

---

# GRNN을 이용한 동영상 움직임 예측 및 대역분할 부호화에 관한 연구

한영오\*

A study on motion prediction and subband coding of moving pictures using GRNN

Young-Oh Han\*

요 약

본 연구에서는 신경회로망의 일종인 GRNN을 이용하여 동영상 대역분할 부호화에 적용하고자 하는 새로운 비선형 움직임 예측기를 제안하였다. 제안된 비선형 예측기의 성능은 가장 일반적으로 많이 사용되는 블록매칭 알고리즘과 비교하였다. 결과적으로 제안된 비선형 움직임 예측기는 블록매칭 알고리즘보다 2-3dB 성능이 우수함을 알 수 있었다. 특히, 제안된 예측기는 클러스터링 과정과 잡음 신호를 둔화시키는 기능이 있어서 원영상의 에지를 잘 보존하는 장점이 있음을 알 수 있다. 이러한 결과는 인간의 시각적 특성에 중요하며 동영상의 대역분할 부호화에서도 우수한 특성을 나타낸다.

Abstract

In this paper, a new nonlinear predictor using general regression neural network(GRNN) is proposed for the subband coding of moving pictures. The performance of a proposed nonlinear predictor is compared with BMA(Block Match Algorithm), the most conventional motion estimation technique. As a result, the nonlinear predictor using GRNN can predict well more 2-3dB than BMA. Specially, because of having a clustering process and smoothing noise signals, this predictor well preserves edges in frames after predicting the subband signal. This result is important with respect of human visual system and is excellent performance for the subband coding of moving pictures.

키워드

GRNN, BMA, subband coding, nonlinear predictor

## 1. 서론

연속된 프레임으로 구성된 동영상을 부호화하기 위해 기존의 동영상 부호화기법인 MPEG2와 같은 하이브리드 DCT 부호화방법을 그대로 적용하기에는 문제점이 있다. 첫째, 예측오차영상을 DCT 부호화기

로 거칠게 양자화하거나 낮은 전송률로 전송하고자 할 때 급격히 변하는 프레임에서 구획화 현상(blocking artifact)이 나타날 수 있으며 둘째, 고주파 대역의 정보손실로 에지저하(edge degradation)가 발생하여 혈관과 같이 중요한 진단정보의 심각한 손실을 초래할 수 있다. 이러한 문제점들을 어느 정도

---

\* 남서울대학교 전자공학과 부교수(youngoh@nsu.ac.kr)

접수일자 : 2010. 05. 03

심사(수정)일자 : 2010. 06. 01

게재확정일자 : 2010. 06. 14

해결할 수 있는 부호화 방법으로 대역분할 부호화 또는 웨이브렛 변환부호화 방식들이 제안되고 있다. 그러나 이러한 부호화 방식에서 가장 중요한 것은 움직임 예측 기법이라 할 수 있다. 그 이유는 움직임 예측을 통한 시간적 중복성(redundancy) 제거는 동영상 데이터 효율적 압축을 위한 중요한 단계이기 때문이다. 다양한 움직임 예측 알고리즘들이 제안되어 왔는데 대표적인 것으로 블록정합 알고리즘(BMA : Block Matching Algorithm)이다[1].

본 연구에서는 동영상 신호의 변화는 사람의 움직임, 카메라의 움직임 및 기타 외부 변화 등 여러 가지 요인에 의한 비정상성(nonstationary) 신호이므로 비선형 예측기의 구현이 용이한 신경회로망을 이용한 움직임 예측 알고리즘을 제안하였다. 사용한 신경회로망 모델로는 GRNN(general regression neural network)[2]이며 이 모델에 대한 비교대상으로 널리 알려진 역전파 학습 알고리즘(backpropagation training algorithm)[3]을 갖는 다층 퍼셉트론을 사용하였다.

본 연구는 동영상의 각각의 프레임을 2차원 주파수 대역으로 분할한 후, 제안된 비선형 움직임 예측 알고리즘을 적용하여 움직임 예측을 수행한 다음, 이러한 결과가 궁극적으로 동영상 부호화에도 효과가 있음을 실험을 통하여 보여주하고자 한다.

