

3)

Integrated Analog Focal Plane Processing for Automatic Alignment

by

Christopher Bruce Umminger

B. S., California Institute of Technology (1988)
S. M., Massachusetts Institute of Technology (1990)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1995

© Massachusetts Institute of Technology 1995

Signature of Author _____
Department of Electrical Engineering and Computer Science
Feb. 6, 1995

Certified by _____
Charles G. Sodini
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by _____
Frederic R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

APR 13 1995 Eng.

LIBRARIES

Integrated Analog Focal Plane Processing for Automatic Alignment

by

Christopher Bruce Umminger

Submitted to the Department of Electrical Engineering and Computer Science
on Feb. 6, 1995, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The speed of signal processing tasks in machine vision is often limited by the large amount of data in an image that must be acquired, transmitted, and processed. One such task is the automatic alignment of parts in manufacturing processes. Applications such as photolithography, PC board fabrication, and high speed web processes, such as color printing, require an alignment sensor that can quickly detect the location of an alignment mark. This research investigated the potential of using analog focal plane processing to build fast alignment sensors. The designs take advantage of the fact that analog circuits generally require less area, power, and time to perform basic computations compared to digital signal processing. Thus, these circuits can be incorporated with the imager on a single integrated circuit chip to compute alignment error. The resulting sensor is faster, more compact, weighs less, and consumes less power than a comparable digital system where the imager is separated from the signal processor.

Two sensors were designed and both use photodiodes that are spatially arranged to mate with the image of a particular alignment mark. Incorporated alongside the photodiodes are circuits which convert the photocurrents into an alignment error signal that is independent of illumination level. One of the sensor designs detects the image of a surveyor's mark (a light and dark check pattern). A coarse alignment signal is obtained by operating the sensor as a four quadrant photodiode. Additional photodiodes detect the edges of the mark to provide a more sensitive fine alignment signal. The other sensor operates with a square grating mark. Coarse alignment is accomplished using edge detection and the fine alignment signal is obtained from a position sensitive moiré fringe pattern formed between the sensor and the image of the mark.

The sensors were fabricated in a 2 μm *n*-well BiCMOS process. Test results of the photocurrent amplifiers confirmed gains ranging from 8 to over 800,000 depending upon the image brightness. The sensors were tested in a mock alignment system and were able to guide alignments with a repeatability of better than 60 ppm of the target size. Sensor bandwidths up to 55 kHz were measured. These results demonstrate that alignment sensors operating at speeds substantially above video rate can be built by incorporating focal plane image processing.

Thesis Supervisor: Charles G. Sodini

Title: Professor of Electrical Engineering and Computer Science

Biography

Christopher B. Umminger was born on July 20, 1966 in New York, New York. Raised near Kansas City, Missouri, he attended the Park Hill Senior High School and graduated in 1984. He received the B.S. degree from the California Institute of Technology, Pasadena, in 1988 and the S.M. degree from the Massachusetts Institute of Technology, Cambridge, in 1990, both in electrical engineering. His master's thesis was entitled "Switched Capacitor Networks for Machine Vision." He is a member of Sigma Xi and Tau Beta Pi.

Mr. Umminger spent the summers of 1986 and 1987 at the C.P.C. Engineering division of General Motors. There he worked with test systems for monitoring the performance of engine control computers. In 1991 he worked at Synaptics, Inc. designing analog integrated circuits for collective computation. His research interests center on analog circuit design applied to signal processing tasks.

Acknowledgments

Funding for this research was provided by the National Science Foundation and the Advanced Projects Research Administration (grant numbers MIP-8814612 and MIP-9117724). Support for myself was provided in part by the Office of Naval Research.

I am indebted to my thesis advisor, Professor Charles Sodini, for all of his assistance. His ability to sort through complex problems and identify their most important and fundamental aspects is a precious talent. My education here at MIT and this thesis have benefitted greatly from his guidance and practical sense. I also appreciate his generous spirit and friendly humor which have so often lightened the burden of graduate study. Professor John Wyatt deserves thanks for all his efforts on behalf of the Vision Group and his advice to me during my time at MIT. Special thanks also go to my other committee members Professor Berthold Horn and Professor Stanley Dubowsky.

Ignacio McQuirk, Steven Decker, Joe Lutsky, and Don Baltus have been excellent colleagues and friends. Their comments and criticisms over the years were invaluable. I owe Dona Lee a special debt of gratitude for her kindness, understanding, dedication, and companionship during the time I struggled with this project. The years I spent residing with my friends on Fourth East, East Campus, are cherished times. I will always remember their comradeship and spirit as well as the cocoa which fueled it all.

Finally, thanks to my parents and the rest of my family who have been so supportive of my education. I deeply appreciate all that they have done for me. This work I dedicate to my grandmother, Elnora Mae Umminger.

— Chris Umminger

*I've served this canon seven years and more,
Yet I am no nearer science than before.
All that I ever had I've lost thereby,
And so, God knows, have many more than I.
Time was when I was fresh and fond of cheer,
Liked decent clothes and other handsome gear;
Now I might wear my hose upon my head,
And where my color was a lively red
My face is wan and wears a leaden look;
If you try science you'll be brought to book.
My eyes are bleared with work on preparations,
That's all the good you get from transmutations.
That slippery science stripped me down so bare
That I'm worth nothing, here or anywhere.*

from 'The Cannon's Yeoman's Tale'
in The Canterbury Tales
by Geoffrey Chaucer
circa 1400

I know things about stuff.

— Zippy the Pinhead

Contents

1	Introduction	19
2	Background	23
2.1	Introduction	23
2.2	The Alignment Problem and Terminology	23
2.3	Application Areas for Automatic Alignment	25
2.3.1	Photolithography	25
2.3.2	Web Processing	27
2.3.3	Other Application Areas	27
2.4	Alignment Methods	28
2.4.1	Moments	28
2.4.2	Edge Detection	30
2.4.3	Correlation	31
2.4.4	Diffraction Gratings	32
2.4.5	Moiré Fringes	33
2.4.6	Comparison of Alignment Systems	37
2.5	Summary	39
3	Surveyor's Mark Alignment Sensor	41
3.1	Introduction	41
3.2	Alignment Method	41
3.2.1	Coarse Alignment	42
3.2.2	Fine Alignment	45
3.2.3	Combining the Two Signals	49
3.3	Current Amplification and Normalization	50
3.3.1	Normalizing (Type A) Current Multiplier	51
3.3.2	Inverting (Type B) Current Multiplier	54
3.3.3	MOS Current Mirror	56
3.3.4	Complete Amplifier	56

3.3.5	Active Cascode	58
3.3.6	Coarse Alignment Amplifier	59
3.3.7	Amplifier Biasing	61
3.3.8	Noise	61
3.3.9	Power	64
3.3.10	Effect of Normalization on Error Signals	64
3.4	Summary	66
4	Test Results: Surveyor's Mark Alignment Sensor	67
4.1	Introduction	67
4.2	Photodiodes	67
4.3	Current Amplifier	70
4.3.1	Fine Alignment Amplifier	70
4.3.2	Coarse Alignment Amplifier	72
4.4	Alignment Error Signals	73
4.4.1	Mock Alignment System	73
4.4.2	Coarse Alignment Signal	74
4.4.3	Fine Alignment Signal	76
4.4.4	Alignment Trajectories	80
4.4.5	Alignment Repeatability	81
4.5	Bandwidth	82
4.6	Noise	85
4.7	Summary	87
5	Square Grating Mark Alignment Sensor	89
5.1	Introduction	89
5.2	Alignment Method	90
5.2.1	Coarse Alignment	90
5.2.2	Fine Alignment	102
5.3	Absolute Value Circuit	113
5.3.1	Absolute Value Circuit Dynamics	116
5.4	Current Amplifiers	120
5.5	Summary	120
6	Test Results: Square Grating Mark Alignment Sensor	121
6.1	Introduction	121
6.2	Absolute Value Circuit	121
6.3	Alignment Error Signals	122
6.3.1	Coarse Alignment Signal	123

6.3.2	Fine Alignment Signal	126
6.3.3	Alignment Trajectories	127
6.3.4	Alignment Repeatability	129
6.4	Bandwidth	130
6.4.1	Fine Alignment	130
6.4.2	Coarse Alignment	131
6.5	Noise	135
6.6	Summary	136
7	Comparison	137
7.1	Introduction	137
7.2	Alignment Methods	137
7.3	Errors	138
7.4	Capture Regions	140
7.5	Implementation	140
7.6	Performance	141
7.6.1	Alignment Signals	141
7.6.2	Bandwidth	141
7.6.3	Noise	142
7.6.4	Power	142
7.7	Summary	143
8	Potential Improvements	145
8.1	Introduction	145
8.2	Photodiodes	145
8.3	Current Amplifier	146
8.4	Square Grating Sensor	148
8.4.1	Analysis	148
8.4.2	Implementation	150
8.5	Rotation Alignment	151
8.6	Focus	154
8.7	Summary	154
9	Conclusion	155
9.1	Summary	155
9.2	Future Work	157
9.3	Contribution	159

List of Figures

2-1	Mark-to-mark alignment.	24
2-2	Mark-to-sensor alignment	24
2-3	Direct referencing wafer alignment system.	25
2-4	Indirect referencing wafer alignment system.	26
2-5	Alignment using image moments of a circular mark.	29
2-6	Mark-to-mark alignment using image moments.	29
2-7	Edge detection alignment with a bar and window.	30
2-8	Alignment using correlation templates.	31
2-9	Nikon LSA diffraction system for alignment.	32
2-10	Fresnel grating pattern.	33
2-11	Moiré pattern from rotated line gratings.	34
2-12	Moiré pattern from line gratings of different pitches.	34
2-13	Moiré pattern from circular gratings.	35
2-14	Moiré pattern from circular gratings with unequal pitches.	36
2-15	Hyperbolic moiré pattern from misaligned circular gratings.	36
2-16	Moiré pattern from Fresnel gratings.	36
3-1	Layout of the photodiodes in the surveyor's mark sensor.	42
3-2	Four quadrant photosensor.	43
3-3	Coarse alignment error vector field with trajectory.	44
3-4	Coarse alignment error vector field with 10 degree rotation.	45
3-5	Fine alignment edge detection.	46
3-6	Fine alignment error vector field.	47
3-7	Fine alignment error vector field with rotation.	49
3-8	Differential current amplifier.	51
3-9	Gilbert normalizing multiplier.	52
3-10	Frequency response of the normalizing multiplier.	54
3-11	Gilbert inverting multiplier.	55
3-12	Frequency response of the differential current amplifier.	57

3-13	Frequency response of the loaded current amplifier.	58
3-14	Frequency response of the loaded current amplifier with active cascode.	59
3-15	Normalizing current amplifier with four inputs.	60
3-16	Amplifier noise spectrum without active cascode.	62
3-17	Amplifier noise spectrum with active cascode.	63
3-18	Small-signal model of the active cascode noise source.	63
3-19	Coarse alignment error vector field after normalization	65
4-1	Cross section of photodiode structure.	68
4-2	Current amplifier input/output characteristic.	71
4-3	Current amplifier nonlinearity.	72
4-4	Coarse alignment x -direction error signal along the x -axis.	74
4-5	Coarse alignment x -direction error signal near the origin.	75
4-6	Coarse alignment error vector field.	76
4-7	Coarse alignment error vector field with ten degree rotation.	77
4-8	Fine alignment x -direction error signal along the x -axis.	77
4-9	Fine alignment error vector field.	79
4-10	Fine alignment error signal with rotated target.	79
4-11	Alignment trajectories with error vector field.	81
4-12	Fine alignment repeatability.	82
4-13	Sensor bandwidth.	84
4-14	Noise spectrum without active cascode.	85
4-15	Noise spectrum with active cascode.	86
5-1	Patterns from square gratings with a large misalignment.	91
5-2	Moiré patterns from square gratings with a small misalignment.	91
5-3	Layout of the photodiodes in the square grating sensor.	92
5-4	Edge falling across a row of photodiodes.	92
5-5	Modified absolute value input/output characteristic.	95
5-6	Edge signal plotted versus edge position.	97
5-7	Dividing the square grating into quadrants.	98
5-8	Regions where edge crossings occur with overlaid square gratings.	98
5-9	Area element of the sensor used in approximate analysis.	100
5-10	Predicted coarse alignment error signals.	101
5-11	One-dimensional sensor responsivity function.	104
5-12	Fourier transform view of moiré fringe generation.	105
5-13	Signal flow diagram for a basic phase detector.	107
5-14	Two-dimensional sensor responsivity function.	108
5-15	Amplitude factor versus pitch mismatch.	111

5-16	Absolute value circuit.	114
5-17	Absolute value circuit input/output characteristic.	115
5-18	Absolute value circuit dynamics.	117
5-19	Absolute value circuit large signal transients.	118
6-1	Absolute value circuit input/output characteristics.	122
6-2	Coarse alignment quadrant currents for x -axis displacement.	123
6-3	Coarse alignment x -direction error signal along the x -axis.	124
6-4	Coarse alignment error vector field.	125
6-5	Effect of rotation on the coarse alignment error signal.	126
6-6	Fine alignment error signal and vector field.	127
6-7	Alignment trajectories with error vector field.	128
6-8	Alignment trajectories and error vector fields with added bias current.	128
6-9	Fine alignment repeatability.	129
6-10	Fine alignment signal bandwidth.	131
6-11	Coarse alignment small-signal bandwidth.	132
6-12	Coarse alignment transient times.	133
6-13	Noise spectrum.	135
8-1	Improved normalizing amplifiers.	147
8-2	Bootstrap circuit around a photodiode.	148
8-3	Rotated square gratings.	152
8-4	Rotated edges across photodiode array.	153

List of Tables

2-1	Specifications of some reported alignment systems.	38
4-1	Current amplifier DC specifications.	70
7-1	Summary of measured performance.	143

Introduction

Many manufacturing processes require accurate position alignment between a part or product and the machine operating on it. Fiducial marks are often used to aid in this alignment process. Manually positioning such marks can be both tedious and slow. Thus, many ways of automating this operation have been developed in order to improve alignment speed and reliability. One of the most important components of any automatic alignment system is the sensor which detects the alignment error. A fast and accurate sensor is essential for quick alignment and high product throughput.

Optical alignment systems are used for many manufacturing tasks, especially in the production of semiconductor microcircuits. Photolithography is a notable example although automatic alignment is also used in die slicing and bonding, as well as wafer probing for testing [1]. Similar alignment problems to those in photolithography occur in the manufacture of flat panel displays, multilayer PC boards, and some web processes such as color printing. All of these tasks require several layers of patterned material to be aligned precisely [2].

Systems for automatically performing alignment vary greatly between manufacturers. Alignment marks range from bars and crosses to diffraction gratings and Fresnel zone plates. The detection systems can be as simple as a photodiode and slit or as complicated as a digital image acquisition and processing system. The type of alignment system chosen often depends upon the precision desired. Systems based on the diffraction properties of light are often used when the desired registration accuracy approaches the wavelength of light. Less sophisticated imaging systems can be used when the registration requirements are less stringent [3].

Many systems use digital signal processing to determine alignment error. This approach has several strengths. Foremost among these is versatility. A software algorithm for image processing can be readily modified for a particular application. Another advantage is the computational accuracy afforded by digital computing. Finally, because such systems are comprised of standard imagers and signal processors, they are relatively inexpensive. These systems also have some disadvantages. Chief amongst these is the speed at which such systems can com-

pute the alignment error. The camera alone typically takes 17 ms (video rate) to acquire and transmit the image. This time results from the information bottleneck that occurs when the two-dimensional image is transformed into a one-dimensional data stream. Additional time is consumed by the subsequent digital signal processing. Although this computation can be pipelined to obtain a high throughput of image information, it is difficult to keep the latency small. This delay can be an important limit on the bandwidth of the alignment system. Another disadvantage of such systems is the limit on alignment accuracy imposed by the pixel size of the imager. Although subpixel accuracy is possible, it can be difficult to detect positions much better than one fourth to one tenth of the pixel size [1, 4, 5, 6, 7] although others claim that accuracies of one to four percent are possible [8, 9].

This research investigated the possibility of using integrated analog focal plane processing to build sensors for automatic alignment. The sensors described in this thesis use photodiodes to detect the alignment marking and circuits integrated with the photodiodes to compute an alignment error signal. Analog circuits are used because they generally require less time, area, and power to perform basic functions when compared to digital signal processing. Computations that can be done easily with analog circuits include the arithmetic operations of addition, subtraction, averaging, multiplication, logarithm, exponentiation, and absolute value as well as light sensing, time and spatial integration, filtering, and thresholding. Unfortunately, analog processing also has some serious disadvantages. Analog circuits are inflexible once built, have limited computational accuracy (typically on the order of eight bits), are expensive in low volume production, and may take more effort to design [10, 11, 12]. Another difficulty with focal plane processors is the small amount of computation that can be built into the tight area constraints of an imaging pixel.

The compactness of analog computation allows the relatively simple signal processing for alignment to be combined with the optical sensor on the same integrated circuit. This reduces the communication bottleneck between the imager and the signal processor and allows for parallel computation with the image information. These advantages combined with the speed of analog computation enable the design of a fast alignment sensor. A single integrated circuit package and associated optics can acquire the image, process it, and output the appropriate error signals to guide an alignment servomechanism. Such a sensor can be faster and more compact, weigh less, and consume less power than digital solutions.

The thesis begins in Chapter 2 by presenting some background information on alignment, clarifying terminology, and discussing various approaches to automatic alignment that have been proposed and/or implemented in the past. Some application areas are discussed with an emphasis placed on photolithography. Various alignment techniques are summarized including methods based on image moments, edge detection, correlation, diffraction gratings, and moiré fringe patterns.

These alignment techniques can be broken down broadly into two classes. The first category

consists of techniques that search for a single feature, or small number of features, in an image to locate the target. Examples include edge detection, correlation, and moment methods. The second class consists of multiple feature techniques that examine the changes in large patterns to determine position. These techniques combine information from the large number of features that comprise the pattern. The primary examples are grating techniques that use moiré fringes and the diffraction patterns of light to detect misalignment. In this work two sensors were designed to explore how analog focal plane processing could be adapted for these two different approaches to alignment.

Chapter 3 presents a sensor design that uses edge detection to position a target called the surveyor's mark. The sensor has two modes of operation. A coarse alignment signal is generated by comparing the brightness in quadrants of the image. After initial positioning with this signal, edge detection is used to provide a more sensitive fine alignment signal for completing the alignment. These two methods are explained in the first half of the chapter. The latter half discusses the circuits and implementation issues that arise in transforming the detected image into an output alignment error signal.

The surveyor's mark sensor was fabricated and its test results are presented in Chapter 4. Measurements were made characterizing the photodiodes and amplifiers. A mock alignment system was also constructed to test the sensor when imaging the target mark. Error signals and alignment trajectories were measured for different target positions using this system. Closed loop alignments were completed and the resulting trajectories and repeatability are discussed. The chapter concludes with data on the sensor bandwidth and noise characteristics.

The second sensor design uses a more complicated square grating target and a moiré fringe alignment technique. Chapter 5 presents the theory behind the operation of this sensor. Coarse alignment is accomplished using edge detection and final positioning is achieved by detecting a moiré fringe formed between the sensor and the target grating. Most of this chapter is devoted to developing models for the coarse and fine alignment signals. An absolute value circuit important for the edge detection is discussed near the end. This sensor uses current amplifiers similar to those discussed in Chapter 3 for the surveyor's mark sensor.

Test results from the square grating sensor are presented in Chapter 6. The absolute value circuit is characterized first. This is followed by data from the sensor operating in the mock alignment system. Results are shown for the alignment error, trajectories, repeatability, as well as measurements of the sensor bandwidth and noise.

The two sensor designs are compared in Chapter 7 and a discussion of the strengths and weaknesses of the implemented alignment methods is made. The discussion is framed around the distinction drawn between single feature methods and pattern methods for alignment. Various measures of performance such as repeatability, bandwidth, and noise are also discussed. The sensors work well, but there is room for improvement. Chapter 8 covers some possible changes to the designs that could make their performance even better. The work is summarized and

concluding comments are made in Chapter 9.

This thesis examines some of the potential advantages of focal plane processing and what is required to capitalize on them. These ideas are applied to the particular problem of automatic alignment because it appears to fit well with the strengths afforded by focal plane processors. The resulting sensors have definite advantages over other methods, primarily in their speed. In the course of designing these sensors, the issues that limit the speed are explored and some of the constraints of focal plane processing are further illuminated. Test results from these sensors confirm their operation and characterize the performance.

