

논문 2015-52-4-21

글자 영상을 위한 학습기반 초고해상도 기법

(Learning-based Super-resolution for Text Images)

허 보 영*, 송 병 철**

(Bo-Young Heo and Byung Cheol Song[Ⓢ])

요 약

본 논문은 글자 영상을 효과적으로 확대 (up-scaling)하기 위한 학습 기반 초고해상도 (super-resolution; SR) 기법을 제안한다. 제안 기법은 크게 학습 단계와 합성 단계로 나뉜다. 학습 단계에서 다양한 HR (high-resolution) /LR (low-resolution) 글자 영상 쌍들을 수집한다. LR영상들은 양자화를 하고, 충분히 많은 수의 HR-LR 블록쌍들을 추출한다. 양자화된 LR블록을 기준으로 블록 쌍들을 소정의 개수의 클래스들로 구분한다. 클래스 별로 최적의 2D-FIR 필터 계수를 계산하고, 양자화된 후 색인용 LR 블록과 함께 사전에 저장한다. 합성 단계에서 입력 LR 영상 내 각 블록을 양자화한 후 사전 내 양자화된 LR블록들과 정합하여 가장 근사한 블록에 대응하는 FIR 필터계수를 선정한다. 마지막으로 선택된 FIR필터로 HR 블록을 합성하여 최종적인 HR영상을 생성한다. 또한, 우리는 잡음이 있는 글자 영상에 대응하기 위해 학습과정에서 잡음 세기에 따른 복수개의 사전들을 제작한다. 입력 LR 영상의 잡음 레벨에 맞는 사전을 선택하여 HR영상을 합성한다. 실험 결과는 제안 기법이 종래 기법보다 잡음이 없는 환경에서는 물론 잡음이 있는 환경에서 우수한 주관적/객관적 화질을 가짐을 보인다.

Abstract

The proposed algorithm consists of two stages: the learning and synthesis stages. At the learning stage, we first collect various high-resolution (HR)-low-resolution (LR) text image pairs, and quantize the LR images, and extract HR-LR block pairs. Based on quantized LR blocks, the LR-HR block pairs are clustered into a pre-determined number of classes. For each class, an optimal 2D-FIR filter is computed, and it is stored into a dictionary with the corresponding LR block for indexing. At the synthesis stage, each quantized LR block in an input LR image is compared with every LR block in the dictionary, and the FIR filter of the best-matched LR block is selected. Finally, a HR block is synthesized with the chosen filter, and a final HR image is produced. Also, in order to cope with noisy environment, we generate multiple dictionaries according to noise level at the learning stage. So, the dictionary corresponding to the noise level of the input image is chosen, and a final HR image is produced using the selected dictionary. Experimental results show that the proposed algorithm outperforms the previous works for noisy images as well as noise-free images.

Keywords : text, super-resolution, noise, multi-dictionary, quantization.

* 학생회원, ** 평생회원, 인하대학교 전자공학과
(Department of Electronic Engineering, Inha University)

Ⓢ Corresponding Author (E-mail: bcsong@inha.ac.kr)

※ 본 연구는 2012년도 정부 (교육부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행되었으며, 미래창조과학부 및 정보통신산업진흥원의 ICT/SW 창의연구과정의 연구결과로 수행되었음. (NIPA-2014-H0502-14-3027)

Received ; December 12, 2014 Revised ; February 26, 2015

Accepted ; March 26, 2015

I. 서 론

최근 TV의 대형화가 이루어지고 스마트폰의 해상도가 높아짐에 따라 저 해상도의 영상을 고 해상도 영상으로 변환하는 스케일링 기법의 중요성이 커지고 있다. 그 중에서도 학습기반의 초고해상도 기법이 많이 연구되고 있는데, 일반적인 학습기반의 초고해상도 기법은 잡음이 없는 일반적인 영상(사람, 동물, 풍경 영상)을 대상으로 한다. 최근에는 오래된 문서를 파일로 보관하

는 일이 많아지고 있고, 영상 장치를 이용하여 글자가 포함 된 영상을 얻기도 한다. 하지만 오래 된 문서의 특징 상 파일로 저장되는 문서는 잡음이 섞이기 쉽다. 또한 영상 장치에 저장된 문서들은 저장의 한계 등으로 인해 잡음이 쉽게 생기고 해상도가 떨어지는 경향을 보인다. 이러한 영상에 일반적인 학습기반의 초고해상도 기법을 적용하면 효과적인 스케일링 결과를 기대할 수 없다.

일반적인 학습 기반의 초고해상도 기법은 크게 학습 단계와 합성 단계로 나누어진다.^[1~3] 학습 단계에서 HR/LR 영상 쌍을 가지고 유사한 블록들을 하나의 클래스(class)로 묶어서 사전들을 생성한다. 합성 단계에서는 입력 LR 영상 블록과 가장 유사한 블록을 사전에서 찾아 HR 합성에 이용한다. Freeman의 예제 기반 알고리즘이 가장 대표적인 학습 기반 초고해상도 기법이다.^[1] Freeman은 HR/LR 영상의 주파수 성분을 추출하여 HR 영상 합성에 이용하였기 때문에 영상의 디테일이 살아 있는 것을 볼 수 있다. 하지만 만족스러운 결과를 얻기 위해서는 최소 수십만 개 이상의 HR/LR 블록 쌍이 필요하기 때문에 블록 쌍을 저장하는 사전의 크기가 매우 커야한다. 또한 정합 과정에서 많은 연산량이 필요하다. 그리고 학습에 사용하는 영상은 여러 영상을 사용 할 수 있기 때문에 예지가 많은 글자 영상의 특징을 살려서 스케일링하기는 어렵다. 특히 잡음이 없는 영상을 가지고 사전을 구성하다보니 잡음이 있는 글자 영상에 대해서는 만족스러운 결과를 얻을 수 없다.

