

Hemispheric Vision with Resolution Enhancement

by

Dean M.H. Anderson

B.A.Sc., The University of British Columbia, 1996

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

THE FACULTY OF GRADUATE STUDIES

(Department of Electrical and Computer Engineering)

We accept this thesis as conforming
to the ~~required~~ standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 1998

©Dean M.H. Anderson, 1998

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical and Computer Engineering
The University of British Columbia
Vancouver, Canada

Date Dec 22/1998

ABSTRACT

Representing the whole world around a given point is an important goal in computer vision and many other applications. Most camera lens systems have been designed and optimized to provide excellent perspective images; however, this leaves out all the information behind and around the camera. Previous methods of obtaining a hemispheric (half-world) and omnidirectional (full-world) view of the world have involved fisheye lenses as well as pan and tilt cameras.

In this thesis, an omnidirectional vision system using two parabolic reflectors is described along with a technique and apparatus to obtain a higher resolution from an imaging system. The system has two parabolic mirrors for imaging. It allows a substantially hemispheric or half omnidirectional view of the world. Two of these sensors placed back to back would provide an omnidirectional view of the world from a single viewpoint. The results from the hemispheric, double parabolic mirror system that we built are shown in the thesis.

The main disadvantage of mapping such a large field of view onto a single sensor is the loss of resolution. We therefore describe a method to obtain better resolution using several image frames with control over the displacement of these frames. This involves meshing together several image frames displaced from each other by known amounts. The images are obtained from a variable angle prism normally used for image stabilization in camcorders. The results are promising.

Table of Contents

| | |
|--|------------|
| Abstract | ii |
| List of Tables | vi |
| List of Figures | vii |
| Acknowledgements | x |
| 1 Introduction | 1 |
| 1.1 The Optical Problem | 3 |
| 1.2 The Resolution Problem | 7 |
| 1.3 Overview of the Thesis | 9 |
| 2 Hemispheric Imaging | 11 |
| 2.1 Proposed System | 11 |
| 2.2 System Parameters | 16 |
| 2.3 Advantage of Using Parabolic Reflectors | 20 |
| 2.4 Limitations | 21 |
| 2.5 Mapping Hemispheric Image into a Perspective Image | 22 |
| 2.6 Results | 23 |
| 2.6.1 Computing Requirements | 25 |

| | | |
|----------|---|-----------|
| 2.7 | Conclusion | 25 |
| 3 | Super-Resolution | 27 |
| 3.1 | Theory | 27 |
| 3.1.1 | Approach | 27 |
| 3.1.2 | One Dimensional Analysis | 28 |
| 3.1.3 | Image Formation | 32 |
| 3.1.4 | Methods for Reconstruction | 33 |
| 3.1.5 | Registration Problem | 34 |
| 3.1.6 | Restoration, Inverse Filtering | 36 |
| 3.1.7 | High Resolution Reconstruction | 37 |
| 3.2 | Simulation | 38 |
| 3.2.1 | Reconstruction with Known Shifts | 38 |
| 3.3 | Final Remarks on Super-resolution Problem | 46 |
| 4 | A Sub-Pixel Resolution System | 47 |
| 4.1 | Description of VAP | 47 |
| 4.2 | Applied Super-Resolution | 57 |
| 4.2.1 | Centroid Calibration | 60 |
| 4.3 | Control of the VAP | 62 |
| 4.4 | Experimental Results using ES-750 | 66 |
| 4.4.1 | Initial Tests | 66 |
| 4.4.2 | Open-Loop Results | 67 |
| 4.4.3 | Sampling at Subpixel Intervals | 76 |
| 4.5 | Issues and Problems | 80 |
| 5 | Conclusions | 83 |
| 5.1 | Future Work for Combined System | 84 |

| | |
|---|-----------|
| 5.2 Summary of Results | 86 |
| Bibliography | 88 |
| Appendix A Aberrations (see Chapter 2) | 92 |

List of Tables

| | | |
|-----|--|----|
| 2.1 | Mirror Specifications | 20 |
| 4.1 | USAF 1951 Resolution Test Chart (lppm) | 76 |

List of Figures

| | | |
|-----|--|----|
| 1.1 | Image from a Fisheye Lens | 4 |
| 1.2 | Orthographic Projection | 5 |
| 1.3 | Orthographic View of Ariel (NASA web site) | 7 |
| 1.4 | OMNICAMERA(Tm) | 8 |
| 2.1 | COnc Projection Image Sensor(COPIS) | 12 |
| 2.2 | View of COPIS Sensor | 12 |
| 2.3 | Two Mirror Hemispheric Imaging System | 13 |
| 2.4 | Convex Parabolic Reflector | 14 |
| 2.5 | Two Mirror Hemispheric Vision system | 15 |
| 2.6 | Two Mirror Hemispheric Vision System | 17 |
| 2.7 | Hemispheric Scene taken with Double Parabolic System | 24 |
| 2.8 | Perspective Image from Hemispheric View(Fig. 2.7) | 24 |
| 2.9 | Perspective Image found by Deconvolving Fig. 2.7 | 24 |
| 3.1 | Ideal pixel PSF array with non-zero separation | 28 |
| 3.2 | Spatial Signals of the Sensor and Image | 30 |
| 3.3 | Averaging of a higher resolution image sequence | 31 |
| 3.4 | Original 256x256x8bit Lena Test Image | 38 |
| 3.5 | Simulation Model | 40 |

| | | |
|------|--|----|
| 3.6 | Low Resolution 64x64x8bit Lena Image Frame (obtained by applying H2 to Fig. 1) | 41 |
| 3.7 | High Resolution Reconstructed Image (no deblurring applied) . . . | 42 |
| 3.8 | Wiener Filtered 256x256x8bit Lena Test Image (from Noisy Frames) | 44 |
| 3.9 | Image Correlation between Frames | 45 |
| 4.1 | Complete VAP Assembly (US 5481394) | 48 |
| 4.2 | Front and Rear Plates of VAP (US pat. 5481394) | 50 |
| 4.3 | Large Prism | 51 |
| 4.4 | One Dimensional Structural View of VAP | 51 |
| 4.5 | Geometry of VAP movement (US pat. 5481394) | 53 |
| 4.6 | Restriction of VAP movement (US pat. 5481394) | 54 |
| 4.7 | VAP actuator unit (US 5481394) | 55 |
| 4.8 | Side View of VAP Actuator (US 5623305) | 56 |
| 4.9 | Our Front End Differential Position Amplifier | 58 |
| 4.10 | VAP Deflection Angle versus Drive Voltage (US 5623305) | 59 |
| 4.11 | Open Loop representation of VAP | 62 |
| 4.12 | VAP Bode Plots (US 5623305) | 64 |
| 4.13 | Closed Loop representation of VAP | 65 |
| 4.14 | Dual Coil Wrapping of VAP Actuator (US 5623305) | 66 |
| 4.15 | Our Control Circuit | 67 |
| 4.16 | Vertical Displacement with Sinewave Input | 68 |
| 4.17 | Horizontal Displacement with Sinewave Input | 69 |
| 4.18 | Horizontal Direction VAP Frequency Response (0.5 V Sinewave) . | 69 |
| 4.19 | Vertical Direction VAP Frequency Response (0.5 V Sinewave) . . . | 70 |
| 4.20 | Square Wave Response | 70 |
| 4.21 | Displacement with Compensating Op Amp | 71 |

| | | |
|------|---|----|
| 4.22 | VAP run open loop in one direction | 72 |
| 4.23 | Hysteresis of VAP | 73 |
| 4.24 | VAP Drift over Time | 74 |
| 4.25 | VAP Displacement versus Applied Voltage | 75 |
| 4.26 | Zoom of Raw USAF 1951 Test Image | 77 |
| 4.27 | Image Reconstruction | 78 |

ACKNOWLEDGEMENTS

I would like to acknowledge Ray Burge, Dr. Greg Grudic, Henry Wong, Dr. Peter Lawrence and Dr. Tim Salcudean for their help, equipment use, and technical assistance. In addition, I would like to thank the technical staff who provided me with enormous assistance and patience (Don Dawson, Bruce Dow, Tony Leugner, Al Prince, Rob Ross and Lloyd Welder). Finally, I would like to thank Ray Burge again for his enthusiasm and exceptional help throughout the entire project.

Chapter 1

Introduction

Hemispheric sensors are being developed[Y⁺95, Nay96] to allow imaging with a field of view of 360 degrees. This means that every lateral direction of a scene can be viewed. The idea is to capture the whole scene around a fixed object and create a panorama or an omnidirectional view of the world at that point.

These sensors are extremely useful in such applications as video conferencing and surveillance. With the hemispheric view, several perspective images can be viewed of the world about the camera at the same instant. A perspective image is one which is linear and undistorted, as in normal eyesight. In a video conferencing application, a single sensor could sit on the table and the receiver at the other end could pick which participant to view at a given time or choose to view some or all of them at once. Similarly, in surveillance, the ideal situation is to have a camera that cannot be compromised. No one can sneak up on or dodge an omnidirectional camera. With normal cameras the security is limited if the person being watched knows where it is pointing. The most popular solution to this security problem for normal non-omnidirectional cameras is to mount them inside a silvered dome so

that no one knows if they are being watched.

More applications for omnidirectional cameras will materialize from the benefit of seeing the whole scene around a fixed point. They have a larger field of view than the unaided eye and as such can image more of the surroundings than a casual observer or a conventional camera. With zooming capability, the possibilities for use increase. In this thesis we will discuss a new design for an omnidirectional camera. We then discuss a system that can be used to overcome the main limitation of the camera, which is a lack of resolution.

The first attempts at omnidirectional imaging involved a fish eye lens[H⁺86]. These lenses were built as an extreme wide angle lens and are still popular today for many applications. More recent methods have involved pan and tilt cameras[KA96]. These are cameras mounted on a platform containing two joints that can be panned and tilted to capture the scene at each angle. The multiple scenes so derived are pieced together to create either a panorama(when there is only panning and no tilting of the camera) or an omnidirectional view. The pan and tilt is popular in speed independent applications such as real estate imaging. In this application, the client can then look at each scene around different rooms or landscapes without visiting the site. Clearly, it is not important for the views to be in real-time.

Recent omnidirectional cameras are using mirrors to perform the imaging. They are formed as standard conic sections and offer the best promise for less distortion of the reconstructed perspective images. Our system is part of this subset of omnidirectional sensors, the advantages of which will become apparent to the reader later in this thesis.

1.1 The Optical Problem

Conventional imaging systems are very limited in their field of view. Cameras are designed to image plane perspective scenes with no distortion in the image plane. Attempts to recover a hemispheric scene originated with fish eye lenses. We propose a system with two parabolic mirrors to accurately image a scene about a complete hemisphere. In addition, a resolution recovery scheme is proposed to compensate for the loss of information inherent in imaging such a large field of view. Putting two of our cameras back to back allows almost a full omnidirectional view.

Fish Eye Lens

The common fish eye lens, such as that seen on door peepholes, is a lens designed for a wide field of view. The problem with these lenses is that it is difficult to make two lenses identical in diameter and focal length that have the same distortion pattern[Nal96]. It is difficult to determine how the distortion affects the image. Perspective images cannot be recovered from a fish eye lens without knowing the exact geometry of the lens or the exact distortion caused by the lens, which varies from lens to lens and takes considerable time to measure. Fish eye lenses are also notoriously expensive to manufacture. A scene imaged from a fisheye lens([SA94]) is shown in Fig. 1.1.

Pan and Tilt

Another approach for the imaging of a hemispheric scene has been the use of a pan and tilt camera. The camera is essentially mounted on a platform that can be pan and tilted to acquire images from a full hemisphere. The images are then

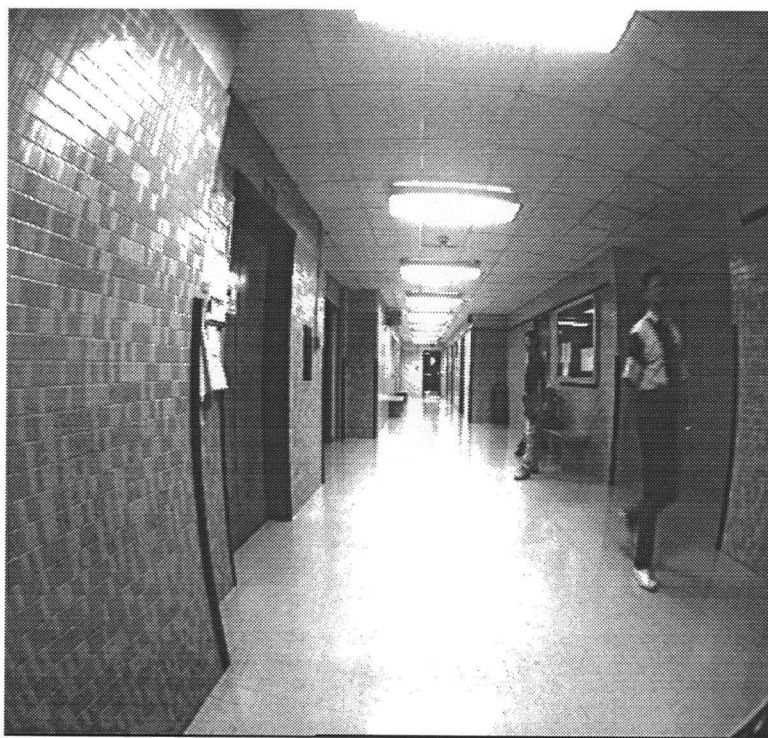


Figure 1.1: Image from a Fisheye Lens

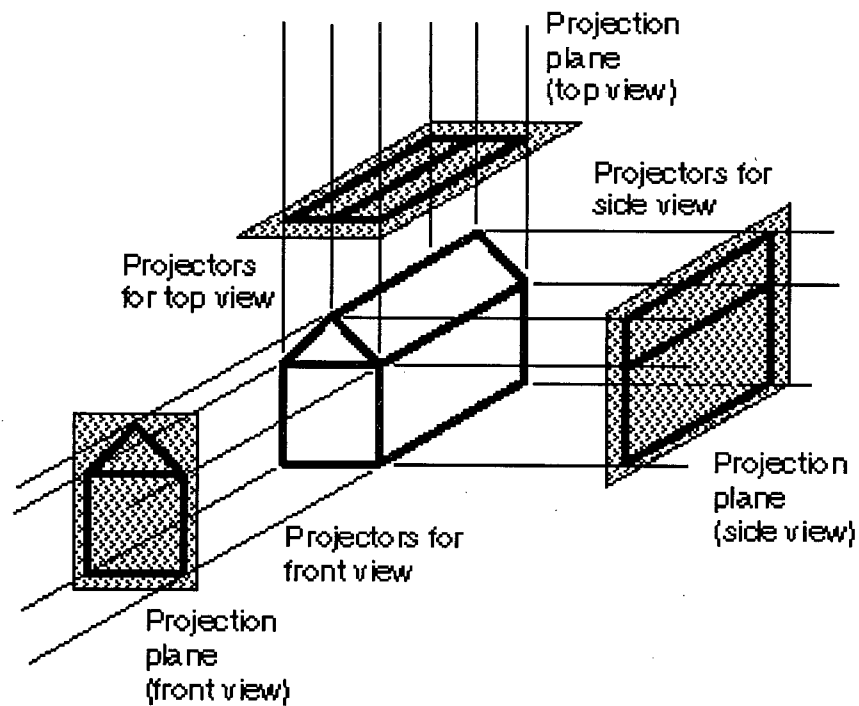


Figure 1.2: Orthographic Projection

stitched together to recreate the hemispheric scene. This method suffers from serious problems. Stitching the scenes together is difficult because the overlapping image boundaries must be identified for each scene. The mechanical nature of the mount used for panning and tilting can also create accuracy problems. There can be calibration problems in keeping the center of the imaging surface at the center of movement. The two main disadvantages of this system are the required high mechanical accuracy and the slow speed. A stepping motor must be used and it cannot be driven fast or backlash will plague the response.

Omnicaamera

The Omnicaamera(Tm) (Fig. 1.4) was developed by Shree Nayar at Columbia University[NPB98]. He investigated omnidirectional catadioptric (which means a combined mirror and lens system) sensors using single mirrors with a conical cross section across the optical axis. The goal was to find systems with a single view-point. The proposed solution was to use a single parabolic primary mirror. Above the primary mirror, a camera system with an orthographic lens is used to image the parallel lines from the lower mirror. An orthographic lens is a telecentric (centred on the optical axis) lens that images the full side of the object with the line of sight perpendicular to that object. Fig. 1.2 shows several orthographic projections of an object[Spe84]. This type of view is commonly used in satellite imaging of planets (Fig. 1.3).

In the Omnicaamera, the lower mirror in his system performs the same purpose as the one we describe later, which is to allow for a single viewpoint. The orthographic lens must be the same size as the lower mirror in order to get the full image. Because of the imaging geometry, the camera must be mounted outside the dome (see Fig 1.4). This is a disadvantage because the camera must be large to accomodate the orthographic lens size and is therefore easily detectable, damaged or knocked out of alignment due to its position. In addition, the orthographic (telecentric) lens is expensive and requires a fair amount of calibration to set up with the camera and align with the optical axis.

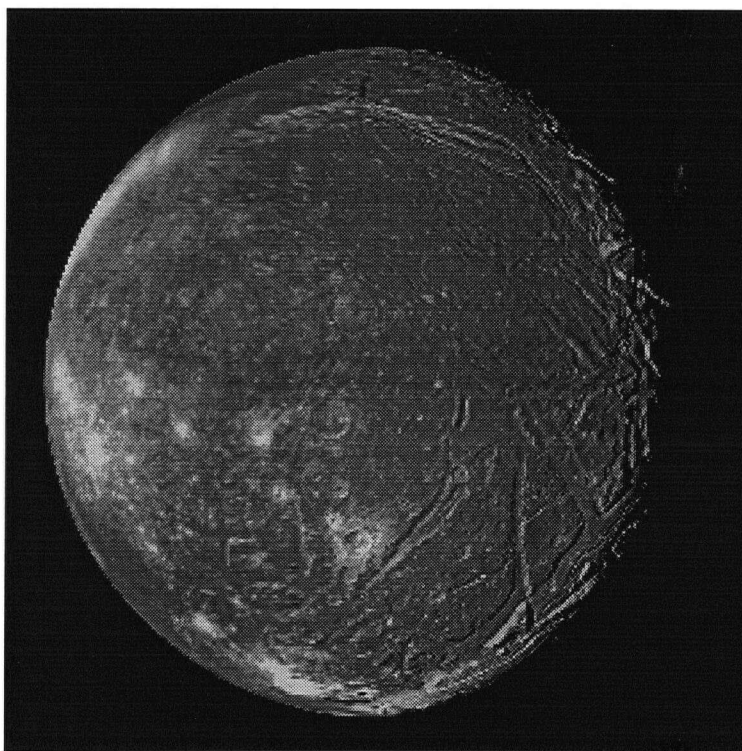


Figure 1.3: Orthographic View of Ariel (NASA web site)

1.2 The Resolution Problem

In many imaging situations it is desirable to get a higher resolution than is obtained through the standard optics. This is the especially the case in wide field of view systems. Every optical system, however, will have a limiting resolution based on the optics or the detector grid spacing and size. Even the most expensive components degrade the resulting imaged view of the real world to some extent. For further resolution improvement, it becomes necessary to choose a different approach than just purchasing the best optics and detectors.

Resolution is an often misunderstood term. The common misinterpretation involves thinking it is the pixel size of an image. A fixed 320x320 picture does not necessarily

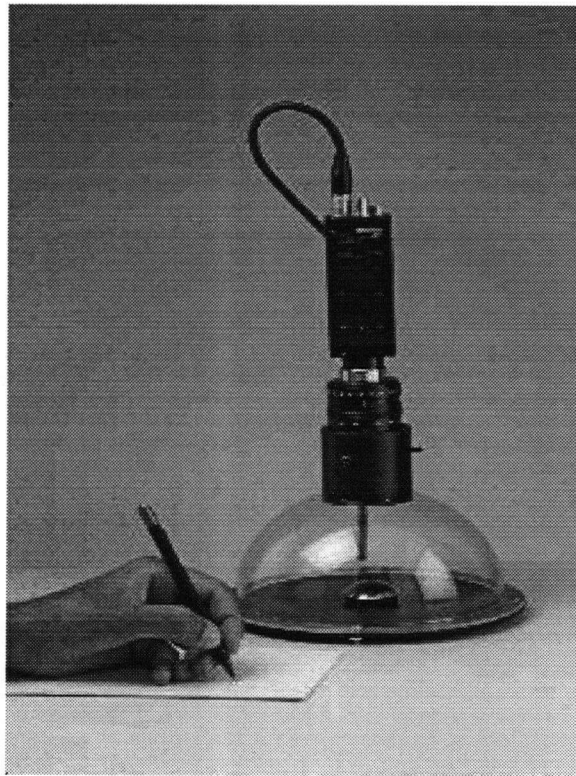


Figure 1.4: OMNICAMERA(Tm)

have its resolution doubled along each dimension when it is resized to 640x640. An image manipulation program simply doubles the pixel size of the original image. There is no new information to be obtained. Resolution of a picture is measured by the bandwidth of the signal from the sensor (lines per mm or otherwise)[Gro86].

