Wavelet-Based Super-Resolution Reconstruction: Theory and Algorithm

Hui Ji and Cornelia Fermüller

Center for Automation Research, Department of Computer Science,
University of Maryland, College Park
{jihui, fer}@cfar.umd.edu

Abstract. We present a theoretical analysis and a new algorithm for the problem of super-resolution imaging: the reconstruction of HR (high-resolution) images from a sequence of LR (low-resolution) images. Super-resolution imaging entails solutions to two problems. One is the alignment of image frames. The other is the reconstruction of a HR image from multiple aligned LR images. Our analysis of the latter problem reveals insights into the theoretical limits of super-resolution reconstruction. We find that at best we can reconstruct a HR image blurred by a specific low-pass filter. Based on the analysis we present a new wavelet-based iterative reconstruction algorithm which is very robust to noise. Furthermore, it has a computationally efficient built-in denoising scheme with a nearly optimal risk bound. Roughly speaking, our method could be described as a better-conditioned iterative back-projection scheme with a fast and optimal regularization criteria in each iteration step. Experiments with both simulated and real data demonstrate that our approach has significantly better performance than existing super-resolution methods. It has the ability to remove even large amounts of mixed noise without creating smoothing artifacts.

1 Introduction

The problem of obtaining a *super-resolution* image from a sequence of low-resolution images has been studied by many researchers in recent years. Most super-resolution algorithms formulate the problem as a signal reconstruction problem. Essentially these algorithms differ in two aspects: one is in how the images of the sequence are aligned; the other is in how the high-resolution image is reconstructed from the aligned image frames. Both issues are critical for the success of the super-resolution reconstruction. In this paper we take a flow-based approach to image alignment ([1, 2, 3, 4]). The focus of the paper is on the later problem (Reconstructing HR from aligned LR images).

Iterative back-projection methods ([1, 5]) have been shown to be effective for high-resolution image reconstruction. It is known, however, that the deblurring process, which is part of this approach, makes it very sensitive to the noise. Thus, the requirement of very accurate image alignment estimates limits its practical use. Various regularization methods have been proposed to deal with the noise issue. However, these methods either are very sensitive to the assumed

noise model (Tikhonov regularization) or are computationally expensive (Total-Variation regularization). See [6] for more details.

Our contributions in this paper are two-fold. First, we model the image formation procedure from the point of view of filter bank theory. Then based on this new formulation, we provide an analysis of the limits of the high-resolution reconstruction. The conclusion is that in general full recovery is not possible without enforcing some constraints on the recovered images. At best we could reconstruct the image convolved with a specific low-pass filter (namely $\frac{1}{4}(1,1) \otimes (1,1)$ for the case of the Box-type PSF).

Second, based on our new formulation, we present a robust wavelet-based algorithm to reconstruct the image. The iteration scheme in our algorithm is inherently more robust to noise than that of classic back-projection methods ([1, 5]), since the projection matrix of our new back-projection scheme has a better condition number. We will show that, both in theory and experiments, it has better performance in suppressing the error propagation than other back-projection iteration schemes.

Furthermore, our algorithm allows us to include a wavelet-based denoising scheme in each iteration of the reconstruction which effectively removes the noise without creating smoothing artifacts. The advantage of our denoising scheme over regularization methods is that it is nearly optimal with respect to the risk bound. That is, it has the theoretical minimal error in removing noises of unknown models. Its effectiveness in removing mixed noises and relatively large amounts of noise is demonstrated in experiments. It is worth mentioning that our denoising scheme adds very little computational burden compared to other complicated regularization methods. Briefly, our method could be described as a generalized iterative back-projection method with a fast and optimal regularization criteria in each iteration step.

Wavelet theory has previously been used for image denoising and deblurring from static images ([7, 9]). However, it has not been studied much with respect to the super-resolution problem. In recent work wavelet theory has been applied to this problem in this sector i [10], but only for the purpose of speeding up the computation. Our contribution lies in an analysis that reveals the relationship between the inherent structure of super-resolution reconstruction and the theory of wavelet filter banks. This relationship is fully exploited by using various techniques from wavelet theory in the iterations of the reconstruction.