본 논문에서는 잡음이 없는 글자 영상에 특화된 초고해상도 기법을 먼저 제안한다. 이 기법은 단일 사전들을 구성하는 학습 기반의 초고해상도 기법으로 학습 단계와 합성 단계로 나누어 생각할 수 있다. 학습 단계에서는 글자 영상의 이진적인 특성을 이용하여 영상을 양자화하여 블록을 추출한다. 추출한 블록에서 패턴을 추출하고, 유사한 패턴들을 하나의 클러스터로 구성한다. 이때, 클러스터별 필터 계수 역시 양자화하여 사전에 저장하기 때문에 메모리 문제를 완화하고, 연산량 역시 줄일 수 있다. 합성 단계의 처음은 학습 단계와 마찬가지로 영상을 같은 bit 영상으로 양자화하여 블록 및 패턴을 추출한다. 그리고 미리 저장 된 사전의 필터 계수 정보를 이용하여 입력 LR 영상과 유사한 패턴을 가지는 블록을 찾아서 HR 영상을 합성한다. 하지만 잡음이 없는 환경에서 영상을 학습하여 사전을 구성하기 때문

에 잡음이 있는 입력 영상에 대해서 좋은 결과를 기대할 수 없다.

그래서 우리는 잡음이 있는 글자 영상을 위한 복수개의 사전에 기반을 둔 초고해상도 기법을 추가적으로 제안한다. 합성 단계에서 잡음이 있는 저해상도 영상에서 잡음 정도를 추측한 후, 이를 바탕으로 잡음을 제거한 저해상도 영상과 잡음이 없는 고해상도 영상의 쌍들을 이용하여 잡음 정도에 따라 복수개의 사전들을 생성한다. 합성 단계에서는 입력 영상의 잡음 정도를 추정하고 이를 바탕으로 잡음을 제거 한 입력 영상을 똑같은 잡음 정도로 구성된 사전에서 정보를 가져와 합성한다. 잡음 정도에 따라 각각 다른 사전을 생성하여 합성에 이용했기 때문에 하나의 사전만을 이용할 때 보다 더 좋은 효과를 기대할 수 있다.

본 논문의 II절에서는 잡음이 없는 글자 영상을 위한 초고해상도 기법을 제안하고, III절에서는 잡음이 있는 글자 영상을 위한 초고해상도 기법을 제안한다. IV절에서는 합성 실험을 통해 II절과 III절에서 제안한 기법들의 성능을 검증하고, 마지막 V절에서는 결론을 맺는다.

II. 잡음이 없는 글자 영상을 위한 초고해상도 기법

잡음이 없는 글자 영상을 위한 초고해상도 제안 기법

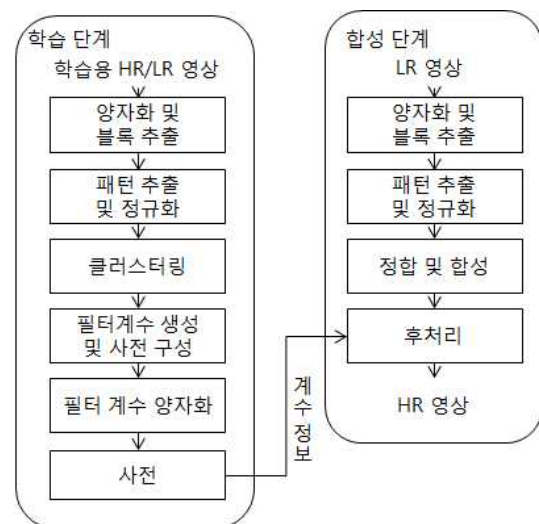


그림 1. 잡음이 없는 글자 영상을 위한 초고해상도 기법 알고리즘.

Fig. 1. Block diagram of the super-resolution for noise-free text images.

은 그림 1과 같이 학습 단계와 합성 단계로 나누어서 생각할 수 있다. 학습 단계에서는 높은 bit 영상을 낮은 bit 영상으로 전환하여 블록을 추출한다. 그 후 패턴을 추출하기 위해 라플라시안 연산자를 적용한다. 합성 시에 사용할 사진을 대표하는 필터 계수 역시 양자화를 통해 메모리를 줄여서 저장한다. 합성 단계에서는 학습 단계와 마찬가지로 영상을 양자화 하여 블록을 추출하고, 정합 과정을 거쳐 입력 LR 영상 블록과 가장 유사한 블록을 사전에서 찾아 HR 영상을 합성한다.

1. 학습 단계

가. 양자화 및 블록 추출

글자 영상의 밝기 히스토그램을 살펴보면 글자 영역과 배경 영역의 밝기값으로 이분화되는 것을 알 수 있다. 이런 특징을 이용하여 글자를 양자화하여 블록을 추출하면 잃어버리는 정보가 매우 적으면서 연산량을 줄여 고속화 할 수 있다. 그림 2는 8bit로 양자화한 영상과 4bit로 양자화한 영상을 7×7의 블록으로 추출하여 확대한 영상이다. 그림 3의 (a)와 (b)를 비교해보면 영상을 높은 bit에서 낮은 bit로 변환하여도 패턴의 모양에 변화가 없음을 알 수 있다.



