

An Iterative Super Resolution Algorithm Based On Adaptive FIR Wiener Filtering

Kyle Xiang ZHANG, Yuk-Hee CHAN and Wan-Chi SIU

Centre for Multimedia Signal Processing
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong SAR

Abstract

Aiming at image/video super resolution applications, this paper presents a spatially adaptive FIR wiener filter for super resolution reconstruction. Missing high resolution samples can be estimated as the weighted sum of nearby low resolution samples. In the proposed algorithm, the optimal weighting coefficients for each of nearby low resolution samples are determined with a distance-based correlation model of samples and iteratively refined according to the most updated estimates of the missing high resolution samples. Both objective and subjective measurements in our simulations show that the proposed algorithm can produce a better result as compared with some conventional algorithms across different noise conditions.

I. Introduction

Super Resolution (SR) is one of the major methods to realize spatial resolution up-sampling. It aims to reverse the image formation process and then re-sample the restored signal at a higher sampling rate. SR has been studied over two decades and many algorithms[1-17] have been proposed.

Iterative back projection (IBP) [12] is one of the successful conventional SR methods. It simulates the image formation process to form a low resolution (LR) image from an estimated high resolution (HR) image. The difference between the resultant LR image and the corresponding observed LR image is then used to improve the estimate of the HR image. Accordingly, the HR image can be iteratively restored to minimize the difference between the observed LR images and the simulated ones. Based on IBP, a robust super resolution algorithm (RSR) [13] is further developed. It introduces a median estimator to solve the problem caused by the presence of outliers during the reconstruction process.

Papoulis and Gerchberg's POCS algorithm [14,15] is another successful iterative solution for SR. In this algorithm, the original HR image is interpreted as the solution of a projection function. This function iteratively projects estimated solutions onto several non-empty and closed-convex sets which are understood to be the prior information of the final solution.

Recently, a SR algorithm based on normalized convolution was proposed by Tuan Q. Pham et. al. [16]. This algorithm performs a robust polynomial fit over a local neighborhood based on a structure-adaptive applicability function. Each pixel is assigned a robust certainty value based

on its intensity difference from its neighbors to control its influence in the polynomial fitting process.

Hardie [17] proposed a fast SR algorithm by using an adaptive wiener filter. In particular, a HR sample is estimated to be the weighted sum of its nearby LR samples. The weights are derived based on a simple space-invariant intensity correlation model among pixels, which does not work properly when noisy and edge regions are encountered.

In this paper, an iterative SR algorithm based on adaptive FIR Wiener filtering is proposed. This algorithm iteratively compensates for the cross-correlation estimation error caused by the presence of outliers in a local region. It is generally assumed that samples are gradually less correlated when they are more distant away from each other. However, the presence of noise and edges may break this assumption. By considering this factor, an iterative process is proposed to optimize the correlation estimation and softly minimize the effect of local outliers during the reconstruction process.

The remainder of this paper is organized as follows. In Section II, an observation model is described as the basis of the development of the proposed algorithm. Section III presents a detailed description of the proposed algorithm. Section IV provides some simulation results for performance evaluation. The conclusion is drawn in Section V.

II. Observation Model

The observation model is used to model the image formation process. It can be mathematically described as

$$y = D(b * G(x)) + n \quad (1)$$

where $G(\bullet)$ is a geometrical transformation operator, b is a point spread function (PSF), $D(\bullet)$ is a down-sampling operator and n is an additive random noise.

III. Proposed Algorithm

To start the process, all LR pixels in successive LR images are registered onto a high resolution grid. Through a spatial sliding window covering a $3d \times 3d$ region of the high resolution grid, where d is the ratio of the target resolution of the reconstructed HR image to the original resolution of a LR image, all LR pixels registered in the spatial sliding window are collected to estimate the $d \times d$ HR pixels in the center of the window.

