

RESEARCH

Open Access

Multiframe adaptive Wiener filter super-resolution with JPEG2000-compressed images

Barath Narayanan Narayanan*, Russell C Hardie and Eric J Balster

Abstract

Historically, Joint Photographic Experts Group 2000 (JPEG2000) image compression and multiframe super-resolution (SR) image processing techniques have evolved separately. In this paper, we propose and compare novel processing architectures for applying multiframe SR with JPEG2000 compression. We propose a modified adaptive Wiener filter (AWF) SR method and study its performance as JPEG2000 is incorporated in different ways. In particular, we perform compression prior to SR and compare this to compression after SR. We also compare both independent-frame compression and difference-frame compression approaches. We find that some of the SR artifacts that result from compression can be reduced by decreasing the assumed global signal-to-noise ratio (SNR) for the AWF SR method. We also propose a novel spatially adaptive SNR estimate for the AWF designed to compensate for the spatially varying compression artifacts in the input frames. The experimental results include the use of simulated imagery for quantitative analysis. We also include real-video results for subjective analysis.

Keywords: Super-resolution; JPEG2000 compression; Adaptive Wiener filter; Spatially adaptive

1 Introduction

Multiframe super-resolution (SR) is a post processing technique designed to reduce aliasing and enhance resolution for detector-limited imaging systems [1]. As described in [2,3], SR methods generally fuse a set of low-resolution (LR) frames with a common field of view to form a high-resolution (HR) image with reduced aliasing. SR methods assume the presence of motion between the frames that can be estimated or known with sub-pixel accuracy. The motion allows each frame to capture certain unique samples of the scene, effectively increasing the sampling frequency of the imaging sensor. SR can be applied to produce a single output frame or produce video output by employing a moving temporal window of frames [4]. SR techniques have proven to be highly successful in providing meaningful resolution enhancement for images and videos under appropriate conditions.

SR research has grown significantly in recent years [1]. However, the majority of SR research focuses on the use of raw uncompressed image data obtained directly from an imaging sensor. But in many practical imaging applications, the acquired video frames must be stored using

limited file size or compressed in order to be transmitted through a band-limited channel. One such powerful compression method is Joint Photographic Experts Group 2000 (JPEG2000) compression [5]. Recently, studies have shown that JPEG2000 is a good choice for high-quality and high-resolution videos [6]. In 2004, the motion picture industry, specifically Digital Cinema Initiatives, announced JPEG2000 as the standard for digital delivery of all motion pictures [7]. In light of this, as well as the emergence of some important new classes of SR algorithms, important questions are raised regarding how to best incorporate the benefits of both SR and JPEG2000 image compression. For example, how does compression before SR compare to compression applied after SR? Also, how does SR performance degrade with compression ratio using JPEG2000? What can be done to improve the performance of SR methods with JPEG2000? We shall attempt to address these and other questions here.

In previous work, some SR techniques have been applied to compressed imagery and video signals. For video signals, most of the work focuses on processing LR imagery that has been compressed using Moving Picture Experts Group (MPEG) and H.26x methods. In [8,9], a compression method of sub-sampling video prior to MPEG4 compression and SR-applied post decompression

* Correspondence: narayananb1@udayton.edu
Department of Electrical and Computer Engineering, 300 College park,
University of Dayton, Dayton, Ohio 45419, USA

shows signal-to-noise ratio (SNR) improvement over merely MPEG4 compressing the original full resolution signal. In [10], P- and B-frames in an H.264 compression technique are sub-sampled prior to compression. At the decompression end, SR is used to resolve the P- and B-frames using the I-frames as the training samples. In addition to compressed video, SR techniques have been applied to independently compressed images. In [11], SR is applied to imagery after JPEG compression and shows degradation to the SR performance when compression is high. In [12,13], SR techniques designed to be robust to JPEG compression artifacts are described. A downsampling-based video coding method is proposed in [14]. There, SR is used to restore the downsampled frames to their original resolutions. An SR algorithm specifically for video with atmospheric turbulence and MPEG-4 compression is described in [15]. In [16,17], SR techniques are applied in the transform domain as part of compression/decompression. Finally, the work in [18] describes an SR technique designed to restore details in imagery that have been degraded due to JPEG2000 compression. With the exception of [18], the combination of SR and JPEG2000 has not been widely studied. We believe this is an important area to investigate.

In this paper, we provide a novel study of the performance of SR with JPEG2000 compression. We employ relatively new SR technique based on the adaptive Wiener filter (AWF). The AWF is a computationally efficient SR method, suitable for real-time implementation, with generally good performance [4]. This method has been selected because of its computational simplicity and best-in-class performance, as demonstrated in [4]. Here, we investigate several architectures for combining AWF SR with JPEG2000 compression. These include systems that apply compression prior to SR and compression after SR processing. We also investigate the use of individual-frame compression, as well as motion-compensated difference-frame compression. We study how SR performance is impacted by a wide range of compression ratios (CRs). Based on our findings, we make some practical and important recommendations and observations regarding the joint use of JPEG2000 and SR. We also show that by modifying the SNR present in the correlation model used by the AWF SR method, the compression artifacts can be better tolerated. Furthermore, a novel spatially varying SNR model is proposed and demonstrated to specifically target the adverse effects of spatially varying compression artifacts.

Applications that we believe are well suited to the joint use of SR and compression include airborne and satellite imaging [19-22]. In these applications, the dominant inter-frame motion is the result of camera platform motion. Thus, the motion can be well modeled with a global motion model. This allows for accurate sub-pixel motion

estimation for super-resolution. There is also a strong need for compression in these applications, in order to store and transmit the acquired data through band-limited channels. We have observed that a video that is well suited to multiframe SR is also likely a good candidate for difference-frame compression. In this case, the correlation between registered frames is exploited for compression, and the high-frequency differences are exploited for SR.

The remainder of this paper is organized as follows. Section 2 presents the basic AWF SR algorithm along with the proposed spatially varying SNR estimation method. Several architectures for combining AWF SR and JPEG2000 compression are presented in Section 3. The experimental results are provided in Section 4. Finally, conclusions are offered in Section 5.

2 AWF SR algorithm

The AWF SR method is introduced in [4]. We provide some of the key algorithm details here for the reader's convenience. We begin with the observation model and then describe the AWF SR algorithm.

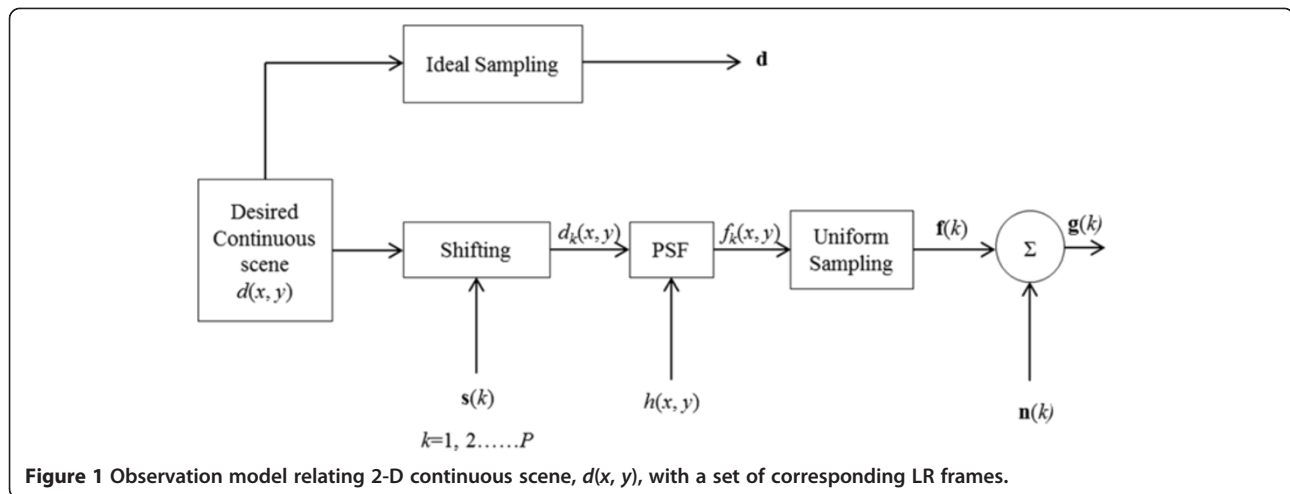
2.1 Observation model

The AWF SR method is based on the observation model depicted in Figure 1. Here, we have P LR frames that are related to the desired continuous scene, $d(x, y)$, through a shift, point spread function (PSF) blur, and sampling with additive noise as shown. In the case of translational motion, the shift and PSF models commute, allowing us to equivalently use the observation model in Figure 2. Here, we have switched the motion and PSF blocks and combined the motion and sampling steps into a single nonuniform sampling operation. The details on the commutation of the motion and PSF blur are addressed in [20]. The PSF blur in Figure 2 yields the intermediate image

$$f(x, y) = d(x, y) * h(x, y), \quad (1)$$

where $h(x, y)$ is the PSF and $*$ is 2D convolution. This blurred image is assumed to be sampled nonuniformly based on the motion parameters and the detector pitch of the focal plane array (FPA) as described in [4]. These samples are represented in lexicographical notation as $\mathbf{f} = [f_1, f_2, \dots, f_N]^T$. With additive noise, these samples are given by $\mathbf{g} = \mathbf{f} + \mathbf{n}$, where \mathbf{n} is an $N \times 1$ array of noise samples. The noise will be assumed to be zero-mean independent and identically distributed Gaussian noise with a variance of σ_n^2 .