그림 2. (a) 8bit 양자화 블록, (b) 4bit 양자화 블록
Fig 2. (a) 8bit quantized block, (b) 4bit quantized block.

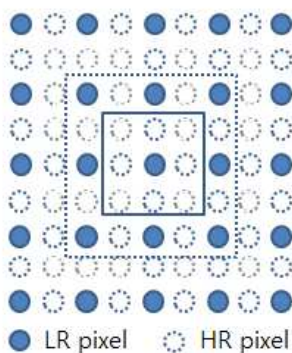


그림 3. 영상 내 저해상도/고해상도 블록 추출
Fig 3. Extraction of LR/HR blocks in an image.

영상의 양자화는 다음의 식 (1)을 통해서 할 수 있다.

$$img_b(m, n) = \frac{img_o(m, n)}{2^{o-b}} \quad (1)$$

img_b 는 목표 bit로 양자화된 영상이고, img_o 는 입력 영상으로 o, b 는 각 영상의 bit수를 나타낸다. (m, n) 은 영상 내 좌표값을 의미한다.

양자화 된 HR/LR영상 쌍은 $M \times M$ 크기를 갖는 블록 쌍으로 추출한다. 이 때, HR 영상 블록은 그림 3과 같이 LR 영상 블록에 대응되는 HR 영상의 중앙 영역에서 추출된다.

나. 라플라시안 패턴 추출 및 정규화

글자 영상은 에지가 많은 특징을 가지고 있다. 이 특징을 활용하기 위해 LR 영상 블록에 전형적인 라플라시안 연산자를 적용하여 패턴을 추출한다. 글자 영상은 에지의 특징을 가지는 영상으로 볼 수 있기 때문에 특징을 충분히 반영한 패턴 추출을 위해서이다. 또한, 추출된 모든 패턴에 대해서 정규화 작업을 수행하여 정규화 된 LR 블록인 LRN 을 생성한다. 이는 패턴은 같지만 밝기값으로 인한 차이를 제외시키기 위해서이다. LR 블록의 평균을 μ , 표준편차를 σ 라고 할 때, LR블록 내 각 화소값들에 대한 정규화는 식 (2)와 같이 정의된다.

$$LRN(m, n) = \frac{LR(m, n) - \mu}{\sigma} \quad (2)$$

LRN 은 블록 분류에 이용된다.

다. LRN 기준 클러스터링

양자화를 통해 추출된 HR/LR 영상 블록 쌍들은 일정한 기준에 따라서 몇 개의 클러스터들로 분류한다. 본 논문에서는 SSE(Sum of Squared Errors) 기반의 K-means clustering을 이용하여 몇 개의 클러스터들로 분류하였다. 같은 클러스터로 분류된 블록들의 평균값을 이용하여 각 클러스터의 대표값으로 사용한다.^[4]

라. 클러스터별 필터 계수 생성 및 사전 구성

이 단계에서는 클러스터별로 필터 계수를 생성하여 사전을 구성한다. 필터 계수는 해당 클러스터에 대응되는 저해상도 블록이 입력 될 경우 HR 영상 블록을 합성할 때 사용된다. 필터 계수 계산에는 분류된 각 클러

스터의 HR/LR 영상 블록 쌍들을 이용하며. HR 영상 블록과 LR 영상 블록과의 관계를 정의한 식(3)은 아래와 같다.

$$I_{HR}(i, j) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} W_{k,i,j}(m, n) I_{LR}(m, n) \quad (3)$$

식 (3)에서 (m, n)은 입력 LR 블록의 화소위치, (i, j)는 출력 HR 블록의 화소 위치를 나타내고, k는 합성과정 시 선택된 클러스터 번호이다.

위 식은 $HR = W \times LR$ 처럼 간단한 행렬식으로 나타낼 수 있다. Gauss-Jordan 소거법을 이용하여 $W = HR \times LR^{-1}$ 과 같이 계수 항으로 정리하여 나타낼 수 있다^[5].

최종적으로 구해진 계수는 식(4)와 같은 형태로 저장된다.

$$W_k = \begin{bmatrix} W_{k,0,0} & W_{k,0,1} & & W_{k,0,MN-1} \\ W_{k,1,0} & W_{k,1,1} & \cdots & W_{k,1,MN-1} \\ W_{k,2,0} & W_{k,2,1} & & W_{k,2,MN-1} \\ & \vdots & \ddots & \vdots \\ W_{k,IJ-1,0} & W_{k,IJ-1,1} & \cdots & W_{k,IJ-1,MN-1} \end{bmatrix} \quad (4)$$

W의 첫 행은 HR블록의 첫 번째 화소를 생성하기 위한 계수 모임이다. 각행의 계수와 입력 LR 블록 정보를 이용하여 HR 화소를 생성한다.

마. 필터 계수 양자화 및 저장

필터 계수 역시 글자 영상이 이진 영상에 가깝다는 특성을 이용하여 양자화시키면 실수 계수를 정수로 저장하기 때문에 적은 메모리만 사용할 수 있다. 필터 계수의 양자화는 식(5)를 이용한다.

$$W_{k,q} = R(2^{bitq} \cdot W_k) \quad (5)$$

식 (5)에서 bit는 양자화 정도를 조절하는 파라미터이고, R은 반올림 함수이다.