Let the number of the registered LR pixels in the window be L and the total number of the HR pixels to be estimated with

these L LR pixels be $N=d \times d$. The N estimated HR pixels estimated at iteration t are lexicographically ordered to form vector $\hat{x}^{(t)} = [\hat{x}_1^{(t)}, \hat{x}_2^{(t)}, \dots, \hat{x}_N^{(t)}]^T$. Specifically, $\hat{x}_m^{(t)}$ is the estimate of the m^{th} HR pixel obtained by

$$\hat{x}_m^{(t)} = \frac{\sum_{k=1}^L \omega_{mk} \psi_{mk}^{(t)} y_k}{\sum_{k=1}^L \omega_{mk} \psi_{mk}^{(t)}} \quad \text{for } m=1, 2, \dots, N \quad (2)$$

in iteration t , where $\bar{y} = [y_1, y_2, \dots, y_L]^T$ are the registered LR pixels in the window, and $\omega_{mk} \psi_{mk}^{(t)} / \sum_{k=1}^L \omega_{mk} \psi_{mk}^{(t)}$ is the normalized weight of the k^{th} LR pixel when estimating $\hat{x}_m^{(t)}$.

A LR pixel's weight is contributed by $\psi_{mk}^{(t)}$ and ω_{mk} . Specifically, coefficients ω_{mk} are elements of matrix

$$W = \begin{bmatrix} \omega_{11} & \omega_{21} & \cdots & \omega_{N1} \\ \omega_{12} & \omega_{22} & \cdots & \omega_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{1L} & \omega_{2L} & \cdots & \omega_{NL} \end{bmatrix} = R^{-1} \cdot P \quad (3)$$

where R is the auto-correlation matrix of $\{y_k | k=1, \dots, L\}$, and P is the cross-correlation matrix of $\{y_k | k=1, \dots, L\}$ and $\{x_m | m=1, \dots, N\}$.

While ω_{mk} is determined by (i) the correlation among LR pixels and (ii) the correlation between LR pixels and HR pixels, $\psi_{mk}^{(t)}$ is an iteration-variant term used to rectify ω_{mk} . In general, the correlation of two pixels is modeled as a function of their distance[17]. This simple model is space-invariant and does not reflect the real situation in the regions containing noise and edges. Accordingly, ω_{mk} cannot weight y_k properly in these regions. $\psi_{mk}^{(t)}$ is used to improve the weighting effect by rectifying ω_{mk} . Its details will be addressed in the last part of this section.

By considering that the observed LR pixel y_k is corrupted with an additive zero mean random noise n and its noisy free version is $f_k = y_k - n_k$, where n_k is the noise magnitude for sample y_k , we have

$$R = E\{\bar{y} \cdot \bar{y}^T\} = E\{\bar{f} \cdot \bar{f}^T\} + \sigma_n^2 \cdot \mathbf{I} \quad (4)$$

$$\text{and} \quad P = E\{\bar{y} \cdot \bar{x}^T\} = E\{\bar{f} \cdot \bar{x}^T\} \quad (5)$$

where $E\{\bullet\}$ is the expectation operator and σ_n^2 is the local variance of the noise.

In theory, f_k are samples of $\{f(i,j)\}$, a blurred version of the original HR image $\{x(i,j)\}$, and hence we have $f(i,j) = x(i,j) * b(i,j)$, where b is the blurring function and (i,j) denotes a position in the HR grid. Accordingly, the cross-covariance of $x(i,j)$ and $f(i,j)$ and the auto-covariance of $f(i,j)$ can be, respectively, determined by

$$r_{f\hat{x}}(i,j) = r_{xx}(i,j) * b(i,j) \quad (6)$$

$$\text{and} \quad r_{ff}(i,j) = r_{xx}(i,j) * b(i,j) * b(-i,-j) \quad (7)$$

In general, the covariance between two HR pixels can be modeled as

$$r_{xx}(\Delta i, \Delta j) = \sigma_x^2 \cdot \exp(-\gamma (\sqrt{\Delta i^2 + \Delta j^2})^2) \quad (8)$$

where $(\Delta i, \Delta j)$ is the distance of the two HR pixels, γ is a tuning parameter and σ_x^2 is the local variance of the HR image. In particular, σ_x^2 is estimated as

$$\sigma_x^2 = \max\left\{\left(\sigma_y^2 - \sigma_n^2\right) / C, 0\right\} \quad (9)$$

where σ_y^2 is the variance of $\{y_k | k=1, \dots, L\}$ and

$$C = \iint \exp(-\gamma (\sqrt{\Delta i^2 + \Delta j^2})^2) \cdot (b(i,j) * b(-i,-j)) \cdot didj \quad (10)$$

With $r_{xx}(\Delta i, \Delta j)$ on hand, elements of $E\{\bar{f} \cdot \bar{f}^T\}$ and $E\{\bar{f} \cdot \bar{x}^T\}$ in eqns. (4) and (5) and hence W can be determined with $r_{ff}(\Delta i, \Delta j)$ and $r_{f\hat{x}}(\Delta i, \Delta j)$, where $(\Delta i, \Delta j)$ is the distance of the two involved pixels in the HR grid.

