

웨이브렛 변환을 이용한 관심영역의 부호화

Transmission of the Region of Interest in Images Using Wavelet Transform

이 수 중 (Lee, Soo Jong)* · 이 완 주 (Lee, Wan Ju)** · 김 용 규 (Kim, Yong Kyu)***

목 차

- I. 서론
 - II. 본론
 - III. 결론
-

Abstract

Region-of-Interest is the region within the image selected for the users needs. The development of multimedia has made the expectation of image telecommunication higher, but the usage of the image, image transmission time, and image storage create problems. When transmitter or the receiver stops transmission at some point, we can still see the general image and the ROI maintains better image quality if the ROI is specified beforehand. In this paper, three methods are proposed and constructed for the transmission of ROI. In the first method, the ROI and the background are separated and then encoded as described above. The second method is to encode without separating the ROI and the background. The masked region is scaled and the coefficients are increased, then the region is transmitted first. The third method is the loseless coding of the ROI. For loseless coding, real number tap cannot be restored perfectly due to the rounding error, so the method of using integers is used. The proposed method shows a better performance than EZW even in case of ROI's PSNR at quality of 40 dB.

Key words: Wavelet Transform, EZW, ROI

* 협성대학교 컴퓨터공학과 조교수

** 용인대학교 컴퓨터정보학과 부교수

*** 성결대학교 정보통신공학부 교수

I. 서론

현재의 정지영상 압축표준인 JPEG(Joint Photographic Expert Group)[1]은 이산 여현 변환(discrete cosine transform: DCT)[1]을 기반으로 하여 mid-bit-rate와 high-bit-rate 에는 좋은 성능을 보이고 있으나 0.25 bpp 이하의 low-bit rate 압축에서는 현저한 블록화 현상(blocking effect)으로 인하여 영상 정보로서의 가치를 기대할 수 없다. JPEG의 단점을 보완하여 정지영상 뿐만 아니라 향후 요구되어지는 새로운 요구인 동영상 처리와의 호환을 통해 영상 압축 및 전송의 새로운 표준으로 자리 잡기 위해 국제 표준화 기구(ISO) 및 국제 전기 기술 위원회(IEC)등 공동위원회 산하에서 JPEG과 JBIG(Joint Bilevel Image Expert Group)을 대체할 국제 표준화로 JPEG 2000을 추진 중이다. JPEG 2000의 주요 목적은 기존의 JPEG의 한계인 사진, 그래픽, 의료 영상, 원격 감지 영상(인공위성 사진) 및 텍스트 등의 다양한 특성(bi-level, gray-level, color)을 갖는 정지화상을 하나의 시스템 하에서 주문자의 다양한 요구에 대해 하나의 구조를 적용하는 것이다. 또한 0.25 bpp 이하의 low bit-rate 전송을 가장 큰 목적으로, 정지화상을 고밀도로 압축하여 전송하고 복원시 원본의 화질에 손색없는 영상 화질의 향상을 제공하여 대용량의 데이터베이스 구성 및 검색, 네트워크 상에서의 점진적인 화상전송, 상업용 사진에 대한 저작권 보호, 그리고 MPEG-4 (Motion Picture Expert Group)와의 호환 등이다.

영상 인식 분야(computer vision)에서 발전되어온 다해상도 신호 해석(multiresolution signal analysis)[4][5]은 움직임 추정에서 물체 인식에 이르기까지 여러 가지 문제를 해결하는데 중요한 역할을 하였다. 이는 입력신호를 서로 다른 해상도로 표현하여 공간 주파수 특성에 따라 각각의 해상도에 적합한 처리를 하는 것으로 제로트리(zerotree)를 이용한 부호화 방법은 웨이브렛 변환된 영상에서 상위 단계의 계수값이 양자화 임계값(threshold)보다 작다면 공간적으로 대응하는 하위 단계의 계수값도 작을 것이라는 가정하에 부호화 하는 방법으로 매우 효율적인 압축 성능을 보인다. Shapiro[6] 웨이브렛 분해된 같은 방향의 부대역간의 자기 유사성(self-similarity)을 이용한 연속적인 근사 양자화(successive approximation quantization: SAQ)를 사용한 EZW(embedded zerotree wavelet) 알고리즘을 제안하였고, Said and Pearlman은 Shapiro의 방법을 개선하여 더욱 압축 효율을 높인 SPIHT(Set Partitioning in Hierarchical Trees) 알고리즘을 제안하였다[7]. Embedded 영상 부호화기는 임의의 순간에 부호화와 복호화를 멈출 수 있는 특성을 가지고 있기 때문에 정확한 압축률의 조정이 가능하고, 점진적인 전송(progressive transmission)이 가능해 정지영상 및 동영상 부호화에 이용될 수 있다.

본 논문에서는 이러한 웨이브렛을 이용한 영상 부호화의 점진적인 전송의 특징을 살려 영상의 어느 특별한 부분만을 임의적으로 선택하여 원하는 비트율에서 원하는 속도로 영상 통신을 할 수 있는 관심영역(Region of Interest)의 부호화에 대한 연구를 하였다. 영상의 어떤 특별한 부분을 상세하게 보기 위해서는 그 부분에 더 많은 비트를 할당함으로써 고해상도의 영상을

획득할 수 있다. 이것은 영상의 관심 영역만 추출하여 그 부분에 보다 많은 비트를 할당하고 그 영역을 먼저 전송함으로써 영상의 대강의 형태를 알 수 있다는 장점을 가지고 있다. 이것의 응용으로는 인간의 얼굴, 쓰여진 단어, 부분적인 배경 및 군에서의 목표물 인식 등 네트워크(network)를 통한 영상 통신의 경우에 매우 다양하게 사용자의 요구에 응할 수 있는 시스템에 이용될 수 있다. 몇몇 응용에서는 무손실 압축이 요구된다[8]. 예를 들어, 의료 영상을 이용하는 원격 의료 환경에서는 정확한 진단을 위해서 점진적인 전송과 함께 무손실 압축이 요구된다. 그러나 이러한 요구는 매우 큰 저장장소를 요구하게 된다. 이것을 해결하기 위한 하나의 방법으로 관심 영역에서는 완전복원과 함께 빠른 전송이 요구되고 나머지 영역에서는 사용자의 요구에 의한 효율적인 압축 방법이 요구된다. 완전복원을 이루기 위해서는 웨이브렛 필터의 실수 탭은 유한 허용 오차로 인한 절사오류(truncation error)로 인해 완전 복원이 불가능해 정수를 사용하는 필터를 사용하여 완전복원을 가능하게 한다. 정수 웨이브렛 필터는 계산이 간단하고 속도가 빨라 소프트웨어와 하드웨어의 구현이 쉽다. 이러한 정수를 이용하는 웨이브렛은 Said and Pearlman의 집합 분할 정렬(Set Partitioning in Hierarchical Trees: SPIHT)의 S+P 알고리즘이 알려져 있고, 그 외에 JPEG 2000에 제출된 알고리즘 중에서 RICOH의 CREW에서 사용된 변환인 S(sequence), TT(Two-Ten) 변환이 알려져 있다[9].