To model the PSF, we follow the approach in [23] that models diffraction and detector integration. For diffraction-limited optics, the spatial cut-off frequency is $\omega_c = 1/(\lambda \times f$ -number). Here, λ is the wavelength of light, and f -number is the ratio of the focal length to the effective diameter of the optics. Another critical parameter for an imaging system is the detector pitch, p , for the focal plane array. The



pitch is the spacing between the detector elements, and one divided by the pitch is the sampling frequency. The Nyquist criterion dictates that to avoid aliasing, we must have $1/p > 2\omega_c$. However, due to a complex trade space in imaging system design, many imaging systems do not meet the Nyquist criterion. This issue is addressed well in [24]. Resolution in such undersampled imaging systems may be thought of as limited by the detector array, rather than optically limited [2]. These systems may benefit from multiframe SR processing, such as the AWF SR method.

2.2 AWF SR processing

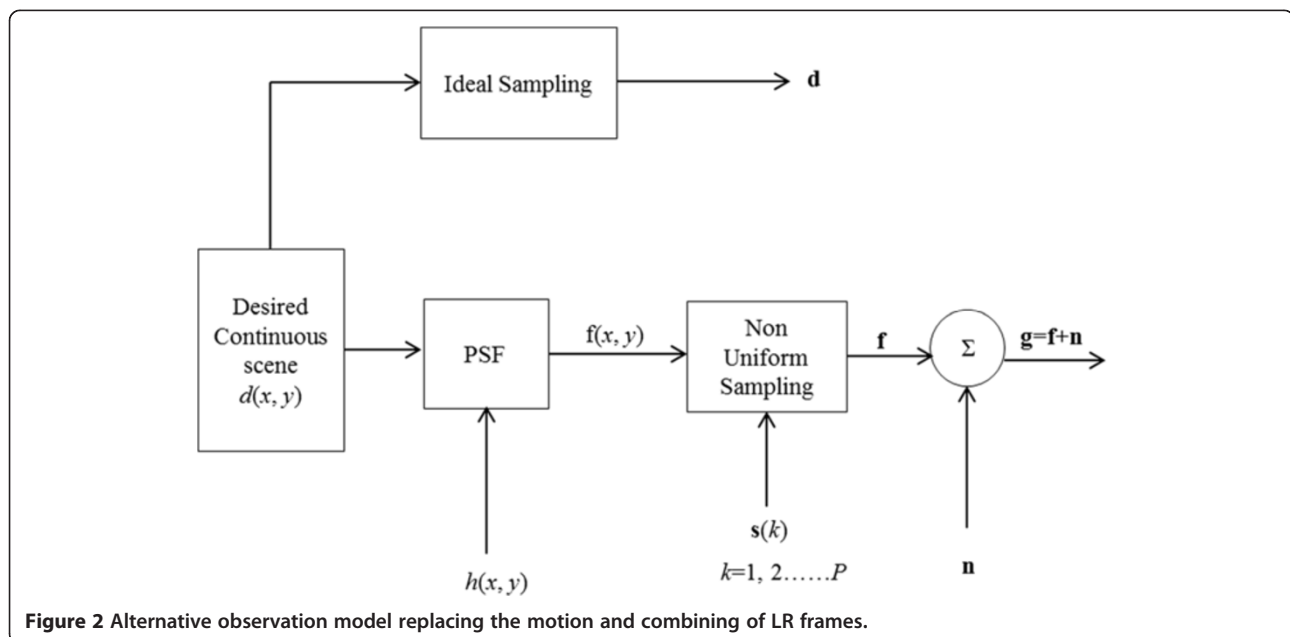
The AWF SR processing of the observed data to produce an HR image estimate is illustrated in Figure 3. Let the resulting HR image be upsampled by a factor of L in each spatial dimension relative to the LR input frames.

Registration is used to populate a common HR grid using samples from all of the LR frames, forming the HR array \mathbf{g} . The AWF SR algorithm then uses a moving window that passes over the HR grid of nonuniformly sampled data in \mathbf{g} . Let the samples in the small moving observation window about the i 'th HR output pixel be denoted \mathbf{g}_i . The output of the AWF is an estimate of the desired HR image and is given by a weighted sum for each HR pixel as follows:

$$\hat{d}_i = \mathbf{w}_i^T \mathbf{g}_i, \tag{2}$$

where \mathbf{w}_i is a vector of weights. The minimum mean squared error (MSE) weights are employed and these are given by

$$\mathbf{w}_i = \mathbf{R}_i^{-1} \mathbf{p}_i, \tag{3}$$



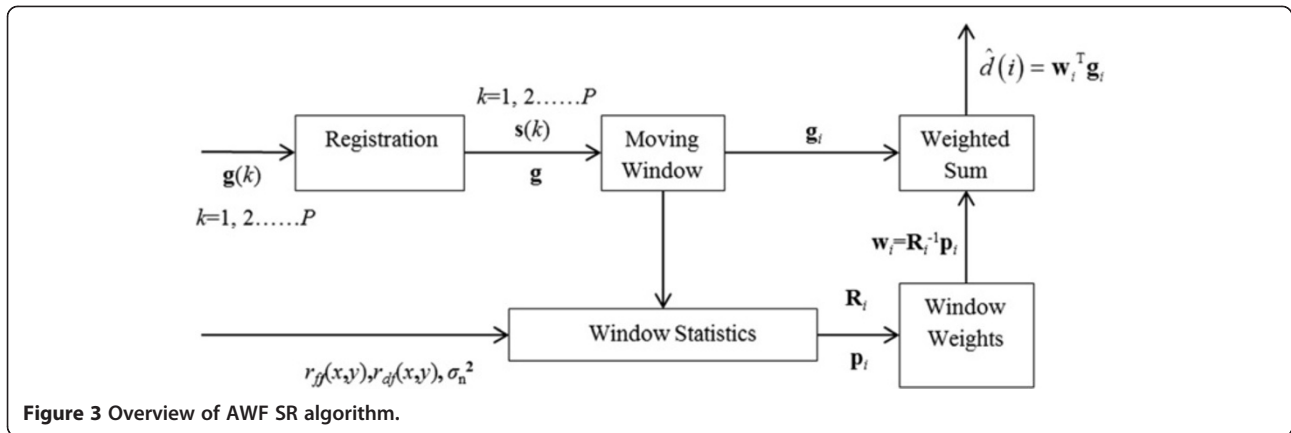


Figure 3 Overview of AWF SR algorithm.

where $\mathbf{R}_i = E\{\mathbf{g}_i \mathbf{g}_i^T\} = E\{\mathbf{f}_i \mathbf{f}_i^T\} + \sigma_n^2 \mathbf{I}$ is the autocorrelation matrix for the observation vector, and $\mathbf{p}_i = E\{d_i \mathbf{g}_i\} = E\{d_i \mathbf{f}_i\}$ is the cross-correlation vector between the desired sample and observation vector.

The correlations employed are model-based, and they vary with the spatial distribution of the samples in \mathbf{g}_i . In particular, the continuous desired image autocorrelation is assumed to be

$$r_{dd}(x, y) = \sigma_d^2 \rho^{\sqrt{x^2 + y^2}}, \quad (4)$$

where σ_d^2 is the desired image variance, and ρ controls the drop in correlation as a function of distance. From this, it can be shown that the other key correlation functions can be computed as follows:

$$r_{df}(x, y) = r_{dd}(x, y) * h(x, y) \quad (5)$$

and

$$r_{ff}(x, y) = r_{dd}(x, y) * h(x, y) * h(-x, -y). \quad (6)$$

Based on the distances between the observed samples in \mathbf{g}_i as they appear on the HR grid, we evaluate Equation 6 to populate \mathbf{R}_i . Similarly, given the distances between the desired sample position and the observed samples, we evaluate Equation 5 and populate the vector \mathbf{p}_i . This allows us to compute \mathbf{w}_i using Equation 3. Note that by estimating the HR image in this fashion, the AWF is simultaneously performing nonuniform interpolation, deconvolution, and noise reduction, all with a single weighted sum operation. This sets it apart from other fast SR

methods that perform nonuniform interpolation and restoration as independent processing steps. The combined approach of the AWF has computational and performance robustness advantages [4].

2.3 Spatially varying SNR model

The AWF SR correlation model tuning parameters impacting performance are ρ and the SNR given by σ_d^2 / σ_n^2 . We will show that by decreasing the assumed global SNR, we are able to make the AWF SR more robust to the effects of compression. However, the compression artifacts tend to be spatially varying. For example, ringing artifacts are often produced in flat areas adjacent to strong edges. To better mitigate the impact of this kind of spatially varying artifact on AWF SR, we propose employing a correlation model with a spatially varying SNR.

