

# Noise Removal using Support Vector Regression in Noisy Document Images

Hee Hoon Kim<sup>1</sup> · Seung Hyo Kang<sup>2</sup> · Jai Hyun Park<sup>3</sup> · Hyun Ho Ha<sup>4</sup> · Dong Hoon Lim<sup>5</sup>

<sup>1</sup>Korea Science Academy; <sup>2</sup>Korea Science Academy

<sup>3</sup>Korea Science Academy; <sup>4</sup>Korea Science Academy

<sup>5</sup>Department of Information Statistics and RINS, Gyeongsang National University

(Received June 18, 2012; Revised July 17, 2012; Accepted July 23, 2012)

---

## Abstract

Noise removal of document images is a necessary step during preprocessing to recognize characters effectively because it has influences greatly on processing speed and performance for character recognition. We have considered using the spatial filters such as traditional mean filters and Gaussian filters, and wavelet transformed based methods for noise deduction in natural images. However, these methods are not effective for the noise removal of document images. In this paper, we present noise removal of document images using support vector regression. The proposed approach consists of two steps which are SVR training step and SVR test step. We construct an optimal prediction model using grid search with cross-validation in SVR training step, and then apply it to noisy images to remove noises in test step. We evaluate our SVR based method both quantitatively and qualitatively for noise removal in Korean, English and Chinese character documents, and compare it to some existing methods. Experimental results indicate that the proposed method is more effective and can get satisfactory removal results.

Keywords: Cross-validation, grid search, support vector regression, noise removal.

---

## 1. 서론

오늘날 점점 늘어나는 대량의 문서를 자동으로 처리하기 위해 종이문서(paper document)를 직접 손으로 입력하지 않고 스캐너나 디지털 카메라와 같은 광학기기를 이용하여 얻어진 문서영상(document image)은 다양한 분야에서 활용되고 있다 (Don, 2001). Figure 1.1은 현재 운영하고 있는 디지털 도서관의 문서영상의 예로서 1945년 이전 신문에서 검색용어 “김구”를 사용했을 때 나타나는 화면을 보여주고 있다.

문서 영상에 대한 연구는 주로 문자의 인식분야에 집중되어 진행되고 있다. 최근 인터넷 포털 사이트들은 사용자들에게 문서 영상 검색서비스를 제공하고 있다. 이를 위해 광학 문자 판독 장치(optical character recognition; OCR)를 이용한 문서영상의 인식은 필수요소이지만 낮은 인식률로 인해 그 만족도

---

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(No.2012-0002709).

<sup>5</sup>Correspondence author: Professor, Department of Information Statistics and RINS, Gyeongsang National University, Jinju 660-701, Korea. E-mail: [dhlim@gnu.ac.kr](mailto:dhlim@gnu.ac.kr)



Figure 1.1. Example of document image

는 기대에 못 미치는 실정이다. 문서를 디지털로 처리하는 과정에서 잡음(noise) 혹은 왜곡(distort) 등에 의해 인식자체가 불가능한 경우가 발생하고 있다. 따라서, 문서영상에서 문자를 인식하기 위한 전처리 과정인 잡음제거(noise removal)는 잡음제거 성능에 따라 문자 인식의 처리 속도와 인식률에 크게 영향을 미치므로 이에 대한 집중적인 연구가 요구되는 실정이다.

지금까지 우리가 접하는 자연 영상(natural image)에서 잡음을 제거하는 전통적인 방법으로는 평균 필터(mean filter), 중앙값 필터(median filter) 그리고 가우시안 필터(Gaussian filter)를 사용하는 방법 등이 있다 (Gonzales와 Woods, 1992). 이 공간 필터들은 단순한 영상에서는 어느 정도 효과적이거나 다양한 영상특징을 갖고 있는 복잡한 영상에서 잡음제거에는 한계성을 갖고 있다. 이러한 한계를 극복하기 위한 방법으로 웨이블릿 변환(wavelet transform)을 이용한 잡음제거 연구가 많이 이루어지고 있다. 대표적인 방법으로는 VisuShrink 방법 (Donoho, 1995), SureShrink 방법 (Donoho와 Johnstone, 1994), BayesShrink 방법 (Vetterli 등, 2000) 등이 있다.

문서영상은 자연영상과 비교했을 때 다른 특징을 갖고 있다. 자연영상은 문서영상에 없는 희소표현(sparse representation)을 갖고 있다 (Li, 2009). 영상신호를  $y$ 라 하고 구성성분을  $x$ 라 할 때 다음과 같이 표현 가능한 경우  $y$ 는 희소표현을 갖는다고 한다.

$$y = Wx.$$

자연영상에서 웨이블릿 변환을 이용한 잡음 제거가 효과적인 이유는 자연영상이 희소표현을 갖기 때문이다.

최근에 복잡한 자연영상에 대한 효과적인 잡음제거를 위해 서포트 벡터 머신(support vector machine; SVM)과 같은 학습이론(learning theory)에 기반한 알고리즘 들이 많이 개발되고 있다 (Wang과 Fu, 2010; Wang 등, 2010; Vapnik, 1995; Lin과 Yu, 2004; Liu 등, 2006). SVM은 Vapnik (1995)에 의해 개발된 통계적 학습방법으로 구조적 위험 최소화(structural risk minimization)를 통해 해를 구하는 방법이다. SVM의 목적은 크게 분류(classification)와 회귀(regression) 두 가지로 나눌 수 있는데 지금까지 SVM은 주로 패턴 분류(pattern classification)하는데 많이 사용되었고 회귀를 위한 SVM 즉, 서포트 벡터 회귀(support vector regression; SVR)는 전통적인 회귀분석보다 우수한 예측력을 보이면서 최근에 각광을 받고 있는 분야이다 (Li 등, 2007).

