Superresolution reconstruction of downsampled RGB Imagery

Time Aigbe, Ali Algahtani, Jeff Jenkins

Introduction

Imaging systems generally produce spatially undersampled images because the detector arrays cannot be made dense enough to yield a sufficiently high spatial sampling frequency. Multi-frame techniques, such as microscanning (referred to as mscan in this paper), are an effective means of reducing aliasing and increasing resolution in images produced by staring imaging systems. one of the techniques is to estimate a high resolution image, with reduced aliasing, from a sequence of undersampled rotated and translationally shifted frames. Interlacing all shifted frames effectively samples the original scene at higher spatial sampling rate 'Controlled Mscan. However, when a sensor is on a moving platform, we must develop 'Uncontrolled Mscan'.

We need to accurately estimate the subpixel motion of scene w.r.t. FPGA to ensure shifts are uniformly distributed and estimate missing frames (shifts). A Weighted Nearest Neighbor technique was used for estimating the missing shift images. Increasing the mscan level by smaller subpixel shifts are required when one does not possess enough images, which requires us to interlace more images. The MTF blurring of system creates a bottleneck for mscan, but the MTF can be modeled since the camera parameters are known. Applying a Wiener filter can remove blur caused by the MTF as well as improve the resolution of mscan data. A Wiener filter is done in the discrete domain on a sensor output image, but blurring is a continuous process. To avoid discrete aliasing when converting from continuous to discreet, we need to sample at a high enough frequency to meet the Nyquist rate. Mscan increases sampling rate & allows continuous functions to compose a discrete Wiener Filter for restoration. This method can be used in real time FLIR imaging.

Our approach starts by investigating controlled and uncontrolled mscan algorithms in the context of low resolution data collected by Infrared imaging systems. Then, we apply these techniques to higher resolution CMOS / RGB image data. We benchmark the performance of the uncontrolled mscan with the 'ideal'

bilinear interpolation. A GUI was developed to enable users to perform 2D Rigid (translation & rotation) registration on downsampled, transformed synthetic data and create a high resolution 'Superresolution' image. The final step presented here is an evaluation of the effects of parameter selection with respect to the Mean Absolute Error between the original image, bilinear interpolation (ideal) image, and superresolution image.

Controlled Microscanning

A controlled vibrating mirror can be used to produce a sequence of shifted undersampled frames. Each frame is shifted by a subpixel amount relative to other frames according to a pre-programmed shift pattern. The sequence of frames are then interlaced according to the known programmed shift routine to form a high-resolution image that represents the input scene effectively sampled at a higher spatial frequency. This process is called controlled microscanning.

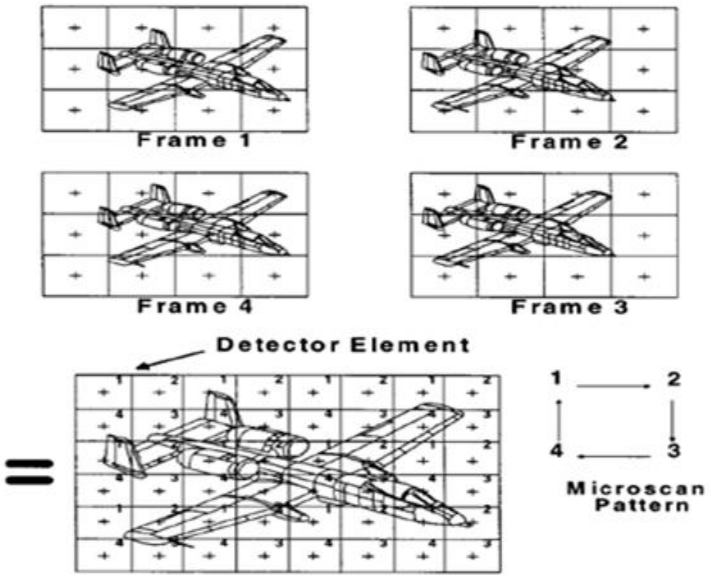


Fig. 3. Illustration of a level 2 microscanning process.

On the other hand, uncontrolled microscanning results from practical applications where the imager is mounted on a vibrating platform, such as an aircraft, and the vibrations associated with the platform can be exploited to create the shifts in the acquired frames Below is an example of the controlled microscanning algorithm performed on an image using a fixed shift (only translation) pattern.

