Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Abhishek Shriram Joshi

Entitled
Image Processing and Super Resolution Methods for a Linear 3D Range Image Scanning Device for Forensic Imaging

For the degree of Master of Science

Is approved by the final examining committee:

Mihran Tuceryan
Chair

Shiaofen Fang

Jiang Yu Zheng

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Mihran Tuceryan

Approved by: Shiaofen Fang 05/24/2012
Head of the Graduate Program Date

Graduate School Form 20
(Revised 9/10)

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

Image Processing and Super Resolution Methods for a Linear 3D Range Image Scanning Device for Forensic Imaging

For the degree of Master of Science

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Abhishek Shriram Joshi

Printed Name and Signature of Candidate

05/24/2012

Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

IMAGE PROCESSING AND SUPER RESOLUTION METHODS FOR A LINEAR 3D RANGE IMAGE SCANNING DEVICE FOR FORENSIC IMAGING

A Thesis

Submitted to the Faculty

of

Purdue University

by

Abhishek Shriram Joshi

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2012

Purdue University

Indianapolis, Indiana

ACKNOWLEDGMENTS

I would like to express my deep and sincere gratitude to my advisor, Dr. Mihran Tuceryan for his guidance and encouragement throughout my Thesis and Graduate studies. Dr. Tuceryan was very helpful and supportive during the entire process.

I also want to thank Dr. Shiaofen Fang and Dr. Jiang Zheng for agreeing to be a part of my Thesis Committee.

Thank you to all my friends and well-wishers for their good wishes and support. And most importantly, I would like to thank my family for their unconditional love and support.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Joshi, Abhishek Shriram. M.S., Purdue University, August, 2012. Image Processing and Super Resolution Methods for a Linear 3D Range Image Scanning Device for Forensic Imaging. Major Professor: Mihran Tuceryan.

In the last few decades, forensic science has played a significant role in bringing criminals to justice. Shoe and tire track impressions found at the crime scene are important pieces of evidence since the marks and cracks on them can be uniquely tied to a person or vehicle respectively. We have designed a device that can generate a highly accurate 3-Dimensional (3D) map of an impression without disturbing the evidence. The device uses lasers to detect the changes in depth and hence it is crucial to accurately detect the position of the laser.

Typically, the forensic applications require very high resolution images in order to be useful in prosecutions of criminals. Limitations of the hardware technology have led to the use of signal and image processing methods to achieve high resolution images. Super Resolution is the process of generating higher resolution images from multiple low resolution images using knowledge about the motion and the properties of the imaging geometry. This thesis presents methods for developing some of the image processing components of the 3D impression scanning device. In particular, the thesis describes the following two components: (i) methods to detect the laser stripes projected onto the

impression surface in order to calculate the deformations of the laser stripes due to 3D surface shape being scanned, and (ii) methods to improve the resolution of the digitized color image of the impression by utilizing multiple overlapping low resolution images captured during the scanning process and super resolution techniques.

CHAPTER 1: INTRODUCTION

In a wide variety of imaging applications it is crucial to capture the fine details in the image. The use of forensic imaging is one such application domain where the fine details in either color two-dimensional (2D) images or the fine details of three-dimensional (3D) depth images of evidence recovered in crime scenes may result in successful prosecution and conviction of criminals. Sometimes the success of such prosecutions depends on the existence of individualizing characteristics of the evidence that can be very small [1][2][3]. The ability to capture these fine details is related to the resolution of the image being captured.

In this thesis we describe the components of a 3D imaging device which is designed for capturing foot print and tire track impressions in crime scenes. The device works by detecting a laser stripe in an image captured by a video camera and moving this laser/camera assembly in a linear motion to generate the full 3D depth image as well as a 2D color texture image. The highly accurate detection and localization of the laser stripe enhances the resolution and accuracy of the depth image. The use of super resolution imaging techniques improves the accuracy and resolution of the captured 2D texture image.

Super Resolution is a vast topic encompassing many aspects of image processing. The particular contributions of this thesis to this topic are the following:

- The development of the image processing techniques for accurately detecting the laser stripe in the captured video frame images. We investigated two methods to accurately detect the position of the laser
    - Peak based detection
    - Edge based detection
- The development of a super-resolution technique which utilizes the linear motion model of the scanning device in order to improve the resolution of the captured 2D color texture image.

Chapter 2 gives a brief background and reviews the relevant prior research. Chapter 3 presents a brief description of the device, the methodology for laser stripe detection, and the methodology of obtaining super-resolution color texture images. Chapter 4 presents experimental results. Finally, Chapter 5 makes concluding remarks and suggests some future work.

CHAPTER 2: BACKGROUND

The resolution of a 2D image is affected by a number of factors such as the density of the sensing elements on the sensor array, the resolving power of the optical pathway, and the motion of the objects in the scene or the motion of the camera. Since 1970's, charge-coupled device (CCD) and CMOS image sensors have been widely used to capture digital images. These sensors are suitable for most imaging applications, but the current resolution and cost may not be sufficient for certain applications that require the ability to capture minute details [4]. Scientists or criminalists often need digital high resolution images, with no visible artifacts when the image is magnified, similar to 35mm analog film.

A direct solution to the problem is to reduce pixel size and increase pixel density by sensor manufacturing techniques. However, reducing the pixel size decreases the amount of light available to each of the pixels. This leads to shot noise that severely degrades the quality of the image. It is estimated that pixel size cannot be reduced beyond 40 μm^2 for a 0.35 μm CMOS process [4]. The current image sensor technology has almost reached this level. Another possible solution to achieve higher spatial resolution is to increase the size of the chip. However, increasing the size of the chip leads to an increase in capacitance. One promising approach to improve resolution is using signal and image processing

techniques to obtain a higher resolution image from multiple lower resolution images. This is called Super-resolution (SR) and is a process of estimating a single high resolution image or video from a given set of low resolution inputs obtained from slightly shifted viewpoints. The major advantage of using signal processing to achieve high resolution is its cost effectiveness because existing low resolution imaging systems can still be utilized. SR image reconstruction is very effective in applications where multiple frames of the same scene can be obtained, and the motion is to obtain such images is not very big or is very constrained.

In super resolution, the lower resolution images represent different looks of the same scene. The technique relies on the fact that if the motion of the camera is sufficiently constrained and there is overlap in the pixels of the images of different views of the scene, this overlapped information can be used to recover sub-pixel level image information to compute a higher resolution image. Thus, the low resolution images are subsampled and shifted with sub-pixel precision. Multiple scenes can be obtained from one camera with several captures or from multiple cameras located in different positions. The scene motions can be obtained using controlled motion in imaging systems, e.g., video sequence obtained from a camera mounted on a linear actuator. If the low resolution images are shifted by integer multiples of pixel units, no new information can be obtained to reconstruct the high resolution image because there would be no overlap between pixels of the low resolution images.

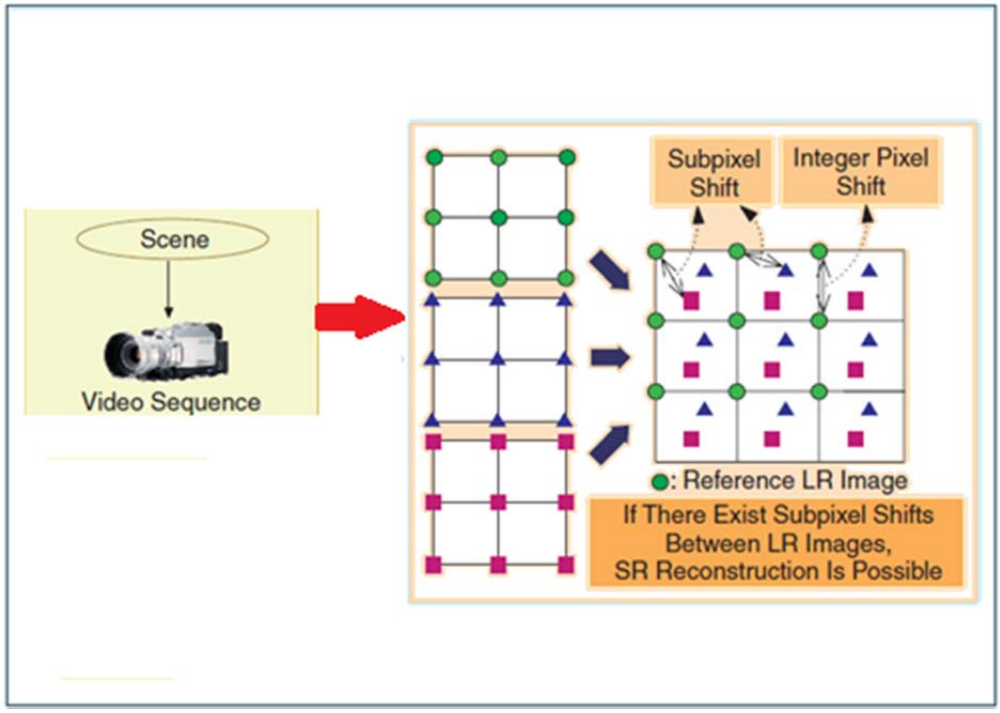

Figure 1 Basics of Super Resolution

During the process of recording a digital image, there is inherent loss of spatial resolution due to optical distortions (out of focus, diffraction limit, etc.), motion blur due to limited shutter speed, noise that occurs within the sensor or during transmission, and insufficient sensor density as shown in Figure 2. Super resolution also covers image restoration techniques that produce high quality images from noisy and blurred images.

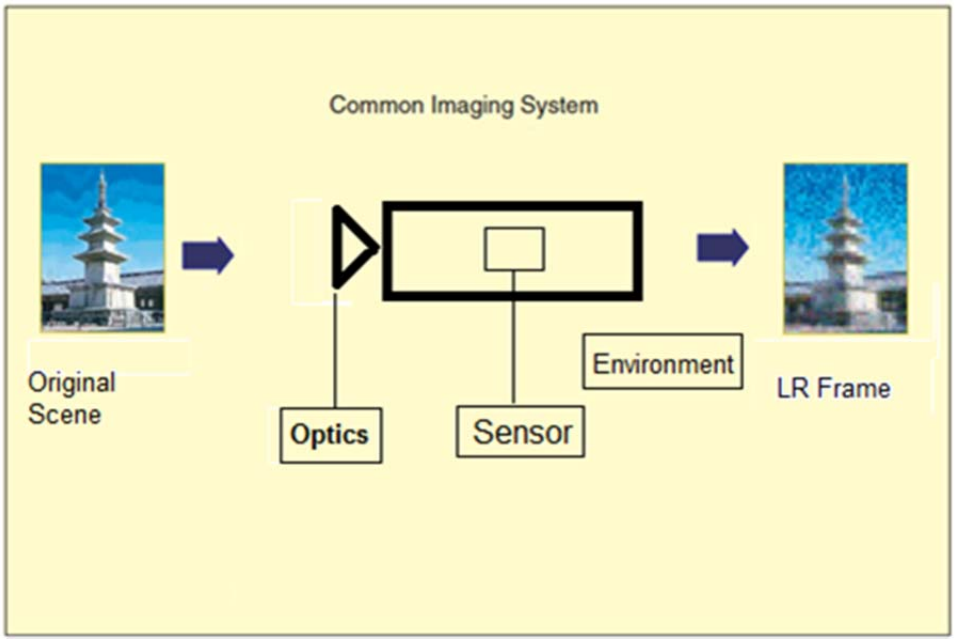

Figure 2 Image Acquisition System

Super Resolution reconstruction is an example of an ill-posed inverse problem [6] as a number of possible solutions exist for a given set of observed images. A common model for SR is stated in the following way: The low resolution input images are the result of projection of a high resolution image onto the image plane, followed by sampling. The main goal is to find the high resolution image which best fits this model given the observed low resolutions images. In the literature, there seems to be two broad approached to Super Resolution image reconstruction:

- Frequency Domain Methods
- Spatial Domain Methods

2.1 Frequency Domain

A large number of super resolution algorithms operate in the frequency domain. They are based on 3 fundamental principles [2]:

- Fourier Transform's shifting properties relating spatial domain translations and frequency domain properties.
- Aliasing relationship between the continuous Fourier transforms (CFT) and the discrete Fourier Transform (DFT).
- The original scene being band limited.