Here, we estimate this local SNR in a novel manner. We first align the LR observed frames based on the SR registration and then average the frames. We then estimate the local variance using a Gaussian weighting function and treat this as the signal variance. Note that the frame averaging tends to reduce compression artifacts and noise, preventing those factors from falsely increasing the local signal estimate. We assume the noise variance is a global constant that is used as a tuning parameter. The ratio of estimated local signal variance to noise variance allows us to form a preliminary local SNR estimate. The final step is to filter this SNR array with a 5×5 2D minimum filter. This minimum filter

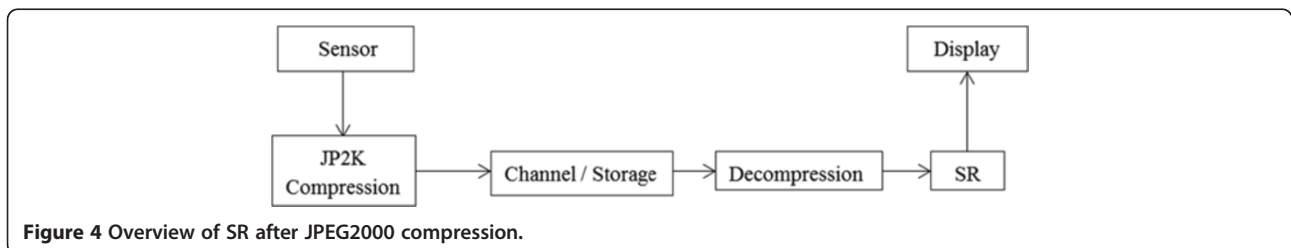
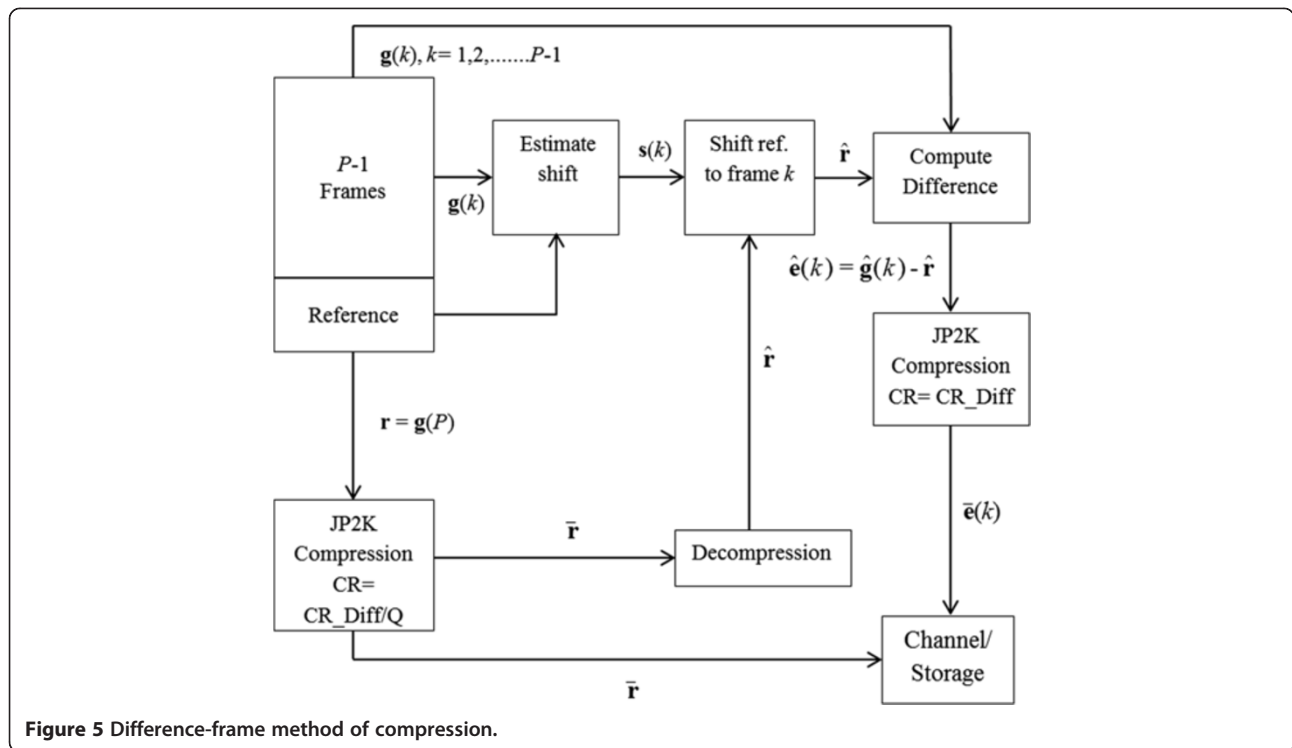


Figure 4 Overview of SR after JPEG2000 compression.



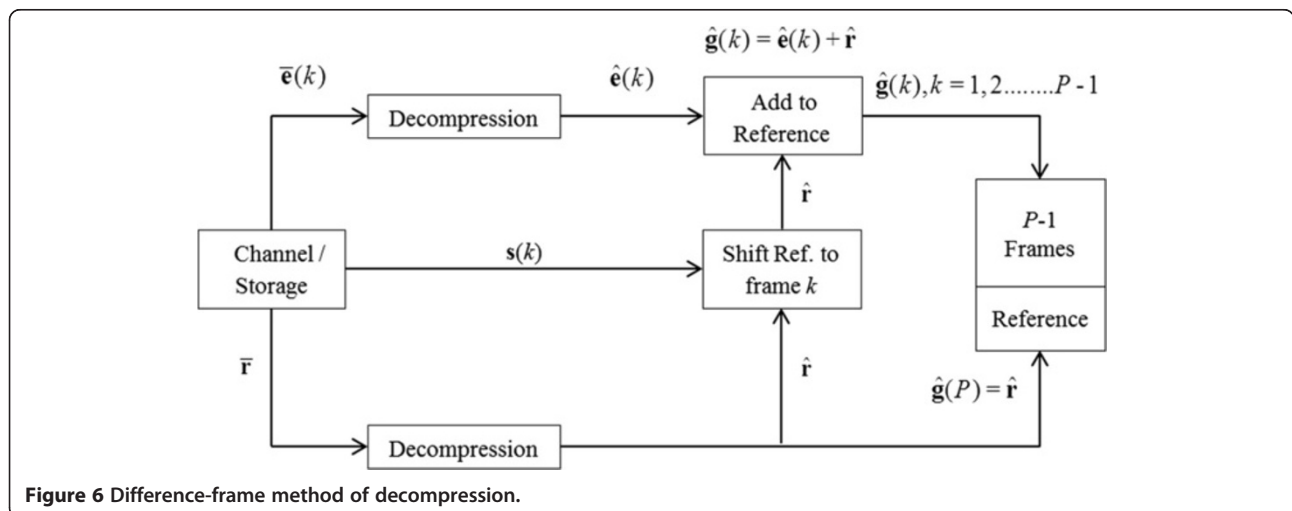
step tends to lower the SNR estimate in flat areas near edges that are most vulnerable to ringing artifacts from compression. Areas of dense texture and detail tend to maintain the high SNR.

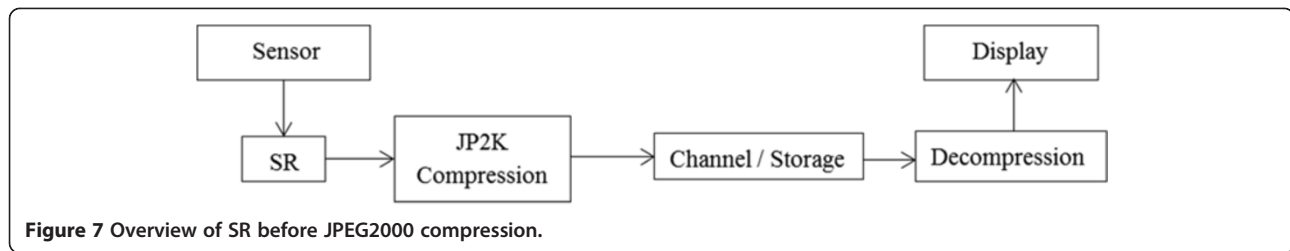
The estimated local SNR is then used with Equations 3 and 4. To keep the computational complexity low for the spatially adaptive AWF SR, we quantize the local SNR value estimated to $K = 20$ levels. The K distinct correlation models give rise to K sets of filter weights. These weights may be pre-computed prior to processing video frames. At each spatial location, the local SNR is

computed and the appropriate filter weight vector is applied.

3 Architectures for SR with JPEG2000 compression

There are number of ways to combine SR algorithm with JPEG2000 compression. One way is to apply AWF SR after performing compression on the LR input frames. We also consider performing AWF SR first and then compressing the resulting SR images. When using compression on multiple input frames, we consider both individual- and difference-frame methods as described below.





3.1 SR after JPEG2000 compression

The most common scenario is that the imagery from the sensor is compressed for storage and/or transmission immediately after it is acquired. Access to the raw uncompressed imagery for SR may not be possible. In this case, SR can only be applied after the compression, as illustrated in Figure 4. Thus, understanding the robustness of SR operating on such compressed imagery is an important problem.

3.1.1 Individual-frame method

Perhaps the most straightforward method for treating the multiframe input with compression is to compress each frame individually and independently. This allows each frame to be decompressed independently, providing an advantage over MPEG-X (1, 2, and 4), for example. We shall refer to this as the individual-frame method. All processing here is done using MATLAB (The MathWorks, Inc., Natick, MA, USA), and JPEG2000 is achieved by using the 9/7 transform, no quantization, and optimal truncation for rate control.

It should be noted that SR is most beneficial for significantly undersampled imaging systems where aliasing is present. For such an imaging system, the individual LR observed frames may not compress well because of high spatial frequency content. However, since a set of

frames suitable for SR must overlap in the field of view, these frames are also likely to exhibit inter-frame correlation. Thus, we also consider compression of registered difference frames, as described in the following subsection.

3.1.2 Difference-frame method

The difference-frame compression method is illustrated in the block diagram in Figure 5. Imagery suitable for multiframe SR is likely to exhibit a great deal of temporal correlation after registration. The global registration used for SR can also serve to aid in compression. Thus, we are proposing global difference-frame compression in this case, rather than the block matching optical flow vectors used in traditional video compression. In this method, the last (i.e., most recent) observed LR frame is considered to be the reference image, denoted as \mathbf{r} . Note that this reference frame is $\mathbf{g}(P)$ as shown in Figure 1. After JPEG2000 compression and then decompression, the reference frame is denoted $\bar{\mathbf{r}}$ and $\hat{\mathbf{r}}$, respectively. We set the CR for the reference image to be $1/Q$ times the CR used for the difference frames, where Q is a tuning parameter. Here, we have found $Q = 8$ to be a good choice.