If one defines 'resolution' as the number of lines visible in the image divided by the field of view, there are three factors that affect the resolution. First is the sharpness of the optics. Second is the number of pixels in the sensor used to cover the field of view. The third factor affecting resolution is the size of each pixel. The smaller the pixel size, the more ideal the sampling but the lower the light sensitivity of the camera. Resolution as defined in this thesis is measured in lines per millimetre.

Recovering the actual real world from a single low resolution image is not possible because the optics and sensor will always have a low pass filtering effect on the image. What is needed is to use multiple images in a constructive manner to recover that lost information.

One approach for recovering high resolution is to look at several low-resolution images of the same specimen. If each image is shifted from each other by a known amount, a higher resolution reconstruction can be done. When the amount of shift is not known, it is possible to model the registration mismatch to determine the relationship between images[KBV90, MS88, SPS87].

In summary, all components of the optical system affect the final image and image resolution. The main degradation however, with today's quality optics, almost always occurs at the detector or sensor[Luk66, Luk67]. This is particularly the case for CCD or CMOS cameras and imaging systems[Nor78]. In this thesis, we will model the detector and use a linear procedure to recover the high resolution by using multiple images.

1.3 Overview of the Thesis

This thesis directly addresses the problems identified in the state of the art by introducing a new optical approach to solving the problems in a hemispheric imaging system.

First, in Chapter 2, a hemispheric imaging system is proposed. The issues involved in imaging a complete scene, including angle of view, blindspot and resolution, will be examined. Finally, our proposed system is described as we constructed it.

The theory for obtaining higher resolution from low resolution frames, also known

as the super-resolution problem, is described in Chapter 3. This is followed by a test of the theory using standard image-processing test images.

Subsequently, in Chapter 4, a method of controlling an optical system to carry out the increased resolution is described. It involves using a prism system with electronic feedback control through PSD sensors to shift individual image frames. The subsequent shifted images have a known mismatch and can be pieced together in a manner to get improved resolution.

The final chapter summarizes the results and briefly describes how the two systems should be integrated for an optimal imaging system.

Chapter 2

Hemispheric Imaging

2.1 Proposed System

The newest systems for hemispheric imaging involve the use of a single mirror with a conic section of revolution. Yamazawa et al. have proposed an omnidirectional image sensor for robot navigation called COPIS (CONic Projection Image Sensor)[Y⁺95] using a conical sensor sitting on top of the robot. The robot with the sensor attached is shown in Fig. 2.1 and a close-up of another version of the same mirror system, not attached to the robot (and mounted in a different dome), is shown in Fig. 2.2. Another researcher, Shree Nayar, has expanded the COPIS research by studying the family of omnidirectional sensors that use one mirror with a conical cross-section[Nay96]. He found that the best results occur with a parabolic reflector. His final system uses one parabolic reflector with an orthographic lens and camera above the mirror (Fig. 1.4).

We propose a system with two parabolic mirrors as shown in Fig. 2.3 and Fig. 2.5. It is similar to Nayar's system but does not require a large and expensive orthographic

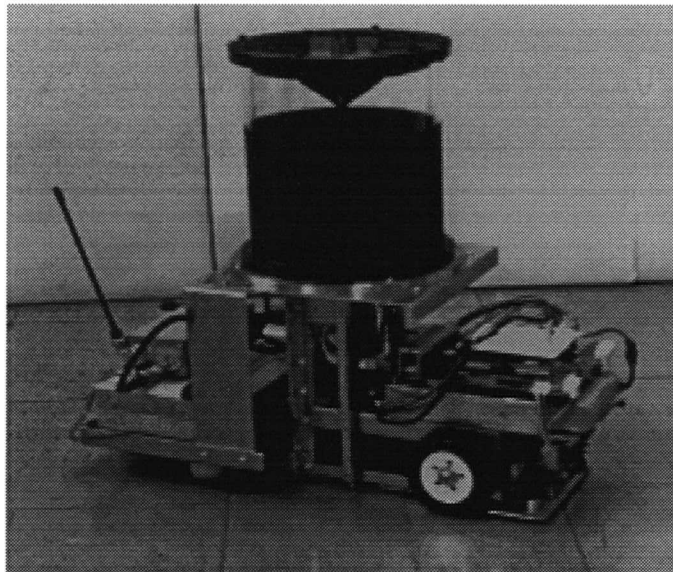


Figure 2.1: COnic Projection Image Sensor(COPIS)

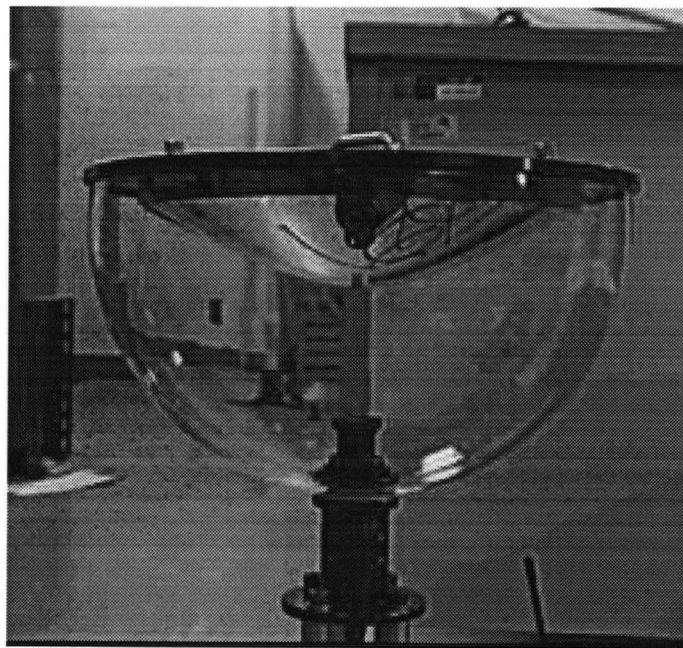


Figure 2.2: View of COPIS Sensor

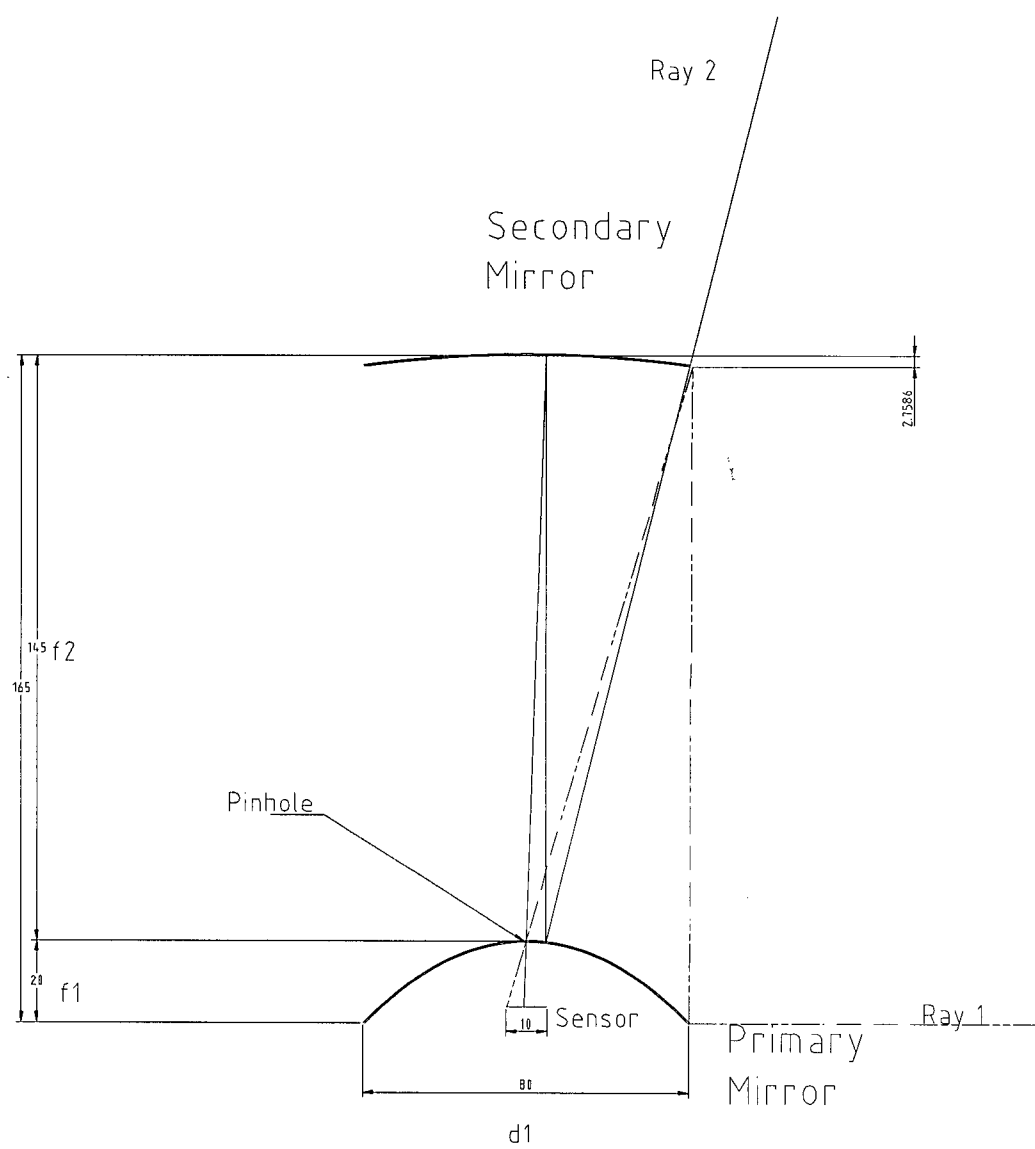


Figure 2.3: Two Mirror Hemispheric Imaging System

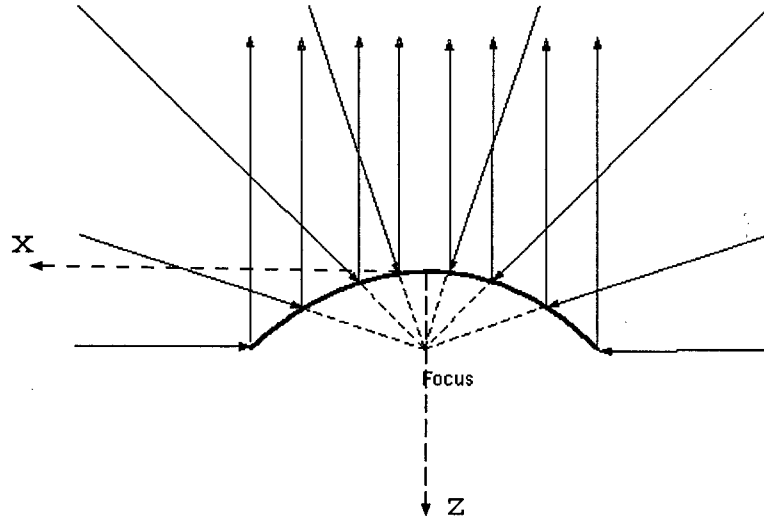


Figure 2.4: Convex Parabolic Reflector

lens. The first mirror or primary mirror in the system is a convex paraboloid. Fig. 2.4 of the primary mirror shows that rays passing through the focus are reflected parallel to the mirror's axis of symmetry. The diameter of the primary mirror is equal to four times the focal length, allowing points on the horizon to be imaged. The equation for the primary is that of a standard parabola, $z_1 = \frac{(h_1)^2}{4f_1}$, where h_1 is the radius $\sqrt{(x_1)^2 + (y_1)^2}$ from the vertex (origin) of the primary.

The upper or secondary mirror is a concave paraboloid with the same clear aperture(diameter) as the bottom mirror, but with a longer focal length. It has the equation, $z_2 = \frac{(h_2)^2}{4f_2}$, where h_2 is the radius $\sqrt{(x_2)^2 + (y_2)^2}$ from the vertex(origin) of the secondary mirror. The parallel rays reflected from the primary mirror are reflected off the secondary and converge at its focus. A pinhole or a small lens located at the vertex of the primary is located at the secondary mirror's focus. This allows light to pass through the primary mirror. A sensor located under the primary mirror detects the image formed by the mirror system.

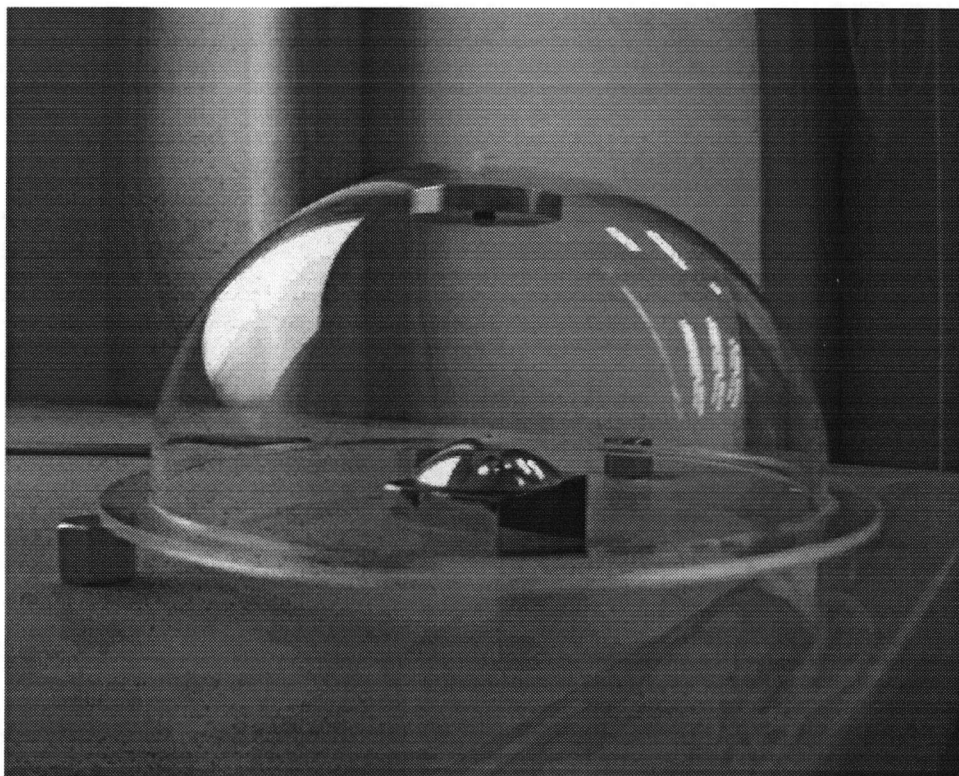


Figure 2.5: Two Mirror Hemispheric Vision system

The pinhole has the purpose of blocking out any stray light which would produce noise in the image. Drilling a pinhole directly on top of the primary parabola is very difficult. Therefore, it is proposed to have a larger aperture in the primary. This aperture is then covered by a sheet with a pinhole in it. This makes calibration easier without the possibility of ruining the primary mirror by drilling an off-center hole. The focal length of the secondary mirror is equal to the distance from the secondary vertex to the pinhole. Although the new pinhole location is not right at the primary vertex, it is still located on the optical axis.

The image formed by the system has a single viewpoint (located at the primary mirror focus) and provides a hemispheric field of view. As outlined by Nayar[Nay96], a single viewpoint allows for demapping of the omnidirectional image into perspective images with the geometry of the scene preserved. A raw CCD or CMOS sensor can be placed below the pinhole in the primary mirror.

Perspective images are formed by first choosing the viewing direction and then using appropriate software to decode the simple mapping done by the mirrors. In addition, by putting two of the systems back to back, a complete omnidirectional system is achieved with minimal blind spots.

2.2 System Parameters

Figure 2.6 shows the entire two mirror system and two rays. In the system, both mirrors have the same clear aperture (defined as the optical diameter of the component). The clear aperture of the primary was chosen so that a ray (ray *a* in the figure), from the horizon (at 90 degrees with respect to the optical axis) would pass through the focus such that it is reflected parallel to the optical axis. This prerequisite requires that the clear aperture of the primary (and subsequently of

the secondary because they share the same clear aperture) be $4f_1$. The parameters involved with Figure 2.6 are defined as follows:

$$h_{bi} : \text{radius of ray b on surface i} \quad (2.1)$$

$$f_i : \text{focal length of mirror i} \quad (2.2)$$

$$\delta_{zXi} : \text{distance from mirror i in z-direction to where ray X hits the mirror} \quad (2.3)$$

$$\rho_s : \text{radius of sensor surface} \quad (2.4)$$

$$\rho_{aS}, \rho_{bS} : \text{radius of ray a, b on sensor surface} \quad (2.5)$$

$$f_s : \text{distance of sensor from P (defined as the sensor focal length)} \quad (2.6)$$

$$\theta : \text{angle subtended from optical axis by ray b} \quad (2.7)$$

The symbol 'P' in the diagram indicates the pinhole at the vertex of the primary. The symbol 'S' in the diagram shows the location of the sensor (CCD) below the primary mirror. The figure shows that there is a blind spot where the environment cannot be imaged. This blind spot caused by the upper mirror is centred about the optical axis of the system. The blind spot diameter, $2 * \rho_{bS}$, can be found by looking at ray b, which was chosen to show how the secondary mirror blocks the imaging.

From ray b, the following equation can be written to determine the blind spot diameter on the CCD surface:

$$\tan \theta = \frac{2f_1}{f_1 + f_2 - \delta z_2} \quad (2.8)$$

where δz_2 is the local (with the local mirror apex as the origin) z coordinate of where ray b grazes mirror 2. This coordinate, δz_2 , is found from the equation for the parabolic surface of mirror 2. Substituting $h_2 = 2f_1$, simplifies the equation for the point in question to $\delta z_2 = \frac{(f_1)^2}{f_2}$. The equation for the blind spot angle (where

this angle specifies the ray which creates the blind spot on the CCD) then becomes:

$$\tan \theta = \frac{2f_1}{f_1 + f_2 - \frac{(f_1)^2}{f_2}} \quad (2.9)$$

Now consider mirror 1 with the focus as the origin. An equation for the radius of a point on the mirror with respect to the axis of symmetry (optical axis) is as follows (standard parabolic equation as defined from the focus):

$$h_1 = \frac{2f_1}{1 + \cos \theta} \sin \theta \quad (2.10)$$

An approximation of how effectively the CCD surface is used can be found by looking at the radius of the extreme rays a and b on the imaging surface. For ray a the following can be written by using similar triangles:

$$\frac{\rho_{aS}}{f_s} = \frac{2f_1}{f_2 - \delta z_{a2}} \quad (2.11)$$

Surfaces 1 and 2 are parabolas. Therefore,

$$\delta z_{a2} = \frac{(2 * f_1)^2}{4f_2} \quad (2.12)$$

where f_s is the distance of the imaging surface to the center of the pinhole or lens at the apex of surface 1. Using similar triangles, a similar equation can be written for ray b:

$$\frac{\rho_{bS}}{f_s} = \frac{h_{b2}}{f_2 - \delta z_{b2}} \quad (2.13)$$

which simplifies using the parabola equation for surface 2 ($z_2 = \frac{(h_2)^2}{4f_2}$) to:

$$\rho_{bS} = f_s \frac{2f_1 f_2}{(f_2)^2 - (f_1)^2} \quad (2.14)$$

| | |
|--------------------------------------|----------|
| Primary Focal Length (f_1) | 1.02 cm |
| Secondary Focal Length (f_2) | 7.24 cm |
| Primary Mirror Diameter ($4f_1$) | 4.06 cm |
| Secondary Mirror Diameter ($4f_1$) | 4.06 cm |
| Mirror Separation (f_2) | 7.24 cm |
| Blindspot Angle (θ) | 14.0 deg |
| Pixel Use Fraction (Eq. 2.15) | 0.35 |

Table 2.1: Mirror Specifications

Recall that ρ_{bS} is the blind spot radius on the CCD.

In one dimension, the number of pixels that are covered can be found from the ratio $\frac{2(\rho_b - \rho_a)}{\text{pixel dimension}}$. So for two dimensions, the number of pixels that are covered in a square CCD of dimension l by l is:

$$\frac{N\pi((\rho_a)^2 - (\rho_b)^2)}{l^2} \quad (2.15)$$

where N is the number of pixels.

Table 2.1 summarizes the mirror parameters for the system.

2.3 Advantage of Using Parabolic Reflectors

Having a single viewpoint is advantageous because all the information from the environment is seen from this reference point. If there were more than one viewpoint, as is the case for COPIS and fish eye lenses (where each scene point is imaged as viewed from slightly different positions), then the world representation would require

mapping all images into a common reference frame. It would also require the depth of points in the scene to be estimated or determined. This makes omnidirectional mapping difficult.