These properties can be used to formulate a set of equations relating the aliased DFT coefficients of the observed low resolution images to sample of the CFT of the unknown scene. By solving this set of equations we can get frequency domain coefficients of the original scene. We can now recover the original scene by computing the inverse DFT. In order to accurately formulate the set of equations it is necessary to have knowledge of the translation motion between frames to sub-pixel accuracy. Restrictions have to be placed on the inter-frame motion that contributes useful data since each low resolution image must have equations independent of each other.

The shifting property of the CFT relates spatial domain translation to the frequency domain as phase shifting given by

$$F_r(u,v) = e^{j2\pi(\Delta x_r u + \Delta y_r v)} F(u,v) \qquad \text{…Eq (1)}$$

If $f(x,y)$ is band limited, then $\exists\, L_u, L_v$ such that $F(u,v) \to 0$ for $|u| \geq L_u f_{s_x}$ and $|v| \geq L_v f_{s_y}$. If we assume that $f(x,y)$ is band limited, we can rewrite the aliasing relationship in the matrix form as,

$$\boldsymbol{Y} = \boldsymbol{\varphi F} \qquad \text{…Eq (2)}$$

$\boldsymbol{Y}$ is an $R \times 1$ column vector with the rth element being the DFT coefficients $\boldsymbol{y_r[k,l]}$ of the observed image $\boldsymbol{y_r[m,n]}$. $\boldsymbol{\varphi}$ relates the DFT of observed data samples of the unknown CFT of $f(x,y)$ contained in the $4L_uL_v \times 1$ vector $\boldsymbol{F}$.

Frequency domain super resolution methods have the advantage of being theoretically simple and computationally easy. They are also well suited for parallel computation as the equations of one observed image are independent of others. One major disadvantage is the limitation imposed by the assumption of a global translation motion model and the space invariant degradation models. It also has limited ability for inclusion of knowledge of spatial domain properties for regularization.

2.2 Spatial Domain Methods

In this class of methods, spatial domain properties are used to formulate the image formation and motion model in order to reconstruct the higher resolution image. The spatial domain observation model accommodates global and non-global motion, optical blur, motion blur, spatially varying point spread function (PSF), non-ideal sampling, and compression artifacts. Spatial domain reconstruction allows inclusion of a-priori constraints.

Let $\boldsymbol{X}$ be the SR image reconstructed from low resolution images $\boldsymbol{y_r}, \boldsymbol{r} \in \{\boldsymbol{1}, \boldsymbol{2}, .., \boldsymbol{R}\}\boldsymbol{X}$ and $\boldsymbol{y_r}$ are related by the following equation.

$$\boldsymbol{y_r} = \boldsymbol{H_r X}$$

$\boldsymbol{H_r}$ incorporates motion compensation, degradation effects and subsampling.

2.2.1 Interpolation of Non-Uniformly Spaced Samples

Motion compensation is used to register a set of low resolution images into a single, dense composite image of non-uniformly spaced samples. This dense composite image is then used to reconstruct a super resolved image using techniques for reconstruction from non-uniformly spaced samples. Degradation can be compensated by applying image restoration techniques. Iterative reconstruction techniques, based in the Landweber iteration have been used [3]. This method is overly simplistic; it cannot reconstruct significantly more content than present in a single low resolution image. Degradation models are limited, and no *a priori* constraints are used.

2.2.2 Iterated Backprojection

Given a SR estimate $\boldsymbol{X}$ and the imaging model $\boldsymbol{H}$, it is possible to simulate the low resolution images $\boldsymbol{Y}$ as $\boldsymbol{Y} = \boldsymbol{HX}$. Iterated backprojection (IBP) procedures update the estimate of the SR reconstruction by minimizing the back projection error between the j^{th} simulated low resolution image $\boldsymbol{Y^j}$ and the observed $\boldsymbol{Y}$ via the back projection operator $\boldsymbol{H^{BP}}$.

2.2.3 Stochastic SR Reconstruction Methods

In this method the SR reconstruction is treated as a statistical estimation problem. They have gained prominence due to their ability to provide a framework for the inclusion of *a priori* constrains necessary for satisfactory solution of the ill-posed SR inverse problem. The observed data $\boldsymbol{Y}$, noise $\boldsymbol{N}$ and SR image $\boldsymbol{X}$ are assumed stochastic. Consider the following equation:

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{N}$$

The Maximum A Posteriori (MAP) approach for estimating $\boldsymbol{X}$ seeks to estimate $\boldsymbol{X_{MAP}}$ for which the a-posteriori probability $\Pr\{X|Y\}$ is a maximum. Formally, $\boldsymbol{X_{MAP}}$ is calculated using the following equation.

$$\begin{aligned} X_{MAP} &= \operatorname{argmax}_X[\Pr\{X|Y\}] \\ &= \operatorname{argmax}_X[\log \Pr\{Y|X\} + \log \Pr\{X\}] \end{aligned} \qquad \ldots \text{Eq (3)}$$

This is achieved by applying Bayes' rule, since, $\boldsymbol{X_{MAP}}$ is independent of $Pr\{\boldsymbol{Y}\}$ and taking logarithms. $\log Pr\{\boldsymbol{Y}|\boldsymbol{X}\}$ is a log likelihood function and $Pr\{\boldsymbol{X}\}$ is the density of X. The likelihood function is determined by the PDF of the noise as $Pr\{\boldsymbol{Y}|\boldsymbol{X}\} = f_{\boldsymbol{N}}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X})$. Markov Random Field image is used to model the prior term $Pr\{\boldsymbol{X}\}$.

Maximum likelihood (ML) estimation has also been used for SR reconstruction [2]. It is a special case of MAP estimation (no prior term). Since prior term is crucial for solving the ill posed inverse problem, MAP estimation should be used instead of ML.

2.2.4 Set Theoretic Reconstruction Methods

Set theoretic methods are popular as they are simple and utilize powerful spatial domain observation model. Methods of projection onto convex sets (POCS) are especially popular as they allow convenient inclusion of information. In set theoretic methods, the space of ST solution images is intersected with a set of constraint sets representing desirable SR image characteristics. These characteristics include properties such as positivity, bounded energy, fidelity of data, smoothness etc., to yield a reduced solution space. POCS refers to an iterative procedure which, given any point in the space SR images, locates a point which satisfies all the convex constrains sets.

Convex sets representing constraints on solution space X are defined.

- Data inconsistency is represented by a set $\{x: |Y - Hz| < \delta_0\}$
- Positivity by $\{x : x_i > 0 \; \forall i\}$
- Bounded Energy by $\{x : ||x|| \leq E\}$
- Compact support $\{x : x_i = 0, i \in A\}$

A projection operator is also determined for each convex constraint. The projection $\boldsymbol{P}_{\propto}$ associated with constraint set $\boldsymbol{C}_{\propto}$ projects a point in the space of $\mathbf{z}$ onto the closest point on the surface of $\boldsymbol{C}_{\propto}$. It is possible to converge to a solution on the surface of the intersection of the K convex constraints sets by iterative application, $\boldsymbol{x^{(n+1)}} = \boldsymbol{P_1 P_2 P_3} \ldots\ldots \boldsymbol{P_K z^n}$. This point is however not guaranteed to be unique since it is dependent on the initial guess.

An alternate SR reconstruction method uses ellipsoid to bind the constraint set [5]. The centroid of this ellipse is taken as the SR estimate. An iterative method is used to find a solution since direct computation is infeasible.

The main disadvantage of using Set Theoretic SR reconstruction is the non-uniqueness of the solution. The solution is highly dependent on the initial guess. Also the convergence rate for this method is very slow and has a high computational cost associated with it. The bounded ellipsoid method ensures a unique solution. However, it cannot be assured to be an optimal solution.

2.2.5 Optimal and Adaptive Filtering

A number of approaches towards SR reconstruction have been proposed using inverse filtering. These techniques have limited ability to include a-priori constraints compared to the powerful framework provided by Bayesian methods or POCS [4]. Some methods have also been used in applications. These methods are in effect LMMSE estimators and do not include non-linear a-priori constraints [2].

2.3 Comparison Between Frequency Domain and Spatial Domain SR Reconstructions

A general comparison of frequency and spatial domain SR reconstructions methods is presented in [2]. We have presented that in Table 1.

Table 1 Frequency Domain vs. Spatial Domain SR

	Frequency Domain	Spatial Domain
Observation model	Frequency domain	Spatial domain
Motion models	Global translation	Almost unlimited
Degradation model	Limited, LSI	LSI or LSV
Noise model	Limited, LSI	Flexible
SR Mechanism	De-aliasing	De-aliasing A-priori info
Computation requirements	Low	High
A-priori info	Limited	No limit
Regularization	Limited	Excellent
Extensibility	Poor	Excellent
Applicability	Limited	Wide
Application performance	Good	Good

From the Table 1 above, it is evident that spatial domain based SR reconstruction methods are better than frequency domain methods. Even though spatial domain based methods are complex and computationally intensive, they provide a degree of flexibility not provided by frequency domain methods.

CHAPTER 3: METHODOLOGY

This Chapter will describe the methods we used in order to solve the disparity image computation and the SR image estimation using the a priori knowledge about the design of our scanner. First, in Section 3.1 we briefly describe the relevant details of the scanning device. Then in Section 3.2 we describe the details of detecting the laser stripes in the captured video frames and thus computing the deformations in it due to the surface shape being scanned. In Section 3.3 we describe our approach to improving the resolution of the color texture image along the scan direction.

3.1 Imaging Device

The imaging device consists of a linear actuator and a high definition (HD) camera, and two line lasers. The camera is mounted on a under carriage which is driven along the length of the actuator by a stepper motor. The imaging geometry of the digitizing device and an image of the prototype are shown in Figures 3 and 4. The speed of the stepper motor can be precisely controlled and remains constant during the entire scan. A 1:10 gearbox translates the stepper motors power to the linear actuator. This ensures the smooth motion of the carriage with extremely minimal variations in its speed.

