

멀티미디어 네트워킹을 위한 압축 신호상에서 동영상 처리: 압축 동영상 비트율 변환

정회원 황 대 환*, 조 규 섭**, 황 수 용***

Manipulation of the Compressed Video for Multimedia Networking : A Bit rate Shaping of the Compressed Video

Dae-Hwan Hwang*, Kyu-Seob Cho**, Su-Young Hwang*** *Regular Members*

요 약

유무선 통신망, 지상파, 위성 등을 포함한 다양한 이종 네트워크 환경하에서의 멀티미디어 서비스는 급속한 인터넷 보급과 병행하여, 이전과는 전혀 다른 새로운 통합 멀티미디어 네트워킹 인프라를 요구하고 있다. 네트워크와 멀티미디어 유형에 관계 없이 사용자에게 이음새 없는 투명한 멀티미디어 통신서비스를 제공하자는 기술적 요구는 국내외에서 GII를 축으로 연구가 진행되고 있으나, 아직 뚜렷한 형상이 완성되지 않은 실정이다. 이 GII의 배경에는 이종 네트워크, 이종 미디어 그리고 다양한 이종 프로토콜이 포함되는 응용 서비스가 상호 투명하게 연결되어야 하는 복잡한 기술적 측면이 있으며, 멀티미디어의 네트워킹은 가장 핵심적인 연구항목이 되고 있다. 본 논문에서는 통신 대역폭의 QoS를 보장할 수 없는 인터넷 그리고 유무선을 통한 디지털 TV 방송에 직접 적용할 수 있는 압축 동영상 비트율 변환 알고리즘을 제안한다. 동영상 비트율 변환은 상기의 네트워크 환경하에서 동영상 스트리밍 서비스를 위해 필요한 기술이다. 기존의 비트율 변환은 해당되는 압축 동영상 디코더와 인코더를 쌍으로 연결하여, 선형 PCM 신호로 변환하고, 목적하는 비트율로 맞추어 주는 방법이 일반적이거나, 그 구현의 복잡성, 수반되는 처리지연 등으로 실시간 멀티미디어 네트워킹에 단점을 갖고 있다. 본 논문에서 제안한 알고리즘은 화질과 구현의 복잡도를 고려하여 비트율 변환을 압축영역에서 행하며, 특히 실시간 처리에 주안점을 두었다. 제안한 알고리즘은 ISO/IEC의 MPEG-2 동영상을 대상으로 소프트웨어로 구현하였으며, 화소영역의 처리방법을 포함하는 기존의 비트율 변환방법에 비해 구현의 복잡도, 처리시간에 장점을 갖는다.

ABSTRACT

Interoperability and inter-working in the various network and media environment with different technology background is very important to enlarge the opportunity of service access and to increase the competitive power of service. The ITU-T and advanced countries are planning ahead for provision of GII enabling user to access advanced global communication services supporting multimedia communication applications, embracing all modes of information. In this paper, we especially forced the heterogeneity of end user applications for multimedia networking. The heterogeneity has several technical aspects, like different medium access methods, heterogeneous coding algorithms for audio-visual data and so on. Among these elements, we

* 한국전자통신연구원(hdh@etri.re.kr)

**성균관대학교

***서울산업대학교

논문번호 : 010167-0709, 접수일자 : 2001년 7월 9일

*본 연구는 정보통신부 출연과제의 일환으로 수행되었습니다.

have been itemized bit rate shaping algorithm on the compressed moving video. Previous manipulations of video has been done on the uncompressed signal domain. That is, compressed video should be converted to linear PCM signal. To do such a procedures, we should decode, manipulate and then encode the video to compressed signal once again. The traditional approach for processing the video signal has several critical weak points, requiring complexity to implement, degradation of image quality and large processing delay. The bit rate shaping algorithm proposed in this paper process the manipulation of moving video on the completely compressed domain to cope with above deficit. With this algorithms, we could realized efficient video bit rate shaping and the result of software simulation shows that this method has significant advantage than that of pixel oriented algorithms.

I. 서론

멀티미디어 네트워킹은 사용자가 만족하는 품질을 제공하며, 통신하도록 하는 것이다. 이 멀티미디어 네트워킹에 대한 연구는 네트워크의 유형에 관계없이 사용자에게 서비스를 제공하겠다는 Global Information Infrastructure(GII)로 이어지고 있으며, 이것은 미들웨어를 정의하는 것으로 개략적인 모습을 보이고 있다[1-5]. 그러나 사용자의 서비스를 이음새 없이 연결하는 중재자로서의 미들웨어를 실현하기 위해서는 네트워크 측면에 두는 비중 못지 않게 멀티미디어가 갖고 있는 특성을 이해하고, 여기에 내재된 이질성을 해결하여야 한다[6-9].

멀티미디어는 크게 오디오, 영상 그리고 데이터 등의 모노 미디어로 분류되며, 여기서 좀더 깊게 들어가면 코딩 방법, 목적 비트율 그리고 표현 방법에 따라 같은 미디어에서도 수 많은 형태로 분기된다. 또한 각각은 고유한 특성을 지녀, 서로 호환이 불가능하여, 멀티미디어 네트워킹은 단순히 음성에 대한 전화 서비스만을 제공하였던 때와는 전혀 다른 양상을 갖는다. 서로 다른 형태의 미디어가 네트워크에 혼재할 때, 이것은 상호통신이 불가능한 문제를 낳게 함으로, 미디어간의 이질성을 해결하는 것은 멀티미디어 네트워킹의 핵심기술이라 할 수 있다. 본 논문은 멀티미디어 네트워킹에 수반되는 미디어의 이질성을 해결하기 위한 내용으로 다양한 미디어 중 압축 동영상을 대상으로 통신 대역폭이 다른 이종 네트워크 또는 인터넷과 같이 일정한 통신대역을 보장할 수 없는 환경하에서 직접 적용할 수 있는 실시간 비트율 변환 알고리즘에 대해 제안 한다.

압축 동영상의 비트율 특성은 사용되는 코딩방식과 실제 영상의 내용 등에 영향을 받으며, 이것은 움직임의 큰 변화가 없는 영상전화, 영상회의형과

TV 방송의 광고 그리고 영화와 같이 급격한 장면 및 움직임 변화가 있는 형태로 나누어 볼 수 있다. 동영상 압축 비트율에 영향을 미치는 또 다른 요소는 실제 영상을 압축하는 인코더의 구조이다. 이것은 현재의 동영상 부호방식이 DCT와 움직임 보상 알고리즘을 핵심으로 하는 하이브리드 방식으로 DCT는 8x8 블록 단위로 처리되며, 움직임 보상은 4개 블록의 휘도 성분과, 2개의 색차 성분으로 구성된 매크로블록을 기준으로 동작하는 동일한 구조를 갖지만, 특정한 상위 계층의 부호화 알고리즘에 따라 출력 비트율의 형태가 달라진다.

ITU-T의 동영상 압축 표준인 H.263은 최초의 인트라 프레임 이후, 이전 영상과의 차이값과 움직임 정보로 표현되는 인트라 프레임을 반복하여 생성하는데 반해 ISO/IEC의 MPEG-2는 고정된 수의 I, P, B 프레임으로 이루어진 Group of Picture(GOP)의 구조를 갖으며, 일정한 주기로 I 프레임과 순방향, 양방향 움직임 정보를 각각 포함하는 P, B 프레임을 생성한다.

압축 알고리즘에 따르는 비트율 특성을 살펴 보면 H.263은 디코딩된 화면의 예측에러 누적을 방지하기 위해 최소 132 프레임 당 하나의 인트라 매크로블록을 강제로 인코딩하는 제한이 있으나 입력 영상의 내용을 정확히 표현하는 출력 비트율 특성을 갖는다. VCR과 유사한 기능을 위한 영상처리를 위해 GOP와 같은 독특한 계층구조를 갖는 MPEG-2는 입력 영상의 내용 이외에 주기적으로 반복되는 프레임 구조에 의해 출력 비트율의 형태가 결정된다. 따라서 동일한 내용의 영상을 대상으로 인코딩하여도 H.263의 스트림과 다른 형태의 출력 비트율 특성을 갖는다.

본 절에서는 고정된 비트율로 인코딩된 압축 동영상을 다른 비트율로 변환하는 알고리즘에 대해 기술하기 앞서 연구의 대상이 되는 동영상 비트율의 특성을 살펴본다. 인코더에서 생성되는 압축 비

트율은 본질적으로 가변 비트율의 특성을 가지며 이것은 공간적, 시간적 잉여 정보를 제거하는 하이브리드 코딩 알고리즘을 사용한 결과이다. 그러나 고정된 통신대역을 통해 동영상 전송하기 위해서는 압축부호를 일정하게 생성하여야 하며, 이것은 시간축상에서 비트율의 변동이 큰 부호량을 완충버퍼를 사용하여 제어하고 네트워크로 일정하게 송신함으로써 가능하다. 이것을 항등 비트율 동영상이라 한다. 동영상 비트율 변환 알고리즘을 연구하기 위해서는 대상이 되는 압축부호의 비트율 특성을 알아야 하며, 본 장에서는 이것을 Variable Bit Rate(VBR), Constant Bit Rate (CBR) 그리고 코딩 알고리즘에 따르는 세부 파라미터의 관점에서 분석한다.

인코더에서 생성되는 압축부호의 발생량을 제어하지 않고 그대로 출력하는 것을 가변 비트율 동영상이라 하며, 이와 같은 가변 비트율의 특성을 갖는 압축부호는 출력율을 제어하는 양자화 파라미터 값을 고정하여 얻을 수 있다. VBR 영상은 CBR 영상에 비해 다음과 같은 차이점이 있다. 인코더 및 디코더에 완충버퍼를 사용하지 않으므로 처리지연이 작다. 전체 스트림에 동일한 양자화 값을 사용하므로 화질이 우수하다. VBR 특성상 시간축상에서 비트율의 변동폭이 매우 크다. 그러나 동일한 화질에 기준한 평균 비트율은 CBR에 비해 낮은 값을 갖는다.

VBR은 CBR에 비해 많은 장점을 가져, VBR의 비트율 특성을 정확히 모델링 할 수 있다면, 네트워크 대역할당의 관점에서 이점을 갖으나, 아직 정확한 VBR 모델이 정립되어 있지 않고, 대부분의 네트워크가 항등 비트율의 대역을 제공하므로 주로 CBR 영상이 통신에 사용되고 있다. 일반적으로 통신에 CBR이 사용됨에 따라 본 연구에서도 비트율 변환의 대상을 CBR 동영상으로 정하였다. 그러나 VBR의 낮은 평균 비트율에서 일정한 화질을 갖는 장점은 비트율 변환의 기본 목적이 되며, VBR 동영상의 세부 특성은 비트율 변환에 중요한 분석자료 및 기준 척도로 사용할 수 있다.

프레임을 건너 뛰는 특수한 경우를 제외하면, CBR의 울제어는 매크로블럭 단위로 적용되는 양자화 파라미터에 의해 제어된다. 즉 임의의 시점에 부호 발생량이 과도할 때 양자화 스텝 크기를 증가시켜 부호 발생량을 억제하고, 반대의 경우에는 양자화 스텝 크기를 감소시켜 순간적인 부호 발생량을 증가시킨다. 양자화 파라미터 제어를 통한 울제어는

CBR 영상이 갖는 매크로블럭 단위의 양자화 스텝 값이 변동됨을 의미하며, 이는 인코딩되는 영상의 화질이 울제어에 따라 매크로블럭 단위로 가변되는 결과를 낳는다.