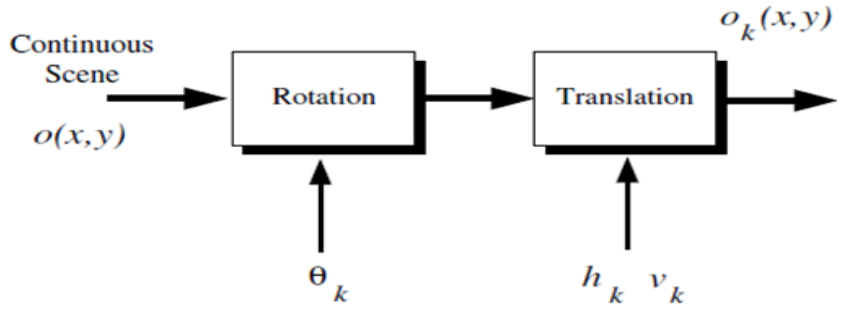
Reference image

Some shifted downsampled images

Reconstructed image

Uncontrolled Microscanning

Uncontrolled Microscanning accounts for translation as rotation fluctuations in imagery that do not arise from a fixed scanning pattern, and are shown in the following block diagram:



These two effects can be represented by the following equation:

$$ok(x, y) = o(x*\cos(\theta k) - y*\sin(\theta k) + hk, y* \cos(\theta k) + x*\sin(\theta k) + vk)$$

for k = 1, 2, 3 …. p. Note that θk k represents the rotation of the k'th frame about the origin (i.e., x = 0, y = 0). The parameters hk and vk represent the horizontal and vertical shift associated with the k'th frame. A level-L controlled mscan requires L^2 frames to be acquired to form HD image.An LXL pattern is used to interlace the staring frames, and yields a NL X NL sized frame, where N is the size of the sq. detector array.

The Sampling Model

The PSF and square detector (assume flat response) functions are convolved with the original image. A comb function representing the sampling lattice is multiplied with a rect function which limits the detector array size. A staring image is acquired by convolving our original image with the PSF as well as the limiting rect function, then multiplying by the rect function times the comb.

Image Registration

We need to choose an algorithm to register the image shift parameters. Gradient approach is good, but only works for shifts on the order of one low-res pixel width. Iterative technique utilized to handle larger shift width.The first image (o_1(x,y)) is always the ‘fixed image’, that subsequent frames (moving images) are

registered to. A Least squares method can solve for h,v shift. Here, we try to move the estimated frame closer to o_1 to minimize the registration estimates – the final estimate obtained by summing all partial estimates

High-Res Image Reconstruction

After determining relative shifts for sequence of low-res frames, the frame are placed in High Def grid using weighted nearest neighbor. WNN takes fixed image & places it in High-Res grid at [0,0]. Series of moving images must be placed in High-Res grid (1/L intervals). For any point on High-Res grid not having a value, find 3 nearest frames & provide them weights inv. Proportional to distance from desired point on High Def grid. Weighted frames are averaged to form the missing frame at each point on the grid.

Continuous observation model

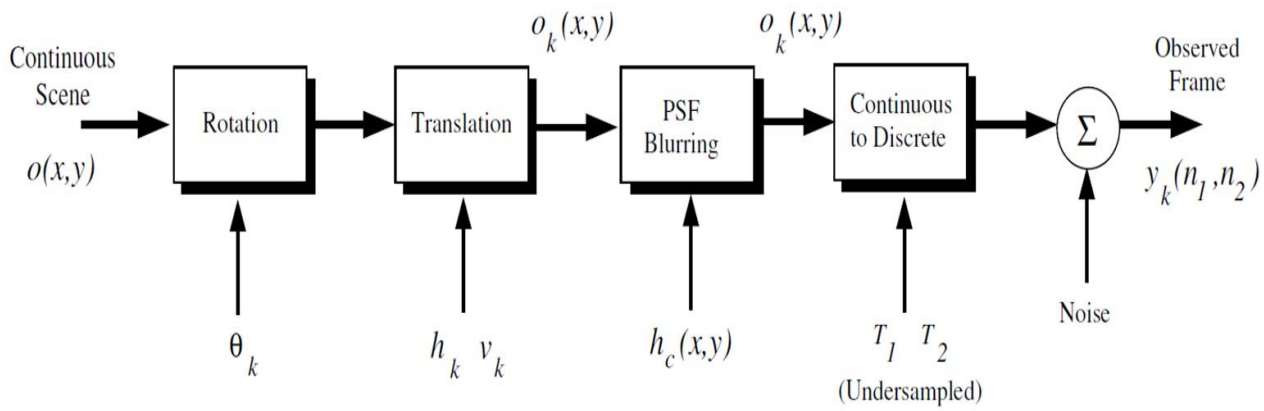


Figure 1. Continuous observation model resembling the physical process of image acquisition.

Continuous Observation Model – Physical Process

The bullet points below capture the physical process captured with the continuous observation model.