Next, the decompressed reference frame is shifted to match each of the remaining $P - 1$ LR frames, and $P - 1$

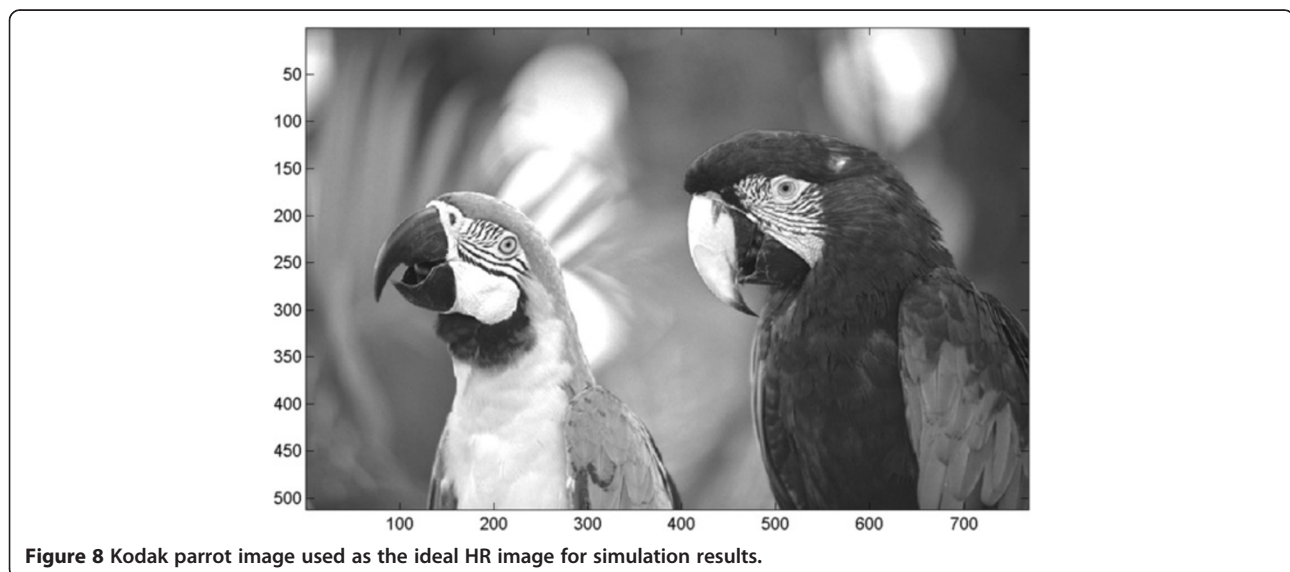




Figure 9 Simulated LR reference frame. This is after individual-frame JPEG2000 compression (CR = 8) and then $L = 4$ bicubic interpolation.

difference frames are computed and compressed. The shifts are estimated from the LR frames the same way they are for SR [4]. The difference frames are denoted as $\mathbf{e}(k)$ for $k = 1, 2, \dots, P - 1$. After compression, these difference frames are denoted as $\bar{\mathbf{e}}(k)$. To be stored and/or transmitted from this compression stage, we have the compressed reference, $\bar{\mathbf{r}}$, along with the compressed difference frames $\bar{\mathbf{e}}(k)$ and shifts $\mathbf{s}(k)$, for $k = 1, 2, \dots, P - 1$.

The process of decompression for the difference frame method is illustrated in Figure 6. The process begins by decompressing the reference and difference frames as shown. Next, the decompressed reference is added to the decompressed difference frames to recover the individual frames. This is given by $\hat{\mathbf{g}}(k) = \hat{\mathbf{e}}(k) + \hat{\mathbf{r}}$, for $k = 1, 2, \dots, P - 1$. Finally, note that the last image is simply the reference and is given by $\hat{\mathbf{g}}(P) = \hat{\mathbf{r}}$.

3.2 JPEG2000 compression after the application of SR

The final architecture considered here applies JPEG2000 compression after SR. This is shown in Figure 7. In this mode, AWF SR is applied directly to the raw uncompressed imagery from the sensor. The resulting SR image is then compressed with JPEG2000 and later transmitted/stored. This mode would be most practicable when SR processing can be done in real time at the sensor. It is also possible to use this mode if SR is applied forensically to stored uncompressed data, the results of which are later disseminated in compressed format.

While the pixel dimensions of the SR image are increased, the aliasing is reduced, making the SR image generally easier to compress. Furthermore, the SR process gets the benefit of working on data with no compression artifacts. In our experiments, we combine



Figure 10 AWF SR with $L = 4$ is applied on the individually compressed LR images with CR = 8. The optimum SNR of 67.40 is used and $\rho = 0.7$.



Figure 11 AWF SR with $L = 4$ is applied on the individually compressed LR images with $CR = 8$. Spatially varying SNR is used and $\rho = 0.7$.

16 LR frames to produce a single SR image with upsampling in each dimension of $L = 4$. In a video-to-video application, the SR video frame rate may be the same as the LR video rate. In this case, we have L^2 the number of SR pixels as we do input pixels. However, the possibility of difference-frame compression on the SR video exists for enhanced compression of the SR video.

4 Experimental results

In this section, we present the experimental results that include simulated LR data and a real-video sequence. The simulated data allow for quantitative analysis, while the real data allow for a subjective comparison in a real application. In addition to comparing the architectures described in Section 3, we also examine the use of the

spatially varying SNR model, as described in Section 2.3, to treat the compression artifacts more robustly.

4.1 Simulated video data

In this section, we begin by presenting and describing the subjective image results and then we present the quantitative results.

4.1.1 Image results

The first set of simulated data is based on 8-bit parrot image from the Kodak database (Rochester, NY, USA) [25], which is shown in Figure 8. This grayscale image contains 512×768 pixels stored at 8 bits per pixel (bpp). We artificially degrade this image to simulate the observation model in Figure 1. In particular, we simulate $P = 16$

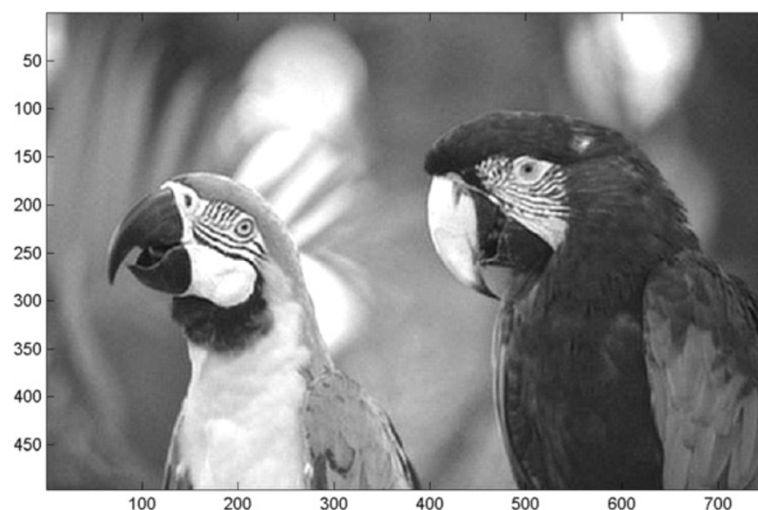


Figure 12 AWF SR with $L = 4$ is applied on the difference frame-compressed LR images with $CR = 8$. The optimum SNR of 315.67 is used and $\rho = 0.7$.



Figure 13 AWF SR with $L = 4$ is applied on the difference frame-compressed LR images with $CR = 8$. Spatially varying SNR is applied and $\rho = 0.7$.

LR frames with random translational shift, blur, and noise. The simulation PSF model is based on parameters matching the real-video data used in Section 4.2. The modeled f-number of the optics is 4, and the detector pitch in the both horizontal and vertical directions is 5.6 mm. We assume a 100 % fill factor [20] and wavelength is $0.55 \mu\text{m}$. The downsampling factor, relating the LR and SR image sizes, is $L = 4$. Finally, white Gaussian noise with a variance of 1 digital unit (DU) is added to simulate low-level electronics noise. Image results are shown in Figures 9,10,11,12,13,14, and quantitative results can be found in Figures 15,16,17 and in Tables 1,2,3.

We report compression level here using the CR value, which is defined as the ratio of input image file size to the output compressed file size. A higher value implies a

smaller file size and reduced image quality. Figure 9 shows the LR reference frame compressed with a $CR = 8$ and then $L = 4$ bicubic interpolation. This is one of the $P = 16$ LR frames generated using the individual-frame compression method as described in Section 3.1.1. The result of applying AWF SR with $L = 4$ on the individually compressed LR frames is shown in Figure 10. Here, $\rho = 0.7$ and the SNR used is 67.40, which maximizes the peak-signal-to-noise ratio (PSNR) [4]. The SR image looks notably sharper than the image in Figure 9. However, some artifacts are still present in the image. We have observed that the individual LR images are rather hard to compress due to aliasing. As a result of the shifts between LR frames, the compression artifacts tend to vary somewhat from frame to frame. This can lead to an

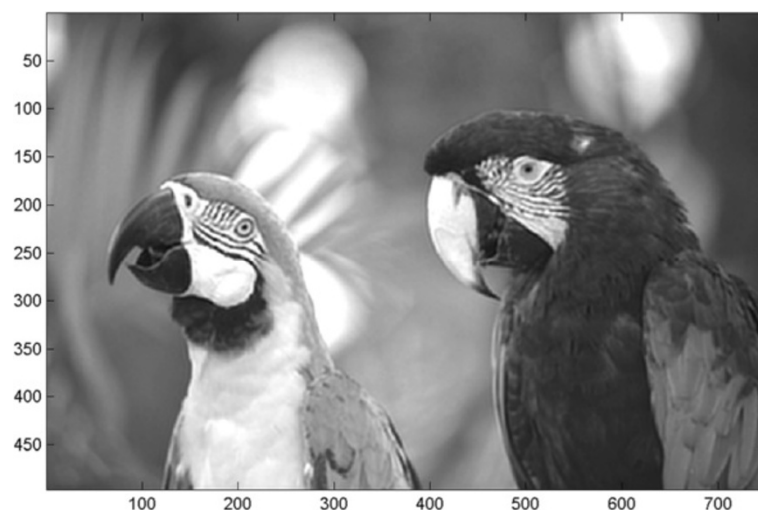


Figure 14 Individual-frame compression with $CR = 8$ applied after AWF SR with $L = 4$ is applied on LR images. The optimum SNR of 67.40 is used and $\rho = 0.7$.

