IMAGE DENOISING THROUGH SUPPORT VECTOR REGRESSION

Dalong Li, Steven Simske

Digital Printing and Imaging Lab
Hewlett-Packard Laboratories
Fort Collins, CO 80528
{dalong.li,steven.simske}@hp.com

Russell M. Mersereau

Center for Signal and Image Processing
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

In this paper, an example-based image denoising algorithm is introduced. Image denoising is formulated as a regression problem, which is then solved using support vector regression (SVR). Using noisy images as training sets, SVR models are developed. The models can then be used to denoise different images corrupted by random noise at different levels. Initial experiments show that SVR can achieve a higher peak signal-to-noise ratio (PSNR) than the multiple wavelet domain Besov ball projection method on document images.

Index Terms— image denoising, support vector regression, wavelet, PSNR

1. INTRODUCTION

Denoising is an important historical and current problem in image processing. Considerable research has been done [1, 2]. The wavelet transform-based approach is one of the most effective for image denoising [3, 4, 5, 6, 7] for photographic images. Indeed, denoising is one of the most important applications of wavelets. Promising results have been reported in these references.

Even as simple an operation as thresholding in the wavelet domain can effectively reduce noise while preserving image edges. For typical photographic images, most of the wavelet coefficients have very small magnitudes, although there are a few large ones that represent important high frequency features of an image such as edges. Since white noise disperses evenly among all wavelet coefficients, removing small wavelet coefficients reduces most of the noise energy while retaining most of the image energy. This sparseness property is useful in image denoising in order to maintain the sharpness of the edges in an image. The wavelet basis used in image denoising should provide a sparse representation. Recently, multiple wavelet basis image denoising methods [7, 8, 9] have been proposed. These algorithms generally provide better denoising results than conventional wavelet thresholding. In this paper, the multiple wavelet basis Besov ball projections (MWBBP) method [7] is compared with the proposed denoising algorithm.

The support vector regression denoising algorithm is a new procedure that is based on a machine learning approach. We formulate image denoising as a regression problem and use support vector regression in solving the problem. In the training phase, support vector regression (SVR) is trained to learn a mapping from a series of noisy training images to the originals. Then, in the test phase, the trained SVR can perform denoising on images that were not in the training set.

The wavelet characteristics of certain types of images, such as document images, are different from those of natural images that have a sparse representation. Therefore, wavelet-domain denoising on these images is not as efficient as it is on natural images. SVR based image denoising can easily overcome such a limitation simply by including examples of non-natural images (such as document image) in the training set.

This paper is organized as follows. Section 2 presents the proposed algorithm. Comparative experiments are shown in Section 3 and Section 4 contains a brief concluding remark.

2. SUPPORT VECTOR REGRESSION BASED DENOISING

Given training data $(X_1; y_1), \ldots, (X_l; y_l)$, where X_i are input attribute vectors (in noisy image) and y_i are the associated output values (in the original image), traditional linear regression seeks a linear function $W^T X + b$ that minimizes the mean square error:

$$\min_{w,b} \sum_{i=1}^{l} (y_i - (W^T X_i + b))^2. \tag{1}$$

where W is the corresponding weight vector for X and b is the intercept (the constant term). If the input data is not linearly distributed, a linear function is inadequate. Support vector machines introduce a kernel function $\phi(x)$ to map the data into a higher dimensional space, where a linear function is adequate. Commonly used kernel functions are linear, polynomial, Gaussian, sigmoid etc. In the high-dimensional space, overfitting can occur. To limit overfitting, a soft margin and a regularization term are incorporated into the objective function. Support vector regression [10] solves the following modified optimization problem:

$$\min_{W,b,\xi,\xi^*} \frac{1}{2}W^TW + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{2}$$

$$\text{subject to} \quad \begin{aligned} & y_i - (W^T\phi(X_i) + b) \leq \epsilon + \xi_i, \\ & (W^T\phi(X_i) + b) - y_i \leq \epsilon + \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0, i = 1, \ldots, l, \end{aligned}$$

where ξ_i is the upper training error, (ξ_i^* is the lower training error) subject to the $\epsilon-$insensitive tube $|y-(W^T\phi(X)+b)| \leq \epsilon$ and ϵ is a threshold. The cost function ignores any training data that is close (within the $\epsilon-$insensitive tube) to the model prediction. This soft margin method has the advantage of tolerating mislabeled samples in the training set. In this objective function, $\frac{1}{2}W^TW$ is a regularization term that smoothes the function $W^T\phi(X_i) + b$ to limit overfitting. Effectively, within the $\epsilon-$insensitive tube, the regularization term constrains the line to be as flat as possible. The flatness is measured by the norm W^TW.

