

Robust Spatial-Domain Based Super-Resolution Mosaicing of CubeSat Video Frames: Algorithm and Evaluation

Debabrata Ghosh¹, Naima Kaabouch¹ & Ronald A. Fevig²

¹ Department of Electrical Engineering, University of North Dakota, ND, USA

² Department of Space Studies, University of North Dakota, ND, USA

Correspondence: Debabrata Ghosh, Department of Electrical Engineering, University of North Dakota, ND, USA.

Received: February 11, 2014

Accepted: March 21, 2014

Online Published: April 22, 2014

doi:10.5539/cis.v7n2p68

URL: <http://dx.doi.org/10.5539/cis.v7n2p68>

Abstract

In this paper, a spatial domain based super-resolution mosaicing system and its evaluation results are presented. This algorithm incorporates two main algorithms, an image mosaicing algorithm and a super-resolution-reconstruction algorithm. Huber-based prior information is used to address the ill-posed nature of the super resolution problem. To test the efficiency of the proposed algorithm, four performance metrics are used: mean square error, peak signal-to-noise ratio, singular value decomposition based measure, and cumulative probability of blur detection. Testing is performed using datasets generated from both a high altitude balloon and an Unmanned Aerial Vehicle (UAV). Results show that the proposed algorithm is highly efficient in real-time applications, such as remote sensing on a small satellite (i.e., CubeSat.) Furthermore, the performance metrics are proven to be accurate in quantitative evaluation.

Keywords: Scale Invariant Feature Transform (SIFT), Regularizer, Huber prior, Peak signal to noise ratio (PSNR), Singular Value Decomposition (SVD), Cumulative Probability of Blur Detection (CPBD)

1. Introduction

There is a push within the spacecraft industry to move toward smaller satellites while preserving performance. Satellite launch costs are quite high, on the order of several thousand US dollars per kilogram, to deliver a spacecraft to low-Earth orbit, which motivates launch providers to effectively utilize the mass that is injected into orbit. There is an increased usage of the excess available mass that can be injected into an orbit that would otherwise be unutilized by a primary spacecraft. This excess launch capability can be utilized by secondary spacecraft, such as small nanosatellites known as CubeSats.

A 1-U CubeSat has a volume of 10 cm × 10 cm × 11 cm and a mass of 1.33 kg or less. These constraints impose severe limits on payload elements, especially imaging systems whose field of view and resolution are dictated by physical dimensions that are fixed for a given remote sensing application. Some of these limitations can be overcome by employing computationally intensive image processing algorithms. We have initiated a new line of research at the University of North Dakota which will culminate in applying efficient super-resolution and mosaicing techniques to images acquired onboard a 1-U CubeSat in order to improve the resolution of those images.

Video frames captured by CubeSat payloads are often limited in quality since the use of high resolution camera is physically challenging and forbiddingly expensive in such applications. Moreover, while capturing an image, a scene may become blurred due to atmospheric effects, finite shutter speed, finite shutter time, and motion of the objects. Similarly, the scene may lose spatial resolution due to sensor array sampling (i.e., because of limited sensor density). Furthermore, images can contain noise because of sensor quality and media turbulence. However, high quality images are desired in many practical applications. Consequently, postprocessing to obtain images with improved spatial resolution from a sequence of low resolution (LR) images of the same scene becomes an attractive proposition. A super-resolution reconstruction algorithm (SR) creates a high resolution (HR) image from a sequence of low resolution (LR) images of a scene taken at different viewpoints. Such reconstruction relies on the ability to estimate motion between frames in the sequence to recover details that are finer than the sampling grid (Protter, Elad, Takeda, & Milanfar, Jan. 2009). Image mosaicing, on the other hand, is the stitching of multiple correlated images to generate a larger wide-angle image of a scene. When image

mosaicing and SR are combined together, it becomes a powerful means of representing all of the information of multiple overlapping images to obtain a high resolution panoramic view of a scene. This method is referred to as super-resolution mosaicing (SRM).

The problem of image mosaicing has been studied by many researchers in recent years. Nie et al. (Nie, Tang, & Xu, 2013) gives an overview of the image mosaicing algorithms available in the literature. Some recent algorithms include an automatic image mosaicing system using a steerable Harris corner detector proposed by Mahesh (Mahesh & Subramanyam, Dec. 2012), a novel algorithm for image mosaicing using linear texture match proposed by Fang (Fang, Yifang, Binwen, & Song, Aug. 2011), and Hierarchical multi-level image mosaicing proposed by Park (S. Park, Ghosh, Kaabouch, Fevig, & Semke, Jan. 2012). SR frequently leads to ill-posed inverse problems because of the influence of blurring operations, the decimation process, and the availability of an insufficient number of LR frames. Despite the difficulties involved in SR computation, researchers have made great progress toward stable algorithms. Park et al. (2003) gives an overview of several SR algorithms available in the literature. Some of the latest contributions include color image super-resolution reconstruction based on sparse representation proposed by Kim (Shen, Zhang, & Fu, Jan. 2013) and morphable model space-based face super-resolution reconstruction and recognition by Zhang (D. Zhang, He, & Du, 2012). Such SR algorithms have been successfully applied to areas such as digital photography, surveillance, and machine vision. However, very little work has been done for SRM. Also, the complexity of the existing super resolution mosaicing algorithms is a serious issue, particularly in practical applications such as real-time surveillance. Furthermore, in present days algorithms' accomplishments and acceptability rely mostly on quantitative evaluation.

Zhang et al. (J. Zhang, Huang, & Fan, Sept. 2007) proposed three performance metrics for quantitative evaluation of various SR methods. One of the metrics depends on the availability of ground truth (GT), which is not feasible in most practical applications such as CubeSat and aerial systems. Another metric's feasibility is also questionable because it assumes the captured images are free of noise. The third metric not only suffers from computational complexity, but also does not guarantee well-distinguishable dynamic range. Finally, the authors did not propose an experimental setup for their evaluation. Tian et al. (Tian, Suzuki, & Koike, Aug. 2010) used HR frames to generate LR frames, which were used as inputs for the SR algorithm. They proposed three objective measures that involve either measuring sub-pixel shifts between frames or detection of interest points or image segmentation. These metrics are computationally complex. In addition, all the metrics require ground truth frames for evaluation which is not possible to obtain in practical applications. Kim et al. (Kim et al., 2011) tested the effect of added blur and global random noise on SR algorithms by using three metrics which are consistent with human perception. However, those metrics do not seem to be superior to simple pixel-based measures like mean square error and peak signal-to-noise ratio. More importantly, they required ground truth frames for their metrics' operation. Several no-reference image blur metrics were proposed in the literature (Chen & Bovik, 2011; Ferzli & Karam, 2009; Kerouh & Serir, 2012) as well. However, they either do not appear to be consistent with human visual perception, or show some degree of discrepancy when blur changes from one type to another.

In this paper, we propose a super-resolution mosaicing algorithm which encompasses both the important aspects of such a system, alignment of the image frames in the sequence into a common coordinate system, and reconstruction of a HR image from those aligned LR frames. Because of problems associated with frequency domain based SR formulations (Babu & Murthy, 2011), we chose to work on a spatial domain based approach. To evaluate the efficiency of this super-resolution mosaicing algorithm, several new performance measurement metrics are used. The structure of this paper is as follows. Section 2 gives a description of the theoretical background of super-resolution reconstruction. This section can be skipped by readers who are familiar with the deterministic regularization and approximation using constrained least square (CLS) to solve the SR problem. Section 3 presents our proposed SRM algorithm. It also discusses the evaluation metrics and the experimental setup, including the extraction of frames from video captured by a high altitude balloon and a UAV along with four performance metrics to evaluate the objective quality of our algorithm. Results related to the datasets are presented in Section 4 along with discussion, demonstrating the abilities of the proposed method. We conclude in section 5, outlining the key findings of this work and scope of future works.

