

동영상 전송을 위한 내용기반 동적 대역폭 조절

(Content-based Dynamic Bandwidth Control for Video Transmission)

김 태 용 * 최 종 수 **
(Tae Yong Kim) (Jong Soo Choi)

요약 본 연구에서는 DCT(Discrete Cosine Transform) 영역에서 불연속 특징을 검출하여 내용기반의 동적 비디오 전송을 위한 트랜스코딩 방법을 제안한다. 이 방법에 의하여 동영상 스트림 각각의 DCT 블록은 내부의 대표적인 불연속의 크기에 따라 다르게 트랜스코딩되어 전송되며, 실험에서 같은 대역폭을 유지하면서 기존의 저주파 통과 필터에 의한 방법보다 내용기반 방법이 더 좋은 비디오 화질을 나타내며 픽셀 영역에서의 방법보다 처리 시간이 빠름을 나타낸다.

키워드 : 영상처리, DCT 변환, 적응형 필터, 트랜스코딩

Abstract In this paper, we propose a content-based MPEG transcoding method using a discontinuity feature in the Discrete Cosine Transform (DCT) domain. A DCT block is transcoded differently depending on the height of dominant discontinuity within a block. In the experiment, we show the result that the video quality of content-based transcoding is better than that of a constant cut-off method and the processing time of the adaptive method is much faster compared with the pixel domain methods in the same bandwidth.

Key words : Image processing, Discrete Cosine Transform, Adaptive Filtering, Transcoding

1. 서론

최근의 인터넷 동영상 전송을 이용한 원격회의, VOD (Video On Demand), 원격교육 등의 실현이 활발히 연구되고 있으나, 서로 다른 ATM, TCP/IP, 무선, PSTN 등의 네트워크 환경으로 인하여 필요한 전송 대역폭을 확실히 보장하기는 어려운 상황이다. 따라서 통과 대역폭에 맞도록 동영상의 전송율을 동적으로 변환하는 기술이 필요하게 되었다[1,2]. 동적인 대역폭은 현재의 비디오 코딩 표준에 의한 스케일 코딩(Scalable Coding) 방법으로 조절 가능하였다[3]. 그러나 이 방법은 표준에서의 변환 계층이 제한되어 이산적인 세 단계의 비디오 화질만을 제공할 수 있어서 세밀한 제어가 어려웠다[4].

현재 제공되는 많은 멀티미디어 서비스는 이미 압축되어 있는 동영상을 사용하므로 보다 세밀하고 통신 환경에 능동적으로 대응하는 트랜스코딩 방법을 적용하여

야 한다. 최근의 동영상 트랜스코딩의 연구는 주로 구현이 쉬운 통신 구조(architecture)나 비트할당(bit allocation) 문제에 초점을 맞추고 있다. Assuncao, Bjork, Nakajima[5-7] 등의 연구에서는 복잡성이 적은 트랜스코딩 구조에서 전송을 조절을 위한 단순한 비트 스케일링 방법이 제안되었으며, Werner[8]는 MPEG 인트라 프레임에서 재 양자화(quantization)를 연구하였다. Jun-Xin, Lu[9,10] 등은 인코딩 단계에서의 양자화 및 할당 비트수의 복잡도를 전송을 조절에 재사용하는 방안을 제안하였다. 그러나 이러한 연구들은 영상 내부의 특징 정보를 적극적으로 반영하지 않아서 특징의 왜곡을 줄이거나 시각적 인식을 향상시키는데 한계가 있었다. 따라서 영상물의 시각적인 질을 급격하게 감쇄시키지 않으면서 동적으로 전송율을 조절하기 위해서는 내용기반 트랜스코딩 방법을 적용하여야 한다.

본 연구에서는 압축된 DCT 영역에서 영상의 화질에 영향을 미치는 주요 특징인 불연속을 찾아내어 이를 바탕으로 대역폭을 변환하는 내용기반 동적 대역폭 조절 기법을 제안한다. 이를 위하여 먼저 하나의 DCT 블록에서 가장 대표적인 불연속 특징을 찾아내고, 한 프레임 내부에서 불연속의 크기에 따라 전송 대역폭에 맞추어

* 본 연구는 교육부 BK21 사업의 지원을 받아 수행되었음

† 종신회원 : 중앙대학교 영상공학과 교수
kimty@cau.ac.kr

** 종신회원 : 중앙대학교 영상공학과 교수
jschoi@imabelab.cau.ac.kr

논문접수 : 2003년 10월 22일
심사완료 : 2004년 4월 7일

저주파 통과 필터를 사용하여 동적으로 트랜스코딩을 수행한다. 이 방법은 기존의 고정 저주파 통과 필터보다 같은 대역폭에서 향상된 화질을 나타내고, 비슷한 화질에서는 적은 대역폭을 소비하며, 처리 시간이 빠르므로 실시간 대역폭 조절 응용분야에 적용 가능하겠다.

제2장에서는 MPEG에서의 대역폭 조절과 기존의 트랜스코딩 방법에 대하여 설명하였으며, 제3장에서는 DCT 블록 내부의 불연속 특징을 분석하는 방법을 설명하였다. 본 연구에서 제안하는 내용기반 동적 대역폭 조절 방법의 세부 내용을 제4장에서 설명하였으며, 실험결과 및 결론을 제5장과 제6장에 각각 서술하였다.

2. MPGE 트랜스코딩(MPEG Transcoding)

본 장에서는 MPEG(Moving Picture Expert Group) 압축 표준에서의 대역폭 조절 개념과 기존 연구에서의 대역폭 조절을 위한 트랜스코딩 방법들을 간략하게 소개한다.

2.1 MPGE 압축에서의 대역폭 조절

동영상 장면은 보통 시간적 공간적으로 중복되는 정보를 포함한다. MPEG 압축 기술은 영상내부의 단순한 움직임 가정하여 연속되는 프레임에서 얻을 수 있는 특성을 활용한다. 이 특성은 픽셀간의 연관성이 많음을 나타내며, 따라서 특정한 위치의 픽셀 값은 같은 프레임이나 인접한 프레임의 주변 픽셀 값을 사용하여 유추하도록 한다. MPEG 압축 알고리즘은 공간상의 유사도와 시간적인 움직임 예측의 방법을 조합하여 높은 압축율을 나타낸다. 한 영상의 8x8 세부 블록에서 DCT 변환(Discrete Cosine Transform)을 사용하여 공간적인 유사도를 분석하여 중복되는 정보를 줄여서 압축한다. 또한 움직임이 거의 없는 장면과 같이 시간적 유사도가 높은 경우에는 프레임간의 움직임 벡터를 사용하여 중복 정보를 제거하여 압축한다.