A sensor that can view the world from a single viewpoint allows the construction of perspective images easily. All that needs to be done is to look and find the mapping from the sensor to the point that is to be examined a certain distance (also called the focal depth) away from the sensor[Nay96]. As shown by Nayar[Nay96], the easiest mirror curvature to obtain a single viewpoint in a catadioptric system is the parabola. In our system, the primary mirror is also a parabola. In addition, the secondary mirror is a parabola that focuses those rays that are reflected parallel to the optic axis through the aperture in the top of the primary. The mapping is simple because the parallel rays reflected from the primary are those rays that are aimed at the focus of the primary. The single viewpoint is evident and perspective images can be found by finding the ray intersections of a given perspective plane on the primary mirror. The final hemispheric image obtained from the CCD is in fact just this but scaled down proportionally by the upper or secondary mirror.

2.4 Limitations

The advantages of our system have previously been examined. In summary, they are as follows: a single viewpoint, a reduced blind spot, a hemispheric field of view, no complex lenses, an omnidirectional field of view with two cameras back to back and compact design.

The limitations come from several considerations. In reality, the CCD has a fast exposure time. The pinhole imaging technique requires long exposures. For some applications, however, real-time imaging is a necessity. In this case, the pinhole must

be replaced by a small lens. In fact, pinhole CCD cameras on the market have a small pinhole lens in them. The lens is inexpensive and does not degrade the image significantly. For our system, using a pinhole lens will not ruin the fundamental imaging properties of the system as a whole. All that needs be done is choose a focal length to image the scene onto the CCD sensor a fixed distance away.

Another problem is with the CCD sensor itself. CCDs and CMOS sensors typically have 512x512 or 480x640 pixels. When the mapping to perspective images is done, the number of pixels is reduced below this number. So to achieve similar image quality on the perspective images as conventional cameras, it is necessary to have a CCD with more pixels. This limitation in CCD resolution also hampers the quality of images obtained from the Omnicamera and COPIS mentioned earlier.

2.5 Mapping Hemispheric Image into a Perspective Image

One of the benefits of having a hemispheric image is that perspective images can be formed in any direction. The image is a compact storage of an entire scene. The perspective image can be formed by specifying the viewing direction, the focal length and the desired image size.

Once the above parameters are set, it becomes easy to produce the desired perspective image. First, the center of the omnidirectional image must be known or determined. After this is known, the radius of the omnidirectional image is found in pixels. Using the desired viewing direction, focal depth (zoom distance away from the viewpoint) and image size, the hemispheric image is stepped across in increments and mapped to a new perspective grid. The resulting data is interpolated to

form a final digital image. A brief description follows:

- Find the radius and the center of the hemispheric view
- Choose the number of steps to cross the image
- Choose the image size and location (angle of view) as well as the focal depth
- Define a matrix of X and Y co-ordinates on the perspective image
- Map the perspective image points from the hemispheric image using the parabola equations and the inputs given above
- Interpolate the points (from the previous step) to the corresponding ones on the hemispheric image to form a perspective image

2.6 Results

Using a standard stock lens, a hemispheric scene was imaged. The image is shown in Fig. 2.7. The results would be better with a better CCD camera with more pixels (such as the recently released MEGA-pixel cameras with 16 times as many pixels). In addition, the fraction of pixels used on the CCD (see Eq. 2.15 derived earlier) was only 0.35. The camera could not be brought close enough inside the primary mirror because there were obstructions on the camera board thereby leaving much of the CCD surface unused. A custom CCD would have been required to fit the mirror system properly.

Using the interpolation algorithm, perspective images were obtained with a software unmapping. These are shown in Fig. 2.8 and Fig. 2.9. Any direction can be chosen

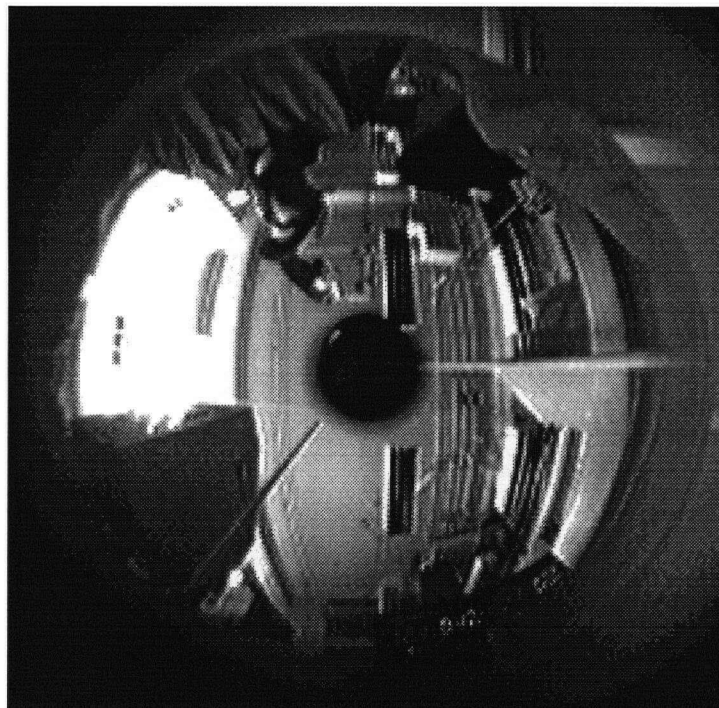


Figure 2.7: Hemispheric Scene taken with Double Parabolic System

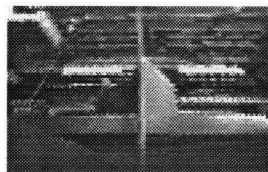


Figure 2.8: Perspective Image from Hemispheric View(Fig. 2.7)



Figure 2.9: Perspective Image found by Deconvolving Fig. 2.7

with similar results. Near the lower extremities, a larger scene area is mapped to the fixed CCD space.

The pipes on the ceiling in Fig. 2.9 are straight as expected, but they are not very clear. This is because of the lack of resolution and the unfinished calibration of the system. The lens did not map the hemispheric image onto the whole CCD surface. In fact, much of the CCD surface was wasted. The results do show the apparatus works and that better results would be obtainable by using a better CCD and having a pinhole lens with a focal length matched to the system so less of the CCD surface was left unused.

2.6.1 Computing Requirements

The unmapping was done using the Matlab(TM) matrix manipulation program. Performing an unmapping to a 70 by 128 pixel image took 544047 floating point operations (flops) on a SPARC 5 with a 85MHz microsparc II processor. This corresponded to a CPU time of 18 seconds. For a m by n image, the computing power order of magnitude in flops is $O(60mn)$. Using the data above, a 640 by 480 picture would take about $1.9E7$ flops. With a fast enough processor and a good compiler, real-time unmapping may be possible at 640 by 480 resolution.

2.7 Conclusion

Although double parabolic systems have been used extensively in astronomical telescopes, they have yet to be used for wide angle imaging. We have proposed a lenless or pinhole lens imaging system for hemispheric imaging. It can be easily manufactured using two inexpensive plastic mirrors and a single image sensor. The design

is much less cumbersome than Nayar's, yet it still provides similar image quality for a CCD with the same number of pixels, and a single viewpoint for distortion free perspective images. An omnidirectional (360 degrees in every direction) system is possible by putting two of our systems back to back, but this would require two sensors. The main application of our system would be for security and video-conferencing. The system would be fully integrated and not require any external mounting of cameras. A preliminary sketch of the system without opto-mechanical components is shown in Fig. 2.3

Chapter 3

Super-Resolution

3.1 Theory

This chapter deals with the problem of resolution enhancement. In the previous chapter, a hemispheric imaging system was described for use with a CCD sensor. The main drawback is with the limited number of pixels on the sensor which means less information to create the perspective images a person would see normally with their own eyesight. The chapter starts with the simplest case, a one dimensional analysis. The more general problem in 2D is then discussed and a simulation performed using standard test images.

3.1.1 Approach

For our resolution simulations, we will assume a knowledge of the high-resolution image. The degrading of the image by the sensor point spread function (PSF) will also be known or approximated. Restoration of images without knowledge of

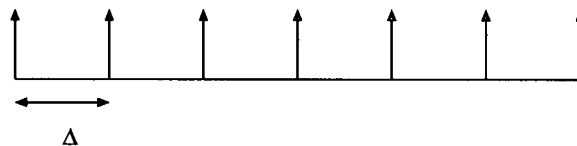


Figure 3.1: Ideal pixel PSF array with non-zero separation

the imaging procedure has been well studied in the literature[MS88, KBV90]. High resolution reconstruction is different from restoration in that a higher final sampling rate and higher detail can be achieved.

For the purpose of our analysis, we will assume that the image incident on the sensor array is continuous. The sensors used for imaging have a point spread function (PSF) with a much larger magnitude than a single object point source. A typical CCD sensor has pixels separated by 13 to 20 μm . Each pixel has a PSF that is very small but non-zero width. We will assume that the resolution of the picture is limited only by the PSF of the detectors. This is a reasonable assumption for CCD cameras as outlined by Gross[Gro86].

3.1.2 One Dimensional Analysis

A linear sensor with ideal pixels is shown in Fig. 3.1. An actual sensor will have pixels with a non-zero point spread function. The non-zero width PSF of an individual pixel is important to consider because it imposes the final limit on the resolution improvement. The resolution cannot be improved by more than this PSF[JR84].

Consider a linear image, y , formed on a pixel array. It will be assumed that each pixel has identical an PSF, h , and that they are separated by Δ . The actual object,

x , is a continuous function. The image can then be written as:

$$y(z) = \sum_{n=0}^{M-1} h(z) * x(z) \delta(z - n\Delta) + v(n\Delta) \quad (3.1)$$

where M is the number of pixels, $*$ is the convolution operator, and v is the observation noise. Fig. 3.2 shows what the spatial signals of the sensor and the image might look like. Notice the large space between the pixel PSFs. Without this space, there could be no resolution improvement with the method we describe later. Luckily, this is a valid assumption as most CCDs have large separations between relatively sharp pixel PSFs[Gro86].

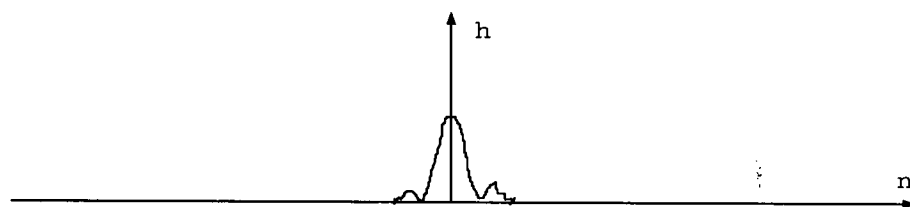
With just a single image, it is not possible to recover the high resolution content of the original scene. If several images are taken, and the displacement of these images relative to each other is known, then it is possible to recover a higher resolution. Because the individual detectors are non-ideal low pass filters, the high frequency information is contained in the original image in the form of aliasing. If the individual detectors were perfect low-pass filters, then it would not be possible to recover the lost details.

Consider now

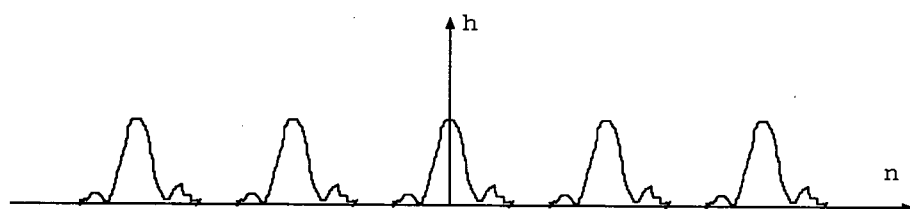
$$y_i(z) = \sum_{n=0}^{M-1} h(z - \alpha_i) * x(z) \delta(z - n\Delta) + v(n\Delta) \quad (3.2)$$

where y_i is the image frame at a displacement of α_i relative to the base frame($\alpha_0 = 0$). Suppose that each frame is displaced from each other by a fraction of the pixel size of the low resolution frame, y_i . If there are L distinct frames taken at different displacements but with α_i less than the pixel size of y (i.e. between the low-res pixels), then it is possible to achieve a 1D resolution improvement by a factor of L .

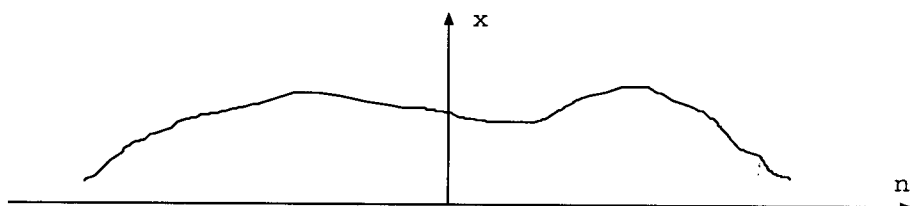
An example of the above follows. Figure 3.3 shows (in a simplified manner) how y_i averages the high resolution pixel sequence x . Let $y_1 = [y_1(0), y_1(1), y_1(2), y_1(3)]$



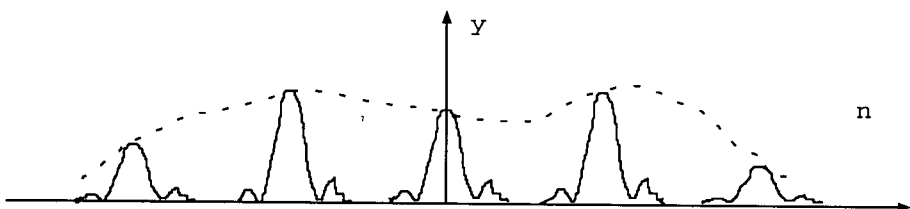
(a) Single Detector Response



(b) Sensor Array Impulse Response



(c) Image Signal



(d) Local Response of Sensor to Image

Figure 3.2: Spatial Signals of the Sensor and Image

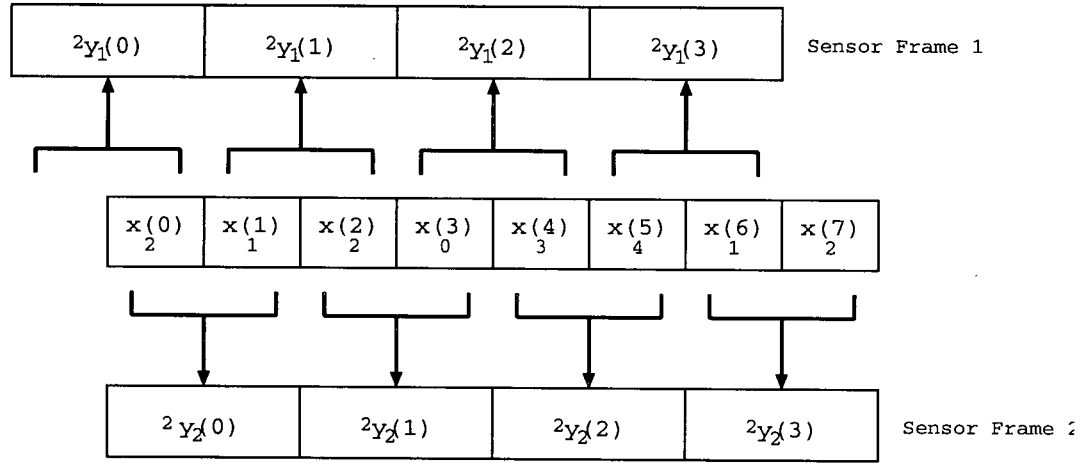


Figure 3.3: Averaging of a higher resolution image sequence

where $y_1(0) = \frac{x(-1)+x(0)}{2}$, $y_1(1) = \frac{x(1)+x(2)}{2}$, $y_1(2) = \frac{x(3)+x(4)}{2}$, $y_1(3) = \frac{x(5)+x(6)}{2}$. Now consider that the sensor frame is shifted by $\alpha_2 = \Delta$ where Δ is the spacing between the pixels of x . We then have $y_2 = [y_2(0), y_2(1), y_2(2), y_2(3)]$ where $y_2(0) = \frac{x(0)+x(1)}{2}$, $y_2(1) = \frac{x(2)+x(3)}{2}$, $y_2(2) = \frac{x(4)+x(5)}{2}$, $y_2(3) = \frac{x(6)+x(7)}{2}$. If and only if $x(-1)$ is known to be zero, then the above gives a solvable set of equations for the high resolution image. Otherwise, shifting the sensor will result in extraneous information being mapped onto the edge pixels, thus corrupting the procedure. If the images, y_i , had larger pixels, then more images would be required to recover the resolution.

An example of the process for 1-D reconstruction with known shifts follows, using an image at the higher level, $x = [2, 1, 2, 0, 3, 4, 1, 2]$, finite in extent such that $x(n) = 0 : n < 0, n > 7$. The low resolution images y_1 and y_2 are obtained as above but with the second image y_2 shifted by one half pixel to the right with respect to y_1 . The result is $y_1 = \frac{1}{2}[2, 3, 3, 5]$, and $y_2 = \frac{1}{2}[3, 2, 7, 3]$. Meshing the series to a larger grid yields $y_{mesh} = \frac{1}{2}[y_1(0), y_2(0), y_1(1), y_2(1), y_1(2), y_2(2), y_1(3), y_2(3)] = \frac{1}{2}[2, 3, 3, 2, 3, 7, 5, 3]$. The same result, y_{mesh} , could be obtained with the simple z

transform filter, $Y_{mesh}(z) = \frac{(1+z^{-1})}{2}$. The inverse filter is then simply $\frac{2}{1+z^{-1}}$ or $\frac{2z}{1+z}$. In the space domain this gives a recursive relationship for $x(n)$ as $x(n) = 2y_{mesh}(n) - x(n-1)$. Using $x(-1) = 0$, the original series $x(n)$ can be obtained perfectly. In practice, however, the actual image cannot have perfectly sharp boundaries. In other words, $x(n)$ would be infinite in extent; and this would cause some error when the above method is used. Interpolation is therefore used for real images as described later.

3.1.3 Image Formation

Imaging always involves the application of non-ideal sensors that degrade the optimal obtainable image. The narrower the PSF of the optics and sensor system, the better the resolution that can be achieved. The degradation of the ideal image $x(j, k)$ by a digital imaging system can be modeled as follows:

$$y_i(j, k) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n)h(j-m, k-n) + n(j, k) \quad (3.3)$$

where $h(j, k)$ is the degradation from the imaging process, $n(j, k)$ the additive noise, M is the number of pixels in the j (horizontal) direction, N the number of pixels in the k or vertical direction, and $y_i(j, k)$ the i th observed low-resolution image. Equation 3.3 can be rewritten in vector-matrix form as:

$$y = Hx + n \quad (3.4)$$

Ideally, the degradation should be minimized by using the best possible optics and sensors. Unfortunately, real detectors can only achieve a given PSF and become very expensive when they are of high quality. As mentioned earlier, we assume

that the sensor PSF is the primary limit on the resolution. Our goal is to recover the high-resolution image (by upsampling the image and increasing the maximum represented spatial frequency), $x(j, k)$, from several low-resolution images. The following summarizes the issues in the imaging system:

- **Sensor PSF:** blurs the image
- **Undersampling:** results in aliasing errors
- **Noise**

3.1.4 Methods for Reconstruction

Reconstruction is possible in both the spatial and frequency domains. Currently, the preferred technique is in the spatial domain[ST90] because lower noise and higher resolution have been achieved spatially. Spatial domain techniques include projection onto convex sets(POCS)[ST90], tomographic backprojection[IP91], interpolation, and minimization of error[TG94]. The frequency domain approach involves spectrum cancellation and eliminating aliasing artifacts[KS93]. The frequency domain method is currently only used when there is a registration problem.

The reconstruction problem has three parts to it. The first part is image registration. The amount of shift between the low-resolution frames must be determined to some accuracy. The second part is the restoration and filtering of the frames to compensate for degradation and noise. Finally, the high resolution image is reconstructed or pieced together using interpolation or other techniques.

Algorithm for high-resolution reconstruction of 2-D images:

1. **Registration:** determine how much each frame is shifted from each other

2. Restoration or Filtering: remove the degradation and noise from each low-resolution frame
3. Reconstruction or Interpolation: use the results of above to piece together a high-resolution image

It turns out that the above three steps are interdependent. There have been attempts to combine all three steps in a global procedure, but so far only steps 1,2 and 2,3 have been combined (in separate instances) successfully[KBV90, TG94] (in the spatial domain).

3.1.5 Registration Problem

Known Shifts

In the ideal case, we would know the exact amount that each low resolution image is shifted with respect to the reference frame. This is only possible when the imaging process is directly controllable. When the sub-pixel shifts are random, the problem is more difficult and becomes one of registration. Only interframe translations of the low-resolution images will be considered. Rapid movement imaging systems, such as aircraft radar, will have to treat the case of translations with frames not spaced at sub-pixel amounts.

Based on the assumption that resolution only depends on the number of pixels, an argument can be made to determine the improvement in resolution. With control over the subpixel shifts, it is possible to achieve a resolution improvement of \sqrt{L} where L is the number of low resolution frames[TOS92]. This is derived as follows. Let the low resolution images have a resolution of $K \times K$, and the high resolution scene a resolution of $N \times N$ ($M=N$). There are then $K \times K$ equations from the

low resolution frames and $N \times N$ unknowns from the individual pixels (in the high resolution scene). This means that the total number of equations that can be formed is $K \times K \times L$. In order to satisfy the requirement that there is the same number of equations as unknowns, it is necessary that $N = \sqrt{L} \times K$. In an ideal world if the pixels were ideal samplers and there is no low-pass filtering by the optics then the maximum theoretical resolution improvement is the square root of the number of distinct sub-pixel frames[TOS92].