2. Mathematical Background

2.1 Formation of LR Images from SR Mosaic Image

The first strategic step toward solving the SR mosaicing problem is to construct an observation model which relates the acquired LR images to the SR mosaic. Several important factors which cause degradation of acquired

image, such as atmospheric blur, camera related blur, noise, and decimation, need to be included in this model (Wenzhong, Yan, & Jian, 2006) as:

$$y_k = DB_k W_k B_k^a R[x]_k + n_k \quad \text{for } 1 \leq k \leq K \quad (1)$$

Where the k th acquired image y_k is assumed to be generated from the SR mosaic x through a series of transformations: atmospheric blur effect B_k^a , followed by warp operation W_k ; camera related blur effect B_k ; and at last decimation effect D . n_k is the additive noise for the k th captured image. The reconstruction of the k th warped SR image from x is represented by $R[\cdot]_k$. Figure 1 illustrates the degradation steps corresponding to Equation (1).

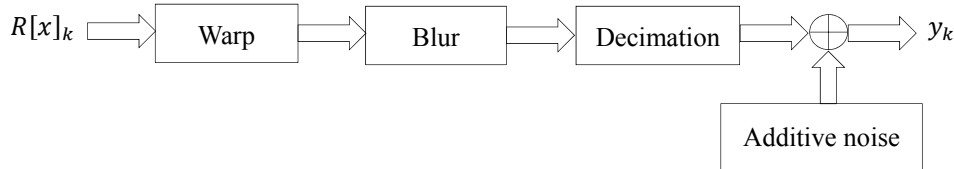


Figure 1. Image acquisition model

2.2 Theoretical Background of SR Mosaicing

Since the goal of the SR algorithm is to generate x from y_k , it is essentially an inverse process, whose reliability depends not only on the availability of multiple LR images, but also idea of several other factors like B_k^a , B_k , n_k . The variables $R[x]_k$, y_k , and n_k are rearranged as column vectors in lexicographic order. Thus, the aforementioned transformation applied to images can be represented as matrix operations (Su, Tang, Wu, Tretter, & Zhou, 2012). An estimate of SR mosaic \hat{x} is performed by optimizing a utility function which minimizes the error between the input LR images and the reconstructed ones (Su, Wu, & Zhou, 2012). A common utility function using a maximum likelihood (ML) estimate is:

$$\hat{x} = \arg \min_x \{ \sum_{k=1}^K \|y_k - DB_k W_k B_k^a R[x]_k\|_2^2 \} \quad (2)$$

However, an inverse problem like the SR mosaicing is ill-posed and requires a regularizer to be added to the solution space to make the inverse problem well-posed. The maximum-a-posteriori (MAP) criterion allows convenient addition of one or more regularizers for SR mosaic estimation (Ben-Ezra, Greenspan, & Rubner, July, 2009). With a smoothness regularizer the MAP minimization of SR mosaicing problem can be expressed as:

$$\hat{x} = \arg \min_x \{ \sum_{k=1}^K \|y_k - DB_k W_k B_k^a R[x]_k\|_2^2 + \lambda \|Lx\|_2^2 \} \quad (3)$$

Or

$$\hat{x} = \arg \min_x \{ \sum_{k=1}^K \|y_k - DB_k W_k B_k^a R[x]_k\|_2^2 + \lambda \sum_{g \in G(x)} \rho(g, \alpha) \} \quad (4)$$

The second term in both Equation (3) and Equation (4) is defined the smoothness constraint. λ is the regularization parameter which controls how much weight should be given to the smoothness constraint. When a small number of LR images are available, large values of λ are required (Pickering, Ye, Frater, & Arnold, 2008). However, it results in overly smooth solutions which may not reflect details in the solution. On the other hand, when large number of LR images is available, small values of λ are used. But, this results in a noisy solution. There is a tradeoff involved while choosing the value of λ . A typical approach for representing the smoothness constraint is to use a discrete 2D Laplacian L as in Equation (3). A better approach involves the use of Huber prior which penalizes the edges and other discontinuities less severely in an image. $G(x)$ is the set of gradient estimates over the cliques and it can also be used as a smoothness constraint. The Huber function $\rho(g, \alpha)$ in Equation (4) is defined as:

$$\rho(x, \alpha) = \begin{cases} x^2, & \text{if } |x| \leq \alpha \\ 2\alpha|x| - \alpha^2, & \text{otherwise} \end{cases} \quad (5)$$

Where α is a threshold parameter which separates quadratic and linear regions of the Huber function. If the smoothness measure is below α , a quadratic term is used to smooth small scale noise. Otherwise, a linear term is used which preserves edges and other discontinuities in the high resolution image (Xiao, Yu, & Xue, Dec, 2007).

The presence of the regularizer guarantees a convex optimization function. Thus unique estimate of \hat{x} can be computed using standard methods such as gradient descent. For the MAP constrained optimization problems like

Equation (3) and Equation (4), solutions at each iteration are expressed as:

$$\hat{x}^{(n+1)} = \hat{x}^n + \beta^{(n)} \{R^T [B_k^{Ta} W_k^T B_k^T D^T (y_k - D B_k W_k B_k^a R [\hat{x}^n]_k)]_{k=1}^K - \lambda^n L^T L \hat{x}^n\} \quad (6)$$

Or

$$\hat{x}^{(n+1)} = \hat{x}^n + \beta^{(n)} \{R^T [B_k^{Ta} W_k^T B_k^T D^T (y_k - D B_k W_k B_k^a R [\hat{x}^n]_k)]_{k=1}^K - \lambda^n G^T \rho'(G \hat{x}^n, \alpha)\} \quad (7)$$

Where,

$$\lambda^n = \left(\frac{\sum_{k=1}^K \|y_k - D B_k W_k B_k^a R [\hat{x}^n]_k\|_2}{K \|L \hat{x}^n\|_2} \right)^2 \text{ for Equation (6)}$$

$$\lambda^n = \left(\frac{\sum_{k=1}^K \|y_k - D B_k W_k B_k^a R [\hat{x}^n]_k\|_2}{K \rho(G \hat{x}^n, \alpha)} \right)^2 \text{ for Equation (7)}$$

If B_k^a and B_k are modeled as some PSF kernels, B_k^{Ta} and B_k^T are modeled as corresponding flipped kernels of B_k^a and B_k . If W_k is the backward warping operator, W_k^T is modeled as the forward operator. D^T is the interpolation operator. β is the step size of the iterative process. The regularization parameter is chosen such that it decreases as the iterative procedure progresses. This is because of the fact that we converge on a better solution by recovering the lost high frequency components.

3. Approach

3.1 Algorithm

The proposed system uses several LR frames as inputs to the SR mosaicing algorithm. The LR frames are first stitched using the mosaicing algorithm and then super resolved using an iterative SR algorithm. Consequently, it would be appropriate to discuss the mosaicing algorithm first and then the super resolution algorithm.

The mosaicing system is a combination of Scale Invariant Feature Transform (SIFT), Best Bins First (BBF), Random Sample Consensus (RANSAC), reprojection and stitching algorithms. Figure 2 shows the flow chart of the mosaicing system.