2. 합성 단계

가. 라플라시안 및 정규화

사전에 대푯값으로 저장되어 있는 LRN 블록과 동

일한 형태로 만들기 위해서 추출된 블록에 대하여 라플라시안 연산을 적용하여 정규화 한다. 이 때 학습 단계에서 사용한 식(2)를 이용하여 최종적으로 LRN_{IV} 블록을 생성한다.

나. 정합 및 HR 블록 합성

HR 블록을 합성하기 위해서 앞의 과정에서 생성한 LRN_{IV} 블록을 사전에 포함되어있는 K개의 대표 블록들과 비교하여 가장 유사한 블록 패턴을 가지는 클러스터를 찾는다. 비교하는 과정은 수행하기 위해 SAD를 사용하여 오차가 가장 작은 클러스터를 합성 시 사용할 클러스터로 선택한다. 클러스터에 저장되어 있는 필터 계수는 정수 형태이다. 따라서 식 (6)과 같이 양자화 과정의 역 과정을 거쳐 HR 블록 합성에 이용한다. \overline{W}_k 는 양자화된 계수 정보이다.

$$\overline{W}_k = W_{k,q} / 2^{bitq} \quad (6)$$

다. 후처리

합성된 HR 영상은 필터 계수를 이용하였기 때문에 완벽하게 고주파를 복원하지 못해 글자 주변에 약간의 블러(blur)와 잡음이 존재한다. 이를 제거하기 위해 Bimodal model을 이용한 후처리 과정을 거쳐 최종적인 HR 영상을 생성한다.^[6]

III. 잡음이 있는 글자 영상을 위한 초고해상도 기법

앞 절에서 소개한 글자 영상에 특화된 초고해상도 기법은 잡음이 없는 환경에서 영상을 학습하여 사전을 구성하였다. 따라서 만약 입력 LR 영상에 잡음이 섞여 있다면 만족스러운 결과를 얻을 수 없다. 앞 절에서 제안한 기법의 단점을 보완하기 위해서 본 절에서는 추가적으로 복수개의 사전에 기반을 두는 잡음이 있는 영상을 스케일링할 수 있는 기법을 제안한다. 제안 기법은 그림 4에서처럼 크게 학습 단계와 합성 단계로 나누어서 생각할 수 있다.

우선 학습 단계에서는 입력 영상이 다양한 잡음 레벨을 가질 수 있는 환경을 가정하여 학습을 통해 잡음의 레벨에 따라 여러 개의 사전으로 분류하였다. 먼저 학

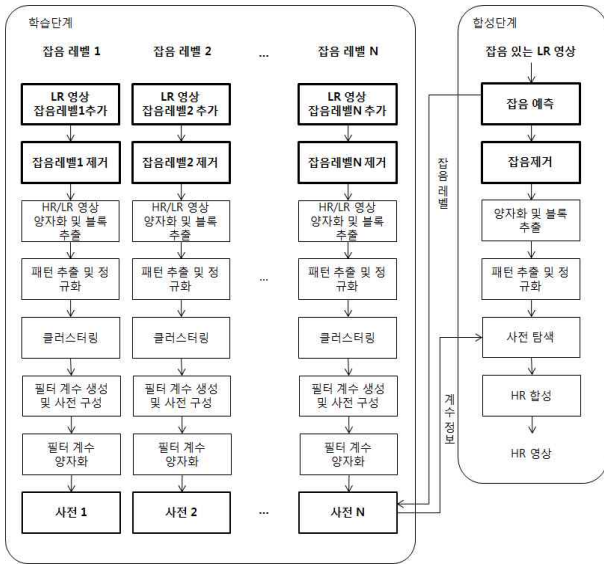


그림 4. 잡음이 있는 글자 영상의 초고해상도 기법 알고리즘
Fig 4. Block diagram of super-resolution algorithm for noisy text images.

습용 HR 영상을 다운샘플링 (Down-sampling)하여 얻은 LR 영상에 강제로 N개의 레벨을 갖는 잡음을 추가한다. 본 논문에서는 평균 0, 분산 $3^2, 6^2, 9^2, 12^2, 15^2$ 을 가지는 가산 백색 가우시안 잡음 (Additive White Gaussian Noise: AWGN)을 추가하였다. 잡음을 추가한 LR 영상에 대해 잡음을 제거하는 과정을 수행한다. 잡음 세기 별로 복수의 잡음 제거된 LR 영상들과 대응하는 HR 영상들이 준비되면, II절의 방식과 동일하게 사전 학습한다.

합성 단계에서는 잡음이 있는 입력 LR 영상의 잡음을 추정하고, 추정된 잡음 세기에 맞는 사전을 선택한다. 이후 과정은 II절의 방식과 동일하다.

IV. 실험 결과

본 절에서는 앞의 II절과 III절에서 제안한 기법들의 성능을 평가하기 위해서 합성 실험을 진행하였다. 먼저 II절에서 소개한 잡음이 없는 글자 영상을 위한 초고해상도 기법의 객관적인 평가를 위해 Bi-cubic과 Lanczos^[7] 방법을 사용하여 영상을 스케일링한 결과의 PSNR (Peak Signal to Noise Ratio)을 비교하였다. 제안 기법의 학습 단계에서는 인터넷 화면과 문서편집기에 작성된 글을 캡처한 총 10장의 1080p급 영상들을 사용하였다. 맑은 고딕과 신명조체의 폰트 10부터 14까지

표 1. PSNR 비교

Table 1. Comparison of PSNR.