본 논문은 현재 주로 사용되는 CBR 영상을 비트율 변환의 대상으로 삼았다. 연구의 목적은 임의의 비트율로 기 압축된 동영상을 압축영역에서 낮은 비트율로 변환하는 것이며, 최적의 화질을 유지하며, 낮은 비트율로 변환시키기 위해서는 화질에 절대적인 영향을 미치는 양자화 파라미터의 합당한 제어가 필요하며, 이것은 대상이 되는 압축영상의 비트율 특성을 정확히 파악하여야 한다. 본 절은 연구의 주제가 되는 CBR 영상이 갖는 비트율 특성과 비트율 변환의 주요 요소가 되는 울제어와 양자화 과정에 대해 설명하였다.

II. 기존 연구

고정된 비트율로 인코딩된 동영상을 낮은 비트율로 변환하는 방법은 압축부호의 영역별로 일부 알고리즘이 연구되었다. 변환방법은 부호에 대한 처리가 전혀 필요없는 완전한 압축부호인 Fixed Length Code (FLC) 영역에서, 디코딩, 인코딩의 전 과정을 거치는 화소영역까지로 그 처리의 단계가 구분된다. 본 절에서는 비트율 변환에 대한 기존연구를 영역별로 세분하여 설명한다[14-23].

FLC 영역의 비트율 변환방법은 동영상을 구성하는 프레임 중 일부를 제거 (dropping) 하는 것이 주로 사용된다. MPEG, H.263의 I, P, B 프레임 중 처음에는 다른 프레임에 영향을 전혀 미치지 않는 B 프레임을 제거하고, 더 낮은 비트율을 요구하는 경우는 P 프레임을 차례로 삭제하는 방법을 사용한다. 이 방법은 압축부호의 프레임 유형 정보를 갖고 있는 FLC만을 분석하고 해당하는 부분을 제거하는 단순한 형태로 구현이 가능하나, 비트율이 프레임 단위로 조절되므로 정확한 울제어가 불가능하며, 특히 디코더의 관점에서 프레임이 제거된 것에 대응되게 부호를 디코딩하여 표시하는 기능이 필요하다. 즉 화면에 정상적으로 영상을 표현하기 위해서는 초당 30 프레임의 변화율을 만족하여야 한다. 따라서 프레임이 제거되어 30 프레임을 채우지 못할 경우는 디코더 또는 영상 출력부에서 모자라는 화면을 이전 프레임으로 대체해야 한다. 따라서 이 기능이 구현되어 있지 않은 경우는 화면의 명멸 (flicker) 또는 화면이 빨리 표현되는 (fast forward)

현상이 나타난다. 즉 프레임 제거를 통한 비트율 변환은 범용 컴퓨터의 소프트웨어로 구현된 디코더에서 인터넷 상의 동영상을 트래픽 변동폭에 맞게 수신하는 응용에는 어느 정도 응용이 가능하나, 그 외 디지털 TV 수신 등을 위한 비트율 변환에는 합당한 방법이 아니다.

DCT 블럭에 대한 계수의 형태 및 위치 정보는 VLC를 해석해 알 수 있다. VLC 영역의 비트율 변환 알고리즘은 VLC 분석을 거쳐 DCT 계수를 조작하는 방법으로 4개의 휘도블럭과 2개의 색차블럭으로 구성되는 매크로블럭 중 색차블럭을 제거하여 컬러영상을 흑백영상으로 변환하여 비트율을 조절하는 방법이 있다. 그 외 컬러 정보는 그대로 유지되거나 DCT의 AC 계수를 삭제하여 비트율을 변환하는 방법으로 첫번째는 저주파에서 고주파의 성분 순으로 배열된 DCT 블럭의 배열 중 고주파에 해당하는 AC 부분을 제거하는 것과 두번째는 AC 성분 중 특정 값 이하의 AC 계수를 삭제하는 방법이다. 이것 역시 간단한 조작으로 비트율 변환이 가능하다는 장점을 갖으나 특정 비트율에 맞는 정확한 율제어는 불가능하며 화질의 변동폭이 크다는 결점을 갖는다.

재양자화를 통한 비트율 변환 방법은 VLD를 거치고 DCT 블럭을 역양자화 한 후, 재양자화, VLC를 하는 것으로 비교적 용이한 조작으로 정확한 비트율을 맞추어 줄 수 있는 장점을 갖는다. 실제 동영상을 인코딩할 때, 율제어는 프레임을 삭제하거나, 양자화 스텝의 크기를 조절하는 방법을 주로 이용한다. 이중 프레임 삭제는 일반적인 방법이 아니며, 후자인 양자화 처리는 동영상 율제어의 핵심이라 할 수 있다. 따라서 양자화 데이터 이상의 영역에서의 비트율 변환도 재양자화를 사용하는 것이 기본임을 감안하면 이 영역에서의 비트율 변환은 구현의 복잡성과 비트율 변환에 따르는 화질과의 절충점에 있다고 할 수 있다. 결국 재양자화를 통한 비트율 변환은 특정한 엔트로피를 갖고 있는 데이터를 최소의 비트로 최적의 화질을 갖도록 하는 양자화 파라미터 값을 찾는 것이라 할 수 있다.

동영상 인코더에서의 양자화 과정은 화질과 비트율을 결정하는 중요한 요소이나 영상의 내용, 움직임 정보와 기타 복잡한 인코딩 파라미터가 혼재하는 상황에서 최적의 양자화 파라미터 값을 찾는 것은 매우 복잡한 과정이 필요하다. 반면에 이미 생성된 압축 동영상을 대상으로 하는 양자화는 기존의 양자화 파라미터 값, 움직임 벡터, 프레임에 할당된

비트 수 등의 사전 정보를 충분히 이용할 수 있으므로 인코더의 경우보다는 최적의 양자화 파라미터 선정 과정이 용이하다. 현재 압축 동영상을 양자화 절차를 통해 낮은 비트로 변환하는 알고리즘은 단순히 양자화 파라미터를 비례식에 의해 결정하는 방법과 *lagrangian cost function*을 이용하는 것이 있으나 전자는 정확한 비트율을 맞추어 주는데 문제점을 갖고, 후자는 이론적으로 최적의 결과를 얻을 수 있으나 수반되는 절차가 매우 복잡하고, 많은 처리시간이 소요되는 맹점이 있다. 재양자화에 의한 비트율 변환은 구현의 복잡성과 화질간의 절충점이 될 수 있으며, 용이한 방법으로 최적의 양자화 파라미터를 선정하는 알고리즘이 제안되었을 경우 이상적인 비트율 변환 방법이 될 수 있다. 따라서 이에 필요한 알고리즘의 연구는 중요한 의미를 갖는다.

MC-DCT 영역에서의 비트율 변환은 DCT 상에서 움직임 보상을 수행하여 보다 정확히 DCT 계수 값을 보정하고 이를 다시 양자화, VLC 하는 방법이다. 움직임 벡터는 변화없이 그대로 사용하거나 DCT 영역에서 움직임 추정을 하여 새로운 값을 생성할 수 있다. 이것은 화소영역의 처리방법에 비해 DCT, DCT-1의 처리가 필요없고, 움직임 보상된 DCT를 사용하여 비트율을 변환하므로 낮은 비트율에서도 일정한 화질을 유지할 수 있으며, 움직임 추정에 DCT 정보를 이용하여 적은 연산만으로 벡터를 생성할 수 있는 장점을 갖는다. 그러나 인접한 영상과 연관성을 갖는 프레임 단위의 동작을 함으로써 화소영역과 동일한 처리지연이 발생하는 실시간성의 문제와 생성된 움직임 정보는 화소영역에 비해 정확성이 떨어지는 단점을 갖는다. 즉 이 방법은 화소영역에 비해 일부 연산과정이 단순해지는 장점을 갖으나 구현의 복잡성, 화질, 지연과 같은 비트율 변환에 대한 성능평가의 척도가 되는 요소를 기준으로 볼 때, 실시간 멀티미디어 통신에 합당한 비트율 변환 알고리즘이 아니다.

화소영역의 비트율 변환은 DCT, DCT-1까지를 포함한 완전한 디코딩 인코딩 처리를 거쳐 요구되는 비트 변환을 수행한다. 화소영역에서의 비트율 변환은 디코더와 인코더를 짝으로 연결하여 필요한 기능을 구현할 수 있는 가장 일반화된 방법이나, 구현의 복잡도가 가장 높은 단점을 갖는다.

III. 압축영역에서의 실시간 비트율 변환

본 장에서 제안하는 비트율 변환 알고리즘의 목

적은 압축 동영상은 최적의 화질을 갖고 목적하는 비트율로 변환하는 방법을 찾는 것이다. 비트율 할당은 특정한 데이터를 표현하는 비트값과 왜곡과의 관계를 체계화한 비트율 왜곡이론 (*Rate-Distortion Theory: RDT*)에 의해 구할 수 있다. 이것은 데이터가 갖는 특성을 분석하여 최적의 화질을 얻는 최소한의 비트를 할당하는 이론으로 동영상의 압축 비트율 생성을 포함한 소스 코딩에 일반적으로 적용된다.

본 절에서는 고정된 비트율로 생성된 압축 동영상 부호를 낮은 비트율로 변환하는 알고리즘을 제안한다. 합당한 비트율을 생성하기 위해 프레임의 공간적, 시간적 특성을 복합적으로 해석하여야 하는 인코더와는 달리 본 연구는 이미 생성된 부호를 대상으로 함으로, 비트율 변환에 참조할 수 있는 정보를 미리 얻을 수 있는 이점을 갖는다. 기 압축된 비트율을 다른 값의 비트율로 변환하는 알고리즘은 앞절에서 설명한 바와 같이 다양한 알고리즘을 사용할 수 있고, 적용 용도에 따라 구현의 복잡도가 다르다. 본 연구는 알고리즘의 주 응용을 통신대역의 변화에 대응하며 실시간으로 대역폭에 부합되는 비트율로 동영상을 변환해야 하는 멀티미디어 통신으로 한다. 이것은 복잡한 장치 및 알고리즘을 이용해 화질에 중점을 둔, 비 실시간형 비트율 변환 방법과는 달리 단순한 구조를 갖고 최적의 비트율로 동영상을 실시간으로 변환하는 알고리즘을 필요로 한다. 멀티미디어 통신을 위한 비트율 변환의 조건인 실시간성, 화질 그리고 구현의 단순성은 기존의 알고리즘을 적용할 경우, 원하는 조건을 만족하기 어렵다. 본 절에서는 이를 해결하기 위해 압축영역에서 최적 양자화 파라미터 값을 찾는 방법으로 접근한다.