MEPG-1 코딩 기술의 중요한 특징 중의 하나는 DCT 계수의 양자화 단계를 조절하여 특정 응용분야에 맞도록 대역폭을 조절할 수 있다는 것이다. DCT 계수를 크게 나누면 높은 압축 성능을 나타내어 저장 및 전송에 용이하지만 화질의 저하를 자저온다. MPEG 표준에서는 매크로 블록 단위로 다른 양자화 값을 사용할 수 있도록 하여 영상 화질개선에 유연성을 제공하지만 적절한 양자화 값에 대한 기준을 선택하는데 어려움이 있다.

표준 MPEG-2 스케일 기능은 계층적으로 압축된 스트림에 대하여 공간 스케일, SNR 스케일, 시간 스케일 등의 기능을 제공하여 서로 다른 대역폭을 요구하는 서비스간의 컨텐츠 공동 사용을 포함한 표시 방법의 유연성을 제공한다. 수신 표시 장치의 해상도가 낮을 경우에

는 원 영상의 최대 해상도를 사용하는 대신에 전체 스트림의 일부분만을 처리하여 적은 공간·시간 해상도의 영상을 재생하거나 화질을 감소시켜 표시하여 준다. 그러나 이러한 스케일 기능을 사용하기 위해서는 전송 및 수신 양쪽에 스케일 기능을 포함시켜야 하는데 현재 보급된 H/W는 해당 기능을 거의 가지고 있지 않다.

2.2 트랜스코딩 기법

앞에서 살펴본 MPEG 압축 표준의 대역폭 조절 및 스케일 기능은 압축할 때 처리하는 기술이지만, 특정한 서버가 가입자의 네트워크 대역폭이 서비스의 양보다 적을 경우에 동적으로 서비스 대역폭을 조절하는 기능을 가질 수도 있다. 이러한 경우 동적 트랜스코딩 기술을 사용하며, 동적 트랜스코딩은 미리 압축되어 코딩된 동영상 스트림을 다른 형식이나 혹은 다른 대역폭으로 환경에 맞게 동적으로 변환하여 코딩하는 기술을 의미한다.

서버 출력 혹은 가입자의 수신 네트워크 자원이 한계에 도달하였을 경우 전체적인 전송 능력은 현저히 저하된다. 이러한 경우 전송 스트림에 시간적 혹은 공간적 트랜스코딩을 수행하여 서비스의 대역폭을 조절할 수 있다.

시간적인 트랜스코딩의 경우, 트랜스코더는 스트림의 프레임 형식을 분석하여 복원에 필요 없거나 중요하지 않은 동영상 프레임을 제거하여 전송하지 않음으로써 대역폭을 줄일 수 있다. 이 방법은 시간적인 전송 프레임 수를 줄임으로써 대역폭을 감소시키며 프레임 표시를 위한 양측단의 지연시간을 유지할 수 있다.

다른 방법인 공간적 트랜스코딩은 DCT 변환된 주파수 영역에서 수행되므로 엔트로피 복원 및 재압축을 수행하여야 한다. 공간적 트랜스코딩은 저주파 필터나 재양자화 방법을 사용하여 공간 해상도를 줄이게 되므로 화질의 감소를 초래한다. 저주파 필터는 DCT 계수의 고주파 성분의 일정 부분을 제거하여 대역폭을 줄이는 방법으로 번짐 현상을 초래한다. 재양자화 방법은 스트림에 기록된 DCT 계수를 큰 상수(적은 단계)로 다시 계산하여 전체적인 대역폭을 줄일 수 있으나 화질의 왜곡을 초래한다. 이와 같은 방법들은 처리가 간단하여 화질의 중요도 및 응용분야에 따라서 적절하게 조합하여 동적인 대역폭 조절에 사용될 수 있으나, 프레임이나 영상 내부의 정보를 사용하지 않고 균일한 파라미터를 사용하므로 내용의 중요도와 무관하게 화질이 감소되는 단점을 가지고 있다[11,12].

3. DCT 영역 처리(DCT Domain Processing)

이제까지 연구에서는 공간영역에서 영상의 특징을 찾는 많은 방법이 제안되어 왔으나, 특징을 얻기 위하여

압축을 복원하여 특징을 분석하고 재압축하는데 많은 시간이 필요하므로 동적 대역폭 조절과 같은 실시간 응용분야에서는 사용하기 어려웠다.

영상의 회전 등의 변환에서 압축의 일부만을 복원한 DCT 영역에서의 처리 연구[13,14]는 공간 영역의 처리보다 50배 이상 빠른 성능을 보였다. 본 장에서는 DCT 영역에서 대표적인 불연속을 찾는 저자의 기존 연구[15]를 간단하게 소개하고, 이를 실시간 응용에 적합하게 변형한 내용을 제안한다.

3.1 불연속 위치 정렬

JPEG이나 MPEG 압축은 8x8 블록의 DCT 변환을 기본으로 한다. 다음 식은 8x8 블록의 FDCT와 IDCT의 수학적 정의를 나타낸다.

$$F(u, v) = \frac{1}{4} C(u)C(v) \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16} \quad (1)$$

$$f(i, j) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(u, v) \cos \frac{\pi u(2i+1)}{16} \cos \frac{\pi v(2j+1)}{16}$$

여기에서, $u, v=0$ 일 때 $C(u), C(v)=1/\sqrt{2}$, 그 밖의 경우는 $C(u), C(v)=1$.

공간영역에서 이상적인 계단 불연속 모델은 8x8 블록 내부의 중앙(식 (1)의 $j=3$ 과 $j=4$ 사이)에서 지역 불연속이 존재하는 것으로 정의하며, 공간상의 영상 밝기 값을 $f(i,j)$ 라 했을 때 수학적으로 다음과 같이 표시할 수 있다.

$$f(i, j) = \begin{cases} a, & \text{if } j=0,1,2,3 \\ b, & \text{if } j=4,5,6,7 \end{cases} \quad (2)$$

여기에서 $-127 \leq a, b \leq 127$ 이다.