Background

2.1 Introduction

Much work has been done over the years in order to automate the alignment process in various manufacturing tasks. There are many different alignment problems and a variety of ways to solve them. This section begins by presenting two basic alignment situations and some associated terminology. Some application areas requiring high speed automatic alignment are discussed with an emphasis on photolithography in semiconductor manufacturing. Several alignment methods are then treated with examples from the literature, most of which also concern the alignment problem in photolithography.

2.2 The Alignment Problem and Terminology

“Alignment” can be defined as the proper positioning of two parts relative to each other. For optical alignment systems this usually means positioning alignment marks. There are really two alignment problems which can be termed “mark-to-mark” alignment and “mark-to-sensor” alignment. Although both seek the same goal, the relative positioning of two parts, they differ fundamentally in implementation.

Mark-to-mark alignment is perhaps the more familiar situation and is depicted in Figure 2-1. A cross and its outline form the fiducial marks in this example, although many other mark combinations are possible. A target mark appears on the part which is to be aligned. It is aligned to a reference mark referred to as the reticle. Typically the reticle is fixed in the optical system, such as crosshairs in a gunsight or microscope. The first lens images the target mark onto the reticle and the image of both is then focussed onto the sensor by the second lens. If a human operator were performing the alignment, for example, then the sensor would be the retina of the eye. In this system the sensor views both marks and neither the sensor nor the marks need to be in a fixed position with respect to the optics. Where the image of the marks

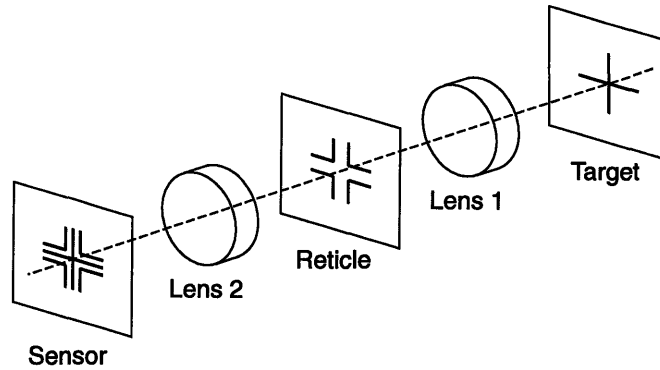


FIGURE 2-1: In mark-to-mark alignment both the target mark and the reticle are imaged onto the sensor. The target mark is aligned to the reticle.

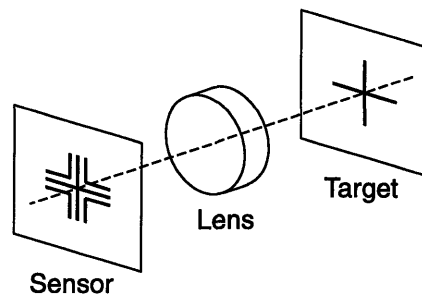


FIGURE 2-2: Mark-to-sensor alignment images only the target mark onto the sensor. The target is aligned to the sensor which can be viewed as incorporating the reticle.

falls on the sensor is not important.

Mark-to-sensor alignment is illustrated in Figure 2-2. Here the sensor itself is the reference and is fixed in relation to the machine, usually on the optic axis. The target mark is imaged onto the sensor and aligned to it. Note that there is no reticle mark although the sensor can be considered as being an implicit reticle mark.

Mark-to-sensor and mark-to-mark alignment are different in important ways. Mark-to-mark alignment requires a more complex optical system. Furthermore, the image processing problem of aligning two marks is more difficult than aligning a mark to a sensor. A mark-to-mark system must be able to detect both the reticle and target marks as well as distinguish between them in order to determine the misalignment. On the other hand, mark-to-sensor alignment requires only that the sensor detect the target mark location. However, in this case the sensor must be aligned to the machine which is not always an easy task. Both sensors presented in this thesis are designed for the mark-to-sensor alignment problem. The reduced amount of information processing required in this case allows one to more easily combine the computation with the imager.

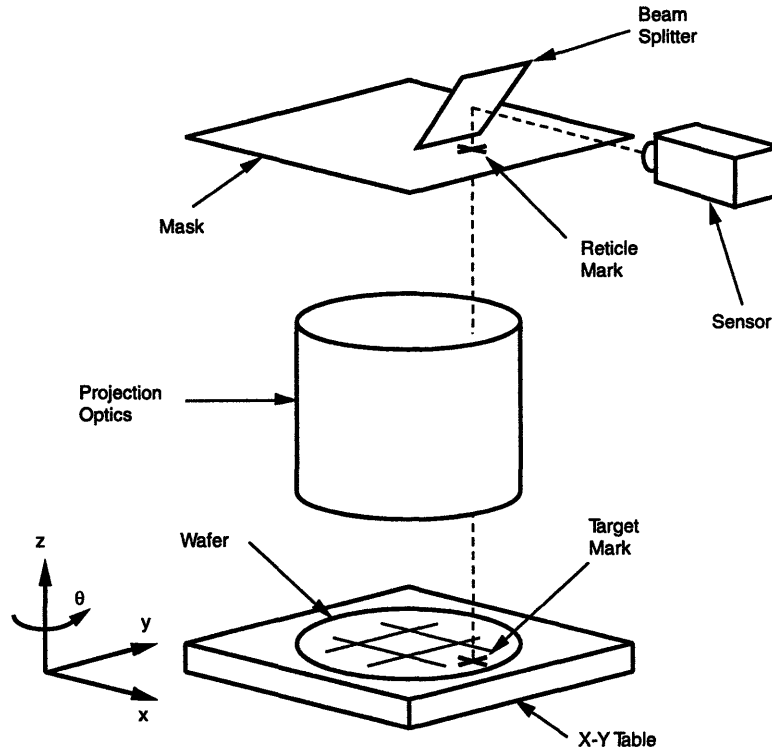


FIGURE 2-3: A direct referencing (mark-to-mark) wafer alignment system for photolithography in the manufacture of integrated circuits. The reticle mark is aligned to the target mark through the projection optics (after Stevenson [3]).

2.3 Application Areas for Automatic Alignment

There are many areas of manufacturing that require automatic alignment. Photolithography in the production of semiconductor microcircuits is one example where alignment is very important and many alignment strategies have been published and implemented. For this reason it is useful to understand some basics about photolithographic alignment systems.

2.3.1 Photolithography

Photolithography is the process by which a pattern is optically transferred to a surface for etching. In the manufacture of integrated circuits it refers to the transfer of a circuit mask pattern to a semiconductor wafer coated with a photosensitive resist. The basic components of this system are shown in Figure 2-3. Before the exposure takes place, the mask pattern must be aligned with the patterns from previously deposited layers. The alignment is critical in this process because of the exceedingly small dimensions involved. Several stages of alignment are used beginning with a mechanical pre-alignment that typically positions the wafer on the x - y table to within $\pm 40 \mu\text{m}$ of the aligned position. This is followed by an optical global pre-alignment that reduces position uncertainty to around $\pm 2 \mu\text{m}$. Lastly, an optical fine alignment

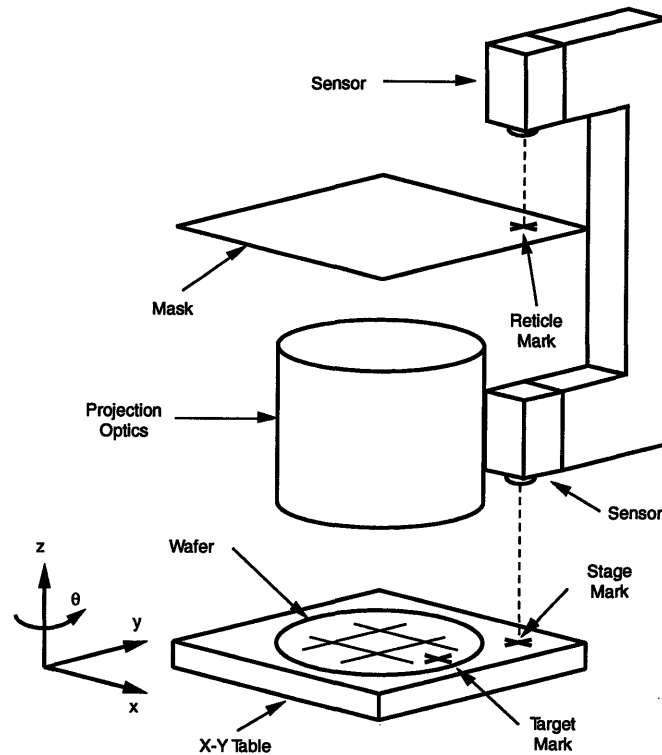


FIGURE 2-4: An indirect referencing (mark-to-sensor) wafer alignment system for photolithography. The reticle and target marks are aligned to separate sensors which are assumed to be rigidly connected in a known and stable way.

positions the part to an accuracy about three to five times smaller than the minimum feature size (about $0.2 \mu\text{m}$ for $0.8 \mu\text{m}$ features). This sequence of alignment operations must be done at least once per wafer. Usually the mask contains the pattern information for only one die site on the wafer. Because there are many dice on each wafer, the machine must expose each one separately and translate the wafer between exposures. Some machines only align the wafer and mask once and use open loop control with high accuracy stepping motors to move between die sites. For very high accuracy applications the fine alignment step is repeated at selected die sites on the wafer before they are exposed [3, 13]. For die-by-die alignment the time spent on the alignment operation becomes an important factor in limiting wafer throughput.

The alignment system illustrated in Figure 2-3 is a mark-to-mark alignment system where the mask contains the reticle mark and the wafer contains the target mark. Such systems are called direct referencing because they align the mask mark directly to the wafer mark. This is to be contrasted with the system shown in Figure 2-4 which uses mark-to-sensor alignment. In this case the mask is aligned to the upper sensor and the wafer is aligned to the lower sensor. The sensors for both alignment operations are assumed to be rigidly fixed to the machine. By aligning each part to its respective sensor, one indirectly aligns them to each other, hence the name indirect referencing for such systems. Many lithography machines use reduction lenses

which reduce the image of the mask by a factor of five or ten. In these systems the reticle alignment is much easier than the wafer alignment and thus a less complex and less expensive alignment system can be used for the reticle. In order to complete the alignment one must know the distance between the aligned position for each sensor, called the baseline, which may drift slightly with temperature and other environmental factors. This can be obtained by placing a permanent target mark on the x - y table. The target is aligned to both sensors and the table is used to measure the distance between the two aligned positions.

These two types of alignment systems each have their advantages. Direct referencing has no baseline and thus no baseline drift as pointed out by Lavine [14] and v. d. Brink [15]. However, a direct referencing system must image through the projection lens which has been optimized for the wavelength of light used for exposure. Indirect referencing systems, although subject to deficiencies of a baseline, can use optics and light sources that are separately optimized for the alignment task [16].

2.3.2 Web Processing

A special purpose high speed alignment sensor may not be of much use if the mechanical portion of the alignment system itself is slow. One application area where this is definitely not the case is web processing. These are manufacturing areas which produce long and thin sheets, or webs, of material. Certain web processes require alignment between the moving rollers and the material in order to synchronize cutting operations or the transfer of patterns. Examples include textile manufacturing and color printing [5].

Performing such alignments in real time is very difficult because the webs are moving at speeds up to 27 m/s [17]. Even acquiring sharp video images of material moving at these speeds is difficult. For example, a web inspection system looking for irregularities in film moving at 4 m/s required a laser scanning at 12 kHz, high speed photomultiplier tubes, and a network of 50 transputers operating at 100 MHz [18]. Because of the difficulty in automating this alignment, adjustments are often made by people who inspect already processed material. Thus, if there is any significant error and it is not rapidly detected, a large amount of web material can be wasted.

2.3.3 Other Application Areas

In addition to photolithography for the manufacture of integrated circuits, automatic alignment is used in many other manufacturing processes. Similar alignment needs occur in the production of flat panel displays and multilayer printed circuit boards. Die slicing, probing, and bonding are additional examples of tasks in integrated circuit manufacture that require alignment [1]. A more detailed examination of the die bonding problem was given by Horn [19]. He proposed using an edge detection scheme to locate the boundary between circuitry and an aluminum layer which frames the die. This frame was essentially an alignment mark. A video

camera was used to obtain images of the die and provide data for the edge detection algorithm. The required signal processing took 250 ms using mid-1970's technology.

A final potential application for such alignment sensors is in the guidance of autonomous vehicles [20, 21]. In some warehouses and manufacturing sites unmanned robotic vehicles are used to move around parts and material. These vehicles sometimes follow marks placed on the ceiling or floor, or else measure their own motion from sparsely placed alignment marks. These marks are necessary to periodically correct for accumulated position errors in a vehicle's guidance system. A fast alignment sensor would permit these vehicles to update their positions while still moving at high speeds.

2.4 Alignment Methods

A wide variety of methods for accomplishing alignment have been investigated and developed. A review of several of these follows.

2.4.1 Moments

One of the most simple alignment methods makes use of the center of area, or other higher order moments, of an image. The center of area of a binary image $b(x, y)$ is found by computing the first moments of the image [22, chapter 3]

$$A = \iint_R b(x, y) dx dy \quad (2.1)$$

$$\bar{x} = \frac{1}{A} \iint_R x b(x, y) dx dy \quad (2.2)$$

$$\bar{y} = \frac{1}{A} \iint_R y b(x, y) dx dy. \quad (2.3)$$

In an idealized binary image with a single mark, as shown in Figure 2-5, the first moments will yield the coordinates of the center of area of the mark. A mark-to-sensor alignment system might define the aligned position to be in the center of the image, in which case a simple error vector such as

$$e_x = \bar{x} - x_0 \quad (2.4)$$

$$e_y = \bar{y} - y_0 \quad (2.5)$$

could be used to guide an alignment system. Because of its symmetry, a circular fiducial mark is claimed by Bose [4] to work best in this kind of a system. Building a detector to find the center of area can be done in several ways. If the mark is a circular spot of light, then one can detect its position using a large area photodiode [23, 24, 25]. Focal plane processors which find the true center of area of an image have also been reported by DeWeerth [26] and Standley [27].

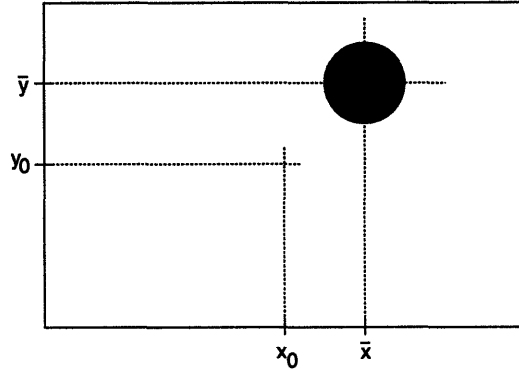


FIGURE 2-5: A moment method might align the centroid of a mark (\bar{x}, \bar{y}) to the center of the image (x_0, y_0) .

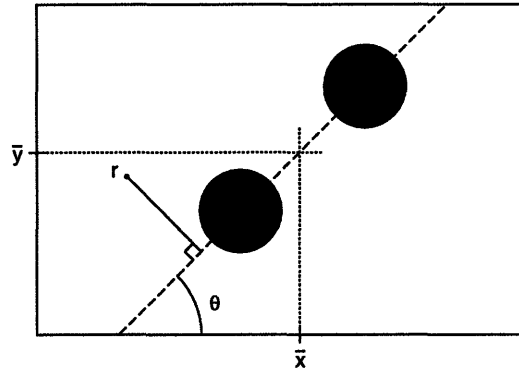


FIGURE 2-6: Higher order moments can be used to compute the center of area (\bar{x}, \bar{y}) for two marks as well as the axis of least inertia (dashed line) connecting them in order to align one mark to the other.

The chief problem with using the center of area is that it may not always correspond to the center of the mark. This can happen if the mark contains a defect, if other marks are present in the image, or if a continuous (as opposed to binary) image is nonuniformly illuminated.

A mark-to-mark alignment scheme is also possible using image moments that would position one circular mark on top of another as depicted in Figure 2-6. By computing higher moments of the image one can find the axis of least inertia of the image. The inertia about an axis is defined as [22]

$$I = \iint_R r^2 b(x, y) dx dy. \quad (2.6)$$

The axis of least inertia will be the line connecting the centers of the two circular marks and will pass through the center of area (\bar{x}, \bar{y}) . By also computing the axis of maximum inertia one can define an alignment measure using the ratio of the maximum inertia over the minimum inertia

$$\text{Alignment measure} \triangleq \frac{\max I}{\min I}. \quad (2.7)$$

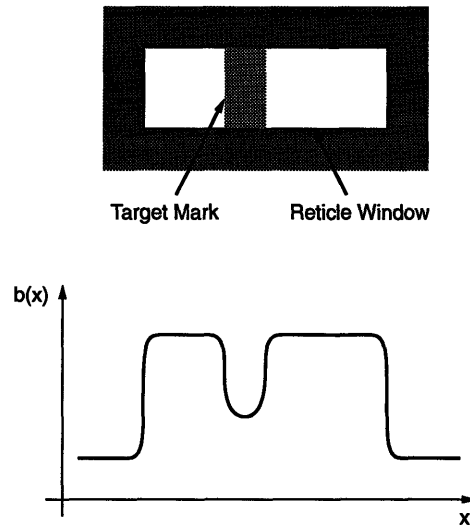


FIGURE 2-7: An edge detection alignment scheme using a bar viewed through an opening in the reticle window. A one-dimensional camera acquires brightness data as a function of position $b(x)$ which is then examined for edges to determine if the bar is in the center of the window.

This ratio will approach one as the marks move into alignment. The sensor proposed by Horn [28] and implemented by Standley has the capability of computing all the moments necessary to determine the center of area and the orientation of the axis of least inertia. However, the accuracy of the calculation is insufficient for most alignment tasks. Errors are quite large for nearly symmetrical images, such as would appear as the marks approach alignment. Directional ambiguity is another problem with this method. There is no way of knowing which direction along the axis of least inertia one must move in order to bring the marks into alignment. The best one can do is to try a direction and see if the marks move closer. If they move apart, then the direction of motion would have to be reversed.

2.4.2 Edge Detection

Several photolithography alignment systems are based upon detecting edges in one dimension. An example of such an edge detection scheme is shown in Figure 2-7. The reticle mark is a clear window on the mask plate through which a single dark bar on the target is viewed. The alignment system places the bar in the center of the window by acquiring brightness data horizontally. This is often accomplished with a linear array camera or a full field imaging camera and post-processing that averages the signal vertically. The resulting one-dimensional brightness function is then processed to determine the locations of the edges of the bar and window. Edge detection systems can be relatively slow because the alignment marks must be imaged, converted into digital signals, and then processed before any motion of the table can begin. Usually the processing is done in a dedicated digital signal processor in order to improve the alignment speed. For example, the SUSS ALX alignment system uses a video camera and a

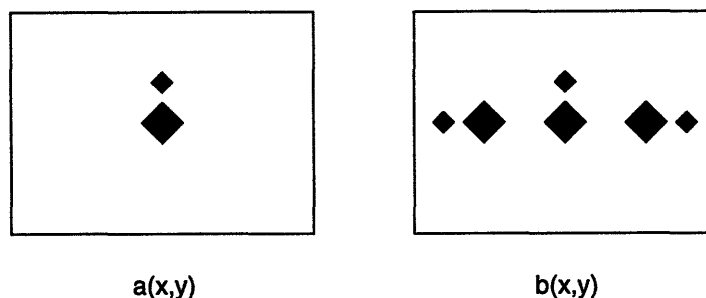


FIGURE 2-8: Correlation methods locate a template image $a(x, y)$ by looking for the position with maximum correlation in a search image $b(x, y)$.

100 MFLOP DSP board to complete alignments in 200 ms [29, 30]. Several wafer steppers have used edge detection in various forms including the Electromask 700SLR and 800SLR [31], Cannon FPA-1550 [32, 33], SVG (Perkin-Elmer) Micralign 600-HT series [34, 35, 36, 37, 38, 39], Optimetrix Model 8010 [40, 41], Hampshire Instruments 5000P [42], some earlier GCA machines [14, 43], as well as machines reported by Bobroff [44] and Akiyama [16].