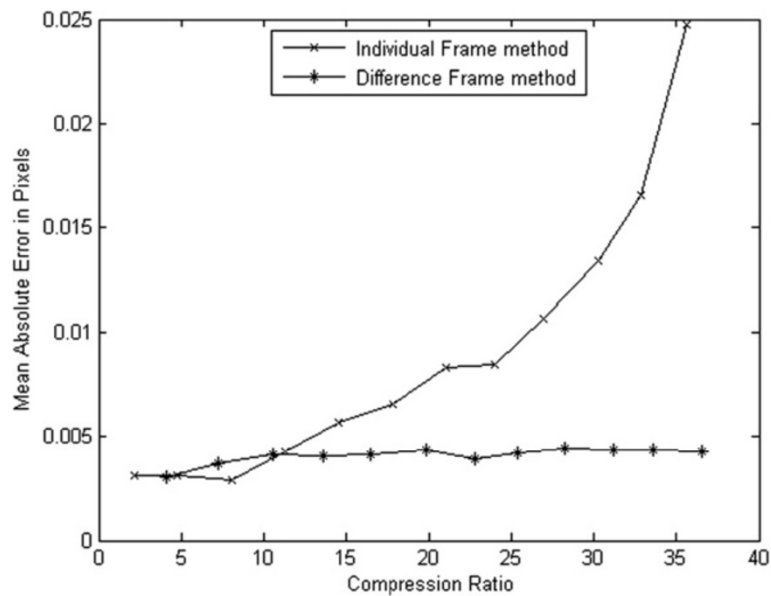


Figure 15 Registration MAE in pixels versus CR for individual- and difference-frame methods using Kodak parrot image.

SR image with artifacts as seen in Figure 10. Of course, one way to reduce these artifacts is to use a lower CR. Another approach is to use the spatially varying SNR model as described in Section 2.3. Figure 11 shows the AWF SR output image obtained using this spatially varying model. This result has lower error and, perhaps more importantly, tends to have visually reduced compression related artifacts.

The next result, shown in Figure 12, is for AWF SR applied to LR frames compressed with the difference-

frame compression described in Section 3.1.2. Note that in the difference-frame compression method, the group of $P = 16$ LR frames is set to have an overall CR of 8, to match the individual-frame method. The optimum SNR of 315.67 is used along with $\rho = 0.7$. This result appears far superior to that obtained using individual-frame compression (for the same overall CR). Here, the redundancy between LR frames is exploited by the difference-frame compression. At the same time, the differences among the LR frames are exploited by SR. We believe

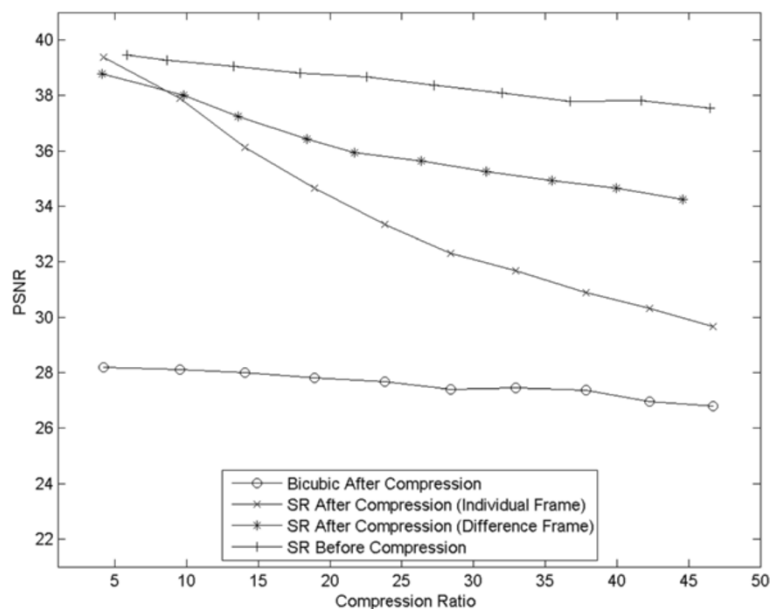
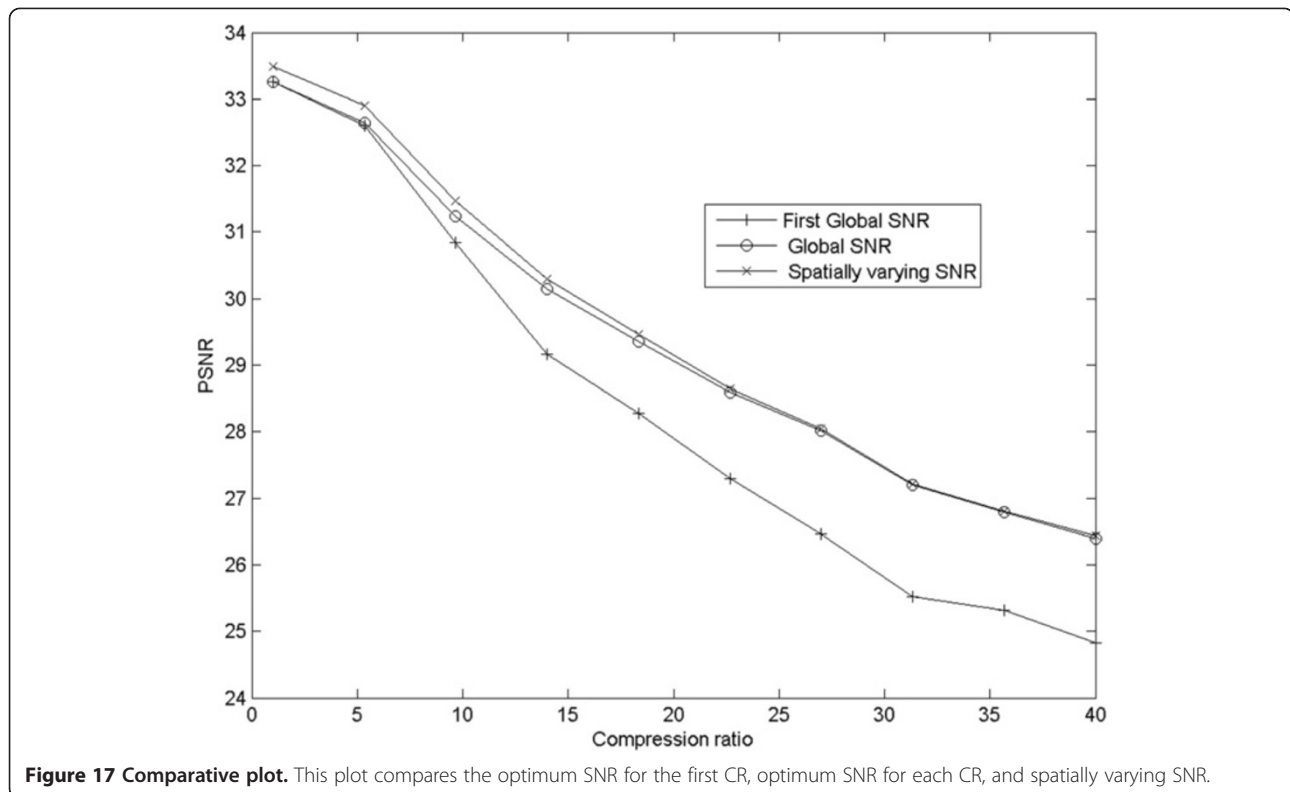


Figure 16 PSNR versus CR for various SR/compression methods. The PSNR optimum SNR is used for the AWF SR with $\rho = 0.7$.



this is perhaps the best way to combine compression and SR. Image data suitable for multiframe SR is probably well suited for difference-frame compression, due to overlapping field of view. Figure 13 shows the result obtained using the spatially adaptive correlation model AWF SR applied after difference-frame compression. As is the case for the individual-frame compression, this result has lower error and visually reduced compression artifacts.

The image obtained using compression after SR, as explained in Section 3.2, is shown in Figure 14. Again, a CR = 8 is used along with an optimum SNR of 67.40. This result also appears to be much better than SR after individual-frame compression. Since we chose $P = 16$ and $L = 4$, the SR image has the same pixel count as the set of LR frames. In a video-to-video application with input and output frame rates that are the same, this approach would face the challenge of an uncompressed throughput increase of $16\times$. Note that difference-frame compression could be applied to the resulting SR video to help deal with the increase. Figure 18 shows a region of interest (ROI) for the original and various processed images, to facilitate close inspection of image detail.

4.1.2 Quantitative results

Let us now turn our attention to the quantitative results for the simulated data. The first quantitative experiment examines the impact of compression on registration,

since sub-pixel registration is a key element of SR. Registering formerly compressed images may be necessary when compression is done prior to SR. Figure 15 shows registration error versus CR for both individual- and difference-frame compressions using the parrot image. The registration error is reported in mean absolute error (MAE) in units of pixels. These results show that registration error is small, even at large compression ratios here. We attribute this to the fact that registration is able to exploit knowledge of the global motion model, and estimation process is highly overdetermined with global motion. Registration based on difference frame-compressed images appears almost unaffected by compression. With individual-frame compression, the registration error does go up, but is still relatively small, even at a CR of 35.

In the next quantitative experiment, we compare the AWF SR method to two other benchmark methods for

Table 1 PSNR obtained using various SR/compression methods

SR techniques	SR after compression		SR before compression
	Individual-frame method	Difference-frame method	
AWF SR	31.73	33.07	33.15
WNN SR	30.13	30.46	29.67
Delaunay SR	30.53	30.98	30.20

This is for JPEG2000-compressed Kodak parrot image with CR = 8 using $\rho = 0.7$.