## II. GRNN을 이용한 움직임 추정 알고리즘

### 2.1. GRNN 움직임 추정 기법

Specht는 연속변수의 추정에 있어서 비선형 회귀 이론에 기초한 그림 1과 같은 GRNN(General Regression Neural Network)을 이용한 움직임 예측기를 제안하였다.

본 연구에서는 1차원 신호에 대한 Specht이 제안한 GRNN을 이용한 2차원 신호인 동영상에 대한 움직임 예측을 위한 방법을 제안하였다. 즉, 연속된 입력 프레임을 정방형의 겹치지 않는 블록으로 나눈다.

$i$  번째 프레임내의 블록은  $p \times p$  화소크기,  $i+1$  번째 프레임내의 블록은  $q \times q$  화소크기를 갖는다.

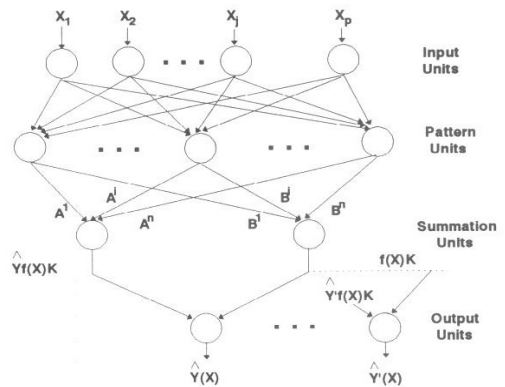


그림 1. GRNN 블록 다이어그램  
Fig. 1 GRNN block diagram

이 때,  $p \geq q$ 로써 그 이유는  $i$  번째 프레임의 블록  $X_i(m,n)$ 이 물체의 움직임 등으로 인하여 급격히 변화하는 영역에 포함된다면  $X_i(m,n)$ 에서  $X_{i+1}(m,n)$ 으로의 예측함수를 근사화하기란 어려운 문제이다. 따라서 변화가 심한 연속 프레임에 대해서  $X_i(m,n)$ 의 크기를 크게 할 필요가 있다.

블록  $X_i(m,n)$ 에서 블록  $X_{i+1}(m,n)$ 을 예측하는 GRNN을 구성하려면 스칼라 변수  $Y^i$ 는 벡터  $Y^i$ 가 되어야 한다. 한편, 연속된 2개의 프레임같이 관측  $(X, Y)$ 의 갯수가 매우 많은 경우 일정한 범위의 관측  $(X, Y)$ 들을 클러스터의 중심부로 표현하도록 하는 클러스터링 기법을 사용할 필요가 있다. 이러한 클러스터링 기법들은 클래스의 수를 알고 클러스터링 하는 방법과 클래스의 수를 모르고 클러스터링 하는 방법으로 크게 분류할 수 있다. 본 연구에서는 연산 속도 및 성능을 고려하여 클래스의 수를 모르고 클러스터링 하는 방법의 하나인 경험적 방법을 단순화하여 사용하였다. 그 이유는 동영상의 경우 프레임간의 변화가 일정하지 않으므로 고정된 클래스의 수를 사용할 수 없으며 또한 부호화시 전송 비트수를 버퍼의 상태에 따라 가변적으로 조절할 수 있어야 하기 때문이다. 관측  $(X, Y)$ 를 클러스터링하면 식 (1),(2)로 표현된다.

$$\hat{Y}(X) = \frac{\sum_{i=1}^M A^i \exp\left(-\frac{D^2}{2\sigma^2} \frac{i}{i}\right)}{\sum_{i=1}^M B^i \exp\left(-\frac{D^2}{2\sigma^2} \frac{i}{i}\right)} \quad (1)$$

여기서,

$$\begin{aligned} A^i(k) &= A^i(k-1) + Y^j \\ B^i(k) &= B^i(k-1) + 1 \end{aligned} \quad (2)$$

이며,  $M < N$ 이 되고  $A^i(k)$ 와  $B^i(k)$ 는 각각  $k$ 번째 관측 후  $Y$ 값의 합과 클러스터  $i$ 에 포함되는 샘플들의 수가된다.

GRNN의 특징은 반복적인 학습을 이용한 역전파 학습 알고리즘과 달리 1회의 처리과정을 통하여 클러스터의 중심 및  $A^i, B^i$ 로 계산한 다음 프레임 예측을 위한 회로망을 구성할 수 있기 때문에 동영상 프레임 예측과 같은 빠른 속도가 요구되는 연산 과정에서 적응적으로 사용할 수 있다.