압축부호의 특성은 할당된 비트율에 대한 왜곡값으로 표현되며, 이 결과는 화질이란 정량화된 결과를 통해 알 수 있다. 일반적으로 영상신호에서 *PSNR*로 나타내는 화질은 프레임 단위의 대표치로 표현되는 값으로 *CBR* 영상과 같이 매크로블럭 단위에서 양자화 값의 변동이 클 경우에는 정확한 화질의 척도가 될 수 없다. 매크로블럭 단위의 급격한 양자화 값 변이는 프레임 내에서 화질의 변동이 심하다는 것을 의미하며, 동일한 *PSNR* 값을 갖는 영상의 경우에도 이것에 의해 분명한 화질의 차이가 발생한다. *CBR*과 동일한 *PSNR* 값을 갖는 *VBR* 영상이 보다 나은 화질을 나타내는 것도 이와 같은 이유이다. 따라서 비트율 변환 알고리즘의 기본 방

향을 다음과 같이 설정한다.

알고리즘의 기본방향

- 매크로블럭의 최적 양자화 파라미터 선정을 통한 비트율 변환
- 객관적 화질 (*PSNR*) 이외에 주관적인 화질 증진
- 프레임 단위의 분석을 통한 최적 비트율 변환
- 최대 33ms (한 프레임) 이하의 실시간 변환

압축 동영상의 비트율 변환을 단순히 설명하면, 목적하는 비트율로 부호를 변환하는 것이다. 그러나 이것을 위해서는 대상이 되는 스트림의 특성 및 구조를 정확히 분석해야 하며, 이 분석 방법에 따라 다른 결과를 낼 수 있다. 최적 비트율 변환을 위한 분석의 단위는 전체 스트림, 프레임 집합, 프레임 그리고 그 이하에 내려 갈 수 있다. 만일 그 단위가 전체 스트림이라면, 압축된 부호 모두를 분석한 후, 스트림의 구조를 변경하는 것을 의미하며, 단순히 비트율 변환의 관점에서 최적의 결과를 가져올 수 있으나, 실시간 응용에는 전혀 사용될 수 없다. 따라서 본 알고리즘은 실시간 응용에 적용이 가능하고 인간의 지각 시스템 (*Human Visual System: HVS*)에 대응되는 프레임 단위의 분석을 통해 비트율 변환을 행한다.

프레임 단위의 비트율 변환의 실제적인 요소는 프레임을 구성하는 매크로블럭의 양자화 파라미터이다. 즉 비트율 변환의 주 알고리즘은 프레임을 구성하는 개별 매크로블럭에 합당한 비트율을 할당하고 이에 해당하는 양자화 값을 찾는 것이다. 영상은 그 구성 형태에 따라 별도의 비트율로 나타내지며, 이것은 영상의 복잡도 또는 활성도 (*Image Activity: IACT*)의 함수로 표현할 수 있다. 즉 *IACT*는 영상의 구성형태를 나타내는 척도가 되며, 이것에 의해 비트를 할당하면 이상적인 결과를 얻을 수 있다. *IACT*에 따르는 비트율 할당모델을 구하기 위해 본고는 여러 유형의 내용을 포함하고 있는 스트림을 *VBR*로 인코딩하여 분석하였다. 또한 *VBR* 모델을 사용한 *IACT*가 *CBR*에서 왜곡되는 것을 막기 위해 양자화 파라미터에 따르는 부호 발생 함수를 구해 이를 정확한 *IACT* 복원에 사용하였다.

*IACT*는 양자화 값 변경의 기본단위가 되는 매크로블럭의 비트율 할당함수를 구하기 위한 것으로

실제적인 비트율 할당함수는 스트림의 구성형태에 종속된다. 본 장에서는 스트림 내용에 종속되는 비트율 할당에 대한 함수를 구하기 위해 다양한 내용을 갖고 있는 압축 동영상을 분석하여 IACT가 매크로블럭의 비트 수와 분명한 관계가 있음을 확인하였고 실제 비트율 변환을 위해 스트림의 내용 및 유형에 따라 비트율 할당함수를 맞추어 가는 적응형 비트율 할당모델(Adaptive Bit-Rate Budget Model)을 제안하였다

3.1 압축 동영상의 출력 비트율 모델링

3.1.1 영상의 활성화도

영상의 특성에 관계없이 Y,Cb,Cr의 동일한 형태로 표현되는 비 압축영역의 동영상과는 달리 압축 영역의 동영상 부호는 DCT 변환 이후, 각 계수 값을 영상의 엔트로피를 나타내는 값으로 사용할 수 있다. 이때 DCT 계수의 특성을 표현하는 것을 IACT라 칭한다. IACT와 출력 비트율과의 관계를 모델링하면, 비트율 할당의 기준으로 삼을 수 있다. 비트율 할당함수를 구하기 위해 본 절에서는 IACT에 대한 매크로블럭의 비트율 생성관계를 분석한다. 매크로블럭의 IACT를 나타내는 방법으로 다음의 네 가지 방법을 택하였다.

1] 매크로블럭을 구성하는 모든 DCT계수의 변이

$$\sum (DC\ coefficient + AC\ coefficient) - (DC\ coefficient + AC\ coefficient)^2$$

2] 매크로블럭을 구성하는 DCT계수 중 AC만의 변이

$$\sum (AC\ coefficient - AC\ coefficient)^2$$

3] 매크로블럭을 구성하는 모든 DCT 계수의 절대값의 합

$$\sum |DC\ coefficient + AC\ coefficient|$$

4] 매크로블럭의 DCT 계수 중 AC에 대한 절대값의 합

$$\sum |AC\ coefficient|$$

그림1은 네가지 방법에 따르는 IACT에 대한 매크로블럭 비트 수와의 관계를 나타낸 것이다. 그림에서 보듯 방법1], 2]의 경우는 서로 상관 관계가 없는 반면, 방법3], 4]는 분명한 연관성을 갖음을 알 수 있다. 또한 3]의 경우보다, 4]가 보다 나은 특성을 보이고 적은 데이터의 분석만으로 표현이 가능하므로 IACT를 나타내는 방법은 4]가 합당함을 알 수 있다. 그러나 방법4]의 경우도 정확한 선형성을 나타내는 비례관계를 보이고 있지 않다. 따라서 본 고에서는 다양한 내용을 포함하는 스트림을 프레임 별로 분석하였다. 이 결과 IACT와 매크로블럭과의 관계는 스트림의 내용보다는 프레임의 유형에 종속된다는 결론을 얻었다. 그림2, 그림3 그리고 그림4에 프레임 유형에 따르는 IACT와 매크로블럭 비트 수의 관계를 보였다.

분석에 사용된 스트림은 MPEG-2 Test Model 5 (TM-5) 인코더를 사용하였으며, 양자화 값을 고정된 VBR 데이터이다. 양자화가 적용되는 최소 단위인 매크로블럭을 표현하는 최적의 비트는 RDT에 의한 값으로 정해져야 하고 각각의 매크로블럭에 필요한 고유한 비트 값이 있다. 고정된 양자화 값으로 표현되는 VBR 영상은 RDT의 관점에서 최적의

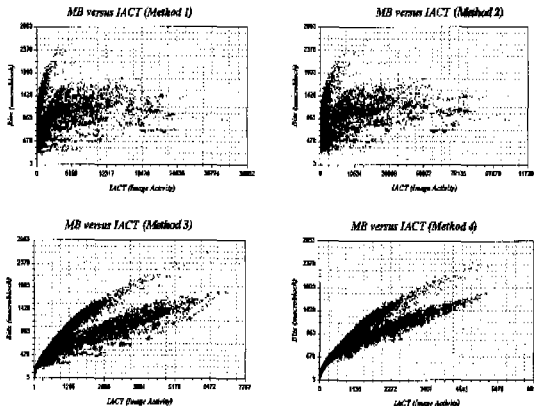


그림 1. IACT의 4가지 방법

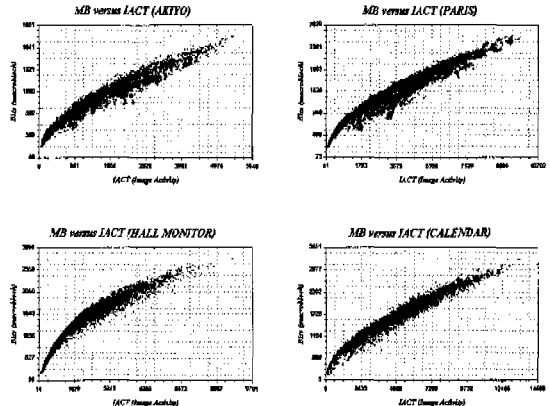


그림 2. IACT와 매크로블럭 비트 수와의 관계 (1 frame)

비트가 할당되었다고 볼 수 없으나, 동일한 양자화 값을 사용한 VBR은 매크로블럭의 특성을 그대로 나타내는 최적에 가까운 비트 할당이 이루어졌다고 가정할 수 있다. 따라서 본 알고리즘에 사용되는 비트를 할당 함수는 VBR의 IACT 특성을 이용한다. 이것은 비트율 변환의 처리 대상이 CBR 영상이지만 영상의 비트율 할당값은 VBR 영상의 장점을 취하기 위한 것이다.

IACT에 따르는 비트율 할당 함수는 프레임 유형 별로 뚜렷한 선형성을 가져 비트율 할당 모델로 사용할 수 있으며, 이 특정 스트림에 대한 모델을 이용하여 프레임을 구성하는 매크로블럭의 비트 할당 값을 구한 후, 변환을 행하면 VBR 영상의 특성을 지닌 결과를 얻을 수 있다. 그러나 비트율 할당 함수는 스트림의 내용에 따라 다소 다른 특성을 갖는다. 따라서 단일 선형 비트율 할당모델을 사용하는 것보다는 압축 동영상의 고유한 IACT에 의해 구간

별로 함수관계를 맞추어 가는 적응형 모델이 바람직 할 것이다.

표 1. 프레임 별 비트율 할당 함수와 평균 IACT

프레임 유형	스트림 종류	프레임 별 비트율 할당 함수 [BITS = α(IACT) ^β]	스트림 종류	프레임 유형	평균 IACT
I frame	Akiyo	21.5(IACT) ^{0.6}	Mother and daughter	I	1094
	Mother and daughter	30.9(IACT) ^{0.6}		P	302
	Paris	34.9(IACT) ^{0.6}		B	173
	Table tennis	24.1(IACT) ^{0.6}		I	941
	Hall monitor	41.5(IACT) ^{0.6}		P	403
P frame	Calendar	10.7(IACT) ^{0.6}	Contour	I	3472
	Contour	20.5(IACT) ^{0.6}		P	1143
	Contour	34.4(IACT) ^{0.6}		B	665
	Table tennis	15.7(IACT) ^{0.6}		I	2500
	Mother and daughter	14.6(IACT) ^{0.6}		P	1051
B frame	Paris	15.5(IACT) ^{0.6}	Hall monitor	R	678
	Table tennis	14.6(IACT) ^{0.6}		I	1836
	Hall monitor	21.2(IACT) ^{0.6}		P	1247
	Calendar	18.4(IACT) ^{0.6}		B	797
	Contour	15.0(IACT) ^{0.6}		I	4189
B frame	Akiyo	18.1(IACT) ^{0.6}	Contour	P	3312
	Mother and daughter	14.6(IACT) ^{0.6}		B	2019
	Paris	15.5(IACT) ^{0.6}		I	2143
	Table tennis	14.6(IACT) ^{0.6}		P	736
	Hall monitor	17.5(IACT) ^{0.6}		B	305
B frame	Calendar	19.1(IACT) ^{0.6}	Contour	I	2384
	Contour	14.6(IACT) ^{0.6}		P	1749
	Contour	16.1(IACT) ^{0.6}		B	1073
	Table tennis	14.6(IACT) ^{0.6}			
	Hall monitor	17.5(IACT) ^{0.6}			