- $O(x, y)$ is the 'true scene intensity image'
    - a.k.a. the reference High Resolution image

- Imager motion between in image acquisition sequence of k frames is modeled as a pure rotation and translation of $O(x, y)$
- The k^{th} frame in the sequence can be expressed as
    - $O_k(x, y) = o(xcos\theta_k - ysin\theta_k + h_k, ycos\theta_k + xsin\theta_k + v_k)$
        - k=1,2,3,….,p.
    - θ_k represents the rotation of the k^{th} frame about the origin ($x = 0, y = 0$)
    - h_k and v_k represent the horizontal and vertical shifts associated w/ k^{th} frame
- The blurring Effect of optics and detector model are modeled
    - $\widetilde{O_k}(x, y) = O_k(x, y) * h_c(x, y)$
        - $h_c(x, y)$ is the continuous system PSF

Low Resolution Observed frame

- $y_k(n_1, n_2) = \widetilde{O_k}(n_1T_1, n_2T_2) + \eta_k(n_1, n_2)$
    - T_1, T_2 are horizontal and vertical sample spacings
    - $\eta_k(n_1, n_2)$ is an additive noise term
    - The dimensions of $y_k(n_1, n_2)$ are $N_1 \times N_2$
    - In lexicographical notation, the full set of p observed Low resolution image data is represented as
        - $= y\left[y_{k,1}, y_{k,2}, \dots, y_{k,M}\right]^T = \left[y_1, y_2, \dots, y_{pM}\right]^T$
            - $M = N_1\, N_2$
            - All observed pixel values are contained within y

Discrete Observation Model

The discrete model is necessary to develop the high resolution reconstruction algorithm, equivalent to the Continuous model

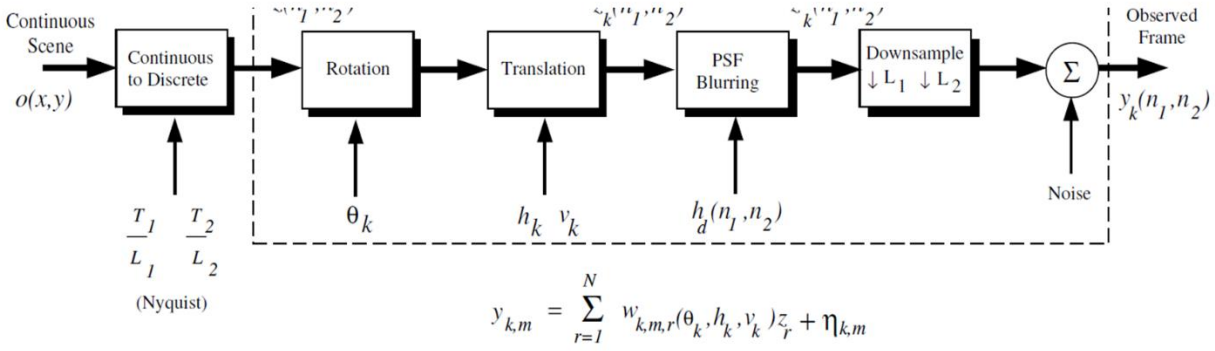


Figure 2. Equivalent discrete observation model illustrating the relationship between the ideally sampled image **z** and the observed frames **y**.

Discrete Observation Model – Philosophy

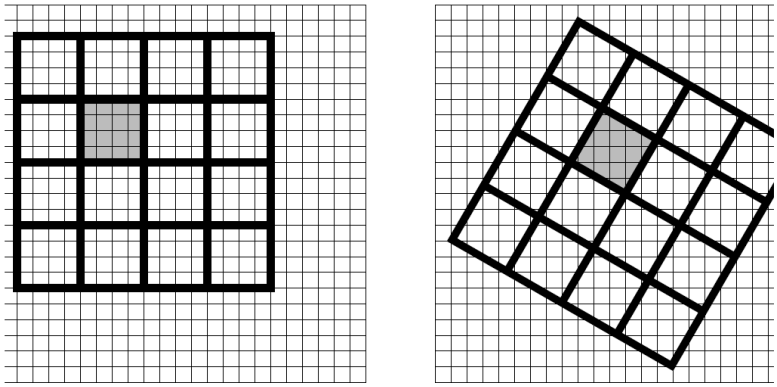
The Continuous model only provides insight into the physical process, but cannot 'enhance' our Low resolution data by itself… We need to relate a discrete High resolution image to the data in y. To do this, we estimate an intensity image $z(n_1, n_2)$ sampled at or above the Nyquist rate with no blur or noise degradation. We assume that the underlying scene z remains constant during acquisition. We also assume that the only frame-to-frame differences in weights result from rotation and translation of each low res frame relative to the high resolution grid.