## 2.2. 제안된 움직임 추정 알고리즘

본 연구에서 제안된 예측 알고리즘은 다음과 같다.

1단계, 임의의 연속 프레임에서  $i$ 번째 프레임을  $p \times p$  화소,  $i+1$ 번째 프레임을  $q \times q$  화소 크기의 블록으로 분할하여 관측치  $(X, Y)$ 의 집합을 구한다.

2단계, 클러스터링 반경  $r$ 을 설정한 다음 식 (3)을 이용하여 입력벡터  $X$ 를 클러스터링한다.

$$X^i(k) = \eta |X - X^i(k-1)| + X^i(k-1) \quad (3)$$

$(X, Y)$ 의 집합중에서 임의의 샘플을 선정하여 그때의  $X$ 를 클러스터링의 중심  $X^i$ 으로 설정한 다음 나머지 모든  $X$ 에 대하여  $|X - X^i|$ 가 반경  $r$ 보다 작거나 같은 모든  $X$ 에 대해 반복적으로 식 (3)을 적용함으로써 클러스터의 중심을  $X^i$  새로운  $X$ 가 들어올 때까지 갱신시켜 나간다. 만약  $|X - X^i|$ 가  $r$ 보다 크면 그 때의  $X$ 는 새로운 클러스터의 중심으로 설

정하여 이러한 과정을 반복 수행한다. 이 때 클러스터  $X^i$ 에 속하는  $X$ 에 대응하는 출력벡터  $Y$ 는 식 (2)를 이용하여 더해짐으로써  $A^i(k)$ 가 계산되고 동시에 클러스터  $i$ 에 속하는  $Y$ 의 개수  $B^i(k)$ 를 구할 수 있다.

3단계, 입력벡터  $X$ 를 제시한 다음, 2단계에서 계산된 클러스터 중심 즉, 그림 2에서 패턴유닛과의 거리  $d$ 를 계산하여 비선형 여기함수를 통과시켜 출력을 발생시킨다.

4단계, 비선형 여기함수를 통과한 각각의 출력에 기대출력  $A$ 를 곱한 다음 모두 더함으로써 패턴유닛의 출력 가중합  $\hat{Y}f(k)$ 는 식 (4)를 이용하여 구한다.

$$\hat{Y}f(k) = \sum_{i=0}^M A^i \exp\left(-\frac{D^2}{2\sigma^2}\right) \quad (4)$$

또한 패턴유닛의 출력함,  $f(k)$ 는 식 (5)를 이용하여 계산한다.

$$f(k) = \sum_{i=0}^M B^i \exp\left(-\frac{D^2}{2\sigma^2}\right) \quad (5)$$

5단계, 식 (4)를 이용하여 입력벡터  $X$ 에 대한 예측벡터  $\hat{Y}(X)$ 를 구한다.

## III. 2차원 대역분할

한편, 1991년 Gharavi[4]에 의해 동영상 대역분할 부호화 방식이 제안된 이후로 최근까지 많은 연구가 진행되고 있다. Gharavi는 동영상의 매프레임을 QMF를 이용하여 7개의 주파수 대역으로 분할한 후, 대역별로 운동추정을 하였다. 이 때 사용된 알고리즘은 블록정합 알고리즘이다. 그 후, 발표된 동영상 대역분할 부호화 또는 동영상 웨이브렛 부호화 방식에서 거의 대부분 블록정합 알고리즘에 기초한 대역별 예측을 수행하였다. 그러나 이러한 블록정합 알고리즘에 근거한 대역별 예측기법은 기저대역인  $B_0$  대역을 제외한 나머지 고주파 대역의 예측에는 성능이 현저히 떨어진다. 대역분할된 영상에서 고주파

대역은 인간의 시각 시스템의 특성상 중요한 시각 정보인 에지성분을 내포하고 있다. 따라서 고주파 대역의 에지성분에 대한 부호화는 정지영상은 물론 동영상 대역분할 부호화 및 웨이브렛 변환 부호화등에서 중요한 과제로 연구되어 왔다[5][6].