PSNR(dB)		
폰트	맑은고딕	신명조
Bi-cubic	21.975	21.018
Lanczos	22.380	21.271
제안기법	25.714	23.597

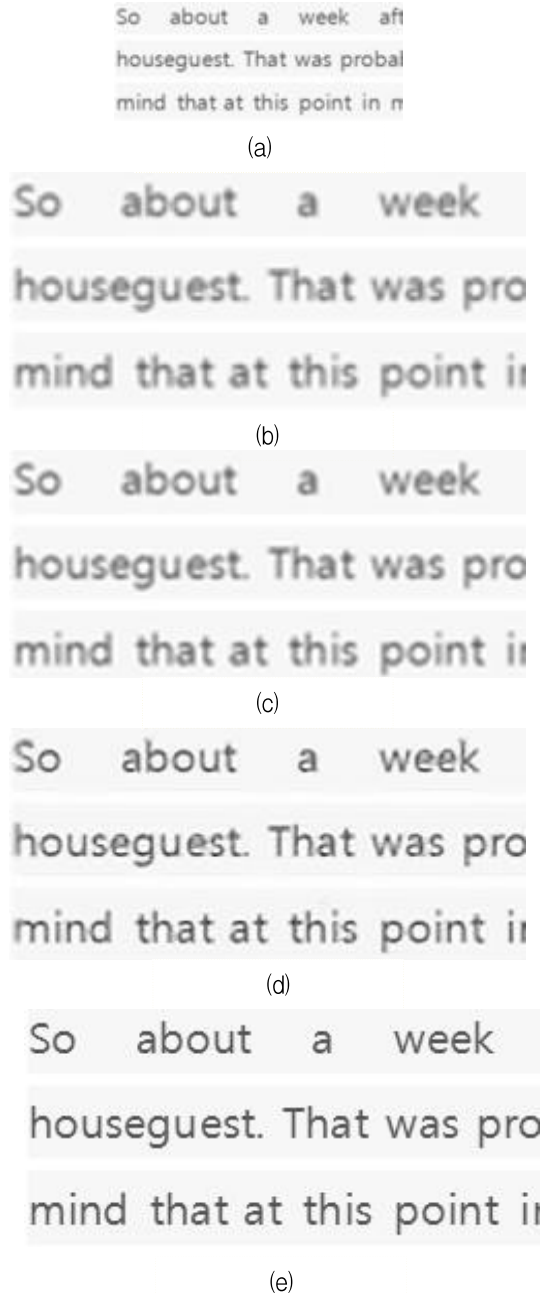


그림 5. (a) 입력 LR 영상, (b) Bi-cubic, (c) Lanczos, (d) 제안 기법, (e) 원본 HR 영상
Fig 5. (a) input LR image, (b) Bi-cubic, (c) Lanczos, (d) proposed, (e) original HR image.

의 문자들이 포함되도록 하였다. 클러스터는 512개로 분류하여 학습하였으며 학습 단계와 합성 단계에서 추출한 블록의 크기는 7×7 이다. 영상의 양자화 정도는 8bit 영상을 입력받아 4bit로 양자화 하였고, 사전의 필터 계수는 32bit float형을 저장 할 때에 5bit int형으로 저장하였다. 스케일링에 사용한 영상은 일반적으로 많이 사용하는 맑은 고딕과 신명조 폰트로 문서 편집기에서 작성한 글을 캡처하여 사용하였다. 스케일 비율은 1:2이다. 본 논문에서는 합성 단계에서 잡음의 세기를 미리 알고 있다고 가정하였으나, 성능이 우수한 다양한 잡음 세기 예측 방법들이 존재하기 때문에 이 중 하나를 채택하여 사용할 수 있다.

표 1에서 보는 것과 같이 제안 기법이 다른 방법들에 비해 평균적으로 약 3dB 높은 것을 알 수 있다. 그림 5는 맑은 고딕 폰트에 대한 입력 LR 영상과 Bi-cubic, Lanczos, 제안 기법 결과 영상의 일부이다. Bi-cubic과 Lanczos 방법을 사용하여 영상을 스케일링하였을 때는 블러 현상이 생겨 화질의 열화가 나타났으며 원본 HR 영상과 비교하였을 때, 예지가 선명하기 못한 것을 알 수 있다. 반면에 제안 기법은 원본 HR 영상에 비해서 다소 블러 현상이 보이지만 Bi-cubic이나 Lanczos 방법의 결과보다는 블러 현상이 적고, 글자의 예지 또한 선명하게 나타나는 효과를 얻을 수 있었다.

하지만 잡음이 없는 환경에서 알고리즘을 수행하였기 때문에 잡음이 있는 글자 영상에 대해서는 좋은 결과를 보이지 못했다. 그 결과는 그림 6과 같다. 그림 6

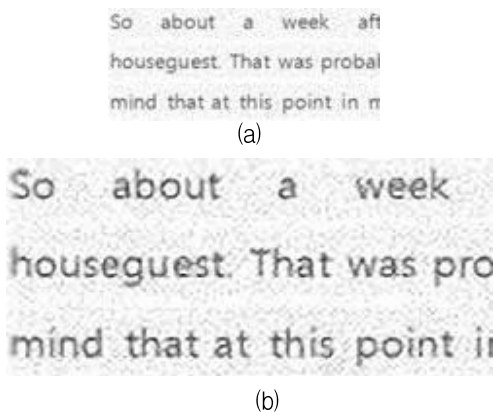


그림 6. 잡음 분산이 9^2 인 경우 (a) 잡음이 있는 입력 LR 영상, (b) HR 영상

Fig 6. In case of noise variance of 9^2 (a) an input noisy LR image, (b) HR image.