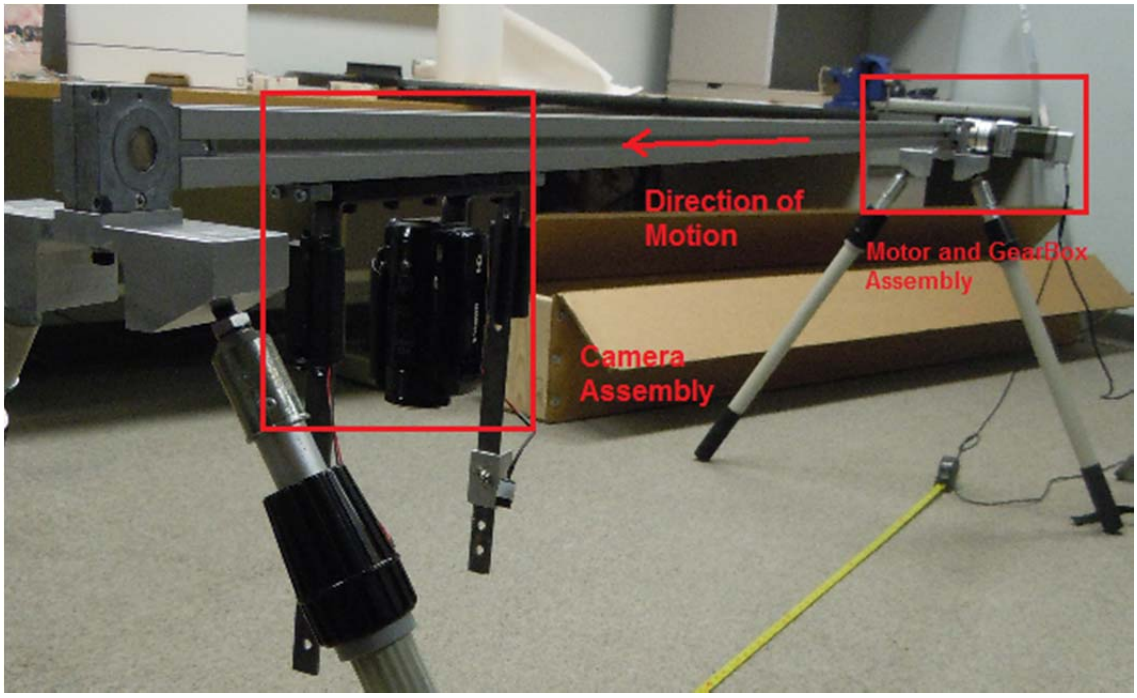

Figure 3 Linear Actuator used to collect data

The camera records HD videos at 1080×1920 pixels resolution. It records videos at 30fps (frames per second). Objects are placed below the assembly and a video is recorded of the scene as it moves over the object being scanned. The objects remain stationary while the camera assembly passes over it. The combination of high resolution, high frame rate, and slow linear motion causes sub-pixel overlap along the direction of motion. Based on different speeds and distance of objects from the camera we get different percentage of sub-pixel overlap. Successive frames from videos are stored and used as input for SR reconstruction. An important point to note here is that since the translation of camera is only along the one axis (Y-axis which is along the length of the actuator), we get sub-pixel overlap only along one axis. Hence, we can get improved resolution only along that axis on this device.

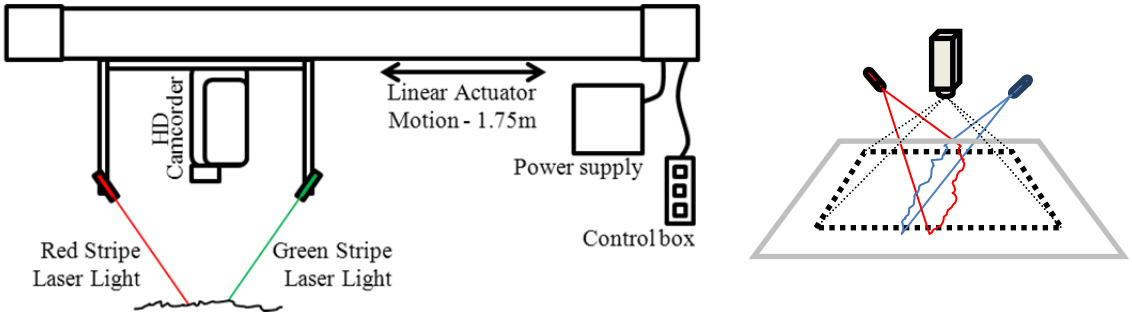

Figure 4 Imaging System

3.2 Laser Detection

As we mentioned in Chapter 1, laser detection forms an essential part of generating 3D impression. It is essential that we accurately detect the position of the laser strip as the accuracy of the 3D impression depends on it. The laser stripe projected typically results in a 4-5 pixel wide stripe and it needs to be localized more precisely in order to have accurate depth values. We investigated two possible approaches to laser detection:

- Detecting the location of pixels along the laser stripe by estimating the peak points along the laser.
- Detecting and localizing the edges of the laser stripe. With image processing methods, typically this can give a much more precise location of the laser stripe and does not have the problems associated with the color saturation in peak detection.

3.2.1 Peak Based Detection

One of the most important properties of any laser is that they are monochromatic, i.e. the entire beam consists of waves composed of only one frequency in the electromagnetic

spectrum. We use this important property to detect the laser in this method. In this method we use the hue of the laser light to detect the laser light.

We have identified hue values associated with the red and green lasers. These values are the used to design a hue based filter that filters out all the information other than the laser strips. In order to detect the laser stripe in the video frame image, we follow the following steps:

1 Set up a region of interest (ROI) for laser stripe:

The region of interest is determined by the geometry of the camera and laser stripe light and is set by the maximum amount of displacement that can occur due to surface height. It is an approach to reduce the search region and thus reduce the computational time for searching the laser light in the image. The yellow lines in Figure 5(c) depict the ROI in this example frame.

(a) (b) (c)

Figure 5 Color based laser stripe detection steps (a) original video frame; (b) region of interest of the original image and laser pixels detected based on hue; (c) Results are super imposed and region of interest is highlighted in yellow.

2 Convert the color representation:

Convert the color representation of the pixels in this region from Red-Green-Blue (RGB) to Hue-Saturation-Value (HSV). Considering only the Hue channel, identify those pixels that where the laser light is reflected on the surface. This identification is done if the pixel's hue value falls within a range $[hue_{min}, hue_{max}]$. This range of hues is defined by the color of the particular laser light (green or red). A dilate morphological image operation is performed in the vertical direction. The pixels resulting from this operation constitute a set of candidate points for the laser pixels. This set of pixels typically forms a thick (5-10 pixels wide) band and the location of the laser stripe needs to be refined for better accuracy. Figure 5(b) shows the candidate points so detected in the example video frame of Figure 5(a).

3 Peak Detection:

The mask generated in step 2 is then applied to the value channel of the corresponding regions of interest for the red and green laser respectively. A [15x1] Gaussian smoothing operator is applied to the value channel to ensure that the peak lies in the center of the laser stripe. This masked value channel is then used to find the peak in intensity which corresponds to the center of the laser strip. We scan each vertical line to find the highest peak. This peak usually corresponds to the center of the laser strip.

Once the position of the laser is detected we can add that information to the disparity image (amount of deformation in the laser light) which is then used to generate the 3D Map.

3.2.2 Edge Based Laser Detection

Peak based laser detection has a major disadvantage. It is highly susceptible to ambient light conditions. The range for hue values change based on the ambient light. Another issue with peak based detection is saturation of the laser light in the camera sensor. When ambient light is low the center of the laser beam tends to be saturated. This makes it difficult to detect the locations of the peaks accurately. Thus, it was decided to use a more robust and reliable laser detection method by making use of edge detection methods which are less susceptible to illumination variations. Moreover, good edge detectors can localize the edges with high accuracy.

Since the laser light is inherently coherent it does not spread much over a short distance. Thus the edges of the laser are clearly visible in most of the video and can be located much more precisely. We make use of this property to accurately detect the position of the laser. The geometric configuration of the edges of the laser light with respect to the camera is the information used in the calibration of the device and subsequent estimation of the surface depth values. The following are the steps to detect the laser stripe.

1 Set up a region of interest (ROI) in the image similar to peak-based laser detection method. A region of interest is defined for each laser.

2 Use an edge detection method to locate the edges of the laser stripe. We used the Sobel edge detection method:

   For detecting the edges using the Sobel operator we apply the following convolution kernels on each column.

$$G_1 = (1 \quad 2 \quad 0 \quad -2 \quad -1) \text{ and } G_2 = (1 \quad 2 \quad 1)$$

We apply the kernels G_1 and G_2 to the image I to get the edge response E. This edge response will then be used to accurately detect the position of the laser in the image plane.

$$E = G_2 * (G_1 * I)$$

E will contain peaks and valleys corresponding to the leading and trailing edges. We apply a threshold to the magnitude of the edge response in order to filter out the weak edges and the edges due to noise.

3 Peak detection:

We apply the peak detection algorithm on the input image to get a rough position of the location of the laser stripe.

4 Laser position detection:

Laser position can now be detected by using the output from step 2 and step 3. We start from the peak and move in two directions, above and below, till we encounter an edge. These positions are marked as the leading and trailing edges of the laser stripe. We can use either the leading or trailing edge as a reliable position of the laser.

3.3 Super Resolution

As we saw in Chapter 2, most of the Super Resolution reconstruction algorithms assume general motion and try to estimate the motion as well as the higher resolution image from the set of low resolution images. However, in our case we have a priori knowledge about the precise motion of the imaging device. We can utilize this knowledge to simplify the process of estimating the higher resolution image. In this section we will discuss about

the imaging device, data pre-processing and actual SR reconstruction by minimizing the error function. In this section we have used two methods for energy minimization.

3.3.1 Data Pre-Processing

The next step after extracting successive frames is to compensate for the distortions introduced due to roll and tilt of the camera. Before any SR reconstruction can take place it is necessary to compensate for the roll and tilt of the camera and bring the frame into an ideal image plane. Figure 6 shows a simplified roll correction. Similarly, the imaging system is used to correct for tilt.

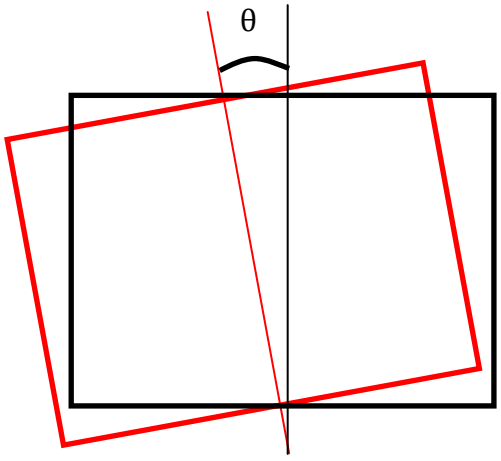

Figure 6 Correction for Image Roll (α) Red: Top view of imaging system Black: Top view of Roll corrected imaging system

Vanishing point based calibration method is used to calculate α and θ. A calibration object is placed along the direction of motion and is captured by the camera. Two frames, at different times, showing the calibration object are extracted. The corners of the calibration object are marked in each frame which can then be used to find the vanishing point.

Camera roll (θ) around its optical axis is given by,

$$\theta = \tan^{-1}\frac{\bar{x}}{\bar{y}} \quad \text{...Eq (4)}$$

where $(\bar{x}, \bar{y})$ are the estimated coordinates of the vanishing point. Camera tilt (α) between optical axis and the line perpendicular to linear axis on optical axis + linear axis plane,

$$\alpha = \frac{\bar{y}}{|\bar{y}|}\tan^{-1}\left(\frac{f}{\sqrt{\bar{x}^2+\bar{y}^2}}\right) \quad \text{…Eq (5)}$$

Using Eq (5) and Eq (6) we can get the values of α and θ. We can now get image into an ideal image plane by applying affine transform.

Using affine transform, all input images are corrected for distortions and brought into the ideal coordinate system. From this point forward all input images are assumed to have been passed through such a correction for roll and tilt.

Since the images have been compensated for roll and tilt it is possible to treat each column as an independent vector aligned with the direction of the linear motion. The problem of SR reconstruction of the whole image can now be treated as a set of 1D SR reconstructions of the image columns along the motion.

