

A collaborative adaptive Wiener filter for multi-frame super-resolution

Khaled M. Mohamed* and Russell C. Hardie

Image Processing Lab, Department of Electrical and Computer Engineering, University of Dayton, Dayton, OH, USA

OPEN ACCESS

Edited by:

Andrey Vadimovich Kanaev,
US Naval Research Laboratory, USA

Reviewed by:

David Mayerich,
University of Houston, USA
Junichi Fujikata,
Photonics Electronics Technology
Research Association, Japan

*Correspondence:

Khaled M. Mohamed,
Image Processing Lab, Department of
Electrical and Computer Engineering,
University of Dayton, 300 College
Park, Dayton, OH 45469-0226, USA
mohamedk1@udayton.edu

Specialty section:

This article was submitted to
Optics and Photonics,
a section of the journal
Frontiers in Physics

Received: 31 December 2014

Accepted: 08 April 2015

Published: 29 April 2015

Citation:

Mohamed KM and Hardie RC (2015)
A collaborative adaptive Wiener filter
for multi-frame super-resolution.
Front. Phys. 3:29.
doi: 10.3389/fphy.2015.00029

Factors that can limit the effective resolution of an imaging system may include aliasing from under-sampling, blur from the optics and external factors, and sensor noise. Image restoration and super-resolution (SR) techniques can be used to improve image resolution. One SR method, developed recently, is the adaptive Wiener filter (AWF) SR algorithm. This is a multi-frame SR method that combines registered temporal frames through a joint nonuniform interpolation and restoration process to provide a high-resolution image estimate. Variations of this method have been demonstrated to be effective for multi-frame SR, as well demosaicing RGB and polarimetric imagery. While the AWF SR method effectively exploits subpixel shifts between temporal frames, it does not exploit self similarity within the observed imagery. However, very recently, the current authors have developed a multi-patch extension of the AWF method. This new method is referred to as a collaborative AWF (CAWF). The CAWF method employs a finite size moving window. At each position, we identify the most similar patches in the image within a given search window about the reference patch. A single-stage weighted sum of all of the pixels in all of the similar patches is used to estimate the center pixel in the reference patch. Like the AWF, the CAWF can perform nonuniform interpolation, deblurring, and denoising jointly. The big advantage of the CAWF, vs. the AWF, is the CAWF can also exploit self-similarity. This is particularly beneficial for treating low signal-to-noise ratio (SNR) imagery. To date, the CAWF has only been developed for Nyquist-sampled single-frame image restoration. In this paper, we extend the CAWF method for multi-frame SR. We provide a quantitative performance comparison between the CAWF SR and the AWF SR techniques using real and simulated data. We demonstrate that CAWF SR outperforms AWF SR, especially in low SNR applications.

Keywords: aliasing, image restoration, super-resolution, under-sampling, correlation model, multi-frame, multi-patch

1. Introduction

Multi-frame super-resolution (SR) is an image processing technique that provides an effective increase in the sampling density of an imaging sensor. This is done by combining pixels from multiple shifted low-resolution (LR) frames onto a common high-resolution (HR) image grid. In this manner, one has more samples per unit area, compared with that provided by the native focal plane array (FPA), with which to reconstruct an HR image. For undersampled imaging systems, this increase in sampling rate can provide a reduction in aliasing

artifacts, which provides improved resolution. SR processing can further improve image resolution by reduction of noise and blur.

The interpolation-restoration SR methods tend to be the simplest, both conceptually and in terms of the computational complexity [1]. A wide variety of such methods have been proposed in the literature [2–6]. Most of these methods begin with registration of LR frames to a common HR reference grid. Then, nonuniform interpolation is used to get uniformly spaced pixel values. This increases the effective sampling rate of the imaging sensor, but doesn't treat noise or other blur. A restoration procedure typically follows the nonuniform interpolation to tackle the effects of the system point function (PSF) and noise. Most such methods decouple the nonuniform interpolation and restoration steps for simplicity. However, the partition weighted sum (PWS) SR method [7] and the adaptive Wiener filter (AWF) SR method [8] perform the nonuniform interpolation and restoration steps jointly. This provides performance and computational advantages in many cases. In particular, the AWF SR method has been shown to provide best-in-class performance among many interpolation-restoration methods [8]. The AWF SR method uses a weighted sum of registered LR pixels from multiple frames, based on a correlation model and knowledge of the subpixel displacements between frames. In addition to multi-frame SR for grayscale imagery, variations of the AWF SR method have been successfully applied to demosaicing, color SR, and SR processing of polarimetric imagery [9–13]. While the AWF SR method effectively exploits subpixel shifts between temporal frames, it does not exploit self similarity within the observed imagery. Recently, a spatially adaptive filtering method for SR has been proposed [14, 15]. This recursive method is based on iterative video block matching and 3D filtering (V-BM3D)[16]. One remarkable advantage of V-BM3D approach is it can exploit nonlocal similarity between blocks within the temporal frames. Very recently, the current authors have developed a multi-patch extension of the AWF method for image restoration [17]. This new method is referred to as the collaborative AWF (CAWF). The CAWF method is also capable of exploring the self similarity during image restoration for its estimate. The CAWF method employs a finite size moving window. At each position, we identify the most similar patches in the image within a given search window about the reference patch. A single-stage weighted sum of all of the pixels in all of the similar patches is used to estimate the center pixel in the reference patch. The weights are determined based on a new multi-patch correlation model.

To date, the CAWF method has only been developed for and applied to single image restoration with no aliasing [17]. In this paper, we develop and demonstrate the CAWF method for multi-frame SR. Like the AWF, it can perform nonuniform interpolation, deblurring, and denoising jointly. The big advantage of the CAWF SR vs. the AWF, is the CAWF can also exploit self-similarity within the image sequence. This is particularly beneficial for multi-frame SR when the signal-to-noise ratio (SNR) is low. We provide a quantitative and subjective performance comparison between the AWF SR and the CAWF SR using simulated and real data. The CAWF SR method shows the capability of delivering high performance, in terms

of objective metrics and subjective visual quality, particularly for low SNR applications.

The rest of this paper is organized as follows. The observation model is defined and discussed in Section 2. Section 3 presents the CAWF SR algorithm. In Section 4, we demonstrate the algorithm efficacy using simulated and real image data. Finally, in Section 5, we offer conclusions.

2. Observation Model

Here we present the observation model used to develop the CAWF SR method. The observation model closely follows that used for the AWF SR method in Hardie [8]. However, we present it again here for the reader's convenience. A block diagram of the observation model is shown in **Figure 1**. It starts with desired continuous image as input, $d(x, y)$. The ideal discrete HR image in lexicographical notation is given by $\mathbf{z} = [z_1, z_2, \dots, z_N]^T$, where N is the total number of HR pixels. It is this image we wish to estimate with the SR processing. In the model, the desired continuous image goes through a geometric transformation to account for intra-frame motion during video acquisition. The multiple outputs can be expressed as

$$d_p(x, y) = T_{\theta_p} \{d(x, y)\}, \quad (1)$$

where T_{θ_p} is a transformation operator associated with the p 'th frame, where $p = 1, 2, \dots, P$. In this paper, we consider only global translational intra-frame motion. After the geometric transformation, the model includes convolution with the system point spread function (PSF). This can be represented as

$$f_p(x, y) = d_p(x, y) * h(x, y), \quad (2)$$

where $h(x, y)$ is the PSF. The PSF model can include a variety of blurring contributors, such as motion blur and atmospheric effects. However, here we follow the approach in Hardie [8] and consider only diffraction from a circular exit pupil and spatial detector integration. We refer the reader to Hardie et al. [18] for details on the PSF model. The next step in the observation model is the sampling block. An important consideration here is the sampling rate relative to the cut-off frequency of the optics. According to the Nyquist sampling theory, the continuous image can only be reconstructed from its samples if

$$\frac{1}{\delta_s} \geq 2\rho_c, \quad (3)$$

where δ_s is the spatial sampling interval in the horizontal and vertical dimensions, and ρ_c is the optical cut off frequency. Note that the optical cut off frequency can be expressed in terms of the wavelength of light and the f -number of the optics. In particular, this is given by

$$\rho_c = \frac{1}{\lambda f/\#}, \quad (4)$$

where λ is the wavelength, and $f/\#$ is the f -number. Sampling at a lower rate can lead to aliasing artifacts in the acquired

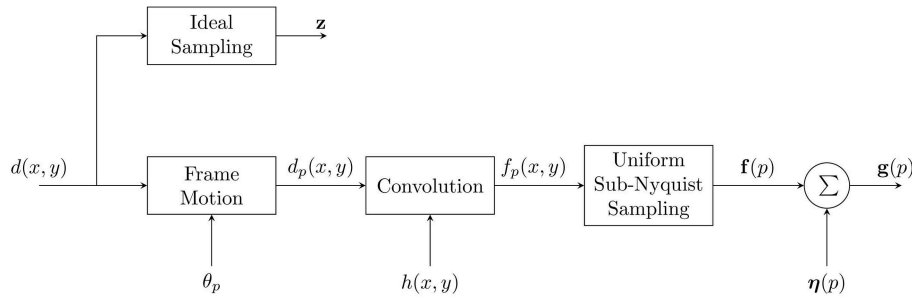


FIGURE 1 | Forward observation model relating the desired continuous image, $d(x, y)$ to the observed LR frames, $g(p)$.

imagery. It also makes an impulse invariant discrete model of the PSF invalid. However, because there are many trade-offs to be made in imaging system design, most imaging systems are designed with some level of undersampling [19]. Let the p 'th sampled frame be represented in lexicographical form as $\mathbf{f}(p) = [f_1(p), f_2(p), \dots, f_J(p)]^T$, where J is the number of LR pixels in each frame. We shall assume that these images are sampled at a rate below the ideal image \mathbf{z} by an integer factor of L in the horizontal and vertical dimensions.