표 2. 비교 방법의 구성

Table 2. Configuration of bench marking methods.

	사전	학습 LR 영상	입력 LR 영상	결과영상 잡음제거
비교방법1	단일	잡음없음	잡음있음	X
비교방법2	단일	잡음없음	잡음있음	0
비교방법3	단일	잡음없음	잡음제거	X

의 (a)와 (b)는 각각 평균 0, 분산이 9^2 일 때의 AWGN를 첨가한 입력 LR 영상과 입력 영상에 따른 결과 HR 영상이다. 그림 6의 (b)를 보면 결과 영상에 잡음이 섞여 있는 것을 볼 수 있다. 실제 PSNR도 23dB가 나와 잡음이 없는 입력 LR 영상에 대한 결과의 PSNR인 25.7dB보다 낮게 나왔다.

따라서 이러한 단점을 보완하기 위해 III절에서는 잡음이 있는 글자 영상을 위한 초고해상도 기법을 제안하였다. 제안 기법의 성능을 비교하기 위해서 II절에서 소개한 기법에 입력 영상의 잡음 유무와 결과 영상의 후처리 유무에 따라 3가지 비교방법을 만들어서 사용하였다. 비교방법의 구성은 표 2와 같다. 학습을 위해서 사용한 영상은 앞의 실험과 동일한 영상을 사용하였다.

하지만 잡음이 있는 환경을 만들기 위해서 학습용 LR 영상에 평균 0, 분산은 $3^2, 6^2, 9^2, 12^2, 15^2$ 5가지 경우에 대하여 AWGN을 추가하고 잡음레벨에 따라 5개의 사전으로 분류하였다. 잡음 제거는 BM3D^[8]기법을 사용하였다. 그 외 클러스터 개수, 블록 크기, 영상과 필터 계수의 양자화 정도는 앞의 실험 조건과 동일하게 진행하였다. 스케일링에 사용한 영상은 문서 편집기에서 작성한 글을 캡처하여 사용하였으며 원본 HR 영상을 다운샘플링하여 LR 영상으로 만들었다. 그 후에 LR 영상에 학습 단계에서 구성한 사전의 잡음 레벨과 마찬가지로 평균 0, 분산은 $3^2, 6^2, 9^2, 12^2, 15^2$ 5가지 경우에 대하여 AWGN을 추가하고, BM3D를 이용하여 잡음을 제거하였다. 객관적인 평가를 위해 비교 방법과 PSNR을 비교하였다.

표 3의 결과를 보면 학습 영상이 잡음이 없는 환경이고, 단일 사전으로 구성된 비교방법 1, 2, 3보다 학습 시 잡음이 있는 환경을 고려하여 복수 개의 사전을 구성한 제안 기법에서 PSNR이 높게 나온 것을 볼 수 있다. 두 글자 모양에서 잡음의 분산 별로 PSNR이 최대 0.5dB 이상 차이가 난다. 하지만 잡음의 분산이 커질수록 PSNR이 낮게 나타난 것을 볼 수 있다. 이는 잡음의 세

표 3. 잡음 세기에 따른 PSNR
Table 3. PSNR results according to multiple noise variances.

PSNR(dB)					
σ^2	3^2	6^2	9^2	12^2	15^2
폰트	맑은 고딕				
비교 방법1	25.220	24.222	23.064	21.861	20.767
비교 방법2	25.370	24.699	23.880	22.982	22.158
비교 방법3	25.404	24.894	24.280	23.628	23.023
제안 기법	25.374	25.376	24.656	23.823	23.035
폰트	신명조				
비교 방법1	23.299	22.564	21.678	20.751	19.779
비교 방법2	23.373	22.843	22.205	21.536	20.815
비교 방법3	23.378	22.958	22.475	21.958	21.445
제안 기법	23.389	23.327	22.790	22.191	21.604

기가 커질수록 잡음 제거 과정에서 생기는 블러 현상이 심하게 나타나기 때문이다. 잡음 분산이 3^2 일 때는 제안 기법과 비교 기법의 PSNR 이 많은 차이를 보이지 않았는데, 이는 잡음의 세기가 약해 잡음의 영향을 상대적으로 덜 받았기 때문이다.

그림 7, 그림 8은 일부 잡음레벨에 대한 입력 영상과 결과 영상의 일부분이다. 비교 방법1은 잡음이 없는 환경에서 학습한 사전을 가지고 잡음이 있는 LR 영상을 스케일링했기 때문에 결과 영상에 잡음이 있는 것을 볼 수 있다. 그림 7와 그림 8의 비교방법 1의 결과 (c)와 비교방법 2의 결과 (d)를 비교하면 (d)의 경우 후처리로 잡음을 제거했음에도 불구하고, 여전히 잡음이 남아있었다. 이 현상은 잡음의 세기가 클수록 더욱 심하게 나타났다. 반면에 비교방법 3은 잡음을 제거한 LR 영상을 입력 영상(그림 7와 그림 8의 (b) 영상)으로 사용하였는데, 제안 기법과 유사한 결과가 나왔다. 하지만 잡음의 제거 과정에서 글자 영역 주변에 생긴 블러와 같은 화질 열화 현상들이 제안 기법보다 많이 생긴 것을 알 수 있다. 또한 제안 기법의 경우 비교방법에 비해 예지 부분이 선명한 것을 볼 수 있다. 이는 학습 단계에서 사용하는 학습 용 LR 영상에 잡음을 추가하여 잡음이 있는

