

# 웨이블릿변환과 신경회로에 의한 칼라 동영상의 실시간 전송

## Real-time Image Transmission on the Internet Using Wavelet Transform and Neural Network

Jeong Ha Kim \*, Hyoung Bae Kim \*\* and Boo Hee Nam \*\*\*

BK21 Dept of Electrical and Computer Engineering, Kangwon National University, Chunchon, Korea

Tel : +82-33-243-4347 Fax : +82-33-242-2059 E-mail: jkim@is.kangwon.ac.kr, boonam@kangwon.ac.kr

### Abstract

본 논문에서는 TCP/IP 기반에서 웨이블릿 변환과 신경회로(Neural Network)를 이용하여 실시간으로 동영상 전송을 위한 알고리즘에 대해 연구하였다. 우선, 서버에서 카메라를 이용하여 영상을 획득한다. 획득한 영상에 대해 2 단계 웨이블릿 변환을 한 후, 인터넷을 통하여 수신 측으로 전송한다. 그 다음, 수신 측에서는 동영상을 보이기 위해, 제한된 시간 안에 수신한 이미지 프레임(image frame)만을 가지고 역웨이블릿변환을 하여 보여주게 된다. 만약 네트워크에 혼잡이 발생하게 되면 소량의 이미지 프레임만을 전송하도록 하였다. 수신 측에서는 네트워크의 혼잡 정도를 송신 측에 특별한 목적으로 제안된 프레임들을 전송함으로써 알려주도록 하였으며, 송신 측에서는 이러한 정보를 수신한 후 역전파 신경회로(back-propagation neural network)를 이용하여 송신할 데이터의 양(data rate)을 조절하여줌으로써 네트워크 속도에 대처하도록 하였다. 이러한 방법을 통해서 우리는 네트워크의 혼잡에 알맞게 조정된 동영상의 스트림(stream)을 보낼 수 있었다.

### I. 서론

TCP/IP(Transmission control protocol/internetworking protocol)는 인터넷을 통한 전송을 정의한 규약이며, 그 효용성은 이미 증명되어왔다. TCP는 모든 커넥션을 이룬 후 데이터를 주고 받는 방식(connection oriented)으로 커넥션 설정과 오류제어, 흐름제어 등을 하므로 데이터 손실이 발생하지 않는 방식이다. 그러나 칼라 동영상 이미지 스트림을 전송한다고 할 때 그 데이터의 양은 네트워크에 부담이 될 수 밖에 없게 된다.

웨이블릿 변환(Wavelet transform)은 이미지 처리에 유용한 방법으로, 1 단계 웨이블릿 변환을 하게 되면 이미지는 근사계수(approximation coefficients), 수평상세계수(horizontal detail coefficients), 수직상세계수(vertical detail coefficients), 대각상세계수(diagonal detail coefficients)의 같은 크기를 가

지는 4 개의 부분으로 분해 된다. 송신 측에서 네트워크를 통해 데이터를 전송할 시에는 근사계수에서 대각상세계수의 순서로 패킷을 구성하여 전송하게 되며, 이렇게 함으로써 제한된 시간 안에 수신한 데이터만을 이용하여 좀더 나은 영상을 복원할 수 있다. 만약, 네트워크의 속도가 매우 낮은 경우가 발생하면, 단지 근사계수부분만을 전송하게 된다.

신경회로는 시스템의 입출력에서 많은 양의 입력/출력 데이터가 이용 가능 하지만 어떻게 입력과 출력이 관련되어 있는지 알 수 없는 경우에 유용하다. 신경회로를 우리가 실행하고자 하는 기능을 갖도록 하기 위해서는 충분한 학습을 시켜야 한다. 본 논문에서는 송신 측에서 인터넷을 통해 영상을 송신할 때 네트워크 속도에 따라 이미지 데이터의 양을 결정하고자 할 때 신경회로를 사용하였다.

### II. 전체 시스템의 구성

그림 1은 본 논문에서 구성한 전체 다이어그램을 나타낸 것이다.

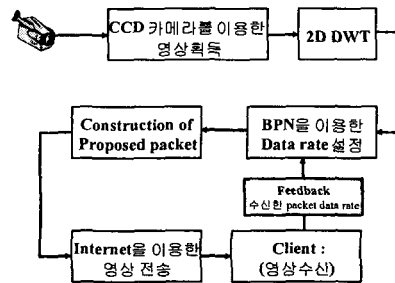


그림 1. 전체 다이어그램

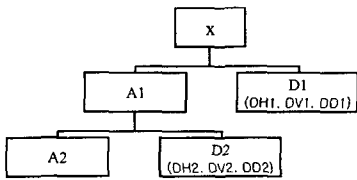
### III. 동영상의 웨이블릿 변환

본 절에서는 CCD 카메라를 이용하여 입력 받은 영상에 대해 웨이블릿 변환을 수행하며, 실시간으로 영상을 전송하기 위한 알고리즘을 제안한다.