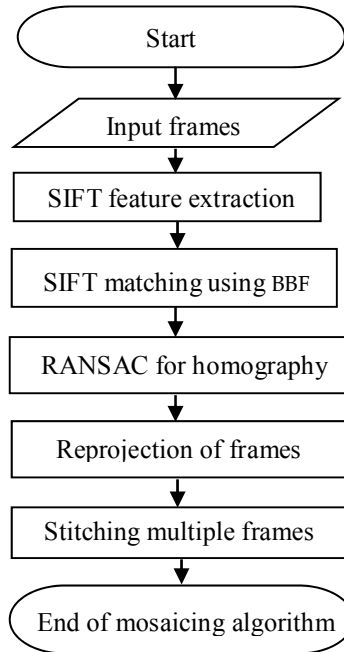


Figure 2. Flowchart of the mosaicing system

The first step of the mosaicing algorithm is the extraction of SIFT features or keypoints from the input frames. Initially, a scale space is constructed using a Gaussian filter with changing scales and an input image. The image is convolved repeatedly with the Gaussian filters and grouped into octaves. An octave corresponds to doubling the value of the standard deviation of the Gaussian filter that we start the octave with. Then the

difference-of-Gaussian (DoG) images are computed from adjacent Gaussian-blurred images in each octave. After the DoG images are obtained, candidate keypoints are identified as local extrema of DoG images across the scales. In the next step, low contrast keypoints and edge response points along the edges are discarded using accurate keypoint localization (Ghosh, Kaabouch, & Semke, 2013). The candidate keypoints are then assigned one or more orientations based on local image gradient directions. A set of orientation histograms is formed over 4×4 pixel neighborhoods with 8 bins in each. Since there are $4 \times 4 = 16$ histograms each with 8 bins, a 128 dimension descriptor vector could be assigned to each keypoint. Since the 128 element keypoint descriptor is represented relative to the orientation(s) assigned to that particular keypoint, the keypoints are invariant to rotation. Furthermore, a very high dimensional descriptor vector makes the keypoints highly distinctive.

The second step evaluates the best matching keypoints between image pairs. This is achieved by identifying the nearest neighbor of a keypoint in the first image from a database of keypoints for the second image (Lowe, 2004). The nearest neighbor is defined as the keypoint with minimum Euclidean distance from a given descriptor vector. However, because of the high dimensionality of the descriptor vectors, matching the feature points by comparing the descriptor vector one by one will require considerably high computation time. Instead, the use of BBF, which is an approximation algorithm, saves significant computation time at the cost of negligible loss of correct matching. This approximation is achieved by using a parameter BBF NN bins, which indicates when the BBF algorithm cuts off the search while looking for the nearest neighbor candidates for a particular feature vector.

Following the initial-matching-points-searching between a pair of images, the RANSAC algorithm is used to remove the outliers and to compute an optimum homography matrix based on certain homography constraints (geometric distance error, maximum number of inliers, etc.). The homography matrix is a 3×3 matrix which designates several transformation parameters between a pair of images. In order to find the homography matrices for a set of frames, one of the frames is assigned as the reference frame and the current homography matrix is multiplied with all the previous homography matrices until the reference frame is reached. Using the homography matrices, frames are projected into a common coordinate system.

Finally, the reprojected frames are stitched to the reference frame to construct the mosaic output. This stitching is achieved by checking each pixel of the mosaic frame to see if it belongs to the warped frame or the reference frame. Accordingly, that pixel in the mosaic frame is assigned the corresponding pixel value from the warped frame or the reference frame. Once we stitch the reference frame and the first reprojected frame, the result is treated as the reference frame for the stitching process, and we continue stitching the rest of the frames.

Our proposed super resolution algorithm is based on the mathematical model discussed in the previous section. Figure 3 shows the flow chart of the SR algorithm. The SR system takes several LR frames as its input, upsamples them, and computes an initial mosaic. Then it uses inverse warping, blurring, and downsampling to reconstruct the LR frames. While inverse warping is achieved by using the same homography matrices computed during the reprojection step of the mosaicing algorithm, blurring is achieved by convolving an image with a kernel which is the flipped version of the kernel used during upsampling the LR frames. These reconstructed frames are then subtracted from the input LR frames to obtain the difference frames. An error mosaic is constructed using these difference frames. The same homography matrices and convolution kernel as that used during upsampling of the LR frames are used. This error mosaic is combined with a regularization operator and used to update the initial mosaic as shown in Equation (7). This process is repeated iteratively in order to minimize the cost functional in Equation (4). Certain stopping criteria could be employed: 1) maximum number of iterations 2) threshold of error between outputs of two successive iterations, etc.

3.2 Experimental Setup and Evaluation

On board cameras are used in both the high altitude balloon and the UAV payloads in order to capture video frames. Captured video clips have a frame rate of 30 fps with 720×480 pixels frame size for the balloon data and 360×240 pixels frame size for the UAV data. Video data are fed to VirtualDub® to extract image sequences. Image sequences generated from balloon data are cropped around the border to get rid of the shadows of the suspending payload. Trimmed balloon image sequence has a frame size of 670×376 pixels. Similarly, image sequences extracted from the UAV data are cropped to 350×236 pixels in order to eliminate the unwanted black border pixels. The image sequence for each dataset is chosen such a way that the frames are in far enough intervals so that successive frames have at least some additional information and such that there is at least 50% overlapping region. Such an overlapping-region criterion is considered to be a requirement for the SIFT algorithm (Ghosh, Park, Kaabouch, & Semke, May. 2012). Ten different 204×114 pixel images with a 0.3:1 decimation are generated for each balloon-datasets. Similarly, ten different 210×142 pixel frames with a 0.6:1 decimation are generated for each UAV-datasets. Each of these decimated frames are convolved with a Gaussian

blurring kernel of size 3×3 pixels and standard deviation varying in the range of $[0.4, 2]$. Decimated frames are also contaminated by additive white Gaussian noise with zero mean and variance changed the range $[0.0002, 0.001]$. The two aforementioned arrangements simulate real captured images which were used to evaluate the behavior of our SR mosaicing algorithm in the presence of blur or noise.

The simulated captured image sequence from each of the datasets is fed to the SR mosaicing algorithm to generate a super resolved mosaic. The atmospheric blur matrix B_k^a is set as identity matrix I , and the camera related blur matrix was set to the following 3×3 space-invariant linear kernel:

$$\frac{1}{256} \begin{bmatrix} 26 & 30 & 26 \\ 30 & 32 & 30 \\ 26 & 30 & 26 \end{bmatrix}$$

