

정지영상 및 동영상에서의 효율적인 블록효과 측정방법

김 문 성* 정 진 구**

A New Effective Measure of the Block Effect in Still Images and Moving Pictures

Moon-seong Kim* Jin-ku Jeong**

요 약

블록 코딩을 기반으로 하는 영상의 압축 부호화는 영상 및 비디오 압축 표준으로 적용되어 왔다. 그런데 영상 복원 시 발생하는 블록효과에 대한 효율적인 측정방법을 찾아보기 힘든 실정이다.

본 논문에서는 서브블록 DCT 부호화를 이용한 복원 영상에 대한 객관적인 블록효과 측정방법을 제안한다. 제안한 방법이 주관적인 순위매김과 일치하는 블록효과 측정방법임을 실험결과로서 입증한다. 또한 제안한 이 새로운 측정방법은 복원 영상에서 쉽게 그리고 효율적으로 블록효과를 측정하는데 널리 사용될 수 있을 것으로 기대한다.

Abstract

Compression coding based on block coding has been applied in image and video compression standard. But there is no block effect measurement due to block based image coding.

In this paper, we propose a objective block effects measurement to reconstructed image using subblock DCT coding. Experimental results show that the block effects measures given by the suggested method agree well with the subjective ranking. This new objective measurement is simple and effective in measuring the block effect in the reconstructed image.

* 대원과학대학 컴퓨터정보통신과 조교수
** 벽성대학 컴퓨터계열 조교수

I. 서론

블록 기반 영상 압축 부호화에서 영상은 서브 블록으로 나누어지고, 변환과 압축은 각각의 블록에서 독립적으로 수행된다. 만일 요구되는 압축 비트율이 상당히 낮다면 블록 효과가 복원 영상에서 나타나게 된다. 서브블록 Discrete Cosine Transform(DCT) 부호화는 H.26x, JPEG, MPEG2 그리고 HDTV를 포함하는 영상 및 비디오 압축 표준에 의해 적용되었다. 일반적으로 신호대 잡음비(SNR)가 높음에도 불구하고 주관적인 평가는 상당히 낮은 경우를 볼 수 있다[1]. 그러므로 신호대 잡음비(SNR)가 에러(error)에 대한 인간의 시각 인지도를 반드시 나타낸다고는 할 수 없다. 인간이 영상을 바라보는 최종의 사용자이므로 영상에 대한 인간의 시각 인지도는 가장 중요한 주관적인 영상의 질(subjective image quality)을 평가하는 것이라고 할 수 있다[2][3].

동일한 영상에 대하여 서로 다른 사람들의 주관적인 민감도의 차이 때문에 믿을만한 주관적인 영상 화질의 평가 방법을 얻기 위해서는 일반적으로 서로 다른 사람들에게 주어진 영상에 대해 주관적인 순위 매김(subjective ranking)에 대한 통계적인 평균이 사용되어야 한다[4]. 많은 실제 상황에서 인간 관찰자의 주관적인 시험을 제거하고 개개인의 관찰자에게서 나타난 영상의 질에 대한 차이를 최소화하는 일이 필요하기 때문에, 블록 효과가 발생했을 때 객관적으로 블록 효과를 측정함으로써 영상의 질을 나타낼 수 있다. 본 논문에서는 복원 영상에서 블록 효과가 발생했을 때 이를 고려하여 영상의 질을 측정하는 방법을 제안한다.

II. 제안 방법

1. 정지 영상에서 평가 방법

시각은 서로 다른 배경에서 작은 에러(error)에 대해

다양한 허용 오차를 가지고 있다. 이에 대한 다양한 주관적인 시험[5][6][7]에서 어떤 가시도 임계치(visibility threshold) 하에서의 에러는 전혀 인식할 수 없다는 것을 보여 주었다. 영상의 특성을 측정하는 방법 중 하나는 영상의 활동도 계수(activity coefficients)를 측정함으로써 이루어질 수 있다. 영상의 한 지점 $f(I,J)$ 에서 활동도 계수는 다음과 같이 정의 되어 질 수 있다.

$$A_n = \max \{ |f(I,J) - f(I-1,J-1)|, |f(I,J) - f(I-1,J)|, |f(I,J) - f(I-1,J+1)|, |f(I,J) - f(I,J-1)|, |f(I,J) - f(I,J+1)|, |f(I,J) - f(I+1,J-1)|, |f(I,J) - f(I+1,J)|, |f(I,J) - f(I+1,J+1)| \} \quad (1)$$

Limb와 Pirsch에 의하여 제안된 활동도 계수(activity coefficients)와 가시 영향(vision effect) 사이의 관계를 기초로 하여, 아래의 그림 1과 같은 관계 함수(relation function)를 나타낼 수 있다[8].