### 3.1 웨이블릿 변환

본 논문에서는 원 이미지로부터 근사계수와 상세계수 성분을 획득하기 위해 DWT(discrete wavelet transform)를 사용하였다. DWT는 근사계수와 상세계수 성분으로 신호를 분해함으로써 다른 해상도를 가지는 다른 주파수 대역으로 신호를 분해한다.

그림 2는 영상 데이터 x를 웨이블릿 변환하여 근사계수와 상세계수 성분으로 분해하는 것을 도식한 그림이다. 그림 2에서 A1은 원 이미지의 1/4 크기의 데이터 크기를 가지게 된다. 본 논문에서는, 원 영상에 대해 두 번의 웨이블릿 변환을 하였다. 그림 4에서는 그림 3의 이미지를 1단계 2D DWT 한 영상을 보여준다.



$$X = A1 + D1$$

$$= A2 + D2 + D1,$$

where  $D1 = DH1 + DV1 + DD1$   
 $D2 = DH2 + DV2 + DD2$

그림 2. 이미지의 2D 웨이블릿 변환



그림 3. 원 영상

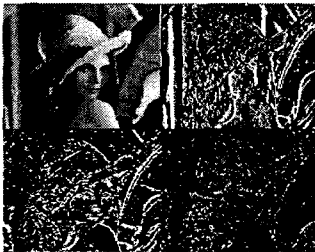


그림 4. 1 단계 웨이블릿 변환 후의 영상

### 3.2 송신 이미지 패킷의 순서

그림 5는 2 단계 2D 웨이블릿 변환을 한 영상을 보여준다.

다. 원 이미지를 2 단계 웨이블릿 변환을 하게 되면 이미지는 각각 원 영상의 1/16 크기의 크기를 가지는 근사계수와 3개의 상세계수 성분 그리고 1/4 크기를 가지는 3개의 상세계수 성분의 총 7개의 부분으로 분해가 된다. 이렇게 분해된 이미지 데이터는 ①: 1/16 크기의 근사계수 ②: 1/16 크기의 수평상세계수 ③: 1/16 크기의 수직상세계수 ④: 1/16 크기의 대각상세계수 그리고 ⑤: 1/4 크기의 수평상세계수 ⑥: 1/4 크기의 수직상세계수 ⑦: 1/4 크기의 대각상세계수, 의 순서로 송신 프레임을 재구성하여 네트워크를 통해 수신측으로 송신하게 된다. 이러한 순서로 이미지 데이터를 송신하는 이유는 제한된 시간 안에 수신된 이미지 데이터만을 가지고도 좀더 나은 영상을 복원할 수 있기 때문이다.

그림 6에서 2 단계 DWT를 실행하여 획득한 데이터를 화살표 방향인 1 번에서 7 번의 순인, 즉 근사계수에서 상세계수 성분의 순으로 패킷의 스트림을 재구성하는 것을 보이며, 그림 8, 9, 10, 11, 12는 각각 1~5, 1~4, 1~3, 1~2, 1의 조각으로 재구성한 영상들을 보여주고 있다..

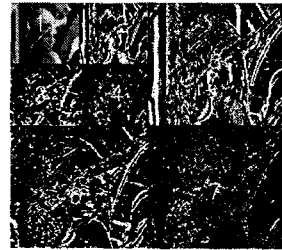


그림 5. 2 단계 웨이블릿 변환 후의 영상

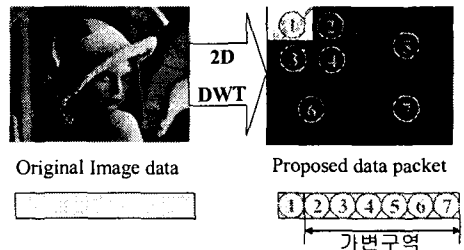


그림 6. 제안된 패킷 스트림 구조



그림 7. 7 번 손실



그림 8. 6,7 번 손실



그림 9. 5,6,7 손실

그림 10. 4,5,6,7 손실



그림 11. 3,4,5,6,7 손실

그림 12. 2,3,4,5,6,7 손실

### 3.3 전송할 데이터량을 결정하기 위한 신경회로

그림 13은 역전파 신경회로를 나타내는 그림이다. 본 논문에서는 역전파 신경회로를 이용하여 네트워크 속도에 따른 송신 데이터 양을 추론하도록 하였다.