2.4.3 Correlation

Image correlation, or template matching, is another popular method used in automatic alignment. It is powerful because it can be easily used for mark-to-mark as well as mark-to-sensor alignment. The technique is based on computing the crosscorrelation integral for two images such as those shown in Figure 2-8

$$\phi_{ab}(x, y) = \iint_{-\infty}^{\infty} a(\xi - x, \eta - y) b(\xi, \eta) d\xi d\eta. \quad (2.8)$$

The image $a(x, y)$ is a template of the mark being searched for in the image $b(x, y)$. The correlation $\phi_{ab}(x, y)$ will be at a relative maximum for the coordinates where the mark in the template occurs in the image $b(x, y)$. Correlation measures which are faster to compute than multiplication, such as the absolute value of difference, can be used and are discussed by Tian [8]. Crosscorrelation of image features rather than image brightness is discussed by Sakou [45].

Two techniques are primarily used for image correlation. In the first, a camera acquires an image of the search area. A template matching processor then computes the crosscorrelation integral in order to locate the mark in the image. The FPA-1550 M III W stepper from Cannon Inc. uses template matching with image processing in some of its alignment operations [32]. The image processing can be relatively slow due to the tremendous amount of computation involved in calculating the crosscorrelation. An image processing system developed by Cognex Corporation reportedly takes 200 ms to search a 576×448 pixel area for a 128×128 size pattern with ± 0.25 pixel resolution using a dedicated microprocessor running at 12.5 MHz [1, 46]. Faster processors available today can probably improve this speed proportionately. Special purpose

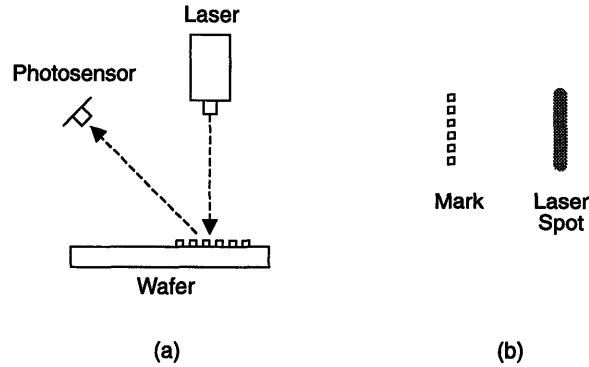


FIGURE 2-9: A simplified diagram of the Nikon LSA alignment system. In the aligned position a photosensor detects a laser beam diffracted from a grating on the wafer surface. (a) Side view of the system. (b) Top view of the grating and laser spot (after [53]).

hardware for template matching is available from LSI Logic Corporation [47] and has also been studied by Ranganath [48].

The second technique for performing correlation uses a single photosensor in a mark-to-mark alignment scheme. The reticle mark is an opaque screen or mask with a window cut out to match the target mark shape. The target mark is moved so that its image on the reticle translates while the photosensor measures the light from the mark which passes through the reticle window. The photosensor signal will be at a maximum when the target mark and the reticle window are in alignment. Typically, the target is scanned in a regular pattern and the photosensor signal is sent to a digital signal processor to determine the location at which the signal was at a maximum. The target is then translated back to the aligned position. Target marks for these systems vary widely. A simple cross, discussed by Hershel, is used on some Ultratech steppers [3, 49] and a series of bars and slits is used on some GCA steppers [50, 51]. Two-dimensional patterns based on Barker codes have also been proposed for this kind of alignment [52].

2.4.4 Diffraction Gratings

Diffraction gratings are useful in photolithography because they are easily made on the wafer surface. Furthermore, they are functional even when containing localized defects and typically require only a photosensor for detection. Figure 2-9 shows a simplified version of the Nikon Laser Step Alignment system [53, 54, 55, 56]. A laser beam is focussed to form an elongated spot on the wafer surface. Normally the beam is specularly reflected away from the photosensor. However, when the spot crosses a row of pits etched into the wafer surface, a diffracted beam from the grating will strike the photosensor. The laser spot is moved over the grating while data from the photosensor is sent to a special purpose digital processor. The processor computes the coordinates where the diffracted beam was strongest and the stage is then moved back to

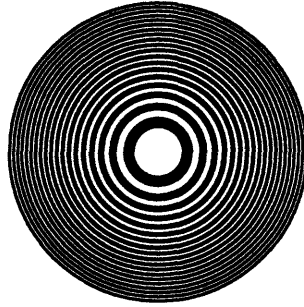


FIGURE 2-10: A Fresnel grating pattern will diffract light into a point similar to a lens.

the aligned position. This is similar to the correlation methods mentioned previously; in effect the laser spot is correlated with the wafer grating. This is an indirect referencing system where the wafer mark is aligned to the photosensor and laser beam. The reticle is aligned to the stepper using a different system. Each dimension of translation is aligned separately in the Nikon system. A method that can align both dimensions simultaneously has been proposed by Kleinknecht which uses multiple gratings and photosensors to improve alignment speed [57]. A system using a combination of two diffraction gratings on the mask and one on the wafer with a scanning glass plate has been proposed by Dubroeuq et al. [58]. A similar system has also been reported by Makosch [59].

A special diffraction grating, called a Fresnel pattern or Fresnel zone plate, is shown in Figure 2-10. This grating will diffract collimated monochromatic light into a point much like a lens does. One-dimensional Fresnel zone plates have been used by Fay [60, 61, 62, 63], McIntosh [64], and Chen [65] to focus light into a line on the wafer. An array of pits is then aligned to this light spot similar to the Nikon LSA method described above. Feldman et al. [66, 67] have proposed a scheme using circular Fresnel zone plates that image the focussed spot onto a quadrant photodiode or pass it through a slit onto a single photodiode.

2.4.5 Moiré Fringes

Moiré fringes, first discussed by Lord Rayleigh, have been used for many years for metrological purposes and more recently in alignment systems. A geometric moiré pattern appears when two similar periodic patterns are superposed to form alternating light and dark regions. These regions are larger than the underlying patterns and are called moiré fringes [68, 69, 70, 71, 72]. Figure 2-11 shows two superposed vertical gratings rotated by a small angle θ relative to each other. The resulting moiré pattern is a set of parallel dark bands that lie along the bisector of the angle between the two gratings. The spacing between the lines is given by

$$D = \frac{P}{\sqrt{2 - 2 \cos \theta}} \approx \frac{P}{\theta} \quad \theta \ll 1. \quad (2.9)$$

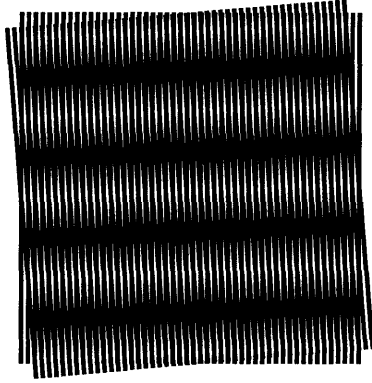


FIGURE 2-11: Moiré pattern from a pair of equal pitch linear gratings with one rotated slightly with respect to the other.

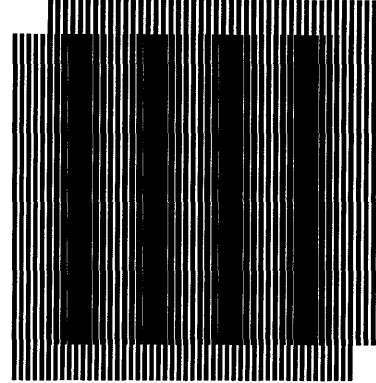


FIGURE 2-12: Moiré pattern from two linear gratings whose pitches differ by ten percent.

The magnification is defined to be the ratio of the fringe spacing to the grating spacing. For small angles it is simply the reciprocal of the angle

$$M \approx \frac{1}{\theta}. \quad (2.10)$$

Figure 2-12 shows the moiré pattern which results from two overlaid vertical gratings with slightly differing pitches $P_1 = P$ and $P_2 = P(1 + \epsilon_p)$. The resulting fringe spacing is given by

$$D = \frac{P_1 P_2}{P_2 - P_1} = \frac{P(1 + \epsilon_p)}{\epsilon_p} \quad \epsilon_p \ll 1 \quad (2.11)$$

with a magnification

$$M = \frac{1 + \epsilon_p}{\epsilon_p}. \quad (2.12)$$

Moiré fringes are very sensitive to changes in the underlying gratings. These two examples show how fringes form with a slight rotation or mismatch in pitch between two overlaid vertical gratings. If the larger pitch grating depicted in Figure 2-12 is translated perpendicular to the grating lines, then the moiré fringes will move M times further in the reverse direction. This magnification of grating motion has been used to extend the capture range as well as improve the sensitivity of alignment systems. Moiré patterns can be applied to both mark-to-mark as well as mark-to-sensor alignment.

One of the earliest applications of moiré patterns to alignment was studied by King and Berry in the early 1970's [73]. They explored the use of gratings made of concentric circles to perform visual mask-to-wafer alignment in proximity printing photolithography. When two identical gratings are superposed in alignment there is no moiré pattern. When these gratings are displaced slightly the pattern shown in Figure 2-13 appears. The fringes form hyperbolas whose foci are located at the centers of the two gratings. The number of fringes corresponds to the number of pitch spacings that separate the centers of the two gratings and serves as a measure of the misalignment. Furthermore, the line connecting the foci of the hyperbolas is

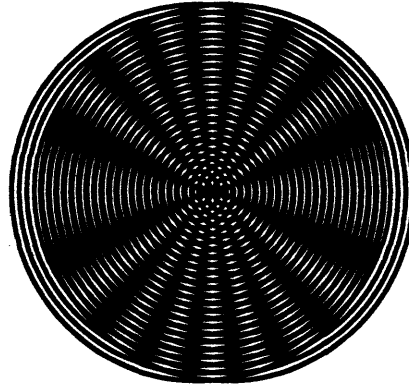


FIGURE 2-13: Moiré pattern generated by overlaying two identical circular gratings. The fringes form hyperbolas about the grating centers and the number of fringes indicates the misalignment.

the line along which the gratings must move to be returned to alignment. It is not apparent from the pattern which direction along this line the correction should be made. This is another example of an alignment scheme with directional ambiguity in the signal. The pattern is easy for a human to perceive and use as an error signal for alignment. A little trial and error quickly determines which direction of motion brings the patterns into alignment. Another problem with this method is the fact that the marks are not rectangular. Rectilinear (manhattan geometry) artwork generators have difficulty creating these marks. It is also difficult to achieve alignment accuracy much greater than one half the pitch spacing of the gratings because the moiré pattern begins to vanish at that separation.

In order to achieve more accurate alignment, King and Berry proposed using two circular gratings with slightly unequal pitches. When two such gratings are in alignment, circular fringes are formed as seen in Figure 2-14. These are similar in character to the fringes formed by the linear gratings shown in Figure 2-12. The fringes are very sensitive to motion of the gratings and thus provide an accurate measure of misalignment. King and Berry used gratings with pitches of 4.0 and 3.8 μm to form fringes with a spacing of 76 μm (20 grating lines). They were able to align the marks to within $\pm 0.2 \mu\text{m}$. Figure 2-15 shows the resulting pattern when the two unequally spaced gratings are misaligned. It consists of cusps rather than hyperbolas and no longer contains directional ambiguity. It is difficult to make a machine recognize the information present in these patterns because of their complexity and it appears that they have not been used in any automatic alignment systems.

The circular Fresnel grating also has an intriguing moiré pattern. A Fresnel grating is formed by alternately blacking out the regions between circles given by the equation

$$x^2 + y^2 = nR^2 \quad (2.13)$$

where n is a natural number and R is the radius of the innermost circle. The moiré pattern

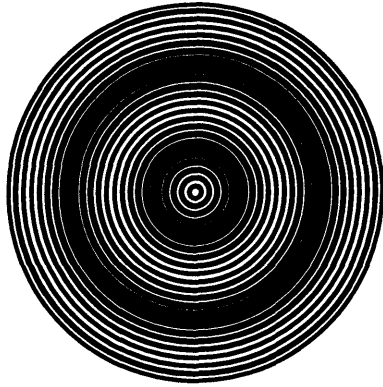


FIGURE 2-14: Circular moiré pattern generated by two overlapping and aligned circular gratings whose pitch spacing differs by ten percent.

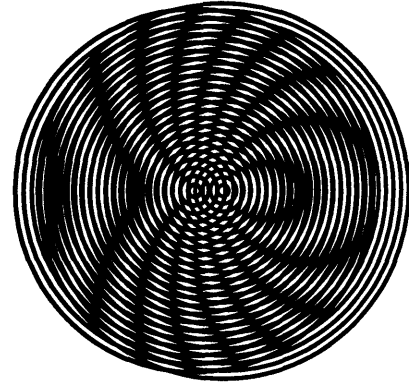


FIGURE 2-15: Hyperbolic moiré fringes generated by two misaligned circular gratings with differing pitches. The bulges to the right show that the smaller pitch grating lies to the right of the larger pitch grating.

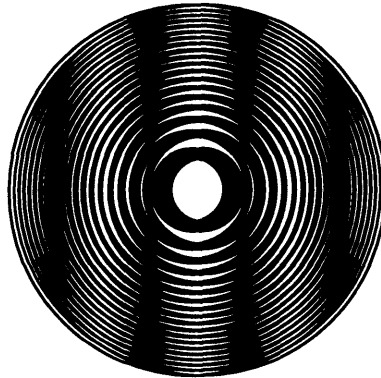


FIGURE 2-16: Moiré pattern generated by overlaying two Fresnel gratings results in parallel fringes perpendicular to the displacement with spacing proportional to the misalignment.

formed by two Fresnel gratings is shown in Figure 2-16. The fringes form a set of parallel lines perpendicular to the line connecting the centers of the patterns. The spacing between the fringes D is related to the separation S between the centers by¹ [69]

$$D = \frac{R^2}{S}. \quad (2.14)$$

This pattern is not as complex when compared to those formed by the circular gratings shown earlier. An alignment system using human operators to observe the fringes in overlaid Fresnel patterns has been reported by Lyszczarz [74]. It does not appear that any automated systems make use of this fringe pattern. A focal plane processing microcircuit that detects line orientation has been reported by Allen [75] that could be of use detecting this fringe pattern. It might be possible to develop a system that detects both the orientation and spacing of a set

¹The derivation in [69] is incorrect by a factor of two.

of parallel lines. However, the Fresnel pattern is probably not widely used in moiré alignment systems for several reasons. First, there remains a directional ambiguity in the fringe pattern. Second, the grating is both non-manhattan and irregularly spaced which makes it difficult to generate. Finally, the sensitivity is not much better than the finest line width drawn at the perimeter of the pattern. When the patterns are misaligned by this amount only two moiré fringes appear at opposite edges of the pattern.

Many alignment systems use diffraction gratings coupled with the moiré effect to create alignment systems. This is also referred to as interferometric or holographic alignment. A treatment of the topic of crossed diffraction gratings was presented by Guild in two books on the subject [76, 77]. Flanders described an early application of moiré diffraction to alignment [78]. Similar systems have been studied by Ishihara [79] and Itoh [80, 81] as well as by several others [82, 83, 84, 85, 86]. The ASM PAS 5500 stepper is one system that combines moiré fringes and diffraction techniques. It uses a set of four linear gratings as an alignment mark which appears both on the wafer and on the reticle. The patterns are overlaid and a photosensor detects the diffracted light as the wafer is moved. The resulting brightness curve as a function of distance has a maximum at the aligned position. The maximum is located by digitally processing the data and then the wafer stage is directed to the aligned position [15, 87, 88, 89, 90, 91]. A similar system was used on the ASM PAS 2500 stepper [92] and developed for the GCA 4800 stepper by Trutna [93]. A moiré fringe technique has also been proposed by Stevenson for flash-on-the-fly exposure alignment [94]. This method uses a checker board pattern on the wafer and a pair of orthogonal linear gratings on the reticle at a forty-five degree angle to the checkerboard pattern. A photosensor is mounted behind each grating and observes the moiré pattern as the checkerboard is moved. Peak detection on one of these signals obtains the position along the direction of motion and a phase comparison between the signals yields a sensitive measure of position in the direction transverse to motion. SVG (Perkin-Elmer) has used a moiré pattern formed between gratings on the mask and wafer to study overlay error [37].

An alignment system reported by Moel [95] uses an interesting combination of diffraction gratings and the moiré effect for a proximity printing x-ray alignment system. In this case two gratings with small and slightly different pitches (1.00 and 1.04 μm) are placed on top of each other. Rather than attempt to image these gratings directly, their zero order diffracted beams are imaged to detect the resulting moiré pattern. This moiré pattern has a 26 μm period that is easily imaged using a standard CCD camera. The signal is digitized and a computer is used to calculate the spatial phase (position) of the fundamental component of the moiré fringe.

2.4.6 Comparison of Alignment Systems

After reviewing the numerous alignment techniques and systems it is useful to see how some of them compare. Table 2-1 shows a summary of several reported alignment systems and lists a few of their figures of merit. The entries in this table were chosen because they

System	Method	Repeatability 3σ [μm]	Capture Region [μm]	Mark Size [μm]	Speed [ms]
Moel [95] 1993	Diffraction	0.018	1	150×50	200 / -
Chen [65] 1992	Diffraction	0.07	4	40×1000	10 / -
Suss [29] 1991	Edge Detection	0.15	20	60×60	200
Hampshire [42] 1989	Edge Detection	0.04	200	100×100	1000
SVG [38] 1987	Edge Detection	0.18	50	103×103	500 / -
Makosch [59] 1987	Diffraction	0.05	10	100×100	300
Cognex [1] 1987	Correlation	0.25 pixel	448 pixels	128×128 pixels	200
ASM [88] 1985	Diffraction	0.08	88	400×400	- / 300
Nikon [53] 1985	Diffraction	0.23	8	60×60	200 / 400
GCA [93] 1984	Diffraction	0.20	4	150×100	400
Electromask [31] 1981	Edge Detection	0.10	56	56×71	25 / 200
SVG [36] 1981	Edge Detection	0.29	40	$4 \times -$	300
Fay [60] 1979	Diffraction	0.05	3	35×35	500
Kleinknecht [57] 1979	Diffraction	0.10	540	270×270	1000
Wilczynski [50] 1979	Edge Detection	0.10	6	254×70	6 / -

Table 2-1: Specifications of some reported alignment systems.

reported alignment speed. Many publications on these systems omit this information because it is generally considered to be of secondary importance to repeatability. All of the entries in the table, except for Cognex, are for photolithography machines. Each report is identified by either its principal author or the manufacturer. This is followed by the main reference number and year of publication. Almost all of the alignment systems in the table use either edge detection or some version of moiré patterns and diffraction gratings for the alignment marks. These are by far the two most popular methods. The repeatability numbers are either alignment repeatability or overlay error for the systems. The capture region is the width over which the alignment can reliably be completed. This is usually small for grating techniques which can generally only capture errors up to the size of the grating pitch. The large capture regions reported by Kleinknecht and ASM are due to separate coarse alignment systems that initially align the gratings. The alignment mark sizes are also reported to give an idea of the area generally consumed by them. Some interesting figures of merit are the ratio of the repeatability to the capture region or to the mark's size. This can get as low as a few hundred parts per million. Finally, the alignment speed is listed in the last column. This has been broken down, where possible, into two components. The first is the time required to compute the alignment error and the second is the time required to mechanically position. This separation can only be made for systems which first compute the error and then move once to remove it. Other systems simply report a total alignment time.

2.5 Summary

Two types of alignment have been distinguished. Mark-to-mark alignment positions a target mark relative to a reticle mark using a separate sensor to determine their positions. Mark-to-sensor alignment positions a mark relative to the sensor which detects it. There are a variety of alignment methods used in industry. These include techniques based upon edge detection, image moments, correlation, moiré patterns, and diffraction gratings. Information published on the topic of automatic alignment primarily concerns the application to photolithography for integrated circuit manufacturing. A review of some of these systems reveals alignment times ranging from about 10 to 1000 ms. A significant fraction of this time is spent processing the alignment signal. Thus, a faster alignment sensor could increase the speed of these systems. Other applications which could benefit from a high speed alignment sensor include textile processing, color printing, and autonomous factory vehicles.

Surveyor's Mark Alignment Sensor

3.1 Introduction

One of the more commonly used alignment techniques is edge detection because it is both relatively easy to implement and because the required mark, such as a stripe or groove, is easy to make. Disadvantages of such techniques are that the capture regions are relatively small and the alignment is sensitive to defects in the edge being detected.