Discrete Observation Model – Procedure

The size of $z(n_1, n_2)$ is $L_1N_1 \times L_2N_2 = N$, where L_1, L_2 are positive integers. $z = [z_1, z_2, \dots, z_N]^T$ is our full set of high resolution frames, and as before in the Continuous observation model, the image is rotated by θ_k and shifted by h_k and v_k, providing us with $z_k(n_1, n_2)$ – an estimate of the high resolution image. h_k and v_k shifts are defined in terms of the low resolution pixel spacings - this requires interpolation due to sampling grid changes via registration. The system PSF is similar to the continuous case, $\widetilde{z_k}(n_1, n_2) = z_k(n_1, n_2) * h_d(n_1, n_2)$. Here, $h_d(n_1, n_2)$ is the equivalent discrete system PSF and the final transformed image is subsampled to the observed frames:

$$y_k(n_1, n_2) = \widetilde{z_k}(n_1L_1, n_2L_2) + \eta_k(n_1, n_2)$$

Visualizing the discrete observation model



(a) (b)

Figure 3. Discrete detector model showing those "virtual" high resolution pixels that contribute to a low resolution pixel for two different registration positions.

Each low resolution pixel is the sum of 'virtual' high resolution pixels within the support of that low resolution detector. This model simulates the integration of light intensity impinging on the span of the low resolution detector.

Uncontrolled Microscanning - Image Registration

This high resolution reconstruction algorithm relies on accurate sub-pixel registration. We do not know the registration parameters, θ_k, h_k and v_k *a priori.* An iterative Gradient based technique is an effective method for estimating sub-pixel translation. Below are some images demonstrating the image registration and Superresoultion GUI used for Uncontrolled Microscanning simulation.

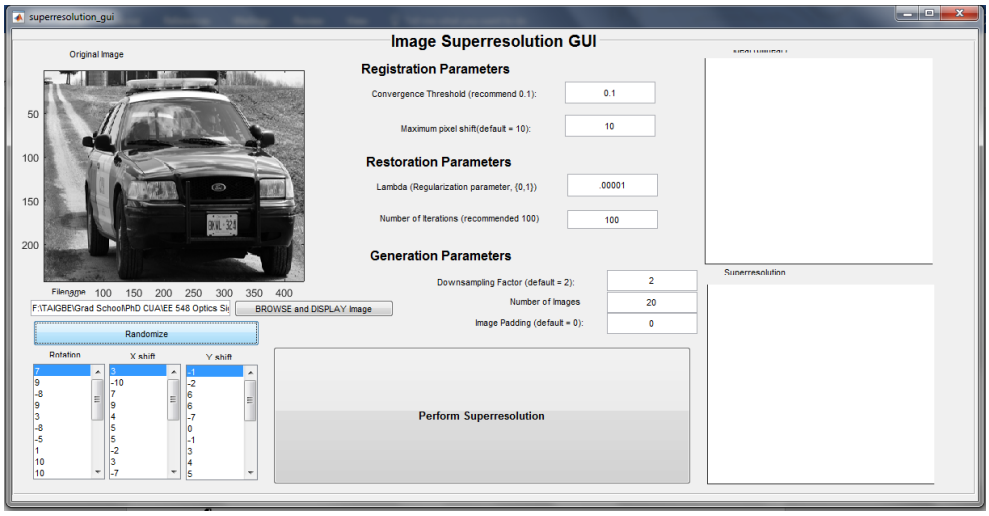


Figure 1: Image upon loading and generation of randomized shift and rotation values.

Figure 2: Image undergoing registration process in GUI (difference between observed image and image after one more downsampled image has updated it)

High Resolution Restoration

With estimates of the registration parameters, the observation model can be completely specified. In light of the observation model in (8), we define the high resolution image estimate to be

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} C(\mathbf{z}), \tag{33}$$

where

$$C(\mathbf{z}) = \frac{1}{2}\sum_{m=1}^{pM}\left(y_m - \sum_{r=1}^{N} w_{m,r} z_r\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\alpha_{i,j} z_j\right)^2, \tag{34}$$

The paper suggested a Gradient Descent approach, but this was very hard to debug (found in the rls_restoration code) and did not converge with our data – especially without a known psf model. The Conjugate gradient method yielded a result in approximately 4 iterations, much less than Gradient descent.