2   Analysis of Super-Resolution Reconstruction

2.1   Formulation of High-to-Low Image Formation

We first formulate the high-to-low image formation process. To simplify the exposition, in the following we only discuss 1D signals with resolution enhancement by a factor 2. Later, without much difficulty, the analysis will be extended to the 2D case with arbitrary resolution increase. Using Farsiu's notation ([6]), the image formation process in the pixel domain can be modeled as

$$y = \sigma[H * X(F(t))] + N, \tag{1}$$

where t is the spatial variable, $X(t)$ is the continuous signal and y is the discrete signal. H is the blurring operator (either optical blurring or motion blurring or both), F is the geometric transform, N is the noise in the low-resolution image, σ is the decimation operator, and "$*$" is the convolution operator. Not considering the noise, the high-resolution (HR) signal x and low-resolution (LR) signal y can be defined as:

$$x = \sigma[X], \quad y = [\sigma[H * X(F(\cdot))]] \downarrow_2 . \tag{2}$$

where $\downarrow_2$ is the downsampling operator with rate 2.

Next we derive the relation between the LR signal y and the HR signal x. Define the difference $E(t) = X(F(t)) - X(t)$, which is also called the *optical flow*. For the simplicity of notation, here we assume a constant flow model $E(t) = \epsilon$ with $0 \le \epsilon < 2$ on the denser grid of the HR image x. Thus, in the LR image the flow is a sub-pixel shift (Recall a 1-unit shift on the coarse grid of y equals a 2-unit shift on the fine grid of x). For the case of $0 \le \epsilon < 1$, the first-order Taylor approximation of Equation (2) can be written as

$$\begin{aligned} y &= [\sigma[H * (X(t + \epsilon))] \downarrow_2 = [\sigma[H * X + H * (\epsilon X')]] \downarrow_2 \\ &= [\sigma[H * X + \epsilon(H * X')]] \downarrow_2 = [\sigma[H * X + \epsilon(H' * X)]] \downarrow_2 \\ &= [\sigma[H * X]] \downarrow_2 + \epsilon[\sigma[H' * X]] \downarrow_2, \end{aligned}$$

The expression above in the pixel domain is then

$$y = [a * x] \downarrow_2 + \epsilon[b * x] \downarrow_2, \tag{3}$$

where a, b are discrete versions of the convolution kernels H and H' respectively. For the case of $1 \le \epsilon < 2$, a similar argument yields

$$y = [a * x(\cdot + 1)] \downarrow_2 + (\epsilon - 1)[b * x(\cdot + 1)] \downarrow_2 . \tag{4}$$

Thus for any LR signal y_k with general optical flow $E_k(t)$, we have the approximation of y_k as either in the form of:

$$y_k = [a * x] \downarrow_2 + \epsilon_k \cdot*[b * x] \downarrow_2$$

or in the form of:

$$y_k = [a * x(\cdot + 1)] \downarrow_2 + (\epsilon_k - 1) \cdot*[b * x(\cdot + 1)] \downarrow_2,$$

where $\cdot*$ denotes the component-wise multiplication operator (to denote general optic flow), and ϵ_k denotes the discrete sample of the optical flow $E_k(t)$. Having available the optical flow values ϵ_k for multiple low-resolution images y_k, we can extract the four components:

$$\begin{aligned} &[a * x] \downarrow_2, \quad [a * x(\cdot + 1)] \downarrow_2 \\ &[b * x] \downarrow_2, \quad [b * x(\cdot + 1)] \downarrow_2 . \end{aligned} \tag{5}$$

As will be shown in the next subsection, the two filters a and b (which are determined by the blurring kernel H and its derivative H') characterize the super-resolution reconstruction.

Let us next look at some examples of filters a and b for different blurring kernels.

Example 1. Consider the box-type blurring kernel $H = \frac{1}{2n}\chi_{[-n,n]}$. Let $E(t) = \epsilon \leq 1$. Then we have

$$y(j) = \int_{-\infty}^{\infty} \chi(2j - t)X(F(t))dt = \frac{1}{2n}\int_{2j-n}^{2j+n} X(F(t))dt = \frac{1}{2n}\int_{2j-n}^{2j+n} X(t+\epsilon)dt.$$

Approximating the integration by quadrature rules, we obtain

$$y(j) = \frac{1}{2n}(\frac{1}{2}(1-\epsilon)x(2j-n) + \sum_{i=-n+1}^{n-1} x(2j-i) + \frac{1}{2}(1+\epsilon)x(2j+n)).$$