본 연구에서는 동영상의 2차원 대역분할을 위해 5/3 short kernel filter[7]를 사용하였으며 그림 2는 주파수 평면에 옥타브로 2차원 대역 분할된 Miss America 영상을 나타낸다.

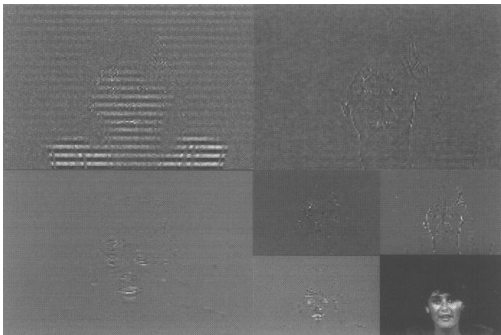


그림 2. Miss America 2차원 대역분할 영상  
Fig. 2 2D subband image of Miss America

#### IV. 실험 및 결과

실험에 사용된 영상은 SIF-625(크기:352x288) 규격을 갖는 Miss America 동영상을 사용하였다. 각 화소당 8bit의 계조도로 구성되었으며 휘도성분인 Y 성분만으로 실험하였다. 또한 예측 알고리즘의 성능을 평가하기 위하여 원영상과 예측된 영상과의 첨두치 신호대 잡음비(PSNR) 및 예측오차의 분산을 사용하였다.

제안된 GRNN을 이용한 비선형 예측기법의 성능을 표준화된 동영상 예측기법인 블록정합 알고리즘(BMA)과 비교, 분석하였으며 또한 대표적인 신경회로망인 역전파 신경회로망(BPN)을 이용하여 동영상 예측한 후, 그 결과와도 비교, 분석하였다.

표 1은 2차원 대역분할에서 예측오차의 분산을 나타낸다. 기저대역  $B_1$  대한 예측오차 분산은 BMA가 가장 적게 나타났으나  $B_1$  대역을 제외한 나머지 고주파대역에서는 GRNN의 예측오차 분산이 가장 적

게 나타났다. 이러한 결과는 동영상 대역분할 부호화 및 웨이브렛 변환 부호화 등에서 고주파대역을 예측한 후, 예측오차를 양자화할 때 GRNN을 이용한 예측기법이 양자화 효율이 가장 우수할 수 있음을 의미한다.

표 2는 예측된 영상과 원영상과의 평균 PSNR 결과이다. GRNN을 이용한 예측기법이 기저대역인  $B_1$  대역에서 BPN 비선형 예측기법보다 3.5dB 정도 성능이 우수하게 나타났지만 BMA 방식에 비해서 4.37dB 성능이 떨어졌다. 그러나  $B_1$  대역을 제외한 나머지 고주파대역의 예측에서는 GRNN 비선형 예측기법이 다른 기법에 비해 2~3dB 정도 성능이 우수함을 보여준다. 그림 4의 (a), (b)는 각각 BMA 와 GRNN을 이용한 예측영상을 나타낸다.

표 1. 2차원 대역분할 영상에서 예측오차 분산  
Table 1. Predicted error deviation of 2D subband image

Variance			
Subband	BPN 비선형 예측기	BMA 비선형 예측기	GRNN 비선형 예측기
B1	105	8.41	44.01
B2	40.33	44.21	30.27
B3	17.77	17.48	10.09
B4	20.01	22.46	10.41
B5	17.70	5.42	0.91
B6	17.46	14.76	1.92
B7	21.71	14.95	4.81

표 2. 2차원 대역분할에서 예측영상의 평균 PSNR  
Table 1. Average PSNR of predicted image in 2D subband

Average PSNR(dB)			
Subband	BPN 비선형 예측기	BMA 비선형 예측기	GRNN 비선형 예측기
B1	28.15	36.03	31.66
B2	28.09	27.71	29.33
B3	27.53	27.63	29.65
B4	28.41	27.83	31.18
B5	29.99	35.14	35.60
B6	24.39	25.33	27.12
B7	17.63	20.65	24.09

그림 3은 Miss America 동영상을 2차원 대역분할한 후, 움직임 예측없이 DPCM 및 PCM 기법만을

이용하여 부호화한 경우 프레임 변화에 따른 PSNR의 변화, BMA 동영상 움직임 예측 알고리즘을 사용한 후 부호화 경우 및 본 연구에서 제안한 GRNN 비선형 예측기를 사용하여 움직임을 예측한 후 예측 오차 영상에 대한 부호화한 경우를 비교하여 보여주고 있다. 그림에서 알 수 있듯이 GRNN 비선형 예측기를 사용하여 움직임을 예측한 후 부호화하는 경우가 가장 효과적으로 부호화를 할 수 있음을 알 수 있다.