Results

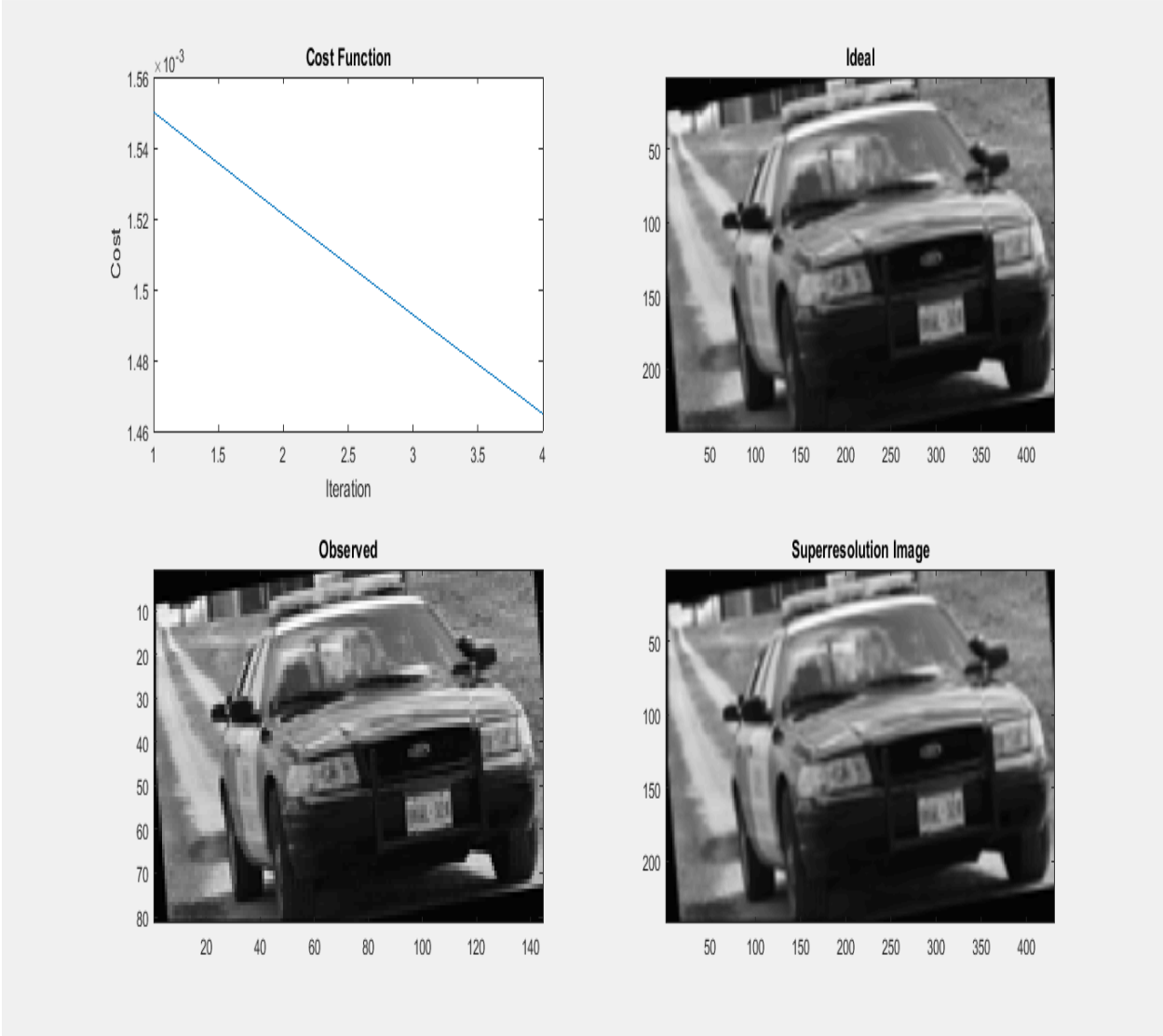


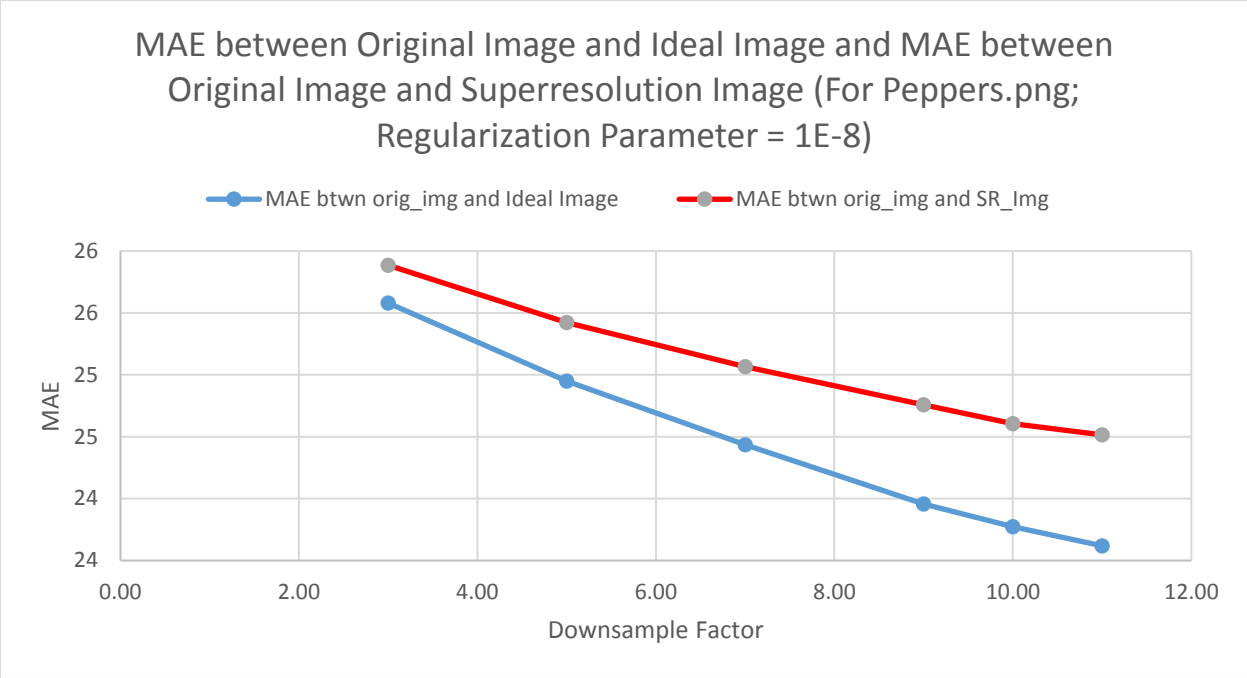
Figure 3: Output of GUI: Cost of the Superresoultion Algorithm (top left); and comparisons between Image after registration (Ideal Image, top right), As observed (with rotation and shift added to original image - bottom left) and after Superrsolution (bottom right).