관심영역의 부호화의 특징은 임의의 모양을 가진 영역이 설정 될 수 있어야 하고, 관심영역은 점진적인 전송에 있어서 나머지 부분보다 더 빨리 전송될 수 있어야 하고, 관심영역의 지정과 비트율이 부호화전에 또는 부호화 중간에 사용자에게 의해 유연하게 조절이 가능해야 한다. 그리고 관심영역의 화질은 그 나머지 부분보다 나아야 하고 관심영역을 형성하고 있는 경계부분이 시각적으로 훼손되지 않아야 하기 때문에 영상의 양자화, 엔트로피 부호화 그리고 관심영역의 형태 부호화는 모두 독립적이다.

본 논문에서는 관심영역의 부호화를 위하여 크게 세 가지의 방법을 통하여 구현해 본다. 첫 번째 방법에서는, 관심영역과 그 나머지영역을 분리하여 관심영역과 나머지 영역을 각각 부호화 하여 각각의 비트스트림으로 전송하는 방법으로 이는 필요에 따라서 관심영역의 부호화는 무손실 부호화 방법을 이용할 수도 있다. 두 번째 방법으로는, 영상에서 영역을 설정한 후에 웨이브렛 변환을 한후 계수들의 크기를 크게 하여 계수들의 재정렬을 통해 관심영역이 먼저 전송되게하고 복원단에서는 원래의 계수값을 회복하는 방법이다. 위와 같은 두 가지의 방법은 손실 부호화 방법인 실수탭을 사용해 부호화 효율을 높인다. 세 번째 방법으로는, 관심영역의 완전복원을 위한 방법으로 SPIHT의 S+P를 이용하여 실험을 하였다. 첫 번째와 두 번째의 실험방법으로는 Shapiro의 EZW 알고리즘이 매우 효율적이고 계산이 간단하여 먼저 구현하여 보고, 이를 개선한 SPIHT의 알고리즘으로 기존의 JPEG 및 EZW, SPIHT의 동일 비트율에서의 관심영역이 있을 때와 없을 때를 인간의 시각적인 관점에서 비교하였다. 또한, 첫 번째 방법에서 마스크를 씌워 잘라낸 부분을 0으로 채웠을 때와 그냥 원래대로 놔두고 부호화를 행했을

때를 비교하여 경계부분의 일그러짐을 관찰하고 효율적인 방법을 제시한다.

II. 본론

1990년대 초에 영상의 압축에 웨이블릿을 이용한 점진적인 영상의 전송(Progressive Image Transmission)에 관한 논문들이 발표되면서 영상의 다해상도 해석(Multiresolution Analysis)에 관한 관심이 고조되었다[44][45]. 영상의 다해상도 해석을 이용한 점진적인 영상의 부호화는 인터넷과 같은 통신 수단의 발달로 인한 사용자의 다양한 욕구를 충족시킬 수 있는 하나의 수단으로, 관심영역(Region of Interest)의 부호화는 점진적인 영상 부호화의 기반 위에 이를 수 있다[10].

점진적인 영상의 전송은 전체 영상을 다 전송하지 않고 그 일부만으로도 대략적인 영상의 재구성이 가능하다. 이런 형태의 부호화는 높은 압축률과 빠른 속도를 요구하는 영상 자료의 통신인 원격 의료 진료, 영상 데이터 베이스, 낮은 대역폭을 통한 시스템에서의 화상회의 등에 적절한 방법으로 수신자에 따라서 화질의 충실도(fidelity)와 전송시간의 고려를 달리 할 수 있는 곳에 이용이 가능하다.

관심영역을 가진 영상의 전송이란, 점진적 전송의 기반 위에 어느 특정한 영역을 사용자가 선택하여 전송할 수 있는 방법으로 단순한 점진적 전송의 방법보다 압축률과 전송시간을 단축시킬 수 있을 뿐만 아니라 효율적인 메모리 관리 등 사용자의 다양한 요구에 응할 수 있다. 관심영역 부호화의 특징은, 관심영역이 그 나머지영역(배경영역)보다 먼저 전송되어야 하며, 관심영역의 화질은 그 나머지 영역보다 좋아야 하고, 전체적인 영상의 비트율과 관심영역의 비트율도 조절이 가능해야 한다. 또한, 영역의 선택에 있어서 직사각형, 원 및 임의의 모양을 선택하여 부호화, 복호화가 독립적으로 수행된다. 이는 마스크를 이용하여 수행되며 마스크에 대한 정보는 비트스트림 헤더(Header)에 부가적인 정보로 제공된다.

관심영역의 부호화는 다음과 같은 특성을 만족해야 한다.

- 점진적인 부호화 방법이다.
- 임의의 모양을 가진 영역 설정이 가능해야 한다.
- 관심영역 영역의 화질은 배경영역보다 좋아야 한다.
- 관심영역을 형성하고 있는 경계부분은 관심영역을 둘러싼 부분과 시각적으로 급격한 차이를 보이지 않아야 한다.
- 관심영역은 영상의 부호화 초기에 또는 부호화 중간에 지정될 수 있어야 한다.
- 관심영역은 배경영역보다 빨리 전송되어 화질이 배경영역보다 먼저 정제(refinement) 되

고, 그 속도는 사용자에게 의해 유연(flexible)하게 조절이 가능해야 한다.