The final step in the block diagram is the noise block. We assume the noise is additive Gaussian noise with zero mean and variance of σ_η^2 . Thus, LR output frame p can be written in a lexicographical form as

$$\mathbf{g}(p) = \mathbf{f}(p) + \boldsymbol{\eta}(p), \tag{5}$$

where $\boldsymbol{\eta}(p) \sim N(\mathbf{0}, \sigma_\eta^2 \mathbf{I})$. Note that for the model in **Figure 1**, the motion and convolution blocks can be commuted for translational motion [8]. The blocks can also be approximately commuted for affine motion under certain conditions [13]. After the commutation, the uniform sampling block can be replaced with a nonuniform sampling block that incorporates samples from all P LR frames on a common grid. This provides a nonuniformly sampled version of $f(x, y)$, but with a denser sampling grid than a single frame provides. The noisy version of these nonuniform samples are represented here in lexicographical notation as $\mathbf{g} = [g_1, g_2, \dots, g_P]^T$.

3. Collaborative Adaptive Wiener Filter for Super-Resolution

3.1. CAWF Overview

The goal of the proposed CAWF SR is to estimate \mathbf{z} from the observed LR frames, $\mathbf{g}(p)$, for $p = 1, 2, \dots, P$. In order to help the reader follow and easily understand the CAWF concept, we define some CAWF parameters in **Table 1**. The CAWF SR algorithm is shown in **Figure 2**. First, subpixel registration is done to place all of the observed LR samples on a common HR grid. The registration is done by using the gradient-based registration technique described in Hardie et al. [20], Irani and Peleg [21], Lucas and Kanade [22]. As mentioned in Section 2, the full set of samples are represented by the vector \mathbf{g} . The CAWF SR

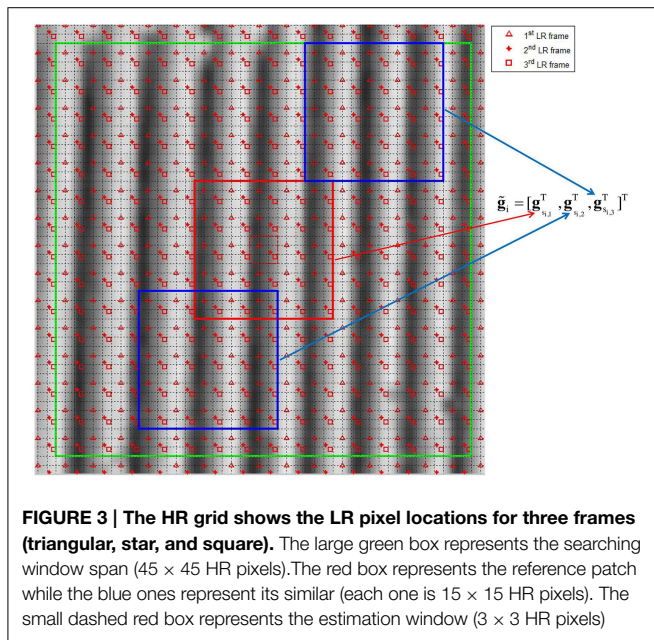
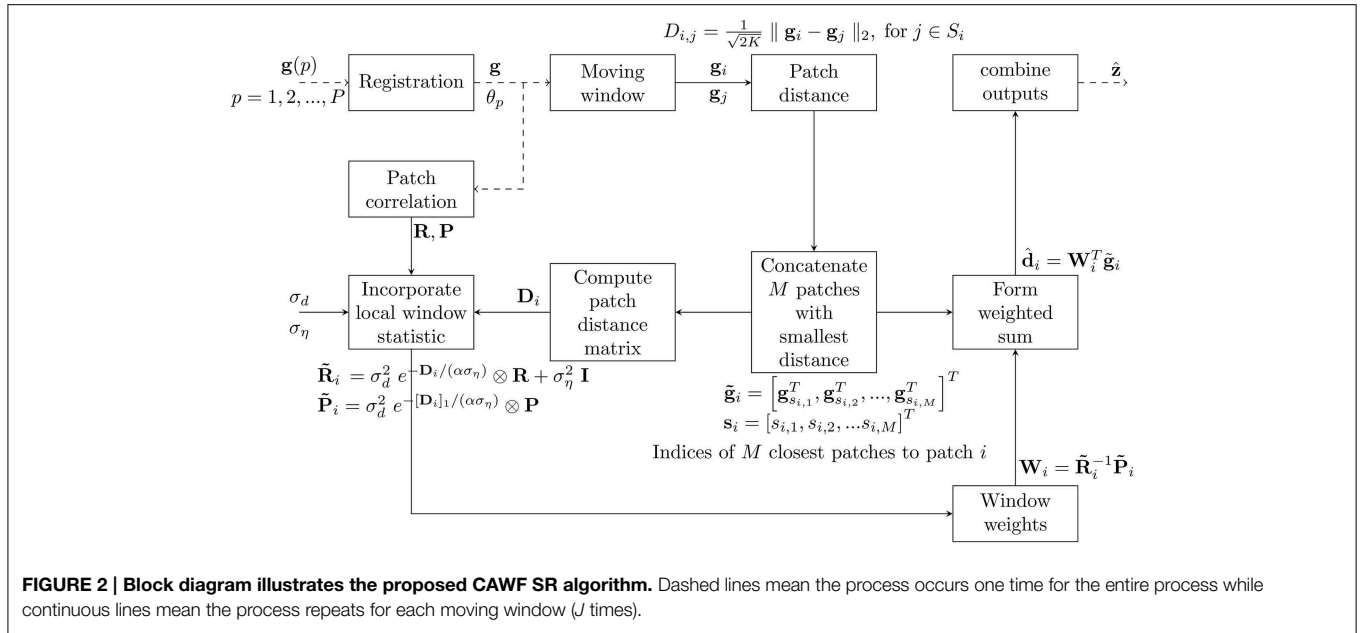
TABLE 1 | CAWF parameters definitions.

Parameter name	Variable	Definitions
HR patch size	$W \times W$	The span in HR pixels for a patch.
HR searching window size	$Z \times Z$	The span in HR pixels where patches can be identified as similar patches.
Number of similar patches	M	Number of patches which are identified as similar patches.
Downsampling and upsampling	L	The factor which is used to up and dawn sampling an image.
Number of pixels in patch	K	The total of LR pixels in HR grid within the patch span.

method employs a $W \times W$ HR pixel spanning moving patch that passes over the HR grid in steps of L HR pixels in the horizontal and vertical dimensions. Let the observed samples that lie within the span of the i 'th patch be denoted with the vector \mathbf{g}_i . Let the number of pixels in each of these vectors be denoted as K . Note that K will be constant because we are considering only translational intra-frame motion, and we move patches only in LR pixel spacings across the HR grid. At each reference patch position, we identify the M most similar patches within a search window about the reference patch of size $Z \times Z$ HR pixels. In particular, for a given reference patch index i , we compute the following distances

$$D_{i,j} = \frac{1}{\sqrt{2K}} \|\mathbf{g}_i - \mathbf{g}_j\|_2, \tag{6}$$

for $j \in S_i$, where S_i contains the window indices within the search window. We select the M patches corresponding to the M smallest distances. All pixels in the similar patches, including the reference patch, are concatenated in one vector. This combined vector is given by $\tilde{\mathbf{g}}_i = [\mathbf{g}_{s_{i,1}}^T, \mathbf{g}_{s_{i,2}}^T, \dots, \mathbf{g}_{s_{i,M}}^T]^T$, where $\mathbf{s}_i = [s_{i,1}, s_{i,2}, \dots, s_{i,M}]^T$. To illustrate this multi-patch selection process, **Figure 3** shows an $L = 3$ HR grid made from three registered LR frames. The solid red box is the reference patch, and the blue boxes represent the identified similar patches. The large green box represents the span of the search window.