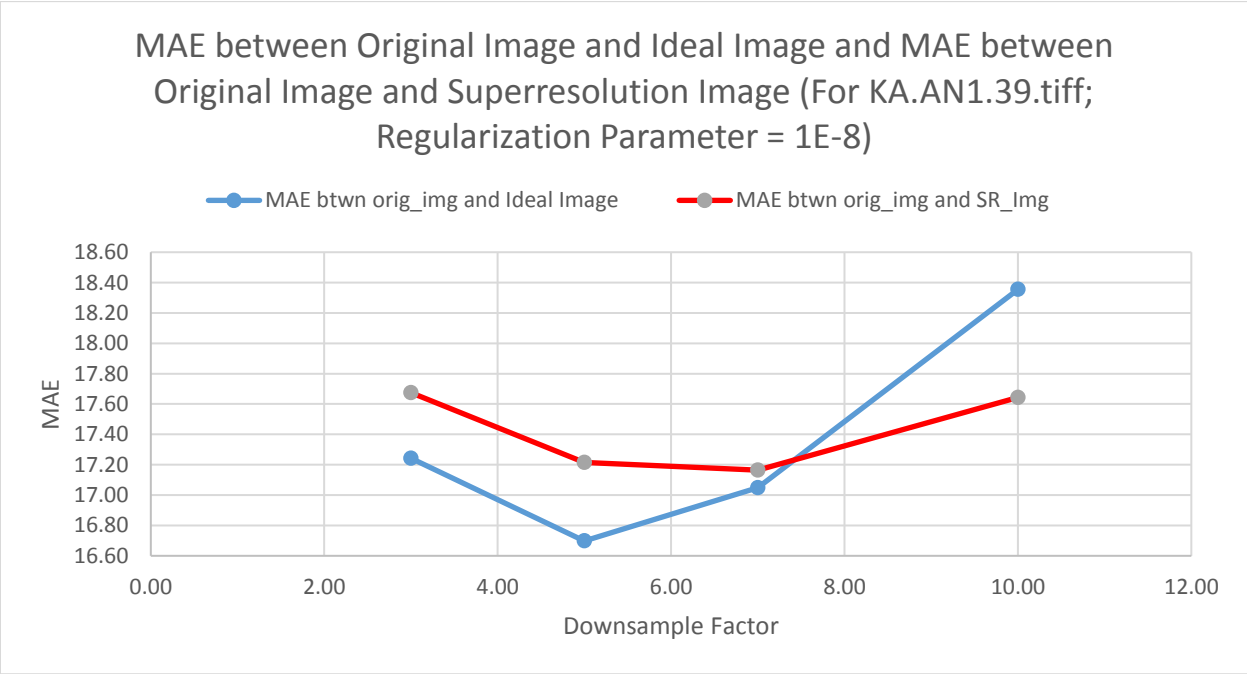
Test Results

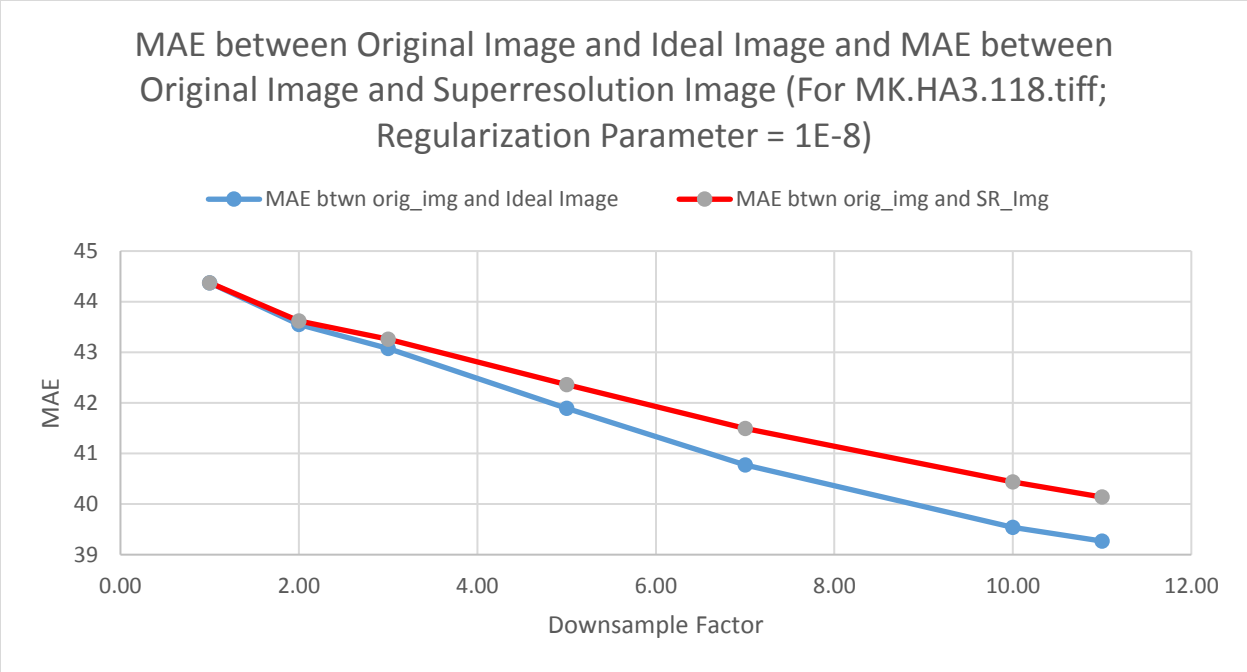
3 images (peppers.png, KA.AN1.39.tiff, MK.HA3.118.tiff, TM.HA2.181.tiff) were applied to the GUI with randomized shift and rotation of the original image and registration and restoration were applied to the images to arrive at superresolution.

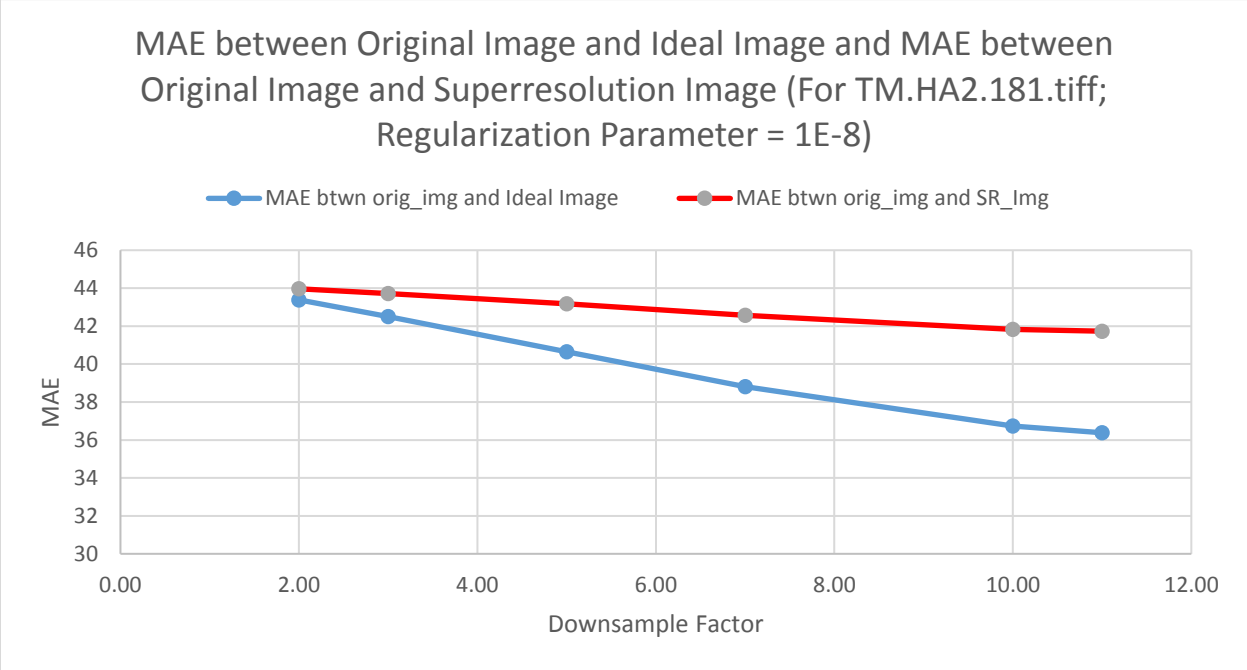
Downsampling Factor	Regularization Parameter	Image Used	MAE btwn orig_img and SR_Img	MAE btwn orig_img and Ideal Image	Delta_MAE
3	1.00E-08	peppers.png	25.8853	25.5801	0.3052
5	1.00E-08	peppers.png	25.4217	24.9485	0.4732
7	1.00E-08	peppers.png	25.0657	24.4336	0.6321
9	1.00E-08	peppers.png	24.7570	23.9576	0.7994
10	1.00E-08	peppers.png	24.6046	23.7728	0.8318
11	1.00E-08	peppers.png	24.5149	23.6189	0.8960
3	1.00E-08	KA.AN1.39.tiff	17.6747	17.2432	0.4315