중심 위치($k=4$)는 이미 알고 있는 이상적인 경계 모델과 비교하여 각종 모델 파라미터를 계산하기 쉬운 위치이므로, 임의의 불연속 위치($1 \leq k \leq 7$)를 중심 위치로 이동시킨 것을 정렬이라고 정의한다. 이러한 위치 정렬을 위하여 기존의 DCT 계수에 보상 주파수를 가감하여 주는데, 이는 공간 영역에서의 위치 이동과 같은 효과를 나타낸다. 보상 주파수는 $c_k(h, v) = F_4(0, v) - F_k(0, v)$ 에 의하여 얻어지는데, 여기에서 F_k 는 불연속의 위치가 k 이고 높이가 $h=|b-a|$ 일 경우의 DCT 계수를 나타낸다. F_3 의 경우를 유도하여 정리하면 다음과 같다.

$$c_3(h, v) = F_4(0, v) - F_3(0, v) = \frac{2}{\sqrt{2}} \left\{ \sum_{i=0}^7 f_4(0, i) \cos \frac{\pi v(2i+1)}{16} - \sum_{i=0}^7 f_3(0, i) \cos \frac{\pi v(2i+1)}{16} \right\} = -\frac{2}{\sqrt{2}} h \left\{ \cos \frac{7\pi v}{16} \right\} \quad (3)$$

나머지 위치의 보상 주파수도 위치 k 를 변화시켜서 쉽게 구할 수 있다. 보상된(이동된) DCT 계수들 $\hat{F}_k(0, v) = F(0, v) + c_k(h, v)$ 에 의하여 구할 수 있으며, 여기에서 $1 \leq v \leq 7$ 이며 $\hat{F}_k(0, v)$ 의 k 는 정렬하기전의 대표적인 불연속의 위치를 나타낸다.

3.2 정렬평가 및 특징추출

위치의 정렬평가를 위하여 이상적인 불연속 모델이 가지고 있는 원점대칭 성질을 사용하였다. 공간 영역에서 대칭은 블록의 가운데 수직 축을 중심으로 플립하고 부호를 바꾼 원점 대칭의 특성을 사용한다. 공간 영역에서 입력 블록을 $f(i,j)$ 라 하고 출력 블록을 $g(i,j)$ 라 할 때, 원점 대칭은 $g(i, j) = -f(i, 7-j)$ 이며, $0 \leq i, j \leq 7$ 이다.

DCT 영역에서는 변환 블록은 식 $G(u, v) = -\cos(\pi v)F(u, v)$ 에 의하여 직접적으로 구해진다. 만약 불연속이 $k=4$ 위치에 있으면 $G(u, v) = -\cos(\pi v)F(u, v) = F(u, v)$ 가 되므로 $\hat{F}(0,2), \hat{F}(0,4), \hat{F}(0,6)$ 은 정렬된 위치에서 모두 0이 되어야 한다. 따라서 불연속의 올바른 위치는 보완 주파수를 사용하여 모델의 위치로 이동하였을 때 다음 식을 최소로 하는 지점으로 선정할 수 있다.

$$D_k = \hat{F}_k(0,2)^2 + \hat{F}_k(0,4)^2 + \hat{F}_k(0,6)^2, \quad (4)$$

높이 h 를 고정시킨 상태에서 위치 k 를 변화시켰을 때 D_k 가 최소로 되는 지점에 대표적인 불연속이 존재하며, 높이 h 를 임의로 선정하였기 때문에 D_k 는 0이 되지는 않지만 위치는 파악할 수 있다.

D_k 는 정렬된(파악된) 위치 k 를 대입함으로써 단순화시킬 수 있다. D_k 는 주기함수이고 우함수이므로 알고 있는 위치 k 와 높이 변수 h 를 대입하여 전개하면 다음과 같다.

$$D_k(h) = \hat{F}_k(0,2)^2 + \hat{F}_k(0,4)^2 + \hat{F}_k(0,6)^2 = (c_k(\tilde{h}, 2) + c_k(h, 2))^2 + (c_k(\tilde{h}, 4) + c_k(h, 4))^2 + (c_k(\tilde{h}, 6) + c_k(h, 6))^2 = (\tilde{h}a_k - ha_k)^2 + (\tilde{h}a_k - ha_k)^2 + (\tilde{h}a_k - ha_k)^2 = (\tilde{h} - h)^2 (a_k^2 + a_k^2 + a_k^2), \quad (5)$$

여기에서 \tilde{h} 는 실제의 높이이고, $c_k(h, v)$ 는 위치 k 에서의 보완 주파수이며, $a_{k,2}, a_{k,4}, a_{k,6}$ 는 k 위치에서 주파수 표현을 위한 상수들이다.

이전 연구[15]에서 제시된 DCT 계수의 수직·수평 성분의 비를 이용하면 근사 경사도 방향(approximated gradient direction)인 불연속의 방향을 예측할 수 있고, DCT 영역에서 이동 매트릭스를 순차적으로 적용하여 회전시킬 수 있다[13]. 회전시켜 불연속을 수직방향으로 일치시킨 블록 내부에서 위의 식(5)가 쌍곡선 모양을 가지고, 전역 최소 값을 가지므로 경사도 하강(gradient descent)[16] 방법을 사용하여 높이를 계산할 수 있다.

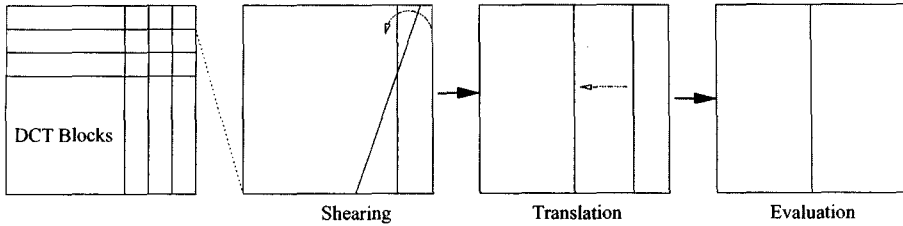


그림 1 DCT영역에서의 불연속 검출을 공간영역의 예로 표현

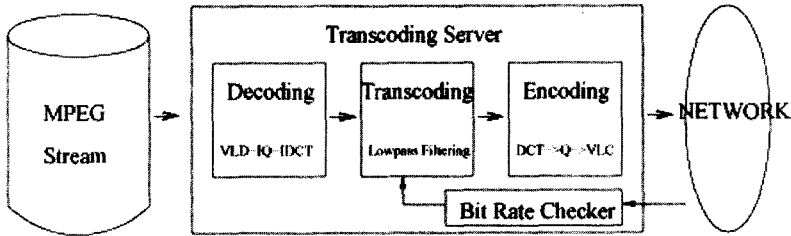


그림 2 동적 스트림 전송서버에서의 디코딩, 트랜스코딩, 인코딩

이 방법은 잡음 불연속을 완화할 수 있고, 검출 속도를 향상시킨다.