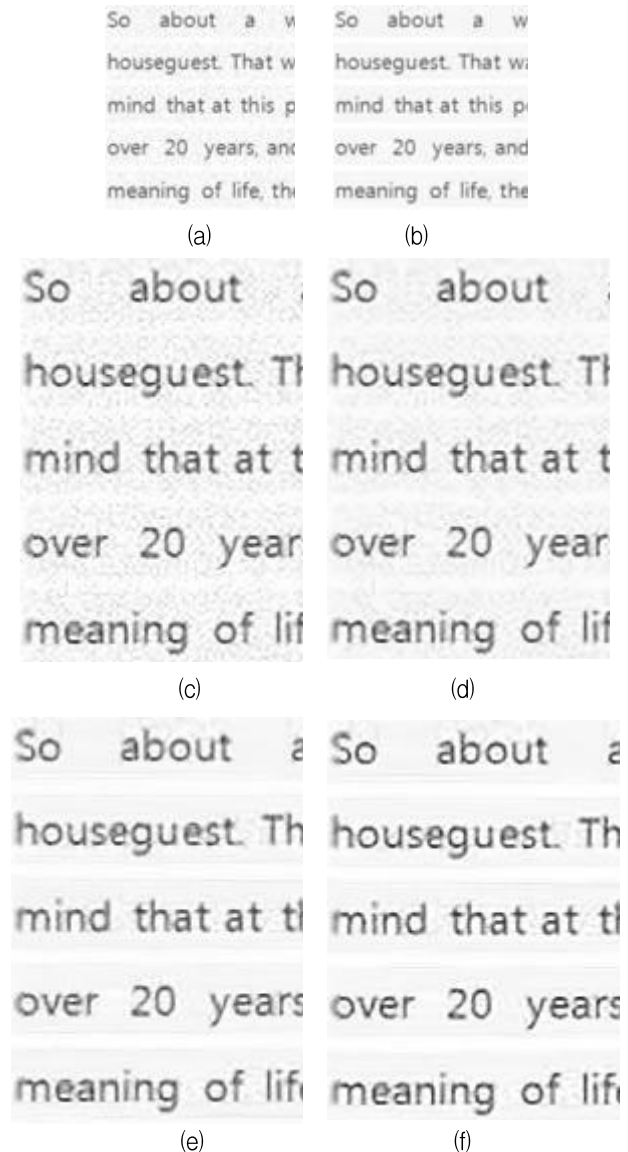


그림 7. 맑은 고딕 폰트, 잡음 분산 6^2 의 입력 영상에 대한 결과 (a) 잡음 LR 영상, (b) 잡음 제거한 LR 영상, (c) 비교방법 1, (d) 비교방법 2, (e) 비교방법 3, (f) 제안 기법

Fig 7. Result for an Input LR image of Malgun Gothic font and noise variance 6^2 (a) a noisy LR image, (b) de-noised LR image, (c) bench-marking 1, (d) bench-marking 2, (e) bench-marking 3, (f) proposed

환경에서 잡음의 세기별로 각각 다른 사전을 구성해서 여러 잡음의 레벨을 가지고 있는 입력 LR 영상에 대해 적절하게 대응 할 수 있도록 하였기 때문이다. 하지만 잡음의 분산이 커질수록 잡음 제거 과정에서 생기는 블러 효과 때문에 마침표, 쉼표 등의 문장 부호는 잘 합성