3.3.2 Least Squares Formulation

Given K images $\{\boldsymbol{X_L^{(n)}}\}_{\boldsymbol{n=1}}^{\boldsymbol{K}}$ of size M_1 *x 1* pixels find the high resolution image $\boldsymbol{X_H}$ of size N_1 *x 1* pixels (where $N_1 > M_1$) which minimizes the squared error function:

$$E(X_H) = \sum_{n=1}^{K}\left\|P_n(X_H) - X_L^{(n)}\right\|^2 + L\delta\|CX_H\|^2 \quad \text{…Eq (6)}$$

Where:

1. $\| . \|$ - Can be any norm, usually L_2.
2. $\boldsymbol{P_n(X_H)}$ is the projection of $\boldsymbol{X_H}$ onto the coordinate system and the sampling grid of image $\boldsymbol{X_L^{(n)}}$.
3. $\boldsymbol{CX_H}$ is called the edge preserving term. It is a penalty term to ensure we do not smooth out the edges. It penalizes the energy function if we try to smooth over the edges. $\boldsymbol{C}$ is a simple [1 -1] kernel designed to get an edge response.
4. $\boldsymbol{\delta}$ is a scalar constant which is used to assign weight to the penalty term. It is usually less than 0.1.
5. $\boldsymbol{L}$ is called a line process. It is a boolean array which turns off the penalty term for certain areas. $l_i = 1$ indicates that there is a discontinuity in the sub-interval $x = [i-1, i]$. The entries of $\boldsymbol{L}$ are set to $(1 - l_i)$ to indicate if there is a discontinuity at that location and if the penalty term needs to be turned off so the estimated HR image is not smoothed over the discontinuity. $\boldsymbol{L}$ ensures that the penalty term is applied only in areas having edges which is determined by an edge detection operator such as the Sobel edge detector.

The projection $\boldsymbol{P_n(X_H)}$ is modeled by four stages:

1. Geometric Transformation
2. Blurring
3. Subsampling
4. Additive noise

Most modern algorithms differ on the optimization technique used for solving the equation, the constraints on $\boldsymbol{X_H}$ added to the system and the modeling of the geometric transformation, blur and noise.

Image Registration Model

1. Geometric Transformation

   Coordinate system of $\boldsymbol{X_H}$ has to be determined in order to have a unique $\boldsymbol{X_H}$. Usually the coordinate system of one of the input image is multiplied by the factor q. The geometric transformation of $\boldsymbol{X_H}$ to the coordinate system of the input images is computed by finding the motion between the input images. High accuracy of motion estimation is crucial for the success of super resolution.

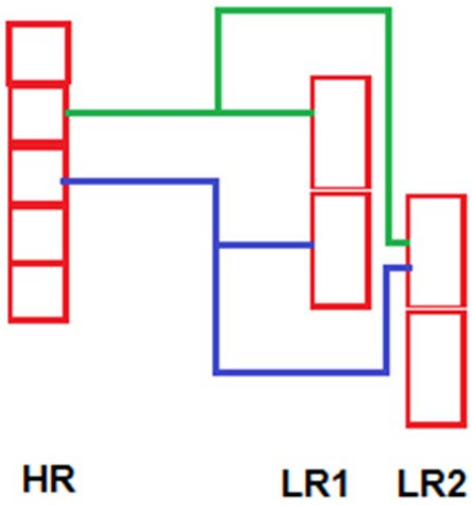

Figure 7 HR Geometric Transformation

The motion between input images is governed by the linear translation of the actuator. Since the camera moves only along the Y-axis it is possible to accurately model the

geometric transformation. Depending on the amount of sub-pixel overlap depending upon the speed of the actuator and height of the object from camera.

2. Blur

   Image blur is usually modeled using convolution. A low pass kernel is used for convolution. We have to take into account blur caused by optics as well as sensors while modeling the blur. In order for the super resolution problem to be uniquely solvable we need to have an accurate kernel.

3. Subsampling

   The subsampling matrix maps the high resolution image $\boldsymbol{X_H}$ to the input low resolution images. Consider a high resolution image $\boldsymbol{X_H}$ of size $N_1 \times 1$. $\boldsymbol{X}$ in this case is the ideal un-degraded image of a continuous scene. It is sampled at or above the Nyquist rate.

4. Additive noise

   It is assumed that the noise is additive and normally distributed with zero mean. It thus becomes possible to find the maximum likelihood solution by minimizing the error function. The assumption of normal distribution of the noise is not accurate in most of the cases, as most of the noise in the imaging process is non-Gaussian (quantization, camera noise, etc.), but modeling it in a more realistic way would end in a very large and complex optimization problem which is usually hard to solve [5].

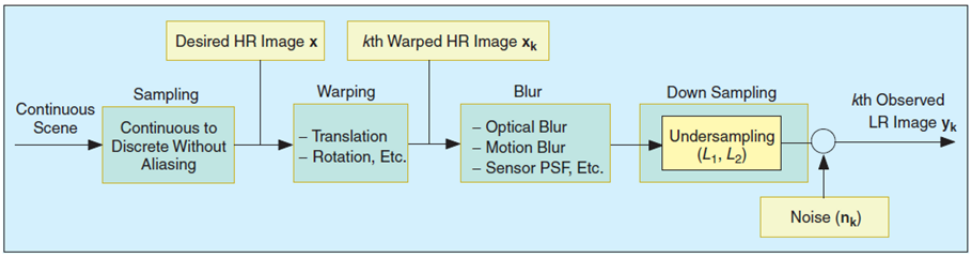

Figure 8 Image Registration Model

3.3.3 Error Minimization

It is critical that the initial estimate be as close to the actual solution as possible to ensure that the Energy function finds the correct solution. One of the important steps to create the initial estimate is to calculate the percentage of sub-pixel overlap between two successive frames.

3.3.3.1 Sub-pixel Overlap Estimation

In order to accurately model the geometric transformation it is necessary to accurately compute the sub-pixel overlap. A profile computed from the video can be used to achieve this. A profile of a video is an image composed by cutting the video along the time axis. This is done by extracting one column at the center of each frame and stacking them along the time axis. This results in a collection of one-dimensional (1D) image data accumulated along the time axis, which can be viewed as a two-dimensional profile image (see Figure 9). This profile can be analyzed to estimate the amount of pixel overlap in the lower resolution images due to the camera motion.

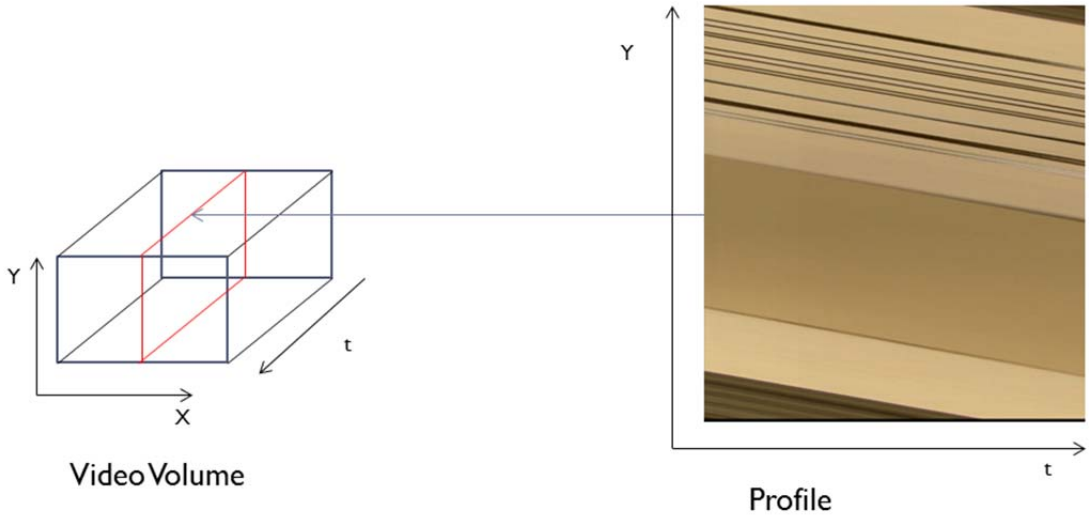

Figure 9 Generating Profile of Video

The lines in the profile represent the edges of the objects in the scene during the camera motion and indicate the amount of sub-pixel overlap. The magnitude and direction of the orthogonal vectors can be used to calculate the sub-pixel overlap.

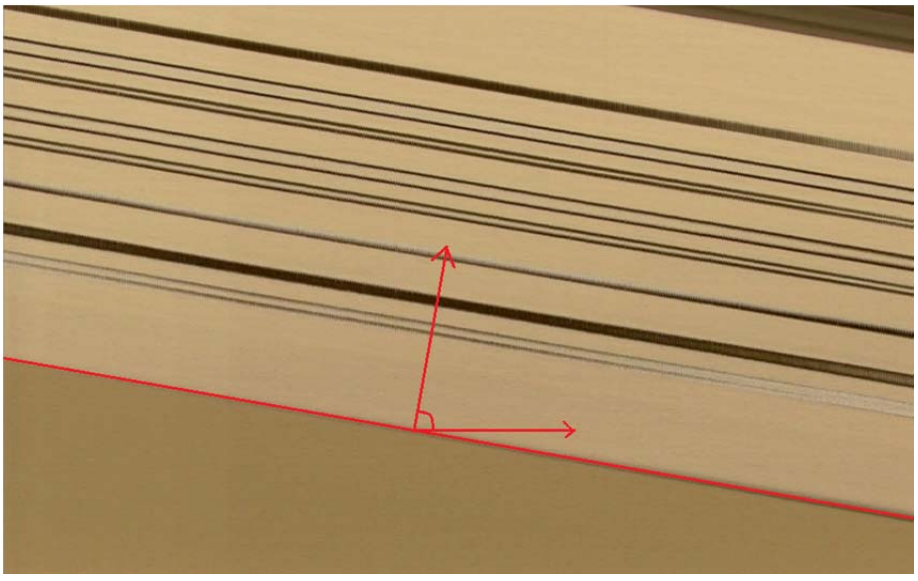

Figure 10 Profile of a video with distance from camera = 20.5’ and speed = 1. θ=84.1

The profile of a video is passed through filters that give the partial derivatives along the X and Y axis. If $\boldsymbol{I}$ is the profile of a video, $\boldsymbol{I_x}$ and $\boldsymbol{I_y}$, represent the partial derivatives of **I** along X and Y axes, respectively.

$$\mathrm{I_x} = \frac{\partial I}{\partial x} \text{ and } \mathrm{I_y} = \frac{\partial I}{\partial y}$$

Magnitude E_{mag} and direction θ can be calculated using the following equations.

$$\mathrm{E_{mag}} = \sqrt{\mathrm{I_x^2} + \mathrm{I_y^2}} \text{ and } \theta = \tan^{-1}\left(\frac{\mathrm{I}_y}{\mathrm{I}_x}\right)$$

$\theta = \pi/4$ represents a sampling rate equal to the speed of translation and implies no sub-pixel overlap. Thus, $\theta \leq \pi/4$ implies there is no sub-pixel overlap and thus no SR reconstruction is possible. On the other hand, $\theta = \pi/2$ implies that the pixels overlap and hence there is no sub-pixel overlap between successive frames. Thus for SR reconstruction, we need $\frac{\pi}{2} > \theta > \frac{\pi}{4}$. Based on various heights and linear speeds, the Table 2 shows the percentage sub-pixel overlap between successive frames.