- 구현 방법은 양자화(Quantization), 엔트로피 부호화 그리고 관심영역의 모양(shape) 부호화는 모두 독립적이다.
- 관심영역의 계수(Coefficient)들의 부호화는 관심영역이 없을 때와 유사한 방법으로 수행된다.
- 칼라 영상의 경우에는 각 성분별로 분리되어 적용된다.

1. 관심영역의 부호화

관심영역의 부호화는 특성상 영상의 점진적인 전송을 기반으로 한다. 그래서 이 분야의 연구 논문 발표들은 대부분 점진적 영상 부호화의 방법으로 문헌적으로 증명된 잘 알려진 기법들을 이용하여 부호화 방법을 제시하게 된다. 또한, 관심영역의 부호화는 점진적인 영상의 부호화에 비해 상대적으로 연구가 부족해서 실제로 비교 대상이 될 수 있는 알려진 방법이 없는 것이 사실이다. 그런 이유로 이 분야의 발표자들은 자체 비교를 하거나 점진적인 영상의 부호화의 방법으로 잘 알려진 부호화 기법들의 결과와 비교를 하게 된다. 이것은 관심영역에 이미 많은 비트를 할당했기 때문에 관심영역의 화질은 관심영역이 없는 부호화 방법에 비해 좋을 수밖에 없어 이것의 비교는 사실 무의미하고 관심영역이 있는 영상의 전체적인 PSNR(peak signal to noise ratio)과 비교를 해야 하는데 이것은 관심영역에 할당된 비트로 인해 전체적인 PSNR은 떨어지게 된다. 따라서, 이것의 판단은 개개인의 특성에 의존할 수밖에 없다. 이런 이유로 사용자의 다양한 요구를 충족시켜주는 시스템이 필요하다. 이러한 시스템은 무선통신과 인터넷의 발달로 인한 영상정보의 다양한 이용과 대용량의 정보를 포함하고 있는 많은 영상의 전송, 영상을 이용한 데이터 베이스의 구축과 같은 시스템에 앞으로 얼마든지 이용될 수 있는 사항이고, 또한, 낮은 대역폭을 갖는 초 저속 전송 매체의 응용에도 응용이 가능하여 앞으로 이 분야의 연구는 동영상에 이용될 수 있는 방법의 연구가 필요하다.

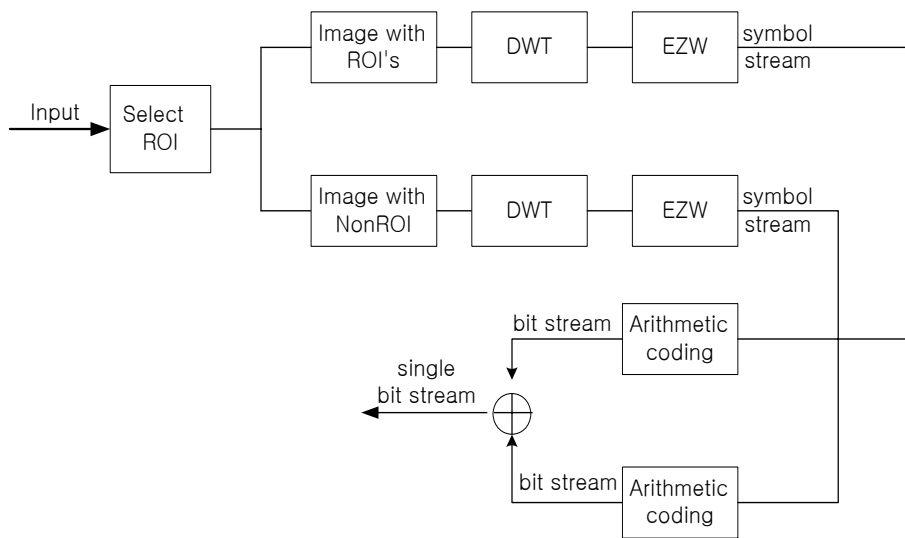
본 논문에서 제안하는 방법은 크게 세 가지로 첫 번째 방법은, 영상의 관심영역과 배경영역을 분리하여 부호화하는 방법으로 여기에는 다시 관심영역으로 지정된 부분을 분리하였을 때 그 부분의 처리를 위한 방법을 두 가지로 나누어 비교하였다. 두 번째 방법으로는, 관심영역과 배경영역의 통합 부호화로 관심영역의 계수들의 값을 배경영역의 값들보다 크게 하여 먼저 전송시키는 방법으로 관심영역의 고 화질을 보장할 수 있는 방법을 연구하였다. 세 번째 방법으로는, 관심영역의 무손실 부호화 방법으로 정수를 이용하는 웨이브렛 알고리즘을 이용하여 구현하였다.

1) 관심영역과 배경영역의 분리 부호화

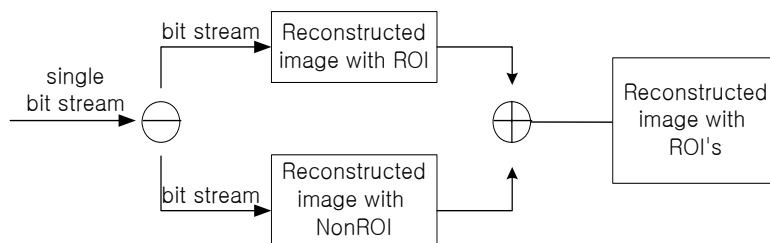
관심영역의 부호화는 가장 단순하게는 관심영역과 배경영역을 별도의 두 장의 영상으로 분리하여 각각 다르게 부호화하는 방법이 있을 수 있다. 그림 1.에 이 방법의 블록도를 나타내었다.

즉, 관심영역이 있는 영상에 대해서는 세밀한 양자화를 수행하여 영상의 질을 높이고, 배경 영상에 대해서는 덜 세밀한 양자화를 수행하여 압축 전송하는 방법이다. 이 방법은 마스크를 이용하여 관심영역과 배경영역을 분리하게 되고 이 마스크에 대한 정보는 부호화시 헤더에 포함되어 복호화기로 보내지게 된다. 이렇게 각각 압축되어진 두 장의 영상은 관심영역에 대한 위치정보와 함께 두 개의 비트스트림을 융합(merge)하여 하나의 비트스트림으로 만들어 부호화기로 전송되고, 복호화기에서는 하나의 비트스트림을 다시 두 개로 분리하여 부호화기에서의 마찬가지로 별도의 두 개의 비트스트림을 복원하여 하나의 영상으로 재구성하게 된다.