$\psi_{mk}^{(t)}$, the parameter used to rectify ω_{mk} , is initialized to be $\psi_{mk}^{(1)} = 1$, and then iteratively updated based on $\hat{x}_m^{(t-1)}$, the most updated estimate of the m^{th} HR pixel, as follows.

$$\psi_{mk}^{(t)} = \begin{cases} 1 & \text{if } t=1 \\ \exp\left(-\left|y_k - \frac{1}{N} \left(\sum_{m=1}^N \hat{x}_m^{(t-1)}\right)\right| / \sigma_n\right) & \text{if } t > 1 \end{cases} \quad (11)$$

Starting from iteration $t=1$, HR pixels are iteratively estimated until they converge to the final estimates. When $t=1$, we have $\psi_{mk}^{(1)} = 1$ and hence the 1st iteration of the proposed algorithm is equivalent to an adaptive Wiener filtering algorithm (different from the one proposed in [17]). When $t > 1$, the algorithm evaluates the intensity difference between a particular LR pixel and the average of the HR pixel estimates in the local region and, in the current estimation, deemphasizes the contribution of the LR pixels the intensity values of which are very different from the local intensity mean of the estimated HR pixels. This helps to reject those outliers and improve the estimation performance in noisy and edge areas.

IV. Simulations

Simulations were carried out to evaluate the performance of the proposed algorithm. A set of 8-bit gray level aerial images shown in Fig. 1 were used as the original HR images in the simulation. Their LR versions were obtained with the observation model shown in Section II. In particular, each original HR image was blurred with a Gaussian PSF, corrupted with additive zero-mean Gaussian random noise of variance σ_n^2 , and then down-sampled by 2 along each dimension. The resultant LR images were then enlarged to reconstruct the original with various SR algorithms including the proposed one. In other words, each HR image was estimated with 1 LR image and the enlarging ratio d was 2. A spatial sliding

window of size 6×6 was used during the reconstruction when the proposed algorithm was used. Tuning parameter γ was selected to be 0.2. Fig. 2 shows the simulation results of several SR algorithms [12-16] for visual comparison.

Table 1 shows the performance achieved with different SR algorithms in terms of *PSNR* under different noise conditions. In particular, *PSNR* is defined as

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{\frac{1}{N_0} \sum_{(i,j)} (x(i,j) - \hat{x}(i,j))^2} \text{ in dB} \quad (12)$$

where x and \hat{x} are, respectively, the original high resolution image and its estimate, and N_0 is the total number of image pixels involved in the comparison.



Fig.1 Twenty aerial images of size 500×500 each (Named as *Image 1* to *20*, from top-to-bottom and left-to-right)

V. Conclusions

This paper presents a SR adaptive FIR wiener filter with cross-correlation refinement. Super resolution reconstruction normally suffers from the presence of outliers. Outliers corrupt the reconstruction of the high resolution image since they could be over considered. The proposed method helps to minimize the degradation caused by the outliers in a local region during the reconstruction process. Simulations demonstrated the effectiveness of the proposed algorithm. It suppresses random additive noise and produces outputs of better subjective and objective quality as compared with some conventional iterative SR algorithms.

VI. Acknowledgement

This work was supported by Centre for Multimedia Signal Processing, The Hong Kong Polytechnic University (PolyU Grant 1-BB9T).