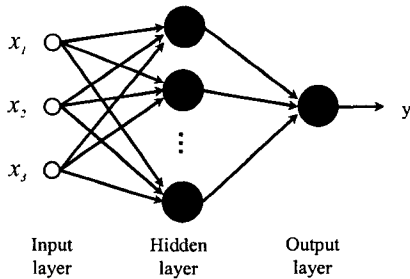


그림 13. 3개의 계층을 갖는 역전파 신경회로

위와 같은 신경회로망의 결선강도를 계산하기 위해선 우선 출력을 구하게 된다. 신경회로망의 출력을 다음과 같이 표현될 수 있다.

$$y = f(\psi); \quad \psi = \sum_{j=1}^M w_j h_j \quad (1)$$

여기서  $h_j$ 는 은닉층의  $j$ 번째 신경소자의 출력으로서 은닉층이 1개인 경우에는 다음과 같이 계산된다.

$$h_j = g(\phi_j); \quad \phi_j = \sum_{k=1}^N c_{jk} x_k; \quad j = 1, 2, \dots, M \quad (2)$$

위의 식을 이용 입력패턴에 대한 출력을 구하게 된다. 다음으로 출력과 목표치의 차이를 구한 후, 마지막으로 오차 값을 역방향으로 전파시키면서 출력층의 연결강도와

은닉층의 연결강도를 변경하게 된다. 식 (3), (4)에서 출력층과 은닉층의 연결강도의 조정을 위한 계산식을 보여준다.

$$w_j(n+1) = w_j(n) + \eta \delta_j(n) h_j(n) \quad (3)$$

$\delta_j$ : 출력층 신경소자의 일반화된 오차

$$c_{jk}(n+1) = c_{jk}(n) + \eta \delta_{nj}(n) x_k(n) \quad (4)$$

$\delta_{nj}$ : 은닉층  $j$ 번째 신경소자의 일반화된 오차

송신할 데이터량의 선택은 네트워크 속도에 영향을 받게 된다. 초당 15프레임을 전송한다고 할 때 50ms 안에 카메라로부터 한 개의 이미지 프레임 입력 받아 2단계 웨이블릿 변환을 하여 인터넷을 통해 전송하여야 하며, 수신측에서는 역웨이블릿 변환을 한 후 화면에 뿌려주어야 한다.

표 1은 각각의 네트워크 속도에서 어떠한 지연 없이 전송할 수 있는 데이터량을 보여준다. 만약 네트워크 속도가 2M bps라고 한다면, 1, 2, 3번의 데이터만으로 패킷을 구성하여 지연 없이 50ms 안에 패킷을 수신측으로 전송할 수 있게 된다.

표 1. 네트워크 속도에 상응한 데이터량

인터넷 속도 Mbps	50ms 당 전송 가능한 데이터량 (bit)	패킷을 구성하는 데이터의 번호	구성된 패킷의 데이터량 (bit)
1	50000	1	28800
1.5	75000	1,2	57600
2	100000	1,2,3	86400
3	150000	1,2,3,4	115200
5	250000	1,2,3,4,5	230400
7	350000	1,2,3,4,5,6	345600
10	500000	1,2,3,4,5,6,7	460800

신경회로를 학습 시키려면 학습 데이터가 필요하게 된다. 본 논문에서는  $x_1, x_2, x_3$ 의 3개의 입력과 1개의 출력인  $y$ 를 사용하여 네트워크를 학습 시켰다, 여기서  $x_1$ 은 송신 패킷의 데이터량이 되고  $x_2$ 는 제안된 시간 내에 수신한 패킷의 데이터량이 되며, 지연의 존재여부가  $x_3$ 이 된다. 만약 지연이 발생하면  $x_3$ 은 0이 되고, 그렇지 않으면 1의 값을 가지게 된다. 출력  $y$ 는 다음 번 송신할 패킷의 데이터량이 되게 하였다.

표 II에서 본 논문에서 사용한 학습 데이터가 주어졌다.

표 2. 학습 데이터

입력	1 1 1 2 2 2 2 3 3 3 3 3 4 4 4 4 4 5 5 5 5 5 6 6 6 6 6 6 7 7 7 7 7 7 7
	1 1 0 2 2 1 1 3 3 3 2 1 4 4 3 2 1 5 5 4 3 2 6 6 5 4 3 2 7 7 6 5 4 3 2
	1 0 0 1 0 1 0 1 0 1 1 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
출력	2 1 1 3 2 1 1 4 3 2 1 5 4 3 2 1 6 5 4 3 2 7 6 5 4 3 2 7 7 6 5 4 3 2

표 II 에서 주어진 데이터를 이용하여 신경회로의 에러가  $10^{-5}$  이하가 되도록 학습시켰으며, 약 10 초 정도의 학습 시간이 소요되었다.