5	1.00E-08	KA.AN1.39.tiff	17.2158	16.6978	0.5180
7	1.00E-08	KA.AN1.39.tiff	17.1645	17.0484	0.1161
10	1.00E-08	KA.AN1.39.tiff	17.6421	18.3546	-0.7125
1	1.00E-08	MK.HA3.118.tiff	44.3730	44.3730	0.0000
2	1.00E-08	MK.HA3.118.tiff	43.6214	43.5473	0.0741
3	1.00E-08	MK.HA3.118.tiff	43.2603	43.0717	0.1886
5	1.00E-08	MK.HA3.118.tiff	42.3601	41.8889	0.4712
7	1.00E-08	MK.HA3.118.tiff	41.4952	40.7719	0.7233
10	1.00E-08	MK.HA3.118.tiff	40.4367	39.5397	0.8970
11	1.00E-08	MK.HA3.118.tiff	40.1371	39.2676	0.8695
2	1.00E-08	TM.HA2.181.tiff	43.9656	43.3825	0.5831
3	1.00E-08	TM.HA2.181.tiff	43.7223	42.5000	1.2223
5	1.00E-08	TM.HA2.181.tiff	43.1845	40.6529	2.5316
7	1.00E-08	TM.HA2.181.tiff	42.5753	38.8106	3.7647
10	1.00E-08	TM.HA2.181.tiff	41.8303	36.7369	5.0934
11	1.00E-08	TM.HA2.181.tiff	41.7326	36.3800	5.3526