본 논문에서는 SVR를 이용하여 문서영상에서 잡음을 제거하고자 한다. 영상에서 SVR 사용은 두 단

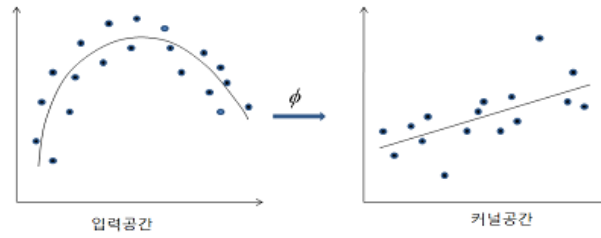


Figure 2.1. Mapping the input space into high dimensional feature space

계 즉, SVR 훈련단계(training phase)와 테스트 단계(test phase)를 거쳐 이루어진다. SVR 훈련단계에서 최적의 SVR 예측모형을 찾고 테스트 단계에서 잡음을 제거하고자 하는 영상에 예측모형을 적용하고자 한다. SVR에서 최적의 모수는  $k$ -조각 교차 타당성( $k$ -fold cross-validation)에 의한 그리드 탐색법(grid search method)을 사용하여 선택한다. 교차타당성은 SVR 추정함수가 훈련영상에 과적합(overfitting)되는 것을 피하기 위해 사용되고 빠른 최적의 해를 찾기 위해 그리드 탐색을 수행한다.

본 논문은 다음과 같이 구성되어 있다. 제 2절에서는 SVR에 대해 간략하게 살펴보고 제 3절에서는 문서영상에서 SVR을 이용한 잡음 제거 방법에 대해 구체적으로 논의하고자 한다. 제 4절에서는 제안된 잡음 제거 방법의 성능을 평가하기 위해 여러 가지 한글문서 영상, 영어문서 영상 그리고 한자 문서 영상에서 기존의 잡음 제거 방법들 간의 정성적이고 정량적인 비교를 하고 제 5절에서 결론 및 향후연구에 대해 논의한다.

## 2. 서포트 벡터 회귀(SVR)

훈련 데이터 집합(training dataset)  $\{(\mathbf{x}_i, y_i), i = 1, \dots, \ell\}$ 이 주어졌다고 가정하자. 여기서  $\mathbf{x}_i \in R^n$ 는 입력벡터이고  $y_i \in R$ 는 목표값을 나타낸다.

SVR은 Figure 2.1과 같이 입력 데이터를 비선형 커널 함수  $\phi(\cdot) : R^n \rightarrow R^m$ 에 의해 고차원의 커널 특징공간(feature space)으로 사상(mapping)시킨 후 새로 얻어진 선형 공간에서 다음과 같은 추정함수  $f(\mathbf{x}_i)$ 을 찾는 것이다.

$$f(\mathbf{x}_i) = \mathbf{w}^T \phi(\mathbf{x}_i) + b, \quad (2.1)$$

여기서  $\mathbf{w}$ 는 특징공간에서 가중치 벡터(weight vector)이고  $\mathbf{w}^T$ 는  $\mathbf{w}$ 의 전치행렬이고  $b$ 는 회귀에서 절편을 나타내는 상수항이다. 식 (2.1)에서  $f$ 가 편평(flatness)하기 위해서는  $\mathbf{w}$ 는 작은 값을 가져야한다. 따라서, 이것은 다음의 놈(norm)을 최소화함으로써 얻어진다.

$$\|\mathbf{w}\|^2 = \langle \mathbf{w}, \mathbf{w} \rangle$$

위의 문제를 다음과 같이 블록 최적화 문제(convex optimization problem)로 나타낼 수 있다.

$$\begin{aligned} & \text{minimize } \frac{1}{2} \|\mathbf{w}\|^2 \\ & \text{subject to } \begin{cases} y_i - (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \leq \epsilon, \\ (\mathbf{w}^T \phi(\mathbf{x}_i) + b) - y_i \leq \epsilon, \end{cases} \end{aligned} \quad (2.2)$$

여기서  $\epsilon$ 는 예측값과 실제값 사이의 오차로서  $\epsilon$ 보다 작으면 무시하고 크면 편차만큼 손실을 반영하기 위해 사용되는 모수이다. 이와 같은 오차를 갖고 있는 손실함수(loss function)를 Vapnik (1995)의  $\epsilon$ -insensitive 손실함수라 한다.