Table 2 Sub-Pixel overlap based on Speed and Distance from Camera

Distance from camera (inches)	Speed (Pulses/sec)	Threshold (Top 5 %)	Theta θ (degrees)	Percentage Overlap
20.5	1	122	84.1657	87.03
	2	110	63.84	41.87
	3	51	52.2	16.00
	4	39	42	-6.67
18	1	74	64.4	43.11
	2	51	55.5	23.33
	3	40	49.4	9.78
	4	43	39.2	-12.89
15.5	1	49	51	13.33
	2	26	44	-2.22
	3	30	40	-11.11
	4	23	37	-17.78

Table 2 shows the percentage sub-pixel overlap for different speed and height settings. It is evident that we have no or negligible sub-pixel overlap for heights greater than 18''. Also the speed of the camera motion is a major contributing factor for sub-pixel overlap between successive frames.

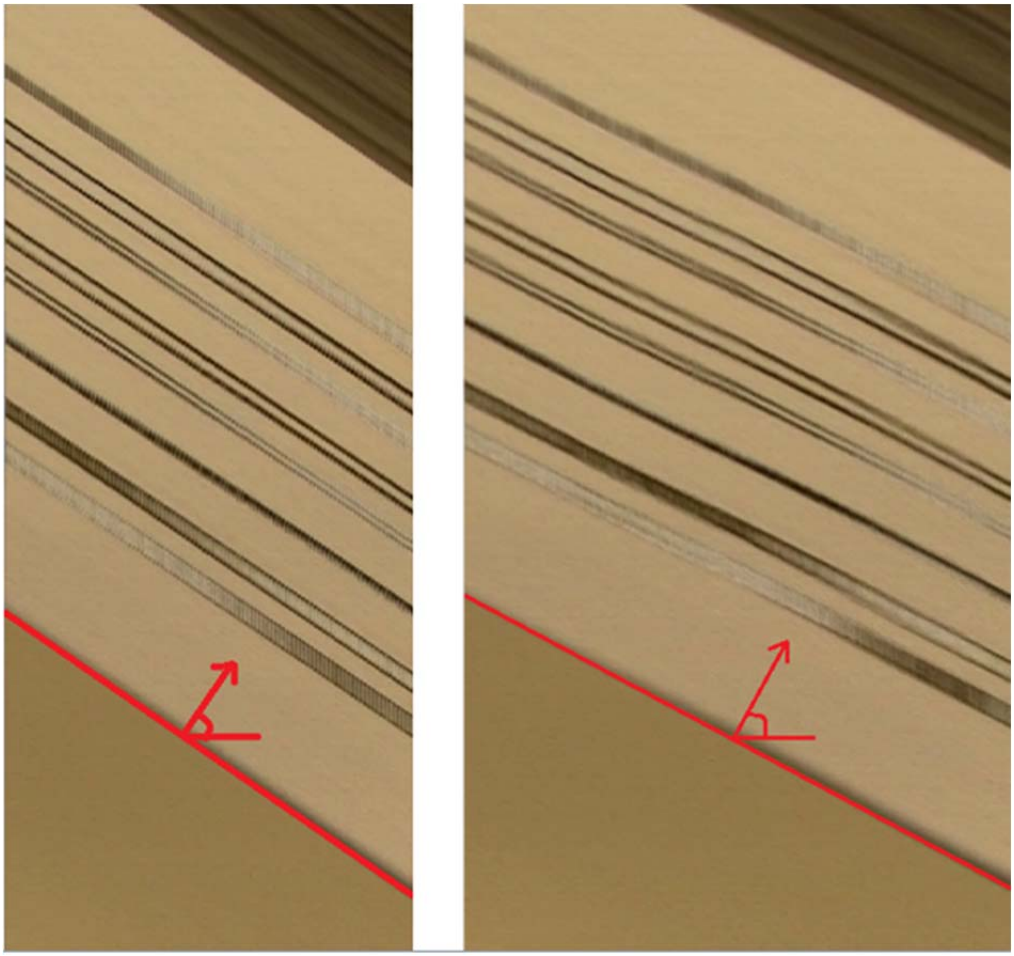

Figure 11 Profile of a video with distance from camera = 20.5' and speed = 3(Right), speed = 4 (Left). θ(Left)=42 θ(Right) = 63

For Energy minimization using Newton's method to converge quickly it is necessary that the initial estimate be close to the solution. Based on the percentage of sub-pixel overlap the higher resolution image is estimated. An overlap of approximately 75% implies that

three low resolution pixels contribute to one higher resolution pixel and thus we can have 3 times improvement in resolution. On the other hand an overlap of approximately 50% implies two low resolution pixels contribute to one higher resolution pixel and thus we can have 2 times improvement in resolution. Initial estimate of each high resolution pixel consists of average of contributing low resolution pixels. Random noise, having varying standard deviation, is added to the initial estimate.

After the initial estimate of the HR image is generated, we refine the estimated solution by minimizing the error using a standard optimization method. The solution for SR reconstruction equation Eq (6) is to find the $\boldsymbol{X_H}$ which essentially minimizes the scalar Energy function E. Two methods, gradient descent and Newton's Minimization were utilized in this thesis.

3.3.3.2 Gradient Descent

Gradient Descent is a first-order optimization algorithm [6]. In order to find the minimum, we take steps proportional to the negative of the gradient of the function taken at the current point. We start with the initial estimate and compute the gradient vector $\boldsymbol{\nabla J}(\boldsymbol{X}^{(1)})$. The next estimate $\boldsymbol{X}^{(2)}$ is obtained by moving some distance from $\boldsymbol{X}^{(1)}$ in the direction of steepest descent i.e. along the negative of the gradient. In general, $\boldsymbol{X}^{(n+1)}$ is obtained from $\boldsymbol{X}^{(n)}$ by the equation Eq (7) [7]

$$\boldsymbol{X}^{(n+1)} = \boldsymbol{X}^{(n)} - \boldsymbol{\gamma} * \boldsymbol{\nabla J}(\boldsymbol{X}^{(n)}) \quad \text{…. Eq (7)}$$

Where $\gamma > 0$, known as step, such that $E\left(X^{(n+1)}\right) > E(X^{(n)})$.

The major disadvantage of using gradient descent is its relatively slow convergence speed. Its rate of convergence is inferior to most other methods. Also a lot depends on the selection of γ, if it is too small, we unnecessarily spend time to get to the minimum. However, if it is too large then we run into the risk of possibly missing the minimum.

From Figure 7 and Figure 8 we can see that the reduction in E value is close to zero after 40 iterations for γ = 0.05 and close to 30 iterations for γ = 0.1.

3.3.3.3 Newton's Method

Newton's method is an iterative method for finding the roots of the equations. It is used to find stationary points of differentiable functions, which are the zeros of the derivative function [7] [6]. Assuming the criterion function can be well-approximated by the second-order expansion around a value $\boldsymbol{X}^{(k)}$.

$$\boldsymbol{J}(\boldsymbol{X}) \cong \nabla \boldsymbol{J}^{\mathrm{t}}\left(\boldsymbol{X}-\boldsymbol{X}^{(k)}\right)+\frac{1}{2}\left(\boldsymbol{X}-\boldsymbol{X}^{(k)}\right)^{\mathrm{t}}\boldsymbol{H}(\boldsymbol{X}-\boldsymbol{X}^{(\mathrm{k})}) \qquad \ldots\text{Eq (8)}$$

Where $\boldsymbol{H}$ is the Hessian matrix of second partial derivatives $\frac{\partial^2}{\partial x_i \partial x_j}$ evaluated at $\boldsymbol{X}^{(k)}$. We choose $\boldsymbol{X}^{(k+1)}$ to minimize the second-order expansion.

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{X}^{(k)} - \boldsymbol{H}^{-1}\nabla \boldsymbol{J} \qquad \ldots\text{Eq (9)}$$

Newton's method is not applicable if the Hessian matrix $\boldsymbol{H}$ is singular.

From Eq (6),

$$E(X_H) = \sum_{n=1}^{K} \left\| P_n(X_H) - X_L^{(n)} \right\|^2 + L\delta \|C X_H\|^2$$

$\boldsymbol{X_H}$ is the high resolution column, of size Nx1, to be constructed. $\boldsymbol{X_L}$ represents the low resolution frame of size Mx1. $\boldsymbol{P_n}$ represents the blurring and subsampling. Blurring is

achieved by using a Gaussian kernel H_G of size NxN, while subsampling is achieved by $\boldsymbol{H_f}$. The resultant matrix will be of size MxN.

$$\boldsymbol{P_n = H_f H_G}$$

Taking the first derivative of E w.r.t X represented by ∇J. Assume K = 2, then we can rewrite Eq (11) as,

$$E(X_H) = \left(\left(P_n(X_H) - X_L^{(1)}\right)^2 + \left(P_n(X_H) - X_L^{(2)}\right)^2\right)^{1/2}$$

We can ignore the square root part of RHS since it will not affect the minimization as E is a scalar. Thus we can write, $\boldsymbol{E(X_H)}$ as,

$$E(X_H) = \left(P_n(X_H) - X_L^{(1)}\right)^2 + \left(P_n(X_H) - X_L^{(2)}\right)^2 + L\delta\left((CX_H)\right)^2$$

$$E(X_H) = \left(P_n(X_H) - X_L^{(1)}\right)^t\left(P_n(X_H) - X_L^{(1)}\right) + \left(P_n(X_H) - X_L^{(2)}\right)^t\left(P_n(X_H) - X_L^{(2)}\right) + L\delta(CX_H)^t(CX_H)$$

Differentiating w.r.t $\boldsymbol{X_H}$ we get $\boldsymbol{\nabla J}$ as,

$$\nabla \mathrm{J} = 2X_H^t\left({P_n}^t P_n + \left({P_n}^t P_n\right)^t\right) - \left(P_n X_L^{(1)}\right)^t - {X_L^{(1)}}^t P_n - \left(P_n X_L^{(2)}\right)^t - {X_L^{(2)}}^t P_n + 4 * L\delta X_H^t \mathrm{C^t C}$$

$$\nabla \mathrm{J} = 4X_H^t {P_n}^t P_n - 2{X_L^{(1)}}^t P_n - 2{X_L^{(2)}}^t P_n + 4 * L\delta X_H^t \mathrm{C^t C}$$

Thus we can generalize for K input images,

$$\nabla \mathrm{J} = 2\mathrm{K}X_H^t {P_n}^t P_n - 2\sum_{i=1}^{K} {X_L^{(i)}}^t P_n \qquad \ldots \text{Eq (10)}$$

Differentiating $\boldsymbol{\nabla J}$ we get the Hessian Matrix $\boldsymbol{H}$

$$\mathrm{H} = 2\mathrm{K}X_H^t {P_n}^t P_n + 4 * L\delta \mathrm{C^t C} \qquad \ldots \text{Eq (11)}$$

From our experiments it was found that the error function minimizes in less than 60 iterations making Newton's Method converge a lot faster than Gradient Descent method. However in case of Newton's Method we have to re-compute $\boldsymbol{H}$ since L changes for each column. It took Gradient Descent 22,271 seconds to reconstruct a HR image of size 2160×650, while Newton's Method took 1,17,000 seconds. Hence Gradient Descent method was chosen for SR reconstruction.

CHAPTER 4: EXPERIMENTAL RESULTS

A number of experiments were performed to validate the SR reconstruction and to study how various factors affected the performance and resolution of SR reconstruction. Many videos were shot for experimentation. In order to study the improvement is resolution we shot 12 videos with different height and speed settings. Below is a summary of all the different settings used:

- Speed settings: 1, 2, 3 and 4 pps (1 pps = 1.32mm/sec)
- We obtained data for 3 different distances from the camera: 20.5'', 18'', and 15.5''.
- The camera was completely zoomed out with fixed exposure and focus. This was done to ensure that the auto focus and auto exposure do not introduce any noise.
- To each data set we added random Gaussian noise to study the performance of SR reconstruction. In each case the mean of the noise was 0 with varying amount of standard deviation. We used noise with standard deviation as 0, 1 and 4.