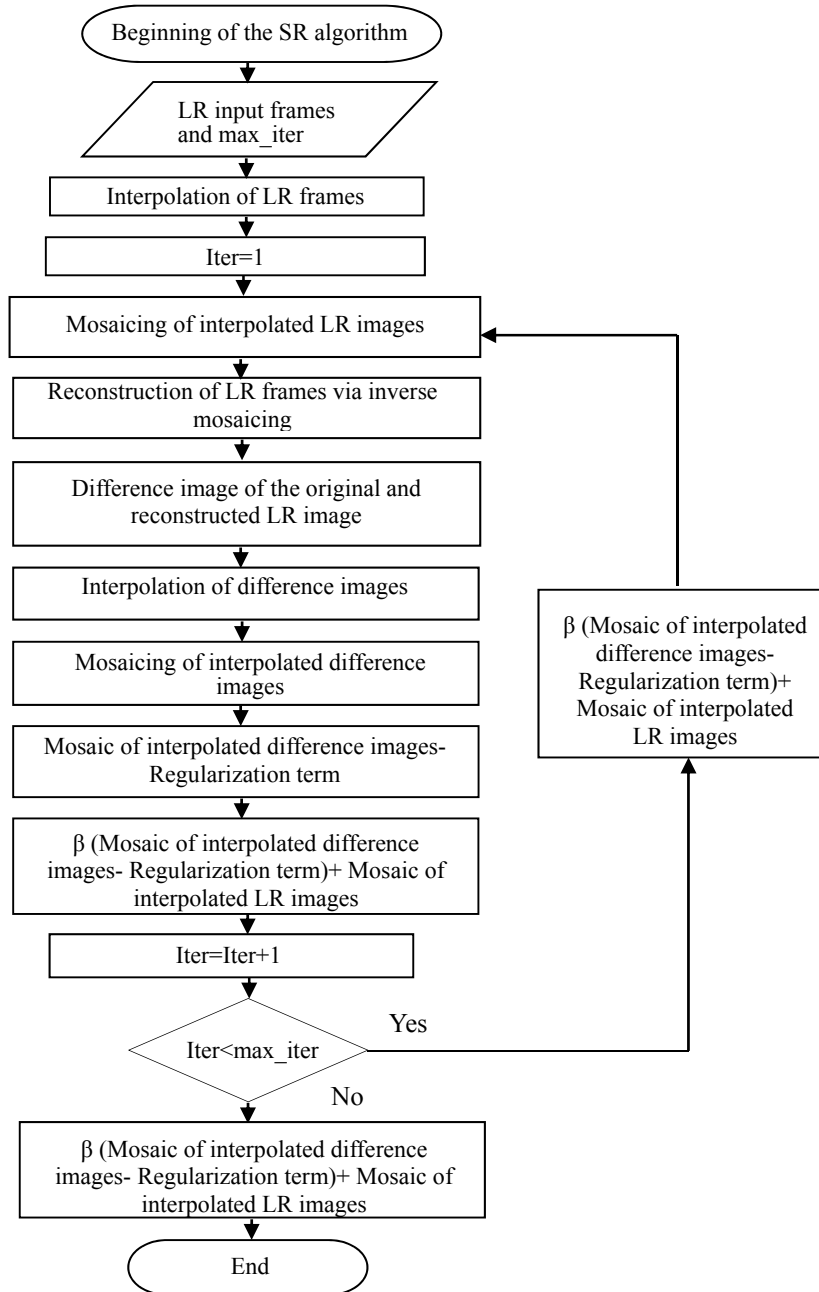


Figure 3. Flowchart of the SRM system

The value of BBF NN bins is chosen as 200 throughout our evaluation. For quantitative evaluation of the generated SR mosaic four performance metrics are used: mean square error (MSE), peak signal-to-noise ratio (PSNR), singular value decomposition-based measure (SVD), and cumulative probability of blur detection (CPBD). While the first three metrics are calculated between the initial mosaic without SR and final mosaic with SR, CPBD is calculated separately both for the initial mosaic without SR and final mosaic with SR.

a) MSE Metric

MSE measures the average of the squares of the errors. This error is the amount by which the final SR mosaic differs from the initial mosaic. Pixel wise difference is computed for MSE measurement as:

$$MSE = \frac{\sum_i \sum_j (I(i,j) - O(i,j))^2}{N}$$

Where, $I(i,j)$ and $O(i,j)$ are the (i,j) th pixel values in the initial mosaic and the final SR mosaic respectively. N is the total number of pixels in each image. MSE thus measures the quality of the SR mosaicing algorithm in terms of its ability to recover high frequency components lost during image acquisition as explained in section 2. Hence, a higher metric value represents a better performance.

b) PSNR Metric

PSNR measures the ratio of the maximum possible intensity value between the initial mosaic and the final SR mosaic to the error introduced by the SR mosaicing algorithm. PSNR is expressed in terms of logarithmic decibel scale as:

$$PSNR = \frac{10 \log_{10}(\max(I(i,j), O(i,j)))^2}{MSE}$$

MSE in the denominator signifies that lower metric value corresponds to better performance of an algorithm.

c) Singular Value Decomposition-Based (SVD) Metric

Hypostatic information, which has good stability, can be expressed by SVD of an image. Thus SVD can be used as one of the primary features of an image (Fei, Lian-Fen, & Yao, Dec. 2007). Any $m \times n$ rectangular matrix, A , can be decomposed into the product of three matrices, a $m \times m$ orthogonal matrix, U , a $m \times n$ diagonal matrix, S , and the transpose of a $n \times n$ orthogonal matrix, V , as:

$$A = USV^T$$

Where $U^T U = I = V^T V$. The columns of U are the eigenvectors of AA^T , the columns of V are the eigenvectors of $A^T A$ and S is a diagonal matrix containing the square roots of eigenvalues from AA^T or $A^T A$ in descending order.

The SVD-based metric measures the difference between the corresponding singular values of the initial mosaic and the final SR mosaic as:

$$D = \text{sqrt} \left[\sum_{i=1}^N (S_i - \hat{S}_i)^2 \right]$$

Where S_i are the singular values of the initial mosaic image and \hat{S}_i are the singular values of the final SR mosaic image. N is the total number of pixels in each image. A higher metric value represents better performance of an algorithm.

d) Cumulative Probability of Blur Detection (CPBD) Based Metric

As SR mosaicing fundamentally increases the resolution of the initial LR mosaic, which requires the use a no-reference image blur metric for quantitative evaluation. Because of its closeness to human blur perception, a metric that calculates the probability of blur detection is necessary such as cumulative probability of blur detection (Narvekar & Karam, 2011). This metric postulates that the blur around an edge is more or less visible based on the local contrast around that edge. Based on this idea, the probability of blur detection at each edge is computed and then pooled over the entire image to obtain a final quality score.

For a given contrast, the probability of blur detection at an edge e_i takes the form of a psychometric function as discussed in (Ferzli & Karam, 2009). This probability is expressed as:

$$P(e_i) = 1 - \exp\left(-\left|\frac{w(e_i)}{wJNB(e_i)}\right|^\beta\right)$$

Where $w(e_i)$ is the width of the edge e_i , $wJNB(e_i)$ is the just noticeable blur (JNB) width, which depends on local contrast C in the neighborhood of e_i . JNB is defined as the minimum amount of perceived blurriness around an edge. β is a parameter which depends on the global contrast of the image. $wJNB$ is measured to be 5 when $C \leq 50$ and 3 when $C \geq 51$ (Ferzli & Karam, 2009). If the actual width of an edge is the same as the JNB width, then $P(e_i) = 63\%$, below which blur is considered to be undetectable. CPBD corresponds to the percentage of edges at which the probability of blur detection is below 63%, i.e. the percentage of edges at which blur cannot be detected. Thus, a higher metric value corresponds to better performance of the algorithm.

4. Result and Discussion

Ten (10) datasets from high altitude balloon video data and 6 datasets from UAV video data, each containing 10 frames, are used to demonstrate the efficiency of the proposed algorithm. With the increase in number of iterations, the algorithm enhances the sharpness of the reconstructed image. Extensive testing shows that the quality does not improve after 7 iterations. Thus the maximum number of iterations was set to 7 throughout the testing. As stated in section 2, the decrease of the regularization parameter with the algorithm iteration indicates that the algorithm is moving toward a better solution. This was verified on different datasets and example of results is shown in Figure 4. As one can see the regulation parameter has a value of 30×10^{-3} for one iteration and decreases exponentially as the number of iterations increases. Note that for iteration 7, λ value does not change significantly from that of iteration 6. This is further confirmed at iteration 8, when the value of λ is almost the same as its value at iteration 7. Therefore in our algorithm we fixed the limited the number of iterations to 7.

