2011

On The Visual Quality Enhancement of Super-Resolution Images

Amr Hussein Yousef
Old Dominion University

Jiang Li
Old Dominion University, jli@odu.edu

Mohammad Karim
Old Dominion University

Andrew G. Tescher (Ed.)

Original Publication Citation

Yousef, A. H., Li, J., & Karim, M. (2011). On the visual quality enhancement of super-resolution images. In A. G. Tescher (Ed.), *Applications of Digital Image Processing XXXIV, Proceedings of SPIE Vol. 8135* (8135OZ). SPIE of Bellingham, WA. https://doi.org/10.1117/12.889291

On the visual quality enhancement of super-resolution images

Amr Hussein Yousef, Jiang Li and Mohammad Karim
Department of Electrical and Computer Engineering
Old Dominion University, Norfolk, VA 23529

ABSTRACT

Super-resolution (SR) is the process of obtaining a higher resolution image from a set of lower resolution (LR) blurred and noisy images. One may, then, envision a scenario where a set of LR images is acquired with a sensor on a moving platform. In such a case, an SR image can be reconstructed in an area of sufficient overlap between the LR images which generally have a relative shift with respect to each other by subpixel amounts. The visual quality of the SR image is affected by many factors such as the optics blur, the inherent signal-to-noise ratio of the system, quantization artifacts, the number of scenels (scene elements) i.e., the number of overlapped images used for SR reconstruction within the SR grid and their relative arrangement. In most cases of microscanning, the subpixel shifts between the LR images are pre-determined: hence the number of the scenels within the SR grid and their relative positions with respect to each other are known and, as a result, can be used in obtaining the reconstructed SR image with high quality. However, the LR images may have relative shifts that are unknown. This random pattern of subpixel shifts can lead to unpleasant visual quality, especially at the edges of the reconstructed SR image. Also, depending on the available number of the LR images and their relative positions, it may be possible to produce SR only along a single dimension diagonal, horizontal or vertical and use interpolation in the orthogonal dimension because there isn't sufficient information to produce a full 2D image. We investigate the impact of the number of overlapped regions and their relative arrangement on the quality of the SR images, and propose a technique that optimally allocates the available LR scenels to the SR grid in order to minimize the expected unpleasant visual artifacts.

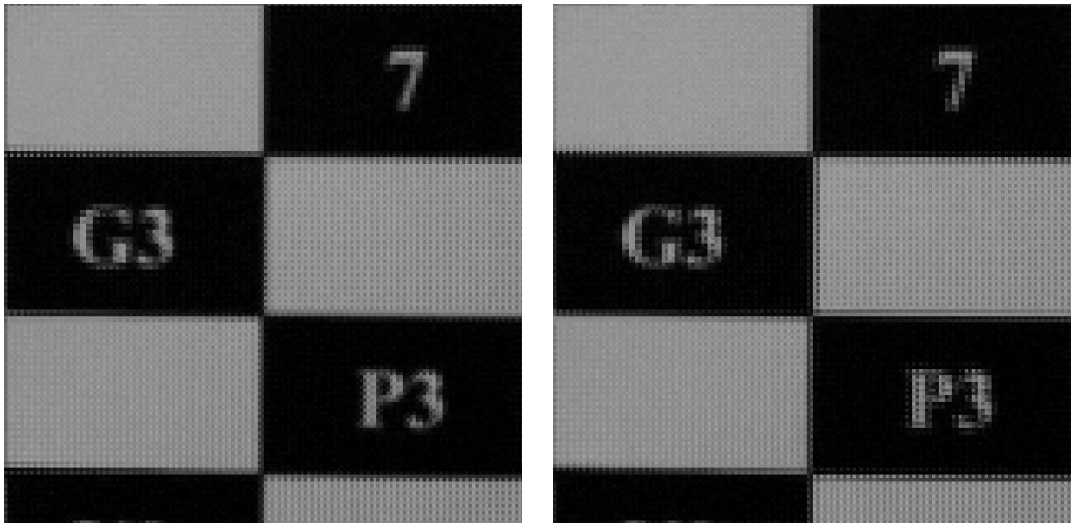
1. INTRODUCTION

Super-resolution is the process of reconstructing a higher resolution image from a set of degraded LR frames that have a slightly different views of the same scene. SR can be classified into many categories :[1] Frequency domain or spatial domain; Iterative or non iterative; Projection based or interpolation based; and Stochastic or non-stochastic. Also, most of SR techniques consists of three major steps: Registration, Reconstruction, and Restoration.[2] Registration is the process of aligning two or more frames with respect to a common reference grid. Reconstruction is used to reproduce HR image from the pixels of LR frames that might be irregularly spaced and restoration involves extracting the reconstructed image form the degradation that are encountered during the image acquisitions process. In our research we have an optical sensor mounted on a moving platform that captures a sequence of frames with some common area between them. These frames might be affected and distorted by the UAV rotational parameters i.e., yaw, pitch and roll which will result in images with shear, tilt or perspective deformations.[3] So, the UAV captured frames should be corrected first by using suitable spatial transformation like the affine or the projective transformation and then be registered within subpixel accuracy. Like many SR cases, the estimated subpixel shifts are random and don't follow a regular pattern causing the spacing between the LR scenels to be nonuniform. One approach called nonuniform interpolation converts the nonuniform spaced samples into uniform ones by interpolating the available pixels and uses them to estimate the samples or the pixels located at the coordinates of the uniform HR grid but this approach is computationally intensive[1] and also the reconstructed image using this technique loses some of the frequency components which in turn affects its visual quality.[4] In this paper, we propose a technique that is simple and effective that can be used to optimally allocate the LR scenels to the uniform HR grid with proper adjustment of their phase shifts and by

Contact: Amr Yousef (aabde008@odu.edu) is a graduate student in the Computational Intelligence and Vision Lab, at Old Dominion University (ODU). Jiang Li (JLi@odu.edu) is an assistant professor at ODU. He is affiliated with Virginia Modeling, Analysis, and Simulation Center (VMASC) and is a member of the IEEE and the Sigma Xi. Mohammad Karim (mkarim@odu.edu) is the vice President for Research of ODU.