The CAWF SR method estimates the $L \times L$ HR pixels in the center of each observation patch. This $L \times L$ estimation window is shown in **Figure 3** as a small red dashed box. All of the pixels in the $\tilde{\mathbf{g}}_i$ are used in a weighted to estimate these HR pixels. This is given by

$$\hat{\mathbf{d}}_i = \mathbf{W}_i^T \tilde{\mathbf{g}}_i, \quad (7)$$

where $\hat{\mathbf{d}}_i = [d_{i,1}, d_{i,2}, \dots, d_{i,L^2}]^T$ is a vector of estimated desired pixels, and \mathbf{W}_i is a matrix of weights. After processing the entire image, all of these outputs are combined to form the estimated

TABLE 2 | The AWF and the CAWF system parameters which are used for both real and simulated data results.

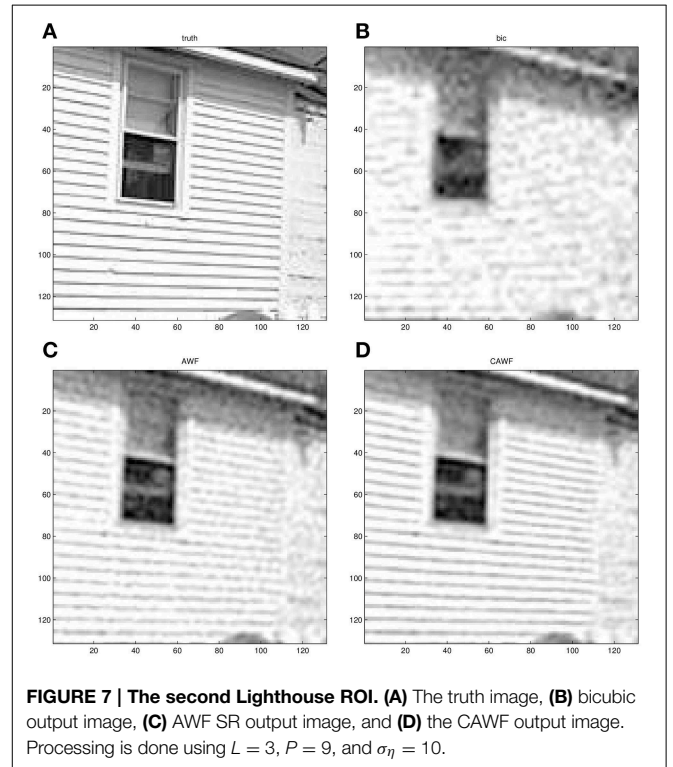
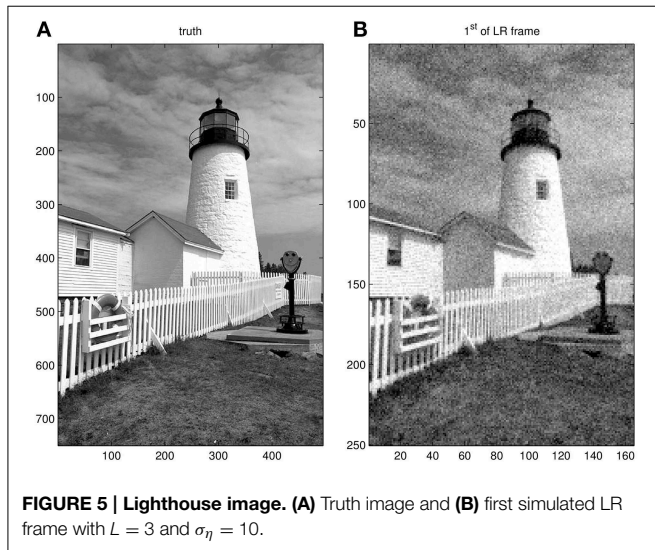
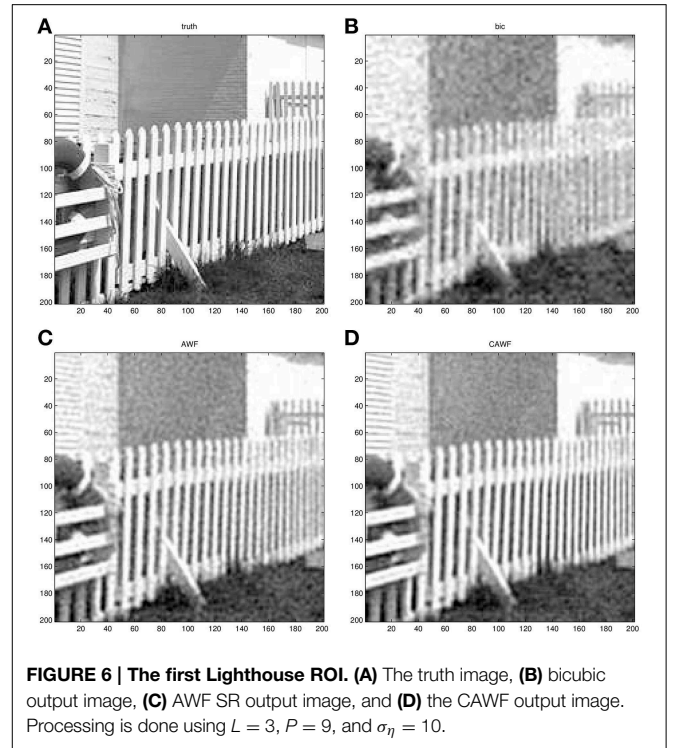
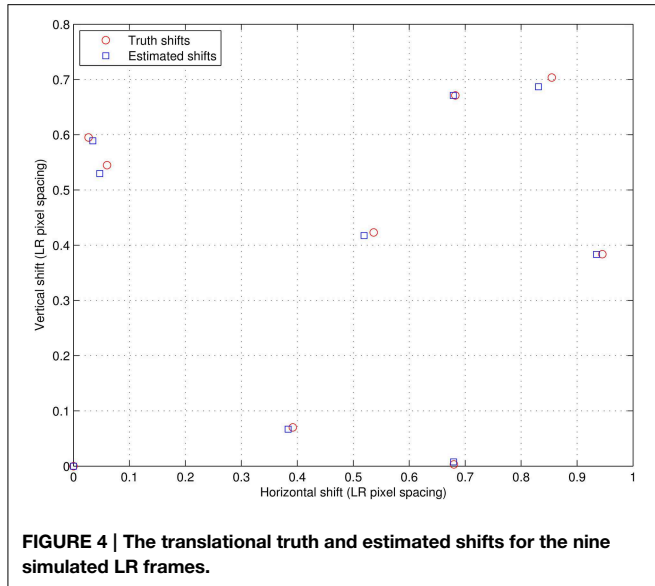
Parameter name	Variable	CAWF values	AWF values
HR patch size	$W \times W$	15×15 pixels	15×15 pixels
HR search window size	$Z \times Z$	45×45 pixels	N/A
Number of similar patches	M	10	1
Autocorrelation decay	ρ	0.75	0.75
Patch similarity decay	α	5	N/A
Patch intensity variance	σ_d^2	1000	1000
F -number	$f/\#$	$f/5.6$	$f/5.6$
Number of frames	P	9	9
Downsampling and upsampling	L	3	3

TABLE 3 | Imaging Source camera (DMK 23U618) specifications.

Parameter name	Values
Pixel pitch	$5.6 \mu\text{m}$
FPA size	640×480
Dynamic range	8 bit
Gain	16.5 dB
Sensitivity	0.015 lx
Shutter	1/64 s

HR image, denoted as $\hat{\mathbf{z}}$. The weights are designed to minimize mean squared error (MSE) [17] and are given by

$$\mathbf{W}_i = \tilde{\mathbf{R}}_i^{-1} \tilde{\mathbf{P}}_i, \quad (8)$$



where $\tilde{\mathbf{R}}_i = E\{\tilde{\mathbf{g}}_i \tilde{\mathbf{g}}_i^T\}$ is a $KM \times KM$ auto-correlation matrix for the multi-patch observation vector $\tilde{\mathbf{g}}_i$, and $\tilde{\mathbf{P}}_i = E\{\tilde{\mathbf{g}}_i \mathbf{d}_i^T\}$ is a $KM \times L^2$ cross-correlation matrix between the multi-patch observation vector, $\tilde{\mathbf{g}}_i$, and the desired vector, \mathbf{d}_i . Note that the vectors involved are treated as random vectors here, and $E\{\cdot\}$ represents the expectation operator. The correlations are modeled using a new nonuniform multi-patch correlation model described in Section 3.2.