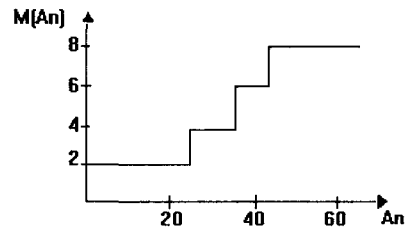


그림 1. 활동도 계수 A_n 과 가시도 임계치 $M(A_n)$ 과의 관계함수
Fig. 1 The relation function between activity coefficient A_n and visibility threshold $M(A_n)$.

위 그림의 관계에서 보면 영상에서 한 지점의 화소점과 그의 주변 지점의 화소점 사이의 차분값이 가시도 임계치의 아래에 있다면 시각에 대해 두 화소점은 서로 다른 점으로 보이지 않는다. 그렇지 않으면 인간의 눈은 두 점에 대해 어떤 차이감을 느낄 것이다. 그래서 우리는 블록 효과의 측정에 대해 다음과 같은 결론을 내릴 수 있다. 서브 블록 부호화에서 블록 효과는 서브 블록들 사이의 독립적인 압축으로 인해서 이웃한 블록들 사이에 인위적인 불연속으로부터 기인하며, 이웃한 서브 블록들의 경계치에서 나타나는데 이 점을 블록 효과 영역(block effect point)이라 부르고 이러한 위치는 영상에 미치는

영향이 결정적이다.

두 개의 서로 다른 서브 블록에 속해 있는 두 지점이 있을 때 그 두 점의 차분값이 허용하는 시각 마스크 효과 (visual mask effect)의 범위에 포함되면 이 두 점의 차이값은 인간에 의해 인식될 수 없다. 그와 같은 점에서는 가시적인 블록 효과가 발생하지 않는다. 그렇지 않으면 인간의 눈은 두 점의 차이점을 인식할 수 있다. 그러나 모든 지점이 블록 효과의 지점이 되는 것은 아니다. 만일 원 영상(original image)에서 이웃한 서브 블록의 경계치(subblock's boundary)에 에지(edge)가 존재한다면 에지 상에 있는 점들은 블록 효과 영역이라고 간주할 수 없다. 블록 효과 영역을 결정하기 위해서 다음과 같은 경우를 고려한다.

- (1) 복원한 영상에서 서브 블록의 경계 지점이 보이고 원 영상의 일치하는 곳에서 경계 지점이 보이면, 그 점은 원 영상에서 실질적인 에지 점(edge point)이다. 그러므로 그 점은 블록 효과 영역이라 할 수 없다.
- (2) 복원한 영상에서 서브 블록의 경계 지점이 보이고 원 영상의 일치하는 곳에 경계 지점이 안 보이면, 그 점은 블록 효과 영역이라 할 수 있다.
- (3) 복원한 영상에서 서브 블록의 경계점이 안 보이고 원 영상의 일치하는 곳에서 경계 지점이 보이면, 그 점은 번짐 에지 지점(blurred edge point)으로 간주된다. 따라서 블록 효과 영역으로 간주될 수 없다.
- (4) 복원한 영상에서 서브 블록의 경계 지점이 안 보이고 원 영상의 일치하는 곳에서 경계 지점이 안 보이면, 그 점은 블록 효과 영역으로 간주될 수 없다.