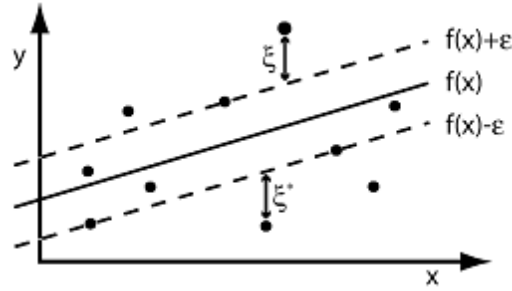


Figure 2.2.  $\epsilon$ -insensitive region, slack variables  $\xi$  and  $\xi^*$

일반적으로 Figure 2.2에서처럼 실제 훈련데이터들은  $\epsilon$ 보다 큰 편차를 갖는다. 이 경우 슬랙변수(slack variable)  $\xi$ 와  $\xi^*$ 을 도입하여 식 (2.2)의 최적화 문제를 다음과 같이 나타낼 수 있다.

$$\begin{aligned} & \text{minimize } \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{\ell} (\xi_i + \xi_i^*) & (2.3) \\ & \text{subject to } \begin{cases} y_i - (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \leq \epsilon + \xi_i, \\ (\mathbf{w}^T \phi(\mathbf{x}_i) + b) - y_i \leq \epsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, \end{cases} \end{aligned}$$

여기서  $C > 0$ 는 모형의 복잡도와  $\epsilon$ 보다 큰 오차 허용정도에 대한 트레이드 오프(trade-off)를 결정하는 상수이다. 여기서  $C$ 값이 크면 오차에 대해 큰 패널티(penalty)를 할당하여 SVR은 낮은 일반화 수준으로 오차를 최소화시킨다. 반면에  $C$ 값이 작으면 오차에 작은 패널티를 할당하여 높은 일반화 수준을 갖게 된다. 만약,  $C$ 값이 무한히 큰 값을 갖는다면 SVR은 어떤 오차도 허용하지 않을 것이고 결과적으로 복잡한 모형(complex model)이 될 것이다. 반면,  $C$ 가 0으로 접근하면 오차는 커지지만 간단한 예측 모형이 될 것이다. 따라서 적절한  $C$ 값 선택은 모형의 복잡도를 조정하는 효과뿐만 아니라 SVR의 성능을 높일 수 있다.

식 (2.3)의 원 문제(primal problem)를 라그랑지 함수(Lagrange function)를 이용하여 다음과 같이 쌍대문제(dual problem)로 변환시킬 수 있다.

$$\begin{aligned} L = & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{\ell} (\xi_i + \xi_i^*) - \sum_{i=1}^{\ell} (\eta_i \xi_i + \eta_i^* \xi_i^*) \\ & - \sum_{i=1}^{\ell} \alpha_i (\epsilon + \xi_i - y_i + \mathbf{w}^T \phi(\mathbf{x}_i) + b) - \sum_{i=1}^{\ell} \alpha_i^* (\epsilon + \xi_i^* + y_i - \mathbf{w}^T \phi(\mathbf{x}_i) - b), \end{aligned} \quad (2.4)$$

여기서  $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ 은 라그랑지 승수(Lagrange multiplier)이다. 따라서 식 (2.4)의  $L$ 을  $(\mathbf{w}, b, \xi_i, \xi_i^*)$ 에 대해 편미분하면 다음과 같은 쌍대 최적화 문제를 도출할 수 있다.

$$\begin{aligned} & \text{minimize } L = -\frac{1}{2} \sum_{i,j=1}^{\ell} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(\mathbf{x}_i, \mathbf{x}_j) - \epsilon \sum_{i=1}^{\ell} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{\ell} y_i (\alpha_i - \alpha_i^*) & (2.5) \\ & \text{subject to } \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) \text{ and } \alpha_i, \alpha_i^* \in [0, C]. \end{aligned}$$

라그랑지 승수  $\alpha_i$ 와  $\alpha_i^*$ 는 위의 이차함수의 해를 나타낸다. 식 (2.5)에서  $\alpha_i$ 와  $\alpha_i^*$  중 0이 아닌 값이 회귀함수의 예측에 사용되는데 이것을 서포트 벡터(support vector)라 한다. 식 (2.5)을 풀면 식 (2.1)의 서포트 벡터 회귀 추정함수는 다음과 같은 해가 얻어진다.

$$f(\mathbf{x}_i) = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b, \quad (2.6)$$

여기서  $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ 는 커널함수이다 (Li 등, 2007).

### 3. SVR을 이용한 잡음제거

#### 3.1. 최적의 모수 선택

SVR의 성능에 영향을 미치는 모수로는 페널티 모수  $C$ , 오차  $\epsilon$  그리고 커널함수  $K(\mathbf{x}_i, \mathbf{x}_j)$  등이 있다. 여기서  $K(\mathbf{x}_i, \mathbf{x}_j)$ 는 흔히 많이 사용하는 다음과 같이 정의되는 RBF(radial basis function)를 사용한다 (Li, 2009; Li 등, 2007).

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}}.$$

본 논문에서  $C$ 와  $\epsilon$ 은 8-조각 교차 타당성에 의한 그리드 탐색법을 사용하여 선택한다. 8-조각 교차 타당성은 단계별로 설명하면 다음과 같다.

- 단계 1. 주어진  $256 \times 256$  영상을  $32 \times 256$  크기의 8조각으로 나눈다.
- 단계 2. 8조각 중에서 임의의 1조각을 제외한 나머지 조각으로 훈련영상을 만들고 그리드 탐색법을 이용하여 SVR 예측모형을 추정한다.
- 단계 3. 단계 2에서 제외된 1조각 영상에 잡음을 추가한다.
- 단계 4. 추정된 예측모형을 단계 3의 영상에 적용하여 잡음 제거된 영상을 얻는다.
- 단계 5. 원 영상과 잡음 제거 영상을 비교하여 다음과 같이 정의되는 평균제곱오차(Mean Square Error; MSE)를 계산한다.