The above argument may lead some to believe that an infinite resolution improvement is possible. Unfortunately, this is not the case because of the zero-crossings in the sensor frequency response. In other words, real pixels have a non-zero point spread function (see Fig. 3.2). The shifted images must be separated by at least the width of the PSF in the spatial domain or there will be overlap. Some high-frequency information will inevitably be lost as a result of the smoothing effect of the sensor response. Noise, as usual, aggravates the problem.

Unknown Image Displacements

When the shifts are unknown, there is a registration problem that must be solved before a high-resolution image can be reconstructed. The formation of the low-resolution images through the system is considered as before. The degradation is usually assumed to be a low-pass blurring or averaging function. Alternatively, the images are modelled by a Gauss-Markov covariance model.

If the original image were known, it would be possible to use matched filtering to determine the best correlation for a given shift. In fact, even though the original scene is not known, matching and correlation can still be used. The high-resolution image is estimated and the shifts varied to produce the best possible reconstruction.

The original image, $x(j, k)$, is not known with certainty, however, so several iterations must be performed with different estimates based on the samples, y_i . Mort and Srinath suggested a maximum likelihood image registration algorithm[MS88] that requires no a priori knowledge of $x(j, k)$.

3.1.6 Restoration, Inverse Filtering

Once the low-resolution frames have been sampled and their relative shift determined, it is necessary to try to remove some of the noise through inverse filtering and restoration techniques. The problem with most inverse filters is that they are unstable and not physically realizable. Pseudoinverse filters attempt to solve this problem by setting the frequency response to zero at ill-posed points. The pseudoinverse filter, however, remains highly sensitive to noise. The preferred technique is Wiener filtering because it removes the noise and some of the blur caused by the low-pass filtering.

When there is no registration problem, the restoration procedure is straight forward and lends itself well to conventional inverse filtering techniques. If there is a registration problem, it is preferable to combine the registration and inverse filtering techniques using an Expectation-Maximization(EM) algorithm[TG94, TK94].

In the EM algorithm, the degradation process is modelled as in Eq. 3.3 and an image covariance model[Jai88]. An initial guess is taken for the shift of each low-resolution frame. Each shift is then incorporated into the covariance model. The resulting modified probability density function of the observed image is then maximized as the shifts are iteratively estimated. Another similar iterative reconstruction technique for nonuniformly spaced samples is proposed by Sauer and Allebach[SA87].

3.1.7 High Resolution Reconstruction

After the optimal shift estimation (if unknown) and restoration have been performed, the final step is to piece the images together. Ideally, a global procedure will be developed to perform all the above taking the interdependences into account. If the shifts of the low-resolution images and a priori knowledge of the final image are known, a recursive approach combining restoration and reconstruction may be the best solution[KBV90]. Otherwise one procedure is to interpolate based on the shifts and amount of noise in the images.

Interpolation can be done using any of the available techniques. Linear interpolation may be the best option, but the choice ultimately depends on the relative shift of each frame. Other interpolation methods include such novel techniques as fractals and more conventional ones involving polynomials.

The reconstruction can be seen as a problem of solving a set of simultaneous equations formulated by Eq. 3.3 for each frame. The equations are linearly independent if the displacement between frames is at subpixel amounts. Since the equations are coupled together with each other, the solution can be simplified by either tomographic backprojection[IP90] or projection onto convex sets[SO89]. Additional constraints (such as finite energy and limited support) have been added to the above methods to account for observation noise[TOS92].



Figure 3.4: Original 256x256x8bit Lena Test Image

3.2 Simulation

3.2.1 Reconstruction with Known Shifts

In order to demonstrate how the resolution recovery procedure is achieved, a simulation was performed. The Lena test image (Fig. 3.4) with 256x256 pixels and 8 bits per pixel was used. The Lena image is a standard reference in image processing. Low resolution samples were taken from the picture. These samples were then be manipulated in an attempt to recreate the high-resolution image. The Lena test image was originally taken with a still camera onto conventional film. The image has since been digitized and standardized for use in imaging processing tests.

The above test image (Fig. 3.4) is the reference image, i . It is impossible to recover

this image perfectly because of the quantization effects and added noise. We consider the model of the sensor shown in Fig. 3.2 where there is space between the pixels where no imaging occurs. A model of the simulation process is shown in Fig. 3.5. Image A is the original image (i), B is the test image after low-pass filtering (yi), and C is the ensemble of low-resolution images shifted with respect to each other. Filter H_2 performs the low pass filtering to obtain a lower resolution image and (G_1) , (G_2) , (G_3) , and (G_4) perform the operations to get the low resolution images.

To obtain the low-resolution images, a simple moving average (MA) filter (see H_2 below) is applied to the known high resolution image. The filter is essentially a low-pass filter that blurs the image onto larger pixels. The MA filter averages the adjacent pixels in a square neighbourhood around each of the high-resolution pixels in question. One such filter is as follows:

$$\mathbf{H}_1 = \frac{1}{16} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad (3.5)$$

For our simulation, however, a weighted moving average, WMA, filter (H_2) gives the closest neighbour pixels more significance. This more accurately reflects the situation on a sensor array. The individual sensors or pixels don't have a flat point spread function in the space domain but rather one that is weighted about its center[Gro86]. The WMA filter is as follows:

$$\mathbf{H}_2 = \frac{1}{36} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & 2 \\ 2 & 4 & 4 & 2 \\ 1 & 2 & 2 & 1 \end{pmatrix} \quad (3.6)$$

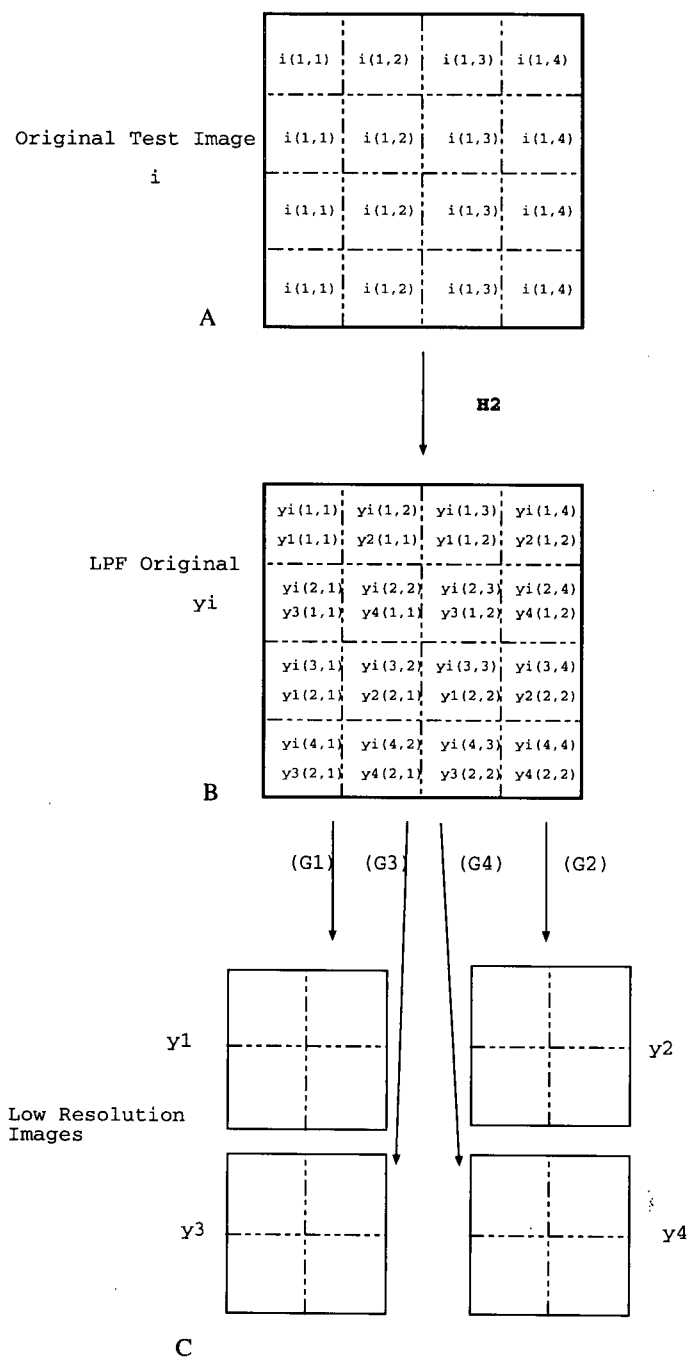


Figure 3.5: Simulation Model

The L low resolution images are formed by subsampling the multiplication of the above filter with the high-resolution image. If the high resolution image is $N \times N$, the low-resolution image is $K \times K$, where $K = \frac{N}{\sqrt{L}}$. In our simulation, $N = 256$ and $L = 16$ such that $K = 64$ and the low-resolution images have pixels $\sqrt{L} = 4$ times as large. One of the low-resolution frames is shown in Figure 3.6.



Figure 3.6: Low Resolution 64x64x8bit Lena Image Frame (obtained by applying H2 to Fig. 1)

In order to complete the simulated process of how the sensor array misses certain parts of the image (see Fig. 3.2) additional subsampling operations $(G_1), (G_2), (G_3)$, and (G_4) are applied to get the final low resolution images. These low-res images are considered the highest resolution images that our imaging system can achieve on its own. From here we work backwards to recover the resolution.



Figure 3.7: High Resolution Reconstructed Image (no deblurring applied)

With the low-resolution images created, the high resolution image, i , (A in Fig. 3.5) can be artificially “reconstructed” by merging the low resolution images (y_1, y_2, y_3, y_4 in Fig. 3.5). Interpolation is used to improve the results as described earlier [Gro86]. Figure 3.7 shows the reconstructed image using a bi-linear interpolation method after 2 dimensional filtering of the image with H_2 and meshing the images to a higher grid. As expected the image appears blurred because in essence it is just the low-pass version of the reference image i . The example we have given is trivial but demonstrates how some high frequency content is lost and just how dramatic the improvement in resolution can be from Fig. 3.6 to Fig. 3.7. A conventional deblurring restoration filter (such as a Wiener filter) could be applied to the image. An alternative method for creating low-resolution shifted frames is suggested in

[KBV90]. It involves interpolating the original image by expanding it with a zero-padded DFT. The higher resolution image can then be sub-sampled and noise added to get the shifted low-resolution frames. The advantage of this technique is that an arbitrarily large number of interframe samples can be obtained by increasing the number of points in the DFT.

Since the low-resolution blurring function, H_2 , is a low-pass filter, some of the high frequency information is lost. This means that the reconstruction problem cannot be uniquely determined and minimization of error techniques converge to several solutions. Our MA filters, H_1 and H_2 , have a non-uniqueness problem because they are singular matrices. No inverse filter exists (because of the singularity), a fact that mandates the use multiple images shifted from each other to recover the resolution.

The above simulation was then repeated but with white Gaussian noise added to each low-resolution image. The noise was then removed using an appropriate Wiener adaptive filter and the high-resolution image reconstructed by interpolation. The resulting image is shown in Figure 3.8. The high-resolution information is visible but the image suffers from greater Gaussian blur. The result is worse because of the added noise.

Registration Problem

In our simulation above, we directly controlled what the sub-pixel shifts were. Suppose, however, that the shifts between frames were not known. With knowledge or an estimate of the ideal image, a correlation analysis can be performed to find the most likely shift.

Two of the low resolution images are expanded by zero padding between samples to the size of the high-resolution original image. The two-dimensional correlation

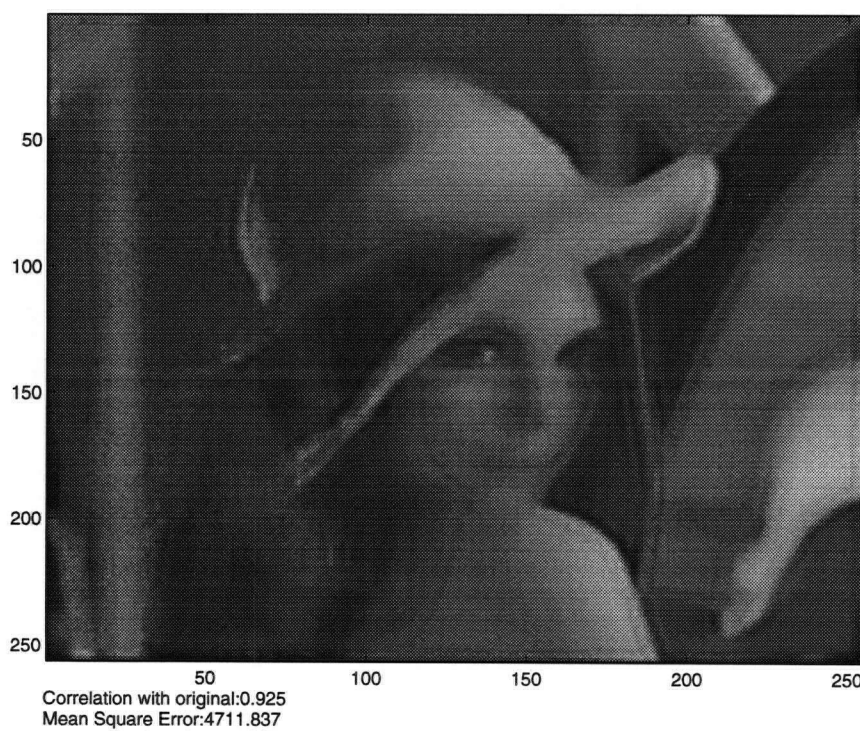


Figure 3.8: Wiener Filtered 256x256x8bit Lena Test Image (from Noisy Frames)

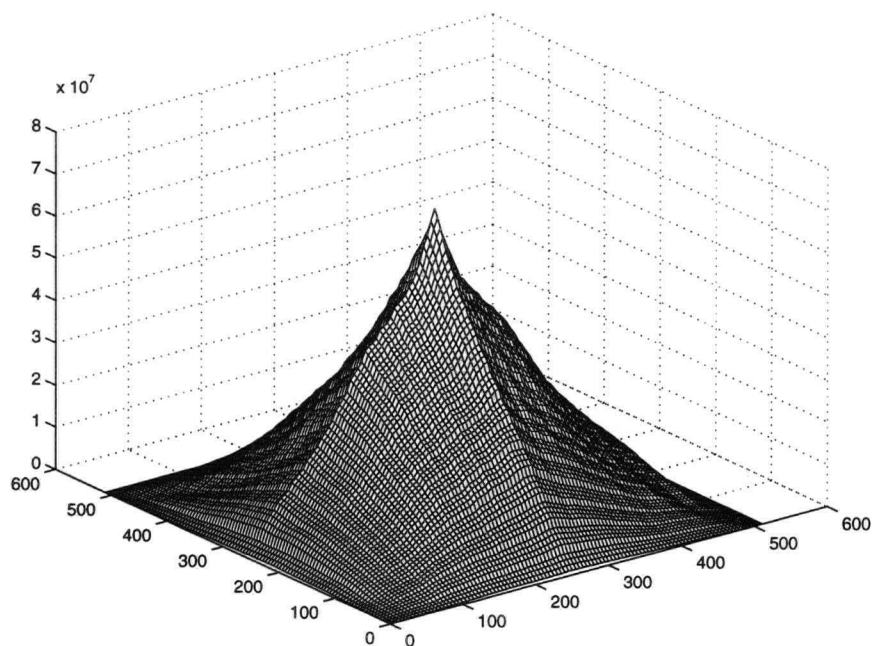


Figure 3.9: Image Correlation between Frames

of each image with the original is then performed and plotted. The shift is found from the maximum of the correlation. The results obtained were consistent with the actual shifts given to the two frames. A plot of the correlation between the tenth low-resolution zero-padded image (shifted 2 pixels in j direction, 1 pixel in k direction) and the original image is shown in Figure 3.9. The maximum of this image was found at (254,255), shifted (-2,-1) from the reference frame at (256,256), which agrees with the shift initially given to it. This simple example shows one way how image registration can be achieved.

3.3 Final Remarks on Super-resolution Problem

When there are multiple low-resolution images at subpixel displacements it becomes possible to upsample and recover higher resolution information. It is apparent that the task becomes much easier with more information about the imaging process and the ideal image. It is also highly desirable to know the exact shift of each frame and make the shifts at sub-pixel levels. In aircraft radar and other difficult to control imaging systems, this is not always possible. If however, a device existed to control the imaging shifts exactly in a controlled manner, high-resolution reconstruction could be achieved without much effort. The whole process is optimized by combining the restoration and interpolation steps when there is no registration problem.

It is clear from the procedure that the algorithm works best when there is much blank space in between pixel PSFs in the sensor array. This does not mean that the pixels on the CCD should be spaced further apart (because this would destroy the response of the CCD on its own), but that the individual detectors (pixels) should have a narrow PSF in the spatial domain.

Conventional high-pass filtering or deblurring techniques should be applied to the final blurred high-resolution reconstruction. The resulting image resolution can thus surpass that possible with the best available sensors and optics. Further, cheaper optics and sensors can be used to produce higher resolution images by taking multiple frames. The main disadvantage is the longer processing time.

Chapter 4

A Sub-Pixel Resolution System

Achieving sub-pixel displacement can be done in one of two ways. The first method is to optically shift the image using a movable refracting element as part of the optics. This can be achieved simply using a prism, an etalon, or even a single refracting plate. The second method involves moving the sensor itself. The CCD could be mounted on a plate, which in turn could be displaced with actuators. The decision of which method to use is based on how easy it is to construct the actuators and how easy it is to set up the position sensing devices.

4.1 Description of VAP

A method of obtaining higher resolution images from video frames is outlined below. The method involves a CANON ES-750 camera with optical image stabilization. The image stabilization is done by driving a variable angle prism based on a vibration input and the current prism angle. The prism itself is used to deflect the image on the CCD surface. Our goal was to use the vari-angle prism (VAP) to

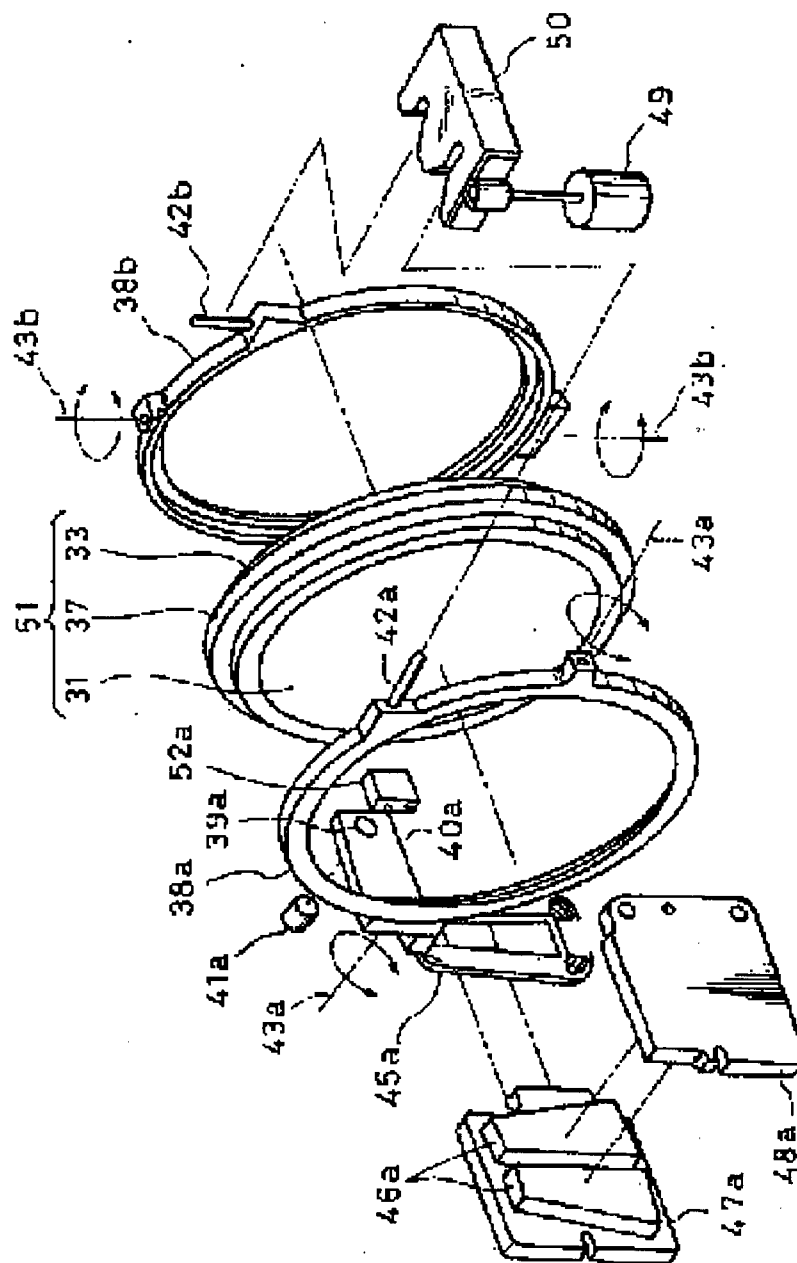


Figure 4.1: Complete VAP Assembly (US 5481394)

control the separation between several images and obtain a higher resolution from this.