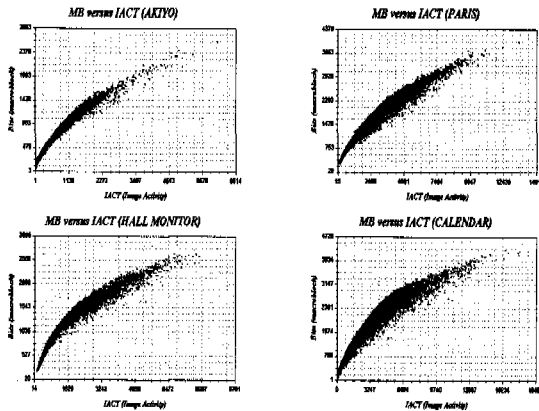


그림 3. IACT와 매크로블럭 비트 수와의 관계 (P frame)

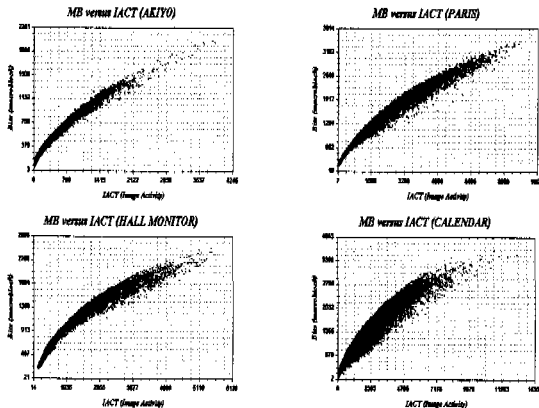


그림 4. IACT와 매크로블럭 비트 수와의 관계 (B frame)

3.1.2 양자화 파라미터와 비트율

앞서 설명한 바와 같이 IACT에 대한 매크로블럭의 관계는 프레임 별의 모델이 합당함을 알 수 있다. 그러나 이것을 비트변환에 직접 적용하기 위해서는 일부 해결되어야 하는 점이 있다. 현재 IACT는 양자화 파라미터를 고정된, VBR 스트림을 분석한 것이다. 그러나 CBR 영상의 IACT를 구하기 위해서는 양자화 파라미터의 변이에 따르는 왜곡을 보상해 주어야 한다. 양자화 값이 고정된 VBR 영상과는 달리 CBR 영상은 프레임내 매크로블럭의 양자화 파라미터가 인코더의 율제어 기능에 종속되어 가변된다. 즉 CBR 영상에 대한 매크로블럭의 IACT를 구하기 위해 역양자화를 거쳐 DCT 계수에

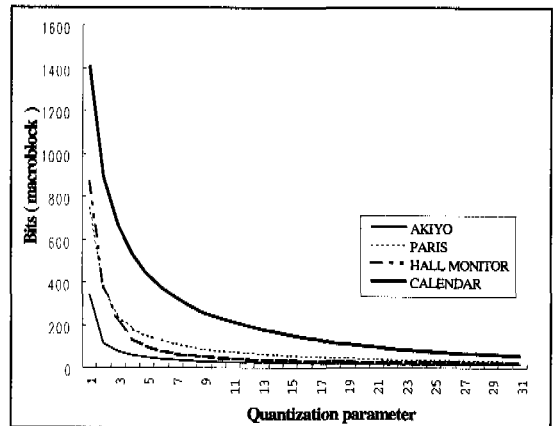


그림 5. 양자화 파라미터와 매크로블럭 비트 수와의 관계

대한 연산의 결과는 정확한 IACT와는 다른 왜곡된 값을 가져 이 왜곡을 처리하기 위해 매크로블럭과 양자화 파라미터의 관계를 정형화 하여 보상한다.

그림5는 양자화 파라미터에 따라 매크로블럭의 비트수가 변화되는 형태를 나타낸 것이다. 그림에서 보였듯이 양자화 파라미터 값에 따라 변화되는 매크로블럭의 비트 수는 명확한 선형성을 갖으나, 실제 VLC 코드로 표현되는 압축 비트는 정확한 선형 함수로 볼 수 없다. VLC는 RLC와 엔트로피 코딩으로 이루어지며 이 결과 발생하는 코드는 양자화 파라미터 값에 따라 생성되는 0의 값을 갖는 DCT 계수의 분포와 코드의 발생 빈도 수에 따라 다르게 표현된다.

IACT는 영상의 내용에 따르는 특성을 나타내는 지표가 되며, 이것을 기준으로 매크로블럭에 비트를 할당하는 것이 본 연구의 기본 알고리즘이다. 최적의 비트 변환을 위해서 영상이 갖고 있는 정확한 IACT를 구해야 하며, 양자화 파라미터에 의해 왜곡된 CBR 영상의 IACT를 원래의 IACT로 보상하는 것이 알고리즘의 성능에 관건이 된다. IACT 왜곡은 매크로블럭 별로 다르게 적용된 양자화에 기인한 것으로 양자화 파라미터의 값이 1 이상인 매크로블럭에 양자화 오류 값을 더해 줌으로써 보상이 가능하다. 양자화 오류와 IACT는 동일하게 데이터의 양으로 표현되므로 서로 더해 줄 수 있으며, 이 과정을 거친 매크로블럭의 IACT는 AC 계수의 확률적 분포에 따라 평균적으로 보상된 값이다. 출력벡터 $\{Y_i; i=1, \dots, N\}$, 입력구간 $\{S_i; i=1, \dots, N\}$, 그리고 양자화기가 q 일 때, 일반적으로 고정된 스텝을 갖는 스칼라 양자화에서 발생하는 양자화 에러는 식(1)으로 표현된다.

$$\bar{d} = E_d(X, q(x)) = E_p(\|X - q(X)\|) = \sum_i \int_{S_i} f(x) \rho(\|x - y_i\|) dx \quad (1)$$

여기서 N 이 충분히 크고 확률밀도 함수 pdf $f(x)$ 는 S_i 내에서 거의 상수의 값을 갖는다고 가정하여 식(2)를 구할 수 있고, $f(x)$ 는 일정구간 내에서 f_i 로 근사화 하여 식(5)를 얻을 수 있다.

$$\sum_i \int_{S_i} \rho(\|x - y_i\|) dx \approx f_i \int_{S_i} \rho(\|x - y_i\|) dx \quad (2)$$

$$P_i = \int_{S_i} f(x) dx \approx f_i V(S_i) \quad (3)$$

$$\sum_i \int_{S_i} f(x) \rho(\|x - y_i\|) dx \approx \sum_i \frac{P_i}{V(S_i)} \int_{S_i} \rho(\|x - y_i\|) dx \quad (4)$$

$$\bar{d} \approx \sum_i \frac{P_i}{V(S_i)} \int_{S_i} \rho(\|x - y_i\|) dx = \frac{1}{V(S_0)} \int_{S_0} \rho(\|x\|) dx \quad (5)$$

Euclidean 일반분포와 $\rho(\mu) = \mu^2$ 인 경우에 qs^k 인 영역에 대한 평균 양자화 에러는 유도과정[24]을 이용하여 식(7)과 같이 된다. 식(8)은 일반 양자화 에러에 HVS 요소를 고려한 것이다.

$$\bar{d} \approx \frac{1}{V(S_0)} \int_{S_0} \rho(\|x\|^2) dx \quad (6)$$

$$\frac{1}{k} \bar{d} = \frac{1}{k} \int_{-\pi/2}^{\pi/2} \dots \int_{-\pi/2}^{\pi/2} \sum_b^k dx_0 \dots dx_{k-1} = \frac{QS^2}{12} \quad (7)$$

$$\bar{d} = 1/N \sum_{i=0}^N \frac{(a_i^2 * QS_i^2)}{12} \quad (8)$$

(where a : HVS parameter and qs : quantization step)

$$\text{Original data} \approx QS * (\text{quantized data}) + \text{distortion} \quad (9)$$

위의 식에서 왜곡은 양자화 에러 $QS^2/12$ 에 해당하고 이 값이 원래의 데이터와 역양자화 처리후의 데이터의 MSE를 의미하므로 식(10)으로 다시 쓸 수 있다. 본 논문에서 사용하는 IACT는 DCT 계수의 절대값을 더한 것이므로 양자화 파라미터에 따르는 IACT의 왜곡을 보상하기 위해 식(10)을 이용하여 다음과 같이 나타낼 수 있다.

$$(B_i - \hat{B}_i)^2 \approx \frac{QS_i^2}{12}, B_i \approx \hat{B}_i + \frac{QS_i}{\sqrt{12}} \quad (10)$$

$$IACT(MB) = \sum_i (a^{-1}(|DCT\ coefficient|)) + \sum_i \frac{qs_i}{\sqrt{12}} \quad (11)$$

3.2 적응형 비트 할당모델

비트의 할당 문제를 IACT의 함수로 모델링하고 양자화 파라미터 차이에 의한 왜곡을 보상하여 줌으로써 CBR 동영상의 비트율 변환이 이루어질 수 있다. 그러나 IACT에 따르는 비트율 할당의 전체적인 특성은 프레임에 대한 함수관계로 표현할 수 있으나, 여전히 스트림의 내용에 따라 함수의 특성이 다르게 나타난다. 이것을 해결하기 위해 분석된 개별 스트림의 비트할당 함수를 평균하는 모델을 사용하는 방법도 한가지 대안이 될 수 있다. 그러나 이것은 스트림 집합의 평균 특성을 나타내므로 정확한 비트율 할당 함수로 사용하는데 문제가 있다. 본 연구는 이것을 해결하기 위해 스트림의 내용에 따라 고유한 함수를 찾아 가는 적응형 비트율 할당

모델을 사용한다.

적응형 비트율 할당모델

- Step 1: 입력 스트림의 프레임 유형 별 IACT 분석
- Step 2: IACT에 따르는 비트율 할당 함수 적용


```
begin procedure
    if( $\Delta_1 < IACT < \Delta_2$ ), use function(A);
    else if( $\Delta_3 < IACT < \Delta_4$ ), use function(B);
    else...
end procedure
```
- Step 3: 적응 함수에 의한 비트율 할당

IACT는 스트림의 내용에 따라 약간씩 다른 특성을 갖으나, 전체 스트림을 대상으로 구간별로 IACT 함수를 구하면, 원래의 값에 수렴할 수 있다. 그림6은 IACT와 비트의 관계를 구간 별로 세분한 것을 보여준다. 모델은 IACT의 구간을 나누는 Δ_{IACT} 를 작게 구분할수록 정확도와 복잡도가 병행하여 높아진다. 그러나 IACT는 영상의 내용에 따라 몇단계로 구분할 수 있으며, 세분된 구간은 거의 동일한 특성을 갖는다. 또한 영상의 화질에 영향을 미치는 중요도를 고려하여 적응형 비트율 할당 함수의 구간 Δ_{IACT} 를 I, P, B 프레임 순으로 차등을 두어 적용하였다. 적응형 비트 할당모델은 입력되는 스트림의 프레임 유형 (I, P, B)과 이것의 개별 매크로블럭에 대한 IACT를 분석하여, 해당되는 비트율 할당함수를 찾는다. 이 함수로 적용한 후, 매크로블럭에 필요한 비트율 할당값을 계산한다.