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left( f(i, j) - \hat{f}(i, j) \right)^2,$$

여기서  $M, N$ 는 각각 영상의 가로, 세로 크기이며  $f(i, j), \hat{f}(i, j)$ 는 각각 원영상과 잡음제거 영상의 픽셀값이다.

- 단계 6. 다른 조각에 대해서도 단계 2-5를 반복한다.

- 단계 7. MSE가 최소가 되는  $C, \epsilon$ 을 최적의 모수로 택한다.

위 단계 2에서 사용된 그리드 탐색법을 단계별로 설명하면 다음과 같다.

- 단계 2.1.  $C, \epsilon$ 의 적절한 범위를 선택한다. 여기서는  $C \in [0.01, 40], \epsilon \in [0.075, 0.25]$ 을 사용한다.
- 단계 2.2. 범위를 적절한 개수의 구간으로 나눈다. 여기서는  $C, \epsilon$  각각에 대해 7개의 로그규모 구간으로 나눈다.

$$\begin{aligned} C_1 = 0.01 < C_2 < \dots < C_7 < C_8 = 4.0, & \log_c C_{i+1} - \log_2 C_i = \ell_c \\ \epsilon_1 = 0.075 < \epsilon_2 < \dots < \epsilon_7 < \epsilon_8 = 0.25, & \log_c \epsilon_{i+1} - \log_2 \epsilon_i = \ell_\epsilon \end{aligned}$$

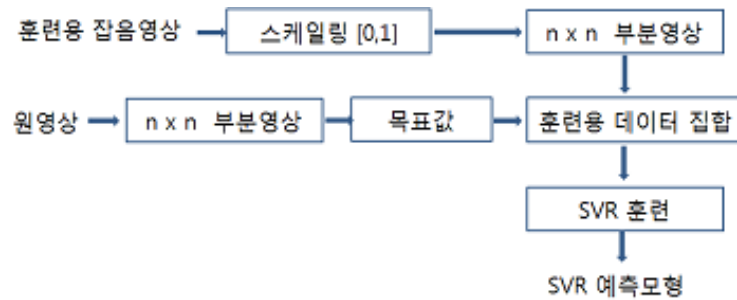


Figure 3.1. Block diagram of SVR training phase

단계 2.3. 각각의  $C_i$ ,  $\epsilon_j$ 에 대해 SVR 예측모형 추정 및 적용 단계를 거쳐 다음과 같이 정의되는 64개 (=  $8 \times 8$ )의 그리드에 대해 PSNR(peak signal-to-noise ratio) 값을 구한다.

$$\text{PSNR} = 10 \log_{10} \frac{\sum_{i=1}^M \sum_{j=1}^N 255^2}{\sum_{i=1}^M \sum_{j=1}^N (f(i, j) - \hat{f}(i, j))^2}.$$

PSNR의 값이 클수록 잡음이 많이 제거됨을 나타낸다. 여기서 실행시간을 줄이기 위해 훈련 영상을 가로로 4픽셀씩 건너뛰어 25%의 픽셀만 이용한다.

단계 2.4. PSNR의 값이 최대일 때의  $C_i$ ,  $\epsilon_j$ 을 중심으로 새로운 구간  $C \in [C_i - 2^{\ell_c/2}, C_i + 2^{\ell_c/2}]$ ,  $\epsilon \in [\epsilon_j - 2^{\ell_c/2}, \epsilon_j + 2^{\ell_c/2}]$ 을 잡는다.

단계 2.5. 단계 2-4를 필요한 정밀도만큼 반복한다. 본 논문에서는 2번 반복하였다.

### 3.2. SVR 예측모형 추정 및 적용

SVR을 이용한 잡음제거는 2단계 즉, 훈련단계와 테스트 단계에 걸쳐 이루어진다. 먼저, 훈련단계에서 SVR 예측모형을 추정하고 테스트 단계에서 잡음영상에 적용하고자 한다.

#### (1) SVR 예측모형 추정

훈련 데이터에 대한 추정된 SVR 예측모형을 얻기 위한 SVR 훈련단계에서 대해 단계별로 설명하고자 한다. Figure 3.1은 SVR 훈련단계에 대한 간략도이다.

- 단계 1. 훈련용 잡음영상을 생성한다. 원영상에 가우시안 잡음이 추가된 잡음영상을 훈련용 잡음영상으로 사용한다.
- 단계 2. 생성된 잡음영상에 대해 [0, 1]로 모든 픽셀값을 스케일링(scaling)한다. 이는 SVR의 계산시간을 줄이기 위해서이다.
- 단계 3. 훈련용 데이터로 구성된 벡터를 만들기 위해  $n \times n$  부분영상을 선택한다. 예를 들어,  $3 \times 3$  부분영상이 선택된다면  $9 \times 1$  벡터가 만들어진다.
- 단계 4. 잡음이 없는 원영상으로부터  $n \times n$  부분영상을 선택한 다음 중심 픽셀값을 목표값으로 설정한다.

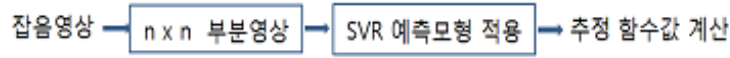


Figure 3.2. Block diagram of SVR test phase

단계 5. 단계 3과 4로부터 훈련용 데이터 집합을 생성한다. 예를 들어,  $3 \times 3$  부분영상이 선택된다면  $\{(\mathbf{x}_i, y_i), i = 1, \dots, \ell\}$ 에서  $9 \times 1$   $\mathbf{x}_i$  벡터가 만들어진다.