본 논문에서는 SPIHT알고리즘을 수정하여 관심영역의 부호화가 가능하도록 하였다. 실험에 사용한 부호화 및 복호화 알고리즘은 아래와 같다.



(a) Encoder



(b) Decoder

<그림 1> 제안된 방법 1의 블록도

부호화 알고리즘. I_{ROI} I_{NonROI}

1. 영상내에서 관심영역의 지정
2. 관심영역과 배경영역을 별도의 영상으로 분리 (I_{ROI} , I_{NonROI})
: 관심영역에 대한 마스크를 생성한다.
배경영역으로 구성된 영상(I_{NonROI})내에는 관심영역이 포함된다. I_{NonROI} 내에 포함된 관심영역을 원래의 영상으로 그대로 두고 부호화하는 방법이 있을 수 있고, 0으로 채우고 부호화하는 방법이 있을 수 있다. I_{NonROI} 내의 관심영역의 채움방식에 따라 부호화 성능이 약간의 차이를 가져올 수 있다.
3. I_{ROI} 에 대해 비트율을 M으로 설정하여 SPIHT부호화를 수행한다.
결과 비트스트림을 B_{ROI} 라 한다.
4. I_{NonROI} 에 대해 비트율을 m으로 설정하여 SPIHT부호화를 수행한다.
결과 비트스트림을 B_{NonROI} 라 한다. 당연히 $M > m$ 이 된다.
5. 두 개의 비트스트림을 하나의 비트스트림으로 융합(merge)하여 전송한다.

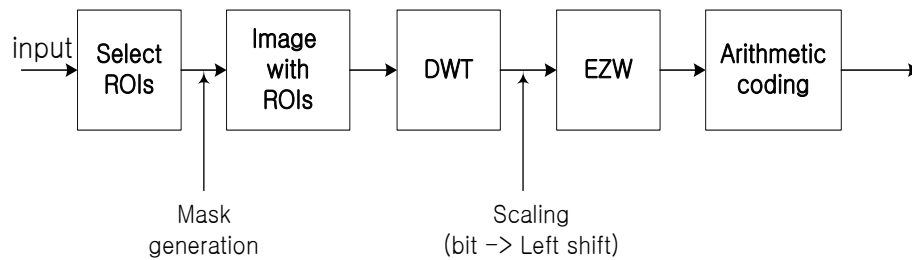
복호화 알고리즘. I_{ROI} I_{NonROI}

1. 전송된 비트 스트림을 B_{ROI} 와 B_{NonROI} 로 분리한다.
2. B_{ROI} 를 복호화하여 I_{ROI} 를 얻어내고, B_{NonROI} 를 복호화하여 I_{NonROI} 를 구성한다.
3. I_{ROI} 와 I_{NonROI} 를 위치정보를 이용하여 하나의 영상으로 재구성한다.

이 방법에 대한 문제점으로는 관심영역과 배경영역을 분리하여 분리된 2개의 영상에 대해 각각 별도의 부호화기, 복호화기를 사용한다는 것이다. 이 문제점을 해결하기 위한 두 번째 방법으로 관심영역과 배경영역을 통합해서 부호화를 해보았다.

2) 관심영역과 배경영역의 통합 부호화

이 절에서는 관심영역과 배경영역을 따로 분리하지 않고 관심영역을 먼저 전송하게 하는 방법을 연구하였다. 그림 2.에 이 방법의 블록도를 나타내었다.



<그림 2> 제안된 방법2의 블록도

EZW에서는 MSE를 가장 크게 줄일 수 있는 계수를 먼저 전송한다. 관심영역에 해당하는 계수들을 먼저 전송한다면 관심영역에 대한 고화질을 보장할 수 있다. 이 방법 또한, 방법 1에서와 같이 관심영역의 지정에 마스크를 이용한다. 마스크를 설정하고, 이를 웨이블릿 변환하게 된다. 또한 마스크는 임의의 영역을 설정이 가능해야 하고, 여러 영역의 설정이 가능해야 한다.

관심영역을 먼저 전송하기 위해서는 관심영역내의 계수들을 크게 해주면 된다. 이것을 스케일링 인자(scaling factor: S)라고 하면 S가 크면 클수록 관심영역의 계수들은 먼저 전송되게 된다. 하지만 EZW의 연속적인 근사 양자화로 인해 불필요한 단계를 거치게 된다. S가 충분히 크지 않다면 관심영역내의 계수들 일부는 관심영역과 분리되어 배경영역과 같이 부호화 된다.

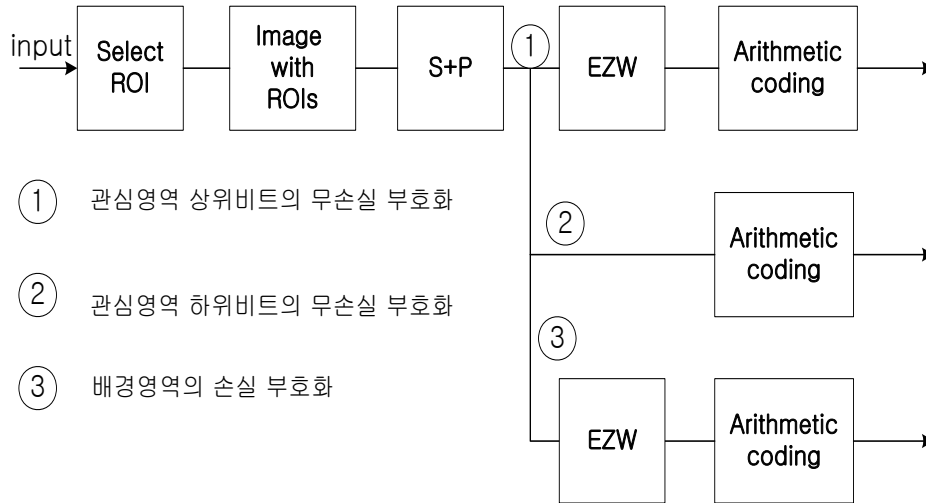
부밴드에서의 관심영역에 속하는 계수가 상위레벨에서 관심영역에 반드시 포함되어 지지는 않는다. 웨이블릿 필터의 특성상 상위레벨에서는 그 값의 분포가 퍼지게 되어 나타난다. 그래서 복원된 영상에서는 관심영역의 부분이 영역인자(Region factor)에 의해 제한을 받게 된다.