Applications of Digital Image Processing XXXIV, edited by Andrew G. Tescher, Proc. of SPIE Vol. 8135,
81350Z · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.889291

using optimum restoration filters, SR image with good visual quality can be obtained. If the LR scenels are not located correctly to the most correct points of the uniform HR grid coordinates, then the visual quality of the reconstructed images will be affected greatly which can be seen in Figure 1 that shows the interlaced LR images without reconstruction and restoration. It can be seen with LR scenels optimum allocation, the right image is much better than the left one. Also, depending on the available number of LR frames we can obtain full SR or partial SR reconstruction when we have insufficient number of LR frames and in the latter case, interpolation should be used with a slight loss in sharpness and the visual quality of the reconstructed image.



(a) Optimal Allocation (b) Non-optimal allocation

Figure 1: The effect of optimal and non-optimal allocation on the visual quality.

The rest of the paper is organized as follows. In Section 2, A quick review of the most efficient technique to register two frames with in subpixel accuracy is presented. Two techniques that can be used to optimally allocate the scenels of LR frames to the HR grid are proposed in Section 3. In section 4, techniques for visual quality enhancements of the reconstructed images are proposed and in Section 5 we present our simulations and results.

2. SUBPIXEL SHIFT REGISTRATION

Subpixel registration is important to many image processing applications such as change detection, nondestructive evaluations and remote sensing.[5]In the latter, a one pixel in the Landsat images corresponds to 80 m distance on the earth, so registration with 0.1 pixel will lead to a resolution of 8 m. The subpixel techniques can be classified into four categories: Correlation interpolation; Intensity interpolation; Differential interpolation; and Phase correlation. In correlation interpolation, the discrete correlation function between two images is calculated and interpolated to fit a surface and the shifts are obtained by searching for the maximum. In intensity interpolation, some parts of the reference image are interpolated according to the subpixel accuracy required and a search is conducted over these parts with a target image. The idea behind differential interpolation is to relate the difference between two consecutive frames to the spatial intensity gradient of the first image.[5]The phase correlation techniques depend on the idea that the phase of the cross power spectrum between two images contains most of the information about the relative displacement between them. In our research, we use one of the most efficient and reliable algorithms based on the phase correlation technique to estimate the shifts between two images within subpixel shift accuracy.[6,7]

For the case of just a translation between two images, the usual technique to address this problem is to compute the cross-correlation between the unregistered and the base images by means of discrete Fourier transform (DFT), and locate its peak.[7] If the image to be registered is $g(x, y)$ and the base image is $f(x, y)$, then the

cross-correlation r_{fg} of $f(x,y)$ and $g(x,y)$ is defined by:

$$\begin{aligned} r_{fg}(x_0,y_0) &= \sum_{x,y} f(x,y)g^*(x-x_0,y-y_0) \\ &= \sum_{\mu,\nu} \hat{f}(\mu,\nu)\hat{g}^*(\mu,\nu)\exp\left[i2\pi\left(\frac{\mu x_0}{M_1}+\frac{\nu y_0}{M_2}\right)\right]. \end{aligned} \tag{1}$$

M_1 and M_2 are the image dimensions; * denotes complex conjugation; and $\hat{f}(\mu,\nu)$ and $\hat{g}(\mu,\nu)$ are the DFTs of $f(x,y)$ and $g(x,y)$ respectively. An efficient algorithm is to start with an initial estimate of the location of the cross-correlation peak. The algorithm used in this paper refines the initial estimate using a nonlinear optimization conjugate gradient routine to maximize $|r_{fg}(x_0,y_0)|^2$. Its partial derivative with respect to x_0 is given by:[7]

$$\frac{\partial|r_{fg}(x_0,y_0)|^2}{\partial x_0} = 2\mathcal{I}\left(r_{fg}(x_0,y_0)\sum_{\mu.\nu}\frac{2\pi\mu}{M_1}\hat{f}^*(\mu,\nu)\hat{g}(\mu.\nu)\exp\left[i2\pi\left(\frac{\mu x_0}{M_1}+\frac{\nu y_0}{M_2}\right)\right]\right) \tag{2}$$

with a similar expression for the partial derivative with respect to y_0. The algorithm iteratively searches for the image displacement (x_0,y_0) that maximizes $r_{fg}(x_0,y_0)$ and can achieve registration precision to within an arbitrary fraction of a pixel.

3. OPTIMUM SCENELS HIGH RESOLUTION GRID ALLOCATION

We propose two approaches that are simple yet effective to optimally allocate the pixels of registered LR frames; i.e. scenels to the HR grid in order to reduce the visual artifacts in the reconstructed image. The two approaches are: (1) Minimum square distance allocation (MSDA) and (2) Mid-point subpixel shifts allocation (MSSA). For simplicity, the subpixel shifts in both approaches will be adjusted such that they will be located in the first quadrant of $\mathbb{R}^2$ space. If d_x^{min} and d_y^{min} are the smallest subpixel shifts in the x and y directions among all the LR frames, then the adjusted phase shift x'_{k_1} and y'_{k_2} for the $k_1k_2^{th}$ frame are given by:

$$\begin{aligned} x'_{k_1} &= x_{k_1} + d_x^{min} \\ y'_{k_2} &= y_{k_2} + d_y^{min} \end{aligned} \tag{3}$$