Or equivalently,

$$y = [a * x + \epsilon(b * x)] \downarrow_2, \tag{6}$$

where a and b are the following low-pass and high-pass filters respectively:

$$a = \frac{1}{4n}(1, 2, \cdots, 2, 1), \quad b = \frac{1}{4n}(-1, 0, \cdots, 0, 1).$$

Example 2. Consider a Gaussian-type blurring kernel H. Using the Cubic Cardinal B-spline $B(t)$ as approximation to the Gaussian function we have

$$y(j) = \int_{-\infty}^{\infty} B(2j - t)X(F(t))dt.$$

Again, by the quadrature rule, we have the approximation

$$y = \sum_i x(2j - i)(a(i) - \epsilon b(i)),$$

where $a = \frac{1}{96}(1, 8, 23, 32, 23, 8, 1), \quad b = \frac{1}{48}(3, 12, 15, 0, -15, -12, -3)$.

2.2   Analysis of the HR Reconstruction

Given multiple LR signals y_k with different motions ϵ_k, theoretically we can obtain two complete sequences $a * x$ and $b * x$. An interesting question arises. Without any assumption on the finite signal x, can we theoretically reconstruct the signal sequence x from these two sequences $a * x$ and $b * x$?

To answer this question, let us express the sequences in another form. Let us write the Z-transform of a signal sequence $x = \{x(i)\}$ as

$$x(z) = \sum_i x(i)z^{-i}.$$

Then with $a(z)$ and $b(z)$ being the Z-transforms of the filters a and b, the Z-transforms of $a * x$ and $b * x$ are $a(z)x(z)$ and $b(z)x(z)$ respectively. Now the question can be answered by investigating whether the polynomial equation

$$(a(z)x(z))u(z) + (b(z)x(z))v(z) = x(z), \tag{7}$$

is solvable for the two unknowns $u(z)$ and $v(z)$. Eliminating $x(z)$ from both sides of the equation (7) yields

$$a(z)u(z) + b(z)v(z) = 1. \tag{8}$$

From the theory of Diophantine equation we know that

Lemma 1. *Given two polynomials $a(z)$ and $b(z)$, Equation (8) is solvable if and only if the greatest common divisor of $a(z)$ and $b(z)$ is a scalar, that is, $a(z)$ and $b(z)$ are co-prime.*

It is observed that $a(z)$ and $b(z)$ in our two examples (Example 1 and 2): both have a common divisor

$$c(z) = (1 + z),$$

which can be seen from the fact that $a(-1) = b(-1) = 0$ and thus $z = -1$ is the root of both $a(z)$ and $b(z)$. Thus for these blurring kernels we cannot perfectly reconstruct $x(z)$ from $a(z)x(z)$ and $b(z)x(z)$. This conclusion holds true not only for our examples, but also for general blurring kernels, as we will show next.

We follow Baker's modeling of the blurring kernel H ([11]). The blurring kernel (Point spread function) is decomposed into two components:

$$H = \Omega * C,$$

where $\Omega(X)$ models the blurring caused by the optics and $C(X)$ models the spatial integration performed by the CCD sensor. Typically Ω is modeled by a Gaussian-type function and C is modeled by a Box-type function. Notice that

$$H' = \Omega' * C.$$

Thus we can express the corresponding discrete filters as:

$$a = \ell * c; \quad b = \tau * c,$$

where c is the discrete version of the spatial integration kernel C, and ℓ and τ are the discrete versions of H and H'. Since $a(z)$ and $b(z)$ have a common divisor $c(z)$, we cannot reconstruct $x(z)$ for general $x(z)$, unless C is a Dirac function, which generally is not true. Based on Lemma 1, we then have the following claim.

Claim 1. *Given multiple LR finite signals y_k, we can not perfectly reconstruct the HR finite signal x without any assumptions on x. At most we can reconstruct $c * x$ for some low-pass filter c. The corresponding Z-transform $c(z)$ of c is the greatest common divisor of $a(z)$ and $b(z)$, which includes the spatial integration filter.*

Notice that c is a low-pass FIR (finite impulse response) filter. To recover x from $c * x$, we have to apply a high-pass filter on $c * x$ and impose some boundary condition on the signal x. Such a deblurring process generally is sensitive to the noise. A reasonable strategy then is to modify our reconstruction goal during the intermediate iterative reconstruction process. Instead of trying to reconstruct x,

we reconstruct $c * x$ in the iterative process, and we leave the recovery of x from $c * x$ to the last step after finishing the iterative reconstruction.