본 절에서는 적응형 비트할당 모델을 이용하여 CBR 압축 동영상의 비트율을 변환하는 알고리즘에 대해 설명한다. 비트율 변환은 높은 비트에서 낮은

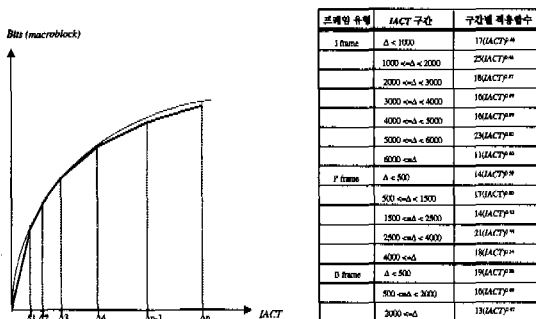


그림 6. 적응형 비트 할당모델과 비트율 할당 함수

비트로의 변환을 의미하며 최적화는 입력 스트림의 분석을 통해 매크로블럭에 왜곡을 고려한 합당한 비트를 할당함으로써 이루어진다. 비트율 변환 알고리즘은 입력 압축영상을 분석하여 전체 스트림, 프레임에 비트를 배분하고, IACT 계산과 적응형 비트율 할당모델을 이용해 매크로블럭에 합당한 비트를 할당하고, 이 비트에 대한 양자화 파라미터를 결정하는 여섯 단계의 과정으로 이루어진다.

비트율 변환 알고리즘

- Step 1: 스트림 분석
- Step 2: 스트림 및 프레임의 비트율 할당
- Step 3: 매크로블럭의 IACT 계산
- Step 4: 적응형 비트율 할당함수에 따라 매크로블럭에 비트 할당
- Step 5: 매크로블럭에 대한 양자화 파라미터 탐색
- Step 6: MSE 비교를 통해 최종 양자화 파라미터 결정

알고리즘의 처음 단계인 스트림 분석은 입력되는 압축부호의 형태를 해석하는 과정이다. 압축부호는 비트변환에 사용될 수 있는 부분과 변경이 불가능한 부분으로 나뉜다. 즉 재양자화를 통해 낮은 비트로 변환하는 본 알고리즘에서는 압축부호의 선택스 파라미터, 움직임 정보 등을 그대로 사용하므로 이것을 제외한 나머지 부분을 비트율 변환에 사용한다.

$$STREAM_{input} = reducible_part(DCTAC\ coefficient) + not_reducible_part(syntax\ parameter + motion\ vector) \quad (12)$$

$$BITS_{frame} = \frac{bit\ rate(stream\ to\ be\ shaped)}{bit\ rate(original\ stream)} \quad (13)$$

$$IACT_K = IACT_K^* + \frac{(2 * QP_K) - 2}{\sqrt{12}} \quad (14)$$

비트율의 변환은 분석되는 압축영상의 구간이 커질수록 화질 측면에서 보다 나은 결과를 낳는다. 즉 향상비트의 특성을 갖는 스트림 전체의 특성을 분석하거나, 상당 수의 프레임을 분석하여 비트를 재할당 함으로서 화질이 향상되는 결과를 얻을 수 있다. 그러나 분석하는 구간이 커질수록 비트변환의 절차가 복잡해지고, 처리지연도 병행하여 증가하게 된다. 따라서 분석은 이를 고려한 합당한 구간이어

야 하며, 본 연구는 한 프레임을 분석 구간으로 삼는다. 이것은 30Hz 단위의 정보인 하나의 영상 프레임이 객관적인 화질과, 화면에 표시되는 기본단위가 되기 때문이며, 멀티미디어 통신에서 동영상 조작처리를 위해 허용되는 지연조건을 고려하여 비트율 변환이 한 프레임 내에서 실시간으로 이루어지도록 하기 위함이다. 이에 따라 스트림 및 프레임에 대한 비트할당은 입력 압축부호의 비트율과 변환하는 압축부호의 비트율을 선형적으로 나눠 구한다.

입력 압축부호의 분석과 스트림, 프레임에 대한 처리 이후, 본 연구의 핵심인 매크로블럭에 대한 비트 할당을 위해 $IAC T$ 를 구한다. $IAC T$ 는 입력 스트림의 매크로블럭을 역양자화 하여 구할 수 있으며, 양자화 파라미터에 의한 오류 보정을 위해 다음 식을 사용한다.

식(14)에서 $IAC T_K^*$ 는 매크로블럭의 AC 계수를 통해 얻은 값이고, $IAC T_K$ 는 양자화 오류가 보정된 것이다. 여기서 양자화 오류에 해당하는 부분은 원래의 $IAC T$ 가 고정된 양자화 파라미터를 기준으로 하였으므로 이것을 고려하여 주었다. 해당 $IAC T$ 처리 후, 실제 매크로블럭에 할당해야 하는 비트 수를 구하여야 한다. 이것은 적응형 비트할당에 사용하며, 해당되는 매크로블럭이 속하는 프레임의 유형과 $IAC T$ 값이 속하는 범위에 따라 할당한 함수를 사용한다. 이 과정을 통해 프레임에 속한 모든 매크로블럭에 대한 비트 할당값을 구할 수 있다. 그러나 할당된 비트 수는 매크로블럭에 이상적으로 할당되어야 하는 기준값일 뿐, 비트율 변환에 직접 사용될 수는 없다. 실제 매크로블럭에 할당되는 비트 수는 알고리즘의 첫단계인 스트림 분석결과와 두번째인 프레임에 대한 비트율 할당을 기준으로 변환에 사용될 수 없는 부분을 제외한 비트 수에 대한 비례식으로 구하여 준다.

$$BIT_{original\ frame} = \alpha(reducible_part) + \beta(not_reducible_part) \quad (15)$$

$$(\alpha = BIT_{original\ frame} - \beta)$$

$$BIT_{frame\ shaped} = \phi + \beta \quad (16)$$

$$(\phi = \alpha/K, K: \text{proportion of shaping})$$

$$\theta = \sum_i H_i BIT(MB_i) \quad (17)$$

$$BIT(MB_0): \dots : BIT(MB_n) = H(IAC T_1): \dots : H(IAC T_n)$$

(H_i , Adaptive bit-rate budget function)

$$BIT_{shape}(MB_i) = \frac{\phi}{\theta} * BIT(MB_i) \quad (18)$$

식(15)~(18)를 이용하여 변환될 비트율에 따르는 프레임내의 개별 매크로블럭에 대한 실제 비트 할당값을 구할 수 있다. $IAC T$ 의 가중치에 따라 매크로블럭에 할당된 비트값은 양자화 과정을 통해 비트 수로 표현된다. 이것은 매크로블럭에 할당된 비트수가 임의의 값을 가져도 이것을 표현하는 방법은 1에서 31까지의 제한된 값을 갖는 양자화 파라미터에 의해 결정되며, 원래의 비트량에 대비한 왜곡의 관계가 최적이 되는 양자화 파라미터 값을 선택해야 이상적인 비트 변환이 이루어진다. 이 양자화 파라미터 선정에 일반적으로 사용되는 것이 그림(7)에 보인 *bisection* 탐색방법이며, 몇단계의 탐색과정을 통해 원하는 값을 얻을 수 있다

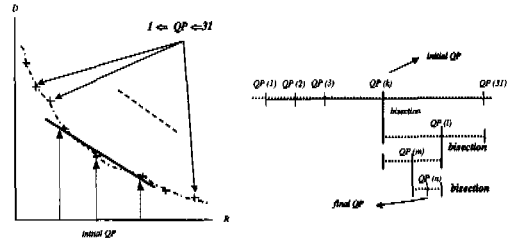


그림 7. bisection 에 의한 양자화 파라미터 값 선정

*bisection*을 이용하는 양자화 파라미터 결정은 선택된 초기값에 따라 수행되는 횟수가 달라지므로, 초기 양자화 파라미터의 선정은 비트 변환 알고리즘의 복잡도를 결정하는 하나의 척도가 된다. 즉 비트 변환을 위한 매크로블럭의 양자화 파라미터 값은 높은 비트에서 낮은 비트로 변환한다는 가정하에 앞에서 1~31의 전체 영역보다는, 원래의 매크로블럭의 양자화 파라미터 값과 31 사이에 있다고 볼 수 있다.

$$QP_{original} \leq QP_{target} \leq 31, QP_{initial} = \frac{QP_{original} + 31}{2} \quad (19)$$

(QP; Quantization Parameter)

그러나 *bisection* 방법은 초기의 양자화 파라미터 값에 따라 가변적으로 여러 단계의 탐색과정을 거쳐야 하므로 탐색의 단계를 줄이기 위한 초기 양자화 파라미터 값 선정에 다음식을 이용한다.

$$BIT(MB_i) = \alpha(\beta * \sigma_i^2 / QS^2 + \epsilon) \quad (20)$$

$$(\alpha, \beta, \epsilon: \text{constant}, QS = 2 * QP, \sigma_i = \sqrt{DCTcoeff_i - DCTcoeff})^2$$

매크로블럭의 양자화 파라미터에 따르는 비트 수

는 식(20)에 의해 나타낼 수 있다. 이것은 DCT 계수의 표준편차를 사용한 것으로, run-length 코드에 의해 비트압축이 일어나는 것을 수식화 한 것이다. 이 수식을 통해 개별 매크로블럭에 대한 양자화 파라미터 값은 다음과 같이 구할 수 있다.

$$QS_i = \frac{\sigma_i^2}{\sqrt{MB_i}}, \quad QP_i = \frac{QS_i}{2} \quad (21)$$

$$BIT_{original_MB} \cdot QP_i = BIT_{target_MB} \cdot QP_{initial} \quad (22)$$

$$(QP_{initial} = \frac{QP_i * BIT_{target_MB}}{BIT_{original_MB}})$$

최적의 양자화 파라미터 값은 단순히 비트할당 모델에 의해 구해진 매크로블럭의 비트 수를 정확히 나타내는 것 이외에 왜곡 측면이 같이 고려되어야 한다. 매크로블럭에 할당된 비트에 근사한 값을 갖도록 하는 양자화 파라미터는 식(22)의 초기 값을 사용하고, 몇단계의 bisection을 통해 구할 수 있다. 그러나 이때의 값은 할당된 비트수와 가장 근접한 비트 수 일뿐, 왜곡은 고려되어 있지 않다. 즉 최종 양자화 파라미터 값은 비트에 대한 측면과 해당 비트에 대한 왜곡을 고려하여 몇가지 양자화 파라미터 중 최적의 값을 선택하여야 한다. 이때 왜곡은 매크로블럭이 영상으로 표현될 때의 화질을 의미하므로 압축영역에서는 정확한 측정 방법이 없다. 따라서 본 논문은 매크로블럭의 왜곡을 판단하는 기준으로 비트율이 변환되기 이전의 매크로블럭이 갖는 DCT 계수와 변환된 후의 DCT 계수의 차이를 왜곡의 척도로 삼는다. 이것은 기존의 PSNR이 원래의 화소와 처리된 후의 화소간을 뺀 것임을 고려하면, 압축영역에서의 왜곡을 측정하는 척도로서 합당하다 할 수 있다.

$$\text{Min}(\sqrt{\alpha[R(QP_i) - \bar{R}(QP_i)]^2 + (1-\alpha)[D(QP_i) - \bar{D}(QP_i)]^2}) \quad (23)$$

$$D(QP_i) - \bar{D}(QP_i) = \frac{1}{n} \sum_{i=1}^n (DCT_{coeffi} - \overline{DCT_{coeffi}})$$

(DCTcoeffi는 원래의 DCT 계수, DCTcoeffi는 비트 변환된 후의 DCT 계수, α 는 양자화 파라미터 결정함수)

최종 양자화 파라미터 선정은 bisection을 통해 선정된 값과 이웃하는 양자화 파라미터 중 식(23)의 값이 적은 것으로 선택한다. 여기서 α 값은 양자화 파라미터의 값을 화질 또는 비트 수의 수렴 중 어느것에 중점을 두느냐를 결정하는 가중치의 의미

를 갖는다.