The following figures show how the Error function minimizes over a number of iterations using Gradient Descent method. They also show the corresponding initial estimate and the reconstructed image for different additive noise levels.

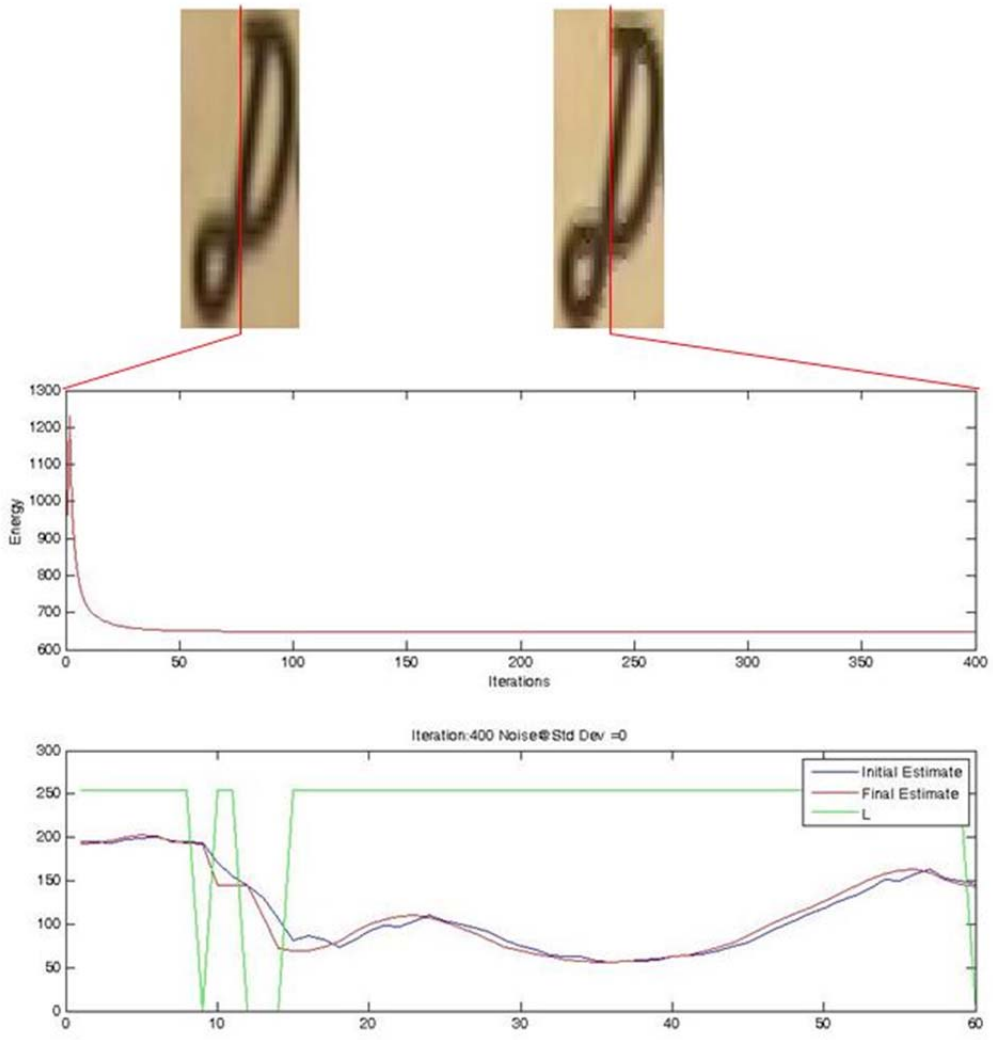

Figure 12 Error Function minimization for Additive Noise with Standard Deviation = 0 from the above figure we can see that the smoothing is not applied over edges. We illustrate in details the estimation of pixel values along one column as shown in the images. The top plot shows the change in energy over 400 iterations. The bottom plot shows 1D profile of the pixel values and discontinuity along the column.

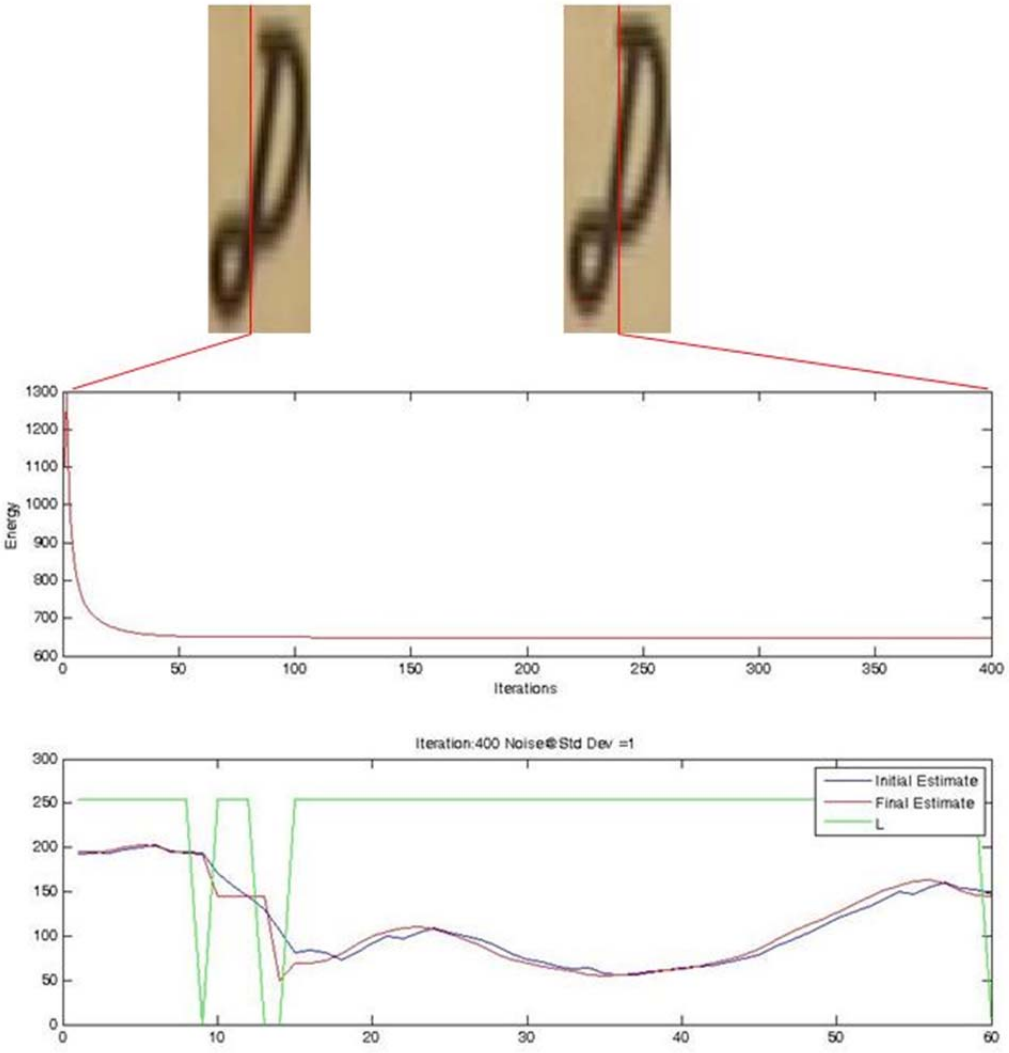

Figure 13 Error Function minimization for Additive Noise with Standard Deviation = 1

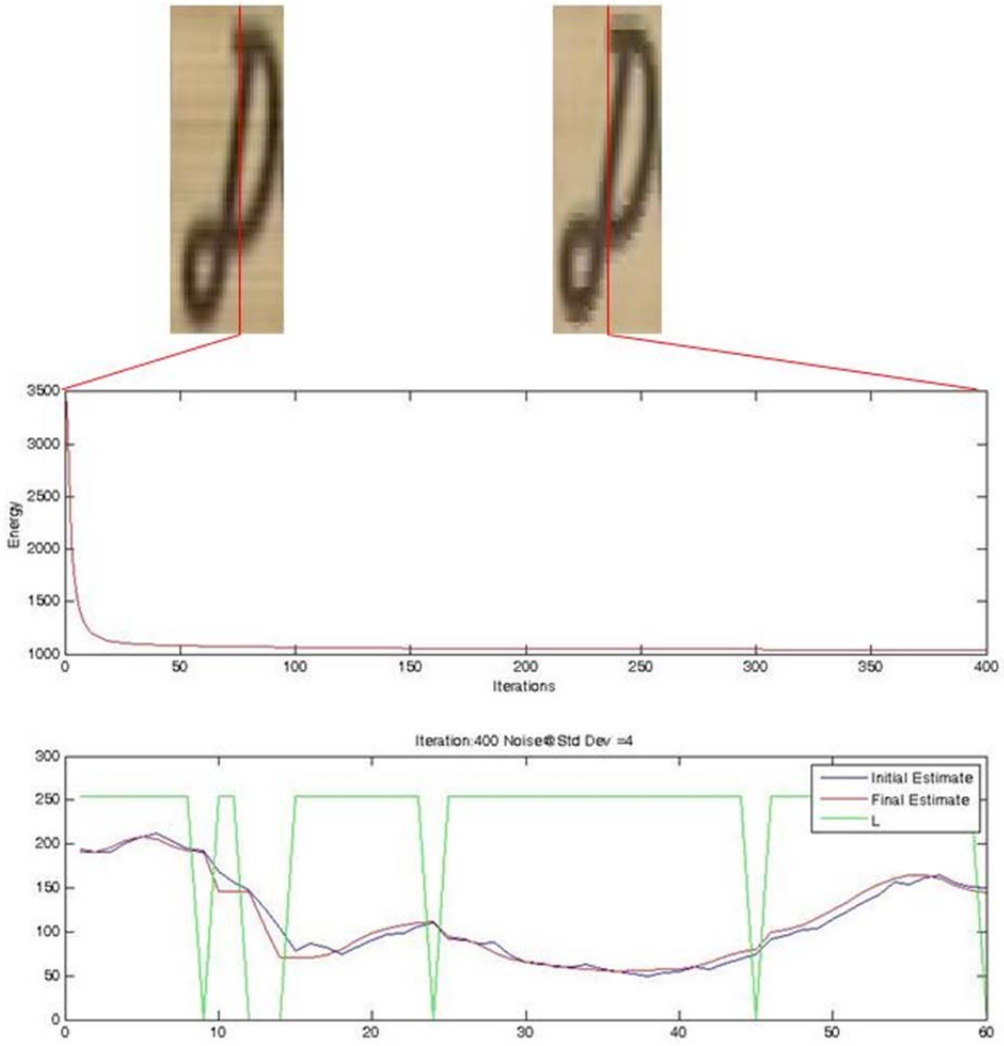

Figure 14 Error Function minimization for Additive Noise with Standard Deviation = 4