Thus the modified HR signal to be reconstructed becomes $\widetilde{x} = c * x$. The corresponding equation for y_k and $\widetilde{x}$ is then:

$$y_k = [\ell * \widetilde{x}] \downarrow_2 + \epsilon_k \cdot [\tau * \widetilde{x}] \downarrow_2 \qquad (9)$$

or

$$y_k = [\ell * \widetilde{x}(\cdot + 1)] \downarrow_2 + (\epsilon_k - 1) \cdot [\tau * \widetilde{x}(\cdot + 1)] \downarrow_2,$$

where the Z-transform of ℓ and τ are $a(z)$ and $b(z)$ divided by their greatest common divisor $c(z)$.

It is worth mentioning that we model the blurring procedure from HR to LR by a first-order Taylor approximation. But our reasoning could easily be extended to the modeling by higher-order Taylor approximations, leading to the same conclusions.

3   Reconstruction Method

3.1   Reconstruction Based on PR Filter Banks

Introduction to PR filter banks. Before presenting our algorithm, we first give a brief introduction to 2-channel PR (*perfect reconstruction*) filter banks (see [12] for more details). A two-channel filter bank consists of two parts: an analysis filter bank and a synthesis filter bank. In our case, the signal $\widetilde{x}$ is first convolved with a low-pass filter ℓ and a high-pass filter h and then subsampled by 2. In other words, we analyze the signal by an analysis filter bank. Then a reconstructed signal $\widehat{x}$ is obtained by upsampling the signal by zero interpolation and then filtering it with a dual low-pass filter g and a dual high-pass filter q. In other words, we reconstruct the signal by synthesizing the output from the analysis bank with a synthesis filter bank. See Fig. 1 for an illustration.
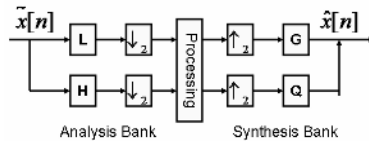

Fig. 1. Two-channel filter bank

Such a filter bank is called a PR filter bank if $\widehat{x} = \widetilde{x}$ for any input $\widetilde{x}$. It is known that (see [12]) the synthesis filters $\{\ell, h\}$ of a perfect reconstruction filter bank have to satisfy the following condition:

$$\ell(z)h(-z) - \ell(-z)h(z) = z^m \text{for some integer } m \qquad (10)$$

and the corresponding synthesis filters amount to

$$g(z) = h(-z); \quad q(z) = -(\ell(-z)).$$

Thus, given any low-pass filter $\ell(z)$, we can find the corresponding high-pass filter $h(z)$ such that we have a PR filter by solving the linear system (10).

Iterative reconstruction scheme. We have available a number of signals y_k and the corresponding estimates of the optic flow values ϵ_k. We also have estimates of the convolution kernels ℓ and τ. Let then ℓ be the low pass filter of our PR filter bank, and we compute the corresponding h (Note h may be different from τ).

Recall that for each LR signal y_k, we have

$$y_k = [\ell * \widetilde{x}] \downarrow_2 + \epsilon_k \cdot * [\tau * \widetilde{x}] \downarrow_2 .$$

Thus $[\ell * \widetilde{x}] \downarrow_2$ amounts to

$$[\ell * \widetilde{x}] \downarrow_2 = y_k - \epsilon_k \cdot * [\tau * \widetilde{x}] \downarrow_2 . \tag{11}$$

Notice that the process of a signal $\widetilde{x}$ passing through a PR filter bank as shown in Fig. 1 can be expressed as:

$$\widetilde{x} = g * [(\ell * \widetilde{x}) \downarrow_2] \uparrow_2 + q * [(h * \widetilde{x}) \downarrow_2] \uparrow_2 . \tag{12}$$

Combining Equation (11) and (12), we obtain the iterative reconstruction of $\widetilde{x}$ from K LR signals y_k as follows: At step $n + 1$

$$\widetilde{x}^{n+1} = q * [(h * \widetilde{x}^n) \downarrow_2] \uparrow_2 + g * \left(\frac{1}{K} \sum_{k=1}^{K} [y_k - \epsilon_k \cdot * (\tau * \widetilde{x}^n) \downarrow_2] \uparrow_2 \right). \tag{13}$$