This chapter describes an alignment sensor based on edge detection. To extend the capture region, an initial coarse positioning scheme is explained that uses a four quadrant photodiode to detect the image of a surveyor's mark. Then a fine positioning algorithm that senses the edges of this mark is discussed. Circuits are presented which compute the error signal from the detected photocurrents. These are based on Gilbert multipliers and both normalize and amplify the photocurrents to generate the alignment error signal. Circuit bandwidth, sources of noise, and power dissipation are considered at the end of the chapter.

3.2 Alignment Method

The sensor uses photodiodes to detect an image of the surveyor's mark. This mark is a simple pattern consisting of a large square divided into four equal sized smaller squares with quadrants one and three colored white and two and four colored black. The layout of the photodiodes is shown in Figure 3-1 with a shaded surveyor's mark overlaying it. The sensor operates in two modes: an initial coarse positioning followed by a more accurate final alignment. The surveyor's mark was chosen because its light and dark quadrants can be used for coarse alignment and the edges separating them can be detected for fine positioning. In both modes of operation the sensor provides output signals measuring the misalignment between the mark's image and the center of the sensor.

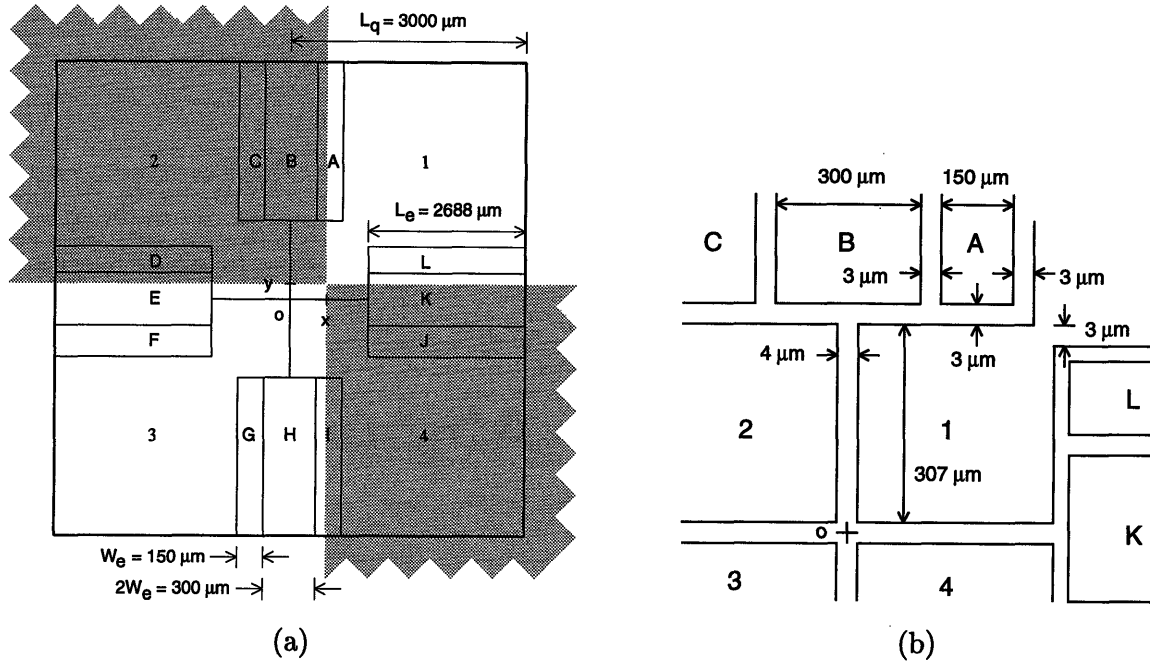


FIGURE 3-1: Layout of the photosensitive areas of the sensor (not to scale). The shaded area represents the misaligned image of a surveyor's mark. (a) Regions 1 through 4 measure brightness for coarse alignment and regions A through L detect edges for fine positioning. (b) A close up view at the center of the sensor showing the isolation gaps between photodiodes.

3.2.1 Coarse Alignment

For coarse alignment the sensor can be viewed approximately as a four quadrant photodetector as shown in Fig. 3-2. The signal collected in each quadrant will be proportional to the light irradiating that region

$$I_n = R \int_{Q_n} E(x, y) dA \quad n = 1, 2, 3, 4 \quad (3.1)$$

where I_n is the photocurrent generated in quadrant n , R is the photodiode responsivity, and $E(x, y)$ denotes image brightness (sensor irradiance). A measure of the alignment error can be formed by taking sums and differences of the quadrant signals

$$e_x = (I_3 - I_1) - (I_2 - I_4) \quad (3.2)$$

$$e_y = (I_3 - I_1) + (I_2 - I_4). \quad (3.3)$$

Figure 3-2 shows the case when the mark is misaligned to the sensor by an amount (x, y) . Assuming that the light regions have an irradiance of E_0 and the darkened regions have an irradiance of ϵE_0 ($\epsilon < 1$), one can show that

$$I_1 = RE_0 (L_q^2 - \rho L_q(x + y) + 2\rho xy) \quad (3.4)$$

$$I_2 = RE_0 (L_q^2 - \rho L_q(L_q - y)) \quad (3.5)$$

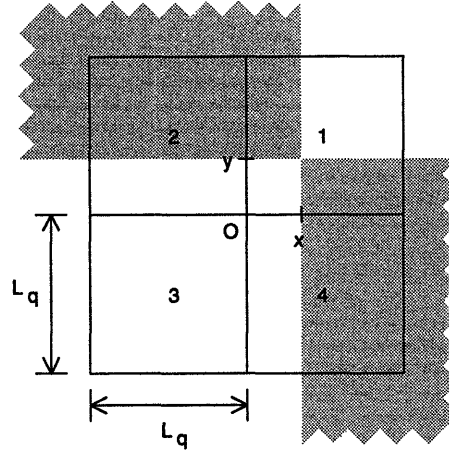


FIGURE 3-2: A quadrant photosensor with an offset image of the surveyor's mark superposed upon it.

$$I_3 = RE_0 L_q^2 \quad (3.6)$$

$$I_4 = RE_0 (L_q^2 - \rho L_q (L_q - x)) \quad (3.7)$$

for $0 < x, y < L_q$ with $\rho \triangleq 1 - \epsilon$. Similar equations hold for displacements into the other three quadrants. In general the error signals can be written as

$$e_x = 2\rho RE_0 (L_q - |y|)x \quad (3.8)$$

$$e_y = 2\rho RE_0 (L_q - |x|)y \quad (3.9)$$

for $-L_q < x, y < L_q$. Note that e_x has some dependence on y . The term $L_q - |y|$ distorts the error signal so that it is not exactly proportional to the misalignment. This distortion becomes evident near the edges of the sensor and especially near the corners where the error signal magnitude becomes small. For small displacements the distortion is less evident. Fig. 3-3 shows a plot of the negative of the error vector field which one might use in a control loop to correct the misalignment. A trajectory is also plotted from a simulation of a proportional feedback system using this error. The distortion in the error signals results in a trajectory that bends slightly.

The coarse alignment signal has a large capture area. If the center of the mark appears anywhere within sensor field of view, then the error signal can be used to drive it to the origin. However, the accuracy of the aligned position is subject to several errors. Mismatch in the photodiodes, illumination variations, and small rotations of the target will all affect the expected error signal.

Mismatch in the photodiodes is potentially a major concern. Given a uniform irradiance across the sensor, the resulting photocurrents from the quadrant photodiodes will have some mismatch. This can arise from small defects in the diodes as well as variations in the doping between them. One can model this as a mismatch in the photodiode responsivity and it will

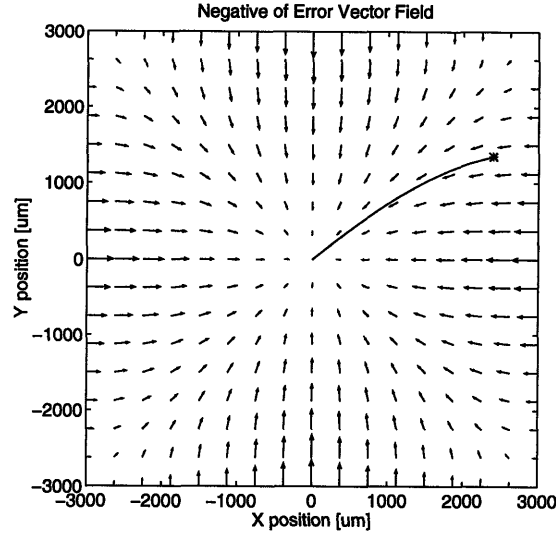


FIGURE 3-3: A plot of the negative of the error vector field for displacements of the mark across the sensor. Superposed upon this is a trajectory formed using proportional control from a starting point at $(2400 \mu\text{m}, 1350 \mu\text{m})$. Positions are in sensor coordinates.

result in an offset in the aligned position. For example, if the responsivity of the photodiodes is modeled as $R(1 + \delta_n)$, where δ_n represents the mismatch, then the null point will shift from the origin by an amount

$$\xi = \frac{\delta_3 - \delta_1}{2} L_q \quad (3.10)$$

in both the x - and y -directions (assuming $\rho = 1$). If the mismatch δ is modeled as a normally distributed random variable with zero mean and standard deviation σ_δ , then the coarse alignment offset will also be normally distributed with a standard deviation

$$\frac{\sigma_\xi}{2L_q} = \frac{\sqrt{2}}{4} \sigma_\delta \quad (3.11)$$

where the offset is expressed as a fraction of the sensor width $2L_q$.

Nonuniformity in the illumination of the mark, or in the irradiance of the sensor, is also a concern. Spatial variations in the target illumination will cause a corresponding nonuniformity in the radiated light from the target and in the irradiance of the sensor. Furthermore, in a lens imaging system there is a reduction in the irradiance at the image plane as one moves away from the optic axis even with uniform radiance from the target [22, 96]. Nonuniform illumination will also cause an offset in the null point of the error signal. If, for example, one assumes the sensor irradiance varies linearly with position as

$$E(x, y) = (1 + bx) m(x, y) \quad (3.12)$$

where b is the slope of the variation, and $m(x, y)$ denotes the image plane irradiance from the mark under uniform illumination, then the resulting offset in the null point will be

$$\xi = \frac{bL_q^2}{2} \quad (3.13)$$

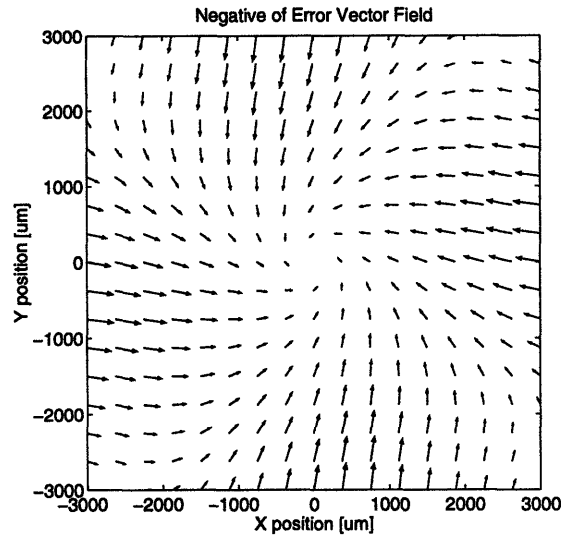


FIGURE 3-4: A plot of the negative of the error vector field when the target mark is rotated by 10 degrees relative to the sensor. Alignment trajectories will spiral into the origin.

in both the x - and y -directions. The quadrant sensor is thus quite sensitive to nonuniform irradiance at the image plane.

Small rotations of the image of the mark relative to the sensor are also important to consider. Because of the difficulty of explicitly finding an analytical solution for Equation 3.1 in this case, the error signals resulting from rotations of the mark were computed numerically. Figure 3-4 shows the result of this computation for a rotation of the target image by 10 degrees. Note that all of the error vectors have been rotated slightly as compared to Figure 3-3. The error vector field shows a curl that causes trajectories to spiral into the origin. As the rotation is increased this effect becomes more pronounced. At a rotation of forty-five degrees the trajectories become orbits. Rotations beyond this point cause the trajectories spiral outward rather than inward. Note that a rotation of the target only affects the path of the trajectory and does not cause inaccuracy in the position of the null point.

Other sources of inaccuracy arise from possible defects in the target mark itself and the required gap between the adjacent quadrant photodiodes. The effect of this gap is to enlarge the null point to a null region the size of the gap width. Theoretically this means that the mark cannot be aligned more accurately than the gap width.

3.2.2 Fine Alignment

One can improve the alignment to the surveyor's mark by modifying the sensor so that it detects the mark's edges in addition to balancing the brightness in quadrants. Edge detection is accomplished by the photosensitive areas labeled A through L in Fig. 3-1. By restricting one's attention to just the edges of the mark, the sensitivity of the alignment to many of the errors that plague the quadrant sensor is reduced. However, the total signal strength, which is

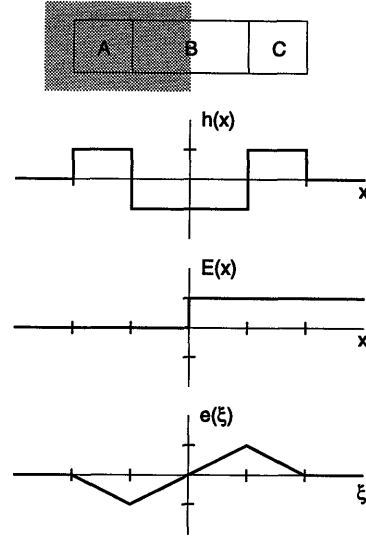


FIGURE 3-5: A fine alignment signal based on edge detection is created by adding the signals from regions A and C and subtracting the signal from region B . The response to a shifted impulse is shown as $h(x)$. The graph of $e(\xi)$ shows the error signal for an ideal brightness edge $E(x)$ shifted by ξ .

proportional to photodiode area, and the extent of the capture region are also reduced.

The regions labeled A , B , and C form a one-dimensional edge detector when their signals are combined as

$$e = I_A - I_B + I_C. \quad (3.14)$$

Region B is sized with twice the area of regions A and C to form an error signal similar to a discrete second difference. Given an irradiance function $E(x)$ across the detector, the error signal can be expressed as

$$e = \int_{-\infty}^{\infty} E(x)h(x) dx \quad (3.15)$$

where $h(\xi)$ is the response of the detector to a shifted impulse $\delta(x - \xi)$. If the irradiance function shifts by an amount ξ , then the error can be expressed as a correlation integral

$$e(\xi) = \int_{-\infty}^{\infty} E(x - \xi)h(x) dx. \quad (3.16)$$

The resulting error for an ideal step edge is illustrated in Fig. 3-5. Note that the error signal is zero when the edge lies midway across photodiode B .

The four edges in the surveyor's mark are detected this way. Signals from the edge detecting regions are combined to form error signals as follows

$$e_x = (I_A - I_B + I_C) - (I_G - I_H + I_I) \quad (3.17)$$

$$e_y = (I_L - I_K + I_J) - (I_D - I_E + I_F) \quad (3.18)$$

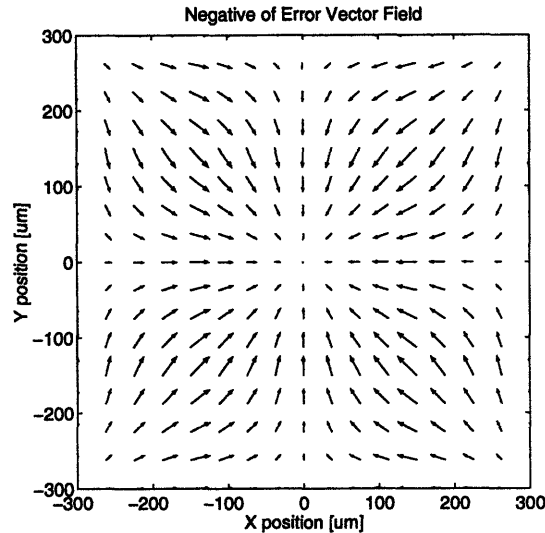


FIGURE 3-6: A plot of the negative of the fine alignment error vector field from the edge detector signal. The vectors point radially inward inside the region ($\pm 150 \mu\text{m}$, $\pm 150 \mu\text{m}$).

The resulting error signal for shifts in the range $-W_e < (\xi, \eta) < W_e$ is

$$e_x = \begin{cases} 2\rho RE_0 L_e (2W_e - \xi) & W_e < \xi < 2W_e \\ 2\rho RE_0 L_e \xi & -W_e < \xi < W_e \\ -2\rho RE_0 L_e (2W_e + \xi) & -2W_e < \xi < -W_e \\ 0 & \text{otherwise} \end{cases} \quad (3.19)$$

$$e_y = \begin{cases} 2\rho RE_0 L_e (2W_e - \eta) & W_e < \eta < 2W_e \\ 2\rho RE_0 L_e \eta & -W_e < \eta < W_e \\ -2\rho RE_0 L_e (2W_e + \eta) & -2W_e < \eta < -W_e \\ 0 & \text{otherwise} \end{cases} \quad (3.20)$$

again assuming the light regions have brightness E_0 and the dark regions have brightness ϵE_0 with $\rho \triangleq 1 - \epsilon$. The resulting error vector field over this region is shown in Figure 3-6. Note that the error vectors within $\pm W_e$ of the origin are directed radially inward.

As with the quadrant error signal, one must consider the effects of mismatch in the photodiodes. This edge detection scheme relies not only upon matching the photodiodes A and C shown in Figure 3-5, but also making the area of photodiode B exactly twice as large. Because photodiodes A and C are identical, the mismatch between them can reasonably be assumed to be zero mean, i.e. having no systematic mismatch. However, because photodiode B has a different geometry a significant mismatch with nonzero mean can be expected. The effect of the mismatch between these photodiodes will again be approximated as a mismatch in their responsivities. For example, the photodiode A will be modeled with a responsivity of $R(1 + \zeta_A)$ where ζ will be considered a random variable with normal distribution and zero mean. Similar responsivities will be used for the other identical photodiodes. On the other hand, photodiode

B will be modeled with a responsivity $R(1 + \gamma_B)$ where the random variable γ will be considered normally distributed with nonzero mean to reflect the expected systematic mismatch. The resulting x -direction error signal for small shifts ξ will be

$$e_x(\xi) = RE_0L_e \left\{ 2\xi \left(1 + \frac{1}{2}(\gamma_B + \gamma_H) \right) - W_e(\gamma_B - \gamma_H + \zeta_G - \zeta_A) \right\} \quad (3.21)$$

assuming that $\rho = 1$. Solving for the offset in the null point one finds

$$\xi = \frac{\gamma_B - \gamma_H + \zeta_G - \zeta_A}{2(1 + \frac{1}{2}(\gamma_B + \gamma_H))} W_e. \quad (3.22)$$

If one assumes that $m_\gamma, \sigma_\gamma \ll 1$, then this simplifies to

$$\xi \approx \frac{1}{2}(\gamma_B - \gamma_H + \zeta_G - \zeta_A) W_e \quad (3.23)$$

which is a sum of random variables with mean $m_\xi = 0$ and variance $\sigma_\xi^2 = (\sigma_\gamma^2 + \sigma_\zeta^2) W_e^2 / 2$. Note that the systematic mismatch in the regions B and H is cancelled because the error signal depends on the difference of these photocurrents. The standard deviation of the offset expressed as a fraction of the total sensor width is approximately

$$\frac{\sigma_\xi}{2L_q} \approx \frac{\sqrt{2}}{4} \frac{W_e}{L_q} \sqrt{\sigma_\gamma^2 + \sigma_\zeta^2}. \quad (3.24)$$

Nonuniformity of illumination will also affect the accuracy of the edge detector. However, because the edge detection photodiodes only occupy a small portion of the total sensor area, their signal is proportionately less sensitive to this source of error. If one models a nonuniform illumination as was done for the quadrant sensor (Equation 3.12), then the x -direction error signal becomes

$$e_x = 2RE_0L_e \left\{ \xi + \frac{b\xi^2}{2} + bW_e^2 \right\} \quad (3.25)$$

for $-W_e < \xi < W_e$ and $\rho = 1$. Under the assumptions that the quadratic term in ξ can be neglected and that $bL_q \ll 1$ one finds an offset of

$$\xi \approx -bW_e^2. \quad (3.26)$$

Note that compared to the result for the quadrant sensor in Equation 3.13 the offset has been reduced by the factor $(L_q/W_e)^2/2$ which is equal to 200 for the dimensions used in the sensor design. Furthermore, if one assumes a linear variation in illumination along the y -direction similar to Equation 3.12, there is no resulting offset in the null point for the x -direction error signal. The edge detector is only sensitive to linear variations in illumination which are transverse to the edge and this sensitivity is dramatically lessened compared to the quadrant sensor.