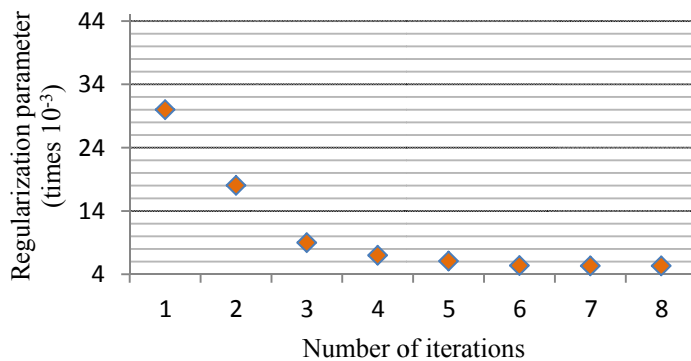


Figure 4. Regularization parameter vs number of iterations

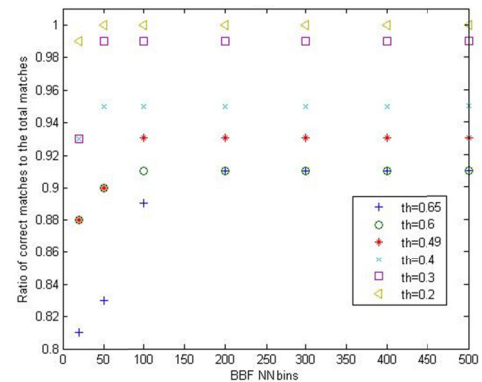


Figure 5. Search performance vs BBF NN bins to search at different distance threshold values for an image pair

As previously mentioned, the value of BBF NN bins is chosen such that the BBF algorithm cuts off the search while looking for the nearest neighbor candidates for a particular feature vector. Lower the value of this parameter higher is the speed of the nearest neighbor search. However, a low value could cost losing a high number of correct matches. Figure 5 shows that a minimum value of 200 should be chosen for BBF bins value in order not to compromise the performance of the algorithm. Note that plots are drawn with varying distance threshold value, which is another parameter for the BBF search algorithm. It indicates the ratio of the distance of the closest neighbor to that of the second-closest neighbor (regarded as incorrect match). It is clear from the graph that the performance saturates at a BBF NN bins of 200. Thus increasing this number would only result in slowing the entire search process without improving the result.

Figure 6 illustrates an example of SR mosaicing results (Figure 6 k) using 10 high-altitude-balloon frames (Figure 6 a, b, c, d, e, f, g, h, i, j) contaminated by Gaussian blurring kernel of $SD = 1.2$. As one can see in these images the algorithm produces SR mosaic with improved resolution using high resolution balloon data. The contents of the SR mosaic are sharper when compared to the contents obtained by simple interpolation operation on any of the input frames. Figure 7 illustrates a SR mosaicing result (Figure 7 k) using 10 UAV frames (Figure 7 a, b, c, d, e, f, g, h, i, j) contaminated by Gaussian noise of variance = 0.0004. As one can see in these images

the algorithm produces SR mosaic with enhanced resolution using UAV data. The SR mosaic depicts more details when compared to the details obtained by simple interpolation operation on any of the input frames.

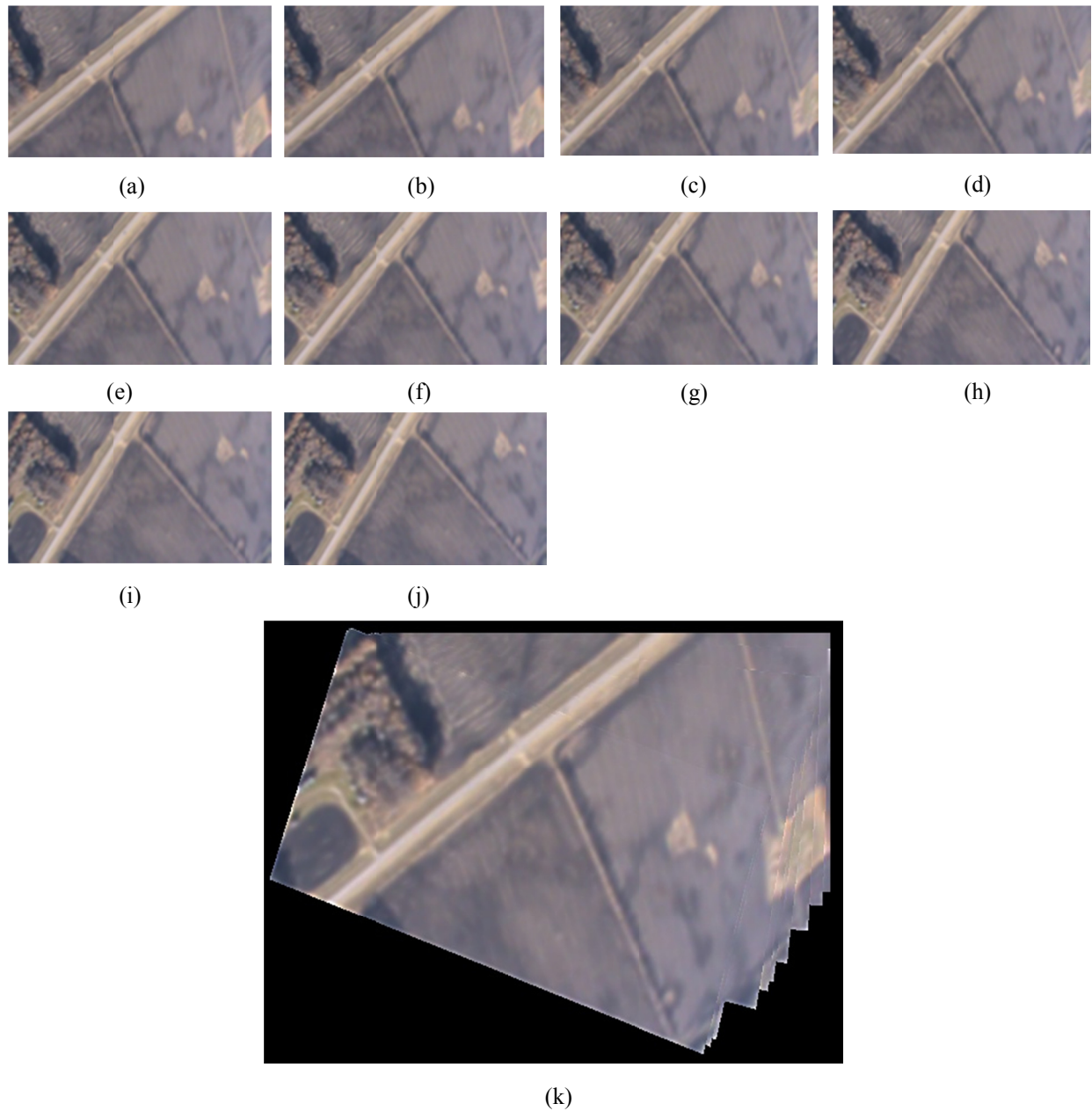


Figure 6. SR mosaicing result using high-altitude balloon datasets affected by Gaussian blur. The input frames are shown in insets (a), (b), (c), (d), (e), (f), (g), (h), (i), and (j). The corresponding SR mosaic is shown in (k)