DCT 영역에서 불연속 특징을 찾는 방법을 요약하면 다음과 같다.

- ① JPEG이나 MPEG 인트라 프레임의 블록으로부터 밝기 요소의 DCT 계수를 습득.
- ② DCT 계수의 수직·수평 성분의 비를 이용하여 필요하면 회전 연산을 수행.
- ③ 식 (3)을 사용하여 DCT 계수를 정렬하고, 식 (4)를 이용하여 위치 k 를 평가.
- ④ 식 (5)을 이용하여 불연속 경계의 높이 h_k 를 계산.
- ⑤ 영상의 모든 블록을 처리할 때까지 ①에서부터 반복 수행.

그림 1은 이러한 일련의 처리 과정을 공간영역의 예로써 나타내었으며, 실제의 처리는 DCT 계수만을 사용하여 수행된다.

4. 내용기반 대역폭 조절(Content-based Bandwidth Control)

본 장에서는 DCT 블록 내부의 대표적인 불연속 특징을 사용하여 저주파 필터의 차단 값을 능동적으로 조절함으로써 동적으로 동영상 스트림을 트랜스코딩하는 방법을 제안한다.

4.1 저주파통과 필터를 이용한 트랜스코딩

그림 2에 트랜스코딩을 수행하는 일반적인 전송 서버의 기능을 단순하게 나타내었다. 서버는 처음 저장장치에서 MPEG 동영상 스트림을 읽어오고, 가변길이 디코

딩(VLD: Variable Length Decoding), 역 양자화(IQ: Inverse Quantization), 역 DCT(IDCT: Inverse DCT) 등의 디코딩 과정을 수행한다. 디코딩된 스트림은 현재의 전송 대역폭을 판별하여 송신 대역폭에 맞도록 저주파 통과필터(LPF: Low Pass Filter), 재 양자화 등의 축소 기술을 사용하여 트랜스코딩된다. 이렇게 변환된 스트림은 다시 DCT변환, 양자화, 가변길이 코딩 등의 인코딩 과정을 통한 후 전송되어 진다.

본 연구에서는 인트라 프레임의 DCT 영역에서 특징을 찾아 변환하므로 공간영역에서 특징을 찾기 위한 DCT-IDCT 과정을 제거할 수 있어 빠른 처리가 가능하다. 필터링 방법으로 DCT 블록에 포함된 잡음이나 미약한 고주파 성분은 제거되고, 전체적인 화질을 크게 저하시키지 않으면서 전송 서버에서 콘텐츠 전달에 필요한 대역폭을 맞출 수 있다.

DCT 블록의 고주파 성분은 고정된 차단 값 C_c [11,12]를 사용하여 다음과 같은 식에 의하여 제거될 수 있다.

$$\overline{AC}_i = \begin{cases} AC_i & \text{if } i < C_c \\ 0 & \text{if } i \geq C_c \end{cases} \quad (6)$$

여기에서 \overline{AC}_i 는 필터링된 i 번째 DCT 계수이고, $0 < i \leq 63$ 이다.

만약 식 (4)의 C_c 가 각 프레임에서 고정 값을 갖는다면 트랜스코딩은 영상의 내용에 무관하게 일괄적으로 적용되며, 트랜스코딩된 블록의 화질은 모든 블록에서 균일하게 저해된다.

이러한 화질의 저하를 영상의 중요도에 따라 적용하

기 위하여 DCT 블록의 대표적인 특징의 크기를 바탕으로 동적으로 필터링을 수행할 수 있는 동적 필터링 차단함수 $C(h)$ 를 사용한다. 특징의 크기는 블록의 명암 대비를 나타내며 명암의 대비가 큰 블록은 사람의 시각 인식에 큰 영향을 미치므로, 불연속 특징이 큰 블록은 많은 주파수 성분을 포함할 수 있도록 필터링하고, 대표적인 특징의 크기가 작은 블록은 인식에 큰 영향없이 고주파 성분을 제거할 수 있다.

동적 필터링 차단함수 $C(h)$ 는 $[0, a]$ 영역에서 높이에 따라 선형적으로 변화한다. 여기에서 a 는 AC 주파수를 통과시킬 수 있는 특징의 최대 높이 값으로서 영상의 최대 명암 대비를 나타낸다. 다른 파라미터 β 는 필터링될 최대 AC 주파수 개수를 나타내며 대역폭 축소를 위하여 사용된다. 필터링 차단 함수 $C(h)$ 의 정의는 다음과 같고, 그림 3과 같은 특성을 갖는다.

$$C(h) = \begin{cases} \frac{\beta-1}{a}h+1 & \text{if } 0 \leq h \leq a \text{ and } a \neq 0 \\ \beta & \text{if } h > a \text{ or } a = 0, \end{cases} \quad (7)$$

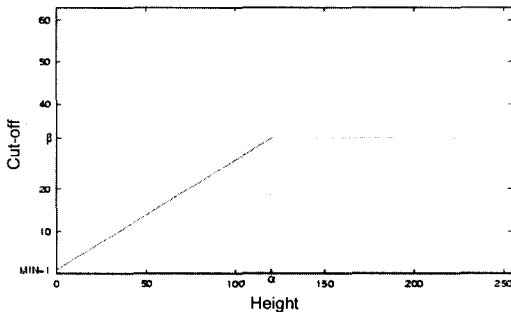


그림 3 파라미터 α, β 변화에 따른 차단 주파수 값

파라미터 α 를 조절함으로써 함수의 기울기를 조절할 수 있고, 이는 필터링 단계를 불연속 높이에 따른 차단 값으로 세밀하게 하거나 거칠게 만든다. 이 높이에 다른 필터링 차단 함수는 보다 나은 화질을 위하여 높이에 따른 비선형 함수로 설계할 수 있으나 빠른 처리 시간과 트랜스코딩된 스트림 대역폭의 선형성을 위하여 그림 3과 같이 선형으로 설정하였고, 고주파 성분의 필터링은 다음과 같이 표현된다.