복원한 영상에서만 경계 지점이 보이는 경우 (2)만이 블록 효과 영역이 된다. 블록 효과를 측정할 수 있는 실행 가능한 알고리즘을 아래에 제시한다. 원 영상을 $X(I,J)$, 복원 영상을 $Y(I,J)$, $I=1,2, \dots, N$, $J=1,2, \dots, N$ 으로 가정한다. 블록 효과 영역의 전체 수를 M 이라 하고, 복원 영상에서 이웃한 서브블록들 사이의 전체 경계점의 수를 T 라 나타낸다. 8×8 서브 블록 영상

부호화에서는 블록 효과 가능 점이 다음과 같은 곳에서 나타날 수 있다.

$$\text{MOD}(I,8) = 1, \text{MOD}(I,8) = 0 \text{ 또는 } \text{MOD}(J,8) = 1, \text{MOD}(J,8) = 0 \quad (I \neq 1, I \neq N, J \neq 1, J \neq N)$$

서브블록 경계점 상에 있는 $\text{MOD}(I,8) = 1, \text{MOD}(I-1,8) = 0$ 또는 $\text{MOD}(J,8) = 1, \text{MOD}(J-1,8) = 0$ 을 만족하는 점들은 서로 이웃하는 점들이다. 블록 효과 영역을 결정하는데 있어서 여기서는 단지 두 가지 경우 즉 $\text{MOD}(I,8) = 1, \text{MOD}(J,8) = 1$ 인 점만을 고려하면 된다.

이러한 과정의 알고리즘을 살펴보면 다음과 같다.

[Step 1] 초기화 : $M = 0, T = 0$

[Step 2] 그림 2에서 보는 바와 같은 순서에 따라 전체 영상을 검색한다.

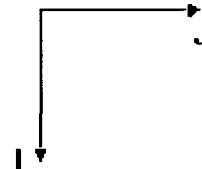


그림 2. 검색 순서
Fig. 2 Scanning order.

만일 검색이 영상의 마지막 지점까지 도달하면 [Step5]로 간다.

[Step 3] 만일 $\text{MOD}(I,8) \neq 1$ 또는 $\text{MOD}(J,8) \neq 1$ ($I=1, J=N, J=1, J=N$)이면, [Step 2]로 되돌아 간다.

[Step 4] 만일 $\text{MOD}(I,8)=1$ 이면, 그 점은 이웃한 서브 블록들의 경계점이다. 그러므로 $T=T+1$ 이라 놓는다. 이와 같은 경우에 J 축과 평행한 선 상에서 블록 효과가 나타날 수 있으므로 아래와 같은 값을 선택한다.

$$V = |X(I,J) - X(I-1,J)|, \quad W = |Y(I,J) - Y(I-1,J)|$$

만일 $\text{MOD}(J,8)=1$ 이면, 그 점은 이웃한 서브 블록의

경계점이다. 그러므로 $T=T+1$ 이라 놓는다. 이와 같은 경우에 I축과 평행한 선 상에서 블록 효과가 나타날 수 있으므로 아래와 같은 값을 선택한다.

$$V = |X(I,J) - X(I,J-1)|, W = |Y(I,J) - Y(I,J-1)|$$

블록 효과를 가지고 있는 영상에서 화소의 활동도 (activity of pixel)는 블록 효과에 의해서 파괴될 수 있다. 이 지점의 활동도 계수(activity coefficient) A_n 은 식 (1)과 원 영상 $X(I,J)$ 을 통해 계산될 것이다. 만일 $V \leq M(A_n)$ 그리고 $W > M(A_n)$ 이면, 이 지점은 블록 효과 영역으로 간주하고 $M=M+1$ 로 놓는다. 다음에 다시 [Step 2]로 넘어 간다.

[Step 5] 전체 블록 효과 영역 총수(M)와 복원 영상에서 서브 블록의 경계점에서 그들이 차지하는 비율(R)을 구할 수 있다.

$$R = (M / T) \times 100\%(2)$$

2. 동영상에서 평가 방법

동영상의 경우 연속한 프레임 상에 블록 효과 영역이 정지영상만큼 나타나지 않는다. 그래서 동영상에서 시공간 주파수 특성과 인간 시각을 고려하였다. 시공간 주파수 특성 $W(u,v,f)$ 는 공간 주파수 특성 $WS(u,v)$ 와 시간 주파수 특성 $WT(f)$ 의 곱으로 나타내 질 수 있다.

$$W(u,v,f) = WS(u,v) \cdot WT(f)(3)$$

$WS(u,v)$ 와 $WT(f)$ 는 다음과 같은 식으로 될 수 있다.