The VAP was chosen (instead of an etalon or moving the CCD) to achieve the resolution shifts because it was a readily available device with pre-made actuators and position sensing. The goal was to expand the functional use of inexpensive camcorders equipped with means for image stabilization without going through the laborious process of building a system from the ground up.

The mechanical structure that we use to deflect the images at subpixel amounts is the Variable Angle Prism shown in Fig. 4.1. The important components are summarized below:

31 Front VAP Surface

37 O-ring

33 Back VAP Surface

46a Permanent Magnets

47a,48a Magnet Casing

45a Voice Coils

43a,b Pitch Axis, Yaw Axis

41a Infrared Diode (IRED)

52a Position Sensing Device (PSD)

39a Window Slit

38ab Front and Back Plates (for moving the vap)

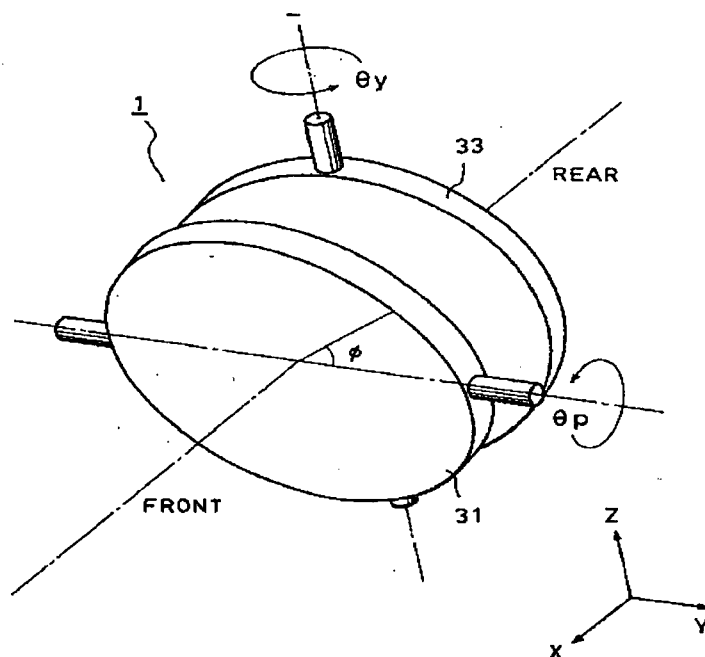


Figure 4.2: Front and Rear Plates of VAP (US pat. 5481394)

Items 42*ab*, 49, 50 are used to hold the VAP in a static position when the locking mechanism is on. Item 40*a* is just an arm attached to the front plate. Please see US patent 5481394 and US Patent 5623305 which describe the VAP as used as a 'Camera Shake Correcting Apparatus Having Damping Coils for Yaw and Pitch'.

The Vari-Angle Prism is formed from two glass surfaces separated and filled by a rubber silicone liquid with index of refraction, n . It is essentially a prism. The Front plate moves about a horizontal axis such that the image is shifted vertically on the CCD. The rear glass surface is confined to movements about a vertical axis, which allows for horizontal image displacement. The pitch angle (θ_p) of the front surface (31) can be changed, while the rear surface (33) controls the yaw (θ_y) as shown in Fig. 4.2.

The prism has the function of bending light rays. The light is refracted through the

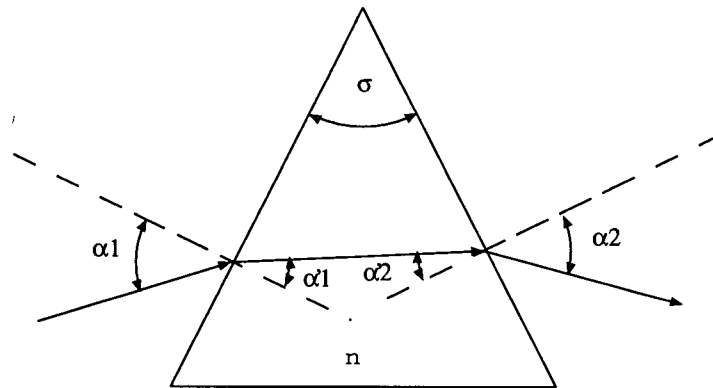


Figure 4.3: Large Prism

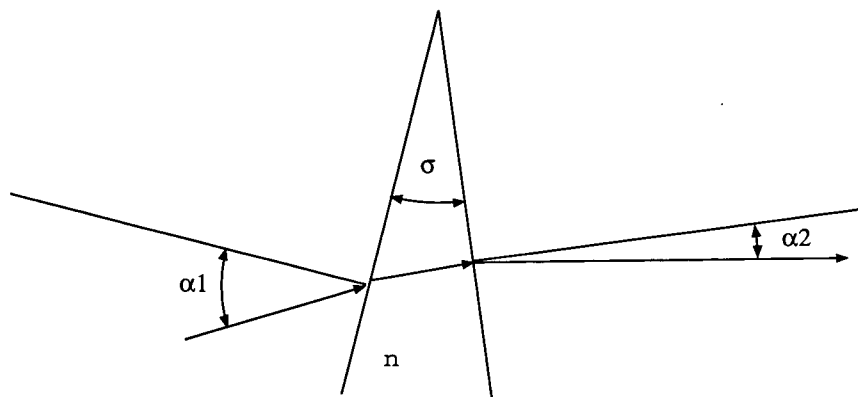


Figure 4.4: One Dimensional Structural View of VAP

prism to the other side. The following equations can be derived with reference to Fig. 4.3 using Snell's law of refraction and simple geometry:

$$\sin \alpha_1 = n \sin \alpha'_1 \quad (4.1)$$

$$\sin \alpha_2 = n \sin \alpha'_2 \quad (4.2)$$

$$\alpha'_1 + \alpha'_2 = \sigma \quad (4.3)$$

$$\epsilon = \alpha_1 + \alpha_2 - \sigma \quad (4.4)$$

where ϵ is the angle between the incident and outgoing light rays, n is the index of refraction and σ is the prism apex angle. When the variable apex angle is small as in Fig. 4.4, the equations above can be simplified to the equation:

$$\epsilon = (n - 1)\sigma \quad (4.5)$$

The deflection of an incoming ray is related to the index of refraction, n , of the prism and the prism apex angle. Equation 4.5 shows that the position of an image on the final CCD screen can be moved about with movement of the prism.

The VAP itself is physically circular and the geometry becomes more complicated when both sides of the VAP (31 and 33 in Fig. 4.2) are moved at the same time as shown in Fig.4.5. Suppose that the rear plate is rotated at θ'_p and the front plate at θ'_y to produce a movement ϕ . The inclination with respect to the angle ϕ for the rear plate is given by θ'_1 . From the geometry, the following can be written for θ'_1 :

$$\tan \theta'_1 = \sin \phi \tan \theta_p \quad (4.6)$$

Similarly, the following can be written for the inclination with respect to ϕ for the front plate:

$$\tan \theta'_2 = \cos \phi \tan \theta_y \quad (4.7)$$

FIG. 3(A)
PRIOR ART

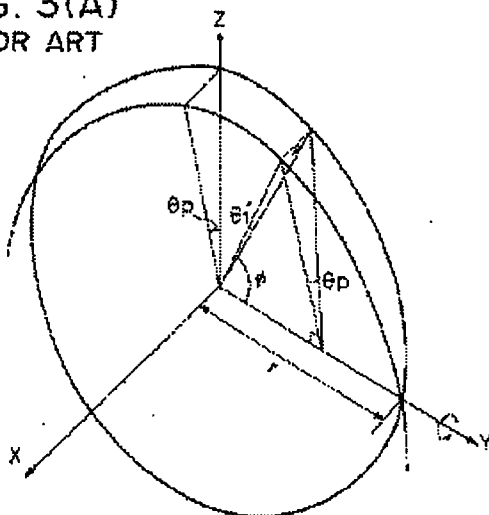


FIG. 3(B)
PRIOR ART

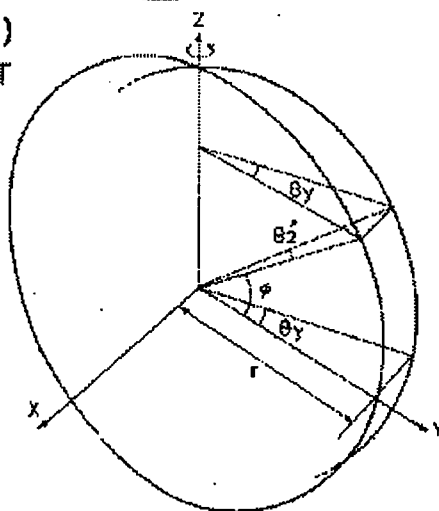


Figure 4.5: Geometry of VAP movement (US pat. 5481394)

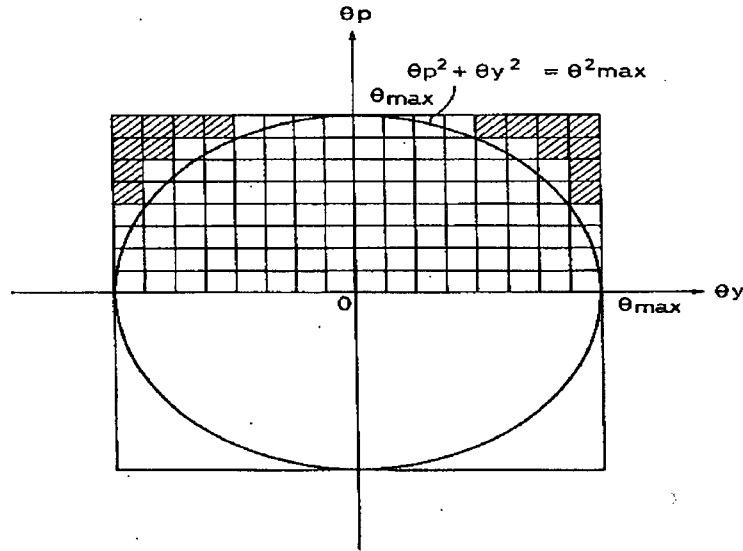


Figure 4.6: Restriction of VAP movement (US pat. 5481394)

The prism apex angle, σ , is the sum of the two above angles:

$$\sigma = \theta'_1 + \theta'_2 \quad (4.8)$$

Therefore σ is larger than both θ_y and θ_p . If we now define θ_{ymax} and θ_{pmax} as the maximum angles that the VAP can be rotated in the respective yaw and pitch directions, then it is clear that for oblique movements, the total prism angle σ can exceed the maximum allowable prism angle σ_{max} , such that the image overshoots the CCD. The solution is to restrict the movements in each direction as shown in Fig. 4.6 and the following equation:

$$\theta_p^2 + \theta_y^2 \leq \theta_{max}^2 \quad (4.9)$$

The VAP has 2 degrees of freedom allowing movement in two dimensions, which is all the movement with physical meaning for an image. All that needs to be achieved now is control of the VAP. This is described in the next section.

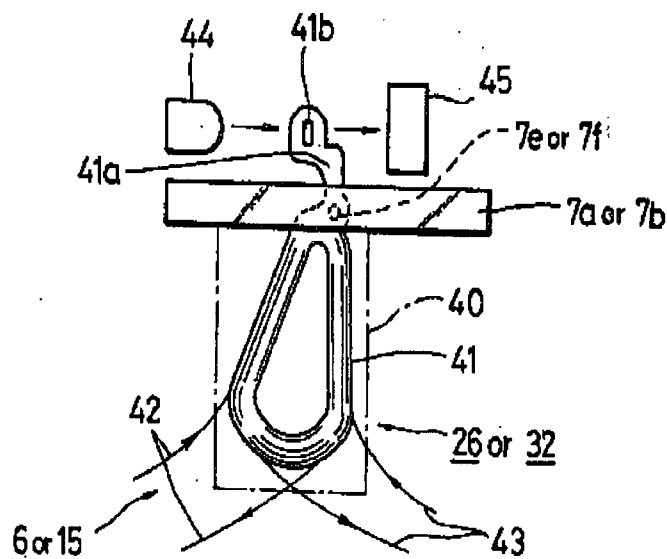


Figure 4.7: VAP actuator unit (US 5481394)

Actuation of the prism surfaces is done through magnetic voice-coils. It is essentially a linear DC motor system. The voice-coils are driven by an external voltage source and move within the permanent magnet stators. Each actuator has a characteristic winding resistance $[V/A]$, R . The actuator coils are wound with pancake wire. Fig. 4.7 and Fig. 4.8 show the actuator unit (with attached photosensors described later). Number 42 on Fig. 4.7 is the actuator winding coil. Numeral 43 on the same figure is an identical coil that will be discussed later. The labels are summarized below:

6 or 15 Coil Casing (see Fig. 4.14)

7ab Actuator Mounts

7ef Hole for Stem 53 or 54

26 or 32 Coil Casing (see Fig. 4.14)

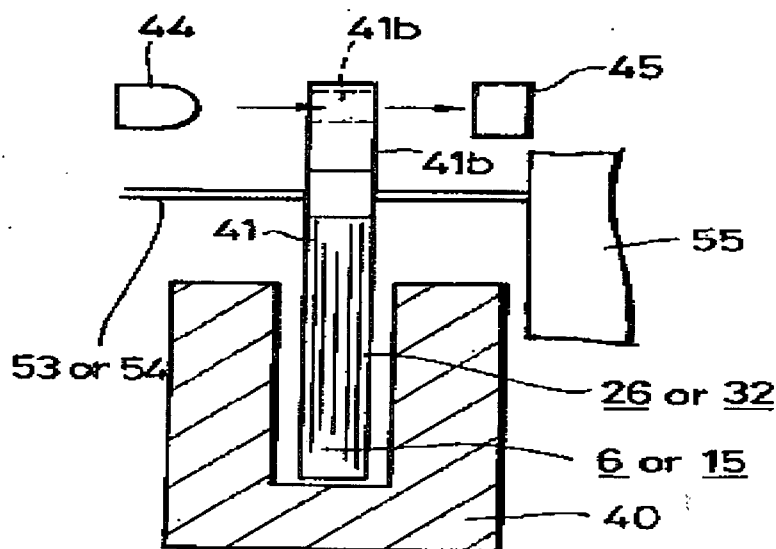


Figure 4.8: Side View of VAP Actuator (US 5623305)

40 Stator

41 Actuator Casing

41b Slit Window

44 Infrared Emitting Diode (IRED)

45 Position Sensing Device (PSD)

53 or 54 Stem Connecting Actuator to VAP Frame

55 Vap Frame

For position sensing, each side has an angle sensor, formed by an infrared photo-diode (IRED) and photo-sensor (PSD). Fig. 4.7 shows the position sensing in detail (44,41b,45). The IRED (44) and PSD (45) combination is fixed in place while a small window slit (41b) on the chassis moves with the VAP outer surfaces. This

moves the dot of light on the photo-sensor allowing the angle of the pitch and yaw to be measured.

We used the instrumentation amplifier circuit shown in Fig.4.9 to amplify the PSD output. INA 118 is an instrumentation amplifier (amplifying the small differential signal from the VAP PSD) set to a gain of 100 through the 560 ohm resistor. OPA27 is just a buffer amplifier. The circuit attached to the non-inverting terminal of OPA27 is a trimming circuit (made more stable by the current references (REF200)) used to tune the instrumentation amp INA118 through the buffer amp (OPA27). It was important to use shielded twisted pair wire to prevent noise from being introduced. V_A is a symbol notifying the voltage feed to the control circuit described later in Fig. 4.15.

4.2 Applied Super-Resolution

Achieving super-resolution in practice is a non-trivial problem. In all real world cases, the actual representation of an image is unknown. For moving camera systems, such as aircraft surveillance, controlling the shift between image frames is next to impossible. It is therefore necessary to consider the full registration problem. Our goal is to show a system that can take several images that are shifted by a known and repeatable amount and use these images to recover information that was lost in the optical process.

In practice, determining the sub-pixel shift is a difficult problem. A test target must be carefully chosen to calibrate the system properly. In addition, an appropriate method for locating the target must be outlined. With the above accomplished, a mapping of the system position to relative shift can be done.

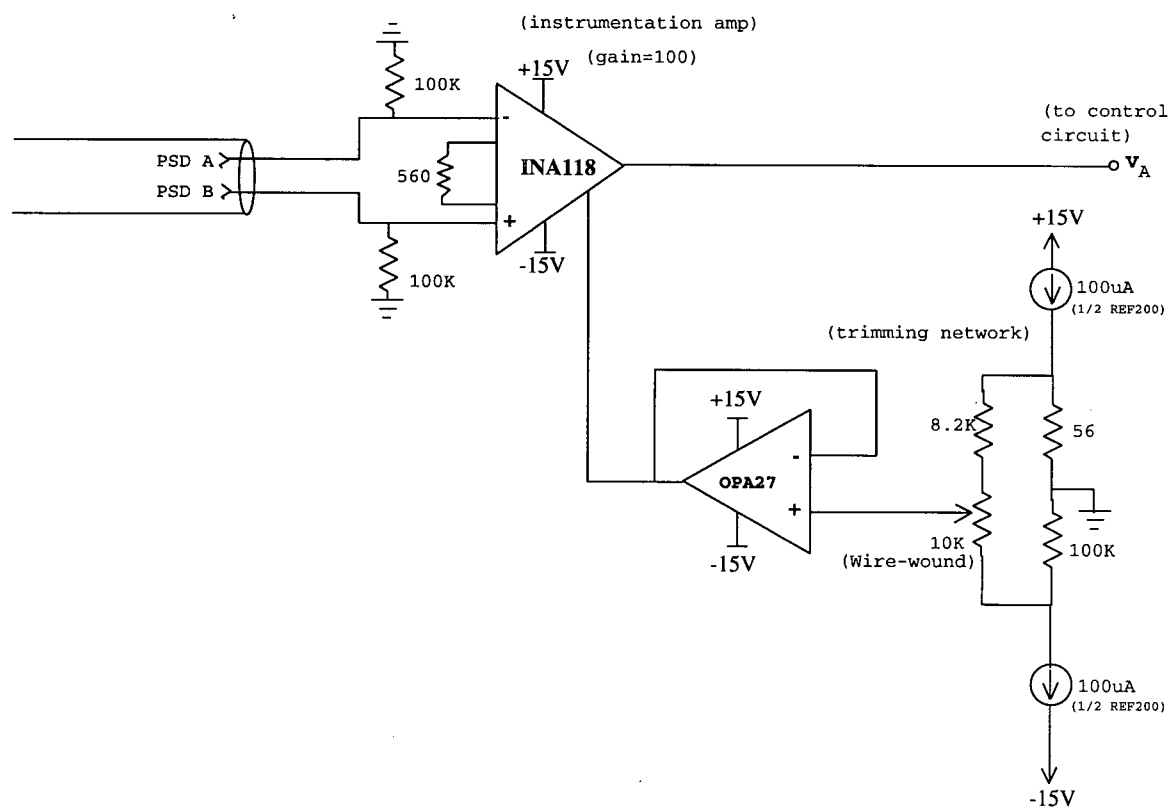


Figure 4.9: Our Front End Differential Position Amplifier

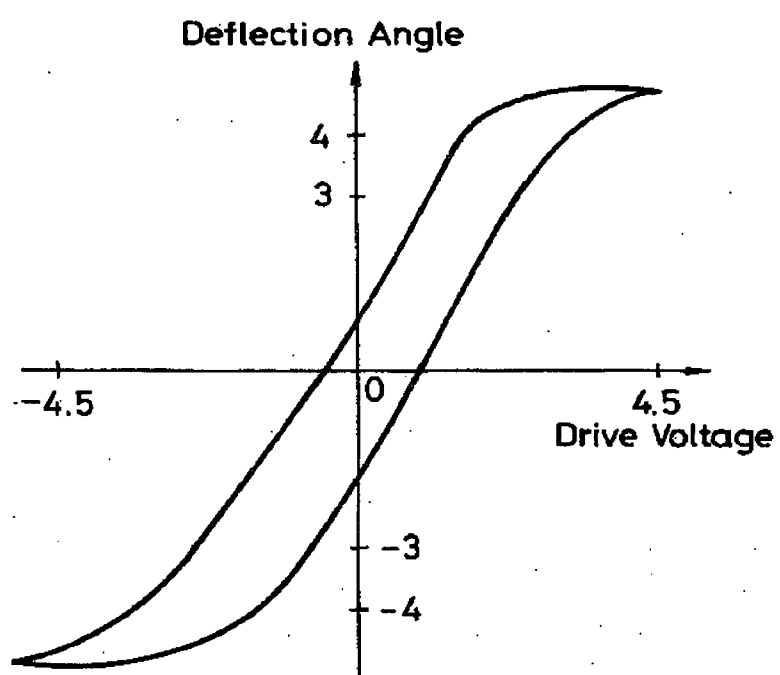


Figure 4.10: VAP Deflection Angle versus Drive Voltage (US 5623305)

4.2.1 Centroid Calibration

In order to calibrate the system, a good test target is required that can be distinguished from the environment. Issues such as external lighting, camera automatic gain control and external vibrations make the consistent determination of the centroid a problem.