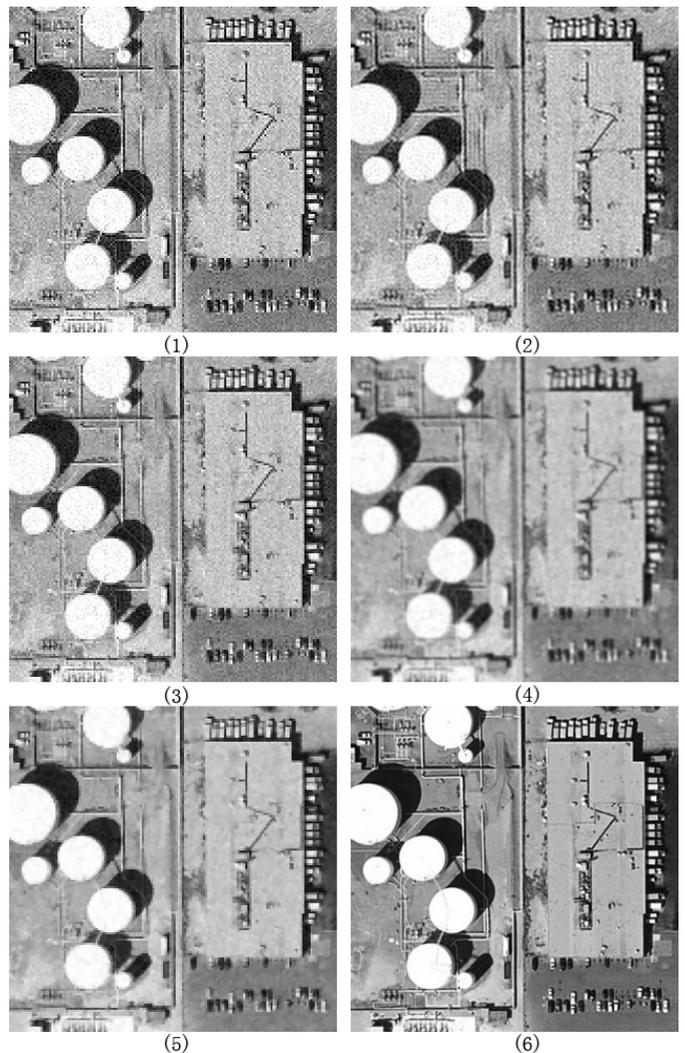


Fig.2 Simulation results of testing image 9 ($\sigma_n^2=100$): (1) IBP[12], (2) POCS[14,15], (3) RSR[13], (4) SANC[16], (5) the proposed and (6) the original.

References

- [1] M. Irani and S. Peleg, "Improving resolution by image registration," *CHIP: Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991.
- [2] S.Farsiu, M.D.Robinson, M.Elad, P.Milanfar, "Fast and robust multi-frame super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327 - 1344, Oct. 2004.
- [3] Y.He, K.H.Yap, L.Chen, L.P.Chau, "A Nonlinear Least Square Technique for Simultaneous Image Registration and Super-Resolution," *IEEE Transaction on Image Processing*, vol. 16, no. 11, pp. 2830 - 2841, Nov. 2007.
- [4] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E.A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Opt. Eng.*, Vol. 37, No. 1, pp. 247–260, Jan. 1998.