Table 1: Input parameter values and corresponding MAE values for the 4 images entered into the GUI and a comparison between the MAE values


MAE between Original Image and Ideal Image and MAE between Original Image and Superresolution Image (For Peppers.png; Regularization Parameter = 1E-8)
MAE btwn orig_img and Ideal Image
MAE btwn orig_img and SR_Img
MAE
26
26
25
25
24
24
0.00
2.00
4.00
6.00
8.00
10.00
12.00
Downsample Factor


MAE between Original Image and Ideal Image and MAE between Original Image and Superresolution Image (For KA.AN1.39.tiff; Regularization Parameter = 1E-8)
MAE btwn orig_img and Ideal Image
MAE btwn orig_img and SR_Img
MAE
18.60
18.40
18.20
18.00
17.80
17.60
17.40
17.20
17.00
16.80
16.60
0.00
2.00
4.00
6.00
8.00
10.00
12.00
Downsample Factor





For all but one image, the MAE of the Superresolution image is higher and less desirable than the MAE of the Ideal image. Future considerations to improve the Superresolution MAE would include increasing the number of downsampled images used for image registration and using a different interpolation method in the Superresolution algorithm (e.g., bicubic, nearest neighbor, and spline) as opposed to bilinear interpolation.

Conclusion

We investigated a Superresolution technique for enhancing a series of undersampled IR images called Uncontrolled microscanning. This technique was applied to RGB imagery (grayscale) by creating synthetic data having random translation and rotation shifts. The synthetic data was registered to a reference image using the algorithm proposed by Hardie et. al, then a high resolution image was reconstructed from the registered images. Due to the larger size of our imagery, it is not likely suited for real-time applications of creating superresolution images from HD cameras.Many aspects of the original algorithm were not relevant since our detector specifications were much better than the target device (IR v. RGB).Tested w/ simulated data; shows a minor improvement to original data so far, more investigation is necessary to discover parameter dependence on reconstruction quality.

References

1. E. A. Kaltenbacher and R. C. Hardie, “Infrared Image Registration and High-Resolution Reconstruction Using Multiple Translationally Shifted Aliased Video Frames”.

2. R. C. Hardie, K. J. Barnard, and E. E. Armstrong, “Joint MAP Registration and High Resolution Image Estimation Using a Sequence of Undersampled Images,” IEEE Transactions on Image Processing, Vol. 6, No. 12, Dec. 1997, pp. 1621-1633.