

# 영상의 국부 주파수 특성에 기반을 둔 deinterlacing

우동헌, 강현민, 엄일규\*, 김유신  
부산대학교 전자공학과  
밀양대학교 정보통신공학과\*

## Deinterlacing based on the local frequency characteristic of image

Dong Hun Woo, Hyun Min Kang, Il Kyu Eom\*, Yoo Shin Kim  
Department of Electronic Engineering, Pusan National University  
Department of Information and Communication, Miryang National University\*  
E-mail : dhwoo@pusan.ac.kr

### Abstract

영상은 주파수 분포로 볼 때, 크게 에지 영역과 평탄 영역으로 구분할 수 있다. 또한 각 영역 안에서도 다양한 주파수 분포를 가지고 있다. 기존의 가중치 합(weighted-sum)기반의 방법으로는 다양한 주파수 변화를 포괄하기가 어렵고, 특히 고주파 성분을 효과적으로 보전할 수 없다. 이런 문제를 해결하기 위해서 본 논문에서는 영상의 국부 영역의 주파수 특성에 기반을 둔 deinterlacing 방법을 제안한다. 제안 방법에서는 우선 영상을 에지 영역과 평탄 영역으로 구분하였다. 그런 다음, 각 영역에 대해 신경 회로망을 할당하였다. 또한 각 영역에 대해서 저주파 성분과 고주파 성분으로 모듈화된 신경 회로망을 사용하여 각 영역 안에서의 주파수 특성의 변화에 잘 반응할 수 있게 하였다. 모의 실험에서 제안 알고리즘은 단일 신경 회로망의 경우나 기존의 다른 알고리즘들보다 개선된 성능을 보여주었다

### I. 서론

기존의 SDTV 의 비월주사 방식의 영상을 PC 용 모니터나 HDTV 의 순차 주사 방식으로 전환해주는 deinterlacing 은 일종의 2 배 영상 보간이다. 따라서 deinterlacing 은 영상 확대 문제에 적용될 수 있으며, 더 나아가, 동영상의 확대, 압축이나 superresolution 으로

적용도 가능하다[3].

기존의 deinterlacing 방법은 크게 움직임 정보를 이용한 것[3][4]와 그렇지 않은 것으로 구분할 수 있다. 움직임 정보를 이용하는 방법들은 좋은 성능을 내지만, 계산량 부담이 크고, 움직임을 잘못 찾았을 때 발생하는 화질의 손상이 크다. 움직임 정보를 사용하지 않는 방법들은 ELA(Edge Lined based Algorithms)[6]로 대표되는 가중치 합 기반 알고리즘들이 대부분이다. 이 알고리즘들은 가중치 합이므로 심각한 손상은 없으나 영상의 고주파 성분을 효율적으로 보전할 수 없다.

신경 회로망을 이용한 영상 보간의 적용할 때, 학습 과정을 통해 미리 보간될 값에 대한 정보를 활용하게 되므로 다른 가중치 합 기반의 알고리즘들과는 달리 고주파의 특성을 잘 살려줄 수 있다. 그러나 영상 보간 및 deinterlacing 문제에서의 신경 회로망의 적용은 패턴 인식에서의 적용과는 차이점이 있다. 즉, 패턴 인식에 있어서는 기대 출력(desired output)이 분류하고자 하는 부류(class)를 나타내는 일종의 코드값이지만, 보간 문제에서 출력은 보간하려는 값이 되므로, 이를 얼마나 정확히 추정하는 것이 중요한 문제가 된다. 따라서 패턴 인식 문제에 있어서는 신경 회로망의 학습 과정에서 어느 한 부류의 패턴에만 편향된 학습이 되지 않도록 하여 다른 부류에 속하는 패턴에 대해서는 다른 기대 출력 코드를 내는 것이 중요하며, 기대 출력값 자체를 얼마나 정확하게 출력하는지 하는 것은 크게 중요하지

않다[5]. 그러나 영상의 다운샘플 된 부분을 추정하는 영상 보간의 문제에서는 기대 출력에 대한 정확한 추정이 중요한 문제가 되므로 상이한 특성의 데이터에 대한 많은 학습은 추정의 정밀도를 떨어뜨린다. 따라서 제안 방법에서는 영상을 에지 영역과 평탄 영역으로 구분하고 각 영역에 독립적인 신경 회로망을 할당함으로써 추정의 정확성을 높였다. 한편, 각 영역에 대해서도 주파수 특성 측면에서 다양한 형태의 패턴이 존재한다. 그러나 이를 더 세분화 하는 것은 알고리즘이 복잡해지고, 각 영역의 학습 데이터가 감소하므로 효율적이지 않다. 따라서 제안 알고리즘에서는 영상의 국부 영역의 고주파 성분과 저주파 성분을 입력으로 가지는 모듈화 된 신경 회로망을 사용해서 한 영역 안의 다양한 변화에 대해 잘 적응하도록 했다.

