Temporal sampling of forward looking infra-red imagery for sub-resolution enhancement post-processing

Serge Dutremble

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
Temporal Sampling of Forward Looking Infra-Red Imagery for Sub-Resolution Enhancement Post-Processing

By
Major Serge Dutremble, CD.
B. S. Collège Militaire Royal de St-Jean
(1982)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Center for Imaging Science of the Rochester Institute of Technology
July 11, 1995

Signature of the Author

Accepted by Dana G. Marsh          October 24, 1995
Coordinator, M.S. Degree Program
M.S. DEGREE THESIS

The M.S. Degree Thesis of Major Serge Dutremble has been examined and approved by the thesis committee as satisfactory for the thesis requirement for the Master of Science Degree.

Dr. John Schott, Thesis Advisor

Dr. Harvey Rhody

Dr. John Mason

Sept 30, 1995
Temporal Sampling of Forward Looking Infra-Red Imagery for Sub-Resolution Enhancement Post-Processing

I, Major Serge Dutremble, hereby grant permission to the Wallace Memorial Library of RIT to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature: ____________________________

Date: 11 July 95
ABSTRACT

Forward Looking Infra Red (FLIR) and RADAR images are examples of low resolution and inherently noisy images. These images are often very sensitive to weather conditions or imaging system characteristics and are usually difficult to evaluate. Interpretation of such images is often restricted to specialists with many years of experience. We propose a quantitative evaluation of two Post-Processing methods of combining temporally close sequential video FLIR images to produce sub-pixel resolution still images. Although we can conceive many different methods to increase the resolution of a video image, this Thesis studies methods based on subpixel interpolation and subpixel separate sampling. The difficulties, requirements and performances of each method are compared and their advantages and disadvantages are rated with respect to their applications. We also investigate theoretically an additional method at Appendix 1 which has been unsuccessful experimentally.
ACKNOWLEDGMENTS

I wish to acknowledge Mr. Tim Gallagher for his help in capturing FLIR video frames.
DEDICATION

I wish to dedicate this Thesis to my two sons, Êrik and David, for their support during my frequent periods of stress.
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................. i

ACKNOWLEDGMENTS ................................................................................................. ii

DEDICATION .................................................................................................................. iii

TABLE OF CONTENTS ................................................................................................. iv

LIST OF FIGURES ......................................................................................................... vii

INTRODUCTION ............................................................................................................. 1

BACKGROUND .............................................................................................................. 2

THEORY .......................................................................................................................... 4

IMAGING SYSTEM CHARACTERISTICS ..................................................................... 4

*Imaging System* ........................................................................................................... 4

Sensor Sub-system ......................................................................................................... 4

Optics ............................................................................................................................... 5

Detector Array ............................................................................................................... 5

Electronic Sub-system .................................................................................................. 7

APPROXIMATIONS ....................................................................................................... 8

*Reality - Nyquist Limit* ............................................................................................... 8

*Physical Detector Size* ............................................................................................... 9

*Electronic Process Blurring - Resampling and Quantization Noise* ......................... 10

*Optical System Blurring* ............................................................................................ 11

FREQUENCY REPRESENTATION ................................................................................. 12

IMAGE REGISTRATION ................................................................................................. 14

INHERENT DETECTOR SAMPLING .............................................................................. 17
LIST OF FIGURES

FIGURE 1 - SCHEMATIC REPRESENTATION OF IMAGE CAPTURE WITH A TYPICAL IMAGING SYSTEM..............4

FIGURE 2 - SAMPLING GRID NOTATION........................................................................................................6

FIGURE 3 - ORIGINAL SAMPLING OF THE ELEMENT (M,N) AND ITS IMPACT ON THE FREQUENCY CONTENT.......9

FIGURE 4 - USUAL DETECTOR SPACING AND SIMPLIFIED DETECTOR SPACING. THE SIMPLIFIED SPACING ALLOWS EASY SYNTHETIC IMAGE GENERATION..............................................................10

FIGURE 5 - SENSOR POINT OF VIEW WHEN FAR FROM THE TARGET (A AND B) AND WHEN CLOSE TO THE TARGET (C). ........................................................................................................................................17

FIGURE 6 - LOW RESOLUTION SIGNAL PRODUCED BY THE DETECTOR SHIFTS.............................................18

FIGURE 7 - SIGNAL PRODUCED BY THE DETECTOR SHIFTS AFTER NEAREST NEIGHBOR INTERPOLATION.......20

FIGURE 8 - RECOMBINATION OF LOW-RES SIGNALS AND RESTORATION OF THE INITIAL HIGH-RES SIGNAL.20

FIGURE 9 - A 2-D, 3 X 3 PIXEL REPLICATION EXAMPLE.............................................................................21

FIGURE 10 - DETECTOR IMPULSE RESPONSE AFTER THE REPLICAION AND RECOMBINATION PROCESS......23

FIGURE 11 - IDEAL 2-D DETECTOR IR FOR METHOD 1 (NEAREST NEIGHBOR INTERPOLATION METHOD).....24

FIGURE 12 - A 2 X 2 SUPER-PIXEL IMAGE CONTAINING FOUR 2 X 2 SUB-SAMPLING FUNCTIONS C_{11}, C_{12}, C_{21} AND C_{22} APPROPRIATELY LOCATED.......................................................................................25

FIGURE 13 - TOTAL FOV (2 BY 2 ELEMENT DETECTOR)..................................................................................26

FIGURE 14 - IMPULSE RESPONSE FOR THE SUB-PIXEL SAMPLING METHOD.............................................27

FIGURE 15 - A 3 X 3 BI-LINEAR INTERPOLATION EXAMPLE WITHIN SUPER-PIXELS.................................46

FIGURE 16 - IMPULSE RESPONSE OF THE NEAREST NEIGHBOR LINEAR INTERPOLATION AND AVERAGING SUB-RESOLUTION ENHANCEMENT PROCESS.........................................................47

FIGURE 17- IMPULSE RESPONSE FOR THE LINEAR INTERPOLATION METHOD...........................................48

FIGURE 18 - TWO PIXELS REPRESENTING TWO DETECTOR ELEMENTS.......................................................49

FIGURE 19 - LINEAR INTERPOLATION PROCESS............................................................................................49

FIGURE 20 - INTERPOLATED SIGNAL.............................................................................................................49

FIGURE 21 - JITTER MATRIX OF (A) A FULL SET OF SUB-IMAGES, (B) A VERTICALLY REGISTERED SET OF SUB-IMAGES AND (C) A RANDOMLY REGISTERED SET OF SUB-IMAGES.................................54
Figure 22 - Interpolation of the pixel indicated by ↑ using the left and right closest neighbor.

Figure 23 - Position of sampling detector vs registration position.

Figure 24 - Chart giving the equivalence between any sub-pixel location and the known registered values.

Figure 25 - Interpolation method based on the distance of the neighboring samples.

Figure 26 - A) Image rebuilt using 5 vertically registered images and 20 interpolated missing images and B) the deblurred enhanced image.

Figure 27 - A) Image rebuilt using 13 randomly registered images and 12 interpolated missing images and B) the deblurred enhanced image.
INTRODUCTION

Forward Looking Infra Red (FLIR) and RADAR images are examples of low resolution and inherently noisy images. These images are often very sensitive to weather conditions or imaging system characteristics and are usually difficult to evaluate. Interpretation of such images is often left to specialists with many years of experience.

The poor quality of images produced by currently available FLIR systems restricts image interpretation to a limited number of highly trained and experienced individuals. The interpretation is often uncertain and sometimes completely inaccurate. It is suggested that substantial improvements in the accuracy of image interpretation could be achieved by using less noisy images having increased spatial resolution resulting in better detail definition. A better, improved image would not only contain increased information, but could also be extremely useful in equipment testing, training of observers, surveillance and information gathering.

Resolution enhancement of low resolution images, while maintaining spectral information, was demonstrated by Munechika et al., 1993 [1] using fusion techniques proving the concept that combining images under certain conditions produced better details. It is also well known that pixel averaging over many images reduces the noise standard deviation so similar combination process for sub-pixel resolution purposes could improve the signal to noise ratio (SNR) of the resulting image.

As resolution enhancement usually permits to clarify some pattern or object which is not “clear” or easily identifiable, we can use the methods presented here for enhancing images produced with many types of imaging systems. The synthetic\footnote{By synthetic, we mean that the image sequences were artificially produced.} images used for demonstrating the method effectiveness are general in nature and, although we discuss FLIR imagery in some detail, the same approaches can be used to enhance imagery from Low Light level TV, Radar or even regular home video.
This thesis will compare two specific methods for combining temporally close sequential video frames of FLIR imagery for the purpose of obtaining a less-noisy, higher-resolution still image with improved details of a target pattern. The methods discussed here will consist of subpixel interpolation and subpixel separate sampling. We investigate the methods theoretically and use the defined algorithms to digitally process synthetic video frame sequences to produce enhanced sub-pixel resolution images. An additional method involving linear pixel interpolation was also considered but was not successful in producing experimental enhanced images. This method is discussed theoretically at Appendix 1.