3.1 Minimum square distance allocation

If the HR grid consists of $M_1 \times M_2$ blocks and every block contains $K_1 \times K_2$ scenels, then the pre-determined uniform HR grid subpixel shifts are $\frac{1}{K_1}$ and $\frac{1}{K_1}$ in the x and y directions respectively which constitutes a uniform pattern or raster. Unlike the uniform pattern, the subpixel shifts of the registered LR frames form a completely random structure. Consider

$$g = \{(0,0),(0,\tfrac{1}{K_2}),\cdots,(\tfrac{K_1-1}{K_1},\tfrac{K_2-1}{K_2})\} \tag{4}$$

is the set of the pre-determined subpixel shifts and

$$f = \{(x_0,y_0),(x_0,y_1),\cdots,(x_{K_1-1},y_{K_2-1})\} \tag{5}$$

is the set of estimated subpixel shifts of the registered LR frames. It is required then to find the optimum mapping or transformation $T: f \to g$ that converts the nonuniform subpixel shifts pattern into a uniform one. The approach here uses the Minkowski distance as a comparison metric to optimally and accurately allocate the elements of the set f to the elements of the set g with minimum distance between the corresponding points in these sets. If $P_{k_1k_2} = (\frac{k_1}{K_1},\frac{k_2}{K_2}) \in g$ and $k_i = 0,1,\cdots,K_i-1$ with $i = 1$ *or* 2 and $Q_{l_1l_2} = (x_{l_1},y_{l_2}) \in f$ and $l_i = 0,1,\cdots,K_i-1$ with $i = 1$ or 2 then the Minkowski distance of order p between $P_{k_1k_2}$ and $Q_{l_1l_2}$ is given by:[8]

$$d(P_{k_1k_2},Q_{l_1l_2}) = \left(\left|\tfrac{k_1}{K_1}-x_{l_1}\right|^p + \left|\tfrac{k_2}{K_2}-y_{l_2}\right|^p\right)^{1/p} \tag{6}$$

Typically, the order p is usually set to 1 or 2. If we measure the Euclidean distance, we set $p = 2$ and to measure the Manhattan distance, we set $p = 1$. For simplicity and computational complexity purposes we use

the Manhattan distance. Assume the HR grid is now empty and we want to fill in the required locations, so we calculate the Manhattan distance between all the points in the set f and only one point in the set g and the point with the minimum distance should be set to this location in the HR grid. In other words, for a given point $P_{k_1k_2} \in g$, the optimum and most close point to it in f is given by:

$$argmin\left\{d(P_{k_1k_2}, Q_{l_1l_2})\right\}_{l_i=0}^{l_i=k_i-1} \tag{7}$$

The search for the nearest location should be done in a zigzag scan to avoid misplacing the elements of the set f to the correct locations of the HR grid. If the total number of LR frames is L, then the total number of searches required to allocate all the LR frames to the HR grid is $L!$.

3.2 Mid-point subpixel shifts allocation

In the MSDA approach, the allocation of LR frames scenels is constrained by the uniform spacing and structure of the HR grid. So, even if with proper adjustment of LR frames phase shift, the reconstructed image will have some undesirable edge artifacts.[9] In the MSSA approach, the scenels will have a relaxed allocation on the HR grid which will be constituted depending on the relative shifts and relative-structure between the different LR scenels providing a smooth reconstructed image. To construct the HR grid in this approach, we first project the scenels locations on the x axis and then divide the inter-distance between the projected scenels to obtain the horizontal lines of the HR grid at the mid-points of the x projected scenels inter-distance. Similarly, the scenels locations will be projected on the y axis and the vertical lines of the HR grid will be located at the midpoints of the y projected scenels inter-distance. The HR grid will be formed by the intersection between the horizontal and vertical lines and every LR frame scenel will located with respect to its relative location to the other LR frames scenels.

3.3 Subpixel shift adjustments

Once the locations of the LR scenels to the HR grid have been determined using either the MSDA or the MSSA methods, the phase shifts of the interlaced LR frames should be spatially shifted using discrete Fourier transform shift theorem so that the interlaced LR scenels will have uniform phase shift differences between them. If $k_1k_2^{th}$ LR image is $s_{k1k_2}(x - x_{k_1}, y - y_{k_2})$ where x_{k_1} and y_{k_2} are the estimated subpixel shifts in the x and y directions then its representation in the spatial frequency domain $\tilde{f}_{k_1k_2}(\nu, \omega)$ is given by:[9]

$$\tilde{f}_{k_1k_2}(\nu, \omega) = \tilde{s}_{k_1k_2}(\nu, \omega)e^{-i2\pi(x_{k_1}\,\nu + y_{k_2}\,\omega)} \tag{8}$$

Also, if the determined HR grid locations are (d_{k_1}, d_{k_2}) then the subpixel shift adjustment can be performed in the spatial frequency domain by re-adjusting the phase of the LR image $\tilde{f}_{k_1k_2}(\nu, \omega)$ as given by:

$$\tilde{g}_{k_1k_2}(\nu, \omega) = \tilde{f}_{k_1k_2}(\nu, \omega)e^{i2\pi((x_{k_1} - d_{k_1})\,\nu + (y_{k_2} - d_{k_2})\,\omega)} \tag{9}$$

4. VISUAL QUALITY ENHANCEMENT OF RECONSTRUCTED IMAGE

After the registered LR frames are adjusted and embedded into the HR grid, the reconstructed image should be restored from the degradations that are encountered during the image acquisition process such as blurring due to system optics, aliasing due to sampling the continuous scene on a finite grid beyond the Nyquist rate and the additive photo-detector noise due to thermal or quantization noise.