The fine alignment signal is much more sensitive to rotations of the mark due to the long and thin shape of the edge detecting photodiodes. A small rotation can easily bring the mark's edges

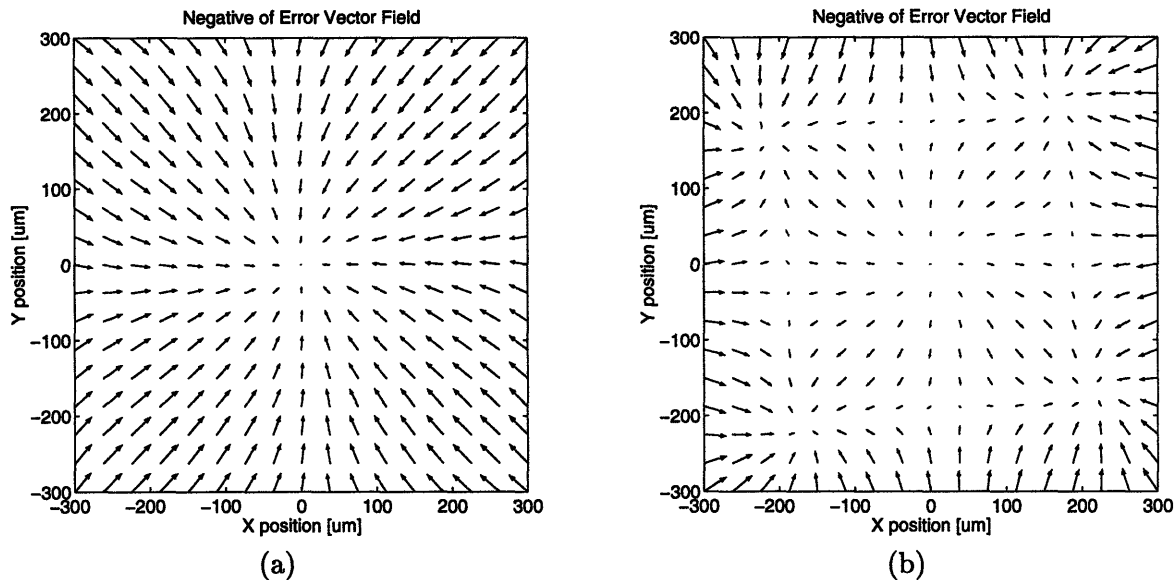


FIGURE 3-7: Negative error vector field for the fine alignment signal with the target rotated by (a) four degrees and (b) seven degrees.

almost entirely outside the field of view of these photodiodes. As was the case for the quadrant sensor, the effect of rotations was studied numerically because of the difficulty in obtaining simple explicit solutions. Figure 3-7 shows the error vectors for two rotations of the mark's image relative to the sensor. Small rotations reduce the strength of the error signal, enlarge the capture region, and add a slight curl to the error vector field. The stable equilibrium point at the origin remains unchanged until a critical angle is reached. Rotation beyond this point causes the equilibrium point at the origin to become unstable and four new stable equilibrium points appear, one in each quadrant. A formula for this angle in terms of the sensor dimensions is

$$\theta_c = \arctan \left(\frac{1}{1 + \frac{L_q}{2W_e}} \right). \quad (3.27)$$

For the sensor dimensions implemented this angle works out to be 5.2 degrees. Figure 3-7(b) shows the effect on the error vector field of a rotation beyond this critical angle. The unstable equilibrium point at the origin as well as the four new stable equilibrium points are evident.

3.2.3 Combining the Two Signals

Although the sensitivity of the fine alignment signal is superior, its capture range is much smaller than the coarse alignment signal. The fine alignment signal is nonzero only when an edge lies within $\pm 2W_e$ of the origin. In order to obtain the benefit of both a large capture region and accurate final positioning, a two step alignment is done. Initial positioning is accomplished using the quadrant signal followed by fine positioning using the edge detecting photodiodes. Thus it is important that the null point of the coarse alignment falls within the capture region

of the fine alignment. This is achieved by adding the signals from the adjacent edge detector areas to the quadrant signals during the coarse alignment step (e.g. signals from regions A and L are combined with region 1).

A design trade-off must be made in the choice of the size of the edge detection region relative to the total sensor area. The ratio W_e/L_q should be small to maximize the sensitivity of the fine alignment signal to displacements of the mark. When this ratio is small it also reduces the error caused by nonuniform illumination. However, such a choice also has repercussions. First, reducing this ratio makes the capture region of the sensor smaller. The capture region must be kept large enough to encompass the null point of the coarse alignment signal after accounting for offsets from photodiode mismatch and nonuniform illumination. Second, the sensor becomes less tolerant of rotations as this ratio is lowered. Finally, a smaller area for the edge detecting photodiodes reduces the resulting photocurrents which lowers the signal-to-noise ratio.

The ratio $W_e/L_q = 0.05$ was chosen for the fabricated design in order to maintain a capture region that could accommodate ten percent worst case mismatch in the signals from the quadrant sensor. Such mismatch will arise from a combination of photodiode mismatch and nonuniform illumination. The choice was a conservative one and was made because there was little information on the expected mismatch between the photodiodes and the expense involved to guarantee good illumination uniformity in the optical set-up.

3.3 Current Amplification and Normalization

The alignment error signals described previously are derived from sums and differences of the currents collected by the photodiodes. Unfortunately these currents are quite small and are proportional to the incident illumination of the target mark. Typically they can range from a few hundred nanoamps down to hundreds of picoamps for a dimly lit target. Such small currents are difficult to work with and an alignment signal proportional to light level is also not desirable. For these reasons current amplifiers were integrated with the sensor to amplify the photocurrents and normalize them such that the output signal is only proportional to the misalignment and not to the amount of illumination. It was also desired that the amplifiers have a low offset and good linearity in order to preserve the alignment signal and a high bandwidth in order to obtain a fast alignment sensor.

Because the photocurrents are single-ended they are easy to add, but difficult to subtract, accurately. For this reason, a differential amplifier is used. The output error signal is taken as a differential output current so that one only needs to sum photocurrents in order to compute it. Common-mode and differential-mode signals will be defined in terms of the single-ended values as

$$\begin{aligned} I_p &= I_{cm} + \frac{1}{2}I_{dm} & I_{cm} &= \frac{1}{2}(I_p + I_m) \\ I_m &= I_{cm} - \frac{1}{2}I_{dm} & I_{dm} &= I_p - I_m \end{aligned} \tag{3.28}$$

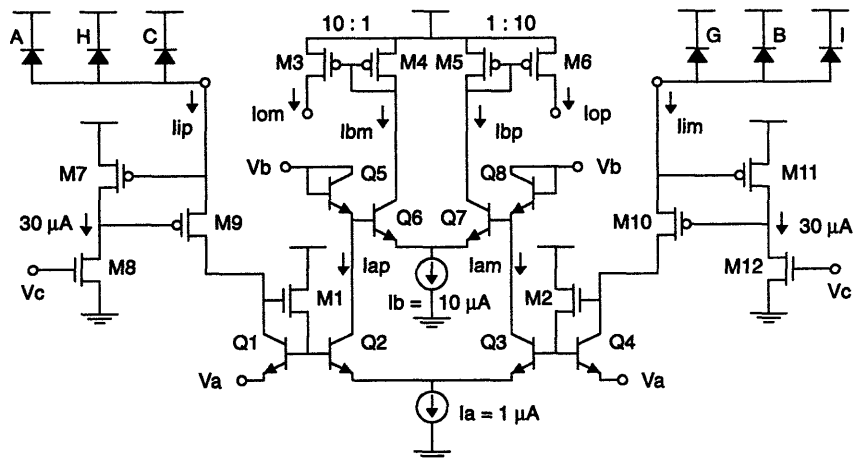


FIGURE 3-8: Differential current amplifier which normalizes and amplifies the detected photocurrents.

where I_p and I_m are single-ended currents, I_{cm} is the common-mode current, and I_{dm} is the differential-mode current.

The entire amplifier for the fine alignment edge detector signal appears in Fig. 3-8. It is shown connected to the six photodiodes used to generate the edge detector error signal e_x of Equation 3.18. The circuit is a differential current amplifier composed of three stages. Transistors M7 through M12 are active cascodes which hold the voltages across the photodiodes at a fixed value in order to improve the dynamic performance. Bipolar transistors Q1 through Q4 form a normalizing, or “type A¹,” multiplier at the input [100]. The MOSFETs M1 and M2 buffer the low current inputs and supply base currents to the bipolar devices. The first stage amplifies and normalizes the input currents to a fixed common-mode level of $0.5 \mu\text{A}$. The second stage amplifier consists of transistors Q5 through Q8 which form an inverting, or “type B²,” multiplier with a current gain of about eight [100]. A final gain of ten is provided by the current mirror M3 through M6. The output signals from the amplifier are differential source currents with a fixed common-mode level of $50 \mu\text{A}$ independent of the image light level.

3.3.1 Normalizing (Type A) Current Multiplier

The Gilbert normalizing (type A) current multiplier is used both to provide gain and to normalize the input signal currents from the photodiodes. The basic form of the circuit is shown in Figure 3-9. Using the translinear principle one can show that

$$I_{op} = \frac{I_a}{I_{im} + I_{ip}} I_{ip} \tag{3.29}$$

$$I_{om} = \frac{I_a}{I_{im} + I_{ip}} I_{im} \tag{3.30}$$

¹The type A amplifier has previously been referred to as a “type B” or “normal” configuration [97, 98, 99].

²The type B amplifier has previously been referred to as a “type A” or “inverted” configuration [97, 98, 99].

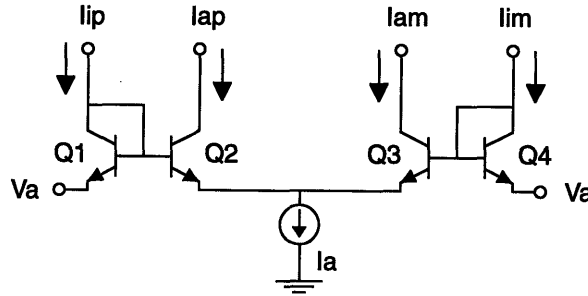


FIGURE 3-9: Gilbert normalizing multiplier (type A form).

assuming the transistors have infinite current gain (β). A useful way to view this circuit is to note that it multiplies the differential-mode input signal by the ratio of the common-mode output to the common mode of the input

$$I_{adm} = G_a I_{idm} \quad (3.31)$$

where

$$G_a \triangleq \frac{I_{acm}}{I_{icm}}. \quad (3.32)$$

Thus the ratio of the differential-mode to common-mode signals is preserved from input to output. By fixing the output common mode to $I_a/2$ the circuit provides a normalized differential output. Spatially uniform changes in the input light irradiance will scale the resulting photocurrents without changing their ratios. This circuit preserves that differential signal without passing through the change in the common-mode signal. Thus the alignment sensor can be used at a variety of illumination levels without having to set any gains, integration times, or aperture stops.

Equation 3.31 implies that the normalizing multiplier can realize any gain set by the ratio of the common-mode input to the common-mode output. However, the effect of finite current gain in the devices places an upper limit of β on the achievable gain and also multiplies Equation 3.31 by the factor $\alpha \triangleq \beta/(\beta + 1)$. Beyond this point all of the input current is diverted from Q1 and Q4 to supply the bases of Q2 and Q3. Surprisingly, Equation 3.31 holds right up to the point where all of the input current is diverted to the bases of the output pair. If the input current level drops below this point, then the normalizing gain G_a in Equation 3.31 is replaced by the fixed gain of β . The operation of the circuit is extended to lower input current levels by replacing the diode connection of the input bipolar devices with MOSFET source followers (M1 and M2 in Figure 3-8). This improved circuit is capable of operation with extremely low input currents.

The combination of bipolar transistors and MOSFETs in this circuit necessitates a BiCMOS process for this sensor. This choice was made for several reasons. First, as indicated above, in order for the amplifier to operate at low currents, a large impedance buffer is required to supply the base currents. The MOSFET devices could be replaced with bipolar transistors, or

Darlington pairs, which would extend the range of operation. However, input currents would still be required for the buffering bipolar devices. Simulations also revealed that using a bipolar device resulted in more nonlinearity at low current levels. This is likely due to the low (and changing) value of β at extremely low currents. Another possibility is to replace the bipolar devices with MOSFETs operating in subthreshold. Such a choice suffers from two problems. First, the devices would have to be made very wide to maintain deep subthreshold operation at currents up to 1 μA . Second, in subthreshold the matching between devices is very poor which leads to a large offset in the amplifier. For these reasons a BiCMOS implementation was chosen because the bipolar devices provide the desired exponential relation along with better matching at low currents whereas the MOSFETs are used to make a high impedance buffer for supplying the base currents.

The dynamic performance of the normalizing multiplier is important to understand because it determines the time response of the entire amplifier. The basic circuit shown in Figure 3-9 is dominated by a single pole from the input capacitance multiplied by the input resistance

$$R_1 = \frac{1}{g_{mQ1}} \quad (3.33)$$

$$C_1 = C_{cs1} + C_{\pi1} + C_{\pi2} + C_{\mu2} \quad (3.34)$$

$$\tau_1 = \frac{kT}{q} \frac{1}{I_{icm}} (C_{cs1} + C_{\pi1} + C_{\pi2} + C_{\mu2}). \quad (3.35)$$

Note that the bandwidth is proportional to the input current I_{icm} . For example, taking $I_{acm} = 500$ nA, $I_{icm} = 5$ nA, and $C_1 = 2$ pF results in a bandwidth of about 15 kHz. If I_{icm} is raised to 50 nA the bandwidth proportionately increases to 150 kHz. This dependence of bandwidth on the common-mode level of the input signal is important to remember. I_{icm} will be determined by the amount of photocurrent generated by the photodiodes and is thus directly proportional to the sensor irradiance. This means a brighter image will result in faster performance.

The frequency response of the enhanced normalizing multiplier with the MOSFET source followers is more complicated than the basic circuit. A small-signal model of the half circuit consisting of Q1, Q2, and M1 reveals that the circuit dynamics are dominated by two closely spaced poles and a zero. The transfer function is

$$\frac{i_o}{i_i} = G_a \frac{1 + \tau_{za}}{\left(1 + \frac{2\zeta}{\omega_n} s + \frac{1}{\omega_n^2} s^2\right)} \quad (3.36)$$

where

$$\omega_n \triangleq \sqrt{\frac{g_{mQ1} g_{mM1}}{K_C}} \quad (3.37)$$

$$\zeta \triangleq \frac{1}{2} \sqrt{\frac{G_a + 1}{n\beta K_C}} \left(C_{a1}(1+n) + nC_{a2} \left(1 + \frac{\beta}{G_a + 1}\right) \right) \quad (3.38)$$

$$\tau_{za} \triangleq \frac{1}{g_{mM1}} C_{a2} \quad (3.39)$$

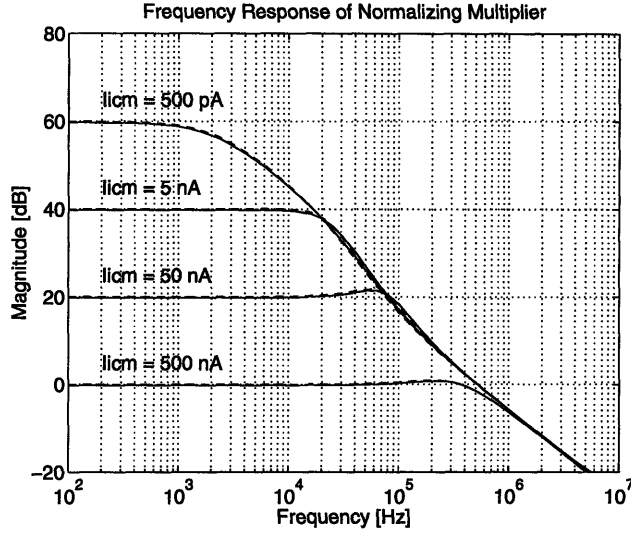


FIGURE 3-10: Frequency response of the normalizing multiplier at various I_{icm} . The transfer function consists of two moderately damped poles followed by a high frequency zero. Solid lines are from simulations; dashed lines are the analytical result.

and

$$K_C \triangleq C_{a1}C_{a2} + C_{a1}C_{a3} + C_{a2}C_{a3} \quad (3.40)$$

$$C_{a1} \triangleq C_{cs1} + C_{gdM1} \quad (3.41)$$

$$C_{a2} \triangleq C_{\mu1} + C_{gsM1} \quad (3.42)$$

$$C_{a3} \triangleq C_{\pi1} + C_{\pi2} + C_{\mu2}. \quad (3.43)$$

The frequency response is plotted in Figure 3-10. Over the range of input current levels shown the damping coefficient of the two poles ranges from about 0.6 to 2.0. Slight peaking due to the complex pole pair can be seen in the response. Note that the bandwidths of the enhanced multiplier are comparable to the basic version of the circuit.

3.3.2 Inverting (Type B) Current Multiplier

The second stage of current amplification is provided by transistors Q5–Q8 configured as a Gilbert inverting (type B) current multiplier. The operation of this circuit is similar to that of the preceding normalizing multiplier. However, the effect of finite β is much more of a problem with this configuration. Figure 3-11 shows the basic topology of this circuit. Note that the output current I_{bp} appears at the opposite side from the input current I_{ap} , hence the appellation “inverting”. An analysis, including finite β , of this circuit yields the following current relationships

$$I_{bp} = K \frac{I_b}{(I_{ap} + I_{am})} \left(I_{ap} + \frac{I_b}{\beta + 1} \right) \quad (3.44)$$

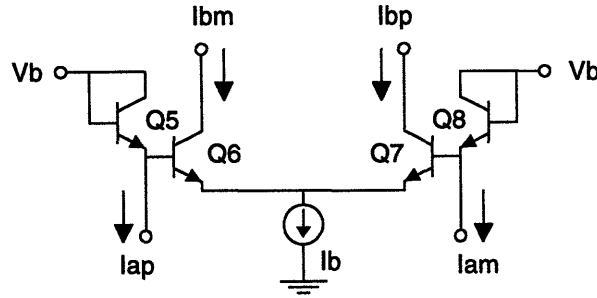


FIGURE 3-11: Gilbert inverting multiplier (type B form).

$$I_{bm} = K \frac{I_b}{(I_{ap} + I_{am})} \left(I_{am} + \frac{I_b}{\beta + 1} \right) \quad (3.45)$$

where

$$K \triangleq \frac{\alpha}{1 + \frac{2I_b}{(\beta + 1)(I_{ap} + I_{am})}}. \quad (3.46)$$

In terms of differential-mode and common-mode currents this can be expressed as

$$I_{bdm} = G_b I_{adm} \quad (3.47)$$

where

$$G_b \triangleq K \frac{I_{bcm}}{I_{acm}}. \quad (3.48)$$

Note that the differential gain is set by the ratio of the common-mode levels, as was the case for the normalizing multiplier, with an additional attenuation factor K . The attenuation factor decreases as the common-mode gain is increased and becomes significant as the input currents become comparable to the base currents of Q6 and Q7. The input currents then have little effect on the V_{be} 's of transistors Q5 and Q8 which are primarily supplying base currents to Q6 and Q7 at this point. This results in lowered differential gain. For example, in the complete amplifier I_{bcm} is simply $I_a/2$ from the output of the normalizing multiplier. With $I_a = 1 \mu\text{A}$, $I_b = 10 \mu\text{A}$, and $\beta = 100$ the attenuation factor K is 0.83. This implies that a 100 percent modulation of the input results in an 83 percent modulation of the output. Note that this modulation is weakly dependent on the transistor β for the process. Because of the attenuation factor, this amplifier is unsuitable for large current gains and is only used to provide a fixed current gain of about eight in this design.

The frequency response of the inverting multiplier is dominated by a single pole followed by a higher frequency zero. The pole is given by the product of the input resistance and the input capacitance

$$R_5 = \frac{1}{g_{mQ5}} || r_{\pi 5} || r_{\pi 6} \quad (3.49)$$

$$C_5 = C_{\pi 5} + C_{\pi 6} + C_{\mu 6}. \quad (3.50)$$

Simulations of the circuit predict $C_5 = 1$ pF and $R_5 = 43$ k Ω yielding a bandwidth of 2 MHz. The zero in the transfer function occurs at

$$R_{zb} = \frac{1}{g_{mQ6}} \quad (3.51)$$

$$C_{zb} = C_{\mu6}. \quad (3.52)$$

With $R_{zb} = 5.2$ k Ω and $C_{zb} = 280$ fF the zero is positioned at 110 MHz, well above the pole and beyond the f_T of transistor Q6. At frequencies above f_T the hybrid-pi small-signal model of the transistor is no longer valid. Thus, the existence of this zero is suspect. Note that this stage is considerably faster than the preceding normalizing multiplier due to the higher current levels.