**BACKGROUND**

Until recently, if we assume good quality optics and electronics, most digital optical systems had their spatial resolution limited by the size of their individual detector window. Technological and cost considerations often make desirable obtaining better spatial resolution through image processing. Cox and Sheppard, 1986 [2] expressed the principle of invariance of information capacity as applied to optical systems. They proposed the concept that spatial resolution increase is achievable at the expense of another parameter. They formulated that the information capacity of an optical system is constant and the spatial frequency bandwidth is not. Their work allowed for a theoretical resolution improvement that would be attainable through image manipulation at the expense of less “important” image parameters. Peleg et al., 1986 [3] worked on improving low resolution camera images through camera motion. The laboratory approach they took commenced by iteratively guessing a high resolution image and then, through simulation of the imaging process, finding simulated low resolution images that would approach closest to the observed images. Minimizing the error between the simulated low resolution images and the observed images was the relevant metric. Although this work uses a “backdoor” approach to sub-pixel resolution, its laboratory results indicate that a marked improvement in detail is achievable. They also conclude that there is need for speed improvement and reliable sub-pixel registration if this approach is to be workable in real cases.

Hutbert’s, 1989 [4] approach relates closely to the subject of this thesis. He separated the imaging system component sources of error in two types; calibratable and non-calibratable. His findings identified that
three non-calibratable sources of error: the blur of the lens, the CCD Array spacing (source of incident image sampling) and the temporal noise on the CCD do not place fundamental limitations on resolution. These sources of error must be removed by other means than by calibrating the imaging system. Two jittering techniques proposed suggest that the accurate registration and combination of low resolution images can result in an increased spatial resolution. The increase is dependent on the frequency content of the image.

Rauchmiller and Schowengerdt, 1988 [5] used sub-resolution targets creating ½-pixel shifts of “point sources” to allow exploitation of sample-scene phase shifts to effectively measure the MTF of the Landsat Thematic Mapper. Their work indicated that a low-resolution optical system records sub-resolution detail information although the details are not readily visible in the images. Boulter, 1990 [6] proposed an approach for increasing the effective resolution of target imagery from mosaic detector arrays and established a basis for processing time sequential images for resolving sub-pixel features. The proposed technique reduced background clutter for the case of moving targets. A combination of images produces an effective higher spatial sampling rate and Nyquist cutoff frequency. Deconvolving the finite size of the detector elements as well as other sources of blurring further improves the image. One key element was the relationship between a high-resolution sub-sampling grid with the initial low-resolution detector array and detector size. Afterwards, further research concentrated on using sub-pixel concepts in very specific fields such as subpixel edge positioning (Cielo and Vaudreuil, 1991) [7] and part sizing (Shilling, 1991) [8]. Recently, the principle of sub-pixel resolution enhancement using image sequences was proven mathematically by Lévesque, 1994 [9] and described for a system with known micro-scanning mechanism. The sub-sampling grid approach was again discussed and the impact of an incomplete grid was discussed. Although it is desirable to have a complete set of low-resolution images recombined to produce the enhanced image, it was also proven that interpolation of the missing grid elements before deconvolution of the detector impulse response was providing acceptable results. Through all sub-resolution related work, the importance of accurate registration at sub-pixel precision is paramount. In most cases reviewed, the exact registration, whether obtained through accurate micro-movement of the detector assembly or through
other means of known “jitter” or target movement, is a major assumption. Registration at sub-resolution accuracy is the basis for successful Sub-Resolution improvement.

THEORY

The following sections will cover the mathematical background behind a typical FLIR imaging process. Each individual sub-system involved in the process is discussed. The treatment of individual images as well as the combination of multiple correlated images is covered and serves as a framework for developing the sub-pixel resolution techniques.

Imaging System Characteristics

![Figure 1 - Schematic representation of image capture with a typical Imaging System.](image)

We have to examine the characteristics and mechanisms inherent to a typical imaging system in order to identify the sources of image alteration. A simplified system is shown at Figure 1 and will be used to define the Imaging Process.

Imaging System

The imaging system can be broken down into separate sub-processes that introduce intermediate image representations of the scene with various degrees of alterations. The imaging system is broken down into two generic sub-systems: the sensor sub-system and the electronic sub-system.

Sensor Sub-system

The sensor sub-system represents the components involved in the image formation and image sampling before it is electronically processed. The following paragraphs define mathematically the image formation
through the sensor sub-system. The sensor sub-system is described in a general sense only. Additional knowledge of a specific sensor can improve any restoration process over the methods discussed in this thesis. As identified at Figure 1, Optics and Detectors are the components of the Optical Sub-system.

**Optics**

A scene $o(x,y)$ is altered by the optical components of the sensor sub-system (Optics) into a two-dimensional continuous image $s(x,y)$. The Optics have an Optical Transfer Function (OTF) equal to $h_0(x,y)$ which causes blurring of the initial scene. In its simplest expression, the detector input image $s(x,y)$ is equal to the convolution of the scene with the OTF of the optics:

$$s(x,y) = o(x,y) * h_0(x,y)$$

(1)

**Detector Array**

The detector component of the sensor sub-system, located at the focal plane of the Optics, samples and further alters the image $s(x,y)$ to form the discrete image $d(i,j)$. Defining the Dirac Delta Function $\delta(x)$ having the following properties:

$$\lim_{b \to 0} rect\left(\frac{x}{b}\right) = \delta(x)$$

(2)

$$\int \int \delta(x,y) dx dy = 1$$

(3)

$$f(x,y) \delta(i,j) = f(i,j)$$

(4)
we can use the image sampling function defined by Boulter, 1990 [6], $c(x, y, i, j)$ in terms of an array of delta functions\(^2\) that specify the center point of each detector element (Figure 2). The coordinates "x" and "y" represent the center of a delta function, and where "i" and "j" denote the corresponding row and column number in the sampled image.

\[d(i,j) = [\rho(x,y) * h_o(x,y) * h_d(x,y)](x,y,i,j) + n_d(i,j)\]  

(5)

again where x and y represent continuous variables and i and j represent discrete variables.

\(^2\) The Dirac Delta Function is assumed whenever there is mention of a Delta Function.
We can also express Equation (5) entirely with discrete images due to properties associated with the Dirac delta function used to define \( c(x, y, i, j) \) and specified at Equation (4):

\[
d(i, j) = o(i, j) * h_c(i, j) + n_d(i, j)
\]

(6)

**Electronic Sub-system**

The electronic sub-system globally includes the transfer to video tape and the image capture process causing some kind of resampling and quantization dependent on the type of equipment used. The resampling occurs when the detector discrete image is recorded on video. The video resolution, being different than the resolution of the detector array, causes a resampling of the image to occur. Quantization, on the other hand, describes an error introduced by replacing the specific gray value returned by the detector and recorded by the video process by a discrete value indexed from 0 (black) to 255 (white). The electronic process, as described in this section, is represented by the impulse response of the electronic process \( h_{ep}(i, j) \). The resulting image, \( g(i, j) \) is the image \( d(i, j) \) convolved by this impulse response with added noise \( n_{ep}(i, j) \):

\[
g(i, j) = d(i, j) * h_{ep}(i, j) + n_{ep}(i, j)
\]

(7)

The resampling and quantization aspect and their impact on some approximations is discussed further in the following section.

Combining Equations (5),(6) and (7), we can see the entire imaging process:

\[
g(i, j) = \left\{ \left\{ o(x, y) * h_{ox}(x, y) \right\} * h_{d}(x, y) \right\} * c(x, y, i, j) + n_d(i, j) \} * h_{ep}(i, j) + n_{ep}(i, j)
\]

(8)

and in discrete form:

\[
g(i, j) = \left\{ \left\{ o(i, j) * h_{ox}(i, j) \right\} * h_{d}(i, j) \right\} + n_d(i, j) \} * h_{ep}(i, j) + n_{ep}(i, j)
\]

(9)
Approximations

We suggest an imaging system representation which is an approximation of an actual system by neglecting the optical system blurring, the impact caused by the ratio of detector size and detector spacing, the electronic process blurring, the resampling and the quantization noise. These approximations are made with full knowledge that they may not seamlessly apply in real situations in order to simplify and generalize the methods we discuss. The electronic and optical system blurring is usually minimal if we use high quality components. As well, high quality equipment will enable adjustments so that the required resampling for the transfer to video tape or during the video capture process has minimal impact. Although the quantization noise is inherent in any digitalization, it is considered negligible if a sufficient number of steps are used and if the image uses the full dynamic range. Furthermore, the detector size/spacing ratio approximation was made initially with the assumption that its effect will be minimal for our test images. In reality corrections must be made for the detector size - spacing ratio as it is less than unity in today's detector array.

Reality - Nyquist Limit

Before we can effectively discuss the impact of the detector size/spacing ratio on the image quality, we must remember the basic sampling theorem which limits the frequencies captured by the detector sampling process. This theorem, known as the Nyquist limit theorem, can also be applied to any resampling process.
Figure 3 - Original Sampling of the element \((m,n)\) and its impact on the frequency content.

Figure 3 demonstrates the effect of the sampling theorem and that after sampling an image using an array with detector spacing \(w_2\), no spatial frequencies above the Nyquist limit, \(\frac{1}{2}w_2\), will be preserved. It is proposed that this fundamental resolution limit may be increased by combining multiple frames from the sensor (temporal oversampling), permitting deconvolution of the sampling aperture and sensor optics to be used to decrease the size of the minimum spatial features that may be resolved.

**Physical Detector Size**

The detector size is usually smaller than the detector spacing and the capture process (detector sampling) usually introduces additional blurring. However, for simplification, we will assume that the detector size equals the detector spacing and that the captured discrete image size is exactly the same as the detector discrete image size as identified at Figure 4. This simplification has minimum impact in the cases where the ratio detector size/spacing is close to unity!
Figure 4 - Usual detector spacing and simplified detector spacing. The simplified spacing allows easy synthetic image generation.

In other words, we say that every pixel in the captured image corresponds exactly to a detector element and that the capture process does not blur the image beyond the effect of detector size. We initially made this approximation with the assumption that it will have little effect on our test image restorations. Experimental results demonstrated that this assumption was correct for our test images but may not be optimal for real FLIR images.