3.2. Nonuniformly Sampled Multi-Patch Correlation Model

This section presents the nonuniformly sampled multi-patch correlation model used to provide the correlations needed for Equation (8). The model is similar to that in Mohamed and Hardie [17], but here applies to the nonuniformly sampled

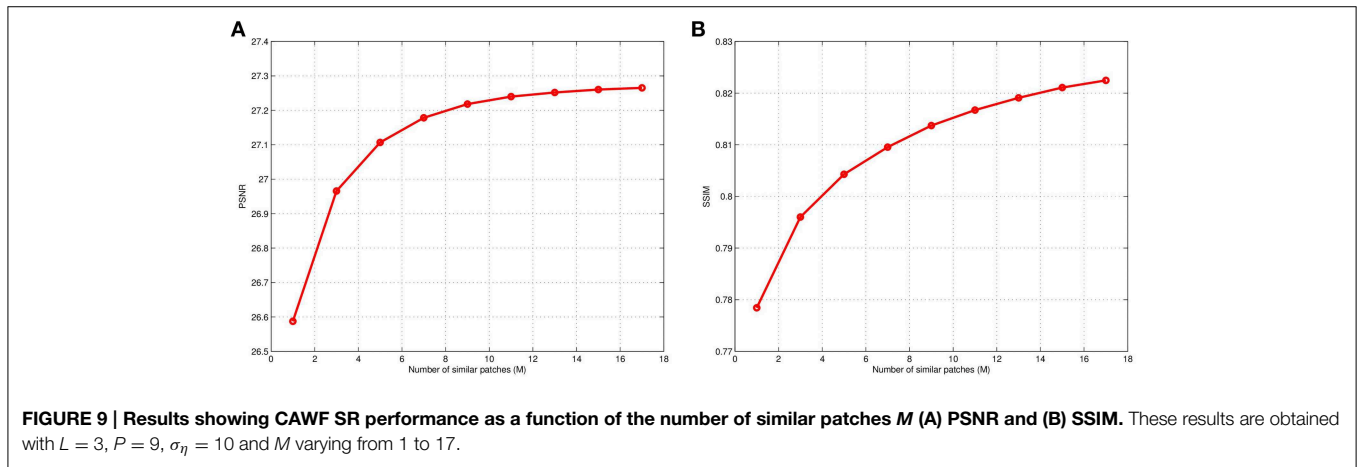
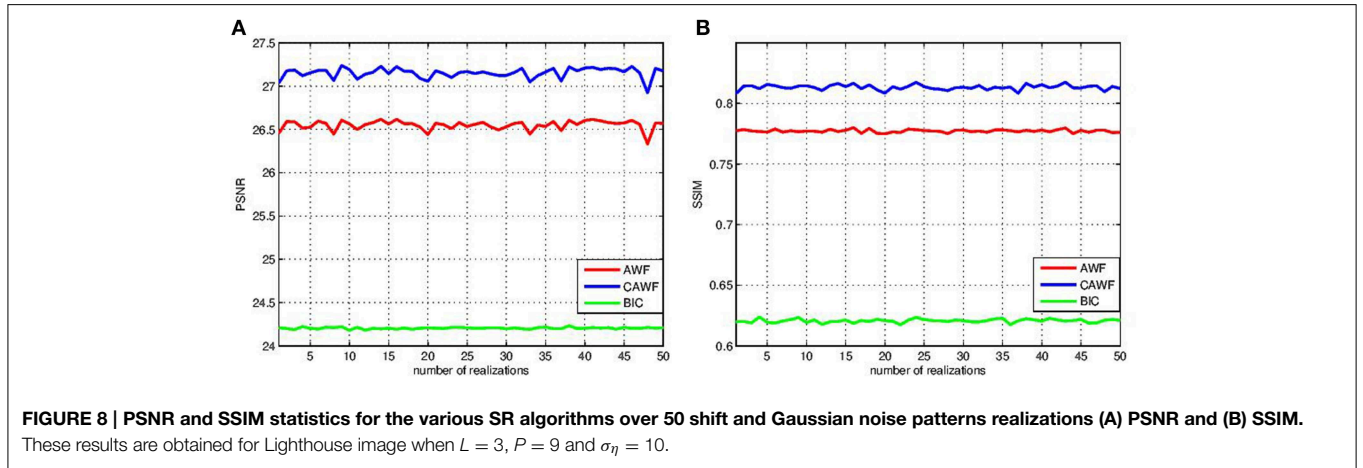
multi-patch observation vector $\tilde{\mathbf{g}}_i$. In particular, the auto-correlation matrix model for $\tilde{\mathbf{g}}_i$ is given by

$$\tilde{\mathbf{R}}_i = \sigma_d^2 e^{-\mathbf{D}_i / (\alpha \sigma_\eta)} \otimes \mathbf{R} + \sigma_\eta^2 \mathbf{I}, \quad (9)$$

TABLE 4 | Quantitative performance of the bicubic, the AWF and the CAWF outputs for 7 Kodak images for $L = 3, P = 9$, and $\sigma_\eta = 10$.

Method	Metric type	Images names						
		Kodim03	Kodim04	Kodim09	Kodim10	Kodim15	Kodim19	Kodim23
Bicubic	psnr	27.85	27.55	26.56	26.85	27.33	24.19	27.67
	ssim	0.63	0.64	0.63	0.65	0.63	0.62	0.65
AWF	psnr	30.17	29.96	29.29	29.24	29.68	26.59	30.29
	ssim	0.76	0.78	0.77	0.79	0.76	0.78	0.78
CAWF	psnr	30.70	30.43	29.74	29.66	30.12	27.21	30.83
	ssim	0.80	0.82	0.81	0.83	0.81	0.81	0.82

Bold numbers means the "highest performance."



and the cross correlation matrix is given by

$$\tilde{\mathbf{P}}_i = \sigma_d^2 e^{-[\mathbf{D}_i]_1 / (\alpha \sigma_\eta)} \otimes \mathbf{P}, \quad (10)$$

where \mathbf{I} is a $KM \times KM$ identity matrix, \otimes is a Kronecker product, σ_d^2 is the variance of desired image, α is the patch similarity decay, \mathbf{D}_i is an $M \times M$ distance matrix between patches, and $[\mathbf{D}_i]_1$ is the first column of the distance matrix. The terms \mathbf{R} and \mathbf{P}

represent the auto-correlation and cross-correlation matrices for a single patch. More will be said about these shortly. In our model, these single patch statistics are “modulated” by the distance matrix to account for interpatch similarity. The parameter α is a tuning parameter to govern the strength of this modulation. The SNR is incorporated into the model with the selection of σ_d^2 and σ_η^2 . The distance matrix is populated using the distance metric in Equation (6) using the M similar patches as follows

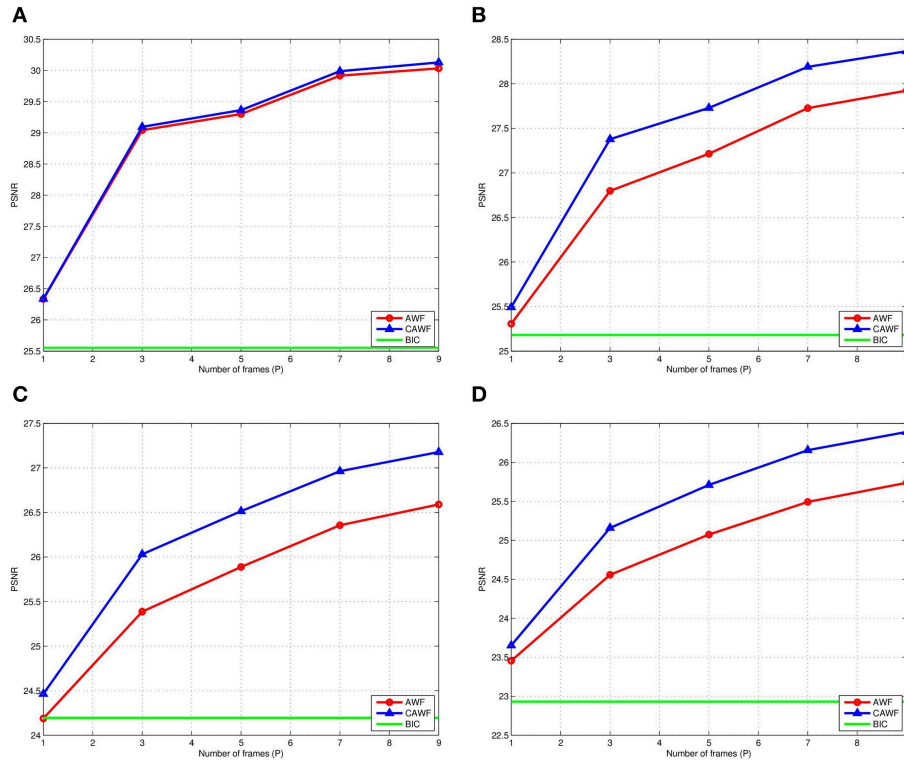


FIGURE 10 | The PSNR for AWF, CAWF, and the bicubic outputs with different levels of noise. (A) $\sigma_\eta = 1$, (B) $\sigma_\eta = 5$, (C) $\sigma_\eta = 10$, and (D) $\sigma_\eta = 15$. These results are obtained with $M = 10$ and the number of frames, P , varies from 1 to 9.