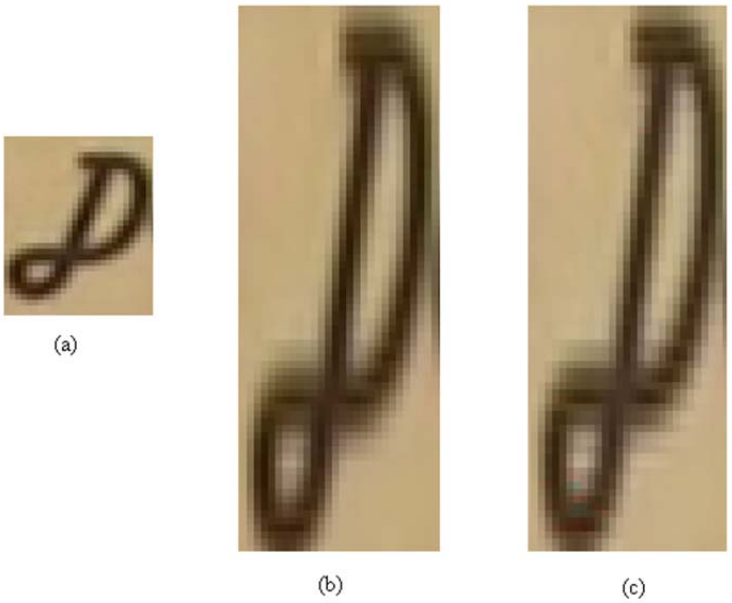

Figure 15 (a) Original LR image (b) Initial Estimate based on 3 LR frames and noise with Std Dev = 1 (c) Reconstructed SR image. As we go from (b) to (c) we can see that the edges are preserved

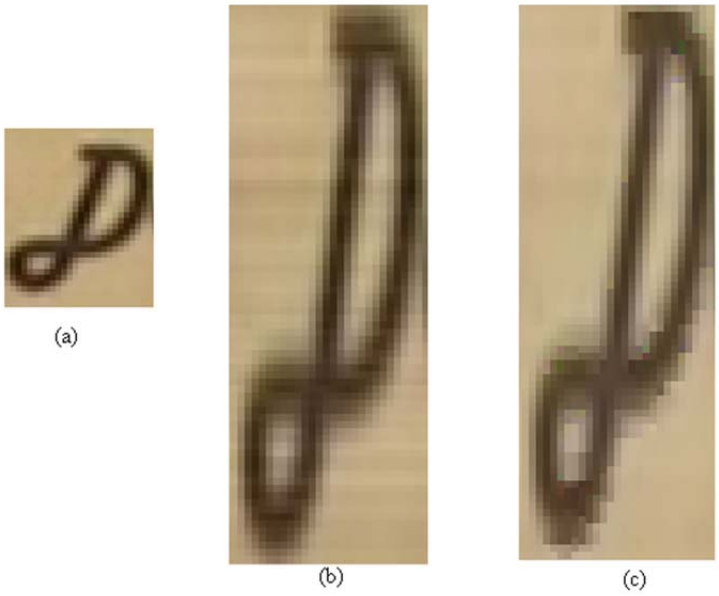

Figure 16 (a) Original LR image (b) Initial Estimate based on 3 LR frames and noise with Std Dev = 4 (c) Reconstructed SR image

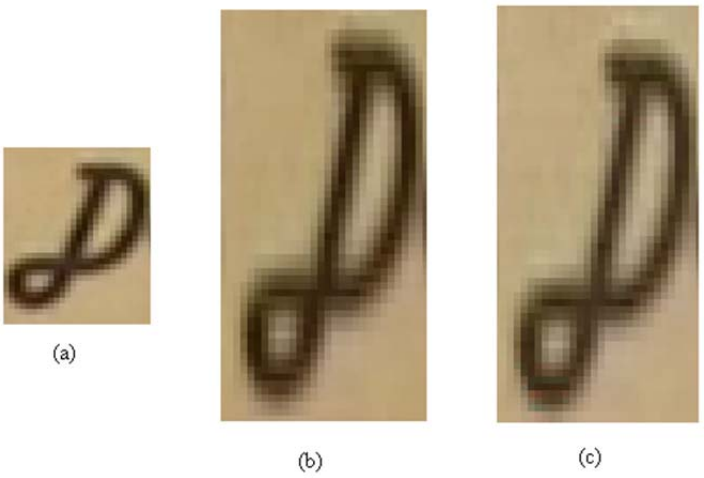

Figure 17 (a) Original LR image (b) Initial Estimate based on 2 LR frames and noise with Std Dev = 1 (c) Reconstructed SR image

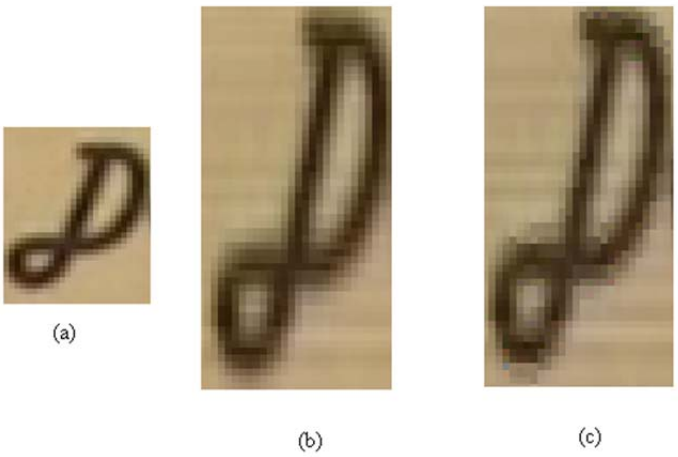

Figure 18 (a) Original LR image (b) Initial Estimate based on 2 LR frames and noise with Std Dev = 4 (c) Reconstructed SR image

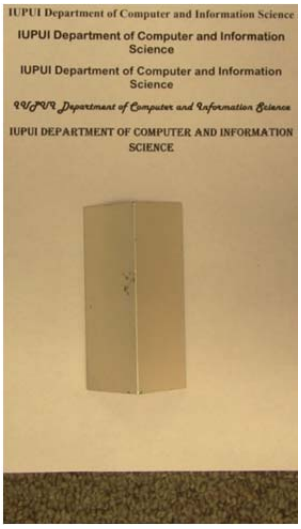

Figure 19 Low resolution (1080*650 pixels) input image

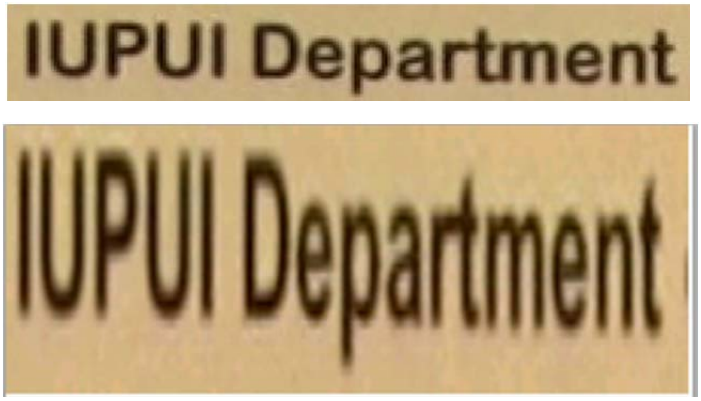

Figure 20 Zoomed in low resolution (Top) and High Resolution Image (Bottom)

Figure 22 shows the reconstructed high resolution image (3240*650 pixels) using 5 low resolution images (1080x650 pixels) corresponding to approximately 80% overlap (Height from camera = 20.5'', Translation Speed=1 pulse per second). It should be noted that super resolution is only in Y direction.

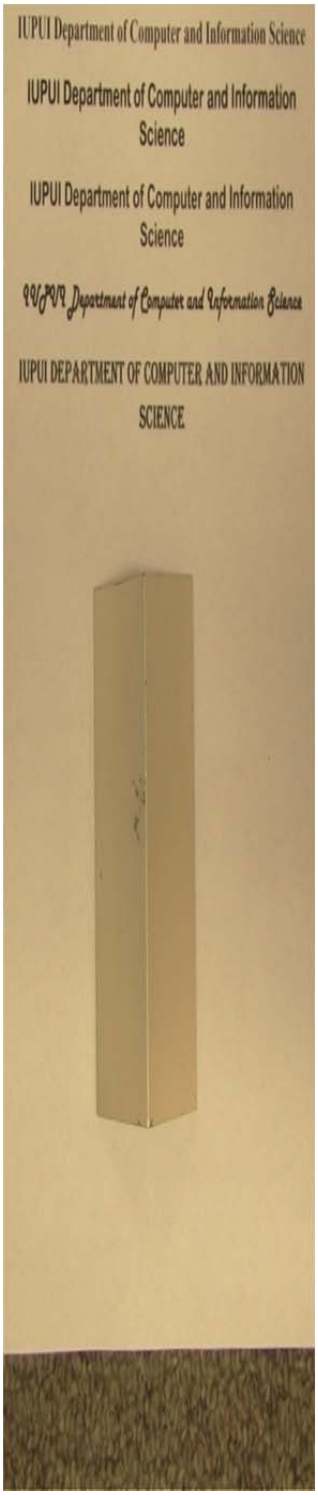

Figure 21 High Resolution Image 3240*650 pixels

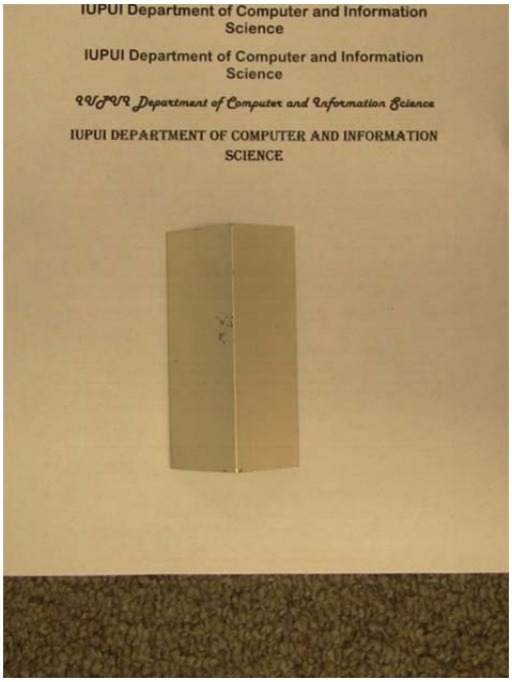

Figure 23 Low resolution (1080*650 pixels) input image

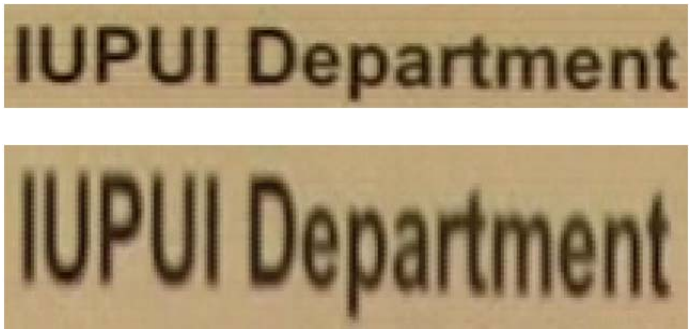

Figure 24 Zoomed in low resolution (Top) and High Resolution Image (Bottom)