**Electronic Process Blurring - Resampling and Quantization Noise**

The electronic process blurring is harder to identify as it is highly dependent on the type of equipment used for video recording and the capture process. It can be assumed to be minimal if high quality equipment is used and if the ratio of resampling density/detector density is close to unity. Additionally we will assume the quantization noise can be neglected. The electronic process is thus considered of minimal effect on the
overall image Quality\(^3\). After this approximation, the imaging system process (Equation (8)) then becomes:

\[
g(i,j) = \{o(x,y)h_d(x,y)\} + n_d(i,j)
\]  \hspace{1cm} (10)

**Optical System Blurring**

Knowledge of the optical system may allow image improvement through deblurring. However, for simplicity, we will assume the Optical system causes no blurring. I.E. \(h_{os}(x,y) = \delta(x,y)\) and \(s(x,y) = o(x,y)\). For many FLIR systems, this is a good approximation since the detector size substantially exceeds the diffraction limited blur spot. The imaging system process is further simplified as:

\[
g(i,j) = [s(x,y)h_d(x,y)](x,y,i,j) + n_d(i,j)
\]  \hspace{1cm} (11)

or, as \(c(x,y,i,j)\) is defined in terms of delta functions, in discrete terms:

\[
g(i,j) = s(i,j)h_d(i,j) + n_d(i,j)
\]  \hspace{1cm} (12)

Similarities with Equation (6) point out that our simplified imaging process reduces the Optical System down to the Detectors as \(g(i,j) = d(i,j)\). The images that will be manipulated in the proposed Sub-Resolution processes are considered to be produced by the detectors.

\(^3\) In fact, the electronic process of image capture usually has some effect which can be measured experimentally and corrected. This additional restoration step, although recognized and identified was not required on any synthetic test image by design and not performed on the FLIR test images within this thesis.
Our approximation that the detector size is equal to the detector spacing permits the generation of simple synthetic images as described in the Results section and points out that we consider the detector size our biggest source of blurring.

**Frequency Representation**

Image processing can sometimes be simplified using data transformations. The Fourier transform usually allows a different, but still exact representation of a linear and shift invariant imaging system process. The images become easier to manipulate due to properties of the Fourier transforms with respect to convolution. We will use the Fourier transform frequency representation of the images in order to first find registration information and second to remove as much as possible the blurring induced by the detector elements. In effect, we will use Fourier transform techniques to find the best estimate of an enhanced image based on a sequential number of low-resolution representations. Through the Fourier transform, we can represent the approximated scene image shown at Equation (12) in the frequency domain as:

\[
G(u,v) = S(u,v)H_d(u,v) + N_d(u,v)
\]  

(13)

where functions identified with capitals are used to represent the images in the frequency domain after the Fourier transform process and where \( u \) and \( v \) are discrete frequency variables.

We would like to find a best estimate \( \hat{s}(i,j) \) of the original signal \( s(i,j) \) from the image \( g(i,j) \) given by our system. We know that if an input has been filtered with some known blur, the best inverse filter is the Wiener-Helstrom filter [6]. The best discrete estimate \( \hat{s}(i,j) \) will only differ from the “ideal” image \( s(i,j) \) by the minimum mean-square error if our system’s image is “deblurred” with the Wiener-Helstrom filter:

---

4 Frequency domain is used to represent a Spatial domain image after Fourier transform.
\[ S(i, j) = g(i, j) * w(i, j) = s(i, j) * h_d(i, j) * w(i, j) + n_d(i, j) * w(i, j) \]  

(14)

or in the frequency domain, after the Fourier Transform:


(15)

where \( W(u, v) \) is the transfer function of the Wiener-Helstrom filter and \( w(i, j) \) is its spatial domain representation after inverse Fourier Transform. The transfer function of the Wiener-Helstrom filter is:

\[ W(u, v) = \frac{H_d^*(u, v)}{|H_d(u, v)|^2 + \frac{|N_d(u, v)|^2}{|S(u, v)|^2}} \]

(16)

where \( H_d^*(u, v) \) is the complex conjugate of \( H_d(u, v) \).

The discrete Fourier Transform of a rectangular sampling aperture (detector element) is described by the product of Sinc functions in the horizontal and vertical dimensions. We use this in Equation (13) to represent \( H_d(u, v) \). We also assume no a-priori knowledge of \( |S(u, v)|^2 \), the Power Spectral Density (PSD) of the image radiance we are trying to estimate, so we set it to a constant as a function of the spectral frequency. Similarly, the PSD of the additive noise, \( |N_d(u, v)|^2 \), is also assumed to be “white” and we can call the term \( \frac{|N_d(u, v)|^2}{|S(u, v)|^2} \), the inverse Signal-to-Noise Power Ratio (Power SNR) in the input image [6]. We can then assume that the inverse Power SNR is also a constant \( \Gamma(u, v) \) which gives us a constrained least-square Wiener-Helstrom filter:
\[
W(u,v) = \frac{H_d^*(u,v)}{|H_d(u,v)|^2 + \Gamma(u,v)}
\]

It must be noted that, although not used in the present work to simplify the computations, we can also use the Power spectrum of the blurred image instead of a constant in the inverse Power SNR as a better estimate for the constrained least-squared Wiener-Helstrom filter. The restored image \(\hat{s}(i,j)\) is obtained by multiplying the discrete Fourier Transform of the degraded image \(g(i,j)\) by the (constrained) Wiener-Helstrom filter \(W(u,v)\) and then taking the inverse Fourier Transform:

\[
\hat{s}(i,j) = FFT^{-1}\{W(u,v)FFT\{g(i,j)\}\}
\]

where \(FFT\) and \(FFT^{-1}\) denote the Fast Fourier Transform and its inverse.

**Image Registration**

In order to properly combine the separate video frames, the area of interest within every frame must be registered to each other at sub-resolution accuracy. This makes the assumption that the area of interest is sensibly more visible (high contrast) than the background! If the target and sensor are stationary, registration will likely be easier than where motion of the target or sensor occurs. In the latter case, every frame will require correction (coarse translation, rotation, scewing and possibly scaling) with respect to a reference frame before registration can be attempted. For simplicity, the registration process used for this thesis will assume no rotational, scewing or scaling effect present between separate frames. These effects and their restoration are discussed at Appendix 2. As a reasonable assumption, inherent sensor jitter will provide desirable uncontrolled sub-resolution micro movements between frames. The registration will be performed using discrete two-dimensional cross-correlation of every frame \(g_k(i,j)\) with respect to the same reference frame \(g_r(i,j)\), where the separate \(g_k(i,j)\) frames will not include the reference frame (I.E.: \(k \neq r\)). The mathematical representation of the two-dimensional cross-correlation is:
\[ \text{Corr}(i, j) = g_k(i, j) \otimes g_r(i,j) = \sum_{\alpha} \sum_{\beta} g_k(\alpha, \beta) g_r(\alpha - i, \beta - j) \]

where \( \otimes \) represents the cross-correlation operator. In the frequency domain, after Fourier Transform, the cross-correlation, due to its similarity with the convolution operator, becomes:

\[ \text{FFT} \{ \text{Corr}(i, j) \} = \text{CORR}(u, v) = G_k(u, v) G_r(-u, -v) \]

The discrete two-dimensional cross-correlation will have a maximum at a location representing the amount of translation required for the reference frame to be as “alike” as possible with the other frame. The digital registration process performed on the images must be accurate to sub-pixel resolution thus requiring that the images used in the registration process have the same number of discrete pixels as the high-resolution image. This can be accomplished with varying success by expanding the area of interest of the low-resolution images to high-resolution through pixel-replication (nearest neighbor interpolation), linear interpolation or other means, before performing the high-resolution cross-correlation. This translation information, retrieved from the cross-correlation function, can be broken down in a super-pixel translation (translation of the low-resolution images for low-resolution registration) followed by a sub-pixel translation (translation within a super-pixel to obtain the high-resolution registration):

\[ HRT = (LRT \times \text{REP}) + \text{SPIT} \]

where \( HRT \) represent the High-Resolution Translation, \( LRT \) represent the Low-Resolution Translation, \( \text{REP} \) represent the number of sub-pixels contained within every super-pixel and \( \text{SPIT} \) represent the Super-Pixel Internal Translation. The image translation expressed by Equation (21) is naturally applied to both dimensions of the images. The use of the nearest neighbor interpolation (also known as pixel replication) introduces a “blocky” aspect which affects the accuracy of the registration. This is why, for the purpose of this thesis, bilinear interpolation was used for all registration as this interpolation method was found to provide exact registration for all test images having a Signal to Noise Ratio (SNR) of 6 dB or better at a 4X.
resolution increase. The following table shows the resulting registration image between two squares at various SNR. The known (3,1) shift of the second image is reported accurately for SNR better than 6dB.

<table>
<thead>
<tr>
<th>No Noise</th>
<th>15 dB SNR</th>
<th>8 dB SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Registration at (3, 1)</td>
<td>Registration at (3, 1)</td>
<td>Registration at (3, 1)</td>
</tr>
<tr>
<td>6 dB SNR</td>
<td>5 dB SNR</td>
<td>3 dB SNR</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Registration at (3, 1)</td>
<td>Registration at (2, 1)</td>
<td>Registration at (-4, -1)</td>
</tr>
</tbody>
</table>

As discussed in detail at Appendix 2, repeating this process with different portions of two images would lead to performing sub-registration which would provide additional information as to the actual image distortion. If all portions of the images register with similar translation, we can say that ONLY translation distortion is present between the two images. If the registrations are showing different translation, we can identify rotation or scaling distortion and increase the accuracy of the registration information. This multiple registration approach was not necessary for the images chosen for this thesis.