$$\mathbf{D}_i = \begin{bmatrix} D_{s_{i,1},s_{i,1}} & D_{s_{i,1},s_{i,2}} & \cdots & D_{s_{i,1},s_{i,M}} \\ D_{s_{i,2},s_{i,1}} & D_{s_{i,2},s_{i,2}} & \cdots & D_{s_{i,2},s_{i,M}} \\ \vdots & \vdots & \ddots & \vdots \\ D_{s_{i,M},s_{i,1}} & D_{s_{i,M},s_{i,2}} & \cdots & D_{s_{i,M},s_{i,M}} \end{bmatrix}. \quad (11)$$

Let us now consider the single patch statistics. Note that

$$\mathbf{R} = E \left\{ \bar{\mathbf{f}} \bar{\mathbf{f}}^T \right\}, \quad (12)$$

is the $K \times K$ auto-correlation matrix of a single noise-free patch obtained from a variance-normalized desired image. The matrix

$$\mathbf{P} = E \left\{ \bar{\mathbf{f}} \bar{\mathbf{d}}^T \right\}, \quad (13)$$

is the $K \times L^2$ cross-correlation matrix for a single noise-free patch obtained from a variance-normalized desired image and the desired pixels in the estimation window. The single patch statistics differ from those in Mohamed and Hardie [17] in that here a single patch is a nonuniformly sampled patch made up of samples from multiple registered LR frames. The relative positions of the samples in the patch depend on the particular interframe registration. Thus, we model these statistics based on continuous-space correlation functions in the same manner as that described in Hardie [8]. In particular, we assume the desired image is a wide sense stationary and we get auto-correlation of normalized noise-free image as

$$r_{\bar{\mathbf{f}}\bar{\mathbf{f}}}^-(x, y) = r_{\bar{\mathbf{d}}\bar{\mathbf{d}}}^-(x, y) * h(x, y) * h(-x, -y), \quad (14)$$

where $r_{\bar{\mathbf{d}}\bar{\mathbf{d}}}^-(x, y)$ is the autocorrelation function of the variance normalized desired image. Following the approach in Hardie [8], this is modeled as

$$r_{\bar{\mathbf{d}}\bar{\mathbf{d}}}^-(x, y) = \rho \sqrt{x^2 + y^2}, \quad (15)$$

where ρ is auto-correlation decay constant. Similarly, the cross-correlation between noise-free signal obtained from a variance normalized desired image and the variance normalized desired image is given by

$$r_{\bar{\mathbf{d}}\bar{\mathbf{f}}}^-(x, y) = r_{\bar{\mathbf{d}}\bar{\mathbf{d}}}^-(x, y) * h(x, y). \quad (16)$$

Since the motion is global translation, spacing between LR pixels is periodic (i.e., the same pattern) for the entire HR grid. With knowledge of the shifts between all LR frames, θ_p , horizontal and vertical distances between all samples in a patch can be computed easily. Then by applying these displacements to Equation (15) and then into Equations (16) and (14), we can fill both correlation matrices in Equations (12) and (13).

Finally, combining (8) with (12) and (13), we can express the CAWF SR weights as

$$\mathbf{W}_i = \tilde{\mathbf{R}}_i^{-1} \tilde{\mathbf{P}}_i = \left[e^{-\mathbf{D}_i / (\alpha \sigma_\eta)} \otimes \mathbf{R} + \frac{\sigma_\eta^2}{\sigma_d^2} \mathbf{I} \right]^{-1} e^{-[\mathbf{D}_i]_1 / (\alpha \sigma_\eta)} \otimes \mathbf{P}. \quad (17)$$

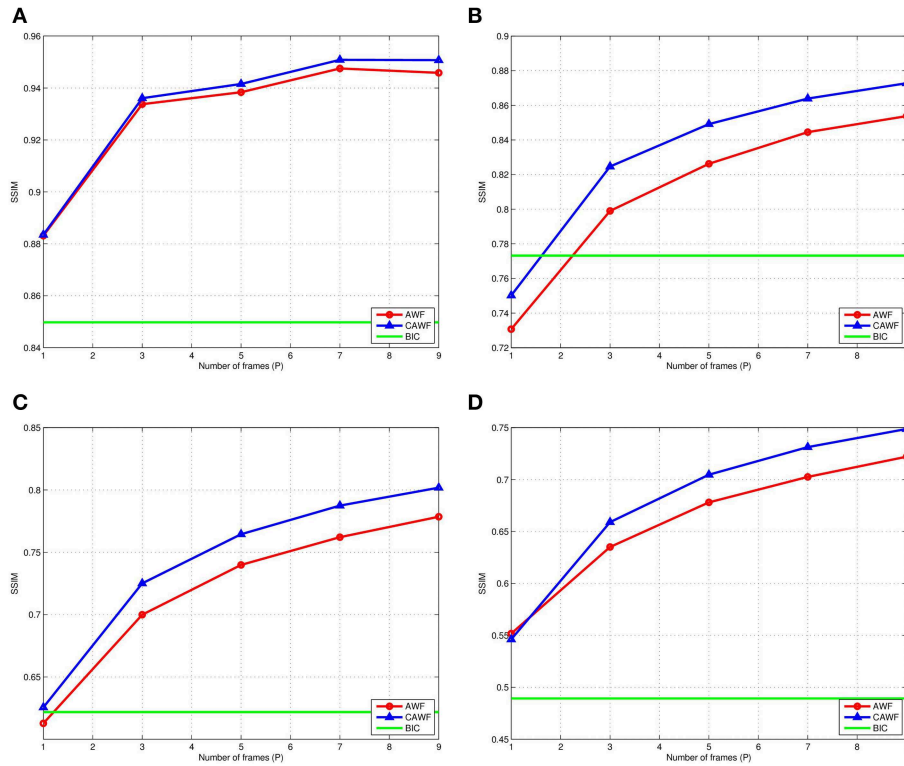


FIGURE 11 | The SSIM for AWF, CAWF, and the bicubic outputs with different levels of noise. (A) $\sigma_n = 1$, (B) $\sigma_n = 5$, (C) $\sigma_n = 10$, and (D) $\sigma_n = 15$. These results are obtained with $M = 10$ and the number of frames, P , varies from 1 to 9.

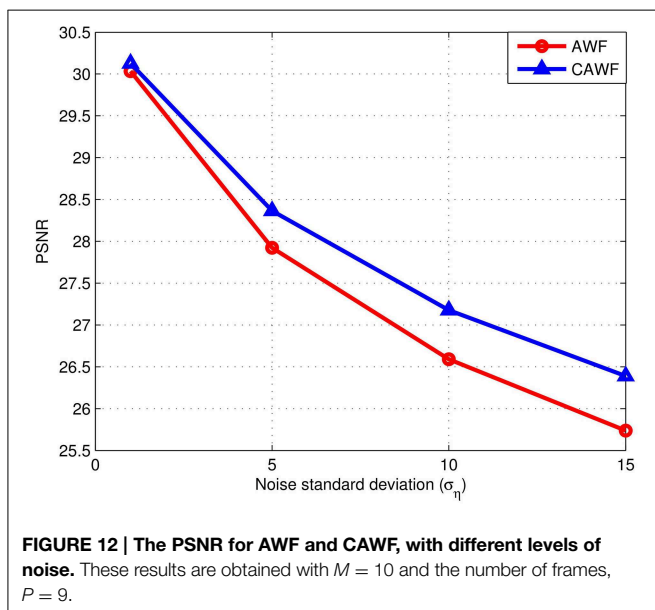


FIGURE 12 | The PSNR for AWF and CAWF, with different levels of noise. These results are obtained with $M = 10$ and the number of frames, $P = 9$.