단계 6. 영상전체에 대해 생성된 데이터에 대해 식 (2.6)의 추정함수식이 얻어진다. 이 추정함수식이 SVR 테스트 단계에서 예측 모형으로 사용된다.

## (2) SVR 모형 적용

추정된 예측모형을 잡음영상에 적용하는 SVR 테스트 단계에 대해 설명하고자 한다. Figure 3.2는 SVR 테스트 단계에 대한 간략도이다.

단계 1. 잡음제거할 잡음영상을 선택한다.

단계 2. 테스트용 데이터로 구성된 벡터를 만들기 위해  $n \times n$  부분영상을 선택한다. 예를 들어,  $3 \times 3$  부분영상이 선택된다면  $9 \times 1$  벡터가 만들어진다.

단계 3. SVR 훈련단계에서 구한 예측모형에 단계 2에서 구한 테스트용 벡터를 적용한다.

단계 4. 부분영상의 중심픽셀을 단계 3에서 구한 추정값으로 대체한다.

단계 5. 잡음영상전체에 대해 추정 함수값을 계산하기위해 단계 2-4를 반복한다.

## 4. 영상실험 및 분석

본 절에서는 제안된 SVR기반 잡음 제거 방법의 성능을 비교하기 위해 공간필터링의 대표적인 가우시안 방법과 대표적인 웨이블릿 변환 방법인 BayesShrink 방법 그리고 SVR의 Cherkassky-Mulier 방법 (Cherkassky와 Mulier, 1998) 등과 비교하고자 한다. 여기서 Cherkassky-Mulier 방법은 원래 영상 데이터에 SVR을 적용한 연구가 아니고 일반 데이터에 대해 다음과 같은 최적의 모수  $C$ 와  $\epsilon$ 을 제안한 연구이다.

$$C = \max(|\bar{y} + 3\sigma_y|, |\bar{y} - 3\sigma_y|),$$

$$\epsilon = \tau\sigma\sqrt{\frac{\ln n}{n}},$$

여기서  $\bar{y}$ 와  $\sigma_y$ 는 훈련 데이터의 반응값의 평균과 표준편차이고  $n$ 는 훈련 데이터의 개수이고  $\tau$ 는 경험에 의해 얻어지는 조절상수로서  $\tau = 3$ 이 사용된다.

본 논문에서 다루는 가우시안 잡음(Gaussian noise)은 두 가지 잡음 즉,  $\sigma^2 = 0.01$ 인 잡음이 적은 경우와  $\sigma^2 = 0.05$ 인 잡음이 많은 경우를 고려하였다.

훈련용 영상으로는 Figure 4.1에 있는 한글, 영어, 숫자, 기호 등이 포함된 논문 영상이 사용되고 테스트 영상으로는 Figure 4.2와 같이 한글영상으로 휴먼테크 한글 홈페이지 영상, 영어영상으로 한국과학영재 학교 영어 홈페이지 영상, 그리고 한자영상으로 반야심경 영상이 사용되었다.

제안된 SVR 방법에서 최적의  $C$ 와  $\epsilon$ 는 잡음의 양에 따라 다르다. 우리는 제 3절에서 논의한 교차 타당성에 의한 그리드 탐색법을 사용하여  $\sigma^2 = 0.01$ 인 경우  $C = 0.025$ ,  $\epsilon = 0.097$ 을 얻었고  $\sigma^2 = 0.05$ 인 경우  $C = 0.029$ ,  $\epsilon = 0.153$ 을 얻었다.

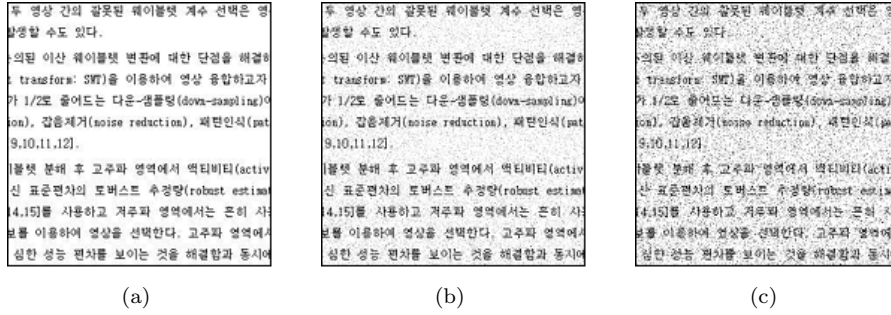


Figure 4.1. Training images used for experiments: (a) noise-free image, (b) noisy image ( $\sigma^2 = 0.01$ ), (c) noisy image ( $\sigma^2 = 0.05$ )



Figure 4.2. Test images used for experiments: (a) Korean image, (b) Noisy Korean image ( $\sigma^2 = 0.01$ ), (c) Noisy Korean image ( $\sigma^2 = 0.05$ ), (d) English image, (e) Noisy English image ( $\sigma^2 = 0.01$ ), (f) Noisy English image ( $\sigma^2 = 0.05$ ), (g) Chinese character image, (h) Noisy Chinese character image ( $\sigma^2 = 0.01$ ), (i) Noisy Chinese character image ( $\sigma^2 = 0.05$ )