**NOTE:** The particular image sequences that were used for this thesis have been specifically chosen to minimize the expected effects caused by target or sensor motion. All synthetic image sequences were designed in order to have all possible sub-pixel translations for a 4X resolution increase. The “real life” FLIR image sequence consist of a temporal sequence of captured video frames. The target is a stationary helicopter adjacent to a stationary military armored vehicle (tank). The sequence of frames are taken with a FLIR sensor mounted on a slowly approaching helicopter. All captured frames show the target as seen when the sensor is far away; thus reducing the impact of target size change as the sensor approaches the target. The registration between sequential frames should only identify translation distortion caused by up-down-left-right movements of the sensor (which would include sensor jitter). An attempt at obtaining an
ideal high resolution image was made by capturing a frame when the sensor was close to the target for comparison purposes with the resulting high-resolution image produced with the enhancement processes. This “ideal” image should be acceptable for qualitative comparisons as the helicopter on which the sensor is mounted is using a fixed descent angle. Quantitative evaluation proved to be unsuccessful as, at close range, the sensor is “on-top” of the target thus having a different point of view than with the images captured when the sensor is far away (changing aspect of image). Figure 5 illustrates the situation.

![Figure 5 - Sensor point of view when far from the target (A and B) and when close to the target (C).](image)

**Inherent Detector Sampling**

The detector averages the image in order to produce a single value. As an array of detectors is used, these values represent pixel elements that are closely correlated spatially. Due to the averaging process, a micro movement smaller than the detector size would produce a different set of values. Assuming a pixel is divided in two sub-pixel elements, Figure 6 shows an ideal 1-D case where only one detector element is used. We can see the different images that a detector can produce when it is shifted at various sub-pixel locations.

The recombination process will affect the detector transfer function when images pixel averaging or interpolation occurs.
Sub-Resolution Concept

The above paragraphs introduced the background theoretical information necessary to perform sub-resolution enhancement. In fact, the methods discussed below use a combination of the techniques identified previously. All sub-resolution approaches reviewed can be summarized with the following general steps:

1. Identify location of the low resolution image pixel at sub-resolution accuracy. I.E.: Perform very accurate registration.

2. Combine a number of low-resolution images following a specific method in order to increase the level of information contained in the resulting high resolution image.

3. Identify the detector Impulse Response as affected by any interpolation and averaging method used.

4. Remove the blurring induced by the detector using its modified Impulse Response.

Figure 6 - Low resolution signal produced by the detector shifts.
Steps 1 and 2 may take different forms depending on the approach taken while step 4 usually involves deconvolution processes.

**Approaches Discussed**

We discuss two different approaches in this thesis. Both involve registration at sub-resolution accuracies, image recombination and deblurring. For simplicity, they are identified as “Method 1 - Sub-Pixel Replication” and “Method 2 - Sub-Pixel Sampling” methods.

**Method 1 - Sub-Pixel Replication (Nearest neighbor Interpolation)**

The first method for sub-pixel resolution enhancement proposes to take every sub-image and perform a $N \times N$ pixel replication, $N \times N$ being the number of sub-pixels contained in a super-pixel in order to obtain the higher spatial resolution for each images. This is best achieved by doing a nearest neighbor interpolation with the four surrounding pixels. The resulting images present the same “blocky” appearance as the low-resolution image they result from. Registration information between one image and all other images is obtained by increasing the spatial resolution of the low-resolution images using linear interpolation$^5$ and performing a high resolution registration. The process follows by shifting the images obtained with nearest neighbor interpolation appropriately and then adding them together thus performing pixel averaging. Deconvolving the detector impulse response as modified by the interpolation and averaging process completes the sub-resolution enhancement. Figure 7 shows initial replicated and registered 1-D low resolution signals where the registration identifies a possible phase shift between each image. Figure 8 shows the recombination process. Low-resolution images produced by shifting the detector at sub-detector

---

$^5$ Linear Interpolation is used for obtaining the registration information but is not used for increasing the spatial resolution of the images before they are combined. It was found experimentally that using linear interpolation to increase resolution in order to find registration information provides accurate results in all cases where the Signal-to-Noise Ratio is better than 15 dB. Nearest neighbor interpolation was found very inaccurate.
spacing location are combined for reconstructing a blurred high-resolution signal. As averaging is
dependent on known sub-pixel locations of the detector, the signal sampling differs both in magnitude and
in phase. We obtain additional information about the details of the target signal if we consider all low-
resolution images combined together. The detector impulse response, ideally a "Rect" function, is blurred
by the image combination process. Deconvolution of the modified detector impulse response completes
the enhancement.

Figure 7 - Signal produced by the detector shifts after nearest neighbor interpolation.

Figure 8 - Recombination of low-res signals and restoration of the initial high-res signal.
From Figure 9, we see that every pixel from the low resolution images would become an \( N \times N \) high resolution super pixel. Registration of all high resolution images ensures the information is properly positioned before the combination of the images is performed. All separate high resolution images are likely to register at a different sub-pixel interval and the images combination will result in a higher resolution image (more pixels) with reduced noise variance. The image details will depend on the actual sub-pixel correlation shift being spread over the entire sub-sampling grid.

**Combination of Multiple Images**

We can express the combination of from many registered output images with the following summation, assuming \( PQ \) images and using the subscript \( k \) and \( l \) to identify each individual image shift condition, based on Equation (13):

\[
G'(u,v) = \frac{1}{PQ} \left\{ G'_{00}(u,v) + G'_{01}(u,v) + \ldots + G'_{0Q-1}(u,v) + \right. \\
G'_{10}(u,v) + G'_{11}(u,v) + \ldots + G'_{1Q-1}(u,v) + \\
\vdots \\
G'_{P-10}(u,v) + G'_{P-11}(u,v) + \ldots + G'_{P-1Q-1}(u,v) \right. \\
\left. + G'_{0-1}(u,v) + G'_{1-1}(u,v) + \ldots + G'_{P-1Q-1}(u,v) \right. \\
\left. + G'_{0Q}(u,v) + G'_{1Q}(u,v) + \ldots + G'_{PQ-1}(u,v) \right. \\
\left. + G'_{P-1Q}(u,v) \right\} 
\]  
(22)

or in summation form:

\[
G'(u,v) = \frac{1}{PQ} \sum_{k,l=0}^{P-1,Q-1} G'_{k,l}(u,v) 
= \frac{1}{PQ} \sum_{k,l=0}^{P-1,Q-1} [S'(u,v)H_d(u-k,v-l) + N'_{k,l}(u,v)] 
\]  
(23)
The factor $\frac{1}{PQ}$ is a normalization factor and the letter $I$ (superscript) is the indication of an interpolated, high resolution image. After some manipulation, we obtain an equation similar to our representation of a single image at Equation (13):

$$
\frac{1}{PQ} \sum_{k,l=0}^{P-1,Q-1} G'_{k,l}(u,v) = \frac{1}{PQ} \sum_{k,l=0}^{P-1,Q-1} \left[ S^I(u,v)H_d(u-k,v-l) + N^I_{k,l}(u,v) \right]
$$

$$
= \frac{S^I(u,v)}{PQ} \sum_{k,l=0}^{P-1,Q-1} H_d(u-k,v-l) + \frac{1}{PQ} \sum_{k,l=0}^{P-1,Q-1} N^I_{k,l}(u,v)
$$

$$
= S^I(u,v)H^+_d(u,v) + \left| N^I(u,v) \right|
$$

Where:

$k$ and $l$ (subscript) identify a specific image horizontal and vertical shift respectively.

$d$ (subscript) identify the detector.

$S^I_k$ is the fourier transform of the signal.

$N^I_{k,l}$ is the fourier transform of the image noise.

$H_d$ is the transfer function of the detector.

$H^+_d$ is the normalized high-resolution transfer function of the detector modified by the combination process.

$|N^I|$ is the mean of the fourier transform of the noise.

$PQ$ is the number of images.

We find that a combination of individual images will tend to average the detector induced noise.

In the spatial domain, the normalized impulse response of the detector as modified by the combination process, $h^+_d(u,v)$, is found by having a one-dimensional "edge" go through all possible sub-pixel shifts of the process. For a 4X spatial resolution increase, the registered edge values are:
and after differentiation, the one-dimensional impulse response is:

Graphically, the detector impulse response after the nearest neighbor interpolation and averaging process is seen at Figure 10.
The two-dimensional image of the detector impulse response for a completely filled 4 X 4 replication and after recombination is seen at Figure 11. The impulse response at Figure 11 can be used in Equation (17) to obtain a constrained least-squared Weiner-Helstrom filter. This filter can then be used to obtain the best estimate of a high resolution enhanced image using Equation (18). If we assume that the detector aperture equals the detector spacing and is "flat", then the recombination of the interpolated images (using the nearest neighbor method) has the same result as resampling. With the advantage of reducing the background noise. A detector with less perfect impulse response would require its combination with the interpolation process for obtaining the actual process impulse response. We discuss the detector impulse response further in our Approximation section below.

![Figure 11 - Ideal 2-D detector IR for Method 1 (nearest neighbor interpolation method).](image)

A similar method was investigated which uses linear interpolation in place of nearest neighbor interpolation. This method failed to produce acceptable experimental results and is described theoretically at Appendix 1.