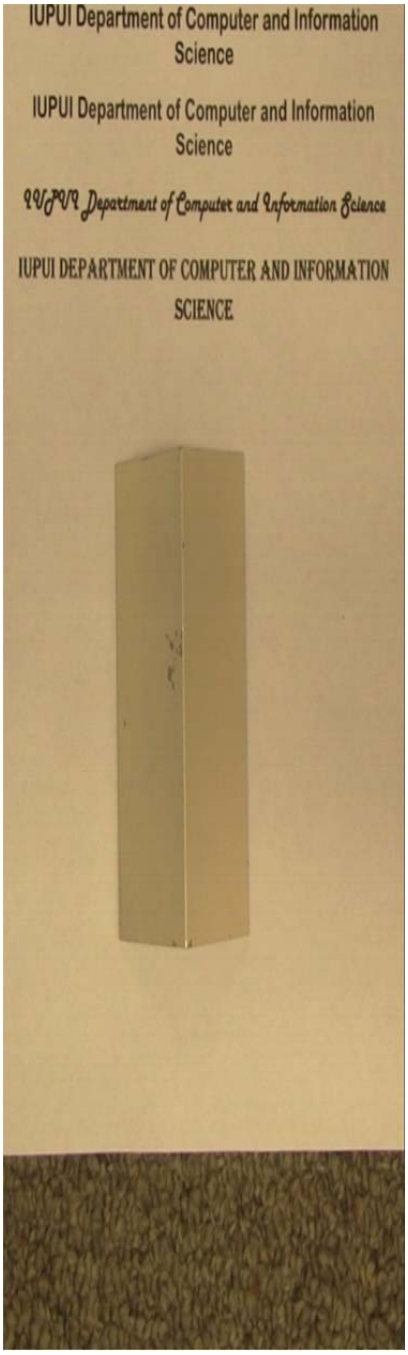

Figure 22 High Resolution Image 2160*650 pixels

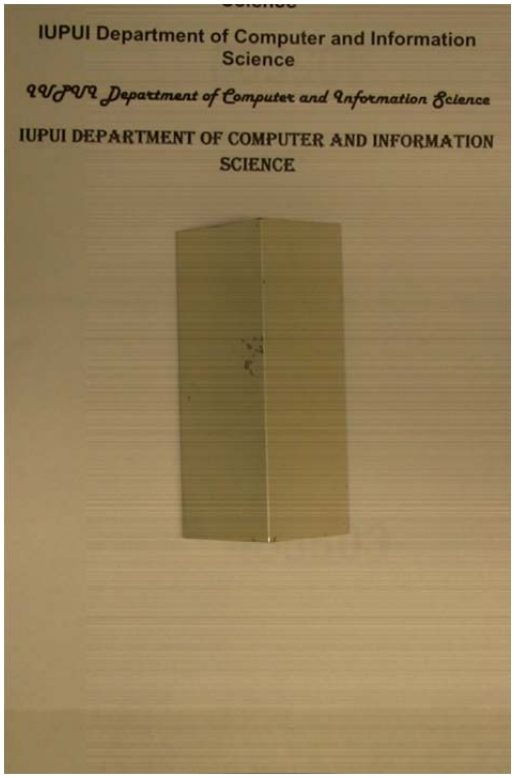

Figure 25 Low resolution (1080*800 pixels) input image

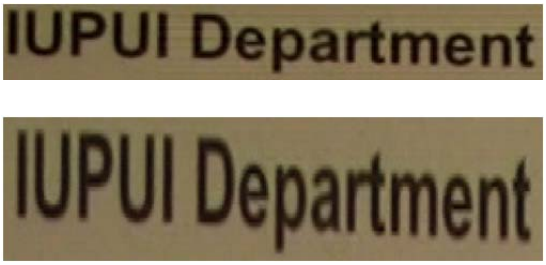

Figure 26 Zoomed in low resolution (Top) and High Resolution Image (Bottom)

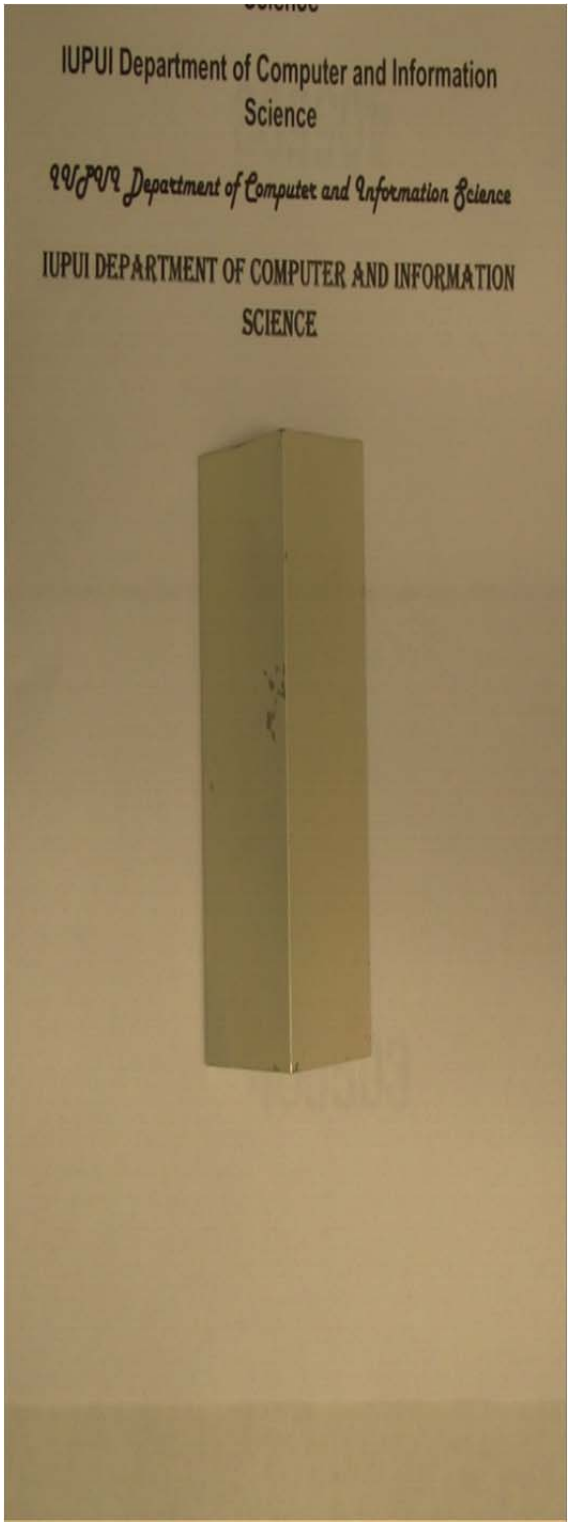

Figure 27 High Resolution Image 2160*800 pixels

CHAPTER 5: CONCLUSION

5.1 Summary

In the first part of this thesis we proposed two techniques for detecting the position of laser in a given frame. We detected the position using peak detection only and by using a combination of peak and edge detection. Accurately detecting the position of the laser forms a critical part in generating good quality 3D impressions. Using this accurate laser position we were able to generate high quality 3D images of the impression.

In this thesis we have also formulated and implemented a method to extract higher resolution images from a 3D impression digitizing device. The method uses the knowledge of the constrained motion model specific to the design of this device in order to estimate the amount of pixel overlap in the successive lower resolution image frames and thus estimate the higher resolution image. The method uses an edge preserving energy minimization approach to compute the final higher resolution image. With increasing requirement for higher resolution images and limitations in improving resolution using hardware techniques has led to the use of signal processing techniques for improving resolution. In this thesis work, we reconstructed high resolution images based on a number of lower resolution images with sub-pixel overlap.

5.2 Discussion

From the results shown in Chapter 4 it is evident that we were successfully able to achieve higher resolution images from a number of low resolution images. Based on the percentage of sub-pixel overlap we were able to successfully achieve a 2-3 times increase in resolution along the direction of motion. An important point to be noted is that because of the linear motion of the device we get sub-pixel overlap along Y-axis. Thus we have improved resolution only along Y-axis.

5.3 Future Work

For the last few years, there have been an increasing number of applications that require high resolution images. With current state of art of sensor technology, it is difficult to increase the resolution using hardware alone without making the product too expensive.

In this thesis work, the value of $\boldsymbol{L}$ in Eq (6) was fixed. $\boldsymbol{L}$ was fixed since we accurately knew the motion estimate and hence our initial estimate was very close to the solution. However, in case of general arbitrary motion we will have to re-compute $\boldsymbol{L}$ for each iteration.

Many new applications require high resolution images. One such application is generating 3-D Impressions of foot prints and tire prints found at a crime scene. There is ongoing research on improving 3D depth information using high resolution texture maps. SR reconstruction can play a vital role in this application.

LIST OF REFERENCES

LIST OF REFERENCES

[1] M. K. P. M. G. K. Sung Cheol Parl, "Image Reconstruction: A Technical Overview," *Signal Processing Magazine, IEEE,* pp. 21-36, 2003.

[2] R. L. S. Seam Borman, "Super-Resolution from Image Sequences – A Review," in *Circuits and Systems, 1998*, 1998.

[3] T. Saito, T. Komatsu and K. Aizawa, "An image processing algorithm for a super high definition imaging scheme with multiple different-aperture cameras," in *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference*, 1994.

[4] A. Patti, A. Tekalp and M. Sezan, "A new motion-compensated reduced-order model Kalman filter for space-varying restoration of progressive and interlaced video," *IEEE Transactions on Image Processing,* pp. 543-554, 1998.

[5] A. Z. a. S. Peleg, *Super-Resolution from Multiple Images having Arbitrary Mutula Motion,* Jerusalem, 2001.

[6] "Gradient Descent," www.Wikipedia.org, [Online]. Available: http://en.wikipedia.org/wiki/Gradient_descent.

[7] P. E. H. D. G. S. Richard O. Duda, "Gradient Descent Procedures," in *Pattern Classification*, Wiley- Interscience, 2001, pp. 224-227.

[8] A. Patti, M. Sezan and A. Murat Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transactions on Image Processing,* pp. 1064-1076, 1997.

[9] E. C. P. a. O. Y. C. W. T. Freeman, "Learing Low-Level Vision," *International Journal of Computer Vision,* vol. 40, no. 1, pp. 25-47, 2000.

[10] W. T. F. a. E. C. Pasztor, "Markov Networks for Superresolution," in *Annual Conference Information Sciences and Systems*, 2000.

[11] S. N. a. T. Kanade, "Limits on Supre-Resolution and How to Break Them," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2000.

[12] A. P. R. M. Y. Altunbasak, "Super-Resolution still and video reconstruction from mpeg-coded video," *IEEE Transactions Circuits System for Video Technology,* vol. 12, no. 4, pp. 217-226, 2002.

[13] Y. H. M. Elad, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Transactions on Image Processing,* vol. 10, no. 8, pp. 1187 - 1193, 2001.

[14] D. R. M. E. S. Farsiu, "Fast and robust super-resolution," in *International Conference on Image Processing*, Santa Cruz, CA, USA, 2003.

[15] M. K. O. a. M. I. S. A. M. Tekalp, "High resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," *ICASSP,* vol. 3, pp. 169-172, 1992.

[16] B. K. R. K. J. S. a. R. H. P. Cheeseman, "Super-resolved surface reconstruction from multiple images," *Maximum Entropy and Bayesian Methods,* pp. 293-308, 1996.

[17] M. E. a. A. Feuer, "Superresolution restoration of an image sequence: adaptive filtering approach," *IEEE Transactions on Image Processing,* vol. 8, no. 3, pp. 387-395, 1999.

[18] M. E. a. A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Transactions on Image Processing,* vol. 6, no. 12, pp. 1646-1658, 1997.

[19] M. I. S. a. A. T. P. E. Eren, "Robust, object-based high-resolution image reconstruction from low-resolution video," *IEEE Transactions on Image Processing,* vol. 6, no. 10, pp. 1446-1451, 1997.

[20] K. J. B. a. E. E. A. R. C. Hardie, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled imagesq," *IEEE Transactions on Image Processing,* vol. 6, no. 12, pp. 1621-1633, 1997.