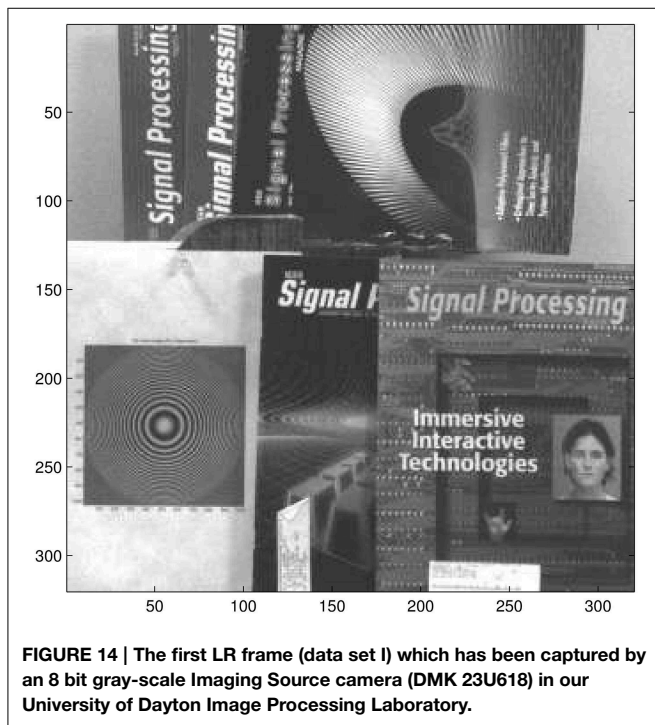
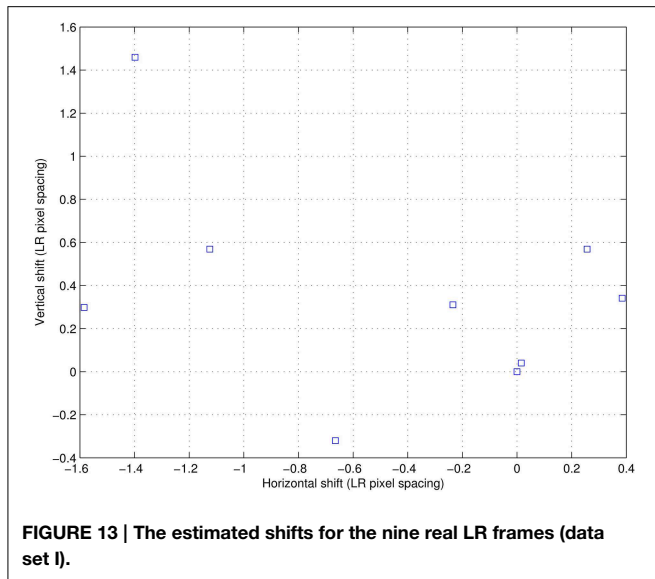
TABLE 5 | Quantitative performance of the bicubic, the AWF and the CAWF outputs for the Lighthouse image when $L = 3$, $P = 9$ and four levels of noise.

Method	Metric type	Noise standard deviation			
		$\sigma_n = 1$	$\sigma_n = 5$	$\sigma_n = 10$	$\sigma_n = 15$
Bicubic	psnr	25.55	25.18	24.19	22.93
	ssim	0.85	0.77	0.62	0.49
AWF	psnr	30.03	27.92	26.59	25.74
	ssim	0.95	0.85	0.78	0.72
CAWF	psnr	30.13	28.36	27.21	26.39
	ssim	0.96	0.87	0.81	0.75

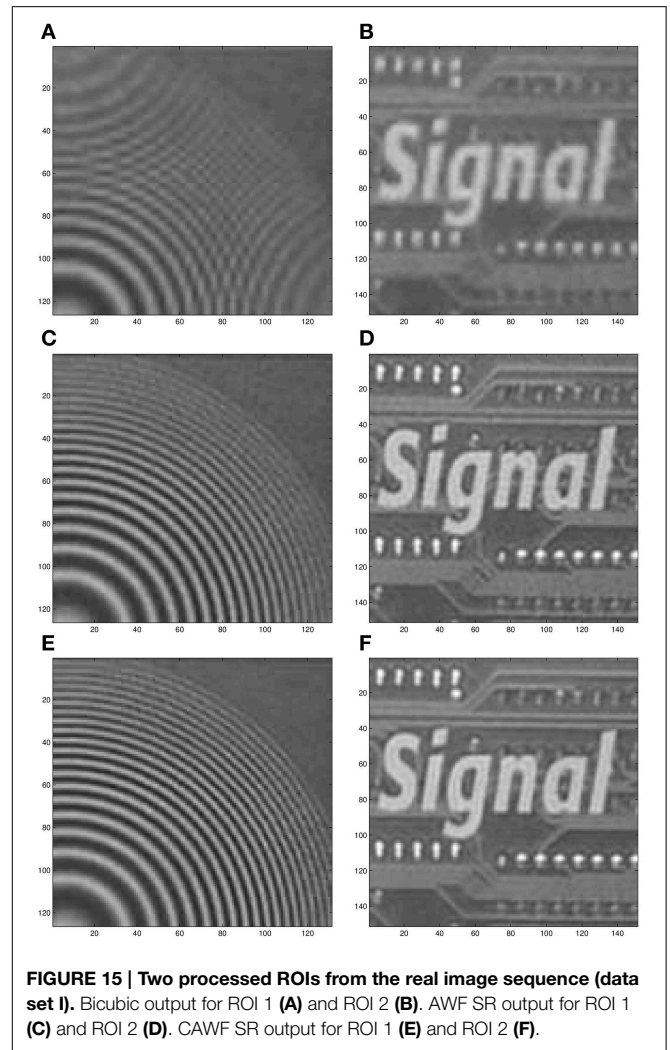
Bold numbers means the "highest performance."

Note the weight matrix, \mathbf{W}_i , is of size $KM \times L^2$ where each column has the weights to estimate one particular HR pixel in the i 'th estimation window. Before applying the weights, we follow the approach in Hardie [8] and normalize each column of weights to

sum to one to avoid possibly creating a grid artifact if processing non-zero mean data. Using the model to compute the required correlations, note that there are two key tuning parameters. One is the auto-correlation decay, ρ , which controls the correlation between pixels within a given patch used to form \mathbf{R} and \mathbf{P} . The other parameter is the patch similarity decay, α , which controls the correlation between patches. The SNR also plays a role in shaping the correlations and ultimately the weights. Note that for a high SNR, the CAWF filter is more aggressive and relies less

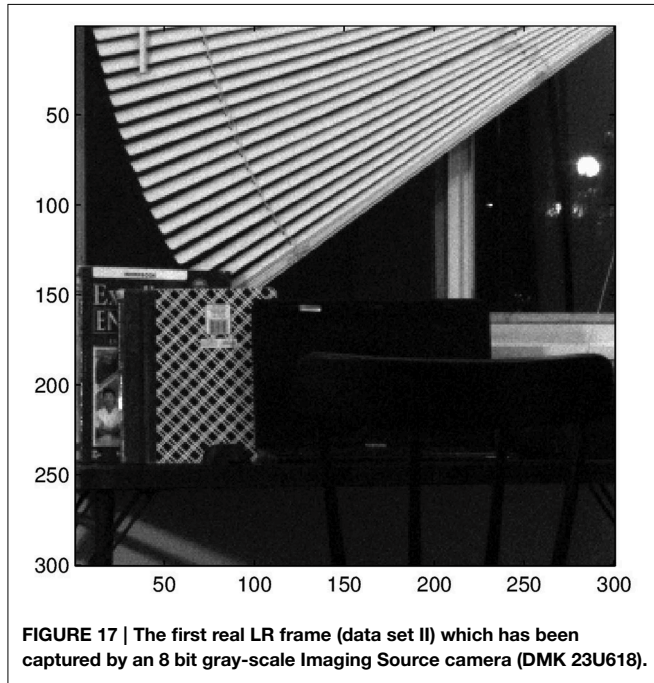
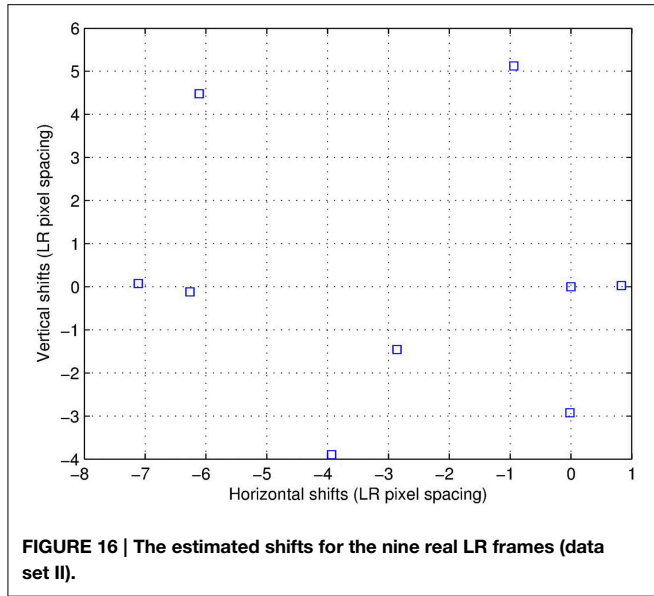


on spatial correlation for noise reduction. In low SNR imagery, the opposite is true. Also, note that when highly similar patches are found, as reflected in the distance matrix, the weights adapt to this and provide more aggressive SR processing. When no similar patches are found, the CAWF effectively reverts to the AWF method, primarily weighting only the reference patch. The relationship between the distance matrix and weights is explored more thoroughly for image restoration in Mohamed and Hardie [17].