The application of SVR in image denoising is straightforward. The input vector is formed by the pixels in a window in a noisy image. The target value is the central pixel value in the noise-free image. When the window shifts to a new position, another sample is added to the data set. The size of the window may be interpreted as the denoising filter point spread function (PSF) support. Usually a 3 by 3 window is chosen. A large value is unnecessary and inappropriate since the correlation between pixels decreases as the pixels are spaced further apart. Thus the pixels on the boundary of a large window provide little information about the central pixel in the window. Moreover, a larger window increases the dimension of the feature vector which means increased time for training/testing.

3. EXPERIMENTS

In the experiments, images from the USC-SIPI Database are used [11]. To illustrate the generalization ability of SVR, the training set is intentionally limited to very few images i.e. one or two images. The test images differ from the training images. Since there are no document image in the database, several scanned document images were used in the experiments. One of those images is shown in Fig. 1. In all the experiments, the size of the neighborhood window is 3×3 for the consideration of computational complexity, a larger window means that the length of the input vector X is longer and that will require longer time in training SVR. The training images are LENA, HOUSE, a document image, and combinations of these images. Gaussian noise is added to the training images with a variance of 0.01 except for the document image, where the noise variance was 0.05 so that the texts would be illegible. In the test images, the noise variance varies from 0.01 to 0.04 in steps of 0.01 for the non-document images while the noise variances in the document images varies from 0.05 to 0.08 in steps of 0.01. The noise levels in the test images are different from those in the training images. LibSVM [12], an implementation of SVR, is used in our experiments.

Table 1. PSNR (dB) comparisons of MWBBP and SVR (using two models) based image denoising

Image	Noisy	MWBBP	SVR	
			Lena	Lenahouse
Aerial	16.82	23.61	22.64	23.43
Boat	16.82	24.69	23.87	23.90
Cameraman	17.12	24.14	22.76	22.84
Baboon	16.70	23.10	22.84	22.94
Peppers	16.93	26.20	24.12	25.02
Texture	16.78	20.52	21.74	21.63
Document	14.37	17.79	18.57	19.13
Average	16.50	22.86	22.36	22.70

Table 2. PSNR (dB) comparisons of two SVR models and MWBBP on natural images and document images

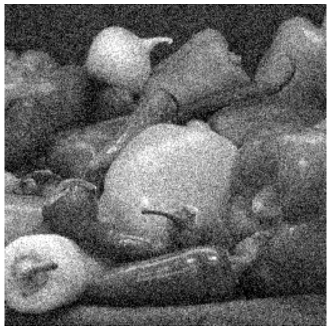
Image	Noisy	MWBBP	SVR		
			Doc	Lena	DocLena
Tiffany	17.77	24.98	23.41	23.25	24.05
Elaine	16.83	24.93	19.26	23.58	21.82
Doc1	14.63	16.86	20.60	16.98	20.35
Doc2	14.50	16.32	19.58	16.07	19.36
Average	15.93	20.77	20.71	19.97	21.40

The peak signal-to-noise ratio (PSNR) can be used to measure the quality of the denoised image, since the original image is available in the simulation. PSNR is defined as:

$$PSNR = 10\log_{10}\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}255^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(f(i,j) - \hat{f}(i,j))^2} \tag{3}$$

where $\hat{f}(i,j)$ is the denoised image, and $f(i,j)$ is the original image. The size of the images is $M \times N$.

Table 1 summarizes the result on some of the test images. For each test image, four noise levels are applied and the averaged PSNR of the four denoised images is computed. There are two SVR models used. One is trained on the LENA image, the other is trained on both the LENA image and the HOUSE image. It can be seen that MWBBP works better on the test images except for the document image and the texture image, whose wavelet domain characteristics are significantly

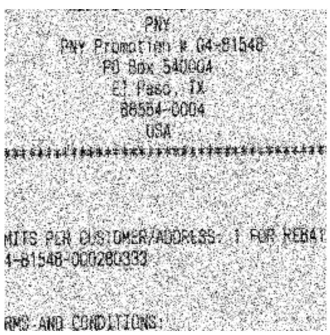


(a)

(b)

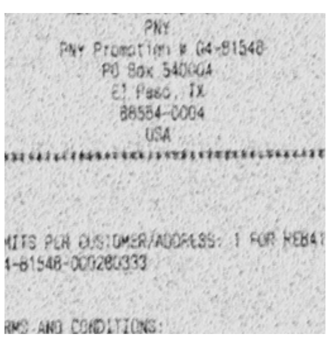
(c)

(d)



(e)

(f)



(g)

(h)