In addition to these degradations, there are discontinuities at the transitions between the interlaced LR frames which will be very observable at the edges and the fine details and less observable at uniform areas of the reconstructed image. Hence, restoration with enhancement should be used to reduce the effect of theses degradations and artifacts within the reconstructed image. We use and compare the effects of applying different enhancement filtering techniques based on stochastic Wiener restoration. These approaches are Stochastic Wiener restoration (SWR) filter, Wiener Characteristic restoration (WCR) filter, and Wiener Gaussian enhancement (WIGE) filter. These filter are used for single image restoration and reconstruction[10] but in our work we

extend them to work for SR problem. We use the fidelity $\mathcal{F}$ as a measure of similarity between the two images $g_1(x,y)$ and $g_2(x,y)$ as defined by[11]

$$\mathcal{F}(g_1,g_2) = 1 - \left(\frac{\|g_1 - g_2\|}{\|g_1\|}\right)^2 \tag{10}$$

where $\|\cdot\|$ is the l_2 norm. Also, depending on the available number of LR frames i.e., scenels, it is possible to obtain full SR in both dimensions of the image or partial SR in a certain dimension and interpolate the missing dimensions scenels. The enhancement filters work as restoration as well as reconstruction filters to generate for missing scenels.

4.1 Stochastic Wiener restoration filter

In spatial domain, the observed image $R_o(x,y)$ is reconstructed through the spatial convolution between the Wiener filter $\Psi(x,y)$, the spatial response of the image display device $\tau_d(x,y)$ and the interlaced image $S(x,y)$ as defined by:

$$R_o(x,y) = S(x,y) * \Psi(x,y) * \tau_d(x,y) \tag{11}$$

The spatial frequency representation of this equation is given by:

$$\hat{R}_o(\nu,\omega) = \tilde{S}(\nu,\omega)\hat{\Psi}(\nu,\omega)\hat{\tau}_d(\nu,\omega) \tag{12}$$

where $\hat{R}_o(\nu,\omega), \tilde{S}(\nu,\omega), \hat{\Psi}(\nu,\omega), \text{and}\hat{\tau}_d(\nu,\omega)$ are the Fourier transform of the output image, interlaced image, Wiener filter, and spatial response of the image display device respectively. We use the symbol tilde "˜" whenever the spectrum of the image is periodic and we use the symbol caret "ˆ" whenever the spectrum of the image is non-periodic. In frequency domain the interlaced image is given by[9]

$$\tilde{S}(\nu,\omega) = \frac{1}{K_1 K_2}\sum_{k_1=0}^{K_1-1}\sum_{k_2=0}^{K_2-1}\tilde{s}_{k_1k_2}(\nu,\omega)\exp\left(-i2\pi(\nu\, x_{k_1} + \omega\, y_{k_2})\right) \tag{13}$$

where $\tilde{s}_{k_1k_2}$ is Fourier transform of the $k_1k_2^{th}$ LR image and x_{k_1} and y_{k_2} are the subpixel shifts in x and y directions. The Wiener filter minimizes the mean-square restoration error (MSRE) e^2 between the input scene $L(x,y)$ and the output image $R_o(x,y)$, as defined by:[12]

$$\begin{aligned} e^2 &= E\left\{\iint |L(x,y) - R_o(x,y)|^2\, dxdy\right\} \\ &= E\left\{\iint \left|\hat{L}(\nu,\omega) - \tilde{S}(\nu,\omega)\hat{\Psi}(\nu,\omega)\hat{\tau}_d(\nu,\omega)\right|^2 dxdy\right\} \end{aligned} \tag{14}$$

Following Amr et al.,[9] the Wiener filter that minimizes the MSRE is given by:

$$\hat{\Psi}(\nu,\omega) = \frac{K_1K_2\hat{\Phi}_L(\nu,\omega)\hat{\tau}^*(\nu,\omega)\hat{\tau}_d^*(\nu,\omega)/\left|\hat{\tau}_d(\nu,\omega)\right|^2}{\left[\hat{\Phi}_L(\nu,\omega)\left|\hat{\tau}(\nu,\omega)\right|^2\right] * \underline{\hat{\amalg}}(\nu,\omega)\sum_{mn}\left|\sum_{k_1k_2} e^{-i2\pi((2\nu-m)\,x_{k_1}+(2\omega-n)\,y_{k_2})}\right|^2 + \sum_{k_1k_2}\hat{\Phi}_{N_{k_1k_2}}(\nu,\omega) * \underline{\hat{\amalg}}(\nu,\omega)} \tag{15}$$

where Φ_L is the power spectral density (PSD) of the input scene, $\hat{\Phi}_{N_{k_1k_2}}$ is the PSD of the noise, and $\underline{\hat{\amalg}}(\nu,\omega)$ is Fourier transform of the sampling function. In frequency domain, the generated Wiener filter will be pointwise multiplied by Fourier transform of the interlaced image S to reconstruct and restore the output image.

4.2 Wiener Characteristic restoration filter

It is sometimes desirable to enhance a specific spatial feature of the input scene which can be formulated by letting $L_c(x,y) = L(x,y) * \tau_c(x,y)$ be the desired representation of the input scene $L(x,y)$, where $\tau_c(x,y)$ is a linear characteristic function. The corresponding MSRE e_c^2 is defined by[10]

$$e_c^2 = E\left\{\iint \left|\hat{L}_c(\nu,\omega) - \hat{R}_o(\nu,\omega)\right|^2 d\nu d\omega\right\} \tag{16}$$