IV. 실험 및 고찰

본 장에서는 비트율 변환 알고리즘 제안에 이어 소프트웨어를 이용한 구현을 통해 기능의 타당성을 검증하고, 기존 방식과의 성능을 비교, 분석한다. 알고리즘 구현에 사용된 압축 동영상은 MPEG 소프트웨어 시뮬레이션 그룹에서 제공하는 TM-5 인코더를 사용하였다. 또한 처리된 영상을 화면에 출력하기 위한 디코더는 인코더와 동일한 소프트웨어 그룹에서 제공되는 표준 소프트웨어를 이용하였다. 그림8은 실험에 사용된 영상을 보여주며, 영상은 움직임이 적은 영상전화, 영상회의형에서 화면변화가 많은 영상까지 다양한 내용을 담고 있다.

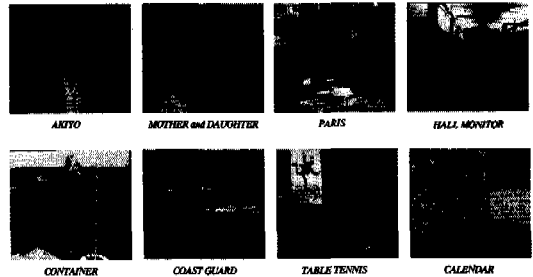


그림 8. 실험용 영상

성능분석은 영상의 화질, 구현의 복잡성 그리고 실시간 처리의 세가지 항목에 대해 이루어졌으며, 디코딩과 인코딩의 전과정을 거치는 화소영역에서의 처리방법을 제안한 알고리즘과의 공통 비교대상으로 삼았다. 영상의 화질은 일반적으로 사용되는 PSNR을 사용하였으나, 비트 변환 알고리즘의 경우에는 객관적인 평가 이외에 인간의 주관적인 판단의 척도 HVS로 하나의 화면 내의 양자화 스텝 변이 값을 이용한 정량화된 방법을 사용하였다. 식(24)와 식(25)는 각각 PSNR과 양자화 스텝의 변이를 나타낸 것이다. 식(24)에서, X_i 는 원래의 화소값, \bar{X} 는 압축신호에 대한 화소값, 그리고 N 은 영상을 구성하는 픽셀의 수이다. 식(25)의 QS_i 는 매크로블럭의 양자화 스텝값, QS 는 전체 양자화 스텝의 평균값 그리고 M 은 하나의 프레임을 구성하는 매크로블럭의 수이다.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}, \quad MSE = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2 \quad (24)$$

$$HVS_n = \frac{1}{M} \sum_{i=1}^M (QS_i - \overline{QS})^2 \quad (25)$$

일반적으로 통용되는 정량적인 성능평가 방법을 사용할 수 있는 화질과는 달리 구현의 복잡도는 분명한 평가의 척도를 통한 분석이 용이하지 않으며, 수식상의 해석도 실제 알고리즘을 구현한 방법에 종속되어 정확한 판단기준으로 삼기에는 문제가 있다. 따라서 일반적인 수치분석에 앞서 동영상 압축 알고리즘에 사용되는 각 부분의 복잡도를 비교한 예를 보인다[25-26].

그림9에 보인 것과 같이 인코더의 복잡도는 움직임 추정부분이 연산의 60~80%, DCT는 10~20% 정도가 소요되어 두가지 처리가 갖는 비중은 전체의 70~90% 이상을 차지하고, VLC, 양자화, FLC 그리고 나머지 부가기능은 극히 일부분에 해당한다. 디코더는 보통 인코더보다 수월하게 구현할 수 있으나 여전히 움직임 보상과 DCT-1은 복잡도의 50% 정도를 차지하는 절대적인 비중을 갖는다. 따라서 압축영상에 대한 알고리즘에 이와 같은 움직임 추정, 보상 그리고 DCT, DCT-1에 대한 처리가 포함되지 않으면 구현에 소요되는 복잡도가 70~80% 이상 절감된다고 볼 수 있다. 물론 그림에 보인 내용은 모든 영상에 공통되는 것이 아니며, 각 요소에 대한 복잡도 역시 구현방법에 따라 상당한 편차를 가질 수 있다. 그러나 이 데이터는 연산의 복잡도에 대한 한가지 기준자료로 사용할 수 있다.

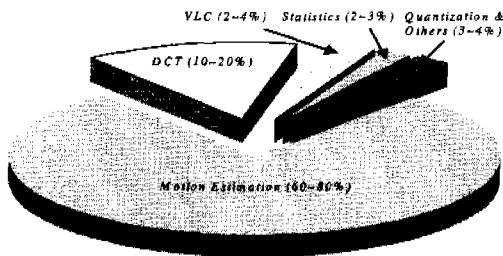


그림 9. 압축 동영상 알고리즘의 복잡도

구현에 대한 개략적인 복잡도를 설명한 위의 내용 이외에 수치적 해석을 사용한 방법이 있다. 이것은 동영상 압축처리의 각 단계를 곱셈과 덧셈에 대한 연산으로 표현한 것으로 보다 객관적인 결과를 얻을 수 있는 한가지 방법이 될 수 있다[27].

인코더, 디코더에서 발생하는 처리지연 역시 수치적 해석보다는 압축 동영상을 사용하는 응용에 따른 제한 값이 있을 뿐이다. 단-대-단 지연의 조건이

엄격한 대화형 멀티미디어 통신의 경우, 종단점 간의 지연은 150~400ms 이내 값을 만족하여야 하며, 이때 동영상 인코더와 디코더에 발생하는 지연은 네트워크의 전파지연과 큐잉지연을 뺀 값이다. 일반적으로 영상통신에 발생하는 단-대-단 지연은 최소한 4 프레임 (132ms) 이상이 발생하는 것으로 받아들여지고 있다. 그러나 이것은 하드웨어로 동영상 압축 코덱을 구현하였을 때를 고려한 것으로 정확한 이론적 근거보다는 현재의 기술이 반영된 수치이며, 구현 방법에 따라 지연 값은 달라질 수 있다. 단-대-단 지연을 표현하기 위해서는 동영상 압축 알고리즘의 각 단계를 처리지연에 대응하는 모델로 정형화하여야 하나, 이것에 대한 모델링은 아직 정립된 것이 없다. 따라서 분명한 수치적 해석은 어려우나, 단-대-단 지연요소를 개략화 하면 다음과 같다.

$$DELAY_{end-to-end} = DELAY_{encoding} + DELAY_{channel} + DELAY_{decoding} \quad (26)$$

$$DELAY_{encoding} = D_E^p + D_E^w \quad (27)$$

(D_E^p : Encoding processing time,
 D_E^w : Encoding waiting time)

여기서 D_E^p 는 동영상을 처리하는데 소요되는 지연으로 인코더를 설계하는 방법에 종속되는 값이며, 한 프레임의 처리에 소요되는 시간인 33ms 정도를 일반적인 지연값으로 사용한다. D_E^w 는 인코더의 프레임 유형에 의해 결정되며, 순방향 움직임만을 사용하는 경우, 0 이 된다.

$$DELAY_{channel} = D_C^b + D_C^t \quad (28)$$

(D_C^b : Channel buffer delay,
 D_C^t : Channel throughput delay)

D_C^b 는 항등 비트율의 통신을 위해 가변적으로 생성되는 동영상 압축부호를 완충버퍼에 일정량 저장하여 송수신함에 따라 발생하는 데이터 완충 버퍼링 지연으로 송수신단 모두 동일한 값을 갖는다. D_C^t 는 데이터가 송신버퍼를 떠나 착신단의 수신버퍼에 입력될 때까지의 시간에 해당하는 통신 지연 값이다.

$$DELAY_{decoding} = D_D^p + D_D^w \quad (29)$$

(D_D^p : Decoding processing time,
 D_D^w : Decoding waiting time)

D_0 와 D_0' 는 인코더에서 사용된 지연요소와 동일한 의미를 가지며, 양방향 움직임 벡터를 사용하는 프레임을 처리할 때만 0 이상의 값을 갖는다. 처리된 영상의 성능분석을 위해 화질, 구현의 복잡도 그리고 지연의 세가지 항목에 대한 정량화된 해석과 실험적인 데이터에 의한 결과를 사용하는 것으로 구분하여 설명하였다.

4.1 성능 분석

4.1.1 화질

비트율 변환에 따르는 화질은 화소영역에서의 처리방법, AC를 제거하는 방법 그리고 본 논문의 알고리즘에 따르는 방법에 의해 각기 다른 비트율로 변환된 영상의 화질을 PSNR로 비교, 평가하였다.

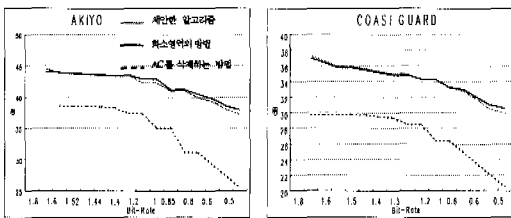


그림 10. 화질 비교

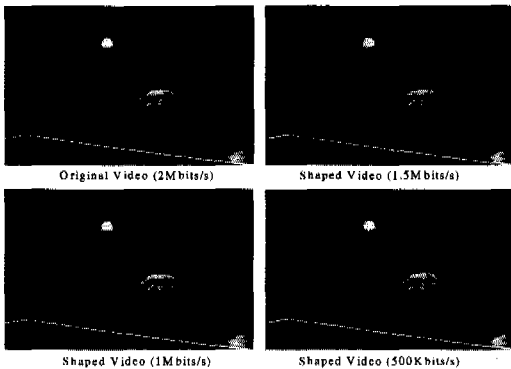


그림 11. 비트율이 변환된 영상

그림10에 보였듯이 비트율 변환에 의한 화질의 차이에 대해서 AC를 삭제하는 방법은 매크로블럭에 포함된 DCT AC 계수를 임의로 제거함에 따라 전반적인 화질은 두 방식에 비해 4dB 이상의 큰 차이를 보였으며, 특히, 낮은 비트율에서는 중요한 AC 정보가 상당부분 제거되어 급격한 화질저하가 발생함을 알 수 있다. 또한 이것은 8x8의 블럭으로 구성된 각 배열의 아래부분을 차례로 제거하는 방법으로 비트율을 변환하므로 목적하는 비트율에 정

확히 맞추어 줄 수 없다는 단점을 지녀, 비트율 변환에 적절하지 못한 방법임을 알 수 있다.