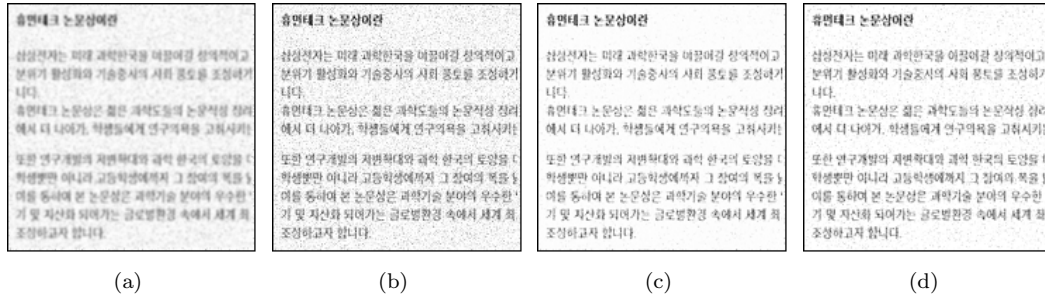


Figure 4.3. Denoised Results for Korean document images( $\sigma^2 = 0.01$ ): (a) Gaussian method, (b) BayesShrink method, (c) Cherkassky-Mulier method, (d) Proposed method

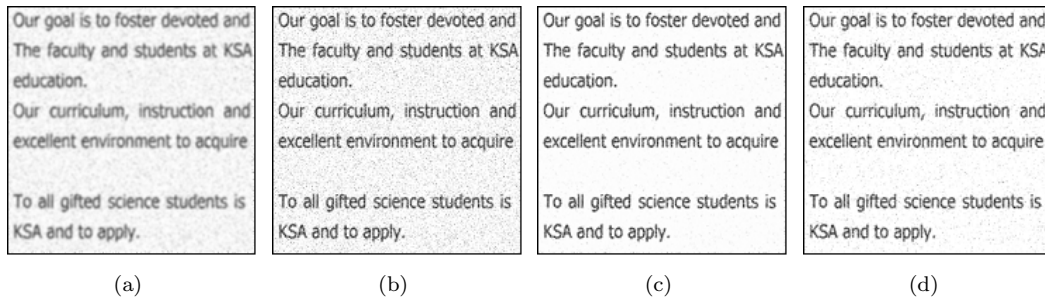


Figure 4.4. Denoised Results for English document images( $\sigma^2 = 0.01$ ): (a) Gaussian method, (b) BayesShrink method, (c) Cherkassky-Mulier method, (d) Proposed method

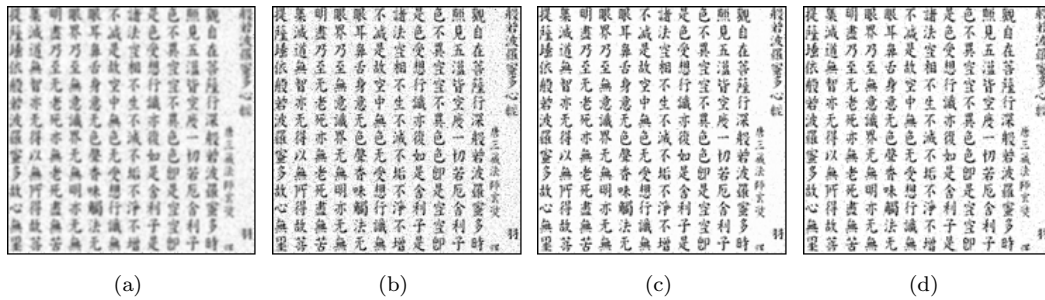


Figure 4.5. Denoised Results for Chinese character document images( $\sigma^2 = 0.01$ ): (a) Gaussian method, (b) BayesShrink method, (c) Cherkassky-Mulier method, (d) Proposed method

최적의  $C$ 와  $\epsilon$ 을 사용한 제안된 방법의 성능을 정량적이고 정성적인 평가로 나누어 비교하고자 한다.

본 영상실험은 VC++ 2008에서 SVR에 대한 라이브러리 LIBSVM (Chang와 Lin, 2011)을 사용하여 실험하였다.

#### 4.1. 정성적인 성능평가

Figure 4.3-4.5는 분산이  $\sigma^2 = 0.01$ 인 한글영상, 영어영상 그리고 한자영상에 대한 여러 가지 잡음제거 방법들의 결과를 보여주고 있다. 위 그림으로부터 가우시안 방법은 열화현상이 일어나서 글자가 희미하게 보임을 알 수 있고 BayesShrink 방법은 열화 현상은 없으나 잡음이 거의 제거되지 않아 남아 있음

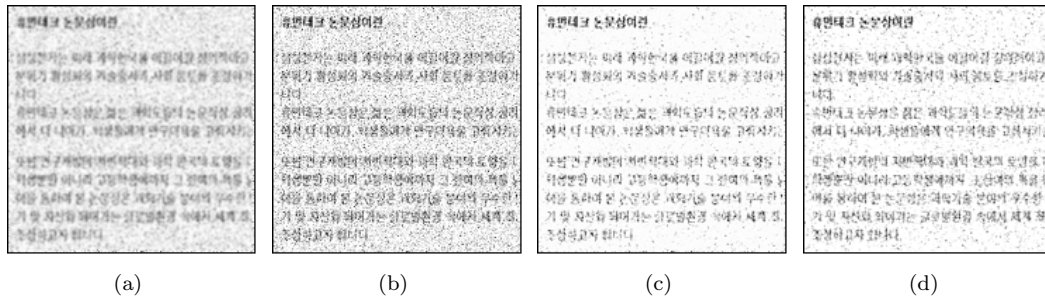


Figure 4.6. Denoised Results for Korean documents( $\sigma^2 = 0.05$ ): (a) Gaussian method, (b) BayesShrink method, (c) Cherkassky-Mulier method, (d) Proposed method