$$WS(u,v) = 2.46 \times (0.1 + 0.25(u^2 + v^2))^{1/2} \times \exp(-0.25(u^2 + v^2))^{1/2}$$

$$WT(f) = 0.134 (1 + (f/0.5)) / (1 + (f/7.8)^2)^{1.2}(4)$$

마스킹(masking)은 공간상의 특성 중 하나인데 동영상의 측정방법에 마스킹을 도입하기 위해서 마스킹 계수(masking coefficient)와 블록의 엔트로피에 관한 연구가 수행되었다. 다음 그림은 마스킹 계수와 엔트로피의 관계를 나타낸다.

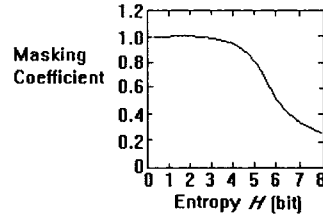


그림 3. 마스킹 계수와 엔트로피와의 관계
Fig. 3 The relation between masking coefficient and Entropy.

위와 같은 특성을 고려해 3D-SNR을 구하는 알고리즘은 아래와 같다.

- (1) 원영상과 복원영상의 차분값에 마스킹 계수를 곱한 $N(x,y,t)$ 를 구한다.
- (2) $N(x,y,t)$ 를 3D-FFT 함으로써 3차원 주파수 값을 $N(u,v,f)$ 를 구한다.
- (3) 시공간 주파수 특성 $W(u,v,f)$ 와 $N(u,v,f)$ 를 곱해 $N'(u,v,f)$ 를 유도한다.
- (4) $N'(u,v,f)$ 를 이용해 3D-SNR을 구한다.

$$SX/2 SY/2$$

$$3D-SNR = 10\log(255 \times 255) / \sum_{U=0} \sum_{V=0} N'(u,v,f) (5)$$

III. 실험 결과

정지영상의 경우 JPEG 알고리즘에 따라 영상을 부호화한 후 복원하였는데 블록효과를 나타내는 영상을 얻기 위해서 다음 식에 따라 양자화 하였다.

$$F^*(u,v) = \text{Nearest integer}(F(u,v) / (Q(u,v) \times nf)) (6)$$

여기서 $Q(u,v)$ 는 정규화배열, nf 는 정규화계수로서

양자화 단계 크기(quantization step size)를 결정하는 요소가 된다.

동영상의 경우 Miss America의 연속적인 프레임들 식(5)에 대하여 이전 프레임과의 상대적인 값을 나타내었다.

표 1. 가시도 임계치 = 2 를 사용한 결과
Table 1. The result used in visibility threshold = 2.

nf	Lenna		Girl	
	R	M	R	M
0.5	3.10	463	1.60	238
2.5	8.00	1193	4.90	729
4.0	10.48	1562	7.65	1140
6	15.79	2354	9.00	1343

본 논문에서 제안한 방법을 이용한 실험 결과와 주관적 순위 매김(subjective ranking)과의 관계를 살펴보면 다음과 같다.

표 2. 블록 현상과 주관적 순위 매김과의 관계
Table 2. The relation between block effect and subjective ranking.

R	주관적인 순위매김
$0 \leq R < 1$	높음 1
$1 \leq R < 4$	다소높음 2
$4 \leq R < 9$	보통 3
$9 \leq R < 12$	다소낮음 4
$R \geq 12$	낮음 5

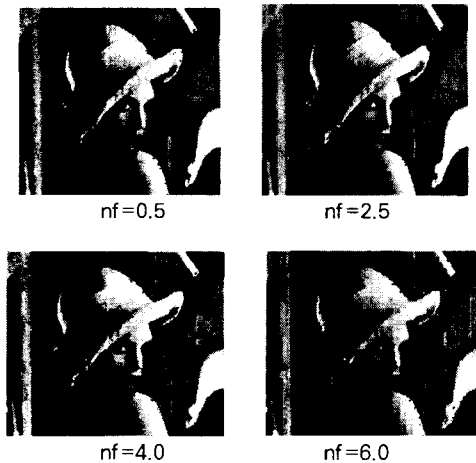


그림 4. 영상 적용 결과
Fig. 4 The result of image application.