본 논문에서는 이러한 스케일링 인자와 영역인자를 이용하여 인간의 시각의 차이에 따른 사용용도를 달리 할 수 있게 한다.

3) 관심영역의 무손실 부호화

의료영상과 같은 특정의 응용분야에서는 관심영역에 대해 원래의 영상이 그대로 전송되어야 하는 경우가 있다. 본 논문에서 사용하고 있는 웨이블릿 필터는 실수필터이다. 이러한 필터를 그대로 사용할 경우, 이론적으로는 원영상의 완전 복원이 가능하나, 실제로 실수계수가 정수계수로 변환되는 과정에서 발생하는 절사에러로 인해 완전복원을 보장할 수 없다. 이런 경우 정수에서 정수로 매핑하는 웨이블릿 변환을 사용하여 완전복원을 구현할 수 있다. 정수 웨이블릿을 사용하는 전체영상의 무손실 부호화에 관한 연구는 활발하게 진행되어 왔다. 반면, 관심영역만의 무손실 부호화에 대한 연구가 상대적으로 부족했다고 볼 수 있다.

본 논문에서는 SPIHT에서 S+P트랜스폼을 이용한 무손실 부호화기를 관심영역 부호화에 적합한 형태로 변환하여 사용하고, 그 결과를 확인한다. 그 블록도를 그림 3에 나타내었다.



<그림 3> 제안한 방법 3의 블록도

S+P변환을 이용한 무손실 부호화에서는 실제로 두 개의 부호화기를 사용한다. S+P변환을 이용하여 얻어진 계수들 중 중요계수를 SPIHT에서 설명한 점진적 전송방법으로 전송하고, 나머지 여분의 정보를 산술 부호화기를 이용한 무손실 부호화기로 전송하는 방법을 사용한다. 즉, 중요정보에 대해서는 EZW부호화 방법을 이용하여 점진적으로 전송하고, 나머지 여분의 정보에 대해서는 무손실 부호화기를 이용하여 전송하게 된다. 이 방법은 계수를 비트 플레인으로 분할하여 상위 비트플레인은 따로, 하위 비트플레인은 또 별도의 산술부호화기를 이용하는 방법이다. 상위와 하위의 기준은 부호화기의 성능을 고려할 때 대략 3비트 정도이다.

관심영역에 대한 무손실 부호화기는 위의 절에서 사용한 부호화기의 웨이브렛 변환을 정수 웨이브렛 변환으로 대체함으로써 얻을 수 있다. 본 실험에서는 S+P변환을 이용하여 관심영역에 대한 무손실 부호화기를 구현한다.

S 변환 계수들은 정수이고 스케일링 성분들은 2의 멱(power)승이기 때문에 그것들은 부호화시 함축적으로 고려될 수 있다. 이런 장점들 때문에 S+P 변환을 갖는 동일한 스케일링 성분들을 사용한다.

여기서 유의하여야 할 것은 위에서 설명한 방법의 S+P변환으로 대체하여 관심영역의 무손실 부호화를 행할 경우, 실제 사용되는 부호화기는 중요정보를 전송하기 위한 EZW부호화기와 여분 정보를 전송하기 위한 별도의 부호화기가 모두 사용된다는 사실이다. 즉, 부호화가 스케일링된 관심계수의 중요정보를 전송하기 위한 EZW부호화기와 SPIHT 알고리즘은 영상을 정확히 복원하기 위해서 모든 비트플레인을 부호화하기 위해서 사용될 수 있다. 그러나 최하위 비트들을 부호화 할 때 효율은 감소하는데 그것은 주로 속도와 메모리 용량 때문이다. 이것은 대개 영상의 손실이 원 영상과 눈으로 구분될 수 없을 때 비트율 때문에 발생한다. 그런 이유

로 SPIHT는 3번째 하위 비트까지 부호화하기 위해서 사용되고 남은 비트플레인을 부호화하기 위해서 고정된 일정하게 분포된 산술 부호화기로 부호화 한다. 계수의 여분정보를 전송하기 위한 산술부호화기, 그리고 스케일링되지 않은 배경영역의 계수를 전송하기 위한 EZW부호화기의 순서로 부호화가 진행된다.

이 경우 관심영역의 계수의 중요정보와 배경영역의 계수의 정보를 전달하기 위한 EZW부호화기에서는 불연속성이 발생되며, 실제로 별도의 코더로 압축되어야 하며, 비트스트림도 서로 성질이 틀린 3개의 비트스트림이 발생하게 된다. 결국 1.1절의 부호화기에 S+P변환을 사용한 것 보다 별다른 잇점이 나타나지 않는다.

2. 실험 및 결과 고찰

프로그램은 Windows 환경에서 사용될 수 있게 설계되었다. 웨이브렛을 이용한 기존의 프로그램들은 대부분 Unix 환경에서 개발되었던 것들이어서 프로그램을 windows 환경으로 만드는 작업에 어려움이 있어서 몇몇 동작은 Windows에서 Dos를 불러서 작업하는 과정이 있다. 그러나 성능의 평가에 이상은 없었으나 EZW의 비트율을 측정했을 때 Lena 512x512 의 1 bpp에서의 PSNR은 문헌에서 알려진 값은 39.55 dB였으나 변환된 프로그램은 39.33 dB로 측정되었다. 이것은 성능의 평가에 영향을 미치지 않고, 또 사용된 영상의 차이일수도 있어 문제가 되지 않는다.