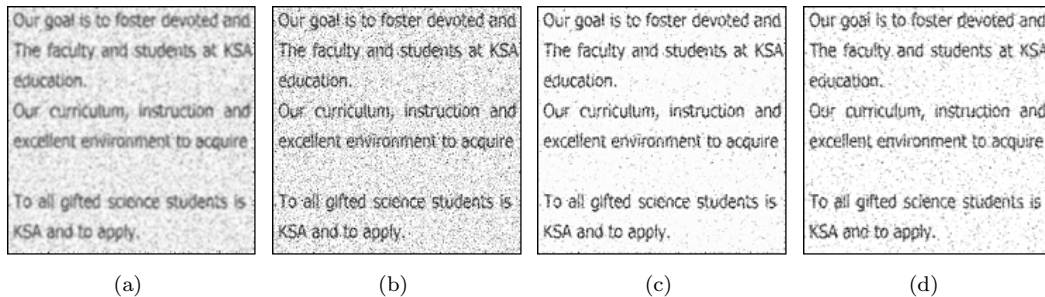


Figure 4.7. Denoised Results for English document images( $\sigma^2 = 0.05$ ): (a) Gaussian method, (b) BayesShrink method, (c) Cherkassky-Mulier method, (d) Proposed method

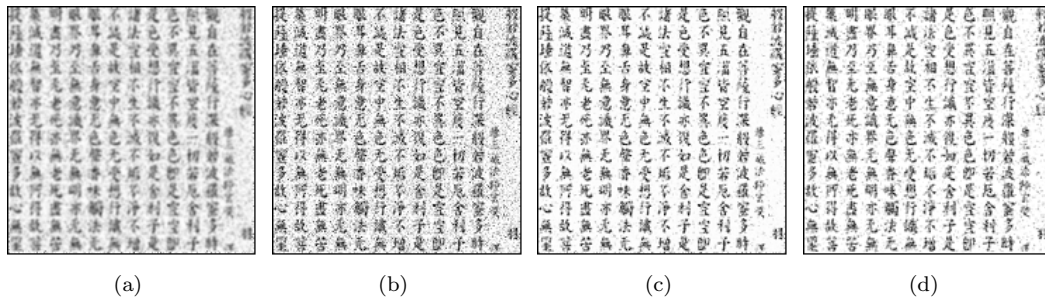


Figure 4.8. Denoised Results for Chinese character document images( $\sigma^2 = 0.05$ ): (a) Gaussian method, (b) BayesShrink method, (c) Cherkassky-Mulier method, (d) Proposed method

알 수 있다. 이처럼 웨이블릿 변환 방법이 문서영상에 효과적이지 못한 이유는 문서영상은 자연영상과는 다르게 희소표현을 갖고 있지 않기 때문이라고 생각한다. SVR에 기반한 Cherkassky-Mulier 방법과 제안된 방법은 열화현상없이 잡음이 잘 제거됨을 알 수 있다. 위의 두 가지 방법 간의 비교는 제안된 방법은 Cherkassky-Mulier 방법과 예지 보전 측면에서 더 나은 것처럼 보인다.

Figure 4.6-4.8은 분산이  $\sigma^2 = 0.05$ 인 한글영상, 영어영상 그리고 한자영상에 대한 여러 가지 잡음제거 방법들의 결과를 보여주고 있다.  $\sigma^2 = 0.01$ 인 영상에서 얻어진 결과보다 많은 잡음에 의해 열화현상이 심하고 문자인식이 어려움이 있음을 알 수 있다. 그러나 제안된 방법은 다른 방법들보다 비교적 잡음 제거력이 좋음을 알 수 있다. 특히, Figure 4.7(d)과 Figure 4.8(d)의 결과 영상은 Figure 4.2(d)와 Figure 4.2(g)의 원영상과 거의 비슷할 정도로 성능이 좋음을 알 수 있다.

**Table 4.1.** PSNR of several denoising methods ( $\sigma^2 = 0.01$ )

잡음제거 방법	잡음 영상		
	한글 영상	영어 영상	한자 영상
가우시안 방법	21.9292	21.0315	19.7613
BayesShrink 방법	21.5984	21.7078	21.3808
Cherkassky-Mulier 방법	25.4736	27.2504	24.4739
제안된 방법	25.6137	26.6635	24.6790

**Table 4.2.** PSNR of several denoising methods ( $\sigma^2 = 0.05$ )

잡음제거 방법	잡음 영상		
	한글 영상	영어 영상	한자 영상
가우시안 방법	19.1127	18.4036	17.8937
BayesShrink 방법	15.1404	15.1084	15.0401
Cherkassky-Mulier 방법	21.1337	21.9220	19.6092
제안된 방법	21.2213	22.0134	20.1342

#### 4.2. 정량적인 성능평가

Table 4.1은 분산이  $\sigma^2 = 0.01$ 인 잡음영상에서 여러 가지 잡음제거 방법들에 대한 PSNR 값을 보여주고 있다.