$$\widehat{AC}_i = \begin{cases} AC_i & \text{if } i < C(h) \\ 0 & \text{if } i \geq C(h), \end{cases} \quad (8)$$

만약 한 DCT 블록의 대표 높이가 0이면 차단 값은 1이 되고 단지 DC 성분만 통과되어 트랜스코딩 된다. 블록의 대표 높이가 α 보다 크면 차단 값은 β 가 되어 β 만큼의 주파수만을 통과시킨다. 따라서 화질의 감소를 초래하지만 β 를 기존 필터링의 고정 값 C_c 보다 2~4배 정도 크게 설정할 수 있으므로 같은 대역폭을 갖는 스

트림으로 변환 했을 경우 동적 필터링 방법이 좋은 화질을 보유하게 된다. α 를 0으로 설정할 경우 $\beta = C_c$ 로 설정할 수 있어 식 (6)의 기존 필터링 방법을 포함할 수 있다.

4.2 압축된 동영상의 대역폭 조절

DCT 영역에서 내용기반 동적 대역폭 조절 방법의 절차를 요약하면 다음과 같다.

- ① 압축된 동영상에서 밝기(luminance) 성분의 DCT 계수를 추출
- ② 대표 불연속의 방향을 예측하고 DCT 영역에서 회전 및 반사 수행(실험에서는 4 방향 사용)
- ③ 대표 불연속이 DCT 블록 중앙에 오도록 식 (3), (4)를 사용하여 정렬
- ④ 대표 불연속의 높이를 식 (5)를 사용하여 계산
- ⑤ DCT 블록을 목표 대역폭에 맞도록 파라미터 α, β 를 조정하고 높이에 따라 트랜스코딩
- ⑥ 영상의 모든 블록을 위의 절차에 따라 처리

이 방법을 수행하기 위하여 입력된 영상 스트림의 인트라 프레임은 부분적으로 디코딩(VLD 및 IQ)되며, 전송 대역폭을 줄이기 위하여 차단 함수 $C(h)$ 를 계산하여 필터링을 수행한다. 이 때 스트림의 헤더와 움직임 벡터는 변화 없이 저장되며 보상기법[5]를 사용하여 연속된 오류(drift error)를 방지할 수 있으며 매우 빠르게 처리할 수 있으나, 참조 프레임(I 프레임)의 화질이 감소했으므로 보상된 프레임(B, P 프레임)의 화질도 저하되며, 저하 정도는 예측(prediction) 방향, 필터링 정도, MC-DCT(Motion Compensation-DCT) 계수 크기에 따라 달라진다. 이 방법은 시간적 트랜스코딩 방법과 결합하면 다양한 대역폭에 적용할 수 있다.

목표(송신) 대역폭 B_{out} 은 그림 4의 축소율 $R(\alpha, \beta)$ 과 입력 대역폭 B_e 를 참고하여 $R(\alpha, \beta) = B_{out}/B_e \times 100$ 식에서 예측할 수 있다. 그림 4의 (a)는 $\beta=64$ 일 때 α 를 $[0, 255]$ 범위에서 변화시킬 때의 100개의 동영상에서 얻은 평균 축소율을 나타내고 있다. α 가 증가할수록 평균 송신 대역폭이 거의 선형적으로 단조 감소함을 나타내고 있다. 그림 4의 (b)는 $\alpha=255$ 일 때 β 가 $[64, 1]$ 범위에서 점차 감소할 때의 송신 대역폭 감소를 나타낸다(실선).

점선으로 표시된 기존의 고정값 필터링은 일정 영역까지 대역폭 감소가 거의 없다가 $20 \geq \beta \geq 1$ 영역에서 급격히 변화함을 보이고 있어 동적인 대역폭 조절이 어려움을 나타내고 있다. 반면에 DCT 블록 내용의 특징을 바탕으로 한 동적 트랜스코딩에서는 파라미터 $0 \leq \alpha \leq 255$ 와 $64 \geq \beta \geq 1$ 를 조절함으로써 송신 대역폭 $B_{out} = B_e R(\alpha, \beta)$ 를 거의 선형적으로 부드럽고 섬세하게 조절할 수 있다.

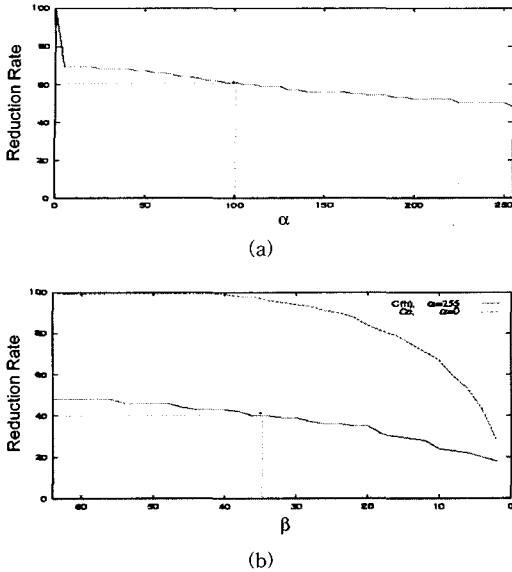


그림 4 차단 함수의 파라미터 α, β 변화에 따른 평균 축소율 : (a) $C(h)$ 의 α 변화에 따른 축소율($\beta=64$), (b) 고정값 C_c 의 변화에 따른 압축율($\alpha=0$)과 $C(h)$ 의 β 변화에 따른 축소율($\alpha=255$)

예로써 송신 대역폭의 60% 축소를 위하여 그림 4(a)의 표시와 같이 $\alpha=100, \beta=64$ 를 선택할 수 있고, $\alpha=255, \beta=35$ 를 사용하여 40%의 축소율을 얻을 수 있다. 시간에 따른 전체적인 대역폭은 과거 일정 시간의 목표대역폭과 축소대역폭의 오차를 누적하여 현재의 목표대역폭을 변화시킴으로서 균일화(regularization) 시킬 수 있다.

