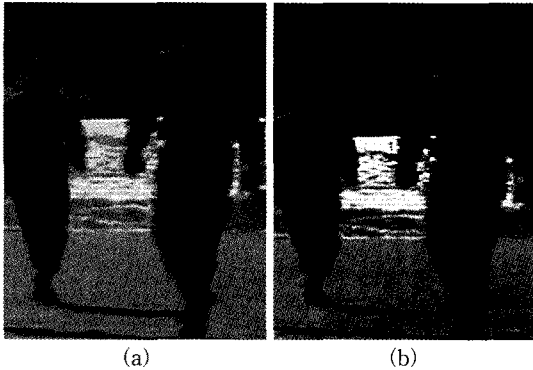


그림 5. Soccer의 28번째 비디오 클립에서 각 알고리즘의 주관적 화질 평가. (a) NVW-SVC (287.6kbps) (b) VW-SVC (281.7kbps)

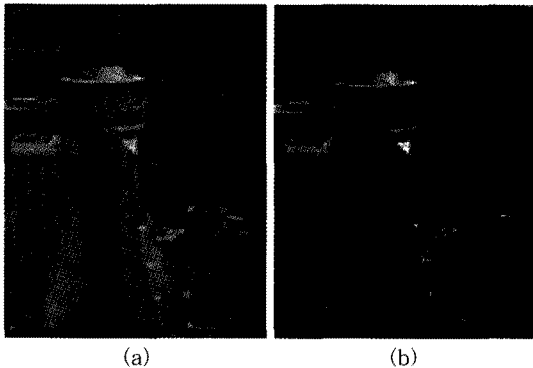


그림 6. City의 76번째 비디오 클립에서 각 알고리즘의 주관적 화질 평가. (a) NVW-SVC (255.6 kbps) (b) VW-SVC (250.5kbps)

이용한 효율적인 스케일러블 비디오 코딩을 위해, RD(Rate and Distortion)단계에서의 효율적인 QP (Quantization Parameter)의 할당 방법 등을 고려할 수 있다.

## 6. 결 론

본 논문에서는 스케일러블 비디오 코딩 알고리즘에서, 인간의 시각적인 특성을 이용하는 새로운 비트 할당 알고리즘을 제안한다. 첫 번째로, 인간의 시각적인 특성을 MB 레벨에서 분석하는 시각적 가중치 모델을 제안하였다. 제안된 스케일러블 비트 할당 알고리즘은 각각의 화질 계층에 존재하는 관심 영역의 크기를 시각적 가중치 모델을 이용하여 조절함으로써, 주어진 인코딩 비트 량에 최적의 비디오 화질을 얻도록 하였다. 기존의 알고리즘과 비교하여, 제안된

알고리즘은 같은 인코딩 비트 량에 화질의 향상을 얻을 수 있었다.

## 참 고 문 헌

- [1] ITU-T and ISO/IEC JTC 1, "Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video," *ITU-T Recommendation H.262 and ISO/IEC 13818-2 (MPEG-2 Video)*, 1994.
- [2] ITU-T and ISO/IEC JTC 1, "Advanced Video Coding for Generic Audio-Visual Services," *ITU-T Recommendation H.264 and ISO/IEC 14496-10(MPEG-4 AVC)*, 2003.
- [3] Text of ISO/IEC 14496-10:2005/FDAM 3 Scalable Video Coding, *Joint Video Team (JVT) of ISO-IEC MPEG & ITU-T VCEG*, Lausanne, N9197, 2007.
- [4] Text of ISO/IEC 14496-4:2001/PDAM 19 Reference Software for SVC, *Joint Video Team(JVT) of ISO-IEC MPEG & ITU-T VCEG*, N9195, 2007.
- [5] A. B. Watson, J. Hu, and J. F. McGowan III, "DVQ: A digital video quality metric based on human vision," *J. Electron. Imag.*, Vol.10, No.1, pp. 1164-1175, 1997.
- [6] P. G. J. Barten, "Evaluation of Subjective Image Quality with the Square-Root Integral Method," *J. Opt. Soc. Amer.*, Vol.7, No.10, pp. 2024-2031, 1990.
- [7] G. Beach, C. J. Cohen, J. Braun, and G. Moody, "Eye Tracker System for Use with Head Mounted Displays," in *Proc. IEEE ICSMC*, Vol.5, pp. 4348-4352, 1998.
- [8] L. Itti, "Quantifying the Contribution of Low-Level Saliency Human Eye Motivement in Dynamic Scenew," *Vis. Cognit.*, Vol.12, No.6, pp. 1093-1123, 2005.
- [9] 이호영, 권순각, 이중화, "H.264 동영상 부호화에서 관심영역의 주관적 화질 개선 방법," *한국멀티미디어학회논문지*, Vol.12, No.7, pp. 913-921, 2009.

- [10] C. Tang, "Spatiotemporal Visual Consideration for Video Coding," *IEEE Trans. on Multimedia*, Vol.9, No.2, pp. 231-238, 2007.
- [11] J. Chen, J. Zheng, and Y. HeY, "Macro Block-Level Adaptive Frequency Weighting for Perceptual Video Coding," *IEEE Trans. on Consumer Elec.*, Vol.53, No.2, pp. 756-781, 2007.
- [12] S. Lee and A. C. Bovik, "Very Low Bit Rate Foveated Video Coding for H.263," in Proc. IEEE ICASSP, pp. 3113-3116, 1999.
- [13] Z. Wang, L. Lu, and A. C. Bovik, "Foveation Scalable Video Coding with Automatic Fixation Selection," *IEEE Trans. on Image Processing*, Vol.12, No.2, pp. 243-254, 2003.



전 선 국

1998년 중국 연변과학기술대학  
전자전산학과(공학사)  
2002년 강원대학교 제어계측 공  
학과(공학 석사)  
2009년 연세대학교 전기전자공  
학부(공학 박사)

2010년~현재 남서울대학교 전자공학과(전임 교수)  
관심 분야: 센서 네트워크, 무선통신, 멀티미디어 통신



하 호 진

1998년 명지대학교 제어계측 공  
학과(공학사)  
2000년 한양대학교 제어계측 공  
학과(공학 석사)  
2009년 연세대학교 전기전자 공  
학부(공학 박사)

2000년~현재 삼성전자 DMC연구소 (책임연구원)  
관심 분야: 멀티미디어 통신, 멀티미디어 신호처리, 컴  
퓨터 네트워크