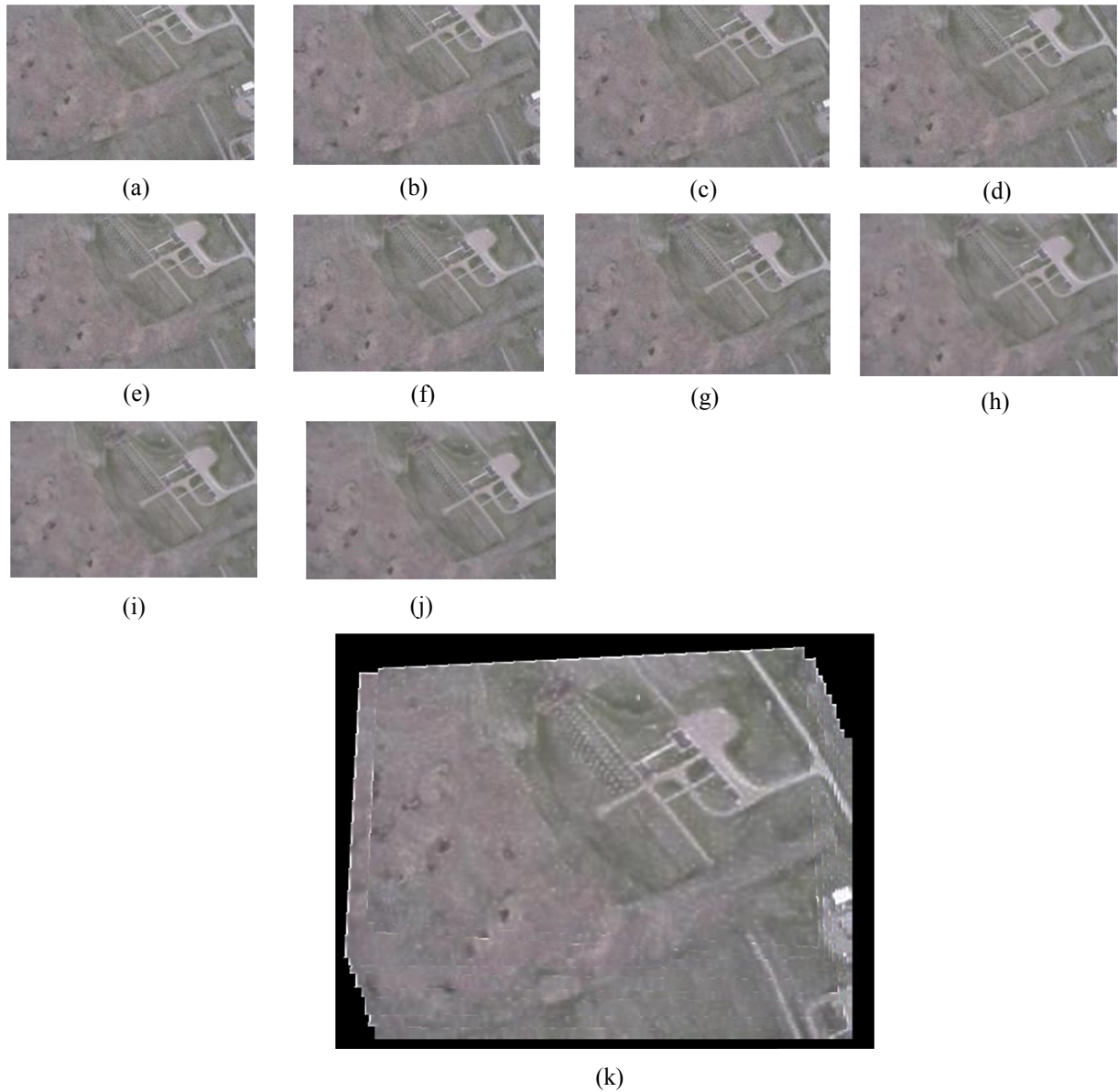


Figure 7. SR mosaicing result using UAV datasets affected by Gaussian blur. The input frames are shown in insets (a), (b), (c), (d), (e), (f), (g), (h), (i), and (j). The corresponding SR mosaic is shown in (k)

Examples of results that compare the mosaic without SR and mosaic with SR are shown in Figure 8 a, and Figure 8 b, which correspond to one of the balloon datasets. As one can see the SR mosaic image is visually sharper than the initial mosaic.

Higher CPBD values are recorded for a mosaic with SR compared to a mosaic without SR, both in the presence of blur and noise. It signifies that the proposed algorithm effectively decreases the blurriness, and hence increases the resolution. Behaviors of metric values versus standard deviation of the blurring kernel for a single dataset were plotted. Examples of results are illustrated in Figure 9 a, Figure 9 b, Figure 9 c, and Figure 9 d for MSE, PSNR, SVD, and CPBD respectively. With the increase in standard deviation of the blurring kernel MSE, SVD, and CPBD values monotonically decrease and PSNR values monotonically increase, as shown in Figure 9 a, Figure 9 c, Figure 9 d, and Figure 9 b respectively. As higher metric values of MSE, SVD, and CPBD indicate better performance of the SR mosaicing algorithm in terms of recovering high frequency components more accurately, behaviors of these plots signify efficacy of the proposed algorithm.

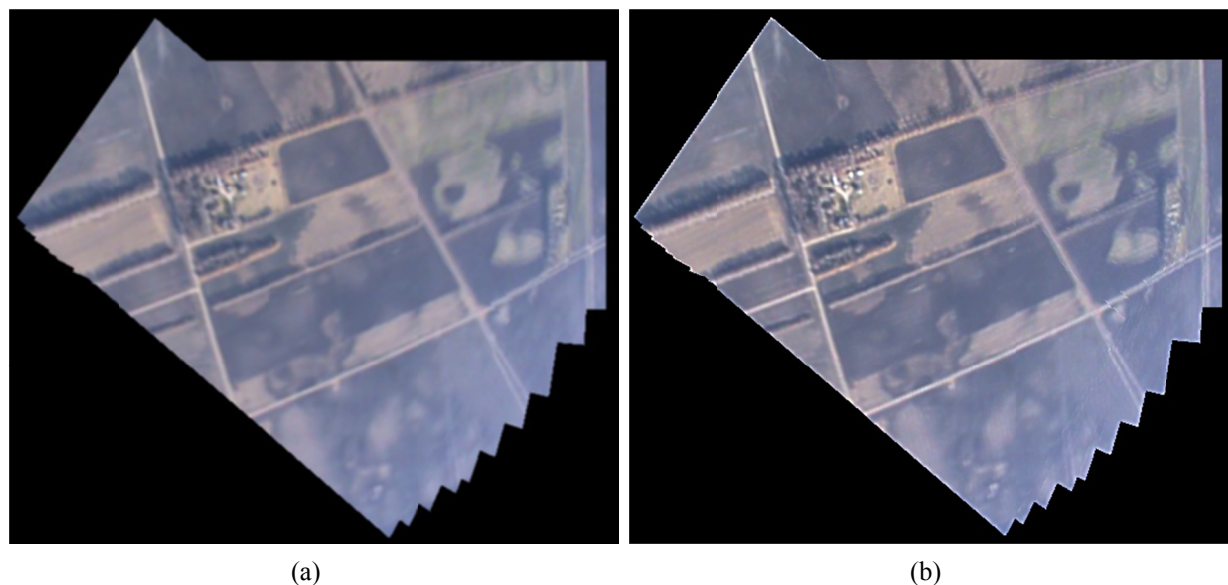


Figure 8. Comparative results of mosaic without SR and mosaic with SR using high-altitude balloon dataset. Mosaic without SR is shown in (a), whereas mosaic with SR is shown in (b)