본 논문에서 실험에 사용된 영상은 Lena 512x512이다. 영상은 실험의 공정성을 위해 UCLA의 Image Communication Lab(ICL)에서 점진적인 영상 부호화 기법들의 평가를 위한 영상으로 사용된 영상들을 사용하였다. Lena 영상은 비교적 저주파 성분이 많은 단순한 영상이다. ICL에서는 문헌적으로 입증된 알고리즘들의 성능평가를 위한 방법으로 PSNR(peak signal to noise ratio)을 사용하였다. PSNR은 인간의 시각에 의해 인식되는 영상에 대한 평가는 정확하게 할 수는 없으나 인간의 시각특성을 반영한 평가를 위하여 Nill이 제안한 WMSE (Weighted mean square error)를 이용하는 것도 큰 의미가 없어서 점진적인 부호화 방법들의 성능을 정량적으로 평가하는데 사용된 PSNR을 이용하여 성능을 비교 분석하였으며 사용된 식은 다음과 같다.

$$PSNR = 10 \log_{10} \frac{255^2}{E[f(i,j) - \hat{f}(i,j)]^2}$$

실험에서 사용된 웨이브렛 변환은 Antonini가 제안한 9-7 탭 필터를 사용하였다. 이 필터는 비교적 적은 탭으로 완전복원 조건을 만족하고 좌우대칭의 구조를 갖고 있어 영상신호의 경계 부분에 대해서 대칭확장이 가능하다. 관심영역과 배경영역의 분리 부호화에서는 마스크를 이용

하여 영상내에서 부호화를 행하게 되는데 이때 이 필터의 영향으로 인한 관심영역의 경계부분이 상위 밴드에서의 위치가 모호하게 된다. 그래서 리전 인자(Region factor)를 도입해 그 경계를 정해준다.

웨이브렛 변환을 이용한 영상의 점진적 부호화 기법으로 대표적인 방법으로는 EZW가 있다. EZW는 제로트리틀을 이용한 연속 근사 양자화를 통하여 부호화하는 방법으로 본 논문에서 구현해 보고 그 개선 방안인 SPIHT와 비교하였다. SPIHT는 현재 점진적인 영상의 부호화 방법으로 가장 좋은 성능을 보여 본 논문에서의 비교대상으로 하였다. 또한, 본 논문에서 사용되는 알고리즘 역시 SPIHT를 수정하여 이용하고 3번째 방법인 관심영역의 무손실 부호화 방법은 S+P 변환을 개선하여 이용하였다. 그림 4는 실험에 사용된 원영상 Lena 512×512이다.



<그림 4> 실험에 사용된 원영상

1) 관심영역과 배경영역의 분리 부호화

이 방법은 가장 간단한 방법으로 영상을 두 장으로 분리 처리하는 방식이다. 입력된 영상을 마스크를 이용하여 관심영역을 지정하고 관심영역과 관심영역을 제외한 부분을 따로 분리한다. 각각에 대하여 웨이브렛 변환을 하여 발생된 계수를 EZW를 통해 심볼스트림을 발생시킨 다음 산술부호화를 하여 비트스트림을 발생시키고 각각 발생된 비트스트림을 하나로 만들어 전송하고 복호화단에서는 다시 두 개의 비트스트림으로 분리하여 각각 복원된 영상을 만들고 이를 다시 합쳐서 복원된 영상을 만드는 방법이다. 이 방법은 다시 두 가지로 나눌 수 있다. 마스크의 영역을 분리하지 않은 것처럼 여기고 계산하는 방법 1-1과 0으로 채우고 계산하는 방법 1-2이다. 방법 1에서는 관심영역의 비트율을 1.2 bpp로 놓고 실험을 하였다. 이것은 실험영상의 관심영역 PSNR이 39~42 dB정도로 인간시각에서는 거의 원영상에 가깝기 때문이다.

첫 번째 방법의 실험의 결과로 관심영역의 크기를 64×64로 했을 때 실험영상의 PSNR을 보

면 기존의 방법인 SPIHT 방법보다는 약간 떨어지지만 EZW보다는 높게 나타난다. 관심영역의 PSNR을 고려한다면 상당히 효율이 높은 방법이라 하겠다. 또한 관심영역을 제외한 배경영역 역시 EZW보다는 화질이 좋았다. 그림 5.에 Lena 영상의 관심영역의 크기를 64×64로 하고 1.2bpp를 할당하여 전체적으로 0.1bpp로 압축된 영상과 EZW를 이용하여 0.1bpp로 압축된 영상을 비교하여 보았다. 관심영역의 크기를 64×64 에서 128×128로 변경했을 때는 더 많은 비트가 관심영역에 소요되어 0.1bpp에서는 현저하게 화질이 저하되었고 크기가 64×64일 경우보다 약 4dB정도 떨어 졌으나 0.25bpp 이상의 비트율에서는 다시 EZW에 비해 좋게 나타났다. 이것은 관심영역의 크기가 낮은 비트율에서는 많은 영향을 미친다는 것을 알 수 있다. 방법 1-1을 이용한 실험의 결과를 표 1.에 나타내었다.

<표 1> 방법 1-1을 이용한 부호화기의 성능비교(ROI bit rate=1.2bpp)

방법 1-1						
Roi size = 64×64					conventional method	
Image	Bit rate(bpp)	ROI+NonROI	ROI	NonROI	EZW	SPIHT
Lena 512×512	0.1	29.50	39.05	29.44	28.93	30.22
	0.25	33.81	39.05	33.77	32.75	34.11
	0.5	37.06	39.05	37.04	36.01	37.21
	1	40.27	39.05	40.29	39.33	40.40
Roi size = 128×128					conventional method	
Image	Bit rate(bpp)	ROI+NonROI	ROI	NonROI	EZW	SPIHT
Lena 512×512	0.1	25.56	40.76	25.28	28.93	30.22
	0.25	32.84	40.76	32.60	32.75	34.11
	0.5	36.68	40.76	36.51	36.01	37.21
	1	40.04	40.76	40.00	39.33	40.40



(a) 방법 1-1, 0.1bpp

(b) EZW, 0.1bpp

<그림 5> 방법 1-1과 기존의 방법과의 비교(Lena 512x512, ROI=64x64 1.2bpp)

두 번째 방법에서는 마스크의 영역을 0으로 채우고 실험을 한 것이다. 표 2.에 그 결과를 나타내었다. 전반적인 실험결과는 방법 1-1과 비교해 별다른 차이가 없었지만 방법 1-1에 비해 PSNR측면에서 약간 떨어지는 결과를 보였다. 방법 1-1이 더 좋은 결과를 보이는 이유는 웨이브렛 필터가 관심영역의 경계부분의 화소를 필터링 할 때 배경영역의 화소도 참조로 하는 특성 때문이라고 추정된다.