Table 2 MAE obtained using various AWF SR/compression methods

Input image	Bicubic after compression (individual frame)	SR after individual-frame compression		SR after difference-frame compression		SR before compression
		Optimum SNR	Local SNR	Optimum SNR	Local SNR	
Parrot	4.60	3.77	3.49	3.25	2.89	2.62
Propeller plane	5.71	4.65	4.43	4.17	3.82	3.50
Lighthouse	9.84	8.19	8.08	6.92	6.70	6.59
Mountain stream	18.49	16.76	16.61	13.74	13.59	13.36
Girl	5.65	4.76	4.59	4.03	3.80	3.51

This is for JPEG2000-compressed images with CR = 8 using $\rho = 0.7$.

the different architectures proposed in Section 3. The results are shown in Table 1. The benchmark methods are Delaunay SR [26] and weighted nearest neighbor (WNN) SR [27]. These methods are nonuniform interpolation-based SR methods with a computational complexity similar to that of AWF SR. We use the Kodak parrot image with CR = 8, and the optimum global SNR is employed for each method. The PSNR results in Table 1 show that AWF SR provided the highest PSNR results for these data. The remainder of the quantitative results focuses on the AWF SR method. However, we have observed the same basic trends using the benchmark methods.

Figure 16 shows PSNR as a function of CR for AWF SR with the different architectures described in Section 3 using the Kodak parrot image. The AWF SR method here uses a global SNR model. The plots clearly show that SR is still beneficial with compression, even out to high CRs. Individual-frame compression prior to SR is very good at low CRs but degrades quickly. The difference-frame compression performance is considerably better than individual frame for higher CRs. SR after compression provides the highest PSNR here. But again, for video-to-video applications, the issue of data throughput must be considered in this approach.

Figure 17 shows how adjusting the SNR in the AWF SR correlation model can improve results as CRs increase. The curves in Figure 17 show PSNR performance as a function of CR for the individual LR frame compression before AWF SR using three different SNR methods. The bottom curve is using a fixed SNR value

of 249.74 for all CRs. While this is the optimum SNR for CR = 4, it does not produce the best results at higher CRs. The middle PSNR curve in Figure 17 uses the optimum SNR for each CR and provides significantly better results. Further improvement is obtained using the spatially varying SNR method, described in Section 2.3.

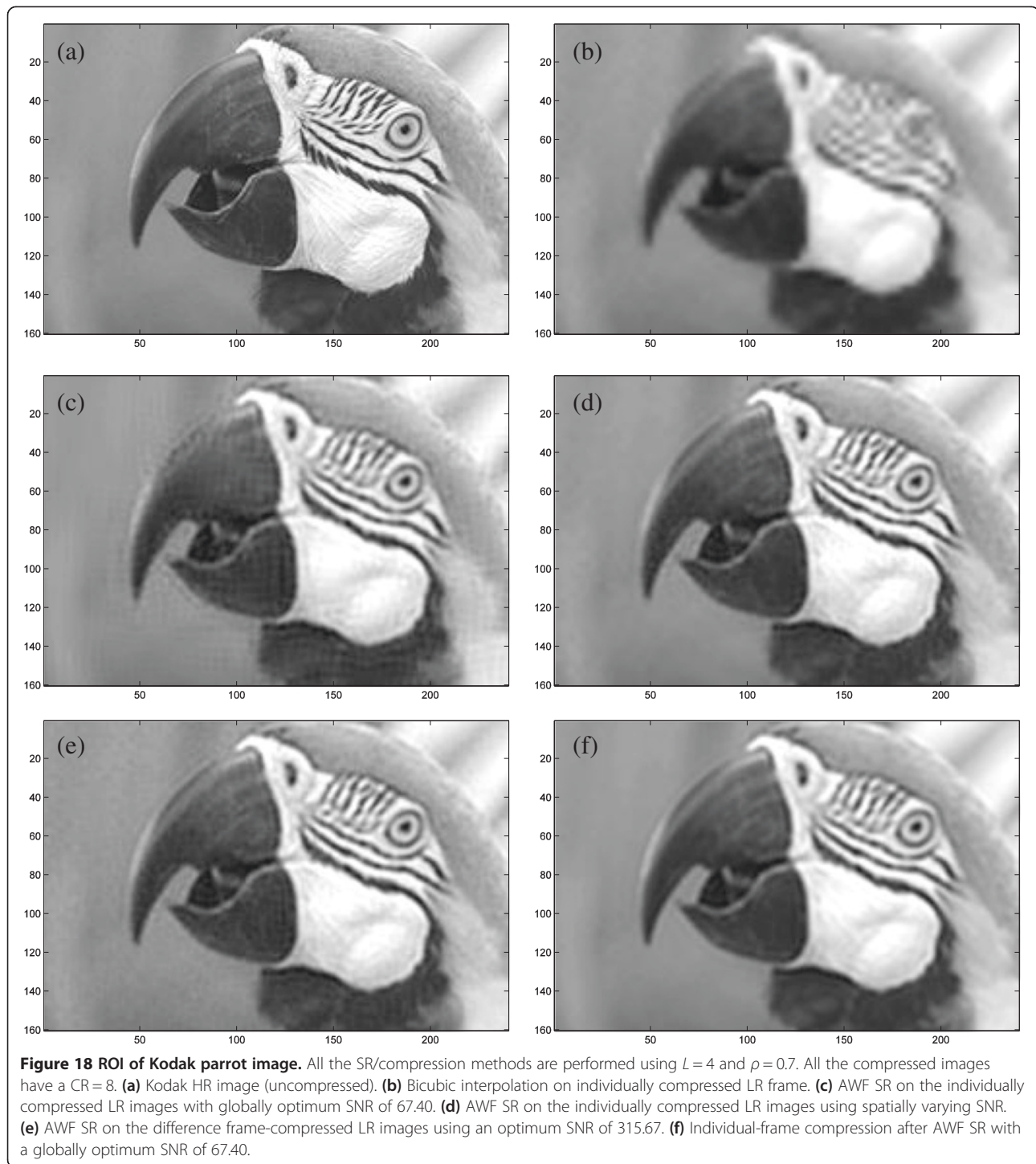
To provide a more comprehensive performance analysis, four additional images are processed. These additional images are also from the Kodak dataset [25] and are shown in Figure 19a,b,c,d. Table 2 lists the MAEs calculated between the true HR image and various processed images for CR = 8. The corresponding PSNR values are shown in Table 3. These results show that SR prior compression consistently produces the lowest error followed by difference-frame method (using spatially varying SNR) before AWF SR. The spatially varying SNR method provides good results for both individual- and difference-frame methods when compared to a global SNR. Note also that even using individual-frame compression prior to SR, we still obtain a higher PSNR than using bicubic interpolation in all cases tested.

In a final simulated data experiment, we apply AWF SR after MPEG compression for the Kodak parrot image sequence with a CR of 8. The resulting MAE value is 3.53, and the PSNR is 32.26. Comparison of these values to the corresponding values in Tables 2 and 3 shows that the JPEG2000 difference-frame compression provides better results here than MPEG. We attribute this in large part to the ability of the global difference-frame method to better exploit the global motion than the block-based motion estimation of MPEG.

Table 3 PSNR obtained using various AWF SR/compression methods

Input image	Bicubic after compression (individual frame)	SR after individual-frame compression		SR after difference-frame compression		SR before compression
		Optimum SNR	Local SNR	Optimum SNR	Local SNR	
Parrot	28.42	31.73	32.02	33.07	33.40	33.15
Propeller plane	26.34	29.17	29.36	30.27	30.42	30.27
Lighthouse	23.56	25.43	25.48	27.00	27.08	26.71
Mountain stream	19.70	20.62	20.67	22.24	22.30	22.10
Girl	27.72	29.99	30.13	31.61	31.78	31.56

This is for JPEG2000-compressed images with CR = 8 using $\rho = 0.7$.



4.2 Real-video data

In this section, the algorithms are demonstrated for a real-video sequence. Figure 20 shows a single frame in the real-video sequence captured using an Imaging Source DMK 21BU04 visible camera (Charlotte, NC, USA). The camera acquires 640×480 8-bit grayscale images with a Sony ICX098BL CCD sensor with 5.6-mm

pitch detectors (New York, NY, USA). The camera is fitted with an F/4 lens with a focal length of 5 mm. The central wavelength is assumed to be $\lambda = 0.55 \mu\text{m}$. Note that this imaging system is theoretically $5.09\times$ under-sampled. In practice, we find that there is very little residual aliasing using $L = 4$, so we believe this is a good choice for SR processing with this sensor. In all of the

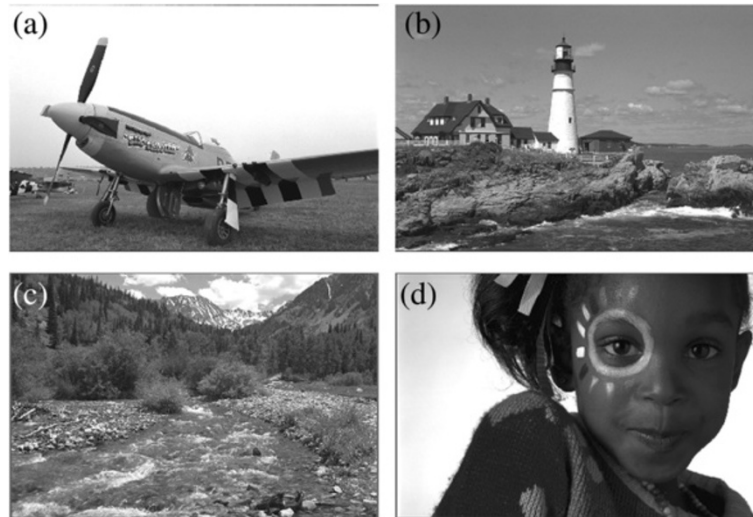


Figure 19 Ideal HR images used for simulation results. (a) Kodak propeller plane image. (b) Kodak lighthouse image. (c) Kodak mountain stream image. (d) Kodak girl image.