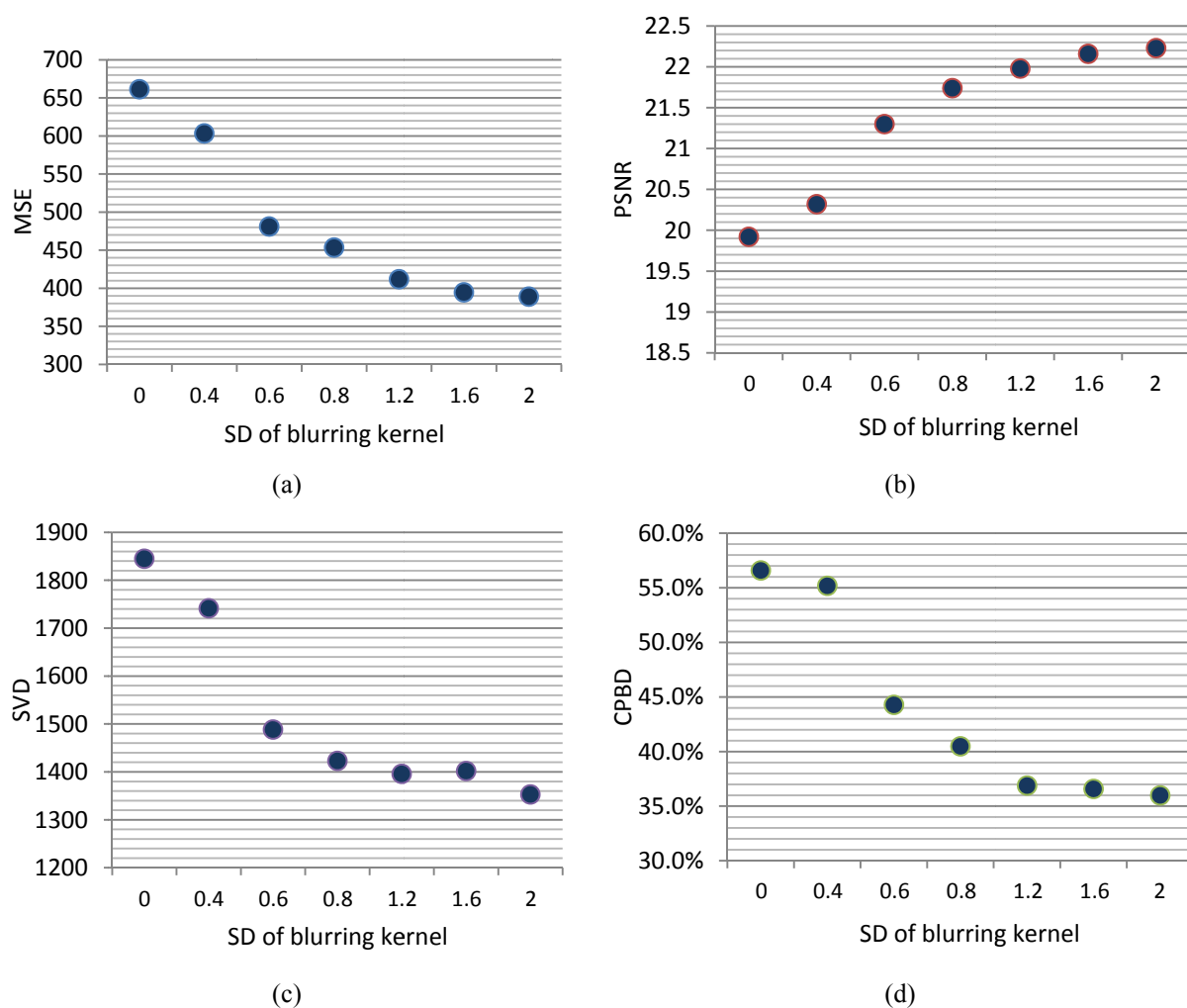


Figure 9. Performance metrics measures against standard deviation of blurring kernel. (a) shows MSE measure, (b) shows PSNR measure, (c) shows SVD measure, and (d) shows CPBD measure

On the other hand, the metrics' behaviors for all datasets are observed to be reverse when the algorithm's performance is evaluated in the presence of additive Gaussian noise. Examples of results are shown in Figure 10 a, Figure 10 b, Figure 10 c, and Figure 10 d for MSE, PSNR, SVD, and CPBD respectively. With the increase in variance of the additive white Gaussian noise MSE, SVD, and CPBD values monotonically increase and PSNR values monotonically decrease as , as shown in Figure 10 a, Figure 10 c, Figure 10 d, and Figure 10 b respectively. As noise affects high frequency components more than the low frequencies, adding more noise implies adding more high frequency components, and hence the algorithm generates even amplified version of those noisy components.

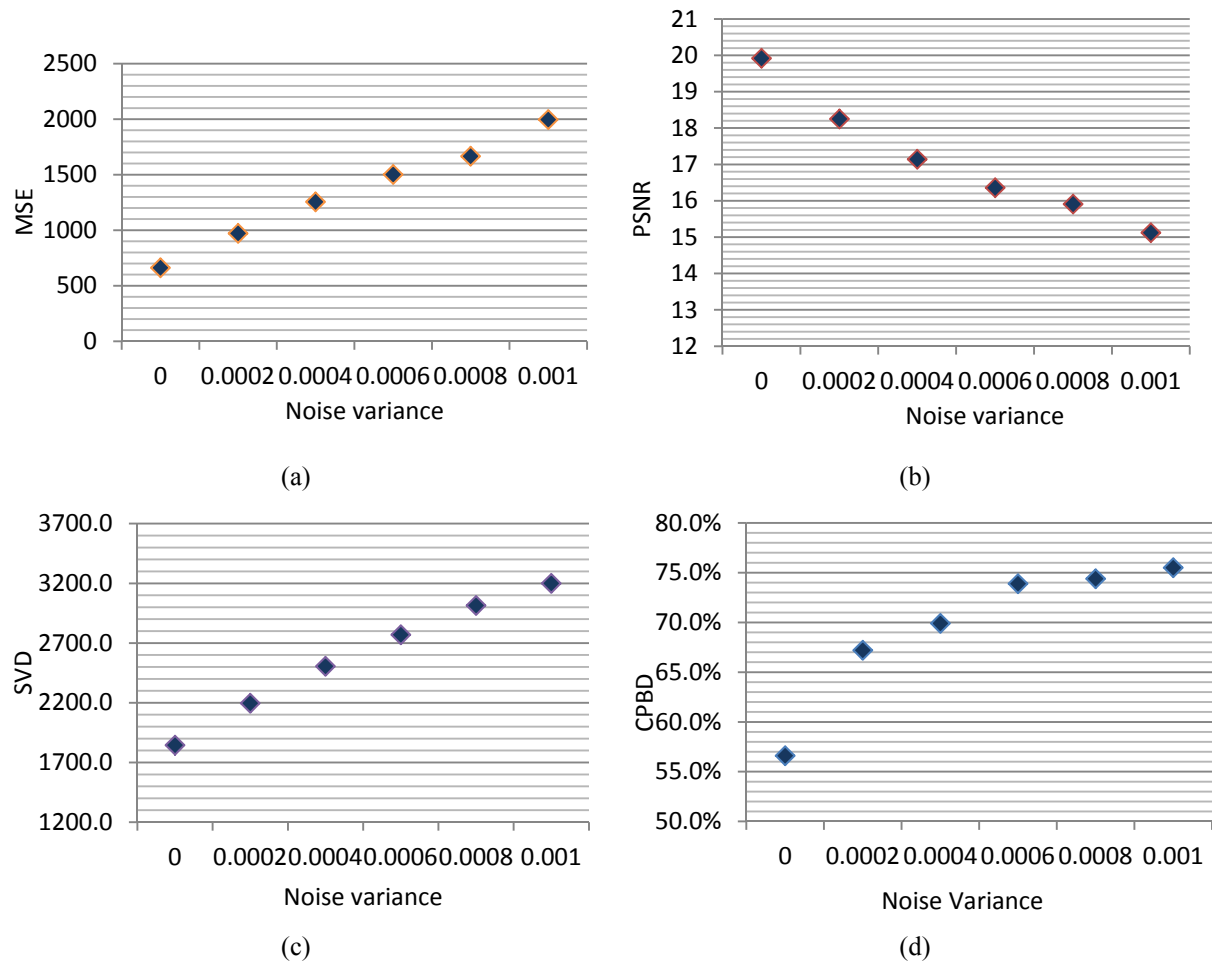


Figure 10. Performance metrics measures against variance of Gaussian noise. (a) shows MSE measure, (b) shows PSNR measure, (c) shows SVD measure, and (d) shows CPBD measure

A few inconsistencies for MSE, PSNR, and SVD metrics (6% of overall observations for each MSE, PSNR metric, and 15% of overall observations for SVD metric) are observed while evaluating all of the 16 datasets. However, the consistency of the CPBD metric proves that it is highly efficient when different distortion types and different distortion levels within a distortion type are evaluated. Furthermore, for applications like super resolution mosaicing of Unmanned Aircraft Systems (UAS), ground truths are not available. Hence, CPBD becomes the preferred choice since it is essentially a no-reference performance metric.