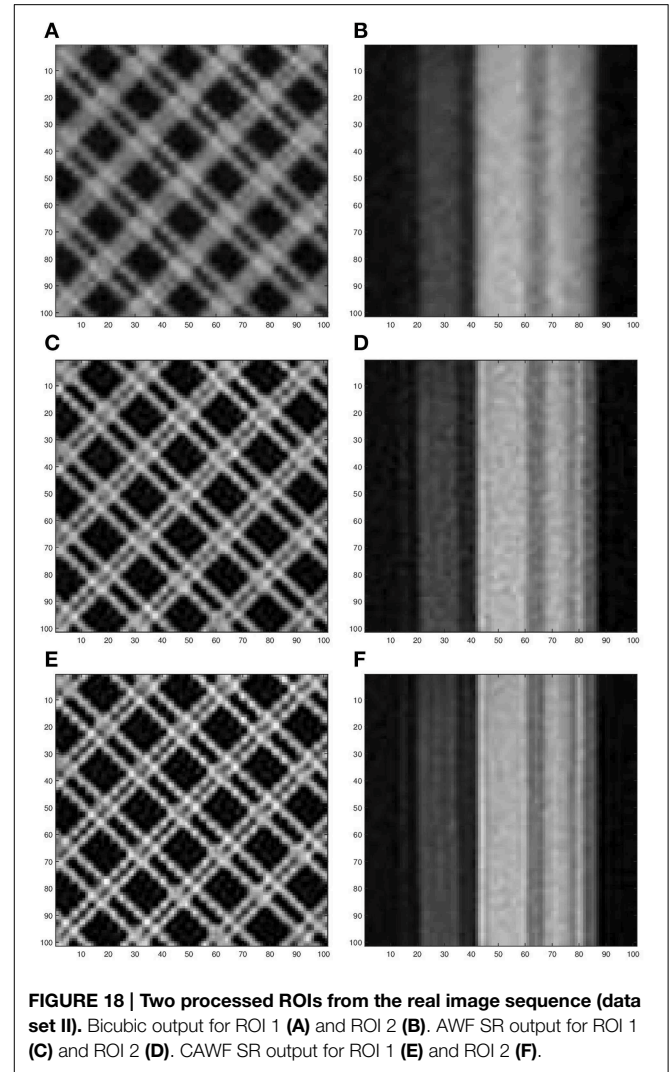
4. Results and Discussion

In this section, the efficacy of the CAWF SR method is demonstrated. We compare against a single frame bicubic interpolation and the AWF SR method, in order to demonstrate the advantages of multi-patch processing. Note that AWF SR is equivalent to CAWF SR when $M = 1$. We present results with both simulated data and real data. The real data represent raw data directly from the imaging sensor. In the case of the simulated data, we begin with a very high resolution and high signal-to-noise ratio image (not from the the sensor used for the real data). We then artificially degrade this ideal image with simulated motion, PSF blur (we follow the approach in Hardie [8] and modeled the diffraction from a circular exit pupil and spatial detector integration), downsampling, and noise, according to the model in Figure 1. The simulation is designed to mimic the real sensor to the greatest degree possible. That is, we simulate the same detector pitch, f-number, downsampling factor, and system PSF, as the real camera. All of the processing, simulation, and sensor parameters used in the are provided in Tables 2, 3.



The simulated data allow for a quantitative performance analysis based on objective truth.

The real camera data are acquired with an 8 bit gray-scale imaging source camera model DMK 23U618. **Table 3** presents the camera specifications [23]. We assume a wavelength of $\lambda = 0.55 \mu\text{m}$ and employ $f/5.6$ optics. Based on Equation (3), the Nyquist detector pitch would be $\delta_s = 1.54 \mu\text{m}$. Thus, the camera, with its $5.6 \mu\text{m}$ detector pitch, is under-sampled by a factor of 3.64. Because of the optical transfer function, there tends to be relatively little energy near the SR folding frequency. Thus, in practice, we find that good results can be obtained with a slightly



smaller upsampling factor of $L = 3$ and this is what we use for the presented results. For both the simulated data and real data, we use the same camera specifications to model the PSF and use $L = 3$ for downsampling and for SR upsampling.

With regard to performance metrics, we use a subjective visual assessment for the real data. For the simulated data, we use the peak signal-to-noise ratio (PSNR) metric, which is defined as

$$PSNR(\mathbf{z}, \hat{\mathbf{z}}) = 10 \log_{10} \left(\frac{255^2}{\frac{1}{N} \sum_{i=1}^N (z_i - \hat{z}_i)^2} \right). \quad (18)$$

We also use the structure similarity index (SSIM), which is defined in Wang et al. [24].

4.1. Simulated Data

The first experiment in the simulated data results is a visual demonstration of the CAWF performance. We use

the uncompressed 8 bit Kodak test image “Lighthouse” (Kodim19) [25] which we have converted to grayscale. We utilize the parameter values in **Table 1** and follow the observation model in **Figure 1** to get the LR frames. We blur the Lighthouse image with the prescribed PSF, shift it by random amounts, down samples these images by $L = 3$ and add Gaussian noise with $\sigma_\eta = 10$. We generate a total of $P = 9$ LR frames in this manner. **Figure 4** shows the nine truth translational shifts as well as the estimated ones. The HR Lighthouse image is shown in **Figure 5A** and the first LR frame is shown in **Figure 5B**. The results for single frame bicubic interpolation, AWF SR and CAWF SR are shown for two regions of interest (ROIs) in **Figures 6, 7**. The first and second HR truth ROIs are shown in **Figures 6A, 7A**, respectively. The outputs of the bicubic interpolation for both ROIs are shown in **Figures 6B, 7B**. The effects of noise can be seen as “blotches” in the flat areas of the image. Aliasing and blurring are clearly evident along the fence. The AWF outputs are shown in **Figures 6C, 7C**, and the CAWF outputs are shown in **Figures 6D, 7D**. Examining the the outputs of AWF and CAWF, it is clear that both methods provide better resolution than bicubic interpolation. However, the CAWF method appears to provide better noise suppression. Quantitative results using seven Kodak test images are provided in **Table 4**. We use the same PSF as before with $L = 3$, $P = 9$, and $\sigma_\eta = 10$. Both PSNR and SSIM results are provided. Note that in average the CAWF PSNR is higher than the AWF PSNR by approximately 0.5 dB, and the SSIM is greater by approximately 0.04 (Bold numbers means the highest performance).

The next experiment shows that the improvement in performance seen with CAWF is observed even with many noise and shift realizations. In particular, we use the lighthouse image and simulate 50 sets of LR frames. Each set is obtained with different realization of random intra-frame shifts and noise samples. The results are shown in **Figure 8**, with PSNR shown in **Figure 8A**, and SSIM shown in **Figure 8B**. From **Figure 8A**, the means of PSNR are 27.15, 26.55, and 24.20 for CAWF, AWF, and single frame bicubic interpolation, respectively. The standard deviations of PSNR for the same methods are 0.0602, 0.0550, and 0.0102. Also, the means of SSIM shown in **Figure 8B** are 0.81, 0.78, and 0.62 for CAWF, AWF, and single frame bicubic interpolation, respectively. The respective standard deviations of SSIM are 0.0021, 0.0012, and 0.0015.

Results are shown in **Figure 9** that illustrate the impact of changing the number of similar patches, M , in the CAWF SR method using the lighthouse image (Kodim19). The PSNR and SSIM of the CAWF SR outputs are shown in **Figures 9A,B**, respectively. For these results, we again use the parameters listed in **Table 2**, but vary M from 1 to 17. Note that $M = 1$ means no patch similarity is utilized, and the CAWF method reduces to the AWF. Note that PSNR increase significantly with M for $M \leq 10$, and there appears to be a “knee” in the curve near $M = 10$. It is for this reason that we have elected to use $M = 10$ in the previous results. A similar patterns is seen with SSIM in **Figure 9B**.

In a final simulated data experiment, we study the impact of noise level and the number of LR input frames used on the AWF and CAWF SR methods. These results are presented in **Figures 10, 11**. The lighthouse image is used with the parameters

in **Table 2**, but with P ranging from 1 to 9. The PSNR results for $\sigma_\eta^2 = 1, 25, 100$ and 225 are shown in **Figures 9A–D**, respectively. A similar set of results, but with the SSIM metric, are shown in **Figure 11**. Clearly, increasing the number of input frame improves performance for both CAWF and AWF. Also, the relative improvement increases with noise level. This trend is shown in **Figure 12**. Note that CAWF SR consistently outperforms AWF SR in these results. **Table 5** shows quantitative results for the bicubic, AWF and CAWF for the four levels of noise when $P = 9$. We conclude these results from **Figures 10, Figure11**.