Relation to other back-projection methods. Applying Equation (12), we can rewrite Equation (13) in the form

$$\widetilde{x}^{n+1} = (\widetilde{x}^n - g * [\ell * (\widetilde{x}^n) \downarrow_2] \uparrow_2 + g * \left(\frac{1}{K} \sum_{k=1}^{K} [y_k - \epsilon_k \cdot * (\tau * \widetilde{x}^n) \downarrow_2] \uparrow_2 \right)$$

$$= \widetilde{x}^n + g * \left(\frac{1}{K} \sum_{k=1}^{K} [y_k - (\ell * \widetilde{x}^n + \epsilon_k \cdot * (\tau * \widetilde{x}^n)) \downarrow_2] \uparrow_2 \right).$$

It can be seen that the iteration scheme presented here falls in the class of back-projection methods. But it has advantages over the usual back-projection iterations. Consider the well-known method by Irani and Peleg [1]. Its iteration can be described as:

$$x^{n+1} = x^n + \frac{1}{K} \sum_{k=1}^{K} T_k^{-1} \left(((y_k - [\ell * T_k(x^n)] \downarrow_2) \uparrow_2) * p \right), \tag{14}$$

where T_k is the geometric transform between y_k and $\widetilde{x}$, and the high-pass filter p is the deblurring kernel. Notice that the two methods differ in the deblurring kernel: one is g with $g(z) = h(-z)$ defined in Equation (10); the other is p in Equation (14), the approximate inverse filter of ℓ.

The requirement on p in (14) is

$$||\delta - \ell * p|| < 1, \tag{15}$$

where δ is the ideal unit impulse response filter. In other words, p should be a good approximation for the inverse of ℓ. In comparison, g in our iteration only needs to satisfy:

$$\ell(z)g(-z) - \ell(-z)g(z) = z^m, \tag{16}$$

that is, $\ell(z)g(-z)$ either has to have no odd-order components or no even-order components. Briefly, Iteration (14) requires a deblurring kernel p such that $\ell * p$ has small coefficients everywhere but the origin. In comparison, Iteration (13) only requires a kernel g with half the coefficients of $\ell * g$ being zero.

It is easy to see that the noise will be propagated exponentially as $O(\|p\|^n)$ in (14) and as $O(\|g\|^n)$ in (13). Generally the flexibility of $g(z)$ makes it possible to design a g that has much smaller norm than p. This leads to much better resistance to noise propagation. Here is an example: Consider $\ell = \frac{1}{4}(1,2,1)$. Then

$$g = (-1/8, -1/4, 3/4, -1/4, -1/8)$$

is a dual PR filter for ℓ with

$$\ell(z)g(-z) = -1 + 9z^{-2} + 16z^{-3} + 9z^{-4} - z^{-6}.$$

It is easy to check that $\|g\|_2$ is around 0.85. The minimum for the norm of all filters with the same length as g is around 1.1. The corresponding p is

$$p = (\frac{1}{2}, -\frac{2}{3}, \frac{4}{3}, -\frac{2}{3}, \frac{1}{2}).$$

This clearly indicates that our iteration scheme is more robust to noise than the usual back-projection scheme.

3.2   Algorithm on 2D Images with Denoising

Next we generalize the algorithm to 2D images. Furthermore we introduce a denoising process during the iterative reconstruction to suppress the noise in the optical flow estimation.

Extension to 2D image. All the previous analysis can be generalized using the tensor product. By an argument similar as for the 1D case, we approximate the LR image I^{LR} with the HR image I^{HR} as follows:

$$I^{LR} = [(a \otimes a) * I^{HR} + u \cdot *((a \otimes b) * I^{HR}) + v \cdot *((b \otimes a) * I^{HR})] \downarrow_2,$$

where "$\otimes$" is the Kronecker tensor product and (u, v) is the 2D optical flow vector. Then the 2D analysis bank is

Low-pass filter: $L = \ell \otimes \ell$,
High-pass filters: $H_1 = \ell \otimes h, H_2 = h \otimes \ell, H_3 = h \otimes h$

and the 2D synthesis filter bank is

$$\text{Low-pass filter:} \quad G = g \otimes g,$$
$$\text{High-pass filters:} \quad G_1 = g \otimes q, G_2 = q \otimes g, G_3 = q \otimes q.$$