5. Conclusion

The proposed super resolution mosaicing algorithm is robust and computationally simple. Use of spatial domain based super resolution makes it practical in real time applications like surveillance and remote sensing. MAP minimization allows the inclusion of a regularizer in order to overcome the ill-posed nature of the super-resolution problem. The ability of the Huber regularizer to preserve edge detail has made it a preferred choice as an optimal regularizer. Objective evaluation using four metrics shows that the proposed algorithm is

effective in CubeSat applications. The CPBD metric, which is based on human blur perception for varying contrast values, exhibits better consistency as a performance metric to evaluate different distortion types. Possible directions of future research include incorporating additional performance metrics which have better correlation to the human visual system. In addition, future research will focus on making the system fully automatic.

References

- Babu, R. S., & Murthy, K. E. S. (2011). A Survey on the Methods of Super-Resolution Image Reconstruction. *International Journal of Computer Applications*, 15(2), 1. <http://dx.doi.org/10.5120/1923-2568>
- Ben-Ezra, A., Greenspan, H., & Rubner, Y. (July, 2009). Regularized super-resolution of brain MRI. Paper presented at the *IEEE International Symposium on Biomedical Imaging*.
- Chen, M., & Bovik, A. C. (2011). No-reference image blur assessment using multiscale gradient. *EURASIP Journal on Image and Video Processing*. <http://dx.doi.org/10.1186/1687-5281-2011-3>
- Fang, C., Yifang, Z., Binwen, Z., & Song, P. (Aug. 2011). Novel algorithm for image mosaic using linear texture match. Paper presented at the *International Conference on Electronic Measurement & Instruments (ICEMI)*.
- Fei, Y., Lian-Fen, H., & Yao, Y. (Dec. 2007). Algorithm for image quality measurement using singular value decomposition. Paper presented at the *IET Conference on Wireless, Mobile and Sensor Networks*.
- Ferzli, R., & Karam, L. J. (2009). A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB). *IEEE Transactions on Image Processing*, 18(4), 717-728. <http://dx.doi.org/10.1109/TIP.2008.2011760>
- Ghosh, D., Kaabouch, N., & Semke, W. (2013). Super-Resolution Mosaicing of Unmanned Aircraft System (UAS) Surveillance Video Frames. *International Journal of Scientific & Engineering Research*, 4(2).
- Ghosh, D., Park, S., Kaabouch, N., & Semke, W. (May. 2012). Quantitative evaluation of image mosaicing in multiple scene categories. Paper presented at the *IEEE International Conference on Electro/Information Technology*.
- Kerouh, F., & Serir, A. (2012). A No Reference Quality Metric for Measuring Image Blur In Wavelet Domain. *International Journal of Digital Information and Wireless Communications (IJDIWC)*, 4(1), 767-776.
- Kim, Y. J., Park, J. H., Shin, G. S., Lee, H., Kim, D., Park, S. H., & Kim, J. (2011). Evaluating super resolution algorithms. Paper presented at the *SPIE Image Quality and System Performance VIII*. <http://dx.doi.org/10.1117/12.874392>
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91-110. <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>
- Mahesh, & Subramanyam, M. V. (Dec. 2012). Automatic image mosaic system using steerable Harris corner detector. Paper presented at the *2012 International Conference on Machine Vision and Image Processing (MVIP)*.
- Narvekar, N. D., & Karam, L. J. (2011). A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD). *IEEE Transactions on Image Processing*, 20(9), 2678-2683. <http://dx.doi.org/10.1109/TIP.2011.2131660>
- Nie, L., Tang, Y., & Xu, Q. (2013). Survey of Image Mosaics Technologies. Paper presented at the *International Conference on Information Engineering and Applications (IEA)*.
- Park, S., Ghosh, D., Kaabouch, N., Fevig, R. A., & Semke, W. (Jan. 2012). Hierarchical multi-level image mosaicing for autonomous navigation of UAV. Paper presented at the *Intelligent Robots and Computer Vision: Algorithms and Techniques, SPIE*.
- Park, S. C., Park, M. K., & Kang, M. G. (May 2003). Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20, 21-36. <http://dx.doi.org/10.1109/MSP.2003.1203207>
- Pickering, M., Ye, G., Frater, M., & Arnold, J. (2008). A transfer domain approach to super- resolution mosaicing of compressed images. *Journal of Physics: Conference Series*, 124(1), 15.
- Protter, M., Elad, M., Takeda, H., & Milanfar, P. (Jan. 2009). Generalizing the Nonlocal-Means to Super-Resolution Reconstruction. Paper presented at the *IEEE Transactions on Image Processing*.
- Shen, M. F., Zhang, L. S., & Fu, H. Z. (Jan. 2013). Color Image Super-Resolution Reconstruction Based on

- Sparse Representation. Paper presented at the *Applied Mechanics and Materials Trans Tech Publications*.
- Su, H., Tang, L., Wu, Y., Tretter, D., & Zhou, J. (2012). Spatially Adaptive Block-Based Super-Resolution. *IEEE Transactions on Image Processing*, 21(3), 1031-1045. <http://dx.doi.org/10.1109/TIP.2011.2166971>
- Su, H., Wu, Y., & Zhou, J. (2012). Super-Resolution Without Dense Flow. *IEEE Transactions on Image Processing*, 21(4), 1782-1795. <http://dx.doi.org/10.1109/TIP.2011.2173204>
- Tian, L., Suzuki, A., & Koike, H. (Aug. 2010). Task-Oriented Evaluation of Super-Resolution Techniques. Paper presented at the *International Conference on Pattern Recognition (ICPR)*.
- Wenzhong, S., Yan, T., & Jian, L. (2006). A fast super-resolution reconstruction from image sequence. *Wuhan University Journal of Natural Sciences*, 11(2), 399-404. <http://dx.doi.org/10.1007/BF02832131>
- Xiao, C., Yu, J., & Xue, Y. (Dec, 2007). A High-Efficiency Super-Resolution Reconstruction Algorithm from Image/Video Sequences. Paper presented at the *International IEEE Conference on Signal-Image Technologies and Internet-Based System*. <http://dx.doi.org/10.1109/SITIS.2007.48>
- Zhang, D., He, J., & Du, M. (2012). Morphable model space based face super-resolution reconstruction and recognition. *Elsevier*, 30(2), 100-108.
- Zhang, J., Huang, C., & Fan, X. (Sept. 2007). On Evaluation Criterion of Image Super-Resolution Reconstruction. Paper presented at the *International Conference on Innovative Computing, Information and Control*.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).