Thus, the Wiener characteristic filter becomes

$$\hat{\Psi}_c(\nu,\omega) = \frac{K_1 K_2 \hat{\Phi}_L(\nu,\omega)\hat{\tau}^*(\nu,\omega)\hat{\tau}_c(\nu,\omega)\hat{\tau}_d^*(\nu,\omega)/\left|\hat{\tau}_d(\nu,\omega)\right|^2}{\left[\hat{\Phi}_L(\nu,\omega)\left|\hat{\tau}(\nu,\omega)\right|^2\right] * \hat{\underline{\underline{\underline{\Pi}}}}(\nu,\omega)\sum_{mn}\left|\sum_{k_1 k_2} e^{-i2\pi((2\nu-m)\,x_{k_1}+(2\omega-n)\,y_{k_2})}\right|^2 + \sum_{k_1 k_2}\hat{\Phi}_{N_{k_1 k_2}}(\nu,\omega) * \hat{\underline{\underline{\underline{\Pi}}}}(\nu,\omega)} \tag{17}$$

For example, the characteristic filter $\hat{\tau}_c(\nu,\omega)$ can be the spatial frequency response $\hat{\tau}_e(\nu,\omega)$ of the Laplacian of Gaussian ($\nabla^2 G$) operator commonly used to enhance the input scene transitions for subsequent edge detections.[10] The spatial response and spatial frequency response of this operator, respectively, are

$$\tau_e(x,y) = \frac{1}{\pi\sigma_e^4}\left(1-\frac{r^2}{2\sigma_e^2}\right)\exp\left[-\frac{r^2}{2\sigma_e^2}\right] \tag{18}$$

and

$$\hat{\tau}_e(\nu,\omega) = (2\pi\,\rho)^2 \exp\left[-2\,(\pi\sigma_e\rho)^2\right] \tag{19}$$

where $r^2 = x^2+y^2$, σ_e is the standard deviation of the Gaussian function, and $\rho = \nu^2+\omega^2$. The selection of the Gaussian standard deviation σ_e normally entails a compromise between high resolution (with a small operator) and suppression of artifacts (with large operator). We choose $\sigma_e = 0.75$ because the spatial response $\tau_e(x,y)$ relative to the unity sampling interval is then the same as that of Marr's model of the smallest operator in early human vision relative to the mean center-to-center distance between photodetectors in the foveal region of the eye's retina.[10]

4.3 Wiener Gaussian enhancement filter

The images that the Wiener filter restores with minimum MSRE normally posses high resolution. However, these images also exhibit visually annoying defects and artifacts due to aliasing, photodetector noise that amplified by the high frequency enhancement inherent in the Wiener restoration, and the ringing near sharp edges (Gibbs phenomenon) caused by the steep roll-off in the throughput response of image gathering restoration. So, it is desirable to combine this filter with an interactive control function for enhancing the visual quality. This control is provided by the Wiener- Gaussian enhancement filter[10]

$$\hat{\Psi}_v(\nu,\omega) = \hat{\Psi}(\nu,\omega)\hat{\tau}_v(\nu,\omega) \tag{20}$$

where

$$\hat{\tau}_v(\nu,\omega) = \exp\left[-2(\pi\sigma_i\rho)^2\right] + \zeta\left(\frac{2\pi\sigma_e\rho}{\sigma_o}\right)^2\exp\left[-2(\pi\sigma_e\rho)^2\right] \tag{21}$$

where $\sigma_o = 0.7$. The standard deviation σ_i controls the roll-off of the Wiener filter, while the standard deviation σ_e and the enhancement factor ζ controls the edge enhancement. The ratio σ_e^2/σ_o^2 is included so that the enhancement with $\zeta = 1$ is directly proportional to the change of the intensity at an edge transitions.[13] The resultant WIGE restoration substantially reduces the ringing of the Wiener restoration at the cost of a barely perceptible loss in sharpness.

5. SIMULATIONS AND RESULTS

In our simulations, we used a set of checkerboard images that are used to simulate the UAV captured images. These images contain the same scene but with different views. By controlling the camera orientation, the different images will contain the distortions that are expected when the UAV experiences yaw, pitch and roll. Figure 2 shows the sets of images that have been used in the simulations. In our simulations, we started with simulated high resolution scene that is blurred by a Gaussian low-pass filter defined as:

$$\hat{\tau}(\nu,\omega) = \exp\left[-\frac{\nu^2+\omega^2}{\sigma^2}\right] \tag{22}$$

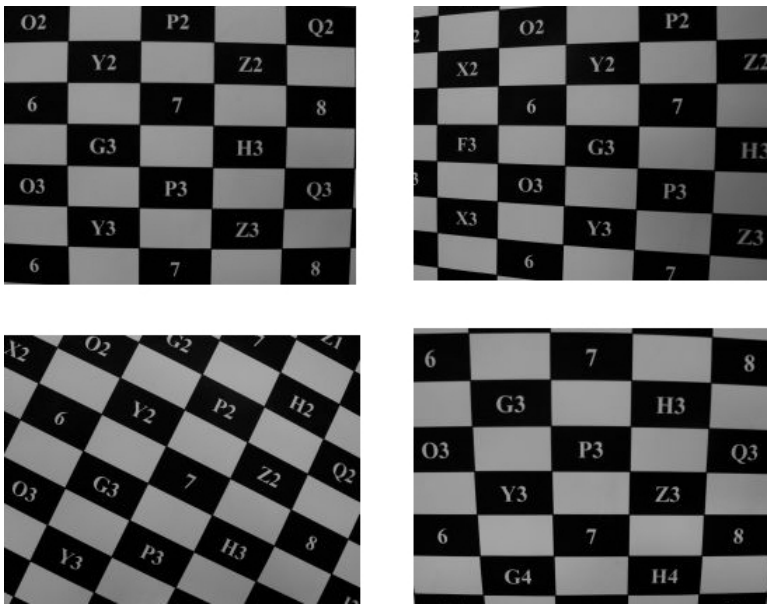


Figure 2: Simulated High Resolution Scenes with different camera orientation to simulate yaw, pitch and roll: (top-left) reference image; (top-right) pitch; (bottom-left) yaw; (bottom right) roll.