SVR에 기반한 Cherkassky-Mulier 방법과 제안한 방법은 다른 두 가지의 방법 즉, 가우시안 방법과 BayesShrink 방법보다 높은 PSNR 값을 갖고 있음을 알 수 있다. 그리고 Cherkassky-Mulier 방법과 제안한 방법 비교에서 제안된 방법은 고려되는 모든 영상에서 Cherkassky-Mulier 방법보다 높은 PSNR 값을 갖고 있음을 알 수 있다. 따라서, Cherkassky-Mulier 방법은 영상에서 잡음제거하는데는 최적의 아님을 알 수 있다.

Table 4.2는 분산이  $\sigma^2 = 0.05$ 인 잡음영상에서 여러 가지 잡음제거 방법들에 대한 PSNR 값을 보여주고 있다.

$\sigma^2 = 0.05$ 인 잡음이 많은 경우에도 제안된 방법은 다른 방법들보다 좋은 성능을 갖고 있음을 알 수 있다.  $\sigma^2 = 0.01$ 인 경우에 얻어진 Table 4.1과 비교에서 BayesShrink 방법의 성능이 다른 방법들 보다 현저히 떨어짐을 알 수 있다.

### 5. 결론 및 향후 연구

오늘날 점점 늘어나는 대량의 문서를 자동으로 처리하기 위해 종이문서를 직접 손으로 입력하지 않고 스캐너나 디지털 카메라와 같은 광학기기를 이용하여 얻어진 문서영상은 다양한 분야에서 활용되고 있다.

문서영상에서 잡음제거와 같은 전처리 기술은 시스템의 성능을 좌우하는 매우 중요한 기술이다.

본 논문에서는 문서영상에서 SVR 훈련단계와 테스트 단계를 거쳐 잡음을 제거하였다. SVR 훈련단계에서  $k$ -조각 교차 타당성에 의한 그리드 탐색법에 의해 최적의 SVR 예측모형을 찾고 테스트 단계에서 예측모형을 이용하여 잡음을 제거하였다.

제안된 방법의 성능을 평가하기 위해 3가지 영상 즉, 한글영상, 영어영상 그리고 한자영상에서 잡음이 많고 적음을 고려하였고 비교대상으로 가우시안 방법, BayesShrink 방법 그리고 Cherkassky-Mulier 방법을 고려하였고 가시적인 방법으로 눈으로 평가하는 정성적인 평가와 PSNR 측도를 가지고 평가하는 정량적인 평가로 나누어 평가하였다.

영상실험결과 SVR에 기반한 Cherkassky-Mulier 방법과 제안된 방법이 다른 방법들 보다 열화현상없이 잡음이 잘 제거됨을 알 수 있었다. 그리고 Cherkassky-Mulier 방법과 제안된 방법 비교에서 제안된 방법이 문자의 에지 보전력이 뛰어나고 높은 PSNR 값을 갖고 있음을 알 수 있었다.

그러나 제안된 방법은 예측모형을 찾기위해 교차 타당성에 의한 그리드 탐색법을 사용함으로써 그리드 탐색점이 증가하는 경우 계산시간이 많이 걸리는 단점이 있다. 따라서 향후 연구에서는 빠른 그리드 탐색법 개발 및 해석적인 방법에 의한 최적의 모수 선택이 요구된다.

## References

- Chang, C. C. and Lin, C. J. (2011). LIBSVM: A library for support vector machines, URL: [www.csie.ntu.edu.tw/~cjlin/libsvm/](http://www.csie.ntu.edu.tw/~cjlin/libsvm/)
- Cherkassky, V. and Mulier, F. (1998). *Learning from Data: Concepts Theory and Methods*, Wiley, New York.
- Don, H. S. (2001). A noise attribute thresholding method for document image binarization, *International Journal on Document Analysis and Recognition*, **4**, 131–138.
- Donoho, D. L. (1995). De-noising by soft thresholding, *IEEE Transactions on Information Theory*, **41**, 613–627.
- Donoho, D. L. and Johnstone, I. M. (1994). Spatial adaptation by wavelet shrinkage, *Biometrika*, **81**, 425–455.
- Gonzales, R. C. and Woods, R. E. (1992). *Digital Image Processing*, Addison-Wesley Publishing Company.
- Li, A., Mersereau, R. and Simske, S. (2007). Blind image deconvolution through support vector regression, *IEEE Transactions on Neural Networks*, 931–935.
- Li, D. (2009). Support vector regression based image denoising, *Image and Vision Computing*, **27**, 623–627.
- Lin, T. C. and Yu, P. T. (2004). Adaptive two-pass median filter based on support vector machines for image restoration, *Neural Computation*, **16**, 192–206.
- Liu, H., Sun, F. and Sun, Z. (2006). Image filtering using support vector machine, *Lecture Notes in Computer Science*, **3972**, 533–538.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*, Springer-Verlag.
- Vetterli, M., Chang, S. G. and Yu, B. (2000). Adaptive wavelet thresholding for image denoising and compression, *IEEE Transactions on Image Processing*, **9**, 1532–1546.
- Wang, X. Y. and Fu, Z. K. (2010). A wavelet-based image denoising using least squares support vector machine, *Engineering Applications of Artificial Intelligence*, 862–871.
- Wang, X. Y., Yang, H. Y. and Fu, Z. K. (2010). A new wavelet-based image denoising using undecimated discrete wavelet transform and least squares support vector machine, *Expert Systems with Applications*, **37**, 7040–7049.