3.3.3 MOS Current Mirror

The final stage of current amplification is accomplished with a PMOS current mirror using devices M3 through M6 in Figure 3-8. This mirror provides both current gain and output source currents with a reasonable amount of voltage compliance. The input devices M4 and M5 are sized long with a W/L of $20 \mu\text{m}/56 \mu\text{m}$ to obtain a large gate drive and thus reduce the effect of mismatch in the threshold voltages. The output devices M3 and M6 are sized to be ten times wider than the input devices to provide a current gain of $G_p = 10$.

The dynamics of this mirror are dominated by a single pole formed from the input resistance and the large gate capacitance

$$R_p = \frac{1}{g_{mM4}} \quad (3.53)$$

$$C_p = C_{gsM4} + C_{gsM3}. \quad (3.54)$$

Simulations estimate these values to be $R_p = 150$ k Ω and $C_p = 8.3$ pF resulting in a pole τ_p at about 130 kHz. Note that this pole is independent of the image irradiance and is usually above the limiting poles of the normalizing multiplier except under very bright conditions. This mirror also has a zero in its transfer function that is substantially above the pole. Neither the pole nor the zero from this mirror will limit the amplifier bandwidth.

3.3.4 Complete Amplifier

The performance of the complete amplifier from input to output can be expressed with the following transfer function

$$\frac{i_o}{i_i} = G_{dc} \frac{1 + \tau_{za}}{\left(1 + \frac{2\zeta}{\omega_n} s + \frac{1}{\omega_n^2} s^2\right) (1 + \tau_p s)} \quad (3.55)$$

where

$$G_{dc} = G_a G_b G_p \quad (3.56)$$

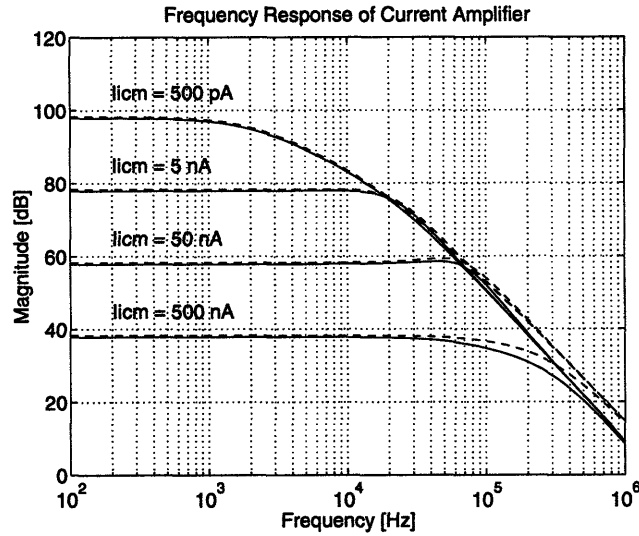


FIGURE 3-12: Frequency response of the complete current amplifier.

and the other parameters are as defined previously. The unimportant high frequency poles and zeros are omitted from this transfer function. A plot of the simulated frequency response is shown in Figure 3-12. The dashed lines show the response modeled by Equation 3.55. The slight divergence above 100 kHz is due to the unmodeled loading effects of the input impedance of the inverting multiplier interacting with C_μ of devices Q2 and Q3. This frequency response shows the performance one can obtain from the unloaded amplifier.

The actual dynamics of the current amplification are substantially altered by the addition of the input photodiodes which have a sizeable depletion region capacitance. The edge detecting photodiodes A, H, and C shown in Figure 3-8 have a total capacitance of about 260 pF and the quadrant photodiodes Q1 through Q4 each have a capacitance of 1.1 nF. The effect of this capacitance is to increase C_{a1} substantially in Equation 3.36 and split the pole pair into two widely spaced poles. The dominant pole becomes the input capacitance multiplied by the input resistance of the normalizing multiplier

$$R_{a1} = (1 + n) \frac{1}{g_{mQ1}} \quad (3.57)$$

$$C_{a1} = C_d + C_{cs1} + C_{gdM1} \approx C_d \quad (3.58)$$

$$\tau_{a1} \approx \frac{kT}{q} \frac{1}{I_{icm}} (1 + n) C_d \quad (3.59)$$

with the second pole moving out to

$$\tau_{a2} = \frac{n}{1 + n} \frac{\beta}{G_a + 1} (C_{a2} + C_{a3}) \quad (3.60)$$

The dominant pole can be clearly seen in Figure 3-13. The bandwidth of the sensor becomes directly proportional to the input current level. The input current level in turn can be expressed in terms of the image irradiance E_0 multiplied by the responsivity R to yield the following

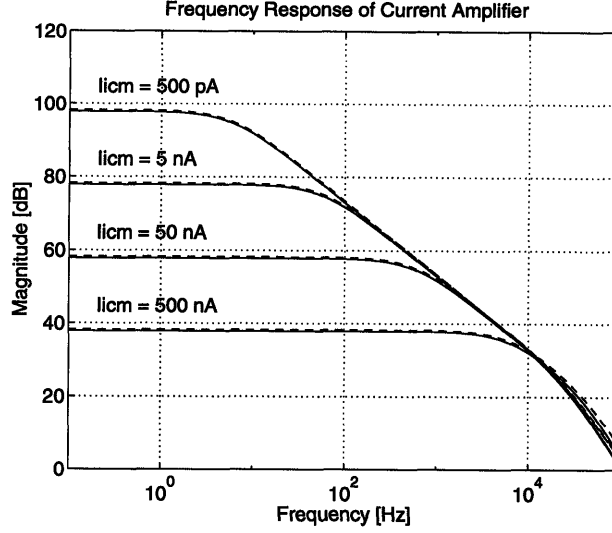


FIGURE 3-13: Frequency response of the current amplifier loaded with input photodiodes.

expression for the bandwidth

$$f_h = \frac{RE_0}{2\pi(1+n)C_{da} \frac{kT}{q}} \quad (3.61)$$

where C_{da} is the photodiode depletion capacitance per unit area. Assuming $E_0 = 10 \mu\text{W}/\text{cm}^2$, $R = 0.13 \text{ A}/\text{W}$, $C_{da} = 15 \text{ nF}/\text{cm}^2$, and $n = 1.2$ results in a bandwidth of 240 Hz. This is an improvement of a factor of four over video rate. Increasing the image irradiance by using stronger illumination or a larger lens will proportionately increase the speed. It is not unreasonable to increase the sensor irradiance by a factor of 10 to 100.

3.3.5 Active Cascode

The chief obstacle to faster operation is the photodiode capacitance. For the illumination conditions assumed above, the edge detector current would be 21 nA. At this signal level the current amplifier would have approximately a 40 kHz bandwidth (see Figure 3-12) if it were not burdened by such a large input capacitance. Regulating the reverse bias across the photodiode and buffering its capacitance from the current amplifier can dramatically increase the bandwidth.

This regulation is provided by the active cascode circuit formed by the devices M7 through M12 in Figure 3-8. Devices M7 and M8 comprise an analog inverter whose output controls the gate voltage of the cascode device M9. The active feedback adjusts the gate drive on M9 to maintain a fixed voltage at the input node. The voltage variation across the photodiode capacitance is reduced by the gain of the inverter. Another way to view this is to calculate the small-signal input resistance of the cascode

$$R_c = \frac{1}{A+1} \frac{1}{g_{mM9}} \approx \frac{n}{AI_{icm}} \frac{kT}{q} \quad (3.62)$$

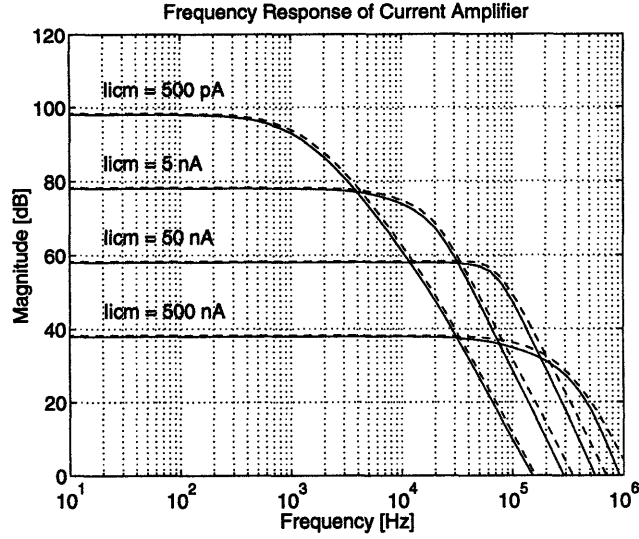


FIGURE 3-14: Frequency response of the current amplifier with the active cascode at the input to buffer the photodiode capacitance.

where $A \triangleq g_{mM7}(r_{oM7}||r_{oM8})$ is the inverter gain. The input resistance is lowered approximately by the gain of the inverter (compare with Equation 3.57). The photodiode capacitance is increased slightly to $C_{da} = 21 \text{ nF/cm}^2$ by the reduction in bias across it. Overall, the pole due to the input capacitance is pushed out to the point where it is no longer clearly dominant. The transfer function from input to output current now becomes

$$\frac{i_o}{i_p} = G_{dc} \frac{1 + \tau_{za}}{(1 + \frac{2\zeta}{\omega_n}s + \frac{1}{\omega_n^2}s^2)(1 + \tau_c s)(1 + \tau_p s)} \quad (3.63)$$

where $\tau_c \triangleq R_c C_d$ and all other parameters remain the same as defined previously. A plot of the frequency response is shown in Figure 3-14. The boost in the bandwidth is roughly a factor of 120 for the device sizes chosen in the implementation.

3.3.6 Coarse Alignment Amplifier

The equations for the coarse alignment error signal require some modifications to the amplifier. Both the x and y error signals are created by summing the same four signal currents. This implies that these currents must be copied before the two error signals can be formed. One possible solution is to copy the photocurrents in a current mirror before normalizing and amplifying them. Unfortunately this would transform them from source currents into sink currents which are unsuitable as inputs to the normalizing multiplier with NPN transistors. Furthermore, both the mirroring and normalization would be done with small photocurrents which is difficult to do accurately and quickly. Instead, the photocurrents are normalized and amplified first, and then copied at the output current mirror in order to form the error signal currents. This requires a multiplier that is capable of normalizing four input currents rather

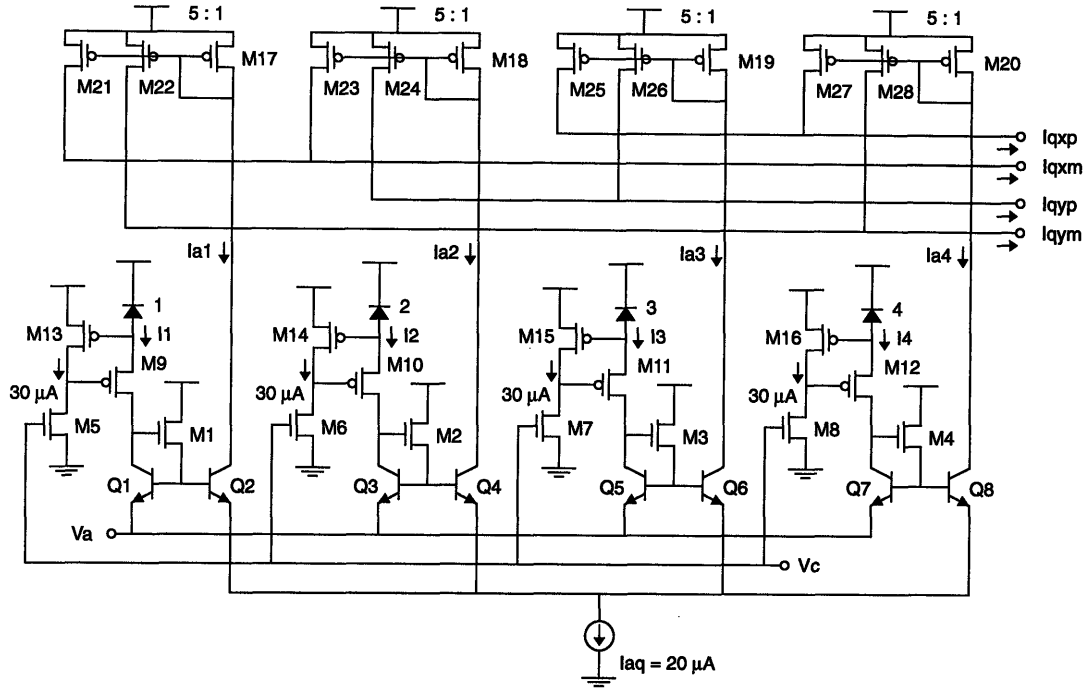


FIGURE 3-15: Normalizing current amplifier with four inputs which is used for the coarse alignment error signal.

than two as was done for the fine alignment signal.

Figure 3-15 shows the complete current amplifier for the coarse alignment signal. The Gilbert normalizing multiplier used for two input currents is extended for four input currents [101]. As in the two input case, the four input circuit relates the output currents to the input currents as

$$I_{an} = \frac{I_{ocm}}{I_{icm}} I_{in} \quad n = 1, 2, 3, 4 \quad (3.64)$$

where now

$$I_{icm} = \frac{1}{4}(I_1 + I_2 + I_3 + I_4) \quad (3.65)$$

$$I_{acm} = \frac{1}{4}I_{aq}. \quad (3.66)$$

The four input normalization circuit preserves the ratios between the input and output currents while at the same time fixing the output common-mode level. It thus rejects variations in the input common-mode signal.

The output from the normalizing amplifier goes directly to the current mirrors. The intermediate gain stage formed by the inverting amplifier is omitted because its multiple input extension is a normalizing divider (inverter) and is incompatible with the desired amplification. In order to maintain the same size input current to the mirror implies that the output devices of the normalizing amplifier must be run at a higher current level than is the case for the fine alignment amplifier. The higher output current of the normalizing multiplier might be expected

to reduce its bandwidth because the larger currents in the output transistors increases the base diffusion capacitance. However, at these low current levels this capacitance is dwarfed by the base-emitter depletion capacitance. Hence the dynamic behavior of this circuit is little affected by the increase in gain. Note that this implies that the intermediate gain stage used in the fine alignment amplifier is probably not needed and can be omitted by running the input stage at higher gain.

3.3.7 Amplifier Biasing

The current amplifiers require both bias currents and voltages. The voltages are provided from off the chip although these could be generated on the chip if desired. The bias currents were also generated from an off-chip current that is mirrored and copied on the chip. An NPN simple current mirror consisting of one input device and three output devices provides the bias currents for the input normalizing multipliers. Two matched bias currents are needed for the x - and y -axis current amplifiers. The third output device is not used but allows the mirror to be laid out in a common centroid configuration to achieve better matching. One of the nice properties of the normalizing multiplier is that the voltage at the emitters of the output devices does not move as long as the input common mode remains fixed. Thus a simple current mirror with moderate output impedance can be used to provide the bias current. On the other hand, this voltage varies in the inverting multiplier as the differential input changes. For this reason a cascoded current source with higher output impedance was used to prevent this variation from causing additional nonlinearity in the amplifier.

3.3.8 Noise

Before examining the noise arising from the complete sensor it is useful to look at the amplifier alone with the active cascode circuit removed. Figure 3-16 shows the results of HSPICE simulations of the amplifier noise spectrum at several input current levels. The spectrum can be divided into two regions. At low frequencies, below the bandwidth cutoff of the sensor, the noise spectrum is flat. Here the shot noise of the Q1 and Q4 collector currents are the dominant noise sources which result in a differential output voltage of

$$v_o^2 = 4qI_{icm}G_{dc}^2R_L^2 \quad (3.67)$$

when the error currents are dropped across load resistors R_L . The figure shows the result for $R_L = 39.1 \text{ k}\Omega$. An additional source of shot noise comes from the collector currents of devices Q2, Q3, Q5, and Q8 bringing the total noise at the output to

$$v_o^2 = 4qI_{icm}\left(1 + \frac{4I_{icm}}{I_a}\right)G_{dc}^2R_L^2. \quad (3.68)$$

This equation describes the low frequency plateaus in the noise spectrum. In the implementation one must also account for the additional shot noise from the photodiodes. The resulting noise

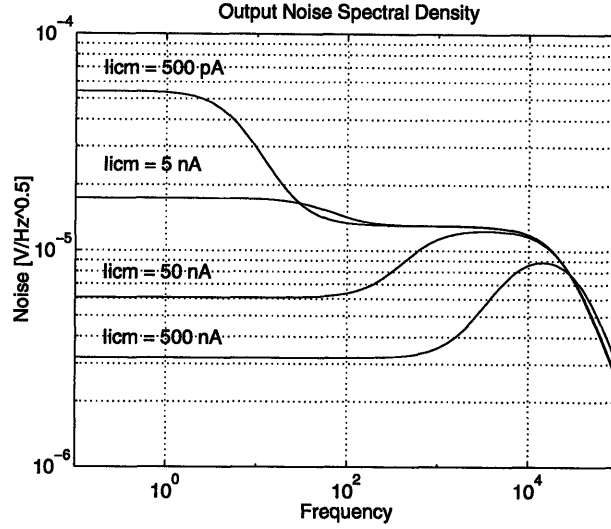


FIGURE 3-16: Simulation results for the output noise spectrum from the amplifier without the active cascode circuit. At low frequencies the collector current shot noise from Q1 and Q4 dominates. At intermediate frequencies the base current shot noise from Q2 and Q3 dominates.

equation is

$$v_o^2 = 8qI_{icm}\left(1 + \frac{2I_{icm}}{I_a}\right)G_{dc}^2R_L^2. \quad (3.69)$$

At frequencies above the dominant pole from the photodiode capacitance, the shot noise from the input devices is shorted to ground. In its place the shot noise from the base currents of devices Q2 and Q3 becomes dominant. Solving for their contribution to the output yields

$$v_o^2 = 4qI_a \frac{G_a}{G_a + 1} \beta \left(\frac{n}{1+n}\right)^2 (G_b G_p R_L)^2. \quad (3.70)$$

At low signal levels ($G_a \gg 1$) the noise is independent of I_{icm} and the equation predicts $v_o \approx 14 \mu\text{V}/\sqrt{\text{Hz}}$. This agrees well with the values shown in the simulations. Near 20 kHz the noise spectrum begins to roll off as the capacitance at the bases of Q2 and Q3 short the noise current to ground.

When the active cascode circuit is included to improve the bandwidth of the sensor, the noise increases as well. Part of this increase arises from the fact that the shot noise discussed above is collected over a larger bandwidth. Furthermore, the active cascode circuit itself adds a substantial amount of noise.

This can be seen in the simulations shown in Figure 3-17. The noise spectrum is dramatically broader and shows a distinctive hump at mid-frequencies. This can be explained by examining the small-signal model of the active cascode shown in Figure 3-18. The noise current source i_{na} represents flicker and thermal noise currents from the inverter devices M7 and M8. The current i_i is the resulting input to the rest of the current amplifier and is added to the photocurrent signal. The transfer function is

$$\frac{i_i}{i_{na}} = \frac{sG_c\tau_c}{1 + s\tau_c} \quad (3.71)$$

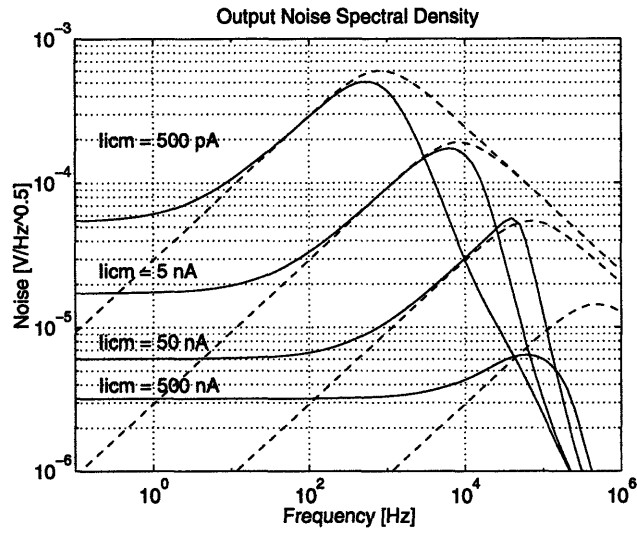


FIGURE 3-17: Simulation results of the output noise spectrum from the amplifier with the active cascode circuit included. Dashed lines are the theoretical result from Equation 3.76.