It is easy to verify that the 2D filter bank defined above is a perfect reconstruction filter bank with the analysis filter bank $\{L, H_i\}$ and the reconstruction filter bank $\{G, Q_i\}$. Then generalizing (13), the iterative equation for the reconstruction of the HR image $\widetilde{I}$ from LR images I_k^{LR} amounts to:

$$\widetilde{I}^{n+1} = \sum_{i=1}^{3} Q_i * [(H_i * \widetilde{I}^{(n)}) \downarrow_2] \uparrow_2$$
$$+ G * \frac{1}{K} \left(\sum_k^K [\widetilde{I}_k^{LR} - u \cdot *((\ell \otimes \tau) * \widetilde{I}^n) \downarrow_2 - v \cdot *((\tau \otimes \ell) * \widetilde{I}^{(n)}) \downarrow_2] \uparrow_2 \right)$$

Recall that here $\widetilde{I}$ is the blurred version of the true I with $\widetilde{I} = (c \otimes c) * I$.

Modified algorithm with denoising process. There always is noise in the estimated flow u, v. However, the deconvolution operator could make the HR image reconstruction very sensitive to such noise. It is known that the noise variance of the solution will have hyperbolic growth when the blurring low-pass filter has zeros in the high frequencies. Thus, denoising is necessary in order to suppress the error propagation during the iterative reconstruction.

To suppress the noise, we introduce a wavelet denoising scheme which subtracts some high-frequency components from $\widetilde{I}^n$. Briefly, we first do a wavelet decomposition of the high-pass response, then apply a shrinkage of wavelet coefficients to the decomposition, and then reassemble the signal.

Our iteration scheme with built-in denoising operator amounts to

$$\widetilde{I}^{n+1} = \sum_{i=1}^{3} Q_i * [\Psi(H_i * \widetilde{I}^{(n)}) \downarrow_2] \uparrow_2$$
$$+ G * \frac{1}{K} \left(\sum_k^K [\widetilde{I}_k^{LR} - u \cdot *((\ell \otimes \tau) * \widetilde{I}^n) \downarrow_2 - v \cdot *((\tau \otimes \ell) * I^{(n)}) \downarrow_2] \uparrow_2 \right).$$

The denoising operator Ψ defined in the equation above is

$$\Psi(H_i * \widetilde{I}^n) = G * [(L * (H_i * \widetilde{I}^n)) \downarrow_2] \uparrow_2 + \sum_{i=1}^{3} [Q_i * \left(\Gamma[H_i * (H_i * \widetilde{I}^n)] \right) \downarrow_2] \uparrow_2,$$

where Γ is the thresholding operator. Here we use the following soft-denoising scheme ([12]):

$$\Gamma(\nu) = \begin{cases} Sign(\nu)(\nu - \mu) & \text{if } |\nu| > \mu \\ 0 & \text{Otherwise.} \end{cases}$$

Briefly the process of super-resolution is:

1. Compute the affine flow between every frame and the key frame.
2. Transform the affine flow to sub-pixel shifts on the finer grid of the HR image.
3. Apply the iteration process described above to the key frame to obtain the HR image.

The algorithm above could easily be adapted to different blur filters by modifying the corresponding dual filters G, Q_i. Also, here we only consider a resolution increase by a factor 2. Any other resolution increase could easily be achieved by changing the two-channel PR filter bank to an M-channel PR filter bank.

Relation to regularization methods. One popular denoising technique used for robust reconstruction is regularization ([13]). Recall that back-projection methods basically find $\widetilde{x}$ by minimizing $\sum_k^K \|y_k - \widetilde{y}_k(\widetilde{x})\|_2^2$, where $\widetilde{y}_k(\widetilde{x})$ is the LR signals derived from our estimated $\widetilde{x}$. Such a least square problem usually is ill-conditioned. One way to increase the stability is to enforce a regularization term and solve:

$$\min_{\widetilde{x}} \sum_k^K \|y_k - \widetilde{y}_k(\widetilde{x})\|_2^2 + \alpha\|\Phi(\widetilde{x})\|,$$

where Φ is some regularization function and α is some pre-defined smoothing factor. If the regularization is a least squares problem, we call it a Tikhonov-type regularization. The advantage is its simplicity and efficiency, the disadvantage is its relatively poor performance. A nonlinear diffusion regularization, like Total Variation regularization usually performs better, but is computational expensive.