In order to minimize the variations in the overhead light, it was found that a white target on black paper worked best. In addition, the camera automatic gain function adjusted the contrast less with this configuration than for a black target on white paper.

The target itself was a 1cm by 1cm square centered in the camera field of view. The camera zoom was set to the widest angle. Using different zooms causes changes in sensitivity to movement of the VAP. In fully zoomed mode, it is clear that small movements in the VAP creates larger displacements in the centroid across pixels because the target takes up more relative area on the sensor, which in turn translates to a larger area on the screen. Because of the significant changes in zoom and VAP temperature during different test conditions, it was found that calibration was necessary before every test.

Two approaches were used to find the centroid. The first was an iterative approach to examine the outer neighbourhood of a suspected spot. The second was to find the centroid of a specified small area of pixels. The difficulty with the first approach lies in the threshold value for a detected spot. On the border, the contrast effect of the camera automatic gain control causes the pixels to be lighter than the background. Using a fixed threshold in a iterative algorithm throws out information on the border. The automatic gain control of the camera causes unwanted effects no matter which algorithm is used, but the local area is less sensitive to it. The best solution

is to use a larger dot to make the boundary have less importance.

Lighting problems also hamper the consistent determination of the centroid. As the brightness and direction of the overhead light shifts, the centroid can move about. For the sub-pixel work we are concerned with, this can cause problems.

Because of the above problems of lighting and camera automatic gain control (AGC) as well as various vibration noise, a statistical approach was at first used to find the centroid. The centroid was found for several hundred image frames.

A sinewave input was initially used to drive the VAP.

$$V_{in} = A \sin \omega t \quad (4.10)$$

$$y_{out} = kV_{in} \quad (4.11)$$

$$y_{rms} = kA_{rms} \quad (4.12)$$

So experimentally, $y_{rms} = \frac{\sum y^2}{N}$, where N is the number of points taken. Obtaining data this way gets rid of the sensitivity to jiggle and averages out the noise effects.

The centroid was actually located using a threshold technique. The algorithm is as follows:

- Check if current pixel is within the threshold for target
- If not, scan through image until one is found
- Recursively check the neighbours of each pixel within the threshold
- When a hit is found (as part of the intended target), multiply X coordinate index by the pixel value
- Add the above value to a running sum for all hits (SUMXXB)

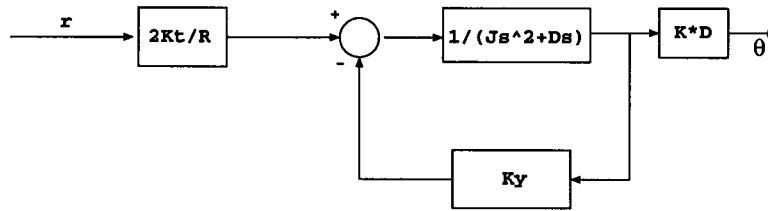


Figure 4.11: Open Loop representation of VAP

- Repeat above two steps for Y coordinates (SUMYXB)
- Add current pixel darkness value to running sum (SUMB)
- Repeat 5 previous steps until no more hits are found
- Calculate centroid location by division (SUMXXB/SUMB, SUMYXB/SUMB)

To eliminate false positives, the user was allowed to select from a number of possible hits. Once the desired centroid was selected, an image of the region in question was displayed on the screen. This made sure the right coordinates were chosen for the test target.

With the test target located, the above threshold technique or a localized area technique was used to calculate the centroid after each movement of the prism.

4.3 Control of the VAP

The VAP can be run open-loop, but doing so results in hysteresis in the displacement/voltage curve. The hysteresis is due to the magnetic nature of the voice coils and the permanent magnets.

Fig. 4.10 of US patent 5623305 shows the hysteresis for the VAP. This same patent also shows how the VAP can be modeled in terms of physical parameters. The

assembly has a particular torque constant, K_t , and inertia J . The silicone liquid inside the VAP has a viscosity resistance $D[g\frac{\text{cm}}{\text{sec}}]$ and the VAP itself has a spring constant, K_y . A block diagram for the VAP is shown in Fig. 4.11. From this diagram, the following equation can be written for the open loop gain:

$$G_{open}(s) = \frac{2K_tKD}{R} \frac{1}{Js^2 + Ds + K_y} \quad (4.13)$$

where R is the winding resistance in Ohms.

This can be simplified to a generalized second-order equation:

$$G_{open}(j\omega) = \frac{K}{1 - (\frac{\omega}{\omega_n})^2 + j2\xi(\frac{\omega}{\omega_n})} \quad (4.14)$$

where n is a normalization and ξ is an attenuation coefficient. Only the form of the equation above is important; we do not need an expression for ω_n or ξ and the patent does not provide one either. A simulation yields Bode diagrams (Fig. 4.12) showing phase and gain versus frequency.

The characteristic curves for the VAP reveal a pole at 100Hz. In addition, running the VAP open-loop prevents control over small angles due to the VAP non-linearities and the hysteresis curve. The liquid inside the VAP changes mechanical (viscosity) and optical properties (index of refraction) significantly with small changes in temperature. The strong temperature dependence of the device as well as the need to have control over small angles requires a more elaborate control system.

In order to improve results, the position of the PSD is fed back. This alleviates the problem of the VAP non-linearity and allows smooth control of the variable angle prism. The phase margin is increased, which reduces the effect of the VAP mechanical properties. A block diagram showing the VAP control is shown in Fig. 4.13 with $G(s)$ equal to the open-loop gain above. The actual circuit we designed and used to feedback the position is shown in Figure 4.15. Symbol V_A is the voltage

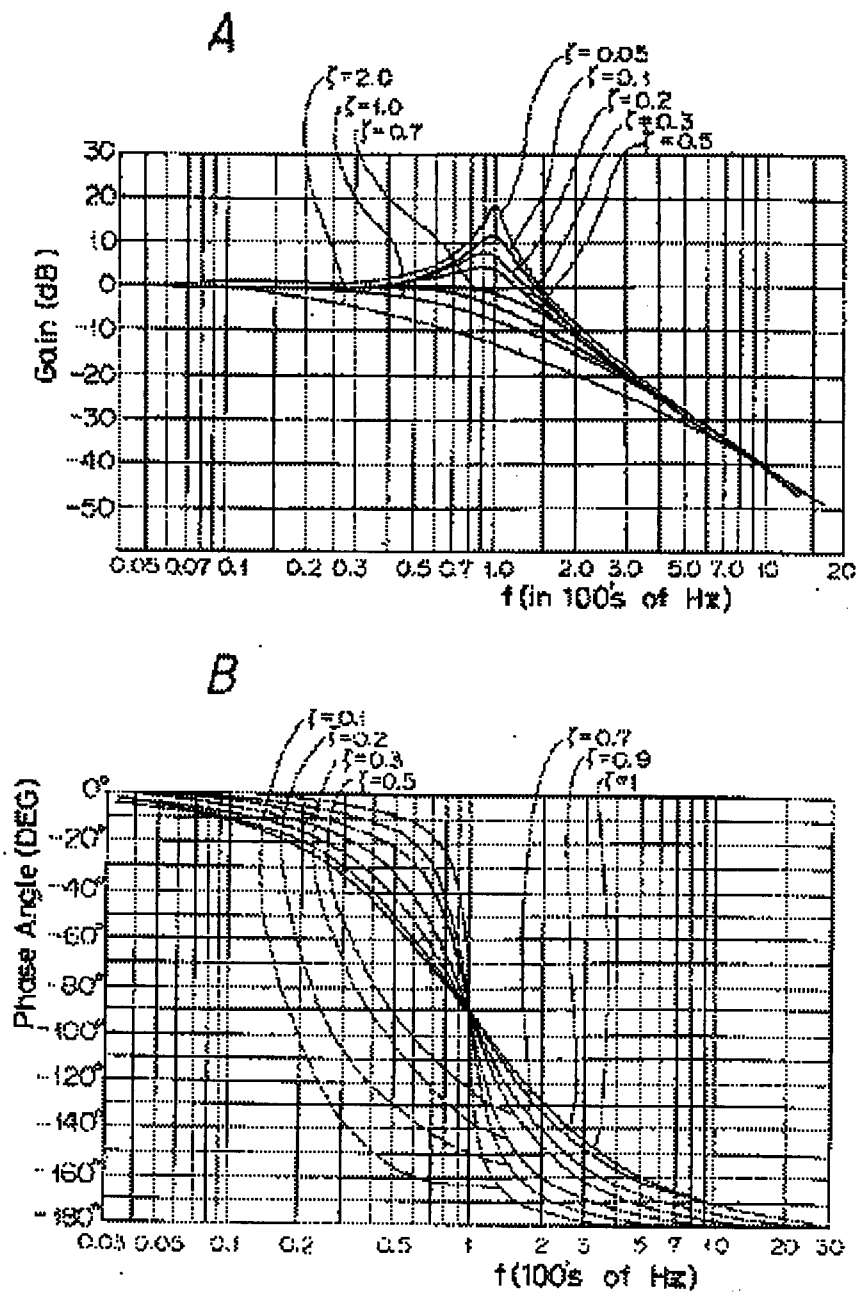


Figure 4.12: VAP Bode Plots (US 5623305)

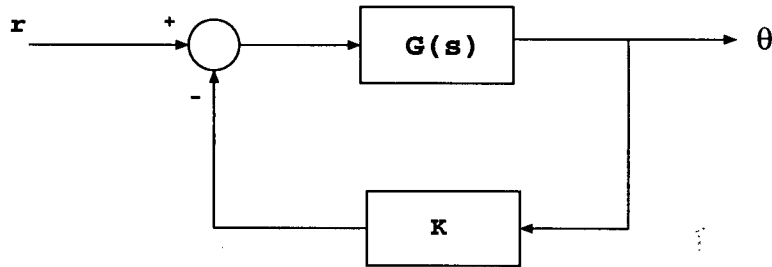


Figure 4.13: Closed Loop representation of VAP

from the front end instrumentation amp (Fig. 4.9). The first operational amplifier TL071 is a high DC gain amplifier tuned through the 100nF capacitor and 39Kohm resistor to roll off at a frequency of 20Hz. Without this rolloff, the parasitic pole at 100Hz causes the VAP to oscillate violently. The feedback loop is composed of the amplifier LF356 and the current mirror of 2N3904 and 2N3906. The mirror is there to get a higher current drive. In addition, there is a couple of 2.2Volt Zener diodes tied as shown in front of the current mirror to prevent the VAP actuator from being overdriven through V_{out} .

The patent mentioned earlier is concerned with correction for image shake. It describes control based on a detected angular velocity from an on-camera gyroscope. All the signals are processed digitally through a microcontroller. In addition, the actuators have a second coil (Fig. 4.14–42,43) on them (physically separated from the first coil by a barrier) to pick up the velocity of the actuator (Fig. 4.14–56). The second coils are called damping or dump coils and are not driven. Their movement through the magnetic field inside the stator induces a voltage. The AC stability is increased by using the second damping coil, which introduces a derivative feedback term to increase the phase margin. In DC motor theory, the technique is known as velocity feedback.

As mentioned earlier, the characteristic pole is at 100Hz with only the position being

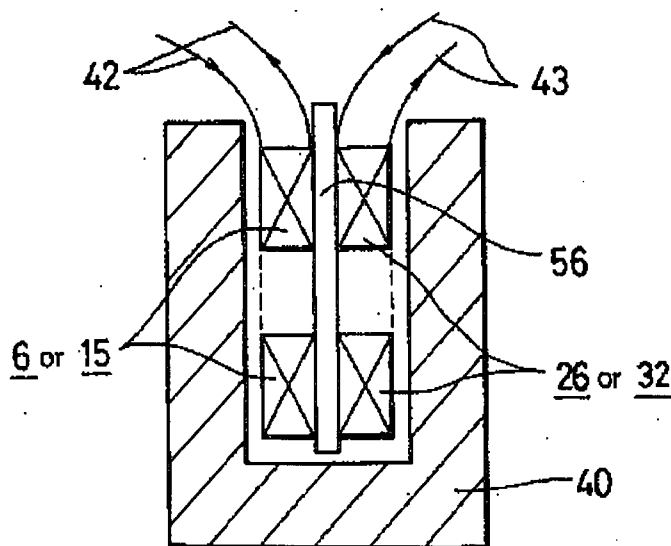


Figure 4.14: Dual Coil Wrapping of VAP Actuator (US 5623305)

fed back proportionally. Adding the derivative feedback helps push the pole to a larger frequency, thus reducing the sensitivity of the system to external vibrations. For our purposes, however, the proportional feedback is sufficient because we are not yet trying to shift the images transparently in real time and are only working in one dimension. In order to stabilize both the vertical and horizontal directions of the VAP, velocity feedback must be used in the control.

4.4 Experimental Results using ES-750

4.4.1 Initial Tests

In essence, there are two resolution problems that we are concerned with. The first is with the CCD surface itself. An image at the CCD surface is sampled by the individual pixels, which is then resampled by the frame grabber at the computer.

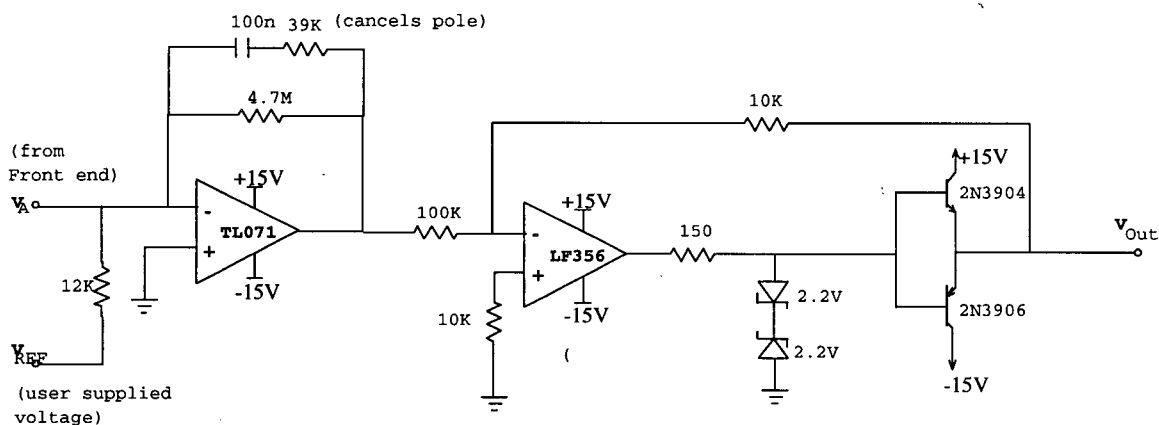


Figure 4.15: Our Control Circuit

The digitized frame suffers loss of detail from quantization effects. The second loss of resolution occurs at the screen. The screen has its own pixels which sub-sample the digitized image when it is displayed. The issue we are concerned with is whether a subpixel movement on the CCD is equal to a subpixel movement in the viewed image on the screen. Luckily, the mapping from the camera image surface to the screen is one to one. Therefore, by studying the subpixel movement on the CRT surface, it is possible to recover the resolution caused by the smoothing effect of the CCD, as discussed earlier.

4.4.2 Open-Loop Results

The first tests on the VAP involved driving it open-loop with a sine wave input and using Equation 4.12 to find the displacements on the CRT. Fig. 4.16 shows the horizontal mean rms displacement for sinewave inputs of several different frequencies. Fig. 4.17 shows the same information but with the control system attached to the vertical (pitch) actuators of the VAP. The results in both cases are linear up to a frequency of 25Hz. At 30Hz, the plot is a flat line indicating no response to

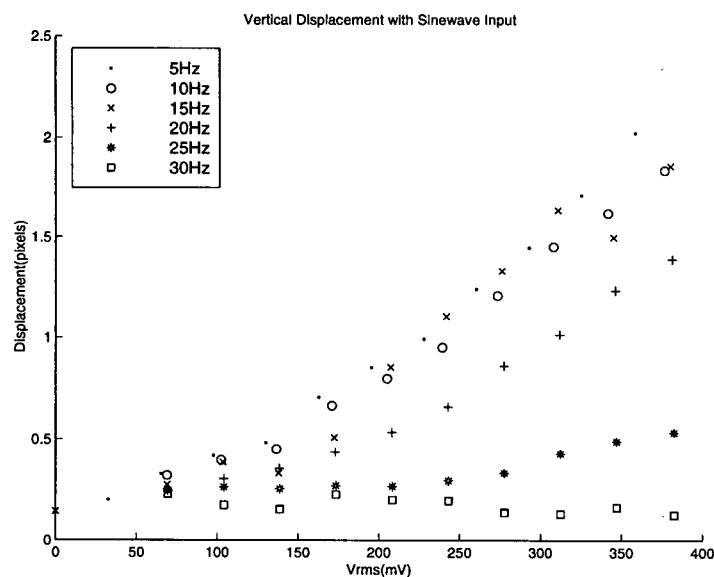


Figure 4.16: Vertical Displacement with Sinewave Input

the signal. This shows that the frequency response deteriorates somewhere around 25Hz.

Next, the frequency response of the VAP was measured to find where exactly the cut-off frequency is. This involved using a 0.5 Volt sinewave with an arbitrarily set frequency. The frequency was changed and the response measured. Fig. 4.18 is a plot of the results. The same experiment was repeated for the vertical VAP plate as shown in Fig. 4.19. The results for both sides coincided. The bandwidth of the device is situated around 20Hz.

A step function frequency response was also performed. The results appear in Fig. 4.20. The square wave response is useful because it shows what would happen if the voltage to the VAP were increased digitally at specific increments. The bandwidth is reduced to approximately 15Hz when a squarewave is used to move the Variable Angle Prism.

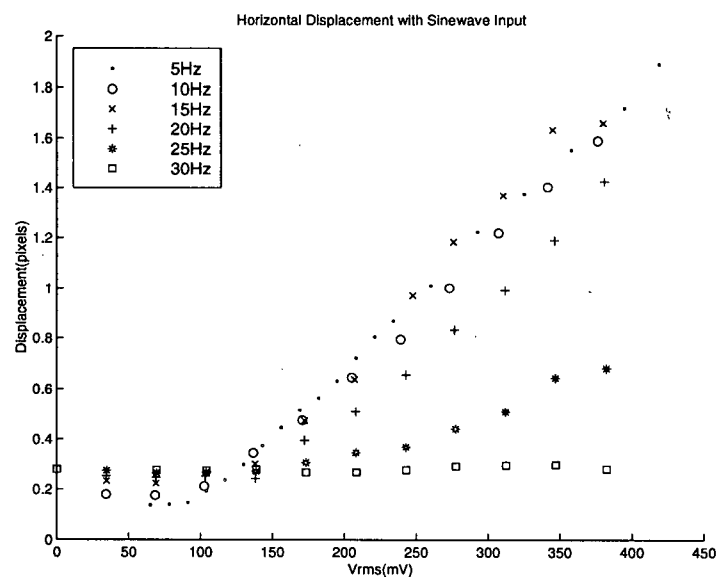


Figure 4.17: Horizontal Displacement with Sinewave Input

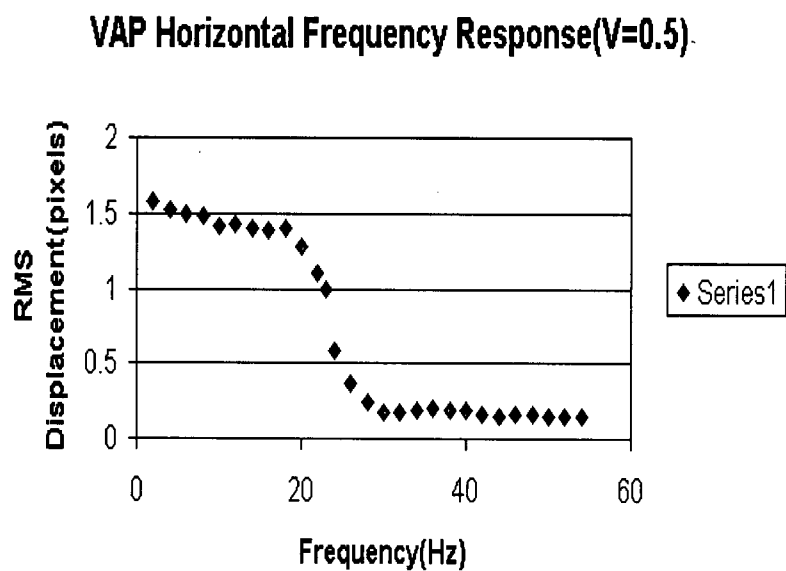


Figure 4.18: Horizontal Direction VAP Frequency Response (0.5 V Sinewave)