<표 2> 방법 1-2를 이용한 부호화기의 성능비교(ROI bit rate=1.2bpp)

방법 1-2						
Roi size = 64×64					conventional method	
Image	Bit rate(bpp)	ROI+NonROI	ROI	NonROI	EZW	SPIHT
Lena 512×512	0.1	29.33	38.61	29.27	28.93	30.22
	0.25	33.72	38.61	33.67	32.75	34.11
	0.5	37.03	38.61	37.01	36.01	37.21
	1	40.29	38.61	40.33	39.33	40.40
Roi size = 128×128					conventional method	
Image	Bit rate(bpp)	ROI+NonROI	ROI	NonROI	EZW	SPIHT
Lena 512×512	0.1	25.37	40.99	25.10	28.93	30.22
	0.25	32.77	40.99	32.53	32.75	34.11
	0.5	36.74	40.99	36.56	36.01	37.21
	1	40.22	40.99	40.17	39.33	40.40

2) 관심영역과 배경영역의 통합 부호화

이 방법은 스케일링을 이용하는 방식으로 계수의 값이 큰 계수부터 먼저 전송되는 방식이다. 부호화 순서는 입력영상에 관심영역을 마스크를 이용하여 영상내에 영역을 설정하게 된다. 관심영역을 포함한 전체 영상을 DWT를 하여 계수를 발생시킨다. 발생된 계수에 대해서 마스크의 영역을 스케일링하여 값을 크게 하고 제로트리 부호화를 통해 관심영역의 계수들이 먼저 전송되게 하는 방법이다. 관심영역을 함께 부호화하는 것은 관심영역을 분리해서 분리하는 방법에 비해 복잡하다. 스케일링 값 S에 의해 발생하는 경우에 대해서 부호화 효율이 달라지기 때문이다. 본 논문에서는 이 S를 Q(Quality) factor라 하고 여기에 R(Region) factor를 도입했다.

실험에서는 Q factor와 R factor의 최적치를 찾기 위해 관심영역을 64×64, 128×128으로 분리하고 그 각각에 대해 Q factor를 $2^1 \sim 2^8$ 까지 변환하여 실험을 하였다. R factor는 Q factor의 각각에 대해 임의로 1, 30, 150으로 나누어 각각에 대해 0.1의 비트율에 대해 실험을 하였다.

실험의 결과로 Lena512 영상에 대해 전체 비트율이 0.1 bpp에서 관심영역의 크기를 64×64에서 128×128으로 변화 시켰을 경우에 대해 표 3.에 나타내었다. 관심영역의 크기 변화는 영역이 64×64를 경계로 전체적인 PSNR 및 관심영역의 PSNR이 증가하다가 그 이상 영역이 커지면 다시 감소하는 현상을 보였다. 사용자의 측면에서 본다면 영역의 크기는 128×128 정도가 좋다고 볼 때 64×64에 비해 약 2dB 정도가 떨어지게 나타났다. 실험적인 측면에서 영역의 크기가

64×64에서 가장 좋은 결과라고 할 때 Q factor가 16일 때 R factor의 변화에 대해 최적의 조건임을 보였고 R factor가 30일 때 최적의 조건임을 보인다. 그림 6.에 전체 비트율을 0.1bpp, 관심영역의 크기를 64×64, R=30으로 고정하고 (a) Q=16, (b) Q=128일 때의 Lena512 영상을 보였다. 관심영역을 코 부분을 설정했을 때 전체적인 PSNR과 관심영역의 PSNR은 (a)가 약 3dB정도 높게 나타났고, 시각적으로 확연히 차이가 있음을 보인다.

<표 3> 방법 2를 이용한 관심영역의 PSNR(Lena512, 0.1bpp)

방법 2						
Lena512, ROI size = 64×64, 0.1bpp						
R factor	Q factor = 16			Q factor = 128		
	F-PSNR	ROI-PSNR	B-PSNR	F-PSNR	ROI-PSNR	B-PSNR
1	28.48	41.58	28.42	23.99	48.39	23.93
30	28.95	39.26	28.89	25.22	36.85	25.16
150	29.41	37.14	29.35	26.50	35.47	26.44
Lena512, ROI size = 256×256, 0.1bpp						
R factor	Q factor = 16			Q factor = 128		
	F-PSNR	ROI-PSNR	B-PSNR	F-PSNR	ROI-PSNR	B-PSNR
1	25.32	30.15	24.44	23.43	30.17	22.41
30	24.37	30.51	23.40	20.20	28.94	19.10
150	23.88	28.13	23.06	17.76	19.52	17.31



(a) R=30, Q=16, 0.1bpp



(b) R=30, Q=128, 0.1bpp

<그림 6> R=30에서 Q factor의 변화(Lena512 image, ROI size=64×64, 0.1bpp)

3) 관심영역의 무손실 부호화

무손실 부호화를 위해 실험에서는 S+P변환을 이용하였다. S+P 변환은 정수를 이용하는 변환으로 완전복원을 이룰 수 있다.

부호화 순서는 입력영상에 대해 마스크를 이용하여 관심영역을 지정하고 무손실 부호화를

위해 DWT대신 S+P 변환을 이용한다. S+P변환에서 얻어진 계수들을 상위비트는 개선된 EZW 방법으로 하위비트는 고정된 산술 부호화를 이용하고 스케일링되지 않은 배경영역은 EZW를 이용하여 각각의 비트스트림을 만들고 다시 하나의 비트스트림으로 합하여 복호기로 전달하고 복호기에서는 다시 세 개의 비트스트림으로 나누어 각각에 대해 복원영상을 만들고 이 세 개를 융합한다.