5. 실험(Experiments)

본 연구에서 제한한 내용 기반 동적 대역폭 조절 방법을 평가하기 위하여 방송에서 취득한 두 개의 동영상을 대상으로 실험하였다. 동영상은 (I B B P B B P B B P B B P B B) 프레임 형식의 MPEG-1 PES (Packetized Elementary Stream, 352x240, 29.97fps, 1743kbps)로 압축되었다. 각 동영상은 50초 분량으로 1500 프레임 중 100 개의 I-프레임을 가지고 있으며, DCT 영역에서 불연속의 회전 시간을 줄이기 위하여 2 방향(90°, 180°)을 검중하여 4 방향을 회전시켰으며, FDCT 및 IDCT는 Loeffler[17]의 알고리즘을 적용하였다. 시험의 대역폭 및 화질의 비교는 각각 100개의 I-프레임을 대상으로 하였다.

영상의 시각적인 질을 평가하기 위하여 RMSE와 PSNR을 사용하였다.

$$RMSE = \frac{[\sum_{i,j} \{f_c(i,j) - f(i,j)\}^2]^{1/2}}{N}, \tag{9}$$

$$PSNR = 20 \log_{10}(\frac{255}{RMSE}), \tag{10}$$

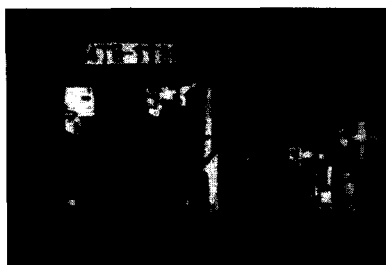
여기에서 원래의 영상을 $f(i, j)$, 재구성(트랜스코딩)된 영상을 $f_c(i, j)$ 로 나타내었으며, RMSE의 작은 값과 PSNR의 큰 값이 재구성된 영상화질의 좋음을 나타낸다.

고정 차단 값 $C_c=4$ 을 사용한 식 (6)의 트랜스코딩과 차단 함수 $C(k; \alpha, \beta)$ 를 사용한 식 (8)의 내용기반 트랜스코딩을 비교하기 위하여 그림 5의 음악 동영상과 그림 6의 스포츠 동영상을 사용하였다. 각 그림의 (a)는 원래의 영상을 나타내며, (b) 영상은 고정 차단 값의 필터를 사용하여 트랜스코딩한 영상을 나타낸다. 이 영상은 블록의 내용과 무관한 고정값을 사용하였기 때문에 명암 대비가 뚜렷한 블록에서 번짐 현상이 발생하였다.

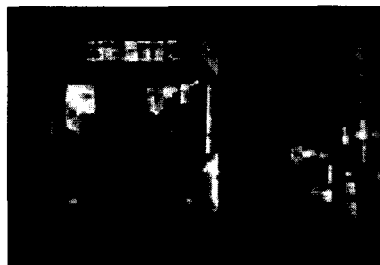
실험에서 각 필터에 의하여 트랜스코딩된 스트림의 대역폭을 동일하게하기 위하여 그림 5에서는 $\alpha=30, \beta=8$ 의 파라미터를 선정하였고, 그림 6에서는 $\alpha=70, \beta=8$ 를 선정하였다. 각 그림의 (c) 영상은 각각 DCT 블록의 대표적인 불연속의 크기를 3장에서 설명한 방법에 의하여 찾아 밝기 값으로 표시하였다. 처리시간의 향상을 위하여 방향 검중을 4방향으로 제한하였기 때문에 기울어진 불연속 일부가 잘 검출되지 않았지만 대부분 적절하게 표시되고 있음을 나타낸다. 그림 (c)의 크기와 설정된 파라미터를 사용하여 내용기반 트랜스코딩된 결과 영상을 각 그림의 (d) 영상에 나타내었다. 내용기반 트랜스코딩은 영상내부의 특징정보를 사용하였기 때문에 결과 영상처럼 화질의 많은 저하를 막을 수 있었다.

그림 5의 트랜스코딩된 영상의 화질을 나타내는 측정값을 표 1에 나타내었다. 맨 위의 행은 고정 차단 값 $C_c=\beta=4, 8$ 을 사용하여 트랜스코딩된 결과의 화질 측정값을 나타내고, 다른 행들은 다양한 α, β 값을 사용한 내용 기반 트랜스코딩 영상의 화질 측정값을 나타낸다. 동적 필터링의 모든 RMSE 결과는 고정 필터링의 결과보다 10% 정도 감소된 결과를 나타내고 있으며, 동적 필터링의 몇 개의 파라미터 쌍 ($\alpha=30, \beta=8$), ($\alpha=70, \beta=12$), ($\alpha=100, \beta=16$)은 고정 필터링과 대역폭은 유사하나 향상된 PSNR을 보여준다.

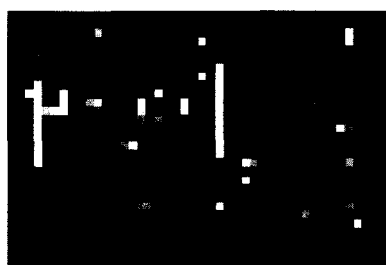
β 가 AC 주파수의 개수를 차단하므로 β 를 크게 하면 대표 불연속이 큰 블록은 많은 주파수로 세밀하게 표현되며, 일정한 대역폭을 유지하기 위해서 불연속이 작은 블록은 더 많이 왜곡된다. 반대로 β 를 작게 하면 특징이 뚜렷한 블록도 약간의 왜곡이 생기지만 다른 블록들의 왜곡도 줄어들게 된다. 이러한 상관성은 PSNR과 같은 측정값으로는 표현되지 않으나 시각적으로 관찰되어 진다.