표.2는 제안한 알고리즘과 화소영역의 방법에 따라 1.5Mbits/s에서 500Kbits/s까지의 비트율로 변환된 영상의 PSNR 평균값과 양자화 스텝의 평균 표준편차를 보였다. 객관적인 화질이라 할 수 있는 PSNR은 표에 나타난 것과 같이 거의 동일한 값을 갖고, 특히 스트림의 형태와 변환 비트율에 따라 제안한 알고리즘에 의해 변환된 영상의 화질이 일부 우월한 성능을 나타내는 부분이 있음을 알 수 있다. 그러나 전체적인 평균을 보면, 화소영역의 방법이 근소하게 나은 특성을 가짐을 알 수 있다. 반면에 본 고에서 또 다른 화질의 지표로 사용한 양자화 스텝의 표준편차는 실험에 이용한 모든 스트림 그리고 변환 비트율에서 화소영역의 처리방법에 비해 낮은 값을 가짐을 알 수 있다. 즉 제안한 알고리즘에 따라 처리된 영상이 갖는 화면의 화질 균일도는 화소영역의 방법에 비해 우수한 특성을 나타내며, 이것과 객관적 화질 그리고 구현의 복잡도를 종합적으로 고려하면, 제안한 방식이 상당한 장점을 갖는 알고리즘임을 알 수 있다.

표 2. 화질 평균

영상의 종류	알고리즘	변환 비트율	객관적 화질 (평균 PSNR)	주파적 화질 (평균 QS 표준 편차)
CALINDAR	화소영역의 처리방법	1.5Mbits/s	25.66dB	10.85
		1.5Mbits/s	25.70dB	7.43
	화소영역의 처리방법	1.0Mbits/s	24.09dB	13.72
		1.0Mbits/s	23.83dB	8.97
	화소영역의 처리방법	750Kbits/s	23.08dB	11.5
		750Kbits/s	22.84dB	9.3
화소영역의 처리방법	500Kbits/s	22.89dB	14.3	
	500Kbits/s	22.59dB	8.45	
PARIS	화소영역의 처리방법	1.5Mbits/s	30.39dB	7.31
		1.5Mbits/s	31.08dB	2.81
	화소영역의 처리방법	1.0Mbits/s	28.31dB	11.29
		1.0Mbits/s	28.58dB	7.03
	화소영역의 처리방법	750Kbits/s	26.77dB	14.32
		750Kbits/s	26.82dB	9.11
화소영역의 처리방법	500Kbits/s	25.66dB	15.59	
	500Kbits/s	25.33dB	10.72	

영상의 종류	알고리즘	변환 비트율	객관적 화질 (평균 PSNR)	주파적 화질 (평균 QS 표준 편차)
AKIYO	화소영역의 처리방법	1.5Mbits/s	43.79dB	1.6
		1.5Mbits/s	43.73dB	1.7
	화소영역의 처리방법	1.0Mbits/s	41.56dB	2.4
		1.0Mbits/s	41.55dB	1.6
	화소영역의 처리방법	750Kbits/s	40.3dB	3.5
		750Kbits/s	39.8dB	2.1
화소영역의 처리방법	500Kbits/s	37.2dB	7.4	
	500Kbits/s	36.6dB	3.8	
Table TENNIS	화소영역의 처리방법	1.5Mbits/s	22.6dB	4.05
		1.5Mbits/s	22.0dB	4.02
	화소영역의 처리방법	1.0Mbits/s	31.12dB	5.06
		1.0Mbits/s	30.91dB	4.82
	화소영역의 처리방법	750Kbits/s	29.41dB	7.13
		750Kbits/s	29.40dB	5.77
화소영역의 처리방법	500Kbits/s	28.41dB	10.27	
	500Kbits/s	28.32dB	9.11	

표3과 표4는 각각 두 방법에 의해 처리된 전체 스트림 그리고 모든 변환 비트율에 대한 평균 PSNR 차이와 평균 양자화 스텝의 표준편차의 차이

를 나타낸다. 표에 보인 바와 같이 평균 PSNR은 -0.44dB에서 +0.13dB의 차이를 보이며, 양자화 스텝의 표준편차는 제안한 알고리즘이 전 구간에서 0.69에서 4.17 만큼 낮은 값을 나타낸다.

표 3. 화질 차이

스트림 종류	평균 PSNR (제안한 알고리즘)	평균 PSNR (화소영역의 방법)	PSNR 차이
AKIYO	40.50dB	40.72dB	0.22dB
COAST GUARD	30.39dB	30.56dB	0.17dB
CONTAINER	34.69dB	35.13dB	0.44dB
HALL MONITOR	34.96B	35.00dB	0.04dB
MOTHER & DAUGHTER	38.07dB	38.49dB	0.42dB
PARIS	29.92dB	27.79dB	-0.13dB
TABLE TENNIS	30.15dB	30.40dB	0.25dB
CALENDAR	23.74dB	23.91dB	0.17dB

표 4. QS 표준 편차

스트림 종류	평균 QS 편차 (제안한 알고리즘)	평균 QS 편차 (화소영역의 방법)	QS 편차의 차이
AKIYO	2.17	3.77	-1.6
COAST GUARD	5.94	7.70	-1.76
CONTAINER	4.79	9.01	-4.22
HALL MONITOR	4.12	7.19	-2.99
MOTHER & DAUGHTER	2.51	3.30	-0.79
PARIS	7.91	12.08	-4.17
TABLE TENNIS	5.94	6.63	-0.69
CALENDAR	8.52	12.59	-4.07

4.1.2 구현의 복잡도

비트율 변환을 위해 본 알고리즘에서 사용하는 모듈은 VLC, VLD, 역양자화 그리고 적응형 비트율 할당함수를 이용하여 매크로블럭 단위로 비트를 할당하고 이 비트를 표현하는 양자화 파라미터 값을 찾는 것으로 구성된다.

이것을 화소영역의 처리방법과 비교하면, 압축 동영상의 디코딩에 소요되는 DCT-1, 움직임 보상부와 화소신호를 다시 인코딩할 때, 사용되는 움직임 추정과 DCT 연산절차가 제외되는 이득을 갖는다. 이것을 단순히 코덱의 복잡도의 관점에 비교해보면, 전체 연산의 60~80% 이상을 차지하는 복잡한 절차가 생략되어, 상당부분의 연산에 대한 개선효과를 갖는다고 볼 수 있다. 전술한 바와 같이 본 알고리즘에서 비트율 변환을 위해 사용하는 알고리즘 중, 가장 큰 비중을 갖는 부분은 적응형 비트율 할당 모델에 따르는 부분과 bisection 방법에 의한 양자화, VLC 과정이다. 이것에 대한 표현은 표5에 보인

것과 같다.

표 5. 비트율 변환 알고리즘의 복잡도

화소영역	연산 과정	픽셀당 곱셈의 수	픽셀당 덧셈의 수
	DCT, DCT ⁻¹	$2(\log_2 N - 3) + 8/N$	$30\log_2 N - 1 + 4/N$
	MC, MC ⁻¹	0	1
	Q ⁻¹	1/β	0
	Q	1	0

제한한 알고리즘	연산 과정	픽셀당 곱셈의 수	픽셀당 덧셈의 수
	DCT, DCT ⁻¹	0	0
	MC, MC ⁻¹	0	0
	Q ⁻¹	1/β	0
	Q	$M \times (1/\beta)$	0
	IACT	0	1/β
	ABBF	$1/(\beta \times 376)$	0
	VLC	0	0

- β : Total number of the DCT coefficient / number of the non zero DCT coefficient
- N : DCT block width (N=8)
- M : Number of bisection (≤6)
- ABBF : Complexity of adaptive Bit-rate Budget Function
- +α : Additional complexity portion

이 표 중 화소영역에 대한 것은 일반적인 하이브리드 코덱에 대한 연산의 복잡성을 간략하게 모델링한 것이다. 그러나, 이것은 모든 연산의 과정을 덧셈과 곱셈으로 단순화 시키고, MPEG-2, H.263의 가장 복잡한 처리과정인 반화소 움직임 추정, 보상, 그리고 양방향 움직임 벡터처리에 대한 연산이 빠져 있다. 이점을 감안하여 두 방법의 복잡도를 비교하면, 본 알고리즘의 VLD, 역양자화는 화소영역과 동일하며, 추가적인 처리로 매크로블럭의 AC 계수당, IACT와 비트율 할당을 위해 각각 1/β, 1/(β x 376)번의 덧셈과 곱셈이 사용되며, bisection에 따르는 추가적인 연산이 포함된다. 그러나, 이 부분 중, 전자의 IACT와 비트율 할당에 대한 연산은 인코더의 화소단위의 율제어에 사용되는 데이터 분석 모델과 대응되므로 실제로는 추가적인 연산으로 볼 수 없다. 즉 본 알고리즘이 별도로 사용하는 것은 양자화, VLC 절차로, 일반적으로 bisection을 통한 방법은 최대 6번의 절차가 소요되나, 본 고에서 제안한 초기 양자화 파라미터 선정방법을 사용하면 평균 2~3회 이내에서 최적 양자화 파라미터 값을 찾을 수 있다. 따라서 화소영역의 복잡도에 비교해 이 부분은 무시할 수 있으며, 결론적으로 전술한 바와 같이 제안한 알고리즘은 기존 방법에 비해 최대 80%이상의 연산의 개선을 가져오는 효율적인 방법이라 할 수 있다.