**Method 2 - Sub-Pixel Sampling**

**Sub-sampling**

As defined by Boulter, 1990 [6], the detector sampling function $c(x, y, i,j)$ can itself be decomposed into a series of unique sub-sampling functions:

$$c(x, y, i, j) = c_{i1}(x, y, i, j) + c_{i2}(x, y, i, j) + \ldots + c_{iP}(x, y, i, j)$$

(25)
where \( P \) and \( Q \) represent the number of sub-sampling function along the \( X \) and \( Y \) coordinate respectively allowing for a total of \( P \times Q \) sub-sampling functions. An example of 4 recombined sub-sampling functions is shown at Figure 12.

![Figure 12](image)

Figure 12 - A 2 x 2 super-pixel image containing four 2 x 2 sub-sampling functions \( C_{11}, C_{12}, C_{21} \) and \( C_{22} \) appropriately located.

Substituting Equation (25) into Equation (11) gives:

\[
g(i, j) = \left\{ \begin{array}{l}
\{ (x, y) \ast h(x, y) \} c_{11}(x, y, i, j) + \\
\{ (x, y) \ast h(x, y) \} c_{12}(x, y, i, j) + \\
\{ (x, y) \ast h(x, y) \} c_{PQ}(x, y, i, j) \\
\end{array} \right\} + n_d(i, j)
\]

(26)

where the output image is expressed as a sum of sub-images with added noise. We could also represent the added noise into separate noise images each attributed to a sub-sampling function as:

\[
n_d(i, j) = (n_{11}(i, j) + n_{12}(i, j) + \ldots + n_{PQ}(i, j))
\]

(27)

Each separate image can be identified separately using the sub-sampling functions as follow:

\[
g_{kl}(i, j) = \{ (x, y) \ast h(x, y) \} c_{kl}(x, y, i, j) + n_{kl}(i, j)
\]

(28)
In order to completely represent an output image as a sum of sub-images, each having their associated added noise, we can combine Equation (26) and Equation (28) to obtain:

\[ g(i, j) = g_{11}(i, j) + g_{12}(i, j) + \ldots + g_{PQ}(i, j) \]  

(29)

High resolution registration information identifies low-resolution images to a sub-sampling function and allows for the combination of all sub-images into a high resolution image in accordance with Equation (25). Figure 13 illustrates this process as four low resolution images correspond to specific sub-sampling functions. Note that the low-resolution images are distributed in accordance with the results of the high-resolution registration.

![Diagram of sub-images registered to sub-sampling functions](image)

Figure 13 - Total FOV (2 by 2 element detector).

Provided accurate sub-pixel registration is obtained, all sub-images \( g_{kl}(i, j) \) can be combined together to obtain an image with enhanced resolution. As Method 2 places the pixels in accordance with the registration information into the sub-sampling function, no averaging occurs. The original detector impulse response is not modified and can be used as is for the process deblurring. Referring to Equations (28) and (29), we identify that if the sub-images are the result of an \( a \times b \) sub-sampling function (or detector arrays), then the appropriate combination of \( P \times Q \) registered sub-images could result in an image \( g(i, j) \) composed of \( aP \times bQ \) elements, showing an increased resolution as if the sampling
function (detector array) was itself actually composed of \( aP \times bQ \) elements. We also see that by definition, the combination of a complete set of independent sub-images have the same number of individual elements than \( g(i,j) \). A linear relationship between the detector output and the incident flux transmitted through the optics is assumed.

The process impulse response involves no interpolation and can be approximated to the impulse response of the detector (detector aperture). Assuming that each pixel of every low-resolution image were exactly represented by a detector element thus setting the initial low resolution impulse response to a delta function, a 4X spatial resolution enhancement as specified by this process will have a 4 pixel square rectangle as impulse response:

![Impulse response for the sub-pixel sampling method](image)

**Figure 14 - Impulse response for the sub-pixel sampling method**

By using pixel replication and high-resolution registration as in the first method proposed, the appropriate sub-sampling function associated with every sub-image is identified. Afterwards, all sub-image pixels are inserted accordingly in the high-resolution grid. Instead of replicating the sub-pixels and causing simple averaging in all sub-pixels, we could have local averaging only in the sub-pixel that will register similarly. Registration that will have information “falling” onto unfilled sub-sampling locations will add to the

---

6 The resolution increase obtained by combining an increasingly large number of images is probably not infinite. This aspect by itself could substantiate further research both theoretically and practically. It is not discussed as part of the scope of this thesis.
information content of the high resolution image other than by just reducing the noise variance. Assuming a completely filled high resolution grid, we can use the detector impulse response at Figure 14 for obtaining a constrained least-square Wiener-Helstrom filter using Equation (17) and a best estimate of the enhanced high resolution image can be obtained with Equation (18).

It is to be expected that no low resolution image will register at some sub-sampling location causing “gaps” in the high-resolution image. These “gaps” must be interpolated which causes additional errors. The error induced be this interpolation can be minimized by using a higher number of low-resolution images which should average the information whenever the sub-sampling functions are registered similarly. The “quality” of the final high resolution image will be improved if the “gaps” are minimized as we would reduce the level of interpolation (guessing) required.

It is clear that proper knowledge of the detector and of the interpolation process is of paramount importance if the blurring effect is to be minimized through image processing. It is also clear that proper knowledge of the sub-sampling functions, in relation with each other, is mandatory if a correct image combination is to be achieved. The resolution achievable is restricted by the accuracy of our registration.

The theoretical treatment of sub-resolution enhancement when the image sequence does not include all possible sub-sampling function is covered in detail by Levesque, 1994 [9] and is summarized at Appendix 3.

RESULTS
This section will describe the implementation details of each method and the restrictions that had to be accepted in order to obtain useful results. The Sub-Pixel Replication Method and the Sub-Sampling Method were used experimentally as described in the previous sections. A series of 6 synthetic test image sequences were processed by each method and the results are displayed in separate tables for easy visual comparison by the reader. Although the image chosen were small in size, it is believed that they allowed the demonstration of the sub-resolution principle and were sufficient for our evaluation.
Implementation Concerns

Both methods described in this thesis were implemented on an INTEL 486 DX-33 platform with 8 Mb of physical RAM using Microsoft Windows for Workgroup 3.11, a 16-bit graphical environment. All software routines were developed using Borland Pascal for Windows 7.0. Due to the hardware and operating software limitations, the size of the initial images were limited to 64 pixels square. This resulted, after a resolution increase of 4X, in limiting the Fast Fourier Transform (FFT) size to 256 pixels square.

The registration process required many images to be manipulated simultaneously in the frequency domain using Fourier techniques and this limitation completely used the system’s capacity as one 256 squared complex array required 2 Mb of virtual memory each. A 512 square complex array would have required 8 Mb of memory to process and was not possible with the hardware available.

Additionally, it was necessary to obtain images containing a “feature” to be enhanced that could fit in a 32 pixel square area. This limitation is again related to the memory restriction of the hardware since the correlation technique utilized for registration used a 32 pixel square window inside a 64 pixel square window. After the 4X resolution increase, this was reaching our complex array size limitation.

Synthetic Images

Due to the above limitations, most of the work was initiated using synthetic image sequences produced artificially instead of real FLIR images. Aldus Photostyler SE 1.1 was used for producing the following synthetic images sequences:

1. **Square.** A 12 X 12 Square was centered inside a 128 pixel square image area. From this “master” image, 25 additional images were produced by shifting appropriately the master image, followed by a 4 X 4 mosaic average which simulated a detector array and a resampling to obtain 32 pixel square images representing a 3 X 3 square appropriately “blurred”. Two noiseless series were made, one having the initial master with values of 255 Digital Count (DC) for the square and 0 DC for the background and one having 255 DC for the square and 127 DC for the background. The series above were repeated with added
Gaussian noise to produce two additional series with approximately a 15 dB SNR which was identified earlier as an acceptable level of noise for our processes.

2. **Pyramid.** A 12 X 12 pixels top view of a pyramid was generated inside a 128 pixel square image area. From this master image, a noiseless and a noisy series of 25 images were produced using the same technique as above.

3. **Bowl.** A 12 X 12 pixels top view of an overturned semi-sphere was generated in the center of a 128 pixel square image area. Again from this master, two series of 25 smaller images were produces simulating what a mosaic detector array would have showed after appropriate shifts of the master image with and without noise.

Additionally, some real FLIR images were captured from a US Army video tape using the ImageLab software available in the Undergraduate Imaging Lab of the Center for Imaging Sciences at RIT. The imagery is fairly noise free and has a feature small enough for our process to use. The "master" image was obtained by capturing and sizing appropriately a close-up of the feature which was also available on the video tape. This master image contains distortion that were impossible to correct and, as such, do not provide a good basis for our quantitative evaluation. As such, the true FLIR image results are included but were not used for our evaluation.

**Evaluation Metrics**

Two metrics were used to compare the methods used on this effort. The first metric is the usual Image RMS Error between the Master image and the Enhanced image for each method. The second metric, obtained from Mason et al., 1994 [10], is the Image Edge RMS Error between the Master image and the Enhanced image for each method.

**Image RMS Error**

This metric is given by calculating the RMS error between two images of the same dimensions, normalizing and calculating the percentage of error with respect to the actual Master image dynamic range:
\[ IRMSE = \frac{100}{\text{DynamicRange}} \sqrt{\frac{\sum_{i=0}^{\text{Width}-1} \sum_{j=0}^{\text{Height}-1} \left( DC_M(i,j) - DC_E(i,j) \right)^2}{(\text{Width} \times \text{Height})}} \]

where the subscript \( M \) refers to the Master Image and the subscript \( E \) refers to the Enhanced image.