SR results, we use $L = 4$ and assume an SNR of 40 and use $\rho = 0.7$ in the AWF correlation model. For results incorporating compression, a CR of 8 is used in all of the results in this section.

Figures 21 and 22 show results for ROIs from Figure 20 centered on the 2D chirp pattern and book titles, respectively. The true chirp pattern is made up of concentric circles with increasing spatial frequency moving away from the center. The SR results with no compression are shown in Figures 21a and 22a. These images provide good representations of the true images. The images formed using $L = 4$ bicubic interpolation after individual-frame compression are shown in Figures 21b

and 22b. Here, moire patterns are clearly visible on the chip due to aliasing, and the lettering on the book titles is degraded. These images are also noticeably blurrier than in Figures 21a and 22a in both cases. The results for AWF SR after individual-frame compression using the optimum global SNR are shown in Figures 21c and 22c. While these results are quite good, some compression artifacts can be seen in the high spatial frequencies. AWF SR after individual-frame compression using the spatially varying SNR is shown in Figures 21d and 22d. The results for AWF SR after difference-frame compression using the optimum global SNR are shown in Figures 21e and 22e. Finally, the results for SR before compression are

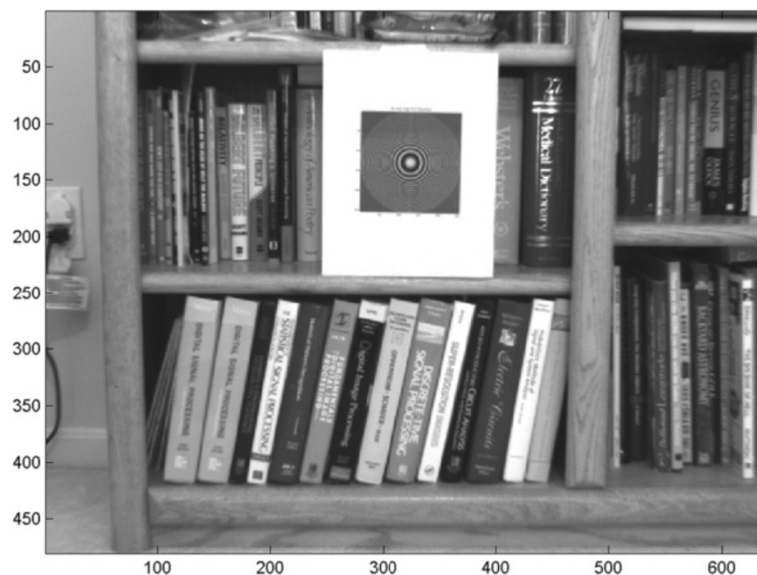


Figure 20 An individual frame (first frame) in the original real-video sequence.

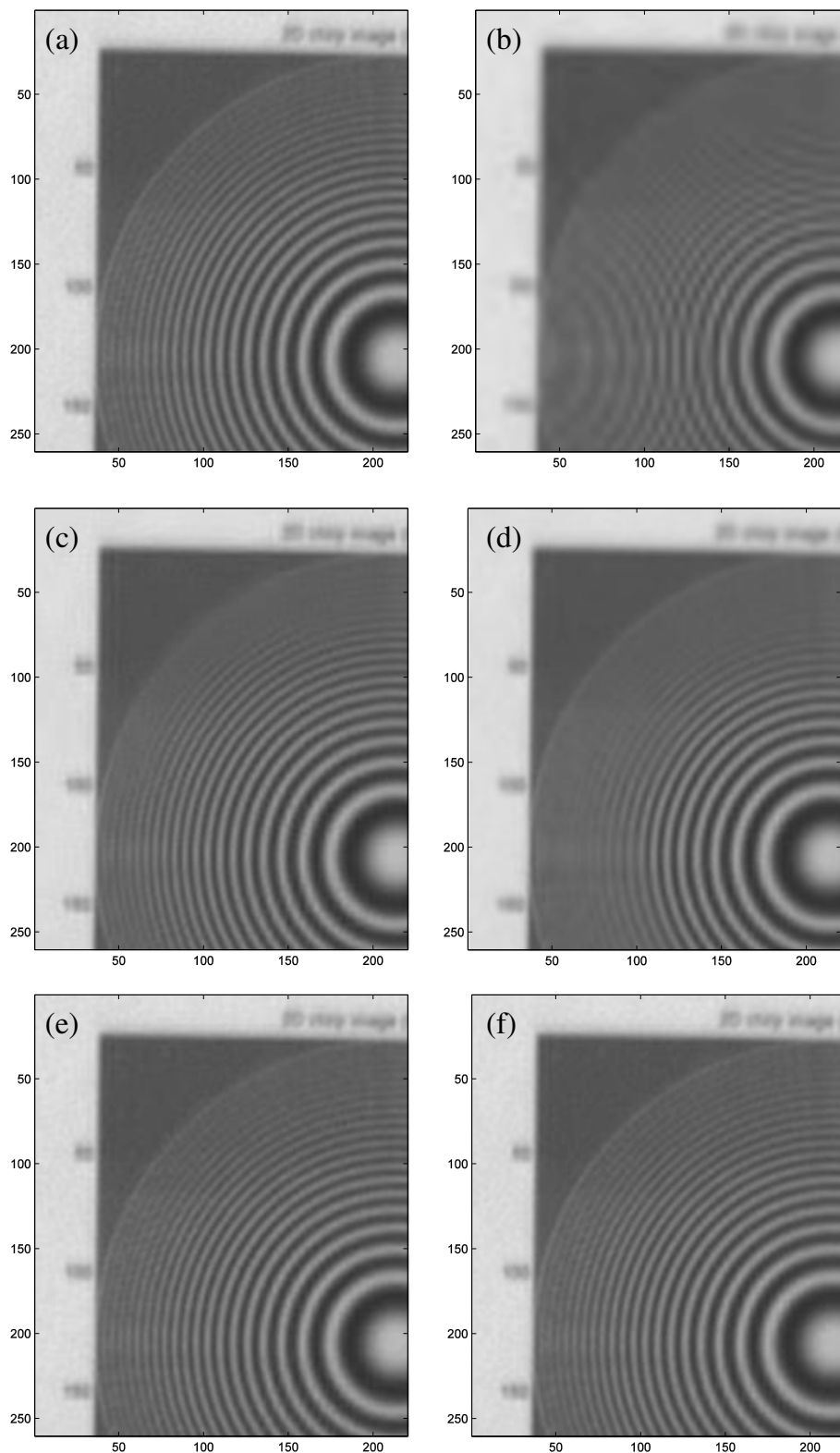


Figure 21 ROI-I of real data for various SR/compression methods using $L = 4$, $\rho = 0.7$, and $CR = 8$. **(a)** Uncompressed AWF SR. **(b)** Bicubic interpolation on individually compressed LR frames. **(c)** AWF SR on the individually compressed frames using global SNR of 40. **(d)** AWF SR on the individually compressed frames using spatially varying SNR. **(e)** AWF SR after difference-frame compression using SNR = 40. **(f)** Individual-frame compression applied after AWF SR.