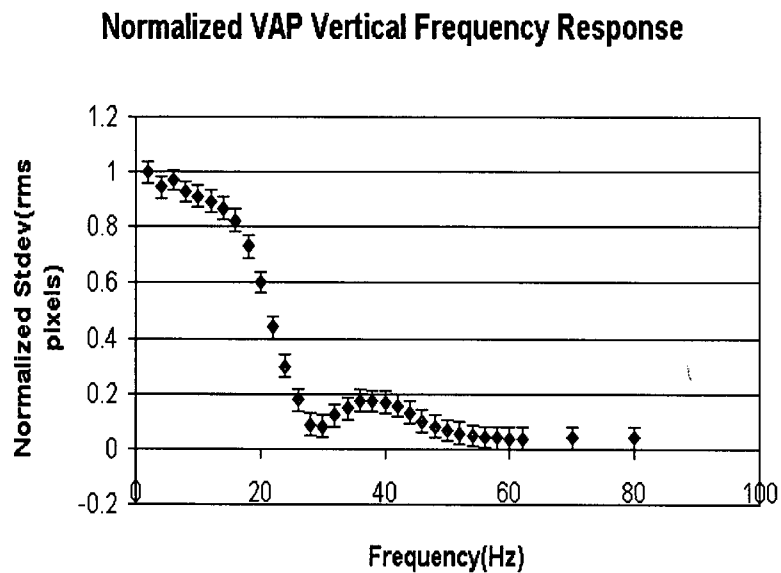


Figure 4.19: Vertical Direction VAP Frequency Response (0.5 V Sinewave)

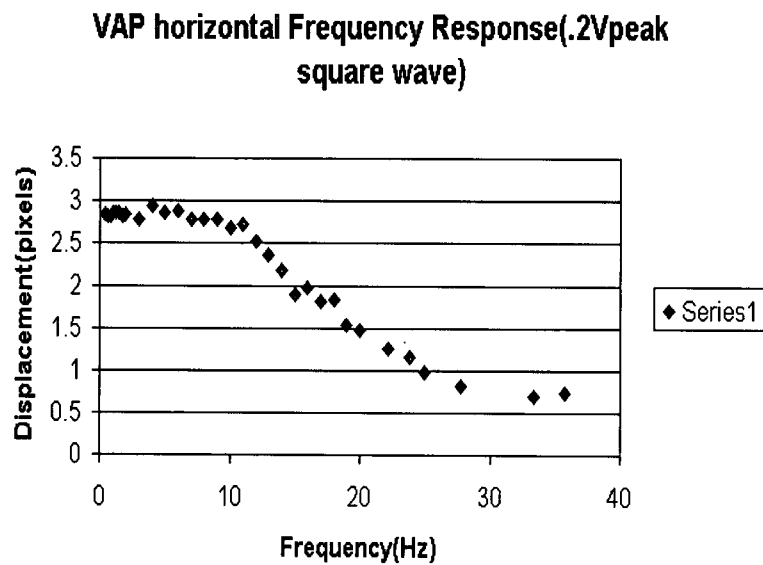


Figure 4.20: Square Wave Response

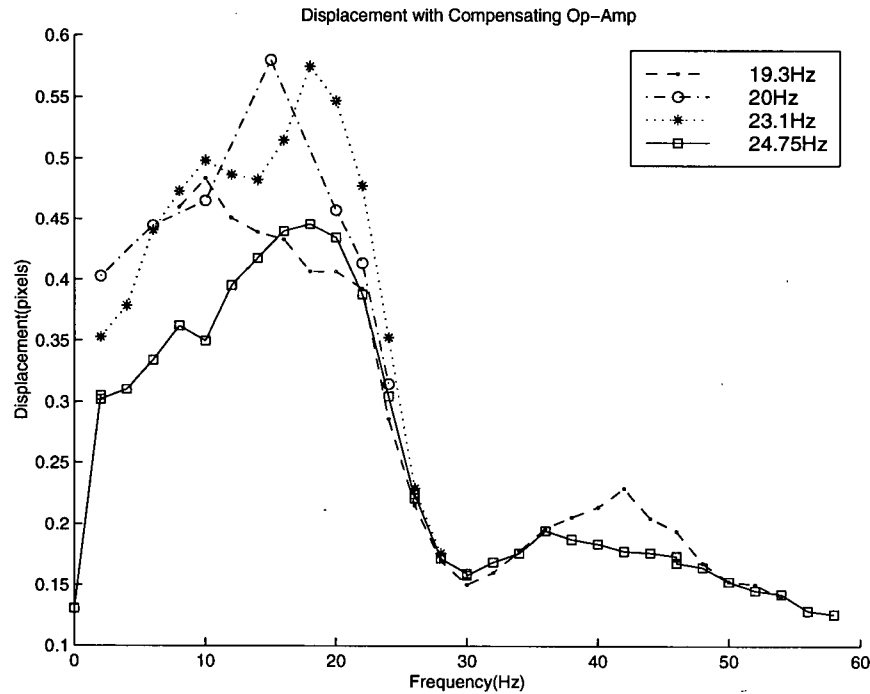


Figure 4.21: Displacement with Compensating Op Amp

With the frequency response determined, we set out to attempt pushing out the response using a compensator circuit with a variable frequency gain inverting op-amp to drive the circuit and attempt to increase the bandwidth. The feedback impedance was changed to adjust the location of the pole. The frequency response was then plotted with each different characteristic frequency, f_c , in Fig. 4.21. The driving op-amp had a gain of $\frac{Z_f}{Z_i}$, with Z_i the input resistor and Z_f the parallel combination of a resistor and capacitor to get the desired frequency f_c . It is apparent that the response deteriorates above 20Hz. The pole at 20Hz was not significantly reduced with the op-amp. This indicates that there are higher order effects involved with the VAP.

The op-amp was then run at DC. In one direction only, the response in Fig. 4.22 was found. The plot is linear in the middle section, but falls off at the edges as

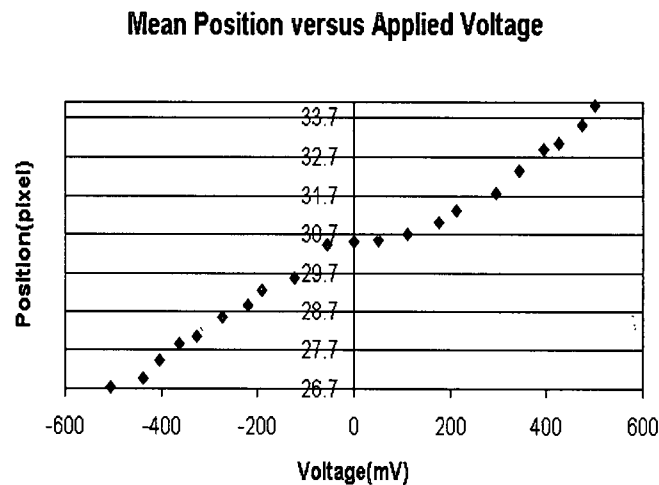


Figure 4.22: VAP run open loop in one direction

expected considering the limiting hysteresis of the actuators.

Unfortunately, the results in Fig. 4.22 were only attainable under strict conditions and were not repeatable between tests. When the VAP was moved back and forth, a hysteresis pattern was immediately apparent. Fig. 4.23 shows the situation for some representative movements of the VAP. The hysteresis is similar to Fig. 4.10 except the scale is different. We are using a different circuit to drive the VAP than that described in the patent so a direct comparison is not possible. The same can be said about the frequency response. The 100Hz pole described in the patent does exist as is clearly apparent when the control falls out of loop and starts oscillating, but it is not the same as the bandwidth that we determined the VAP to have using an external driving circuit.

In addition, the instability of the device was also a concern. A program was written to continuously sample the VAP over a period of 4 hours. The input was a constant magnitude and frequency square wave and the output was the displacement of the test target at each point as shown in Fig. 4.24. As mentioned earlier, the drift is

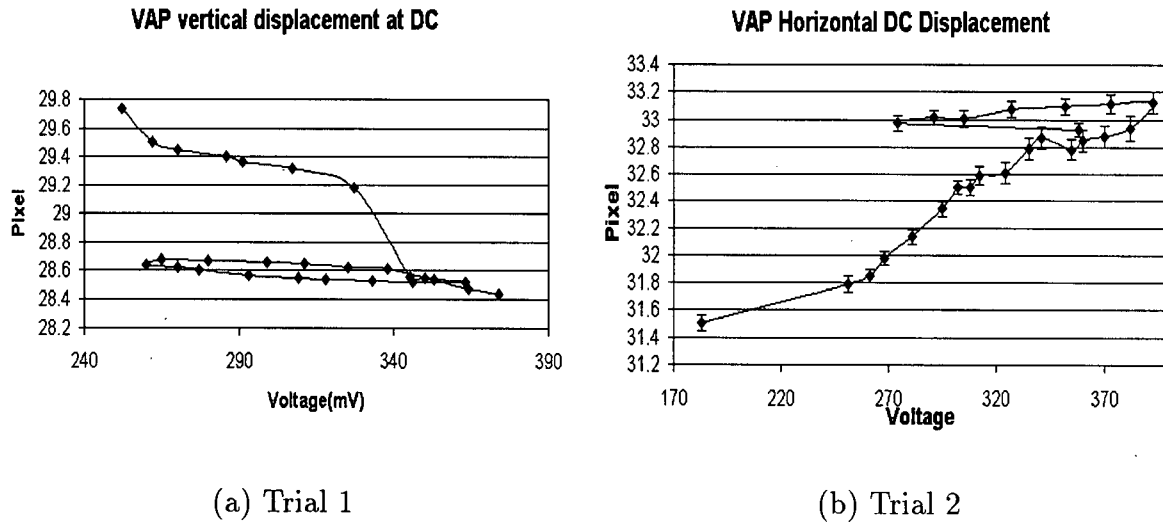


Figure 4.23: Hysteresis of VAP

due to several factors. The extreme points are most likely due to random vibration of the bench where the camera is mounted. In addition, a long term change in the displacement is evident and the separation between the high and low points becomes narrower. The narrowing of the displacement can be explained by a decrease in the index of refraction, n , caused by heating and subsequent expansion of the VAP fluid. The skew is explained by the VAP settling into an equilibrium state.

Feedback

The above graphs and plots indicate a clear need for feedback. The feedback network in Chapter 3 (Figures 4.9 and 4.15) was constructed and attached to the VAP position sensors and actuators directly.

Using feedback improves the situation considerably. The results show that subpixel

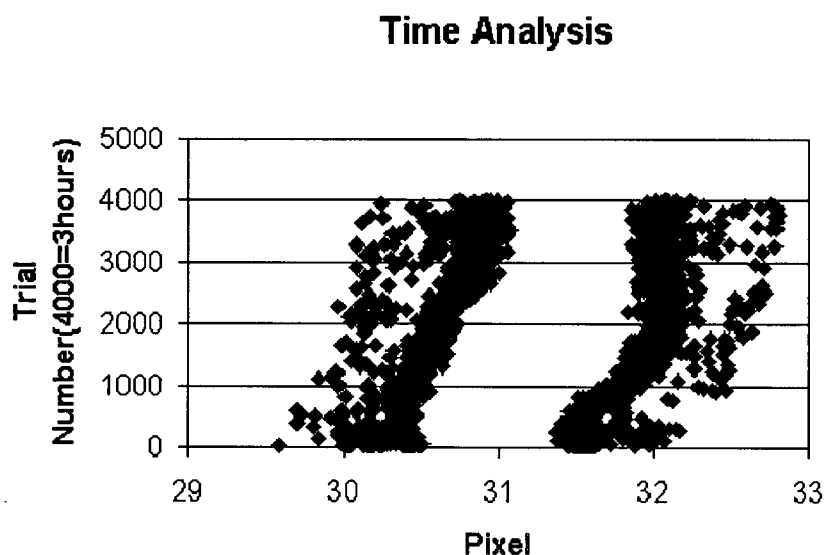
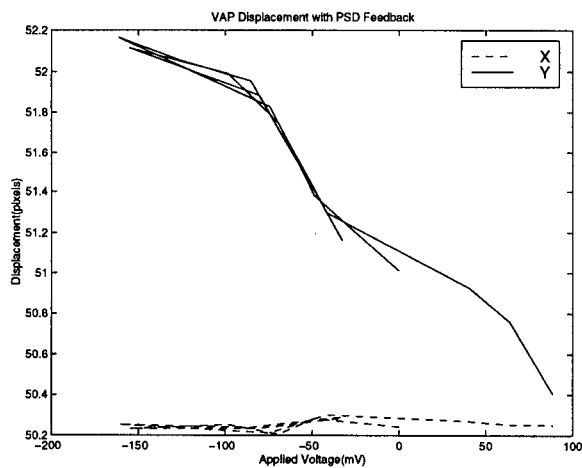


Figure 4.24: VAP Drift over Time

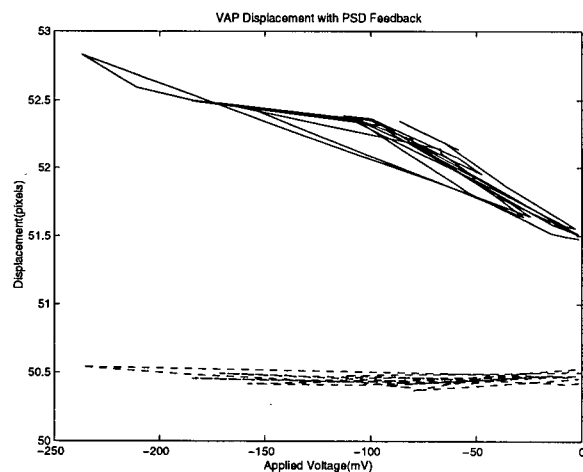
movements can be achieved using the Variable Angle Prism. Initially, the following results in Figs. 4.25a and 4.25b. There is an obvious knee to the graphs. The knee is not due to hysteresis but rather to a spring loading effect in the actuator assembly. Over a smaller range with a single pixel, linear results are obtained as shown in Fig. 4.25c.

With the above linear VAP movement over a pixel range, it becomes easy to sample images displaced at subpixel amounts to each other. The procedure must be done quickly, however, because of the inevitable heating of the unit and subsequent drift. The above mentioned Figs. 4.25c and 4.25d do show some separation when the movement is reversed. An explanation would be the slippage of the window as discussed in the previous Chapter. For this reason, it is recommended that the samples be taken while the movement is in the same direction.

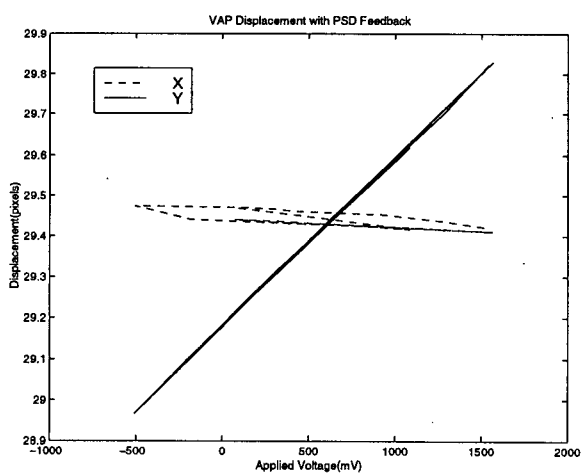
It should be noted that the feedback is stable for frequencies up to 100Hz. Also, a



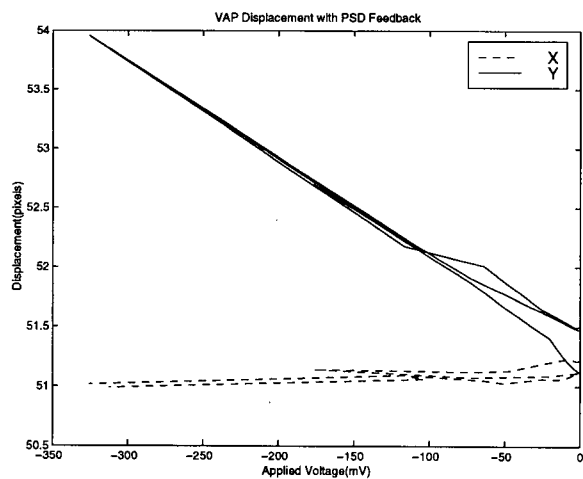
(a)



(b)



(c)



(d)

Figure 4.25: VAP Displacement versus Applied Voltage

| <i>Element</i> | <i>Group</i> | | | | |
|----------------|--------------|------|------|-------|-------|
| | 0 | 1 | 2 | 3 | 4 |
| 1 | 10.0 | 22.0 | 40.0 | 80.0 | 160.0 |
| 2 | 11.2 | 22.4 | 44.9 | 89.8 | n/a |
| 3 | 12.6 | 25.2 | 50.4 | 101.0 | n/a |
| 4 | 14.1 | 28.3 | 56.6 | 113.0 | n/a |
| 5 | 15.9 | 31.7 | 63.5 | 127.0 | n/a |
| 6 | 17.8 | 35.6 | 71.3 | 143.0 | n/a |

Table 4.1: USAF 1951 Resolution Test Chart (lppm)

large magnitude shake can bring the control out of the stable range. In addition, if the bench suffers large vibrations then the VAP will oscillate at the 100Hz parasitic frequency. Such behavior was observed using an oscilloscope.

4.4.3 Sampling at Subpixel Intervals

In this section, actual images were taken with the VAP calibrated using the above methods. In order to step the VAP at equal intervals, a DAC was used to drive the position. For each trial, the VAP is calibrated and the graph of centroid displacement versus applied voltage used to find how much voltage was needed to get fractional increments. All the tests were done with a displacement of one quarter pixel in one dimension (the vertical dimension). The maximum resolution improvement from the four images displaced at quarter pixel amounts is a factor of 2.

The test image is the standard USAF 1951 resolution chart (Table 4.1). Four images were taken displaced from each other by a quarter pixel each.