where σ is the optical-response index that specifies the design of the image acquisition device. This Gaussian filter approximates the spatial frequency response of the image-gathering device. The simulated HR images are down sampled by a factor of 2 and a white noise and blurring are superimposed to them such that the blurred signal-to-noise ration (BSNR) is 30 dB:

$$\text{BSNR} = 10 \log_{10} \left(\frac{\sigma_L^2}{\sigma_N^2} \right) \tag{23}$$

where σ_L^2 is the variance of the blurred image scene and σ_N^2 is the variance of the white noise. The derivation of the Weiner restoration and enhancement filters depend on the estimation of the input scene PSD $\hat{\Phi}_L$. Itakura et al.[14] have shown that the PSD of natural scenes can be approximated by:

$$\hat{\Phi}_L(\nu, \omega) = \frac{2\pi\mu^2\sigma_L^2}{\left[1 + (2\pi\mu\rho)^2\right]^{3/2}}, \tag{24}$$

where $\rho^2 = \nu^2 + \omega^2$ and μ is the scene mean spatial detail. The registration is performed using both spatial transformations and the algorithm proposed by Guizar et. al.[7] to register the images within subpixel accuracy. Table 1 shows the estimated subpixel shifts between the reference image and the other LR frames. For the

	LR 1	**LR 2**	**LR 3**
Subpixel shifts	$(0.06, -0.49)$	$(0.28, 0.18)$	$(-0.04, -0.12)$

Table 1: The estimated subpixel shifts using Guizar et. al.[7]

purpose of simulations, we need to examine the available number of LR frames on the reconstructed image. So, if we have 4 LR frames, we can obtain full SR in both dimensions by using 4 scenels per pixel and by using the MSDA method we can optimally allocate the scenels to the HR grid. Table 2 shows the locations of LR scenels after using the MSDA method. Also, if we have only two of these scenels and suppose they are LR 1 and LR 2 then we can use either the MSDA or the MSSA methods to allocate them to the HR grid and obtain a partial SR in a certain dimension and interpolate the missing pixels on the other dimension. Table 3 shows the results

	Ref image	LR 1	LR 2	LR 3
HR grid locations	$(0, 0.5)$	$(0.5, 0)$	$(0.5, 0.5)$	$(0, 0)$

Table 2: The optimal HR grid locations for the LR scenels (full-SR)

	LR 1	LR 3
MSDA	$(0, 0.5)$	$(0.5, 0.5)$
MSSA	$(0, 0)$	$(0.5, 0.5)$

Table 3: The optimal HR grid locations for the LR scenels (partial-SR)

of using the MSDA and the MSSA methods. The MSSA makes the allocation of the scenels more relaxed to the HR gird than the MSDA as the latter restricts the scenels to a pre-defined points on the HR grid. After the LR frames are allocated to the HR grid using the MSDA or the MSSA, they are sub-pixel shifted using Equation 9 to readjust their phase shifts. Then we use three different Wiener restoration filters introduced in Sections 4.1, 4.2 and 4.3 to enhance and restore the reconstructed images from the degradations encountered during the image-gathering process. We examine the cases of sufficient and insufficient scenels along with using the MSDA and MSSA methods to produce full-SR or partial-SR where the fidelity is used as a comparison metric between the original HR image and the reconstructed ones. The WIGE filter parameters used in this simulation are: $\zeta = 0.2, \sigma_o = 0.7, \sigma_i = 0.3$ and $\sigma_e = 0.8$. Figure 3 shows the different reconstructed images using different enhancement filters and the effect of allocating the LR scenels using either the MSDA and MSSA. Also, Figure 4 shows a fidelity comparison between the different reconstructed images and the original simulated HR scene. It can be seen that the Full-SR MSDA reconstructed images are better that than the ones that have partial-SR. Also, the images with MSSA allocation is much better than the the ones with partial-SR MSDA allocation. In terms of their fidelity, images with MSDA full-SR reconstruction have higher fidelity than the ones with partial-SR reconstruction using either MSDA or MSSA methods. When comparing the performance of different restoration filters along with using the different allocations methods, it can be seen that both the SWR and WCR filters perform nearly the same and they are much better than the WIGE filter reconstructed images. The WIGE filter outputs are more blurrier than the outputs produced by either the SWR or WCR filters. When comparing the fidelity of the reconstructed images through the different restoration filters, it can be concluded that the SWR reconstructed images are little higher than the WCR reconstructed images while the fidelity of SWR and WCR reconstructed images are much higher than the ones with WIGE reconstruction.

6. CONCLUSIONS

In this paper we present two methods (MSDA and MSSA) to optimally allocate the LR scenels to the HR grid in order to minimize the visual artifacts in the reconstructed images. The MSDA full-SR reconstructed images are much better than the MSDA or MSSA partial-SR reconstructed images. Because the MSSA provides a relaxed allocation to the HR grid, the MSSA partial-SR images are much better than the MSDA partial-SR images. To restore and enhance the reconstructed images, we used SWR, WCR, and the WIGE restoration filters. The reconstructed SR images using the different restoration and enhancement Wiener filters have a very good visual quality and minimum square error between the input scene and the reconstructed output. Also, The SWR and WCR have a better performance than the WIGE filter and it is recommended to use the latter when the visual artifacts have a great effect on the quality of the reconstructed image otherwise the other two can be used.