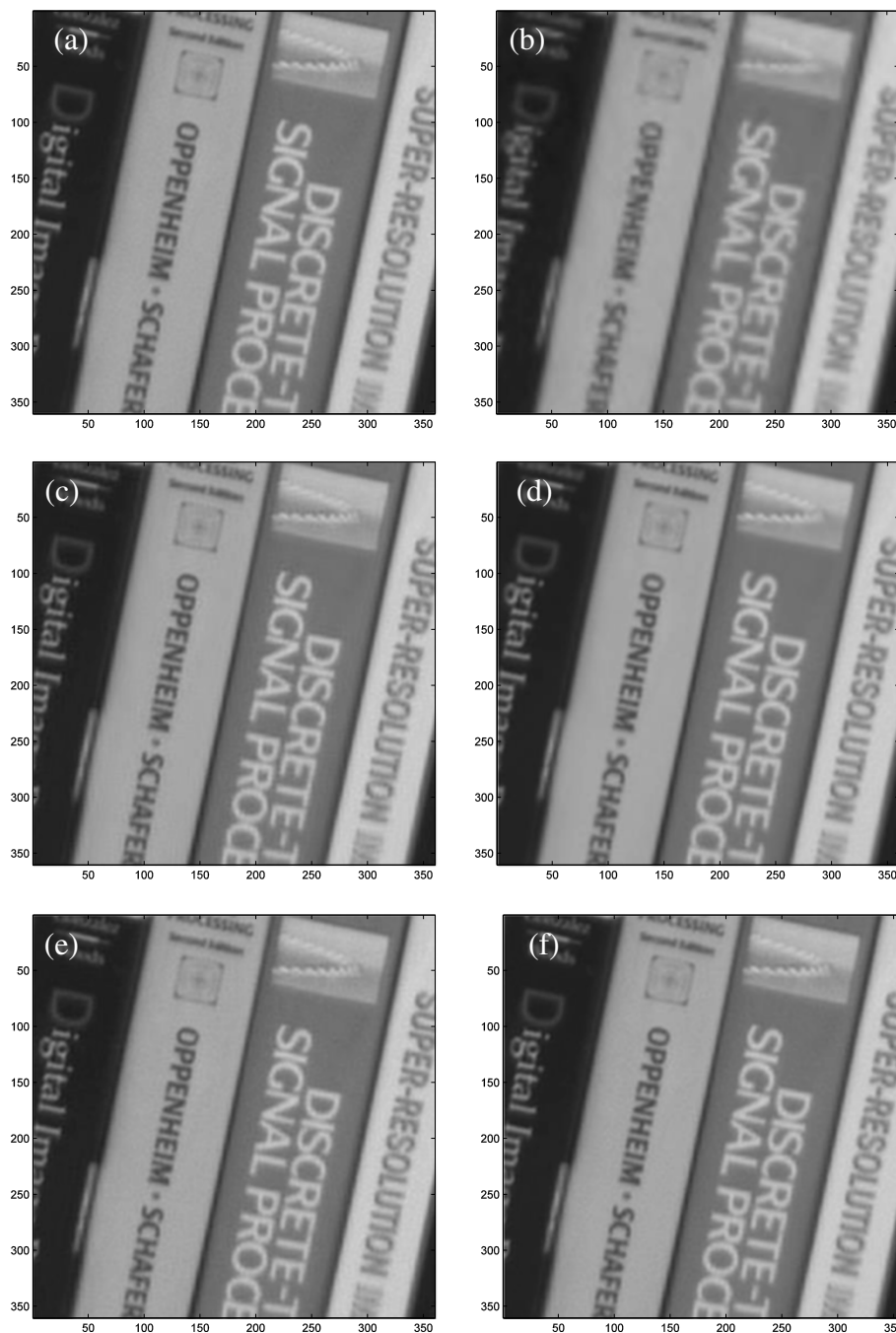


Figure 22 ROI-II of real data for various SR/compression methods using $L = 4$, $\rho = 0.7$, and $CR = 8$. (a) Uncompressed AWF SR. (b) Bicubic interpolation on individually compressed LR frames. (c) AWF SR on the individually compressed frames using global SNR of 40. (d) AWF SR on the individually compressed frames using spatially varying SNR. (e) AWF SR after difference-frame compression using SNR = 40. (f) Individual-frame compression applied after AWF SR.

shown in Figures 21f and 22f. Both of these results look comparable to the uncompressed results in Figures 21a and 22a.

We see in Figures 21 and 22 that the results with the real-video data follow the same pattern seen with the simulated data. In particular, the AWF SR before

compression is the best method among the architectures tested. This is followed closely in performance by AWF SR after difference-frame compression and then AWF SR after individual-frame compression using the spatially varying SNR. AWF SR after individual-frame compression using the optimum global SNR, while perhaps

inferior to the other methods, still outperforms bicubic interpolation by a significant margin both subjectively and quantitatively.

5 Conclusions

The results obtained suggest that SR prior to compression provides the best results when compared to other architectures. However, this may require real-time SR processing at the sensor or storage of full resolution video for later processing. This may not always be feasible. Furthermore, in video-to-video SR applications with SR prior to compression, the data throughput is significantly increased if the frame rate is the same for the LR and SR frames (since the SR frames are upsampled by L). Often, a more practical scenario is to apply SR after JPEG2000 compression of the LR frames. With this approach, we have demonstrated that SR processing still provides improvement, even for relatively high CRs, provided that the SNR in the correlation model is adjusted to account for compression artifacts. We have proposed a novel approach for estimating and using a spatially varying SNR with the AWF SR method to help mitigate spatially varying compression artifacts.

When compression is done prior to SR, we have shown that difference-frame compression is superior to individual-frame compression. Note that a set of frames suitable for multiframe SR must overlap in field of view such that accurate registration possible. With such overlapping frames, there will be a tendency to have a significant inter-frame correlation. With difference-frame compression, this redundancy between LR frames is exploited. On the other hand, it is the differences among the LR frames that are exploited by SR to reduce aliasing; it improves the performance of the system. This provides a potentially practical and efficient approach to combining SR and JPEG2000 compression.

In summary, AWF SR processing can be effectively combined with JPEG2000 compression. Even at relatively high CRs using simple individual-frame compression prior to SR, we see improvement over bicubic interpolation. Difference-frame compression prior to SR provides improved performance without increasing the video throughput. Also, applying spatially varying SNR proposed here can improve the performance of AWF SR algorithm with JPEG2000 compression. One potentially important application for this kind of SR with JPEG2000 compression is in the field of airborne imaging [19-22]. Here, multiframe SR has been shown to be highly applicable [19-22]. At the same time, the need for compression is high due to the large amounts of data typically acquired with such systems.

Competing interests

The authors declare that there are no competing interests.

Received: 21 October 2013 Accepted: 4 April 2014

Published: 25 April 2014

References

1. P Milanfar (ed.), *Super-Resolution Imaging* (CRC, Boca Raton, 2011)
2. RC Hardie, RR Schultz, KE Barner, Super-resolution enhancement of digital video. *EURASIP J. Adv. Signal Process.* **2007**(20984), 1–3 (2007). doi:10.1155/2007/20984
3. MG Kang, S Chaudhuri, Super-resolution image reconstruction. *IEEE Signal Process. Mag.* **20**(3), 19–20 (2003)
4. RC Hardie, A fast image super-resolution algorithm using an adaptive Wiener filter. *IEEE Trans. Image Process.* **16**(12), 2953–2964 (2007)
5. MW Marcellin, *JPEG2000 Image Compression Fundamentals Standards and Practice* (Springer, New York, 2002)
6. M Monaco, *Intra-frame JPEG-2000 vs. inter-frame compression comparison: the benefits and trade-offs for very high quality, high resolution sequences. Paper presented at the SMPTE technical conference and exhibition* (Pasadena, CA, USA, 2004), pp. 20–23
7. K Monaco, *Analog devices video technology enables delivery of digital cinema to big screen* (Analog Devices, Inc, 2006). http://www.analog.com/en/press-release/jun_26_2006ADI_video_tech_enables/press.html. Accessed 27 Feb 2014
8. R Molina, AK Katsaggelos, LD Alvarez, J Mateos, *Towards a new video compression scheme using super-resolution. Paper presented at the electronic imaging, International Society of Optics and Photonics* (San Jose, CA, USA, 2006), pp. 607706.1–607706.13
9. D Barreto, LD Alvarez, R Molina, AK Katsaggelos, GM Callico, Region-based super-resolution for compression. *Multidimensional Syst. Signal Process.* **18**(2–3), 59–81 (2007)
10. F Brandi, RD Querioz, D Mukherjee, *Super-resolution of video using key frames. Paper presented at the IEEE international symposium on circuits and systems* (Seattle, WA, USA, 2008), pp. 1608–1611
11. WT Freeman, TR Jones, EC Paszlor, Example-based super-resolution. *IEEE Comput. Graph. Appl.* **22**(2), 56–65 (2002)
12. Z Xiong, X Sun, F Wu, *Super-resolution for low quality thumbnail images. Paper presented at the IEEE international conference on multimedia and expo* (Hannover, Germany, 2008), pp. 181–184
13. Z Xiong, X Sun, F Wu, Robust web image/video super-resolution. *IEEE Trans. Image Process.* **19**(8), 2017–2028 (2010)
14. M Shen, P Xue, C Wang, Down-sampling based video coding using super-resolution technique. *IEEE Trans. Circ. Syst. Video Tech.* **21**(6), 755–765 (2011)
15. B Fishbain, LP Yaroslavsky, IA Ideses, *Real-time turbulent video super-resolution using MPEG-4. Paper presented at the electronic imaging, International Society for Optics and Photonics* (San Jose, CA, USA, 2008), p. 681106
16. BK Gunturk, Y Altunbasak, RM Merseau, Super-resolution reconstruction of compressed video using transform-domain statistics. *IEEE Trans. Image Process.* **13**(1), 33–43 (2004)
17. M Pickering, GF Ye, J Arnold, A transform-domain approach to super-resolution mosaicking of compressed images. *J. Phys. Conf. Ser.* **124**(1), 012039 (2008)
18. Z Xiang, Y Jie, SD Du, *Super-resolution reconstruction of image sequences compressed with DWT-based techniques. Paper presented at the IEEE international conference on wavelet analysis and pattern recognition* (Beijing, China, 2007), pp. 555–560
19. RC Hardie, KJ Barnard, Fast super-resolution using an adaptive Wiener filter with robustness to local motion. *Optics Express* **20**(19), 21053–21073 (2012)
20. RC Hardie, KJ Barnard, O Raul, Fast super-resolution with affine motion using an adaptive Wiener filter and its application to airborne imaging. *Optics Express.* **19**(27), 26208–26231 (2011)
21. Q He, R Schultz, Super-resolution reconstruction by image fusion and application to surveillance videos captured by small unmanned aircraft systems, in *Sensor Fusion and Its Applications*, ed. by C Thomas (InTech, Rijeka, 2010), pp. 475–486
22. A Camrargo, Q He, K Palaniappan, F Jara, *Super-resolution mosaics from airborne video using robust gradient regularization. Paper presented in SPIE defense, security and sensing* (International Society for Optics and Photonics, Burlingame, CA, USA, 2013), p. 875209-1–875209-7
23. RC Hardie, KJ Barnard, JG Bogner, EE Armstrong, EA Watson, High-resolution image reconstruction from a sequence of rotated and translated frames

and its application to an infrared imaging system. *Opt. Eng.* **37**(1), 247–260 (1998)

24. RD Fiete, Image quality and λ FN/ p for remote sensing systems. *Opt. Eng.* **38**(7), 1229–1240 (1999)
25. The Kodak Image Database, in <http://r0k.us/graphics/kodak>. Accessed 27 Feb 2014
26. S Lertrattanapanich, NK Bose, High resolution image formation from low resolution frames using Delaunay triangulation. *IEEE Trans. Image Process.* **11**(12), 1427–1441 (2002)
27. MS Alam, JG Bogner, RC Hardie, BJ Yasuda, Infrared image registration and high resolution reconstruction using multiple transnationally shifted aliased video frames. *IEEE Trans. Instrum. Meas.* **49**(5), 915–923 (2000)

doi:10.1186/1687-6180-2014-55

Cite this article as: Narayanan *et al.*: Multiframe adaptive Wiener filter super-resolution with JPEG2000-compressed images. *EURASIP Journal on Advances in Signal Processing* 2014 **2014**:55.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