모의 실험에서 제안 알고리즘은 PSNR 측면에서 단일 신경 회로망이나 다른 기존 알고리즘들보다. 개선된 성능을 보였으며, 주관적 화질 평가에서도 좋은 성능을 보여주었다.

## II. 신경 회로망을 이용한 deinterlacing

신경 회로망은 인간의 두뇌 활동을 모델링하였으며 학습 과정을 통해 주어진 입력에 대해 기대 출력이 나오도록 가중치 값을 조절한다. 가장 유명한 신경 회로망은 BP(Back Propagation) 학습 알고리즘을 가지는 MLP(Multi-Layer Perceptron)이며 그 구조는 그림 1 과 같다. 그림 1 에서 보이는 입출력 관계는 식(1)에서와 같이 수학적으로 표현할 수 있다.

$$y = \Gamma[W\Gamma[Vx]] \quad (1)$$

위 식에서  $y$  는 MLP 의 출력 벡터이며,  $w, v, \Gamma$  는 각각 출력층 가중치 행렬, 입력층 가중치 행렬, 활성화 함수(activation function) 행렬이다. 식(1)에서 보이는 것 같이 신경회로망은 일종의 선형 필터에 활성화 함수에 의한 비선형성을 가한 형태를 가지고 있다. 학습은 이 가중치 행렬  $W, V$  을 기대 출력이 나오도록 조절하는 과정이다. 학습 알고리즘은 최급강하(Gradient-descent) 알고리즘에 기반을 둔 BP(Back Propagation)알고리즘을 사용한다[2].

신경 회로망을 영상 보간에 응용할 경우, 신경 회로망의 학습에서 기대출력은 보간할 화소값이 된다. 식

(1)에서의 가중치 행렬  $W, V$  들은 음의 값을 가질 수 있다. 이것은 가중치 값으로 양의 값을 가져야 하는 ELA 를 비롯한 많은 기존의 deinterlacing 방법들과의 차이점이다. 이를 통해 신경 회로망에 기반을 둔 deinterlacing 은 고주파 성분에 대해 효율적인 보간이 가능하므로 영상의 에지의 선명성을 살릴 수 있다. 한편, 신경 회로망은 학습 데이터의 크기가 클수록, 또한 패턴의 형태가 다양할수록 학습의 효율은 떨어진다. 특히 영상 보간의 경우에는 패턴 인식과는 다른 고려가 필요하다. 즉 패턴 인식의 경우는 기대 출력은 하나의 부류(class)를 나타내는 일종의 코드이다. 그러나 영상 보간의 경우는 기대 출력은 실제 보간해야 할 값이다. 패턴 인식의 경우는 서로 다른 부류(class)에 속하는 패턴에 대해서 다른 코드를 출력하면 되며, 기대 출력 그 자체에 얼마나 근접하느냐는 큰 의미가 없다. 따라서 한 부류에 속하는 데이터에 대해서만 너무 편향된 학습을 하지 않는 것이 중요하며 이를 위해 학습 데이터의 입력을 난수적으로 넣어주는 등의 여러 방법을 사용한다. 반면에 영상 보간의 경우는 기대 출력은 보간할 화소값이 된다. 따라서, 신경 회로망의 출력이 실제 출력에 얼마나 근접하느냐 하는 것이 중요한 관심사가 되며, 이런 경우 서로 판이하게 다른 학습 데이터들은 서로의 기대출력 추정의 정밀도를 떨어뜨리게 된다. 일반적으로 영상 보간의 경우에는 크게 에지 영역과 평탄 영역의 주파수 특성이 판이하게 다르므로, 이것이 학습의 효율을 떨어뜨리는 주 요인이 된다.

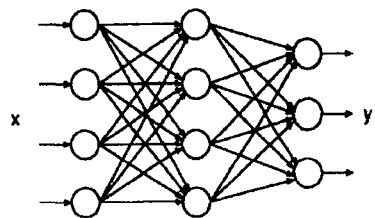


그림 1. MLP(Multi-Layer Perceptron)의 구조

## III. 제안 방법

앞 절에서 언급한 문제를 해결하기 위해 본 절에서는 먼저 영상을 주파수의 특성에 따라 에지와 평탄 영역으로 구분하여 각 영역에 신경 회로망을 적용하는 방

법을 설명하고, 그 다음에 각 영역의 다양한 주파수 변화에 적용할 수 있는 모듈화 된 신경 회로망을 제안한다. 제안 방법에 쓰인 신경 회로망의 입출력 구조는 그림 2 와 같다. 이 구조는 단일 신경 회로망에서 Plaziac 의 방법[1]에 비해 계산량은 작으면서 동등한 성능을 보여주었다.