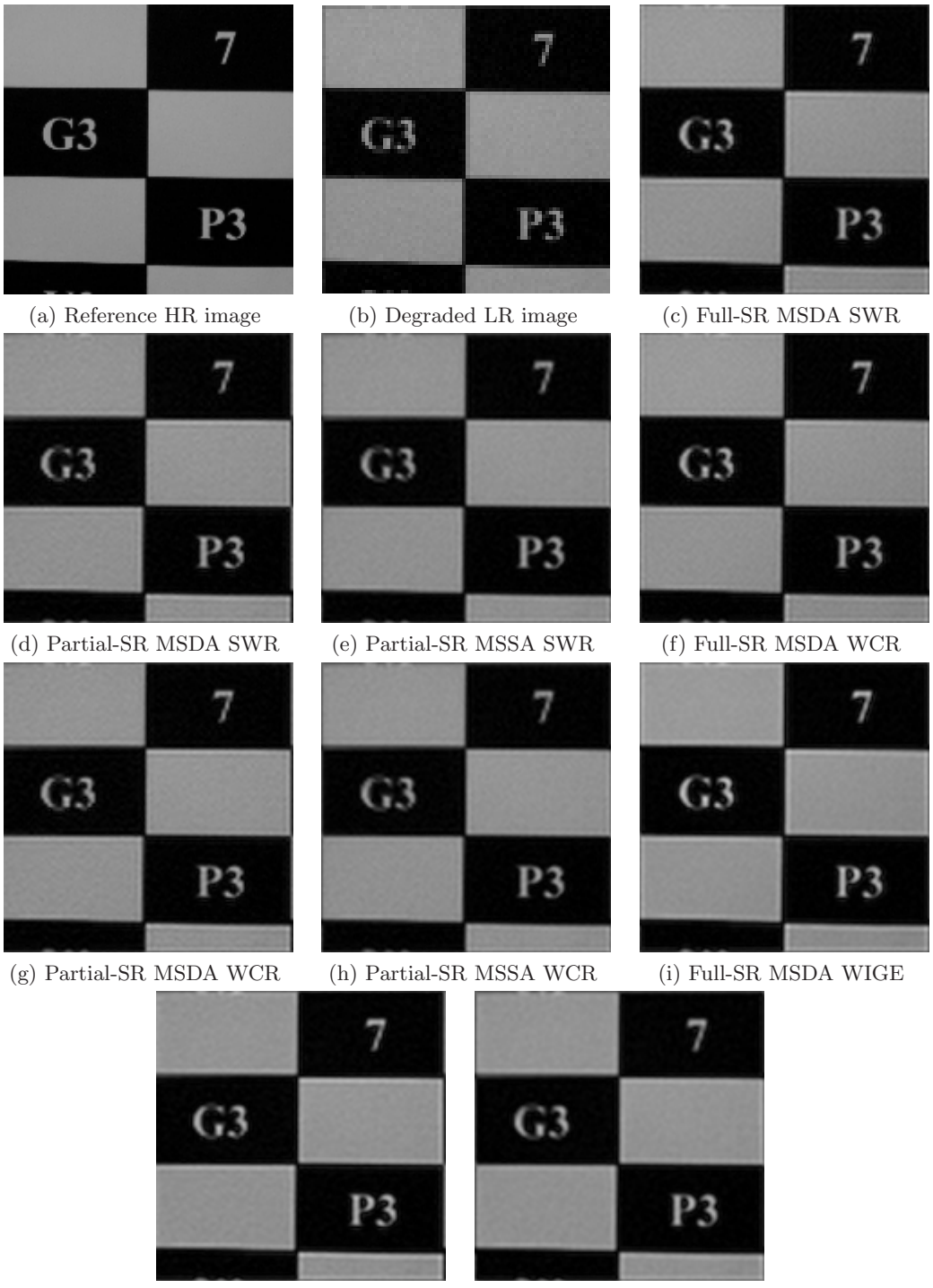
ACKNOWLEDGMENTS

The authors wish to thank the NASA Aviation Safety Program for the funding which made this work possible. In particular, Dr. Li's work was partially supported under NASA cooperative agreement NNL07AA02A.

REFERENCES

1. H. E. Sankaran, "Super-resolution using non-uniform to uniform resampling in spline spaces," master of science, Tamere University of Technology, Tampere, Finland, March 2007.

2. J. Shi, S. E. Reichenbach, and J. D. Howe, "Small-kernel superresolution methods for microscanning imaging systems," *Appl. Opt.* **45**(6), pp. 1203–1214, 2006.
3. A. H. Yousef and Z. ur Rahman, "Super-resolution reconstruction of images captured from airborne unmanned vehicles," in *Visual Information Processing*, Z. ur Rahman, S. E. Reichenbach, and M. A. Neifeld, eds., *SPIE Proceedings* **7701**, p. 77010, SPIE, 2010.
4. S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences: A comprehensive review with directions for future research," tech. rep., Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana, USA, july 1998.
5. Q. Tian and M. N. Huhns, "Algorithms for subpixel registration," *Comput. Vision Graph. Image Process.* **35**, pp. 220–233, August 1986.
6. R. A. Reed, "Comparison of subpixel phase correlation methods for image registration," final report, Aerospace Testing Alliance, Arnold Air Force Base, Tennessee, USA, April 2010.
7. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics Letters* **33**, pp. 156–158, 2008.
8. A. R. Webb, *Statistical Pattern Recognition*, Wiley, 2nd ed., 2002.
9. A. H. Yousef, Z. ur Rahman, and M. Karim, "On the restoration of the microscanned images captured from unmanned airborne vehicles," *Visual Information Processing XX* **8056**(1), p. 80560D, SPIE, 2011.
10. F. O. Huck, C. L. Fales, and Z. ur Rahman, *Visual communications: An information theory approach*, Kluwer Academic Publishers, Boston, Mass., 1st ed., 1997.
11. S. K. Park and Z. ur Rahman, "Fidelity analysis of sampled imaging systems," *Optical Engineering* **38**, pp. 786–800, 1999.
12. C. L. Fales, F. O. Huck, J. A. McCormick, and S. K. Park, "Wiener restoration of sampled image data: end-to-end analysis," *J. Opt. Soc. Am. A* **5**(3), pp. 300–314, 1988.
13. S. E. Reicenbach and S. K. Park, "Small convolution kernels for high-fidelity image restoration," *IEEE Transactions on Signal Processing* **39**(10), pp. 2263–2274, 1991.
14. Y. Itakura, S. Tsutsumi, and T. Takagi, "Statistical properties of the background noise for the atmospheric windows in the intermediate infrared region," *Infrared Physics* **14**(1), pp. 17–29, 1974.