Methods	$\sigma_n^2=50$						$\sigma_n^2=100$					
	[12]	[13]	[14,15]	[16]	[17]	Proposed	[12]	[13]	[14,15]	[16]	[17]	Proposed
<i>Image.1</i>	23.12	24.54	24.18	23.69	24.41	24.89	21.83	23.52	23.40	23.49	24.16	24.48
<i>Image.2</i>	22.64	24.03	23.47	23.26	23.91	24.30	21.40	23.14	22.81	23.08	23.61	24.02
<i>Image.3</i>	24.87	26.28	26.88	26.94	27.51	27.75	23.03	25.13	25.54	26.54	27.02	27.14
<i>Image.4</i>	20.32	22.00	21.56	21.30	21.62	22.16	19.53	21.31	21.12	21.20	21.49	21.95
<i>Image.5</i>	22.64	24.08	23.69	23.42	24.01	24.45	21.54	23.18	23.03	23.23	23.82	24.15
<i>Image.6</i>	24.46	26.21	26.03	25.87	27.11	27.67	22.74	24.75	24.86	25.58	26.81	27.28
<i>Image.7</i>	20.25	21.22	20.94	21.44	21.90	22.15	19.44	20.45	20.54	21.32	21.70	21.91
<i>Image.8</i>	22.17	24.03	23.99	23.22	24.03	24.29	21.08	23.15	23.27	23.04	23.77	24.06
<i>Image.9</i>	21.85	23.42	22.96	22.53	23.14	23.72	20.87	22.61	22.38	22.37	22.91	23.49
<i>Image.10</i>	22.00	23.36	22.62	22.35	22.90	23.51	21.08	22.60	22.07	22.21	22.68	23.24
<i>Image.11</i>	21.08	22.26	21.52	21.05	21.71	22.29	20.28	21.68	21.11	20.96	21.50	22.09
<i>Image.12</i>	23.16	24.78	24.94	24.03	24.91	25.19	22.09	23.85	24.11	23.84	24.57	24.88
<i>Image.13</i>	23.05	25.08	24.95	24.39	25.13	25.60	21.80	23.93	23.99	24.17	24.77	25.19
<i>Image.14</i>	23.15	24.82	24.73	24.18	24.92	25.24	21.94	23.82	23.92	23.98	24.66	24.92
<i>Image.15</i>	22.88	24.38	23.70	23.46	24.21	24.85	21.66	23.40	23.07	23.27	23.87	24.49
<i>Image.16</i>	22.22	23.43	22.99	22.84	23.40	24.00	21.36	22.85	22.43	22.66	23.21	23.70
<i>Image.17</i>	22.15	23.39	22.54	22.57	23.04	23.63	21.19	22.60	22.02	22.41	22.84	23.34
<i>Image.18</i>	23.75	25.31	24.57	24.82	25.53	26.03	22.50	24.20	23.80	24.58	25.22	25.65
<i>Image.19</i>	21.61	22.94	22.22	22.11	22.62	23.06	20.65	22.22	21.72	21.97	22.40	22.78
<i>Image.20</i>	21.51	22.76	22.04	21.74	22.34	22.89	20.73	22.13	21.59	21.62	22.14	22.65
Average	22.44	23.92	23.53	23.26	23.92	24.38	21.34	23.03	22.84	23.08	23.66	24.07

Table.1 PSNR performance of various SR algorithms including IBP[12], RSR[13], POCS[14,15], SANC[16], AWF[17] and the proposed under different noise conditions

[5] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," IEEE Trans. Image Process., Vol. 6, No. 6, pp. 1–14, Jun. 2004.

[6] B.Narayanan, R.C.Hardie, K.E.Barner, M.Shao, "A Computationally Efficient Super-Resolution Algorithm for Video Processing Using Partition Filters," IEEE Transaction on Circuit and System for Video Technology, Vol.17, No.5, pp.621-634, May 2007

[7] N. Nguyen and P. Milanfar, "A wavelet-based interpolation restoration method for superresolution," Circuits, Syst., Signal Process., Vol. 19, No. 4, pp. 321–338, Aug. 2000.

[8] H.F.Shen, L.P.Zhang, B.Huang and P.X.Li, "A MAP Approach for Joint Motion Estimation, Segmentation and Super Resolution," IEEE Transaction on Image Processing, Vol.16, No.2, pp.479-490, February 2007

[9] T. R. Tuinstra and R. C. Hardie, "High resolution image reconstruction from digital video by exploitation on non-global motion," Opt. Eng., Vol. 38, No. 5, May 1999.

[10] H.L.Wu and J.T.Zhang, "Non-parametric Regression Methods for Longitudinal Data Analysis," Wiley series in probability and statistics, 2006

[11] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacement," in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, pp. 742-746, June 1988.

[12] M. Irani and S. Peleg, Improving resolution by image registration, CVGIP: Graphical Models and Image Proc., Vol. 53, pp. 231-239, May 1991.

[13] A. Zomet, A. Ravacha and S. Peleg, Robust super resolution, in Proceedings of the Int. Conf. on Computer Vision and Pattern Recognition, Vol. 1, Dec. 2001, pp. 645–650.

[14] A. Papoulis, A new algorithm in spectral analysis and band-limited extrapolation, Circuits and Systems, IEEE Transactions on , Vol.22, No.9, pp. 735-742, Sep 1975

[15] R. Gerchberg, "Super-resolution through error energy reduction," Optical Acta, 21, pp.709–720, 1974.

[16] T. Q. Pham, L. J. van Vliet, and K. Schutte, "Robust fusion of irregularly sampled data using adaptive normalized convolution," EURASIP Journal on Applied Signal Processing, 2005.

[17] R.C.Hardie, "A Fast Image Supre-Resolution Algorithm Using an Adaptive Wiener Filter," IEEE Transaction on Image Processing, Vol.16, No.12, pp.2953-2964, December 2007