(a) 원래의 영상



(b) 기존 방법의 고정값($C_c=4$)에 의하여 트랜스코딩된 영상



(c) DCT 블록의 대표적인 불연속 크기(밝을수록 높은 불연속을 나타냄)

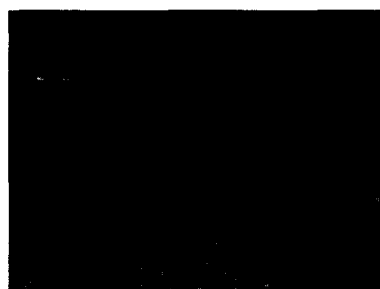


(d) 내용기반 방법에 의하여 트랜스코딩된 영상 ($\alpha=30, \beta=8$)

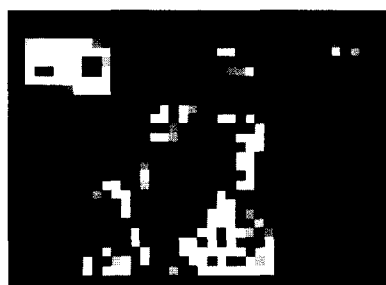
그림 5 음악 동영상의 트랜스코딩



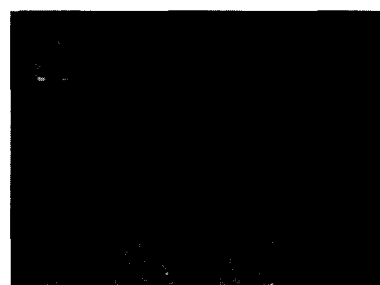
(a) 원래의 영상



(b) 기존 방법의 고정값($C_c=4$)에 의하여 트랜스코딩된 영상



(c) DCT 블록의 대표적인 불연속 크기(밝을수록 높은 불연속을 나타냄)



(d) 내용기반 방법에 의하여 트랜스코딩된 영상 ($\alpha=70, \beta=8$)

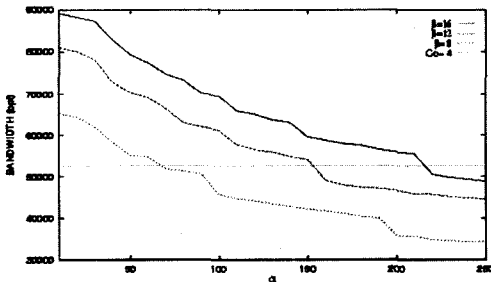
그림 6 스포츠 동영상의 트랜스코딩

표 1 100개의 I-프레임에서 파라미터 α, β 변화에 따른 평균 대역폭(bit per frame), PSNR(dB), RMSE

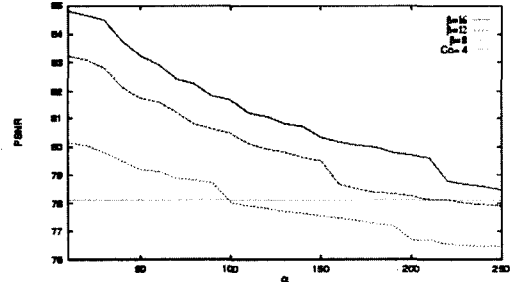
α	β	대역폭	PSNR	RMSE	α	β	대역폭	PSNR	RMSE
0	4	64401	78.22	0.0406	0	8	89398	81.22	0.0311
10	8	66959	79.38	0.0347	5	16	89036	82.39	0.0268
20	8	65972	79.20	0.0352	10	16	88193	82.30	0.0270
30	8	63420	78.82	0.0364	20	16	87034	82.08	0.0275
50	12	68388	79.47	0.0343	40	24	91435	82.55	0.0265
60	12	66310	79.20	0.0352	50	24	86794	81.81	0.0281
70	12	62606	78.66	0.0370	60	24	84269	81.45	0.0289
80	16	68728	79.42	0.0345	60	32	92000	82.51	0.0266
90	16	64999	78.89	0.0362	70	32	88451	81.89	0.0279
100	16	64436	78.77	0.0366	80	32	85870	81.52	0.0288

표 2 MPEG 압축된 1500 프레임에서의 한 프레임 평균 트랜스코딩 처리시간(ms): T_r 는 읽고 패킷을 분해하는데 필요한 처리 시간, T_c, T_q, T_a, T_s 는 각각 기존의 고정값 필터링 처리시간, 재 양자화 처리시간, 내용기반 필터링 처리시간, IDCT+FDCT 처리시간을 나타냄.

비디오 타입	T_r	T_c	T_q	T_a	T_s
음악(352x240)	5.43	37.95	53.28	85.65	367.41
스포츠(320x240)	5.24	32.80	47.75	79.52	342.39



(a) 대역폭의 변화



(b) PSNR의 변화

그림 7 파라미터 α, β 변화에 따른 평균 대역폭과 PSNR

α, β 의 상관성을 직관적으로 분석하기 위하여 그림 6의 영상을 대상으로 파라미터 α, β 의 변화에 따른 화질 측정값의 변화를 그림 7에 표시하였다. 그림 7의 (a)는 다양한 α, β 값에 따른 대역폭(bits per frame)을 나타내었으며, 수평 점선은 고정 필터링을 수행했을 경우의 대역폭을 나타낸다. 도표에서는 특징이 큰 블록을 세밀하게 표현하기 위하여 동적 필터링의 β 를 C_r 보다 2~4배 크게 설정하였다. 그림 7의 (b)는 α, β 의 변화에 따른 PSNR의 변화를 보여 주며, 수평 점선은 고정 필터링에 의하여 트랜스코딩된 영상의 PSNR을 나타낸다.

그림 7의 (a)와 (b)를 비교 분석하면 같은 대역폭에서는 각각의 β 에서 동적 필터링에 의한 트랜스코딩 결과가 높은 PSNR을 가지며, 같은 PSNR에서는 적은 대역폭을 가지는 것을 알 수 있다.

내용기반 동적 필터링 방법의 처리 시간은 표 2에서 나타난 바와 같이 고정 필터링에 의한 방법보다 크게 증가하지 않는다. 일반적인 UNIX 워크스테이션에서 각 동영상의 1500 프레임을 처리하여 한 프레임으로 평균 한 시간을 기능별로 T_r, T_c, T_q, T_a, T_s 등으로 나타내었다. T_r 는 읽고 패킷을 분해하는데 필요한 처리 시간, T_c, T_q, T_a 는 DCT 영역에서 각각 기존의 고정값 필터링 처리시간, 재 양자화 처리시간, 내용기반 필터링 처리시간을 나타내며, T_s 는 픽셀 영역으로 변환하여 처리 후 다시 DCT 영역으로 변화하는데 필요한 IDCT+FDCT 처리시간을 나타낸다. 352x240 크기의 영상에서 내용기반 동적 필터링 처리시간은 85.65ms로써 기존의 고정 필터링 방식의 처리시간 37.95ms나 재 양자화 방

법의 처리시간 53.28ms보다 많지만 IDCT를 수행하고 내용기반 필터링 수행 후 다시 FDCT를 수행해야하는 공간영역의 처리(367.41ms 이상)보다는 훨씬 빠르며, 보통의 MPEG 압축에서 I 프레임이 초당 2 프레임씩 설정되는 것을 감안하면 실시간 처리가 가능한 시간이다.