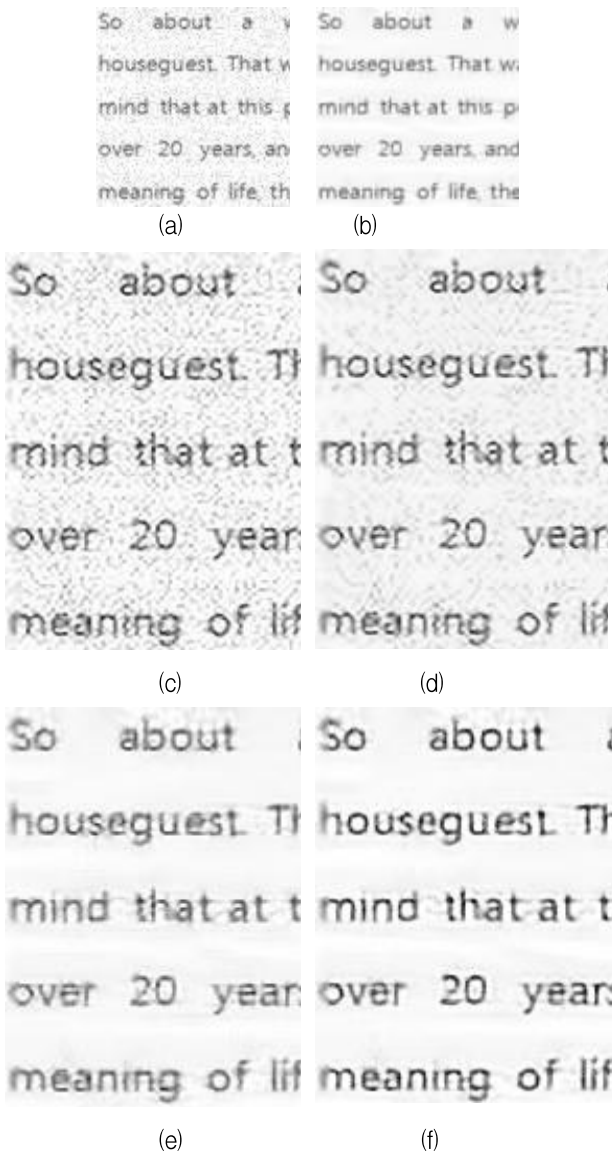


그림 8. 맑은 고딕 폰트, 잡음 분산 15^2 의 입력 영상에 대한 결과 (a) 잡음 LR 영상, (b) 잡음 제거한 LR 영상, (c) 비교방법 1, (d) 비교방법 2, (e) 비교방법 3, (f) 제안 기법

Fig 8. Result for an Input LR image of Malgun Gothic font and noise variance 6^2 (a) a noisy LR image, (b) de-noised LR image, (c) bench-marking 1, (d) bench-marking 2, (e) bench-marking 3, (f) proposed

되지 않았다. 이는 비교방법에서도 나타나는 현상으로 잡음의 강도가 클수록 작은 문자 부호들이 잡음으로 취급되어 같이 제거되기 때문이다.

V. 결 론

기존의 학습 기반 초고해상도 기법은 잡음이 없는 환경에서 학습이 이루어져 잡음이 있는 LR 영상에 대하여 만족스러운 결과를 얻을 수 없다, 또한 일반적으로 만족스러운 결과를 얻기 위해서는 사진을 구성하는 HR/LR 영상 블록 쌍들이 많아야 하는데 이는 많은 메모리를 요구한다. 본 논문은 글자 영상의 특징을 이용하여 영상을 양자화함으로써 메모리를 줄이고 연산량을 줄이고자 하였다. 또한 저장 매체 등에 의해 잡음이 생기는 현상을 고려하여 학습 단계에서 잡음의 레벨에 따라 N개의 사진을 구성하였다. 실험 결과 비교 방법 대비 제안 기법은 글자 주변에 블러와 같은 화질 열화 현상이 매우 적은 것을 알 수 있었다. 또한 글자의 에지 부분이 선명해 전체적으로 비교 기법에 비해 영상이 선명해졌다. 객관적인 평가를 위해 PSNR을 비교하였을 때 제안 기법이 비교 기법 대비 우수함을 볼 수 있다.

REFERENCES

- [1] W. Freeman, T. R. Jones, and E. C. Pasztor. "Example-based super-resolution." *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56-65, Oct. 2002
- [2] J. Park, Y. Kwon, and J. H. Kim, "An example-based prior model for text image super-resolution," *Proc. IEEE ICDAR*, pp. 374-378, 2005.
- [3] K. I. Kim, and Y. H. Kwon. "Example-based learning for single-image super-resolution," *Pattern Recognition*, pp. 456-465, 2008.
- [4] R. P. Duda, P. E. Hart, D. G. Stork, "Pattern Classification 2/e", *WILEY-INTERSCIENCE*, pp. 526-530.
- [5] H. Anton, R. C. Busby, "Contemporary linear algebra," *WILEY*, pp. 51-53.
- [6] H. Luong and W. Philips, "Robust reconstruction of low-resolution document images by exploiting repetitive character behaviour," *International Journal of Document Analysis and Recognition (IJ DAR)*, vol. 11, no. 1, pp. 39-51, 2008.
- [7] C. E. Duchon, "Lanczos filtering in one and two dimensions," *Journal of Applied Meteorology*, vol. 18, pp. 1016-1022, August 1979.
- [8] D. Kostadin et al., "BM3D image denoising with

- shape-adaptive principal component analysis.”
Signal Processing with Adaptive Sparse Structured Representations, 2009.
- [9] S. Yang, M. Wang, Y. Chen, Y. Sun, “Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding,” *IEEE Trans. Image Process.*, vol. 21, no. 9, 2012.
- [10] T. M. Chan et al. “Neighbor embedding based super-resolution algorithm through edge detection and feature selection,” *Pattern Recognition Letters*, pp. 494-502, 2009.
- [11] H. Luong, W. Philips, “Non-local text image reconstruction,” *Proc. ICDAR*, vol. 1, pp. 546-550, 2007.
- [12] R. Walha, F. Drira, F. Lebourgeois, A. M. Alimi, “Super-resolution of single text image by sparse representation,” *Proc. ICVGIP*, pp. 22-29, 2012.

— 저 자 소 개 —



허 보 영(학생회원)
2013년 인하대학교 전자공학과
학사.
2013년~현재 인하대학교 전자
공학과 석사 과정.
<주관심분야 : 영상처리, 선명도
개선>



송 병 철(평생회원)
1994년 한국과학기술원 전기 및
전자공학과 (학사).
1996년 한국과학기술원 전기 및
전자공학과 (석사).
2001년 한국과학기술원 전기 및
전자공학과 (공학박사).
2001년~2008년 삼성전자 디지털미디어연구소
책임 연구원
2008년~현재 인하대학교 전자공학부 교수
<주관심분야 : 영상 신호처리, 영상시스템/SoC>