Wavelet denoising is closely related to nonlinear diffusion regularization. [14] shows that a simple soft-denoising with Haar wavelets ($\ell = \frac{1}{2}(1,1), h = \frac{1}{2}(1,-1)$) is equivalent to Total Variation based nonlinear diffusion ($\Phi(\widetilde{x}) = \|\widetilde{x}\|_1$) for a two-pixel signal. Roughly speaking, the wavelet denoising process in our reconstruction is comparable to some nonlinear diffusion regularization schemes in its ability to suppress the error propagation. However it doesn't have the computational burden of most nonlinear diffusion regularizations, since it only needs a linear wavelet decomposition over one level. In comparison nonlinear regularizations need to solve a nonlinear optimization.

4   Experiments and Conclusion

We compare our algorithm's high-resolution reconstruction to standard methods using both simulated and real data.

Simulated data. We simulated 4 low-resolution images (16×16) from a high resolution image by shifting, blurring and downsampling. The blurring filter is

$$\ell = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

Noise of three types of sources was simulated:

1. Error in motion estimation. It is modeled by local Gaussian white noise with parameter σ. The local covariance matrix is made up by the magnitudes of the image gradients.
2. Noise in pixel formation. We added a Gaussian white noise with parameter γ to the pixel values.

3. Error in PSF modeling. We also checked how error in the PSF modeling influences the performance. The approximated PSF $\hat{\ell}$ used in the reconstruction was

$$\hat{\ell} = \frac{1}{16} \begin{pmatrix} 1\ 1\ 1 \\ 1\ 8\ 1 \\ 1\ 1\ 1 \end{pmatrix}.$$

We compared our wavelet-based method to the popular "POCS" back-projection method ([8]) enforced by Tikhonov regularization (See Fig. 2). It is possible that enforcing Total Variation regularization would give a bit better results. However, it requires solving a nonlinear minimization over each iterative step during the reconstruction, which is very computational expensive. In our implementation, the regularization term is the 2-norm of the Laplacian smoothness constraint with parameter $\alpha = 1$.

Fig. 3 demonstrates how well the wavelet-based method performs for various noise settings. Performance is measured by the SNR (Signal-to-Noise ratio) of the reconstructed image to the true image, which is defined as: $SNR = 20 \log_{10} \frac{\|x\|_2}{\|x-\hat{x}\|_2}$, where $\hat{x}$ is the estimation for the true image x. Fig. 3 clearly indicates the advantage of our wavelet-based method in suppressing the noise. Especially when noise is large, the boost in performance is significant.

Real data. We used an indoor sequence with a paper box (Fig. 4) of 13 image frames. An interesting planar region was chosen manually. Fig. 5 and Fig. 6

(a) Original image   (b) Noisy LR image   (c) Wavelet   (d) Tikhonov

Fig. 2. The HR images (c) and (d) are reconstructed from four LR images by five iterations. (c) is reconstructed by our method. (d) is reconstructed by POCS method with Tikhonov regularization. The motion noise is local Gaussian noise with $\sigma = 0.2$. The image formation noise is Gaussian noise with $\gamma = 0.01$. The approximation $\hat{\ell}$ is used in the reconstruction instead of the true PSF ℓ.
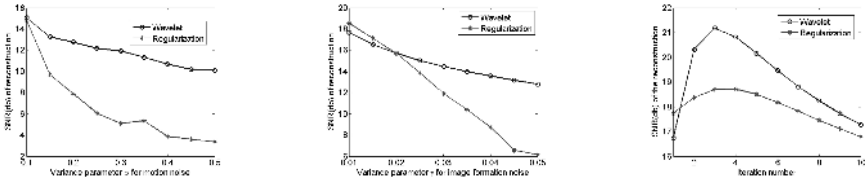
(a) Comp. for flow noise   (b) Comp. for formation noise   (c) Comp. for PSF error