6. 결론(Conclusion)

실시간 대역폭 조절을 위한 다양한 트랜스코딩 방법이 제안되었지만 압축된 스트림이 가지고 있는 공간상의 정보를 활용한 방법은 제안되지 않았다. 본 연구에서는 압축된 DCT 영역에서 블록 내부의 대표적인 불연속 특징을 찾고 이를 기반으로 동적으로 목표 대역폭에 따라 스트림을 트랜스코딩하는 내용기반 동적 대역폭 조절 방법을 제안하였다. 이 방법은 효율적인 처리시간으로 실시간 응용분야에 적용될 수 있으며, 목표대역폭의 세밀한 조절이 가능하고, 같은 대역폭에서 기존 방법에 비해 좋은 화질을 유지함을 실험에서 나타내었다.

대역폭과 트랜스코딩 파라미터 간의 정밀한 연관관계 정립을 위한 연구가 보완되면, DCT 영역에서의 특징 분석 및 동적 대역폭 조절 방법 등이 실시간 응용 기술에 다양하게 변형되어 응용될 수 있을 것이다.

참 고 문 헌

[1] N. Chaddha, "A software only scalable video delivery system for multimedia applications over heterogeneous networks," Proc. International Conference on Image Processing, Washington, DC, pp.404-407, Oct. 1995.

[2] J. Moura, R. Jasinschi, H. Shiojiri-H. and C. Lin, "Scalable video coding over heterogeneous networks," Proc. of SPIE, Vol. 2602, pp. 294-306, 1996.

[3] M. Ghanbari, "Two-layer coding of video signals for VBR networks," IEEE J. Select. Areas Commun., Vol. 7, pp.771-781, 1989.

[4] ISO/IEE 13818-2, "Information technology-generic coding of moving pictures and associated audio information-Part 2: Video," 1995.

[5] P.A.A. Assuncao and M. Ghanbari, "A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit stream," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8, No. 8, pp. 953-967, Dec. 1998.

[6] N. Bjork and C. Christopoulos, "Transcoder Architectures for Video Coding," IEEE Transactions on Consumer Electronics, Vol. 44, No. 1, pp.88-98, 1998.

[7] Y. Nakajima, H. Hori and T. Kanoh, "Rate conversion of MPEG coded video by re-quantization process," International Conference on Image

Processing, 1995.

[8] O. H. Werner, "Requantization for Transcoding of MPEG-2 Intraframes," IEEE Transactions on Image Processing, Vol. 8, No. 2, pp. 179-191, 1999.

[9] JunXin, Ming-Ting Sun, and Kangwook Chun, "Bit-Allocation for Transcoding of Pre-Encoded Video Streams," Visual Communications and Image Processing, pp.164-171, 2002.

[10] Ligang Lu, Shu Xiao, Jack L. Kouloheris, Cesar A. Gonzales, "Efficient and Low Cost Video Transcoding," Visual Communications and Image Processing, pp.154-163, 2002.

[11] Francisco Garcia, David Hutchison, Andreas Mauthe, and Nicholas Yeadon, "QoS Support for Distributed Multimedia Communications," Proceedings of IFIP/ IEEE International Conference on Distributed Platforms, Dresden, Germany, Feb 1996.

[12] Nicholas Yeadon, Francisco Garcia, David Hutchison, and Doug Shepherd, "Continuous Media Filters for Heterogeneous Internetworking," Proceedings of SPIE - Multimedia Computing and Networking (MMCN'96), 1996.

[13] Bo Shen and Ishwar K. Sethi, "Inner-Block Operations On Compressed Images," ACM Multimedia '95, pp. 489-498, 1995.

[14] Brian C. Smith and Lawrence A. Rowe, "Algorithms for Manipulating Compressed Images," IEEE Computer Graphics & Applications, pp. 34-42, 1993.

[15] TaeYong Kim and Joon Hee Han, "Model-based Discontinuity Evaluation in the DCT Domain," Signal Processing, Vol. 81, No.4, pp. 871-882, 2001.

[16] Robert M. Haralick and Linda G. Shapiro, "Computer and Robot Vision," Addison Wesley, pp. 605-606, 1992.

[17] C. Loeffler, A. Ligtenberg and G. Moschytz, "Practical Fast 1-D DCT Algorithms with 11 Multiplications," Proc. Int'l. Conf. on Acoustics, Speech, and Signal Processing, pp. 988-991, 1989.



김 태 용
 1986년 2월 한양대학교 전기공학과(공학사). 1988년 2월 한양대학교 전자통신공학과(공학석사). 1998년 2월 포항공과대학교 컴퓨터공학과(공학박사). 1988년 3월~1999년 2월 한국통신 운용연구단 연구원. 2003년 4월~현재 한국컴퓨터게임학회 총무이사. 2003년 9월~현재 중앙대학교 첨단영상대학원 영상공학과 부교수. 관심분야는 영상통신, 영상처리, 컴퓨터비전, 컴퓨터게임



최 중 수

1975년 2월 인하공대 전기공학과(공학사). 1977년 2월 서울대학교 전자공학과(공학석사). 1981년 2월 Keio University (일본) 전기공학과(공학 박사). 1981년 2월~1981년 9월 Aloka Co. Ltd.(일본) 연구소 연구원. 1981년 9월~1999년 8월 중앙대학교 전자공학과 교수. 1999년 9월~현재 중앙대학교 첨단영상대학원 영상공학과 교수, 대학원장. 1988년~1992년 한국통신학회 영상통신연구회 위원장 역임. 1990년~1992년 대한전자공학회 화상처리 연구회 위원장 역임. 1987년 영상처리 및 이해에 관한 워크샵 대회장 역임. 1993년~현재 한국통신학회 학술,편집,총무이사, 부회장. 관심분야는 영상정보, 컴퓨터비전, 증강현실