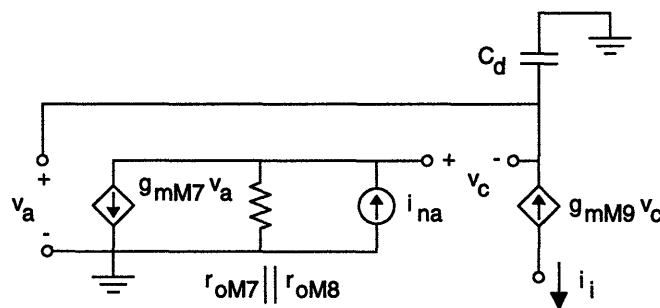


FIGURE 3-18: Small-signal model of the active cascode subcircuit and its noise source.

where $G_c \triangleq g_{mM9}(r_{oM7}||r_{oM8})$ and τ_c is as defined previously. At low frequencies the spectrum is determined by the shot noise sources already discussed. However, as the pole τ_c is approached the noise source i_{na} becomes dominant. The noise source is comprised of thermal and flicker components

$$i_{na}^2 = i_{fl}^2 + i_{th}^2 \quad (3.72)$$

given by

$$i_{th}^2 = \frac{8}{3}kT(g_{mM7} + g_{mM8})\Delta f \quad (3.73)$$

$$i_{fl}^2 = \frac{(K_{FN} + K_{FP})I_{DS}}{C_{ox}L^2f}\Delta f. \quad (3.74)$$

The thermal noise component is a constant which evaluates to about 8.9×10^{-24} A²/Hz. The flicker noise portion depends upon the flicker noise coefficients which are not available for this process. In their absence, estimates of $K_{FN} = 5 \times 10^{-28}$ and $K_{FP} = 1 \times 10^{-29}$ C²/Vs were used. These estimates are based on flicker noise coefficients reported in several references [102, 103, 104, 105, 106]. With $I_{DS} = 30 \mu\text{A}$, $C_{ox} = 0.85 \text{ fF}/\mu\text{m}^2$, and $L = 4.3 \mu\text{m}$ this yields

$$i_{fl}^2 = \frac{9.5 \times 10^{-19}}{f} [\text{A}^2/\text{Hz}]. \quad (3.75)$$

The flicker noise dominates the thermal noise in this circuit. The dashed lines in Figure 3-17 show the flicker noise contribution at the output calculated according to

$$v_o^2 = 2i_{fl}^2 \left(\frac{i_i}{i_{na}} \right)^2 G_{dc}^2 R_L^2. \quad (3.76)$$

The hump shown in the simulations is due to the flicker noise from the active cascode devices which increases until the high frequency poles of the current amplifier cut it off.

3.3.9 Power

The amplifiers dissipate essentially all of the power consumed by the alignment sensor. Most of this power is spent on the four differential output currents which together amount to $400 \mu\text{A}$ and consume 2 mW using a 5V supply. An additional 1.4 mW is consumed by the active cascode inverters. The remaining amplifier stages, bias currents, and photocurrents use roughly 0.6 mW which brings the total power dissipation to 4.0 mW .

3.3.10 Effect of Normalization on Error Signals

The normalizing current amplifiers have a significant effect on the ideal error signals described by Equations 3.3 and 3.18. After amplification the coarse alignment signal becomes

$$I_{qx} = 5I_{aq} \frac{I_3 - I_1 + I_4 - I_2}{I_1 + I_2 + I_3 + I_4}. \quad (3.77)$$

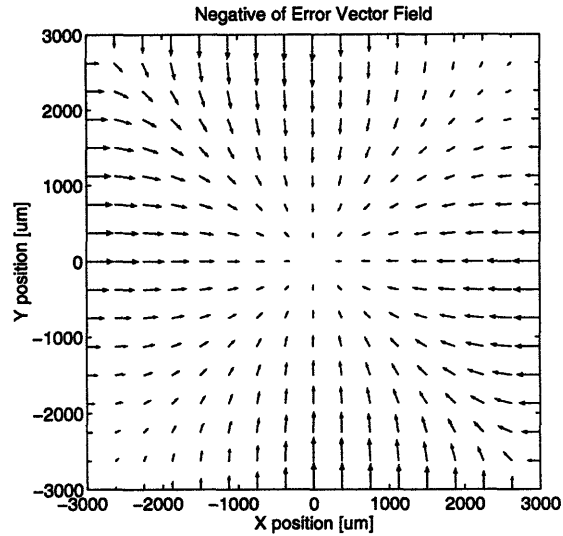


FIGURE 3-19: Coarse alignment negative error vectors after normalization from the amplifiers (with $\rho = 0.9$).

Substituting in Equations 3.4-3.7 for the currents yields

$$I_{qx} = 5 I_{aq} \frac{\rho(L_q - y)x}{L_q^2(2 - \rho) + \rho xy}. \quad (3.78)$$

Note that the irradiance level E_0 drops out leaving an output error signal that is independent of the irradiance. However, the error signal is now inversely proportional to the sum of the quadrant photocurrents. The sum depends on the position of the target and adds additional distortion to the alignment signal as shown in Figure 3-19. Note that the error vectors in quadrants one and three become weaker relative to quadrants two and four. This occurs because displacements of the target towards quadrants one and three increase the total amount of light falling on the sensor, whereas displacements towards quadrants two and four reduce this amount of light. The resulting distortion becomes more apparent the further the displacement is from the origin.

A slightly different situation occurs for the fine alignment error signal. After amplification the edge detector signal is

$$I_{ex} = G_b G_p I_a \frac{(I_A - I_B + I_C) - (I_G - I_H + I_I)}{I_A + I_B + I_C + I_G + I_H + I_I}. \quad (3.79)$$

The denominator of this equation turns out to be a constant equal to $4RE_0L_eW_e(1 + \epsilon)$. Thus

after normalization the error signal becomes

$$I_{ex} = \begin{cases} \frac{1}{2} G_b G_p I_a \frac{1-\epsilon}{1+\epsilon} \left(2 - \frac{\xi}{W_e}\right) & W_e < \xi < 2W_e \\ \frac{1}{2} G_b G_p I_a \frac{1-\epsilon}{1+\epsilon} \frac{\xi}{W_e} & -W_e < \xi < W_e \\ \frac{1}{2} G_b G_p I_a \frac{1-\epsilon}{1+\epsilon} \left(2 + \frac{\xi}{W_e}\right) & -2W_e < \xi < -W_e \\ 0 & \text{otherwise} \end{cases} \quad (3.80)$$

As long as the mark is not rotated and is within the capture region of the edge detectors, the sum of the edge detector photocurrents is a constant independent of the position of the alignment mark. Thus, no additional distortion of the error signal arises due to the normalization.

3.4 Summary

The design for an alignment sensor that detects the position of a surveyor's mark has been presented. The sensor operates in two modes. A coarse alignment signal is generated using quadrant photodiodes to initially position the target mark. Final positioning is achieved using another set of photodiodes that detects the edges of the mark and provide a very sensitive error signal. The photodiode currents are normalized and amplified using BiCMOS circuits to provide an output current that is related to alignment error and independent of illumination level. The bandwidth of this amplifier is improved by using active cascodes to buffer the photodiode depletion capacitance at the expense of added noise.

Test Results: Surveyor's Mark Alignment Sensor

4.1 Introduction

This chapter will present test results from the surveyor's mark sensor. During the course of this work three different test chips were fabricated through MOSIS using the Orbit 2 μm n -well BiCMOS process [107]. The first version was used primarily to test the circuits and photodiodes and only had an imaging area of $1 \times 1 \text{ mm}^2$ in size. A second sensor was then fabricated containing an enlarged imaging area of $6 \times 6 \text{ mm}^2$ and added the active cascode circuit to improve the sensor bandwidth. The final version of the chip made some improvements in the active cascode as well as in the coarse alignment amplifier. The data presented here is from the final design unless otherwise mentioned.

4.2 Photodiodes

The sensor uses PMOS transistor p -type source/drain implants in an n -well for the photodiodes. A cross section of the structure is shown in Figure 4-1. Because these are essentially parasitic photodiodes they are not normally characterized in much detail, if at all. Their responsivity, leakage (or dark) current, and matching characteristics were measured to assure their suitability for use as photosensitive elements.

The responsivity of the photodiode is a measure of how much current is generated per unit of light power at a particular wavelength. Narrow band illumination was provided by a bank of red AlGaAs light emitting diodes¹ (LED's). These diodes have a narrow emission peak at 637 nm with a half width of 20 nm. Uniform illumination was obtained by using LED's with light diffusers and placing them 15 cm away from the chip. The irradiance from this source at

¹HP HLMP-D101.

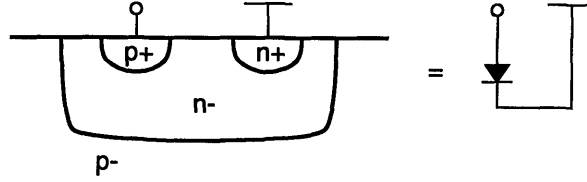


FIGURE 4-1: Cross section of the photodiode structure. It is formed by a PMOS p -type implant in an n -well connected to $V_{dd} = 5$ V.

the sensor surface was about $5.4 \mu\text{W}/\text{cm}^2$. The resulting photocurrents were measured from all of the photodiodes in the sensor with 5 V reverse bias. Data was collected from several different chips. Each chip was swapped with a calibrated photodiode (Newport 1815 power meter with 818-SL detector) to measure the incident light power. The calculated responsivity was 0.140 A/W under this illumination which corresponds to a quantum efficiency of 27%. The responsivity dropped to 0.130 A/W (QE = 25%) at 2.7 V reverse bias (normal operating conditions) and to 0.120 A/W (QE = 23%) with 1.1 V reverse bias (when the active cascode is engaged). Commercial photodiode sensors typically have quantum efficiencies in the range of 70% to 90%.

In order to determine the dark current of the photodiodes, the reverse leakage current was measured under 5 V of reverse bias and without illumination.² These currents were measured for photodiodes of several different sizes in order to find the leakage as a function of the area and perimeter of the diode. The leakage current is then expressed as

$$I_l = J_{RA}A + J_{RP}P \quad (4.1)$$

where A is the device area, P is the perimeter, and J_{RA} and J_{RP} are the corresponding leakage coefficients. A picoammeter with a 1 fA resolution was only able to measure leakage currents on the order of tens of femtoamperes. This leakage appeared to occur almost entirely along the device perimeter. The coefficient J_{RP} was always below 2 pA/cm. The area dependent leakage was less than $100 \text{ pA}/\text{cm}^2$, below the resolution of the measurements. The largest structures on the test die are the quadrant photodiodes which had leakages well under 1 pA.

The leakage currents for other structures on the sensor were examined to determine if they might affect the integrity of the photocurrent signal. Measurements were made for the n -diffusion and the collector diffusion (n -well) to substrate junctions. These junctions also had very small leakage currents with $J_{RP} < 7 \text{ pA}/\text{cm}$ and $J_{RA} < 300 \text{ pA}/\text{cm}^2$. The leakage currents for the photodiodes and these other structures are all far below the levels of the photocurrents obtained under normal illumination.

The matching properties of the photodiodes were investigated using the red LED light source. This illumination proved to be uniform to within 0.2% across the surface of the sensor.

²These measurements were done only on the first test chip.

The photocurrents from each of the photodiodes were measured and their calculated responsivities were compared to determine their mismatch parameters. The effect of photodiode mismatch on the sensor offset can be viewed from two perspectives. The first is a statistical approach where the mismatch parameters are treated as independent random variables. From the measured responsivities one can calculate the sample statistics for the photodiodes. Because the sample sizes are small, there are only four quadrant photodiodes for example, the resulting matching parameters were averaged with the measurements from several chips.

For the quadrant photodiodes the mismatch parameter δ had a sample standard deviation of 0.60%. Using the statistical offset model from Equation 3.11 results in a normalized offset with a standard deviation of only 0.22% of the sensor size for the coarse alignment. This is much smaller than expected. The capture region of the edge detectors was sized to tolerate offsets in the coarse alignment of up to 5% to accommodate both photodiode mismatch and nonuniform illumination.

For the edge detecting photodiodes, the statistics were computed for both the small photodiode mismatch ζ and the large photodiode mismatch γ . Both parameters had comparable standard deviations with $\sigma_\zeta = 0.43\%$ and $\sigma_\gamma = 0.43\%$. It is interesting to note that the matching variances of all the photodiodes are comparable even though they differ greatly in shape and size. The mean of ζ is chosen to be zero by definition so that the mean of γ then expresses the systematic mismatch between these two different sized photodiodes. The resulting mean $m_\gamma = -0.95\%$ reveals that this middle photodiode, although drawn to be twice the area, yields slightly less than twice as much photocurrent. This is perhaps due to photocurrent generated along the diode perimeters. The statistical model of Equation 3.23 predicts that the normalized fine alignment offset will have a standard deviation of 110 ppm due to this mismatch.

Another way to examine the expected offsets caused by photodiode mismatch is to explicitly compute them from Equations 3.10 and 3.21. The standard deviation of the resulting offsets from several chips were calculated using these equations. The resulting coarse alignment offset was 0.27% and the fine alignment offset was only 20 ppm. Compared to the statistical method the coarse alignment value is larger whereas the fine alignment value is substantially smaller. This difference arises because there is significant correlation in the mismatch parameters that is antisymmetric about the center of the sensor. Since the coarse alignment signal is formed by subtracting photocurrents from opposing photodiodes, the effect of this correlation is exacerbated. On the other hand, the fine alignment signal is obtained by summing the edge detector signals from opposing edges which cancels the antisymmetric component of the mismatch. The fine alignment offset of 20 ppm is only the error resulting from photodiode mismatch. More significant errors will be contributed from other sources such as nonuniformity of illumination, amplifier offsets, and the resolving power of the optics.

Table 4-1: Typical DC parameters measured for the current amplifier.

Amplifier	Full Scale Output [μA]	Gain	Output Offset [%]	Nonlinearity [%]
Normalizing	1	0.1–10,000	0.3	0.2
Inverting	8.2	8.2	0.2	0.2
Mirror	100	10	< 0.1	< 0.05
Composite	82	8.2–820,000	0.3	0.4

4.3 Current Amplifier

The current amplifier was designed to normalize the input photocurrents and scale them up to a level suitable for output from the sensor. Several DC measurements were made of these amplifiers to determine their gain, offset, and linearity. The amplifier was tested with the source/measurement units (SMU's) from two HP 4145's. They varied the input currents differentially about a fixed common-mode level I_{cm} by an amount $\pm I_{cm}$ while the output currents were measured. The measurement was repeated over a range of common-mode input levels. Because the amplifier is designed to operate with very small input currents it was not practical to measure the amplifier bandwidth directly. Bandwidth measurements were made indirectly using optical input and are discussed later.

4.3.1 Fine Alignment Amplifier

The fine alignment amplifier was discussed in Section 3.3 and shown in Figure 3-8. A summary of typical DC parameters for the complete amplifier and its components is shown in Table 4-1. The full scale output is the maximum differential current signal that appears at the output of each stage. For the normalizing multiplier and the current mirror this is simply twice the output common-mode current. However, the inverting multiplier has a slightly reduced full scale differential output due to the attenuation factor K in the differential gain. This also reduces the full scale output of the composite amplifier. The gain is the differential current gain from each stage. For the normalizing multiplier this gain is dependent on the common-mode level of the input. Measurements were made at common-mode input levels of 5 μA , 500 nA, 50 nA, 5 nA, 500 pA, and 50 pA to test the amplifier over several decades of input range.

Nonlinearity in the amplifier was measured, but it is not of much importance for the alignment application. The error signals from this sensor are not linear functions of the image position except when the alignment error is close to zero. The small additional nonlinearity from the amplifiers is of little significance. The importance of the amplifier offset depends on the alignment application. This offset will cause the error signal to be zero when the image of the mark is shifted slightly relative to the sensor. If one truly desires the image to be aligned

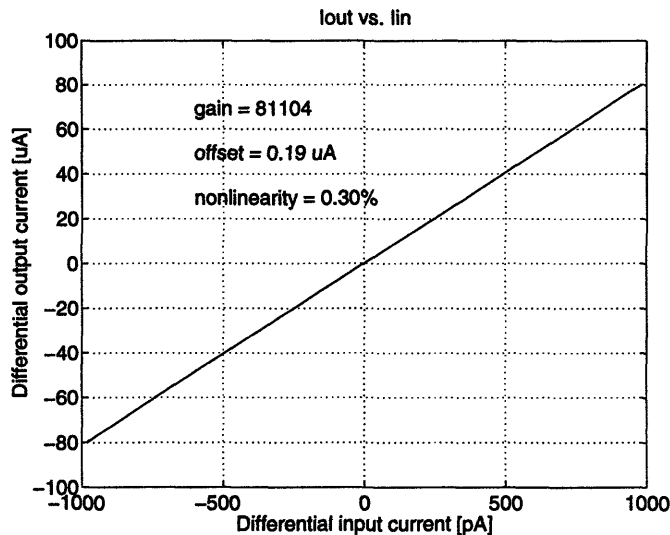


FIGURE 4-2: Typical input/output characteristic of the differential current amplifier. Input common-mode current was 500 pA and output common-mode current was 50 μ A.

exactly to the sensor, then this amplifier offset (and also the photodiode mismatch) will lead to some error. If, however, it is only desired that the alignment be repeatable, then a fixed amplifier offset is not going to be important as long as it does not exceed the capture region for the transition between coarse and fine alignment.

The values for offset and nonlinearity are expressed as a percentage of the full scale output because this is constant as the gain is varied. They are typical worst case values for common-mode inputs down to 500 pA. Below this level the amplifiers continue to operate properly but accurate measurements could no longer be made with our instrumentation. The nonlinearity was computed as the maximum deviation of the data from a best fit line over the middle 80% of the full scale range. The nonlinearity of the composite amplifier is larger than all of the components because of loading effects between the stages. The output impedance of the amplifier is the output impedance of the PFET device in the current mirror which was found to be greater than 7.0 M Ω .

Another measurement was made to determine how well the normalization works as the common-mode input is varied. Ideally, the output should remain constant for a fixed ratio of the input differential signal relative to its common-mode level. This was checked by measuring the variation in the output for an input ratio $I_{idm}/I_{icm} = 0.8$ at the different common-mode levels. On average the output varied by only 60 nA (0.1% of the 60 μ A signal) which is near the measurement resolution of the HP 4145. This shows that the amplifier effectively normalizes the input differential signal while rejecting variations in the common-mode signal.

A typical input/output characteristic is shown in Figure 4-2 for the complete amplifier. For this measurement the common-mode input level was 500 pA corresponding to an amplifier gain of about 80,000. The linearity is excellent with only 0.3% of nonlinearity over 80% of

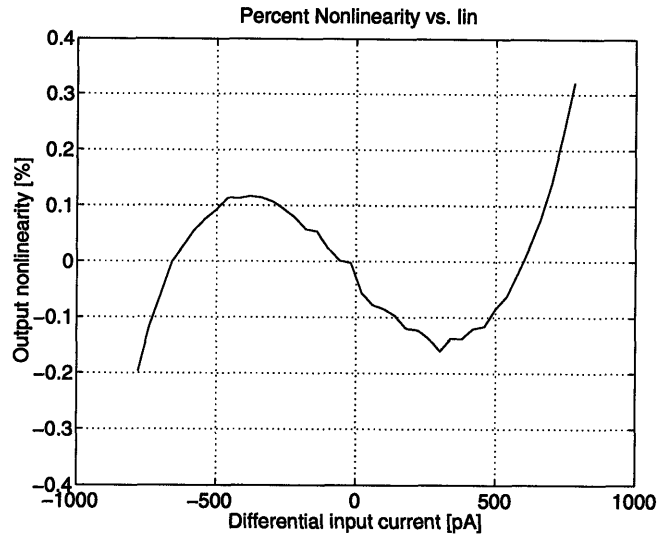


FIGURE 4-3: Nonlinearity of the differential current amplifier. This is the difference between the data and a least squares fit line expressed as a percentage of the full scale difference. Nonlinearity is chiefly due to Early effect in the bipolar devices.

the full scale range. The offset referred to the output is $0.19 \mu\text{A}$ which is about 0.2% of the full scale output. A plot of the amplifier nonlinearity is shown in Figure 4-3. HSPICE simulations indicate that the nonlinearity is chiefly due to Early effect in both the bipolar transistors and, to a lesser extent, the MOSFETs in the current mirror. Gilbert [100] claims that β mismatch between bipolar devices in these amplifiers is also a significant source of nonlinearity. However, in this implementation the Early effect adequately accounts for the observed nonlinearity. Essentially identical results were seen for the other common-mode input levels that were tested. These measurements demonstrate that the current amplifier is capable of accurately scaling differential currents over a wide range of input levels.