Fig. 3. (a) and (b) Compare two methods for various amounts of motion noise and image formation noise. The reconstructed image is obtained by 5 iterations. The x-axis denotes the variance of the noise, the y-axis denotes the SNR of the reconstruction. (c) compares the performance of two methods over the iterations for PSF model error. The x-axis denotes the iteration number, the y-axis denotes the SNR of the reconstruction.
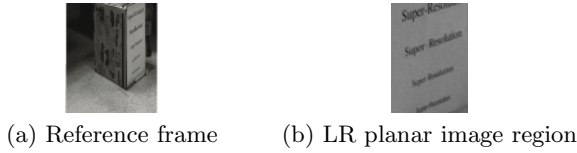
(a) Reference frame   (b) LR planar image region

Fig. 4. The reference image frame from the video and its selected region

(a)Interpolation   (b) POCS   (c) Wavelet
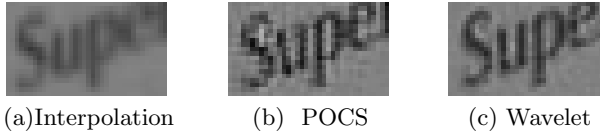
Fig. 5. Comparison of one reconstructed HR region for various methods

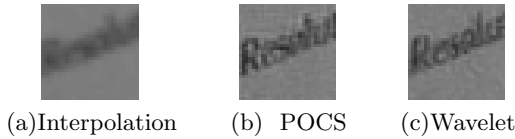(a)Interpolation   (b) POCS   (c)Wavelet

Fig. 6. Comparison of another reconstructed HR region for various methods

show the comparisons of four different methods for different regions. Here the reconstructed HR images double the resolution of the LR images. The HR image in Fig. 5-6(a) were obtained by cubic interpolation from a single LR image. In Fig. 5-6(b) we used the POCS method, where the flow field is estimated by an affine motion model. Fig. 5-6(c) show the results from our reconstruction scheme. The difference can be visually evaluated. Clearly, there is large improvement from (b) to (c) in Fig. 5 and Fig. 6. The letters in Fig. 5(c) and Fig. 6(c) are the clearest, and there are minimal artifacts around the edges.

Summary. We have presented a theoretical analysis and a new algorithm for super-resolution problem based on wavelet theory. It has been demonstrated both in theory and experiments that the proposed method in this paper is very robust to noise without sacrificing efficiency. The reconstruction scheme allows for super-resolution reconstruction from general video sequences, even when the estimated optical flow is very noisy.

References

1. Irani, M., Peleg, F.: Motion analysis for image enhancement: Resolution, occlusion and transparency. Journal of Visual Comm. and Image Repr. **4** (1993) 324–335.
2. Eland, M., Feuer, A.: Restoration of a signal super-resolution image from several blurred, noisy and undersampled measured images. IEEE Transaction on Image Processing (1997) 1646–1658.
3. Bascle, B., Blake, A., Zisserman, A.: Motion deblurring and super-resolution from an image sequence. In: ECCV. Volume 2. (1996) 573–582.

4. Zhao, W., Sawhney, H.S.: Is super-resolution with optical flow feasible? In: ECCV. (2002) 599–613.
5. Tekalp, A., Ozkan, M., Sezan, M.: High-resolution image reconstruction from low-resolution image sequences and space-varying image restoration. In: ICASSP. (1992) 169–172.
6. Farsiu, S., Robinson, D., Elad, M., Milanfar, P.: Robust shift and add approach to super-resolution. In: SPIE. (2003).
7. Chambolle, A., Devore, R., Lee, N., Lucier, B.: Nonlinear wavelet image processing: variational problems, compression and noise removal through wavelets. IEEE Trans. Image Processing **7** (1998).
8. Youla, C.: Generalized image restoration by the method of alternating orthogonal projections IEEE Trans. Circuits Syst. **25** 1978.
9. Chan, R., Chan, T., Shen, L., Shen, Z.: Wavelet deblurring algorithms for spatially varying blur from high-resolution image reconstruction. Linear algebra and its applications **366** (2003) 139–155.
10. Nguyen, N., Milanfar, N.P.: An wavelet-based interpolation-restoration method for superresolution. Circuits, Systems and Signal Processing **19** (2002) 321–338.
11. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. In: CVPR. (2000) 372–379.
12. Mallat, S.: A wavelet tour of signal processing. Academic Press (1999).
13. Weickert, J.: Anisotropic Diffusion in Image Processing. ECMI Series, Teubner, Stuttgart, 1998.
14. Mrazek, P., Weickert, J., Steidl, G.: Correspondences between wavelet shrinkage and nonlinear diffusion. In: Scale-Space. (2003) 101–116.