**Image Edge RMS Error**

This metric is provided by finding the edges for two images to compare. A mask was then produced by using a threshold to set all edge pixels to "1" and non-edge pixels to "0". We can combine both masks with the OR operator in order to obtain a single mask encompassing all edges. Multiply the initial images with the mask produces images having non-zero pixels at the edge area locations. The RMS error between two "edge" images is then calculated as above with the difference that the normalization is performed by dividing with the total number of EDGE pixels instead of dividing by the total number of pixels. The percentage in then calculated with respect to the actual Master image dynamic range:

\[ IERMSE = \frac{100}{\text{DynamicRange}} \sqrt{\frac{\sum_{i=0}^{\text{Width}-1} \sum_{j=0}^{\text{Height}-1} \left( DC_{EM}(i,j) - DC_{EE}(i,j) \right)^2}{(\text{Number Of Edge Pixels})}} \]

where the subscript \( EM \) refers to the Edge Master image and the subscript \( EE \) refers to the Edge Enhanced image. The non-edge pixels will not affect the summations as they are zeroes.
**Quantitative Results**

All synthetic images shown are 64 x 64 pixels except the single blurred low-resolution images which is 16 x 16 pixels. All real FLIR images are 256 X 256 pixels except for the low-resolution image which is 64 X 64 pixels. Some images may have been re-centered to provide a better visual representation. The detector size is the same size as one pixel of the low-resolution image. In all cases, a combination of 16 low-resolution images was used. For the Sub-Sampling method, this constitutes a complete set at a 4X resolution increase. The case where an incomplete set of low-resolution images is used is considered outside the scope of the present work but was investigated by Lévesque, 1994 [9] and is reported at Appendix 3.

**Image of a Square**

The results indicate that method 2 provides a slight increase in accuracy.
The results indicate that, in this case, method 2 produces perfect resolution enhancement.

**Image of a square (50% gray background)**

**Results - Image of a square (50% gray background)**

<table>
<thead>
<tr>
<th>Original Image (Master)</th>
<th>Single Blurred Low-Resolution Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>(64 x 64 pixels image)</td>
<td>(16 x 16 pixels image)</td>
</tr>
</tbody>
</table>

**Method 1 (64 x 64 pixels images)**

<table>
<thead>
<tr>
<th>Registered image combination</th>
<th>Deblurred Image</th>
<th>2.52% Image RMSE (Error Image shown)</th>
<th>2.03% Image Edge RMSE (Error Image shown)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Method 2 (64 x 64 pixels images)**

<table>
<thead>
<tr>
<th>Registered image combination</th>
<th>Deblurred Image</th>
<th>0% Image RMSE (Error Image NOT shown)</th>
<th>0% Image Edge RMSE (Error Image NOT shown)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Image of a pyramid**

The results indicate a marked improvement when method 2 is used.

<table>
<thead>
<tr>
<th>Results - Image of a pyramid</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Image (Master)</strong> (64 x 64 pixels image)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 1 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 2 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
</tbody>
</table>
**Image of a bowl**

The results indicate that method 2 provides a marked improvement in image enhancement accuracy.

<table>
<thead>
<tr>
<th>Results - Image of a bowl</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Image (Master) (64 x 64 pixels image)</td>
<td></td>
<td>Single Blurred Low-Resolution Image</td>
<td>(16 x 16 pixels image)</td>
</tr>
<tr>
<td>Method 1 (64 x 64 pixels images)</td>
<td></td>
<td>Deblurred Image</td>
<td>8.06% Image RMSE (Error Image shown)</td>
</tr>
<tr>
<td>Registered image combination</td>
<td></td>
<td></td>
<td>13.67% Image Edge RMSE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Error Image shown)</td>
</tr>
<tr>
<td>Method 2 (64 x 64 pixels images)</td>
<td></td>
<td>Deblurred Image</td>
<td>4.39% Image RMSE (Error Image shown)</td>
</tr>
<tr>
<td>Registered image combination</td>
<td></td>
<td></td>
<td>3.30% Image Edge RMSE (Error Image shown)</td>
</tr>
</tbody>
</table>
**Image of a square (with 15 dB SNR)**

The results indicate that, when noise is present, method 1 provides a better enhanced resolution image.

<table>
<thead>
<tr>
<th>Results - Image of a square (with 15 dB SNR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Image (Master)</td>
</tr>
<tr>
<td>(64 x 64 pixels image)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 1 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
<tr>
<td>(Error Image shown)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 2 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
<tr>
<td>(Error Image shown)</td>
</tr>
</tbody>
</table>
**Image of a square (50% gray background and 15 dB SNR)**

Method 1 provides a marked improvement over method 2.

<table>
<thead>
<tr>
<th>Results - Image of a square (50% gray background and 15 dB SNR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Image (Master) (64 x 64 pixels image)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 1 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 2 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
</tbody>
</table>
**Image of a pyramid (15 dB SNR)**

Method 2 produces a better image than method 1 in this case.

<table>
<thead>
<tr>
<th>Results - Image of a pyramid (15 dB SNR)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Image (Master)</strong></td>
</tr>
<tr>
<td>(64 x 64 pixels image)</td>
</tr>
<tr>
<td><strong>Single Blurred Low-Resolution Image</strong></td>
</tr>
<tr>
<td>(16 x 16 pixels image)</td>
</tr>
</tbody>
</table>

| Method 1 (64 x 64 pixels images)         |
| Registered image combination             |
| Deblurred Image                          |
| 7.96% Image RMSE (Error Image shown)     |
| 14.92% Image Edge RMSE (Error Image shown) |

| Method 2 (64 x 64 pixels images)         |
| Registered image combination             |
| Deblurred Image                          |
| 3.78% Image RMSE (Error Image shown)     |
| 10.51% Image Edge RMSE (Error Image shown) |
**Image of a bowl (15 dB SNR)**

The results indicate an increased accuracy in the enhanced image when method 2 is used.

<table>
<thead>
<tr>
<th>Results - Image of a bowl (15 dB SNR)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Image (Master)</strong></td>
</tr>
<tr>
<td>(64 x 64 pixels image)</td>
</tr>
<tr>
<td><strong>Single Blurred Low-Resolution Image</strong></td>
</tr>
<tr>
<td>(16 x 16 pixels image)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 1 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Registered image combination</strong></td>
</tr>
<tr>
<td><strong>Deblurred Image</strong></td>
</tr>
<tr>
<td><strong>6.32% Image RMSE</strong></td>
</tr>
<tr>
<td>(Error Image shown)</td>
</tr>
<tr>
<td><strong>13.55% Image Edge RMSE</strong></td>
</tr>
<tr>
<td>(Error Image shown)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 2 (64 x 64 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Registered image combination</strong></td>
</tr>
<tr>
<td><strong>Deblurred Image</strong></td>
</tr>
<tr>
<td><strong>3.35% Image RMSE</strong></td>
</tr>
<tr>
<td>(Error Image shown)</td>
</tr>
<tr>
<td><strong>7.98% Image Edge RMSE</strong></td>
</tr>
<tr>
<td>(Error Image shown)</td>
</tr>
</tbody>
</table>
Image of a true FLIR target

Method 1 provided some results although the enhancement is not immediately visible to the observer. With a 6 dB SNR assumed, method two is not successful in deblurring the high resolution image. This may be caused by registration errors due to a lower SNR. The registration method used is not providing accurate results for images having less than 6 dB SNR. Furthermore, as method 2 “fills in” the sub-resolution grid without image averaging, registration errors can induce high frequencies that will be emphasized during the deblurring process. The error images and values, when produced with the inaccurate master image were very large and are considered misleading, especially in the case of method 2 where high frequencies were visibly induced. The use of these methods on real FLIR images must be furthermore investigated. The effect of Signal-to-Noise ratio must be further defined as it has an important impact on the registration process used. As identified before, all previous work on sub-resolution increase was done with the assumption that registration was perfectly achieved, which is not the case with this sequence of images. The poor results obtained with the available FLIR sequence made them unsuitable for the qualitative evaluation.

<table>
<thead>
<tr>
<th>Results - Image of a true FLIR target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Image (Master)</td>
</tr>
<tr>
<td>(256 X 256 pixels image)</td>
</tr>
<tr>
<td>Single Blurred Low-Resolution Image</td>
</tr>
<tr>
<td>(64 X 64 pixels image)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 1 (256 X 256 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
<tr>
<td>Deblurred Image</td>
</tr>
<tr>
<td>13.21% Image RMSE</td>
</tr>
<tr>
<td>6.41% Image Edge RMSE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method 2 (256 X 256 pixels images)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered image combination</td>
</tr>
<tr>
<td>Deblurred Image</td>
</tr>
<tr>
<td>21.93% Image RMSE</td>
</tr>
<tr>
<td>23.23% Image Edge RMSE</td>
</tr>
</tbody>
</table>
Qualitative Results

The qualitative results were obtained by doing a one page questionnaire asking to indicate for eight image series, which of two images represented more closely a target image. All synthetic images were used and were displayed in random order on the page. The questionnaire can be found at Appendix 4.