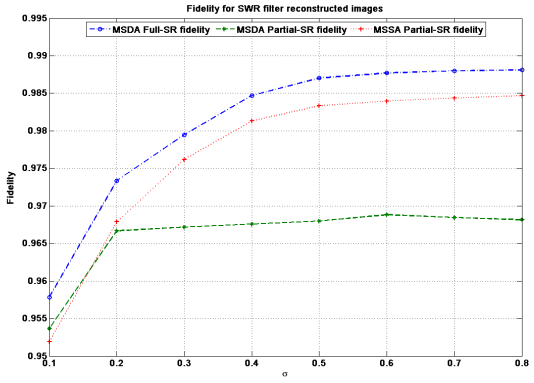
(a) Reference HR image (b) Degraded LR image (c) Full-SR MSDA SWR

(d) Partial-SR MSDA SWR (e) Partial-SR MSSA SWR (f) Full-SR MSDA WCR

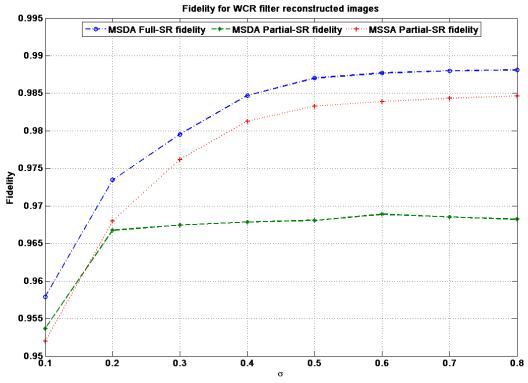
(g) Partial-SR MSDA WCR (h) Partial-SR MSSA WCR (i) Full-SR MSDA WIGE

(j) Partial-SR MSDA WIGE (k) Partial-SR MSSA WIGE

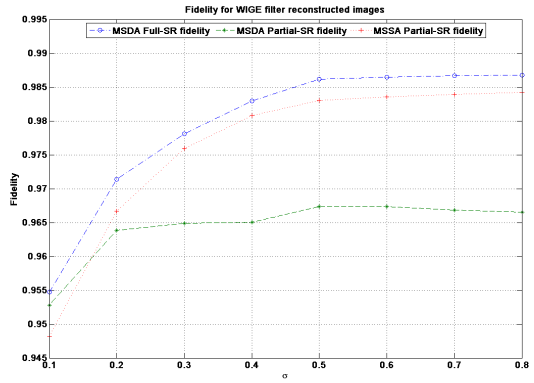
Figure 3: Different restoration outputs at image-gathering device index $\sigma = 0.8$



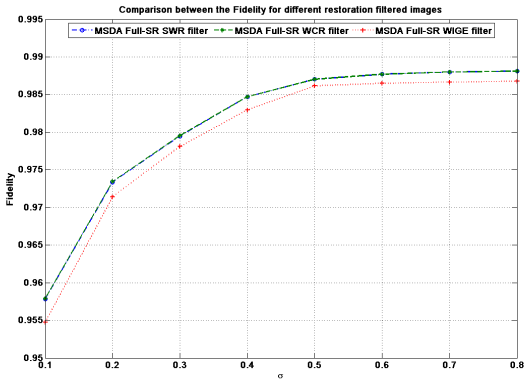
(a) SWR filter reconstructed images fidelity.



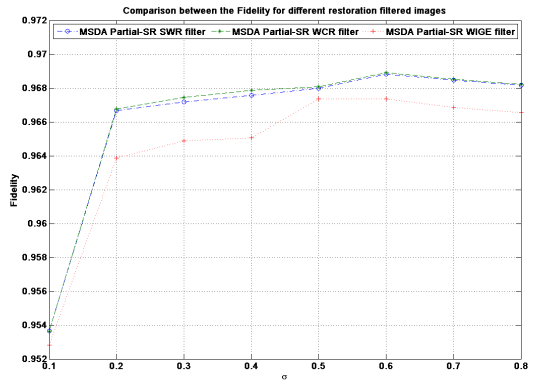
(b) WCR filter reconstructed images fidelity.



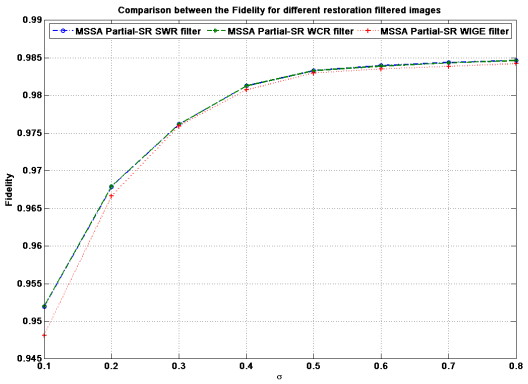
(c) WIGE filter reconstructed images fidelity.



(d) MSDA different full-SR restoration filtered images fidelity comparison.



(e) MSDA different partial-SR restoration filtered images fidelity comparison.



(f) MSSA different partial-SR restoration filtered images fidelity comparison.

Figure 4: Fidelity comparison for different restoration outputs with different image-gathering optical index σ.