IV. 실험 결과

실험에서 우리는 신경회로에 의해 결정된 데이터 량을 가지도록 이미지 패킷을 구성하여 클라이언트로 전송하였다. 그림 15 의 위쪽 그림에서 네트워크 스피드에 따라 네트워크의 속도에 대응하기 위해 송신 측의 데이터 량도 따라서 변동됨을 볼 수 있다. 그림 15 의 아래쪽 그림은 네트워크 속도가 변동함에 따라 수신 측에서 보낸 데이터를 그에 따라 수신함을 보여준다. 또한, 그림 15 에서는 네트워크 속도가 각각 증가하거나 감소하거나, 갑작스럽게 변하는 경우에 대해서도 보이고 있다. 이렇게 신경회로를 이용하면 네트워크 속도에 딱 맞는 데이터 량을 선택하게 될 확률은 99%이상이 되며, 만약 단지 if 문에 의한 단순 조건 문을 이용하게 된다면, 확률은 98%이하로 떨어짐을 확인할 수 있었다. 이와 같이 만일 본 논문에서 제한한 방법을 사용하게 된다면, 네트워크 속도에 맞추어 데이터 량이 변하여, 실시간으로 끊김 없이 영상이미지를 전송할 수 있었다. 이와 같은 실험을 통해 제안한 방법이 타당함을 보였다.

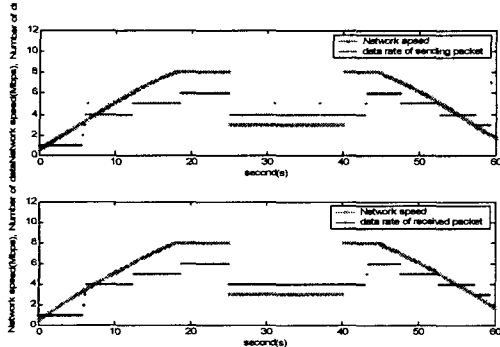


그림 15. 실험결과

V. 결론

본 논문에서 우리는 웨이블릿 변환과 신경회로를 이용하여 인터넷에서의 실시간으로 동영상상을 전송하기 위한

방법을 제시하였으며, 만족할만한 결과를 획득하기 위해 알고리즘을 실행하였다. 카메라로부터 획득한 영상 프레임은 2 단계 웨이블릿 변환을 거쳐 서버 쪽에서 클라이언트로 네트워크를 통해 전송하였다. 그 다음, 클라이언트에서는 단지 제한된 시간 안에 받은 이미지 패킷만으로 역 웨이블릿 변환을 하여 동영상을 보여주었다. 네트워크 상태에 만약 혼잡이 생긴다면 단지 각각의 이미지 단편만을 송신하도록 하였다. 반대로 만약 네트워크가 속도가 아주 좋다면 각 이미지의 모든 데이터를 클라이언트로 전송할 수 있게 될 것이다. 수신 측에서는 특별히 제한된 목적의 프레임을 송신 측에 보내줌으로써 네트워크의 혼잡 정도를 송신 측에 알리게 되며, 송신 측에서는 역 전파 신경회로를 통해 결정된 송신할 이미지의 데이터 량을 조절함으로써 수신 측의 경고에 응답할 수 있었다. 이러한 방법을 통해 우리는 네트워크 조건에 적합한 동영상의 스트림을 구성해서 보낼 수 있었다.

ACKNOWLEDGMENT

This paper supported by BK21(Brain Korea 21) Project of Kangwon National University.

참고문헌

- [1] A. Ben-Artzi, A. Chandna, and U. Warrier, "Network management of TCP/IP networks: present and future," IEEE Network, Volume 4, Issue 4, pp. 35 -43, July 1990,
- [2] P. P. Dang, and P. M. Chau, "Robust image transmission over CDMA channels," IEEE Transactions on Consumer Electronics, Volume 46, Issue 3, pp.664-672, Aug 2000.
- [3] P. M. Corcoran, "Mapping home-network appliances to TCP/IP sockets using a three-tiered home gateway architecture," IEEE Transactions on Consumer Electronics, Volume 44, Issue 3, pp. 729 -736, August 1998.
- [4] R. A. Devore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," IEEE Tr. On information theory, pp. 719-746, March 1992.
- [5] Juan Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," IEEE Transactions on Image Processing, Volume 10, Issue 11, pp.1647-1658, Nov 2001.
- [6] A. N. Sarlashkar, M. Bodruzzaman, and M. J. Malkani, "Feature extraction using wavelet transform for neural network based image classification," Proceedings of the Thirtieth Southeastern Symposium on System Theory, pp. 412 -416, 8-10 March 1998.
- [7] R. Hecht-Nielsen, "Theory of the backpropagation neural network," IJCNN., International Joint Conference on Neural Networks, Vol.1, pp. 593-605, 18-22 Jun 1989.
- [8] W. Chang, B. Bosworth, G. C. Carter, "On using backpropagation neural networks to separate single echoes from multiple echoes," ICASSP-93 IEEE International Conference on Acoustics, Speech, and Signal Processing, Volume 1, pp.265-268, 27-30 Apr 1993.