그림 4는 14번째 프레임의 전체 복원영상 및 눈 부분만을 확대하여 보여주고 있다.

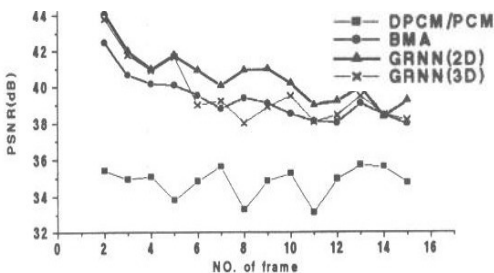


그림 3. 프레임 변화에 대한 PSNR  
Fig. 3 PSNR to frame change



그림 4. 복원영상(14번째 프레임)  
Fig. 4 Reconstruction image of 14th frame

### V. 결론

본 연구는 동영상 대역분할 부호화 및 웨이브렛 변환 부호화 등에서 사용될 수 있는 신경회로망을 이용한 비선형 예측기를 제안하였다. 심장조영상 신호를 2차원으로 대역 분할하여 제안된 알고리즘을 적용한 결과 다음과 같은 실험결과를 얻을 수 있었다.

첫째, Specht이 제안한 신경회로망 GRNN을 동영상 예측에 적용할 수 있도록 그 방법을 제시하였으며 둘째, 클러스터링 반경은 고주파 대역간 예측에는  $r=30$ , 기저 대역 간 예측에는  $r=10$  정도가 적당함을 알 수 있었으며 셋째, 대역분할된 동영상 예측시 GRNN을 이용한 비선형 예측기법은 BMA 방식보다 기저대역을 제외한 고주파대역에서는 예측오차의 분산이 상대적으로 작게 나타므로 보다 효과적으로 부호화할 수 있다는 결과를 얻을 수 있었다. 향후 제안된 동영상 예측기법을 다양한 동영상 부호화에 적용하여 부호화 성능에 대한 효과를 구체적으로 검증할 예정이다.

### 참고 문헌

- [1] S. Kappagantula and K. R. Rao, "Motion compensated interframe image prediction," IEEE Trans. Comm., Vol. COM-33, pp.1011-1014, Sept. 1985
- [2] D. F. Specht, "A general regression neural network," IEEE Trans. Neural Networks, Vol. 2, No. 6, pp. 568-576, 1991.
- [3] K.Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward network are universal approximators," Neural Network, Vol. 2, pp. 359-366, 1989.
- [4] H. Gharavi, "Subband coding Algorithm for video applications: videophone to HDTV-conferencing," IEEE Trans. Circuit Syst. Video Tech., Vol 1, No. 2, pp. 174-183, June 1991.
- [5] O. Johnson, O. V. Shentov and S. K. Mitra, "A technique for the efficient coding of the upper bands in subband coding of images," Proc. ICASSP, pp. 2097-2100. April 1990.
- [6] P. H. Westerink, J. Biemond and D. E. Boekee, "An optimal bit allocation algorithm for subband coding," Proc. ICASSP, pp. 757-760. April 1988.
- [7] D. L. Galland A. Tabatabai, "Subband coding of images using symmetric short kernel filters", Proc. ICASSP, Vol. 3, pp. 450-461, 2005.

## 저자 소개



### 한영오(Young-Oh Han)

1886년 2월 : 연세대학교 전기공학과 졸업 (공학사)

1989년 8월 : 연세대학교 대학원 전기공학과 졸업(공학석사)

1995년 8월 : 연세대학교 대학원 전기공학과 졸업(공학박사)

1996년 3월 ~ 현재: 남서울대학교 전자공학과 부교수

※ 주 관심분야 : 디지털 신호처리, 무선센서 네트워크, U-Health