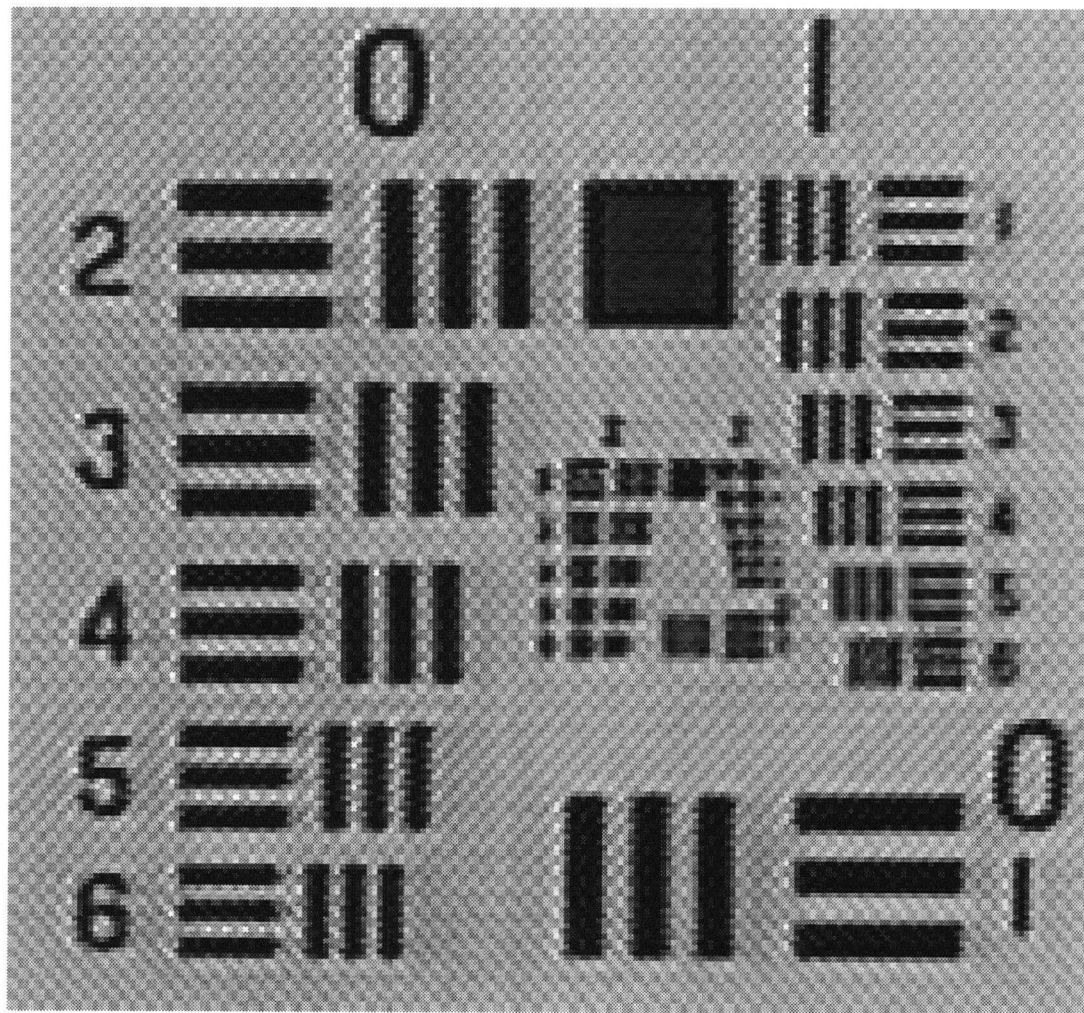


Figure 4.26: Zoom of Raw USAF 1951 Test Image

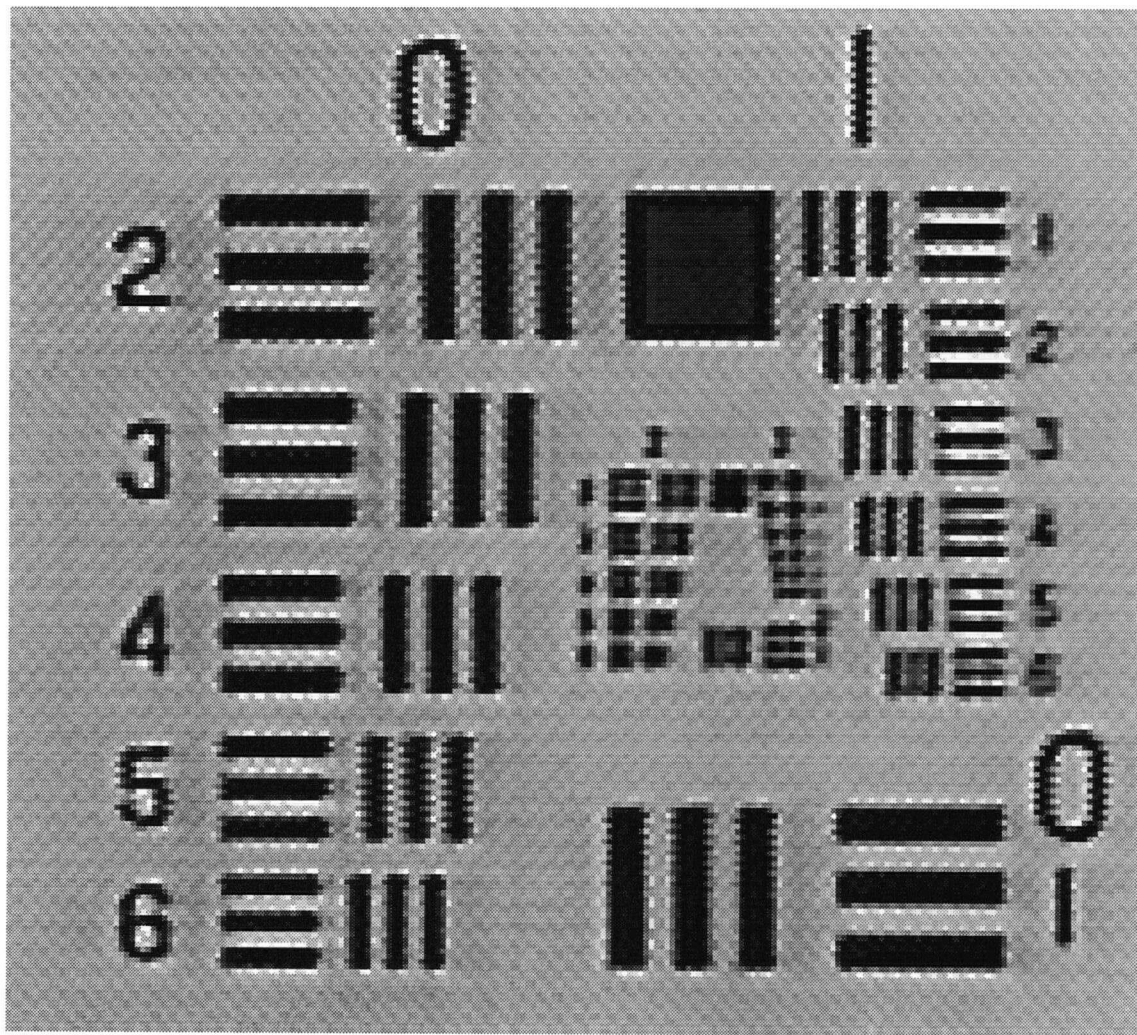


Figure 4.27: Image Reconstruction

It is difficult to tell where the resolution cuts out on the above image. Zooming in on the middle section shows where it deteriorates (Fig. 4.26). Examination of the zoomed image shows that the last clearly distinguishable group is Group 1 Element 5 giving an uncorrected resolution of 31.7 lppm. Group 1 is on the far right of the image.

The high resolution reconstruction was done by meshing the image and fitting it to the higher resolution grid. This was done using Matlab(TM) and the result is shown in Fig.4.27. Note that there are some jaggies present in the horizontal direction. These are due to the fact that the image was not perfectly aligned. The 2nd element of Group 2 can be distinguished in this image putting the resolution at 44.9 lppm. Unfortunately the numbers are not clear on the image. Group 2 is the leftmost group in the center of the picture (to the right of the large Group 0). Element 2 is at the top of this group and element 1 of Group 2 is in the bottom left corner of the two center groups (note that Element 1 of Group 0 is in the bottom left corner of the outer two groups (0 and 1)). The resulting resolution improvement is about 40 percent.

The theory mentioned earlier predicts the maximum resolution recovery from the above method at double the resolution of the low-res pictures. The experimental results were lower than this predicted maximum value for two main reasons. First, there was a delay of thirty minutes from when the VAP was calibrated to when the above reconstruction was done. This allowed for some drift in the VAP reference (from the PSD) because of temperature changes in the VAP fluid. Second, there is noise introduced through the image acquisition which corrupts the images and can not be removed through our sampling process. In addition, the actual theoretical limit is measured by the PSFs of the actual CCD pixels. There was insufficient information to determine what the form and spacing of these PSFs was. The doubling of

resolution figure is just an upper bounds for four samples displaced at quarter-pixel amounts.

Computing Requirements

With four 480 by 640 pictures, the above reconstruction took 4294833 flops in Matlab(TM). The number of flops for reconstruction, using L images each of size m by n , is on the order of magnitude $O(30mn + 2 * L * m * n)$. This maybe too slow for implementation in real time. For real time operation, a faster processor or multiple processors may be required. In addition, the picture size might have to be reduced.

4.5 Issues and Problems

The patent literature for Canon's proprietary VAP technology is geared towards image stabilization. They have constructed the device for rapid correction of high frequency camera shake. It incorporates a gyroscope and microprocessor for very fast correction. The microprocessor allows digital control and filters which makes the calibration of the system much easier.

We effectively bypassed the microcontroller used by Canon. We do not need gyroscope inputs because it is assumed that the camera is held stable. This cannot be the case, however, for a hand-held camcorder. In bypassing the microcontroller, we used analog circuitry for ease of initial measurements. Unfortunately, the VAP was not designed for long term DC stability. In addition, the instrumentation amps used to pick up the signals are subject to drift. The drift can be minimized, but since we are concerned with very small inter-pixel shifts, the measurements must

be done rapidly. In a final system, continuous calibration will be required as well as digital control of the entire system. Examples of the drift were shown previously (Fig. 4.24).

The control is ultimately limited by the position sensing apparatus. Unfortunately, Canon does not publish the specifications for the infrared diodes and PSD sensor used in either the patent literature or the service manual. Modern PSDs have very good resolution and can track laser spots to an accuracy of $1\mu m$. In this device, however, the accuracy is clearly limited by the mechanical window. Typical infrared diodes have a broad spectrum with a viewing angle between 20 and 40 degrees. Without the mechanical window, the IRED would spray itself onto the PSD making measurements difficult to obtain. The mechanical window, however is fairly large on the order of $1 - 2mm$. Both the precision and accuracy of the angle sensor is therefore limited primarily by the mechanical window (Numeral 41b, Fig. 4.7).

Another problem with the VAP assembly is the crosstalk from one side to another. When the angle is changed in the yaw direction, there is movement in one of the pitch directions. The converse is also true. According to the diagram of the VAP assembly in Fig. 4.1, the back plate is not completely independent from the front. The axis of the opposite plate can rotate on movement of the other plate because nothing holds the plates static and the axes aren't perfectly centered. Therefore, the control must hold the one side when the opposite side is moved. In a microcontroller system as implemented for the actual VAP controller in the camera, this could be directly implemented in the microcontroller itself. Unfortunately, when we connected our controller to both sides at the same time, the VAP would initially stay stable until one side was moved. Afterwards, it would oscillate uncontrollably at the 100Hz parasitic frequency described in the patent literature. This is due to the noise and

crosstalk reducing the phase margin to an unacceptable level. A microcontroller based control is recommended to control both sides concurrently.

In a future incarnation of a VAP image displacement device for the purposes of sub-sampling to increase resolution, it is recommended that a new angle sensing device be constructed. Instead of an IRED, a laser diode could be mounted on the chassis where the existing window slit is. Also, more elaborate digital control using a microprocessor is required to improve stability and to get the images before drift occurs.

Chapter 5

Conclusions

This thesis has summarized the results of work in image processing. We have also given a simulation of how a higher resolution is achievable from several lower resolution frames.

A device has been described that displaces images. This device is called a Variable-Angle Prism (VAP). The displacement is linear with the pixel range.

We have also shown how to control the VAP to get images displaced at subpixels to each other. We then used these images to get a higher resolution in the one dimension. The extension to 2 dimensions is nontrivial but possible.

We have shown that the improvement in resolution from the above method is about 40 percent. This is less than the doubling of resolution that is theoretically possible using 4 images. Further improvement would require digital control with a microcontroller to pick out the images rapidly and precisely.

5.1 Future Work for Combined System

Ideally the system would incorporate both the Vari-Angle Prism and the hemispheric camera. The loss of resolution caused by the sampling of the CCD could be compensated quite nicely by dithering the images with the VAP. The main resolution problem in the imaging system is at the CCD. If conventional film were used, there would be much less resolution lost. Unfortunately, current CCDs have pixel sizes of 10-20 micrometers, much larger than the resolution achieved with a photographic film. The advantage of using CCDs over film is the continuous sampling ability and easy conversion into a standardized digital form.

Unfortunately, the VAP is currently too large in its present form to fit under the primary mirror. It could be placed in front of the mirror but this would complicate matters enormously. The demapping would have to take into account the fact that the rays pass through the VAP twice. A simulation would have to be done and a calibrated table look-up formed to determine the amount of shift. This is not an efficient solution and the preferred method would be to have a miniature VAP or prism deflection system made to fit the hemispheric camera.

After the VAP project was complete, we discovered that Sharp came up with a device for deflecting images with a single refracting element. This device did not exist when the research for this thesis was performed. The reader who wants to do more work in this area is strongly urged to check out their patent. It is titled 'Imaging Apparatus Having Improved Resolution due to Control of the Inclination Angle of a Refracting Plate in More than One Direction' (US patent 05637861). It maybe a better solution than the VAP because it is far more simple and much less sensitive to temperature.

In addition, the mirror has too much material inside obstructing the placement of

the CCD camera. The preferred mirrors need to be built using a mold technique so that they are very thin and strong. The cost is higher than for the aluminum ground mirrors that we used, but the bulk manufacturing prospects are great and subsequent mirrors are very cheap once the mold is built.

Another problem is that the particular camera that we used has a potentiometer on the surface that prevents proper mounting inside the primary to coincide with the focal length of the lens. In addition, the lens does not fit inside the mirror hole. A custom lens needs to be built to fit exactly inside the hole on the top of the primary mirror or preferably the hole needs to be increased or decreased to fit a stock lens of choice. This involves choosing the right clear aperture of lens along with a focal length that will project the image to the CCD surface correctly. It is my experience, however, that a standard CCD camera will not work with the apparatus. A CCD camera needs to be built from scratch using raw CCD sensor components. Such a task is not trivial.

Ideally, the camera would be digital. Our experiments with the camcorder used a frame grabber that resampled the NTSC format. The camera outputs the signal in analog form which is then chopped up by the ADC on the frame grabber. Using a digital camera would eliminate additional conversion loss and degradation by quantization noise.

Another big problem is the blind spot. Right now, the view can only be called substantially hemispheric. Perhaps in a later incarnation, the blind spot in the center of the scene could be reduced further by putting a lens system in a hole with a diameter the same size as the blind spot in the upper mirror (h_{b2} in Fig. 2.6). This would be projected onto the unused centre spot that currently exists on the CCD because of the geometry. If the image from beyond the secondary mirror is limited in scope to the same size as the blindspot, then it will not interfere with the

hemispheric data. Such a modification would allow more of the actual hemispheric scene to be imaged.

Future work with the VAP involves digitizing the control and connecting both front and back sides of the device together to shift the images in the horizontal and vertical directions and obtain 2 dimensional image resolution improvement. A microprocessor is required to coordinate the movements properly and assure loop stability.

Finally, we have mentioned the possibility of incorporating the two technologies of omnidirectional vision and super-resolution into a single device. The prospects are very good as they complement each other nicely. Many more experiments need to be done, however, to develop a full working model that is commercially viable.

5.2 Summary of Results

In this work we have shown a system for hemispheric or omnidirectional vision. It is compact and has a very large field of view of 360 degrees in one hemisphere. The geometry and issues concerning the camera have all been examined and a design proposed for minimizing some of the intrinsic optical problems. A resolution recovery method was also shown and built using a Variable Angle Prism to effect image shifts.

The hemispheric vision system above was in fact built and demonstrated to work. There are certain things that need to be done to improve the resolution of the images, however. These include, but are not limited to, making a custom digital CCD camera with a large number of pixels, rebuilding the mirrors with an injection moulding process so that they are very thin (one thousandth of an inch thick is possible), building custom optical mounts to hold the mirrors in place, and finally

choosing a new lens matched to the custom CCD and primary mirror aperture. Much work needs to be done before the system can be marketed commercially. The market for such a device has not been determined, but it may grow exponentially as knowledge of omnidirectional sensors and their applications is determined.

Also in this thesis, a image displacement apparatus known as a Variable-Angle prism was outlined. It displaces rays of light and can be used for sub-pixel movements. Using several images displaced from each other at sub-pixel amounts, a higher resolution can be obtained for the final image of the original scene. The VAP was used successfully to obtain resolution improvement in one dimension. Subpixel displacements were also confirmed using the device and our custom control hardware. Super-resolution was demonstrated successfully.

Bibliography

- [Gro86] D. Gross. Super Resolution from Sub-pixel Shifted Pictures. *Master's thesis*, Tel-Aviv University, October 1986.
- [H⁺86] E. Hall et al. Omnidirectional Viewing Using a Fish Eye Lens. In *Optics, Illumination, and Image Sensing for Machine Vision*, volume Vol. 728, pages pp. 250–256, Chicago, 1986. SPIE.
- [IP90] Michal Irani and Shmuel Peleg. Super Resolution From Image Sequences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 38(1):pp.115–120, January 1990.
- [IP91] M. Irani and S. Peleg. Improving Resolution by Image Registration. *CVGIP: Graphical Models and Image Processing*, Vol. 53:pp.231–239, May 1991.
- [Jai88] A.K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, New York, 1988.
- [JR84] L.S. Joice and W.L. Root. Precision Bounds in Superresolution Processing. *Journal Optical Society of America*, pages pp149–168, Feb 1984.

- [KA96] A. Krishnan and N. Ahuja. Panoramic Image Acquisition. *Proceedings of IEEE Conference on Computer Vision and Image Acquisition*, pages pp.379–384, June 1996.
- [KBV90] S. P. Kim, N. K. Bose, and H. M. Valenzuela. Recursive Reconstruction of High Resolution Image From Noisy Undersampled Multiframe. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 38(6):pp1013–1027, June 1990.
- [Kor88] Dietrich Korsch. *Reflective Optics*. Academic Press, INC., Boston, 1988.
- [KS93] S. P. Kim and Wen-Yu Su. Subpixel Accuracy Image Registration by Spectrum Cancellation. *IEEE Procs.*, Vol. V(6):pp153–156, 1993.
- [Luk66] W. Lukosz. Optical Systems with Resolving Power Exceeding the Classical Limit. *Journal Optical Society of America*, 56:pp. 1463–1472, November 1966.
- [Luk67] W. Lukosz. Optical Systems with Resolving Power Exceeding the Classical Limit, II. *Journal Optical Society of America*, 57:pp. 932–941, July 1967.
- [MS88] M.S. Mort and M.D. Srinath. Maximum Likelihood Image Registration with Subpixel Accuracy. *Applications of Digital Image Processing XI, SPIE*, Vol. 974:pp38–45, 1988.
- [Nal96] V. Nalwa. A True Omnidirectional Viewer. *Technical report, Bell Laboratories, Holmdel, NJ 07733, USA*, February 1996.
- [Nay96] Shree K. Nayar. Catadioptric Omnidirectional Cameras. *Technical Report, Dept. of Computer Science, Columbia University*, October 1996.

- [Nor78] A. Nordbryhn. The Dynamic Sampling Effect with CCD Imagers. In *Applications of Elec. Imaging Systems*, volume 143, pages pp42–51. SPIE, 1978.
- [NPB98] Shree Nayar, Shmuel Peleg, and Rafi Brada. Omnidirectional Imaging Apparatus. *US Patent 5760826*, page 24, June 2 1998.
- [SA87] Ken D. Sauer and Jan P. Allebach. Iterative Reconstruction of Band-Limited Images from Nonuniformly Spaced Samples. *IEEE Transactions on Circuits and Systems*, Vol. 34(12):pp.1497–1506, December 1987.
- [SA94] S. Shaw and J.K. Aggarwal. A Simple Calibration Procedure for Fish-Eye(High Distortion) Lens Camera. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages pp. 3422–27. IEEE, 1994.
- [SO89] H. Stark and P. Oskoui. High-resolution Image Recovery from Image-plane Arrays, Using Convex Projections. *J. Opt. Soc. Am. A*, Vol. 6:pp1715–1726, 1989.
- [Spe84] W.P. Spence. *Engineering Graphics*. Prentice Hall, NJ, 1984.
- [SPS87] Danny Keren S. Peleg and L. Schweitzer. Improving Image Resolution Using Subpixel Motion. *Pattern Recognition Letters*, Vol. 5:pp223–226, March 1987.
- [ST90] M. Ibrahim Sezan and A. Murat Tekalp. Adaptive Image Restoration with Artifact Suppression Using the Theory of Convex Projections. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 38(1):pp.181–185, January 1990.

- [TG94] Brian Tom and Nikolas Galatsanos. Reconstruction of a High Resolution Image from Registration and Restoration of Low Resolution Images. *Image Processing, 1994 International Conference*, Vol. 3:pp.553–557, 1994.
- [TK94] B.C. Tom and A.K. Katsaggelos. Multi-Channel Image Identification and Restoration Using the Expectation-Maximization Algorithm. In *Applications of Digital Image Processing XVII*, volume Vol. 2298, San Diego, July 1994. SPIE.
- [TOS92] A.M. Tekalp, M.K. Ozkan, and M.I. Sezan. High-Resolution Image Reconstruction From Lower-Resolution Image Sequences and Space-Varying Image Restoration. *IEEE Proc. ICASSP-92*, Vol. 3:pp.169–172, 1992.
- [Y⁺95] K. Yamazawa et al. Obstacle Detection with Omnidirectional Image Sensor HyperOmni Vision. *IEEE Internal Conference on Robotics and Automation*, pages pp.1062–1067, May 1995.

Appendix A

Aberrations (see Chapter 2)

Traditional optical design concerns maximizing the quality of mapping a planar object onto a planar sensor. There are several types of aberrations, called the Seidel aberrations. These aberrations are third order effects and include the phenomena of spherical aberration, coma, astigmatism, field curvature, and distortion. They are represented by the coefficients A, B, C, $-(C+2D)$, and E respectively[Kor88]. Our two mirror system (Fig.2.6) is similar in form to a double parabolic collimator described by Korch[Kor88] and has the same Seidel aberration coefficients[Kor88]:

$$A = 0 \quad (\text{A.1})$$

$$B = -(\tau_1)^{-2}(c_1^2 + c_2^2) \quad (\text{A.2})$$

$$C = 2\tau_1(2t_1 + \tau_1 d_1)B + 2\left(\frac{d_1}{\Omega_1}(c_1^2 + c_2^2) + c_1 - c_2\right) \quad (\text{A.3})$$

$$D = \frac{1}{2}C - c_1 + c_2 \quad (\text{A.4})$$

$$E = -\tau_1^2 t_1(3t_1 + 2\tau_1 d_1)B + \tau_1(t_1 + \frac{\tau_1 d_1}{2})C + \tau_1 t_1 D - \tau_1^2 d_1(d_1 c_2^2 + c_1 + c_2) \quad (\text{A.5})$$

where $\tau_1 = 1 - \nu_1 t_1$, $\nu_1 = \frac{1}{s_1}$, c_1 and c_2 the vertex curvatures of the mirrors($c_1 = \frac{1}{2f_1}$),

t_1 the entrance pupil distance ($t_1 = f_1$ in our system), d_1 the mirror separation, s_1 the object distance and Ω_1 the ratio of the clear aperture of mirror 2 to mirror 1.

Because A is zero, there is no spherical aberration in our system. This comes by virtue of using parabolic reflectors. Please see Korch[Kor88] for a more detailed expression of the coefficients. For our system with an object at 5 metres the coefficients are as follows:

$$B = 0.067 \frac{1}{\text{cm}^2}$$

$$C = 2.48 \frac{1}{\text{cm}}$$

$$D = 1.06 \frac{1}{\text{cm}}$$

$$E = 5.3$$