Fig. 1. (a) The noisy PEPPERS image, Gaussian noise variance is 0.01, $PSNR = 20.11$ dB. (b) The denoised image by MWBBP, PSNR = 27.88 dB. (c) The denoised image by Lena SVR model, PSNR = 26.54 dB. (d) The denoised image by doc SVR model, PSNR = 27.92 dB. (e) A noisy document image, PSNR = 12.85 dB. (f) The denoised image by MWBBP, PSNR = 16.43 dB. (g) The denoised image by the LENA SVR model, PSNR = 16.10 dB. (h) The denoised image by doc SVR model, PSNR = 19.67 dB.

different from those of natural images. The results also show that the expanded training set yields a better model for denoising as the PSNR of the images denoised using the Lena-House SVR model are generally higher than those by the SVR trained on the LENA alone. This is not surprising. Although our test dataset was modest, we performed statistical analysis on the difference between MWBBP and the SVR results in Table 1. We first computed the differences between MWBBP and the SVR results as a primary statistic, then performed a z test on the set of δ so acquired (compared to a null hypothesis of zero-mean). For the LENA set, the one-tailed p-value was 0.035; for the LENAHOUSE set, the one-tailed p-value was 0.056. These results support a statistically relevant difference between the MWBBP and SVR results.

To illustrate that a specific model improves the denoising quality, we trained another SVR model from a document image that resembles the test document images. As shown in Table 2, the more specific SVR model resulted in a larger performance difference both in terms of PSNR and visually on the document test images. On the other hand, the document SVR does not work well on natural images. However, the performance on the natural images is improved by using another SVR model that was trained on both the document image and a photographic image (LENA image).

Fig. 1 shows the comparative results on the PEPPERS image and one of the test document images. Though there are no significant differences in the results on the PEPPERS image, the result by the SVR trained on the document image clearly achieved a better denoised document image. In the MWBBP result, there are many visible distortions around the text, especially in the upper part of the image. In the region where there is no text, the SVR result is much cleaner.

4. CONCLUSION

In this paper, we have applied SVR to a new application; namely, image denoising. Simulations show that SVR can learn a generally useful model, even one trained on a very small data set (one or two images). The learned models have been tested on a variety of images (texture, aerial, nature, document). Some initial experiments already suggest that SVR-based image denoising achieves better performance on non-natural images such as document images than wavelet-domain based approaches such as the multiple wavelet basis Besov ball projections method. When the model is trained on a larger data set, as with other machine learning algorithms, the result will generally be improved. A more specific training set can generate a better model that usually can produce better denoising on the images that are similar to those in the training set. This suggests that SVR-based image denoising may need an additional image classification step to compare the input image and the images in the training sets. However, wavelet domain methods do not have such limitations. Moreover, the computational cost associated with SVR is higher than that of the wavelet-domain methods.

5. REFERENCES

[1] O. M. Lysaker, S. Osher, and X. Tai, "Noise removal using smoothed normals and surface fitting," *IEEE Trans. Image Processing*, vol. 13, no. 10, pp. 1345–1357, 2004.

[2] M. Ghazel, G. H. Freeman, and E. R. Vrscay, "Fractal image denoising," *IEEE Trans. Image Processing*, vol. 12, no. 12, pp. 1560–1578, 2003.

[3] A. A. Bharath and J. Ng, "A steerable complex wavelet construction and its application to image denoising," *IEEE Trans. Image Processing*, vol. 14, no. 7, pp. 948–959, 2005.

[4] J. Zhong and R. Ning, "Image denoising based on wavelets and multifractals for singularity detection," *IEEE Trans. Image Processing*, vol. 14, no. 10, pp. 1435–1447, 2005.

[5] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.

[6] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Processing*, vol. 9, no. 9, pp. 1522–1531, 2000.

[7] H. Choi and R. G. Baraniuk, "Multiple wavelet basis image denoising using Besov ball projections," *IEEE Signal Processing Lett.*, vol. 11, no. 9, pp. 717–720, 2004.

[8] P. Ishwar, K. Ratakonda, P. Moulin, and N. Ahuja, "Image denoising using multiple compaction domains," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc. — ICASSP '98*, Seattle, WA, May 1998, pp. 1889–1892.

[9] S. P. Ghael, A. M. Sayeed, and R. G. Baraniuk, "Improved wavelet denoising via empirical Wiener filtering," in *Proc. SPIE Vol. 3169, p. 389-399, Wavelet Applications in Signal and Image Processing V, Akram Aldroubi; Andrew F. Laine; Michael A. Unser; Eds.*, A. Aldroubi, A. F. Laine, and M. A. Unser, Eds., Oct. 1997, pp. 389–399.

[10] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Verlag, New York, 1995.

[11] *University of Southern California - Signal & Image Processing Institute - The USC-SIPI Image Database*, http://sipi.usc.edu/services/database/Database.html.

[12] C. C. Chang and C. J. Lin, *LIBSVM: a library for support vector machines, Software available at*, 2001, http://www.csie.ntu.edu.tw/ cjlin/libsvm.