실험방법은 각 실험영상에서 관심영역의 크기를 64×64와 128×128로 하고 전체 비트율을 0.1에서 2까지 변화하여 실험하고 그 결과를 표 4.에 나타내었다. 영역의 크기가 작을 때는 낮은 비트율에서도 어느 정도의 화질이 보장되지만 영역의 크기가 커졌을 때 낮은 비트율에서는 영상정보로서의 가치가 없었다. 이는 관심영역에 무손실을 위해 많은 비트를 할당했기 때문에 낮은 비트율에서는 관심영역만이 나타나는 결과를 보였다. 비트율이 높아졌을 때에는 오히려 관심영역의 크기변화에 대해 영상은 PSNR의 측면에서 점점 차이가 줄어들었다. 그 실험 영상을 그림 7.에 나타내었다. 관심영역의 무손실을 위해 할당되는 비트율은 약 5 비트 정도가 할당됐고 낮은 비트율에서의 완전복원은 관심영역의 크기에 크게 좌우되는 것을 볼 수 있었다.

<표 4> 관심영역의 무손실 부호화

방법 3			
ROI size = 64×64			
Image	Bit rate(bpp)	ROI+NonROI(dB)	Lossless region(bit)
Lena 512	0.1	23.91	5.46
	0.25	32.17	
	0.5	35.18	
	1	39.18	
	2	43.71	
ROI size = 128×128			
Image	Bit rate(bpp)	ROI+NonROI(dB)	Lossless region(bit)
Lena 512	0.1	5.97	4.74
	0.25	5.97	
	0.5	33.28	
	1	38.18	
	2	42.80	



(a) 방법 3, ROI=64x64, 1bpp (b) 방법 3, ROI=128x128, 1bpp
 <그림 7> 방법 3을 이용한 ROI의 크기변화에 따른 화질 비교(1 bpp)

III. 결론

관심영역의 부호화에 대한 아이디어는 두 가지로 나누어진다. 그 중한가지는 영상을 분리해서 부호화 하는 방법이고 또 하나는 스케일링 방식을 이용한다. 영상을 두 개로 분리해서 각각 독립적으로 부호화를 행하여 다시 융합하는 과정을 거치면 된다. 분리하는 방법으로는 마스크를 사용하였으며 마스크에 대한 정보는 부가 정보가 된다. 마스크는 임의의 모양과 두 개 이상의 영역을 지정할 수 있도록 설계되었다. 실험에서 증명되었듯이 이 방법들은 EZW에 비해 좋은 성능을 나타내었다. 다만, 기본 알고리즘으로 사용된 SPIHT의 한계를 넘지는 못하였으나 관심영역의 지정과 관심영역에 대한 비트율의 조정이 가능한 점은 사용용도를 달리 했을 때 매우 좋은 부호화 효율이라고 볼 수 있다. 관심영역의 크기는 부호화 효율에 큰 영향을 미친다. 특히 완전복원을 위해서는 관심영역의 크기에 따라서 모든 정보가 관심영역에 집중되어 나머지 정보는 전혀 식별이 불가능할 수도 있다. 관심영역의 지정에 따른 전체적인 부호화 효율을 높이기 위해서는 관심영역과 배경영역에 상호 절충이 필요하다.

영상을 통합해서 부호화 하는 방법은 분리하는 방법에 비해 좀더 복잡한 과정을 거친다. 영상의 경계부분의 처리에 대한 문제로 스케일링 인자(Q factor)와 리전 인자(R factor)를 도입해 관심영역에 속하는 부분을 조절하도록 하였다. 스케일링을 크게 할수록 관심영역의 화질은 좋아졌으나 전체적인 화질은 떨어지는 결과를 보였다. 리전 인자도 크게 할수록 같은 결과를 보였다. 실험을 통해서 본 결과로는 영역을 분리해서 부호화 하는 방법이 더 좋은 결과를 나타내었다. 그러나 관심영역의 확실도를 위해서는 통합해서 부호화 하는 방법이 이용될 수도 있다.

무손실 부호화 방법의 수행방법은 방법 1의 분리해서 부호화 하는 방법과 유사하다. 다만,

관심영역의 완전복원이 가능 하느냐의 차이이다. 무손실 복원을 하기 때문에 전체적인 화질은 떨어지나 관심영역은 무손실 복원이라는 장점을 가지고 있다. 위와 같은 방법들은 각기 사용자에게 따라 용도를 달리해 사용될 수 있다.

제안된 방법들은 원격 의료 진료, 영상 데이터베이스의 구성 및 검색 그리고 인터넷의 영상 통신을 위한 하나의 도구로 이용될 수 있다.

향후 연구되어야할 문제로는 관심영역의 지정이 하나 이상일 때 웨이브렛 변환으로 인한 상위레벨에서의 겹침 문제를 해결하고, 인간 시각적인 측면에서의 평가방법에 대한 연구가 필요하고 객체 지향적인 프로그램을 개발하여 초 저속 전송매체에서의 이용이 가능한 방법을 구현하는 것이다.

참고문헌

- [1] K. R. Rao and P. YIP, *Discrete Cosine Transform - algorithms, advantages, applications*, Academic Press, San Diego, 1990.
- [2] G. Karlsson and M. Vetterli, "Theory of two-dimensional multirate filter banks," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol.38, No.6, 1990, pp.925-937.
- [3] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect reconstruction technique," *IEEE ASSP Mag.*, Vol.4, 1987, pp.4-20.
- [4] Gerald Kaiser, *A Friendly Guide to Wavelets*, Birkhäuser, 1995.
- [5] C. Sidney Burrus, Ramesh A. Gopinath, *Introduction to Wavelets and Wavelet Transforms*, Prentice-Hall, New Jersey, 1998.
- [6] J. M. Sapiro, "Embedded image coding zerotrees of wavelet coefficients," *IEEE Trans. Signal Proc.*, Vol.41, 1993, pp.3445-3462.
- [7] J. M. Shapiro, "An embedded hierarchical image coder using zerotrees of wavelet coefficients," Proc. Data Compression Conf., Snowbird, Utah, IEEE Computer Soc. Press, 1993.
- [8] O. Egger and M. Kunt, "Embedded Zerotree based Lossless Image Coding", Proc. of ICIP. 1995.
- [9] V. N. Ramaswamy, K. R. Namuduri and N. Ranganathan, "Performance Analysis of Wavelets in Embedded Zerotree based Lossless Image Coding Schemes", *Proc. of Intl. conf. on Image Processing*, October 26-30, 1997, Santa barbara, California.
- [10] J. M. Sapiro, "Embedded image coding zerotrees of wavelet coefficients," *IEEE Trans. Signal Proc.*, Vol.41, 1993, pp.3445-3462.