4.2. Real Data

In this section, two sets of real data are processed and presented. In both data sets, we get the LR frames by panning the camera described in **Table 3** across a scene. We utilize nine frames as inputs ($P = 9$). We employ the parameters listed in **Table 2** for processing. Also we assume the noise is additive zero-mean Gaussian noise with variance of σ_η^2 . Using the sample variance in a flat region of the scene, we estimate the noise standard deviation for both data sets to be $\sigma_\eta = 5$.

The estimated intra-frame translational shifts for the first data set are shown in **Figure 13**. The first full (320×320) LR frame from this dataset is shown in **Figure 14**. Aliasing can be seen in this imagery, especially in the chirp pattern shown in the left lower corner. To better illustrate the SR results, two ROIs are selected. These processed ROIs are shown in **Figure 15**. The bicubic results are shown in **Figures 15A,B**. The AWF results are shown in **Figures 15C,D**. Finally, the CAWF results are shown in **Figures 15E,F**. As with the simulated data, the AWF and CAWF methods appear far better than the bicubic interpolation in terms of aliasing reduction and resolution enhancement. However, the CAWF outputs appear to show more detail than the AWF outputs do. This can be seen in the high frequency regions of the chirp ROI and the parallel lines above and below the word “signal.”

For the second real data set, the estimated intra-frame translational shifts are shown in **Figure 16**. **Figure 17** shows the first full (300×300) LR frame from this second dataset. Also two ROIs are chosen, one shows a portion of the notebook in the scene, and the other is part of the window mullion. These processed ROIs are shown in **Figure 18**. The bicubic results are shown in **Figures 18A,B**. The AWF results are shown in **Figures 18C,D**. Finally, the CAWF results are shown in **Figures 18E,F**. We believe that the performance of CAWF SR method using the real data is consistent with that obtained with the simulated data, and that CAWF SR does offer a notable improvement over AWF SR.

5. Conclusion

We have proposed a novel CAWF method for multi-frame SR. The first step of the CAWF SR method is to register all LR frames one a common HR grid. Then, for each LR pixel position on the HR grid, we identify the M most similar patches within a given search window about the reference patch. A single-stage weighted sum of the pixels in the similar patches is used to estimate the

$L \times L$ HR pixels in estimate window at the reference patch. The weights are based on a new nonuniformly sampled multi-patch correlation model. The CAWF SR method can be viewed as an extension of AWF SR [8] using mutli-patch processing. Using multiple similar patches to form the estimate allows the CAWF technique to exploit self-similarity to greatly increase the noise robustness over the AWF SR method. In light of the CAWF method, The AWF method can be viewed as a special case where $M = 1$ and only the reference patch is included in the weighted sum to form the estimate. Like the AWF SR, CAWF SR can perform nonuniform interpolation, deblurring, and denoising jointly. The big advantage of the CAWF, vs. the AWF, is improved performance in low SNR data. From the real and the simulated

image data, the CAWF SR outperforms AWF SR for all scenarios and cases tested. Although the most benefit is seen when the noise level is high. We believe the CAWF SR approach might be beneficial in several other applications as well, including those where the AWF is successful [9–13].

Author Contributions

KM coded the CAWF SR algorithm, collected all of the data used here, and prepared all of the experimental results. RH suggested the CAWF SR approach as an extension of the AWF SR method. Both KM and RH contributed to the writing and editing of the current manuscript.

References

- Hou HS, Andrews H. Cubic splines for image interpolation and digital filtering. *IEEE Trans Acoust Speech Signal Process.* (1978) **26**:508–17. doi: 10.1109/TASSP.1978.1163154
- Nasir H, Stankovi V, Marshall S. Singular value decomposition based fusion for super-resolution image reconstruction. *Signal Process.* (2012) **27**:180–91. doi: 10.1016/j.image.2011.12.002
- Szydzik T, Callico GM, Nunez A. Efficient FPGA implementation of a high-quality super-resolution algorithm with real-time performance. *IEEE Trans Consum Electron.* (2011) **57**:664–72. doi: 10.1109/TCE.2011.5955206
- Song BC, Jeong SC, Choi Y. Video super-resolution algorithm using bi-directional overlapped block motion compensation and on-the-fly dictionary training. *IEEE Trans Circuits Syst Video Technol.* (2011) **21**:274–85. doi: 10.1109/TCSVT.2010.2087454
- Mancas-Thillou C, Mirmehdi M. An introduction to super-resolution text. In: Chaudhuri B, editor. *Digital Document Processing. Advances in Pattern Recognition.* London: Springer (2007). p. 305–27.
- Lin F, Fookes C, Chandran V, Sridharan S. Super-resolved faces for improved face recognition from surveillance video. In: Lee SW, Li S, editors. *Advances in Biometrics. Vol. 4642, of Lecture Notes in Computer Science.* Berlin; Heidelberg: Springer (2007). p. 1–10.
- Barner KE, Sarhan AM, Hardie RC. Partition-based weighted sum filters for image restoration. *IEEE Trans Image Process.* (1999) **8**:740–5. doi: 10.1109/83.760341
- Hardie R. A fast image super-resolution algorithm using an adaptive wiener filter. *IEEE Trans Image Process.* (2007) **16**:2953–64. doi: 10.1109/TIP.2007.909416
- Hardie RC, LeMaster DA, Ratliff BM. Super-resolution for imagery from integrated microgrid polarimeters. *Opt Express.* (2011) **19**:12937–60. doi: 10.1364/OE.19.012937
- Narayanan BN, Hardie RC, and Balster EJ. Multiframe adaptive Wiener filter super-resolution with JPEG2000-compressed images. *EURASIP J Adv Signal Process.* (2014) **2014**:55. doi: 10.1186/1687-6180-2014-55
- Karch BK, Hardie RC. Adaptive Wiener filter super-resolution of color filter array images. *Opt Express.* (2013) **21**:18820–41. doi: 10.1364/OE.21.018820
- Hardie RC, Barnard KJ. Fast super-resolution using an adaptive Wiener filter with robustness to local motion. *Opt Express.* (2012) **20**:21053–73. doi: 10.1364/OE.20.021053
- Hardie RC, Barnard KJ, Ordenez R. Fast super-resolution with affine motion using an adaptive wiener filter and its application to airborne imaging. *Opt Express.* (2011) **19**:26208–31. doi: 10.1364/OE.19.026208
- Milanfar P. Super-Resolution imaging. In: *Digital Imaging and Computer Vision.* CRC Press (2011). Available online at: <http://books.google.com/books?id=fjTubMnvOkqC>
- Danielyan A, Foi A, Katkovnik V, Egiazarian K. Image and video super-resolution via spatially adaptive blockmatching filtering. In: *Proceeding of International Workshop on Local and Non-local Approximation in Image Processing (LNLA).* Lausanne (2008).
- Dabov K, Foi A, Egiazarian K. Video denoising by sparse 3D transform-domain collaborative filtering. In: *15th European Signal Processing Conference (EUSIPCO).* Poland: Poznan (2007).
- Mohamed K, Hardie R. A collaborative adaptive Wiener filter for image restoration using a spatial-domain multi-patch correlation model. *EURASIP J Adv Signal Process.* (2015) **2015**:6. doi: 10.1186/s13634-014-0189-3
- Hardie RC, Barnard KJ, Bognar JG, Armstrong EE, Watson EA. High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system. *Opt Eng.* (1998) **37**:247–60. doi: 10.1117/1.601623
- Fiete RD. Image quality and FN/p for remote sensing systems. *Opt Eng.* (1999) **38**:1229–40. doi: 10.1117/1.602169
- Hardie R, Barnard K, Bognar J, Armstrong E, Watson E. High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system. *Opt Eng.* (1998) **37**:247–60. doi: 10.1117/1.601623
- Irani M, Peleg S. Improving Resolution by Image Registration. *CVGIP Graph Models Image Process.* (1991) **53**:231–9. doi: 10.1016/1049-9652(91)90045-L
- Lucas BD, Kanade T. An iterative image registration technique with an application to stereo vision. In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Vol. 2, IJCAI'81.* San Francisco, CA: Morgan Kaufmann Publishers Inc. (1981). p. 674–9.
- Sony. *The Imaging Source Technology Based on Standards* (2014). Available online at: http://www.theimagingsource.com/en_US/products/cameras/usb-cmos-ccd-mono/dmk23u618/
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* (2004) **13**:600–12. doi: 10.1109/TIP.2003.819861
- Kodak. *Kodak Lossless True Color Image Suite* (1999). Available online at: <http://r0k.us/graphics/kodak/>. [Online: Accessed 15-Nov-1999].

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Mohamed and Hardie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.