4.1.3 실시간 처리

제안한 알고리즘은 비트율 변환을 위해 프레임 단위의 IACT 해석을 하고, 이 결과에 의해 몇 단계 양자화 절차를 거쳐 매크로블럭에 대한 최적 양자화 파라미터를 결정한다. 이것을 화소영역의 처리방법에 소요되는 지연과 비교하면 전체지연은 식(26)으로 나타낼 수 있고, 이중 인코딩 지연은 MPEG-2의 B 프레임을 사용할 때 다음과 같이 표현할 수 있다.

$$DELAY_{encoding} = D_C^P(\geq 33ms) + D_C^W(\approx 0) \tag{30}$$

통신에 따르는 $DELAY_{channel}$ 중 D_C^P 는 본 알고리즘에 동일하게 적용되므로 고려하지 않는다. Common Intermediated Format(CIF)의 해상도를 갖는 영상에 대한 송수신 완충버퍼 용량을 450Kbits를 설정하고, 점유도를 50%로 가정하면 송수신 완충 버퍼링에 따르는 지연은 다음과 같이 쓸 수 있다.

$$D_C^P = D_{Txbuffering} + D_{Rxbuffering} \tag{31}$$

$$D_C^P \approx \frac{225kbts}{2Mbps} + \frac{225kbts}{2Mbps} \approx 225ms$$

마지막으로 디코딩 지연을 계산하면, $DELAY_{decoding}$ 는 식(29)와 같이 D_D^P 와 D_D^W 로 구성되는데 이중 D_D^P 는 33ms 이상의 값을 갖고, D_D^W 는 양방향 움직임 정보를 사용하는 B 프레임이 있으므로 다음과 같다.

$$DELAY_{decoding} = D_D^P(\geq 33ms) + D_D^W(\approx 66ms), \tag{32}$$

$$DELAY_{decoding} \geq 99ms$$

식(30), 식(31) 그리고 식(32)에서 구한 값을 모두 더하면 화소영역에서 발생하는 전체지연을 구할 수 있다.

$$DELAY_{total} \geq 33ms + 225ms + 99ms \geq 357ms \tag{33}$$

반면에 제안한 알고리즘에서 발생하는 지연은 영상을 복부호 처리하지 않으므로 인코딩, 디코딩 지연은 없고, 특히 입력 CBR 데이터의 비트율 특성을 그대로 유지하면서 변환하므로 송수신 완충버퍼가 필요없어 버퍼링 지연은 0가 된다. 즉 전체지연은 한 프레임을 수신하여 분석하는데 사용되는 지연, 양자화 그리고 VLC에 소요되는 값으로 식(34)와 같다.

$$DELAY_{end\ to\ end} = DELAY_{encoding(0)} + DELAY_{channel(0)} + DELAY_{decoding(0)} + DELAY_{frame}(\approx 33ms) + DELAY_{msQ} + DELAY_{msVLC} \tag{34}$$

이중 $DELAY_{msQ}$ 와 $DELAY_{msVLC}$ 는 최적 양자화 값을 찾기 위한 *bisection* 과정에 필요한 지연으로 소프트웨어로 구현할 경우에는 무시하지 못할 값을 갖으나, 하드웨어로 알고리즘을 구현할 경우, 무시할 만한 값으로 처리가 가능하다. 화소영역의 지연은 MPEG-2의 일반적인 인코딩 방법을 사용한 것을 전제로 하여, B 프레임에 대한 처리지연을 포함시켰다. 그러나 MPEG-2가 B를 포함하지 않을 경우 전체 지연은 66ms를 뺀 값이 된다. 이제 화소영역의 비트율 변환 방법과 본 알고리즘에 소요되는 전체지연을 비교해 보면, 화소영역은 B 프레임을 사용할 때와 사용하지 않을 때, 각각 357ms, 291ms 이상의 지연값을 갖는다. 반면에 본 알고리즘은 B 프레임의 존재 유무에 관계없이 33ms에 수렴하는 지연치를 가져 화소영역의 처리방식에 대비하여 처리지연이 크게 감소되는 결과를 낳는다.

V. 결론

인터넷 같이 네트워크의 대역폭에 대한 QoS를 보장할 수 없는 상황에서 압축 동영상을 전송하기 위해서는 대역폭의 변동에 따라 유연하게 동영상의 비트율을 맞추어 주어야 한다. 그 외 압축부호의 비트율 변환은 다양한 용도로 사용할 수 있으나, 비트율 변환에 사용되는 알고리즘은 급격히 변동하는 대역폭에 비트율을 정확히 맞추어주는데 어려움을 갖거나, 복잡한 연산과정과 처리에 많은 시간이 소요되어 실시간 통신에는 부적합한 알고리즘이 대부분이다.

본 논문은 실시간 멀티미디어 통신에 적합한 비트율 변환 알고리즘을 제안하였다. 제안한 알고리즘은 비트의 변환을 압축영역에서 행하며, 기존 방법에 비해, 변환된 비트율의 수렴속도 및 정확도, 알고리즘에 사용되는 연산, 그리고 실시간 처리에 장점을 갖는다. 제안한 알고리즘은 압축영역에서 재양자화를 통해 비트율을 변환하며, 이 재양자화는 프레임 구성하고 있는 매크로블럭의 양자화 파라미터를 다른 값으로 바꾸어 주는 것이다. 동영상을 이루고 있는 각 요소는 합당한 값으로 표현되어야 하는 고유한 엔트로피가 있으며, 이 값으로 표현되었을 때, 최적의 결과를 얻을 수 있다. 본 알고리즘은

매크로블럭에 대한 비트 할당을 위해 영상의 내용을 나타내는 척도가 되는 IACT를 사용하였고, 이 IACT를 이용하여 매크로블럭에 비트를 할당하는 비트 할당함수를 구하였다. 그러나 비트 할당함수는 매크로블럭의 비트를 할당하는 기준이 될 수 있으나, 프레임과 영상의 내용에 따라 편차가 있는 모델은 모든 유형의 영상에 적용하기에는 문제를 갖는다. 이 문제를 해결하기 위해 프레임의 유형과 IACT의 범위에 따라 고유한 함수를 사용하고, 입력 스트림의 매크로블럭에 대해 합당한 비트함수를 사용하는 적응형 비트할당 모델과 이를 기준으로 비트율을 변환하는 알고리즘에 대해 제안하였다.

참고 문헌

- [1] Heinrich J. Stuttgart, "Network Evolution and Multimedia Communication," *IEEE Multimedia*, pp. 42-59, Fall 1995.
- [2] L. C. Wolf, C. Griwodz, R. Steinmetz, "Multimedia Communication," *Proceedings of the IEEE*, VOL. 85, NO. 12, pp. 1915-1933, Dec. 1997.
- [3] M. Decina, "Convergence of Telecommunications and Computing to Networking Models for Integrated Services and Application," *Proceedings of the IEEE*, VOL. 85, NO. 12, pp. 220-227, Dec. 1997.
- [4] M. H. Sherif, "Multimedia Networks and the Public Switched Telephone Network," *IEEE Communications Magazine*, pp. 92-94, Jan. 1996.
- [5] R. Wittmann, M. Zitterbart, "Towards Support for Heterogeneous Multimedia Communications," *Proceeding of Distributed Computing System*, pp. 336-341, 1997.
- [6] M. Hogan, S. M Radack, "The Quest for Information Technology Standards for the Global Information Infrastructure," *Standard-View*, VOL. 5, NO. 1, pp. 30-35, Mar. 1997.
- [7] C. Dabrowski, W Majurski, W Mccoy, S Wakid, "Modeling NII Services: Future Needs for Standards and Interoperability," *StandardView*, VOL. 2, NO. 4, pp. 283-217, Dec. 1994.
- [8] H. Vin, "Heterogeneous Networking," *IEEE Multimedia*, pp. 84-87, April/June 1997.
- [9] K. Asatani and S. Nogami, "Trends in the Standardization of Telecommunications on GII, Multimedia, and Other Network Technologies and Services," *IEEE Communication Magazine*, pp. 32-46, June 1996.
- [10] C. C. Bisdikian, F. A Schaffa, "The Multimedia gateway: An Evolution towards Switched Video Services," *Proceedings of Local Computer Network*, pp. 54-58, 1996.
- [11] I.Mizunuma, C. Shen, M. Takegaki, "Middleware for Distributed Industrial Real-Time Systems on ATM Networks," *17th IEEE Real-Time Systems Symposium*, pp. 32-38, 1996.
- [12] P. Haskell, D. G. Messerschmitt, K. Nichols, E. Pol, R. Morley, P. Jardtetzky, "Multimedia Consumer Applications on a Heterogeneous Network," *Proceedings of 1st IEEE International Workshop on Community Networking*, pp. 229-237, 1994.
- [13] S. Mccanne, E. Brewer, R. Kata, L. Rose, "Toward a Common Infrastructure for Multimedia-Networking Middleware," *Proceeding of NOSSDAV*, pp. 39-49, 1997.
- [14] N. Yeadon, F. Garcia, D. Hutchison, D. Shephered, "Filter: QoS Support Mechanisms for Multipeer Communications," *IEEE Journal on Selected Areas in Communications*, VOL. 14, NO. 7, pp. 1245-1262, SEP. 1996.
- [15] A. Balachandran, A. T. Campbell, M. E. Kounavis, "Active Filters: Delivering Scaled Media to Mobile Devices," *Proceedings of NOSSDAV*, pp. 125-134, 1997.
- [16] S. Li, B. Bhargava, "Active Gateway: A Facility for Video Conferencing traffic Control," *Proceedings of COMPSAC*, pp. 308-311, 1997.
- [17] O. H. Werner, "Generic Quantizer for Transcoding of Hybrid Video," *Proceedings of the 1997 Picture Coding Symposium*, Sep. 1997.
- [18] V. Ratnakar, M. Livny, "RD-OPT: An Efficient Algorithm for optimizing DCT Quantization Tables," *Proceeding of DCC*,

pp.1-11, 1995.

[19] G. Keesman, R. Hellinghuizen, F. Hoeksema, G. Heideman, "Transcoding of MPEG bitstreams," *Signal Processing: Image Communication*. pp. 481-500, 1996.

[20] P. J. Brightwell, "Flexible Switching and editing of MPEG-2 Video Bitstreams," *IEE Conference Publication NO. 447*, pp.547-552, 1997.

[21] P. A. A. Assuncao, M. Ghanbari, "Optimal Transcodind of Compressed Video," *Proceeding of Image Processing*, pp.739-742, 1997.

[22] P. A. A. Assuncao, M. Ghanbari, "Fast Computation of MC-DCT for Video Transcoding," *Electronics letters*, pp.284-286, 1997.

[23] P. A. A. Assuncao, M. Ghanbari, "Transcoding of MPEG-2 Video in the Frequency Domain," *Proceeding of ICASSP*, pp.2633-2636, 1997.

[24] J. H. Conway, N. J. A. Sloane, *Sphere packing, Lattices and Groups*, Springer-Verlag, New York, 1988.

[25] T. Urabe, H. Afzal, Grace Ho, P. Pancha, M. E. Zaeki, "MPEGTool: An X Window Based MPEG Encoder and Statistics Tool," *Proceeding of ACM Multimedia*, pp.510-514, 1993.

[26] K. Tarlreja, P. V. Rangan, "Editing Techniques for MPEG Multiplexed Streams," *IEEE International Conference on Multimedia Computing and Systems*, pp.278-285, 1997.

[27] Jorn Sotermann, "Methodologies used for evaluation of video tools and algorithms in MPEG-4," *Signal Processing Image Communication* 9, pp.343-365, 1997.

황 대 환(Dae-hwan Hwang)

정회원



1986년 2월 : 성균관대학교
전자공학과(공학사)

1988년 2월 : 성균관대학교
전자공학과 통신공학
(공학석사)

1999년 2월 : 성균관대학교
전자공학과 통신공학
(공학박사)

1991년 2월~현재 : 한국전자통신연구원 멀티미디어
통신팀장, 책임연구원

<주관심 분야> 멀티미디어 네트워킹, 디지털 신호
처리

조 규 섭(Kyo-seob-Cho)

정회원



1974년 2월 : 성균관대학교
전자공학과(공학사)

1976년 2월 : 성균관대학교
전기공학과 통신공학
(공학석사)

1987년 2월 : 성균관대학교
전자공학과 통신공학
(공학박사)

1977년~1992년 : 한국전자통신연구원 (연구위원)

1992년 2월~현재 : 성균관대학교 전기전자컴퓨터공
학부 교수

<주관심 분야> 데이터 통신

황 수 응(Su-young Hwang)

정회원

홍익대학교 전자공학과(공학사)

명지대학교 전자공학과 전자공학(공학석사)

명지대학교 전자공학과 (공학박사)

현재 : 서울산업대학교 전자공학과 교수

<주관심 분야> 생체신호, 디지털 신호처리