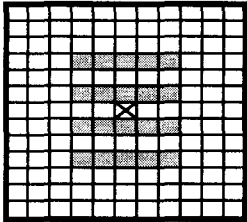


그림 2. 제안된 신경 회로망의 입력 구조

### 3.1 에지 영역과 평탄 영역에 분리 적용

제안 방법에서는 영상의 평탄 영역과 에지 영역을 구분하여 각각 독립적인 신경 회로망을 적용한다. 그림 3 는 제안된 알고리즘의 구조를 보여준다. 그림에서 보는 바와 같이 먼저 영상의 입력 화소는 국부 분산에 의해 에지 영역인지 평탄 영역인지를 구분된다. 각 영역에 대해 신경 회로망은 독립적으로 각 영역에 속하는 화소들에 의해 학습되고 테스트된다. 이와 같은 구조를 통해 신경 회로망은 주파수 측면에서 유사한 특징을 가지는 패턴들만을 집중적으로 학습하게 되므로 앞 절에서 언급한 학습효율의 저하를 개선할 수 있다. 또한 이 방법에서 신경 회로망의 입력은 화소값에서 국부 평균을 뺀 값을 사용한다. 이를 통해 같은 패턴이라도 다른 DC 값을 가지므로 생기는 학습의 지체를 피할 수 있다.

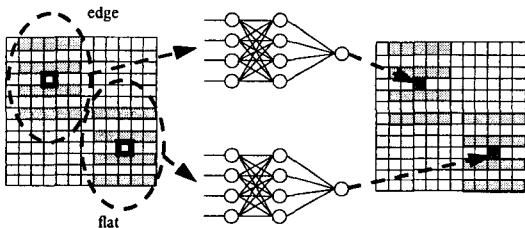


그림 3. 에지와 평탄 영역으로 분리되어

### 적용된 신경 회로망

#### 3.1 모듈화 된 신경 회로망

앞 절에서와 같은 방법을 사용해서 구분된 각 영역에서도 다양한 형태의 패턴들이 존재한다. 이러한 패턴들을 앞의 방법에서와 같이 더 세분화 된 영역으로 나누는 방법은 알고리즘의 복잡성을 증가시킨다. 또한 각 영역 안의 다른 데이터들은 어느 정도 유사성이 있기 때문에 이들을 완전히 분리하는 것은 사용할 수 있는 학습 데이터의 감소를 초래하므로 학습의 효율을 저하시키게 된다. 이 문제들을 극복하고 한 영역 안의 다양한 패턴 변화들을 효율적으로 학습하기 위해 그림 4 과 같은 구조의 모듈화 된 신경 회로망을 제안한다. 그림 4 의 신경 회로망의 저주파 모듈과 고주파 모듈의 입력의 구조는 그림 5 와 같다. 그림에서 신경 회로망의 출력은 'x'로 표시되어 있다. 제안된 신경 회로망의 저주파 모듈의 입력은 그림 5 에서 보는 바와 같이 국부 영역에서 9 개의 그룹이다. 이 그룹들은 각각 그 안의 화소들의 평균값을 나타낸다. 반면, 고주파 모듈은 그림에서 보는 바와 같이, 국부 영역에서의 화소값에서 국부 평균을 뺀 값이다. 이처럼 제안된 신경 회로망의 두 모듈은 서로 독립적인 구조를 가지며 출력 노드에서 결합되어 있다. 따라서, 주어진 입력에 대해 각 주파수 모듈이 독립적으로 주어진 입력에 대해 기대 출력을 내도록 학습을 하되, 출력 노드를 함께 공유하므로, 다른 모듈과 상호 보완 작용을 할 수 있다. 즉 입력 영상의 화소값을 직접 입력하는 경우보다 고주파 성분들과 저주파 성분들을 분해해서 볼 때, 입력 패턴의 특성이 분명하게 드러나게 된다.