The following results were tabulated after asking a population of 48 untrained observers:

<table>
<thead>
<tr>
<th>Image Series</th>
<th>% of population that found Method 1 better</th>
<th>% of population that found Method 2 better</th>
</tr>
</thead>
<tbody>
<tr>
<td>Square</td>
<td>54</td>
<td>46</td>
</tr>
<tr>
<td>Square (50% gray background)</td>
<td>12</td>
<td>88</td>
</tr>
<tr>
<td>Square (15 dB SNR)</td>
<td>96</td>
<td>4</td>
</tr>
<tr>
<td>Square (15 dB SNR and 50% gray background)</td>
<td>98</td>
<td>2</td>
</tr>
<tr>
<td>Bowl</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Bowl (15 dB SNR)</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Pyramid</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Pyramid (15 dB SNR)</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

The qualitative results indicate that, visually speaking, method 2 is consistently better for images that include soft edges. For images with high contrast and sharp edges, method 1 was better when noise was present and Method 2 was better with the image which has medium contrast (50% background). Although the resulting images differed between methods, the qualitative evaluation provided a high level of indecision for the Square.
Discussion

The methods described in this thesis were simplified in order to process only a certain type of image sequences. This was necessary in order to obtain speedy results.

Assuming that appropriate features can be localized at different areas of the image, the image distortion caused by rotation or scaling can be corrected by registering these additional features with respect to the master image and comparing the registration data obtained. From these results, rotation and scaling information can be retrieved and the images can be appropriately corrected. Although it was not necessary to correct our test images for rotation or scaling, it should be a mandatory step on “real life” images unless the images were obtained under a control situation that specifically disallow distortion other than translation.

Noise in the images can adversely affect the registration process as used in all methods investigated. It was found that images with a Signal-To-Noise ratio of less than 6 dB do not register accurately when the cross-correlation operator is used. For example, using a possible shift range of 4 sub-pixel locations, a SNR of 10 dB failed to register images with known sub-pixel shifts of ±1 and ±3 and instead reported sub-pixel shifts of ±2, ±4 or multiples of ±2 sub-pixel shift. Degrading the images to less than 5 dB resulted in completely erratic registration.

The following table summarizes the Quantitative and Qualitative results:

<table>
<thead>
<tr>
<th>Image</th>
<th>Best Quantitative method</th>
<th>Best Qualitative method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Square</td>
<td>2</td>
<td>1 or 2</td>
</tr>
<tr>
<td>Square (15 dB SNR)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Square (50% gray background)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Square (50% Gray background and 15 dB SNR)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pyramid</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Pyramid (15 dB SNR)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bowl</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Bowl (15 dB SNR)</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
The qualitative results seem to agree with the quantitative results in all cases except the square. This discrepancy is not determinant as the quantitative error results for the Square were very low and very close, and it is understandable that the qualitative results became a question of preference for the observer. The enhanced images and the master images were:

<table>
<thead>
<tr>
<th>Master Image</th>
<th>Enhanced Image - Method 1</th>
<th>Enhanced Image - Method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Master Image" /></td>
<td><img src="image2.png" alt="Enhanced Image - Method 1" /></td>
<td><img src="image3.png" alt="Enhanced Image - Method 2" /></td>
</tr>
</tbody>
</table>

Note that all images are cropped to 32 X 32 pixels from 64 X 64 pixels to show more of the area of interest. It is visible that method 2 restored the square more uniformly than method 1. However, method 1 enhancement was identified as more similar to the master image as the average square gray value was closer to the master image. All error calculated in the dark areas are not visible to the observer and did not influence a decision towards method 2. In any case, both methods produced very close error values for the two metrics used and the qualitative evaluation indicated a higher level of indecision for these images.

**CONCLUSIONS**

Considering the result discrepancy noted for one image sequence, it is fair to state that generally method 1 will provide better sub-pixel resolution enhancement for noisy images representing sharp edges and high contrast. Possible applications do not include FLIR images where an increase in image detail is desired. Typical application would be to sharpen an image where the smallest detail element is expected to be larger than the desired super pixel size but which can be positioned within sub-pixel accuracy. Sub-pixel localization of edges should be easily performed using sub-pixel nearest neighbor interpolation method.

Method 2 should be used for enhancing the resolution of images having soft edges whether they include noise or not. Although not confirmed with the treatment of the FLIR image, we believe based on the synthetic image results, that it should be more suited to the resolution enhancement of high contrast FLIR
images which contain slow varying gradients with good SNR. A typical application would be the detail enhancement of shapes with lower feature gradient. I.E. Soft edges.

It was demonstrated that images having a SNR less than 6 dB do not enhance well with the methods proposed due to the inaptitude of obtaining accurate registration information.

Real FLIR image processing was inconclusive as the image sequence used did not contain an accurate master image and the SNR of the FLIR images could not be identified to be better than 6 dB. The SNR of real FLIR image may be too low for processing using the method presented. However, additional work is required to ascertain this.

**RECOMMENDATIONS**

It is recommended that additional work be done to ascertain the impact of noise on the registration process as defined in this thesis. Additionally, different registration methods could be investigated to see if a better method can be defined. This is a problem largely documented which is often eluded to by using micro-scanning techniques. Most of these techniques are successful under controlled experimentation and are not sufficient in real life situations.

The methods defined should be used to perform sub-resolution enhancement of more Real FLIR sequences. Better knowledge of these real FLIR images can help identifying problems in the proposed methods.

Scale and rotation distortion should be included in the registration process in order to avoid errors in registering real images. Although synthetic images could be made to obtain appropriate sequences, these distortion are hard to control on real images.
REFERENCES


Appendix 1 - Method using Bi-Linear Interpolation

Sub-Pixel Linear or Higher Order Interpolation

This method initially suggested that instead of the pixel replication method, an interpolation approach may provide more acceptable results using linear, cubic spline or other higher order interpolation method. We will try to use the same approach as with the nearest neighbor interpolation and apply it with Bi-linear interpolation. A two-dimensional (Bi-)linear interpolation is shown at Figure 15.

![Figure 15 - A 3 x 3 Bi-Linear interpolation example within super-pixels.](image)

We can easily use Bi-Linear interpolation to augment the spacial resolution of each images of our sequence. In fact, this was used in both our successful methods as a better means to find accurate registration. After performing this registration, the high-resolution registered images are then combined as in the method 1 (Sub-Pixel Replication) by averaging every registered sub-pixel, still improving the SNR. We can assume that, similarly to the nearest neighbor interpolation method, the detector impulse response would be modified by the linear interpolation and averaging process and could then be used to deconvolve the high-resolution image.

We find that the detector impulse response can be calculated again by having a one-dimensional “edge” go through the process. For a 4X spatial resolution increase, the registered edge values are:
and after differentiation, the one-dimensional impulse response is:

<table>
<thead>
<tr>
<th>IR</th>
<th>0</th>
<th>4</th>
<th>12</th>
<th>24</th>
<th>40</th>
<th>47</th>
<th>40</th>
<th>24</th>
<th>12</th>
<th>4</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td>0</td>
<td>22</td>
<td>65</td>
<td>130</td>
<td>217</td>
<td>255</td>
<td>255</td>
<td>217</td>
<td>130</td>
<td>65</td>
<td>22</td>
</tr>
<tr>
<td>(Normalized)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

graphically, the detector impulse response after the nearest neighbor linear interpolation and averaging sub-resolution enhancement process is shown at Figure 16.

![Graph](image-url)

**Figure 16 - Impulse response of the nearest neighbor linear interpolation and averaging sub-resolution enhancement process**
The two-dimensional image of the impulse response for a completely filled interpolation grid is seen at Figure 17:

![Image of impulse response](image)

**Figure 17- Impulse response for the linear interpolation method**

This method should reduce the edge contrast between super-pixels. Although promising, this approach does not work! By looking more into the conceptualisation of using linear interpolation, we see that it introduces information into the image that is not initially provided by the detectors. Our original system assumes the detector elements to have a flat response as represented at Figure 18. This being the ideal representation of the detector element impulse response. Linear interpolation, which is in effect a convolution with a triangle function ($tr$) (Figure 19) between pixels, changes our system by altering the detector impulse response. As the values in-between the known pixel values are interpolated, we introduce errors which are not produced by our system. Although the linear interpolation produces a smoother image (Figure 20), many samples are replaced by erroneous values. It is suggested that this is the reason why the “enhanced” blurred high resolution image produced using Linear interpolation can not be deblurred using such a modified detector impulse response.
Figure 18 - Two pixels representing two detector elements.

Figure 19 - Linear interpolation process.

Figure 20 - Interpolated signal.

The exact approach used in the nearest neighbor interpolation method is not applicable if we use linear interpolation. As this method was originally expected to work in a similar fashion, it was tried as defined above. The Impulse response was calculated and deblurring was attempted on the test images. The
experimentation showed that deblurring using the detector Impulse response does not work and that any other high order linear interpolation method should provide similar results.

On the other hand, we have to remember that we simplified our detector to have the detector width equal to the detector spacing. The use of linear and higher order interpolation may prove to be useful in interpolating the values between detector elements in cases where the detector width is smaller than the detector spacing, which is usually the case in today's detector assemblies.
Appendix 2 - Image Registration if Distortions are Present

There is five basic types of image distortion which can be easily corrected if present. They are Image Shift, Image Scale, Image Skew, Image Perspective and Image Rotation. In this thesis, we specifically designed the synthetic images to contain no distortion except for a sub-pixel shift detected only when the spatial resolution is increased through interpolation. In real life images, there is always a strong possibility to have to manipulate images which in addition to the sub-pixel shift, will have some or all other distortions present at some degree. This appendix will discuss the five distortions types identified above and their correction through mathematical manipulation. These corrections are merely coordinate transformations as discussed in “Techniques for Image Processing and Classification in Remote Sensing”, Robert A. Schowengerdt, 1983, pp. 106-110. The coordinate transformations are outlined in the following paragraphs:

Shift

\[ x = a_0 + x' \]
\[ y = b_0 + y' \]

The values of \( a_0 \) and \( b_0 \) are zero degree distortion parameters. These shift parameters are provided by our registration analysis. The image can accurately registered by performing a \( a_0 \) pixel shift in the \( x \) axis and \( b_0 \) in the \( y \) axis.