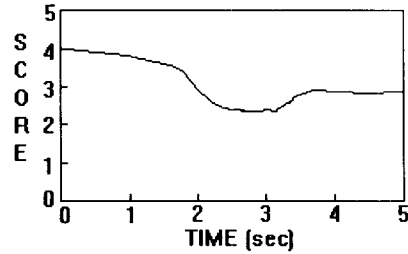


그림 5. 동영상에 적용한 변화도
Fig. 5 The change graph for moving picture application.

IV. 결론

본 논문에서는 정지영상과 동영상으로 나누어 블록효과가 발생했을 때 영상의 질을 평가하는 방법을 제안하였다. 정지영상의 경우 복원영상에서 이웃한 서브블록의 경계점들의 차분값이 가시도 임계치를 초과하고 원영상에서 같은 위치의 점에서의 차분값의 가시도 임계치를 넘지 않을 때 그 지점을 블록효과 지점으로 간주한다. 영상의 서브블록들 사이에서 경계점들의 총수에 대한 블록 효과 영역의 비율이 복원영상에서 블록효과에 대한 객관적인 평가기준으로 사용될 수 있다.

이와 같은 평가방법을 사용함으로써 얻어진 평가결과는 주관적인 순위매김과 잘 일치함을 확인하였다. 영상의 복잡도(detail of the image)는 가시도 임계치에 영향을 주는 매우 중요한 요소이지만, 다른 요인들 즉, 배경의 명도, edge의 길이, 크기 등도 역시 가시도 임계치에 영향을 주는 요인들이 될 수 있다.

동영상의 경우 연속된 프레임에 대한 질을 평가하는데 있어서 주관적인 평가는 인간 시각의 특성 때문에 완만한 경사를 가지고 변하는데 비해, 본 논문에서 제안한 방법은 주관적인 평가의 비슷한 곡선을 나타내었는데, 그 이유는 3D-SNR이 인간 시각의 특성을 고려했기 때문이다.

참고문헌

[1] Shanika A. Karunasekera and Nick G. Kingsbury "A Distortion Measure for Blocking Artifacts in Images Based on Human Visual Sensitivity" IEEE Trans. on Image Processing.

[2] F.X.J Lukas and Z.L. Budrikis, "Picture quality prediction based on a visual model", IEEE Trans. Commun., vol. COM-30, no. 7, pp. 1679-1692, July 1982.

[3] M.Miyahara, "Quality assessment for visual services", IEEE Commun. Mag., pp. 51-60, Oct. 1988.

[4] Okamoto et al. "A Study on subjective and Objective Evaluation Method for Coded Moving Picture Quality" International Picture Coding Symposium, pp. 519-523, 1996.

[5] Sakata et al. "Spatio-Temporal Frequency Response of Sine-Wave and Television Noise by Human Vision", Proc. of IEICE, J65-B, pp. 989-996, 1982.

[6] B.Chitprasert. "Human Visual Weighted Progressive Image Transmission" IEEE Trans. Com., vol. 38, no. 7, pp. 1040-1044, 1990.

[7] Kubota et al. "A Three Dimensional Noise Weighting Function and Its Application to HDTV Transmission", Proc. of IEICE, J69-B, pp. 51-56, 1986.

[8] Limb and Peter pirsch, "Design of DPCM Quantizers for Video Signals Using Subjective Tests", IEEE Trans. Commun., vol. COM-29, no. 7, July 1981.

저자 소개



김 문 성
 1993. 2 한남대학교 물리학과 졸업
 1995. 2 한남대학교 대학원 컴퓨터공학과 졸업(공학 석사)
 2002. 8 전북대학교 대학원 컴퓨터공학과 졸업(공학 박사)
 1997. 3 ~ 현재 : 대원과학대학 컴퓨터정보통신과 조교수
 관심분야 : Intelligent Control, Image Signal Processing, Computer Simulation



정 진 구
 1992. 2 한남대학교 컴퓨터공학과 졸업
 1994. 2 한남대학교 대학원 컴퓨터공학과 졸업(공학석사)
 1995. 8 ~ 현재 : 벽성대학 컴퓨터계열 조교수
 관심분야 : 시스템시뮬레이션, 컴퓨터그래픽스, 멀티미디어, 전자상거래