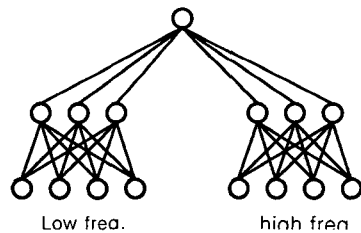


그림 4. 영상의 저주파 성분과 고주파 성분에 의해 모듈화 된 신경 회로망

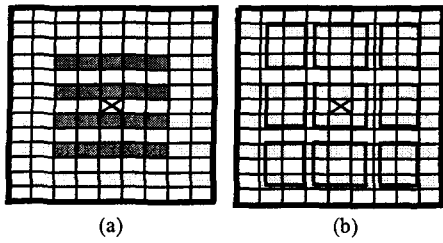


그림 5. 제안된 신경 회로망의 입력 구조  
(a) 고주파 모듈 (b) 저주파 모듈

#### IV. 모의 실험

제안된 알고리즘의 성능 평가를 위해서, 본 논문에서는 tiffany, truck, tank, airplane, aerial, zelda, couple, mandrill, elaine, goldhill, crowd, bridge 의 12 개의 영상이 학습을 위해 사용되었다. 객관적인 성능 평가를 위해서 제안 알고리즘과 다른 deinterlacing 알고리즘[1][6]의 PSNR 을 구했다. 표 1 은 그 결과를 보여준다. 표 1 에서 NN(2ch)은 에지 영역과 평탄 영역으로 나눈 뒤 각 영역에 MLP 를 적용한 경우이며 Modular NN 은 각 영역에 제안된 모듈화 된 신경 회로망을 적용한 것 뜻한다. 표에서 보는 바와 같이 제안 알고리즘들이 단일 신경 회로망의 경우나 다른 알고리즘들의 경우보다 개선된 PSNR 성능을 보여준다. 그림 6 는 lena 영상에 대한 제안 알고리즘의 주관적인 비교를 보여준다. 비교를 위해 선형(linear) 필터의 결과 이미지가 같이 보여졌다. 두 영상을 비교해 볼 때, 선형 필터의 이미지에서는 모자챙 부분에서 블록 현상이 발생하는 반면에 제안 알고리즘은 부드럽게 보간하는 것을 볼 수 있다.

#### V. 결론

본 논문에서는 영상의 국부 주파수 특성에 기반을 둔 deinterlacing 방법을 제안했다. 제안 알고리즘에서는 영상을 에지 영역과 평탄 영역으로 구분하여 독립적인 신경 회로망을 할당함으로써 상이한 학습 데이터로 인한 학습의 효율 저하를 개선하였고, 또한 각 영역에 주파수 특성으로 모듈화 된 신경 회로망을 적용함으로써 각 영역 안의 다양한 변화에 잘 적용할 수 있게 하였다. 모의 실험에서 제안 알고리즘은 단일 신경 회로망을 이용한 방법이나 다른 알고리즘들에 비해 개선된 성능을 보였다.

표 1. 제안 알고리즘과 다른 알고리즘들과의 PSNR 비교

	ELA	Linear	Median	Single NN	NN (2ch)	Moduar NN
Lena	33.59	38.03	36.78	38.28	38.55	38.64
Boat	30.70	36.88	33.45	36.77	36.86	37.02
Peppers	33.75	33.51	32.73	35.60	36.09	36.12
Barbara	24.04	33.44	27.93	32.22	33.20	33.51

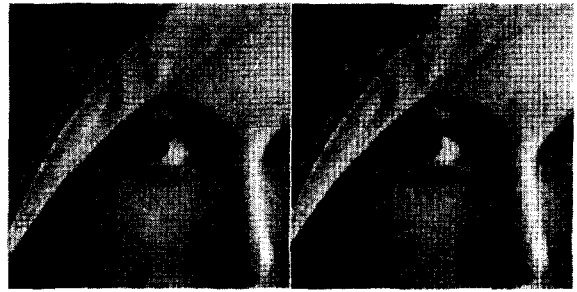


그림 6 주관적 비교 (a) 제안 방법 (b) 선형 필터

#### 참고문헌

- [1] Nathalie Plaziac "Image interpolation using neural networks", IEEE trans. image processing, Vol. 8, No. 11, pp. 1647-1651, Nov. 1999.
- [2] J. M. Zurada, Introduction to artificial neural system, PWS Publishing Co., 1995.
- [3] B. Martins et. al. "A unified approach to restoration, deinterlacing and resolution enhancement in decoding MPEG-2 video", IEEE trans. circuit and systems for video technology, Vol. 12, No. 9, pp. 803-811, Sep. 2002.
- [4] Y. Y. Jung et. al. "An effective de-interlacing technique using motion compensated interpolation" IEEE trans. consumer electronics, Vol. 46, No. 3, pp. 460-466, Aug 2000.
- [5] S. W. Lee " Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network," IEEE trans. Pattern Analysis and Machine Intelligence, vol. 18, pp.648-652, June 1996.
- [6] G. D. Haan, and E. B. Bellers, "Deinterlacing - an Overview", In Proc. IEEE, Vol. 86, No. 9, pp. 1839-1857, Sep. 1998.