Scale

\[ x = a_1 x' \]
\[ y = b_1 y' \]

The values of \( a_1 \) and \( b_1 \) are first degree distortion parameters. These scale parameters can be obtained by performing a piecewise registration. If we assume that no other distortion is present, the various registration values will be mapped to the scale equations. The image can accurately registered by using the scale equations for each pixels to relocate them appropriately. \( a_0 \) pixel position scaled in the \( x \) axis and \( b_0 \) in the \( y \) axis.

Skew

\[ x = x' + a_2 y' \]
\[ y = y' \]

The value \( a_2 \) is another first degree distortion parameter. This skewing parameter is again found by performing a piecewise registration and mapping the different results to the skewing equations.

Perspective

\[ x = a_3 x' y' \]
\[ y = y' \]

The value \( a_3 \) is a second degree distortion parameter. This skewing parameter is again found by performing a piecewise registration and mapping the different results to the skewing equations.
Rotation  
\[ x = a_1 x' + a_2 y' \]
\[ y = b_1 x' + b_2 y' \]
\[ a_1 = b_2 = \cos \theta \]
\[ a_2 = -b_1 = \sin \theta \]

If the rotation equations are mapping the results of a piecewise registration, then we can say that we have rotational distortion. The distortion can be removed by using the rotation equation to relocalize the pixels appropriately.

The combination of various distortion types can be a difficult task to identify and can be difficult to implement. This is why, it was found outside the scope of this thesis. However, as accurate registration is of paramount importance in sub-resolution enhancements, further research on accurate registration using various computational techniques would be beneficial.
Appendix 3 - Partial Solution of the Sub-Resolution Problem

The majority of the information contained in this appendix is an adaptation from original work done by M. P. Lévesque in his May 1994 paper titled: "A Super-Resolution Technique For Micro-Scanned Image Sequences".

Both methods described in this thesis assume that a complete set of low-resolution images, or sub-images, is available, i.e., that there is an image available for every possible jitter position \((i,j)\). This implies that 16 images are required for a 4 X 4 resolution enhancement. However, it is possible that not all of these images are available. In such cases, it is possible to find partial solutions to the sub-resolution enhancement problem using a special sub-resolution algorithm. This situation is discussed further in the following section.

The Jitter Matrix

In order to be able to apply the method to every possible case of jittered image sets, let us define the jitter matrix. In fact, this jitter matrix tells us the sampling theory. Three cases of image sets are presented at Figure 21, the full set of registered images, the set identifying one-dimensional scanning and a set containing random holes, I. E., where no sub-image registered at some sub-pixel location. In the jitter matrix, a zero means that there is no image available for this registered sub-pixel position while a one means that there is an image available.

This convention can be refined by using fractions or integers greater than one. The fractions meaning that for these sub-pixel positions, the image has been obtained by interpolation and the integers greater than one meaning that many images are available for these sub-pixel locations. When many images are available for the same sub-pixel position, the noise can be reduced by the image averaging and the interpolation weighting factors can be modified to take into account the fact that the available information for this sub-pixel position is more reliable. However, this is an interpolation problem and is not overly discussed here.
In the two methods discussed in this thesis, we did not use a jitter matrix as we ensured every low
resolution image registered to a single sub-pixel location for the entire jitter matrix. In effect, it was like
having a complete jitter matrix filled with no duplication.

<table>
<thead>
<tr>
<th>i</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 21 - Jitter matrix of (a) a full set of sub-images, (b) a vertically registered set of sub-images
and (c) a randomly registered set of sub-images.

The sub-pixel algorithms developed previously are applicable only to the full set of sub-images shown at
Figure 21. If the jitter matrix is not completely filled, the data set is not well conditioned and the
reconstructed image cannot benefit from the deblurring process. We cannot deconvolve an object smaller
than the impulse response without creating artefacts and bouncing effects. An image reconstructed from a
complete set is smooth (no sharp details are present) and it is straightforward to deconvolve it, but this is
not the case when sub-images are missing.

The jitter matrix is useful to indicate when to pre-process the image set in order to get a pseudo complete
set of registered sub-images. It tells us if sub-images are missing and how to interpolate them using the
closest neighbors.
**Interpolation of Missing Images**

The method we have to interpolate the missing images is based on a linear interpolation of the closest neighboring images, i.e., the images having the most similar registration. There is room for more work on this problem and we can expect some improvement from a better interpolation method.

The method developed uses the linear interpolation of the closest known samples. The main problem is to determine which are the closest neighboring pixels. Figure 22 shows two different cases in one dimension. The pixel for which interpolation is required is marked by a vertical arrow. The first case is obvious and the interpolation is shown at Equation (32).

\[ I_2(n) = \frac{1}{2} I_1(n) + \frac{1}{2} I_3(n) \]  

(32)

where we use the format \( I_j(n) \) and:

- \( I \) is the sub-image
- \( j \) (subscript) is the registration location within a high resolution pixel \( (4 \geq j \geq 0) \)
- \( n \) is the high resolution pixel identifier

![Figure 22 - Interpolation of the pixel indicated by ↑ using the left and right closest neighbor.](image)

However, the second case shows a particular situation. In that case, the closest right pixel that should be \( I_6(n) \) (registered at sub-position 6) is in fact the pixel having the index \((n+1)\) of the sub-image having the farthest left registration value i.e. \( I_1(n+1) \). In other possible cases, the closest left pixel can be the pixel index \((n-1)\) of the sub-image having the farthest right registration value. This can be confusing but Figure 23 and Figure 24 explain this situation well. In the case of Figure 22b, the interpolation can be done using Equation (33).
\[ I_4(n) = \frac{2}{3} I_3(n) + \frac{1}{3} I_1(n + 1) \] (33)

Figure 23 - Position of sampling detector vs registration position.

The gray lines at Figure 23 represent the pixels we want to interpolate while the black lines represent the left and right samples we use in the interpolation.

<table>
<thead>
<tr>
<th>( I_0(n) )</th>
<th>( I_0(n+1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 24 - Chart giving the equivalence between any sub-pixel location and the known registered values.

In Figure 24, the two upper lines give the pixel index and the registration value corresponding to the sub-pixel locations -5 to 9. The bottom line represents a jitter matrix line repeated three times showing the registered sampling cycle.
The conventional linear-interpolation method determines the weighting factor by considering the distance to the neighboring sample. Thus, the method that uses the distance a and c produces the interpolation as shown at equation (34).

\[ B = w_l A + w_r C \]

where the left and right weighting factors \( w_l \) and \( w_r \) are \( w_l = \frac{c}{a + c} \) and \( w_r = \frac{a}{a + c} \).

![Diagram of interpolation method based on the distance of the neighboring samples.](image)

Figure 25 - Interpolation method based on the distance of the neighboring samples.

This method gives good results for a 1D interpolation. However, in 2D, many neighboring samples are available and a strategy must be established to select which samples can be used for the interpolation. This has not been done yet. We use just a multi-pass interpolation: a first pass that uses the horizontal neighbor and a second pass that uses the vertical neighbors. This process was good enough for the demonstration but we think it can be improved.

**Enhancement of Incomplete Sets of Registered Images**

The two cases of the incomplete images at Figure 21 have been investigated. Figure 26(a) shows the image rebuilt using five vertically registered images while Figure 26(b) shows the deconvolved enhanced image. Similarly, Figure 27 shows the rebuilt and enhanced images that used 13 randomly registered images corresponding to Figure 21.
Figure 26 - a) Image rebuilt using 5 vertically registered images and 20 interpolated missing images and b) the deblurred enhanced image.

As can be seen in Figure 26, when only a 1D set of registered images is available, the resolution is increased only on the considered axis. On the other axis, the pixels are only interpolated. In fact, we cannot expect anything else.

The case presented in Figure 27 is more interesting. The resolution is increased on both axes but the result is degraded by the presence of aliasing. This is normal since the number of pixels of the enhanced image is higher than the degrees of freedom of the image sequence. However, we can expect that a better interpolation method could only generate a little blur (on some pixels) and minimize the aliasing effect. In fact, two different scenarios were used during the interpolation step and two different aliasing patterns were obtained. Thus we can conclude that the interpolation method should be the object of future investigation.
Figure 27 - a) Image rebuilt using 13 randomly registered images and 12 interpolated missing images and b) the deblurred enhanced image.

The beauty of the interpolation method used along with the enhancement method is that it is possible to enhance the resolution of an image using any set of registered images with the same method. The interpolation process can be automatized and no ad-hoc solution are required to enhance a specific set of registered images.

Another advantage of this method is that, in the worst case, the enhanced image obtained cannot be worse than the original low-resolution images (if you have only one image available). On the other hand, in the best case, the enhanced image can show details that were not visible in the original images. But, in every case, the method is always applicable and can improve the image quality.
Appendix 4 - Qualitative Investigation Questionnaire

Qualitative Evaluation of Enhanced Image Results

For each row, please circle a or b, for the image which best represent the first image shown.
Circle the ROW NUMBER if you see no difference. Alignment is not important.

1. a. [Image 1] b. [Image 2]
2. a. [Image 3] b. [Image 4]
3. a. [Image 5] b. [Image 6]
4. a. [Image 7] b. [Image 8]
5. a. [Image 9] b. [Image 10]
6. a. [Image 11] b. [Image 12]
7. a. [Image 13] b. [Image 14]
8. a. [Image 15] b. [Image 16]