4.3.2 Coarse Alignment Amplifier

The coarse alignment amplifier was discussed in Section 3.3.6 and illustrated in Figure 3-15. The input normalizing amplifier is operated at a $5 \mu\text{A}$ common-mode level and the final output current has a nominal common-mode output of $50 \mu\text{A}$. Note that these bias conditions are somewhat different than is the case for the fine alignment amplifier. DC measurements were made by setting two of the four input currents to zero and differentially varying the remaining two currents about a fixed common-mode level. The measurements were repeated for various input combinations and input common-mode current levels of 500 nA , 50 nA , 5 nA , and 500 pA .

The aspects of DC performance, such as gain, gain variation, offset, and nonlinearity were measured as was described above for the fine alignment amplifier. The differential current gain ranged from 95 to 95,000 depending upon the input current level. This gain was slightly less than the ideal value of 100 to 100,000 because the output common-mode current was actually

closer to $48 \mu\text{A}$ in the experiments than the $50 \mu\text{A}$ nominal value. Gain variation for a fixed differential signal was about 0.4% of the full scale signal and the nonlinearity was 0.6%. Offset referred to the output was only 0.1% of full scale, which was at the measurement resolution. Both gain variation and nonlinearity were slightly worse for the four input amplifier than was the case for the two input version. This may be due to the fact that, although a symmetrical layout was used, it was not strictly common centroid. Surprisingly, the offset values were superior. This may be a coincidence due to the small number (four) of chips tested. Typical input/output characteristics of this circuit appeared much the same as in Figures 4-2 and 4-3. The only noticeable difference is the slightly higher differential gain for the same input current level.

4.4 Alignment Error Signals

4.4.1 Mock Alignment System

In order to characterize the alignment sensor it was necessary to image an alignment marking upon it and measure the resulting error signals. These measurements were made by constructing a mock alignment system. The alignment sensor was mounted inside an empty video camera body allowing the use of inexpensive standard C-mount video camera lenses. Most of the experiments used a Cosmicar C31204 2/3 inch format zoom lens with an adjustable focal length from 12.5–75 mm. The zoom lens allowed easy adjustment of the optical magnification. In front of the zoom lens was attached a #3 close-up lens for 33 cm working distances. The camera body was attached to a Kaiser RS1 copy stand to allow adjustment of the camera height.

On the copy stand platform was placed a computer controlled x - y table. The table was constructed out of two Aerotech ATS15010 linear positioning stages. These stages have a total travel of four inches and are driven by stepper motors with a 10,000 step/inch resolution. The absolute accuracy of the system was specified to be $12.2 \mu\text{m}$ with a repeatability of $3.1 \mu\text{m}$. In practice the bidirectional repeatability appeared to be only $12.5 \mu\text{m}$ (5 steps) due to hysteresis in the positioning system. The magnification of the optics was chosen such that the total four inch travel mapped onto the six millimeter imaging area of the sensor. Normalized to the sensor field of view, the specified positioning accuracy is 120 ppm with a repeatability of 30 ppm. The table stepper motors were controlled by two Aerotech Unidex U100M motor controllers.

The camera and alignment sensor looked down upon an alignment marking mounted on the positioning table. Since the primary objective of this system was to characterize the sensor, a large alignment mark measuring eight inches across was used. This has several benefits. First, the optical system does not require magnification which permits the use of standard video camera optics. Second, the resolution, repeatability, and accuracy required of the positioning system are correspondingly relaxed. Third, vibrations in the mechanical system are less of a constraint. Fourth, the precision with which the target itself is made can be reduced. The

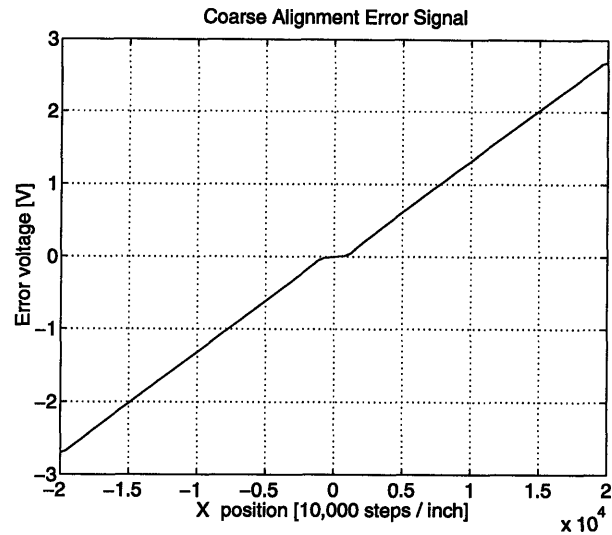


FIGURE 4-4: Coarse alignment x -direction error signal as the target is moved along the x -axis over the full field of view. The slope is about $140 \mu\text{V}/\text{step}$ except in the zone near the origin.

alignment target was laser-printed on eight and one-half by eleven inch white bond paper. Finally, for all of these reasons the mock alignment system could be built at moderate expense.

The sensor was connected by a ribbon cable to a test board which supplied power as well as bias currents and voltages. The error signals, output as differential currents, were converted to single-ended voltages on the test board using 741 operational amplifiers and resistors. The current to voltage conversion was done using precision $39.1 \text{ k}\Omega$ resistors matched to 0.1%. The resulting error voltages were measured using the source/measurement units of an HP 4145 controlled by an HP 715/33 Apollo workstation. The workstation could communicate with the motor controllers and the HP 4145 over a GPIB interface. Thus it was possible to issue motion commands and measure the resulting position error signals to characterize the sensor entirely under computer control.

4.4.2 Coarse Alignment Signal

A one-dimensional slice of the x -direction error signal is shown in Figure 4-4. For this measurement the alignment target was translated along the x -axis while the coarse alignment x -direction error voltage was measured. The signal is roughly proportional to displacement from the origin with a swing of $\pm 2.69 \text{ V}$. Because the error currents are dropped across $39.1 \text{ k}\Omega$ resistors, a full scale differential signal would be 3.91 V . This is not achieved because the mark does not have perfect contrast due to light reflecting from the darkened regions. From Equations 3.78 the resulting attenuation is given by $(1 - \epsilon)/(1 + \epsilon)$. Using the ratio of the maximum achieved signal to the maximum possible signal one can solve for the value of ϵ which turns out to be 0.185. Apparently a significant amount of light is reflected from the darkened regions of the target. A confirmation of this large value of ϵ was made by taking the ratio

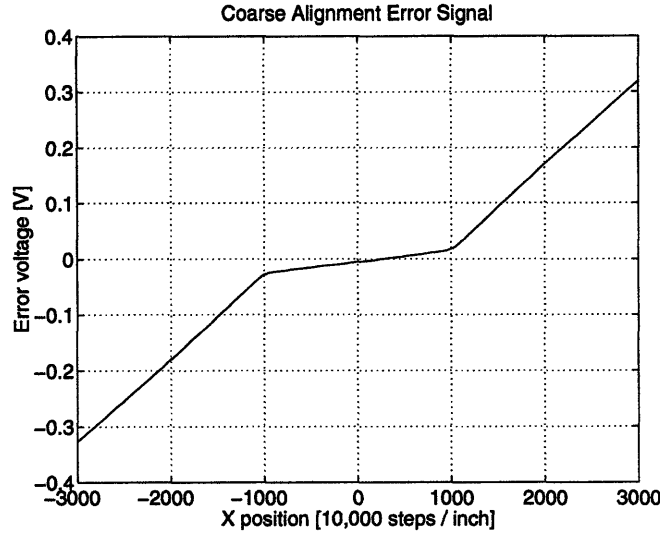


FIGURE 4-5: Coarse alignment error signal near the origin showing the reduced slope of $20 \mu\text{V}/\text{step}$ due to the presence of the edge detecting photodiodes.

of photocurrents from darkened and fully illuminated photodiodes. This was done using the square grating sensor with the surveyor's mark target and yielded a contrast of 0.173, fairly close to the estimate obtained above. The square grating sensor was used because its test board in the video camera provided easy access to these currents. The contrast varied by about ± 0.01 depending upon the position of the light source.

The slope of the error signal is about $140 \mu\text{V}/\text{step}$ over most of the range of motion. Near the origin the signal flattens due to the presence of the edge detecting photodiodes. Figure 4-5 shows a close-up of this behavior. The slope increases slightly to $156 \mu\text{V}/\text{step}$ in the range of 1000 to 2000 steps because of the additional signal from the outer edge detecting photodiodes. The slope drops dramatically to only $20 \mu\text{V}/\text{step}$ between ± 1000 steps of the origin. This occurs when the edges of the mark fall on the center edge detecting photodiodes, leaving only the small region at the center of the sensor contributing to the coarse alignment signal.

Due to the small slope near the origin a substantial offset is possible. The null point can easily occur anywhere in the range of ± 1000 steps of the origin. The offset seen in the figure is chiefly due to nonuniformity in the irradiance of the sensor. Measurements of the irradiance at the target typically revealed a variation of five to ten percent across the field of view. Furthermore, inaccurate centering of the sensor behind the lens of the camera (accurate to maybe 2 mm) adds additional nonuniformity in the sensor irradiance that is difficult to quantify. A white sheet was placed as the target to get a crude idea of the illumination effect on the offset. Ideally a white sheet would yield uniform irradiance of the sensor and thus zero error signal. The observed error signal is due to a combination of offsets from nonuniform illumination, photodiode mismatch, the current amplifiers, and the current-to-voltage converters on the test board. All of these should be quite small, on the order of 0.3% of the full scale signal, except the

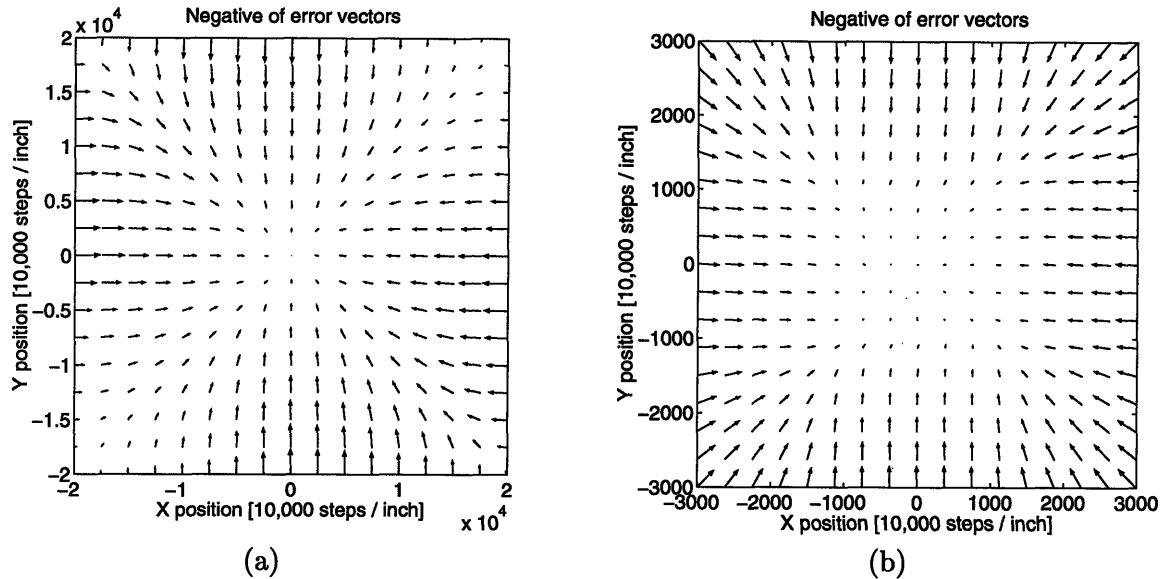


FIGURE 4-6: Negative error vector field from the coarse alignment signal (a) over the full field of view and (b) near the origin showing the low gain region inside ± 1000 steps.

illumination offset. Typically the white sheet offset was about 50 mV, or about 1.3% of the full scale signal, and was very sensitive to the position of the illuminating source. A better lighting arrangement and more precise control of the sensor position behind the lens should reduce this offset considerably. Nevertheless, this problem was anticipated by designing the size of the fine alignment capture region to tolerate a coarse alignment offset of up to five percent, easily accommodating the observed values. If a smaller capture region were used with better lighting uniformity, then a fine alignment error signal with larger slope could be obtained.

The coarse alignment error signals were also measured at a grid of displacements over the sensor's field of view. These measurements were then used to construct a map of the negative error vectors shown in Figure 4-6 (a). The resulting vector field compares well with the expected result seen in Figure 3-19. The slight differences between these two figures are chiefly due to the presence of the edge detecting photodiodes (not modeled in the coarse alignment analysis) and the effect of the previously mentioned image contrast. A closer view of the origin is shown in Figure 4-6 (b). Here the low gain in the region ± 1000 steps is quite evident. The effect of a ten degree rotation of the mark is shown in Figure 4-7. As the simulations predicted (see Figure 3-4), the error vectors are all rotated slightly. The resulting curl in the vector field will cause trajectories to spiral into the origin. One such trajectory is also shown in the figure.

4.4.3 Fine Alignment Signal

Measurements of the fine alignment error signal also agreed well with the expected results. The x -direction error signal along the x -axis is shown in Figure 4-8. The signal clearly shows the expected sawtooth shape depicted in Figure 3-5. The signal strength peaks at about ± 1.05 V

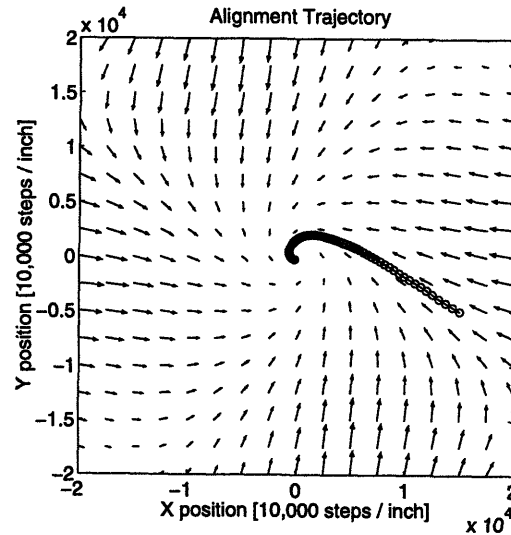


FIGURE 4-7: Negative error vector field from the coarse alignment signal with the mark rotated by ten degrees relative to the sensor. Note the resulting spiral in a measured alignment trajectory.

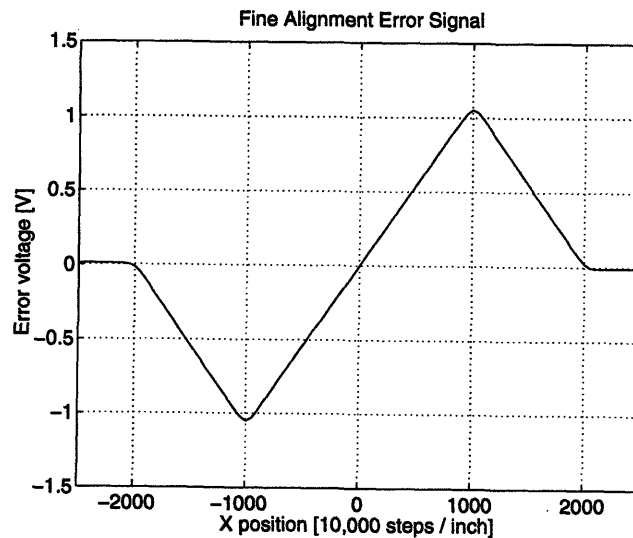


FIGURE 4-8: Fine alignment x -direction error signal as the target is moved along the x -axis. The capture region extends to ± 2000 steps and the signal has a slope of 1.1 mV/step through the origin.

at positions ± 1000 steps. These peaks occur when the edges of the mark lie at $\pm W_e$ from the center of the sensor. Equation 3.80 predicts a signal of

$$V_{ex} = (39.1 \text{ k}\Omega)(50)(0.82)(1 \text{ }\mu\text{A}) \frac{1 - 0.173}{1 + 0.173} = 1.13 \text{ V} \quad (4.2)$$

assuming the same value for ϵ from the coarse alignment data. The slope of the signal through the origin is 1.1 mV/step corresponding to a peak value of 1.1 V. This is in good agreement with the value predicted above. The actual peak value of 1.05 V is slightly less due to rounding of the peak discussed below. With a slope of 1.1 mV per step a voltmeter with 1 mV resolution can detect displacements of a single step of the table. This is an absolute motion of 2.5 μm (25 ppm) of the target. The corresponding motion of the image is 0.15 μm .

Ideally the peaks at ± 1000 steps would be sharp and have discontinuous slope. In actuality they are rounded over a distance of about 70 steps which corresponds to 10.5 μm on the sensor. This rounding could arise from several sources. First, the 3 μm gaps between the photodiodes would ideally cause a flat region 3 μm wide. Second, blurring of the image will round the peak. The sharpness of the image is limited to the minimum resolvable distance set by Rayleigh's criterion [96]

$$\Delta l = 1.22 \frac{f\lambda}{D} \quad (4.3)$$

where f is the focal length of the lens, D is the aperture diameter, and λ is the wavelength of illumination. The ratio f/D is the $f\#$ of the lens. For $f\# = 2.5$ and $\lambda = 637 \text{ nm}$ the minimum resolvable distance is about 2 μm . The effects of blurring and the photodiode gap are both substantially smaller than the observed rounding. Rounding of the peak will also occur if the image edge and the edge of the photodiodes are not exactly parallel. A small rotation of the mark could cause this as was seen in the previous chapter. The precision to which the target rotation could be adjusted was about a tenth of a degree. A rotation error of this magnitude would cause a 10 μm displacement across the 6 mm width of the sensor and a corresponding broadening of the peak. Finally, aberration in the lens causing pin cushion distortion will bend the image of a line. This could also contribute to the rounding of the peak, although no investigation was done of the lens distortion. The most likely cause of this broadening is probably a slight rotation of the mark relative to the sensor.

Figure 4-9 shows the negative error vectors in the fine alignment capture area. The vector field agrees well with the expected field shown in Figure 3-6. Because the error signals are linearly proportional to displacement inside ± 1000 steps, the vector field in this region points radially inward. The entire capture region extends out to ± 2000 steps.

The effect of rotation of the mark is shown in Figure 4-10. The fine alignment error signal is much less tolerant of rotations. It was shown earlier that a rotation in excess of 5.2 degrees causes the null point of the error signal to split into multiple equilibrium points. Part (a) of the figure shows the x -direction error signal along the x -axis when the mark is rotated by 5.8 degrees. Note the slightly negative slope at the origin. The actual critical angle in this case

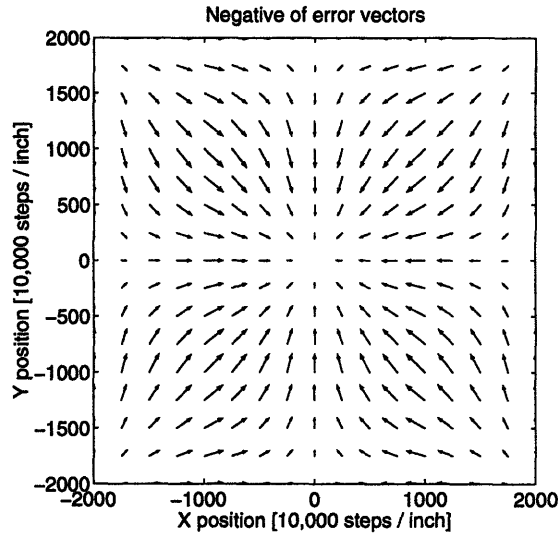


FIGURE 4-9: Measured negative error vector field for the fine alignment signal.

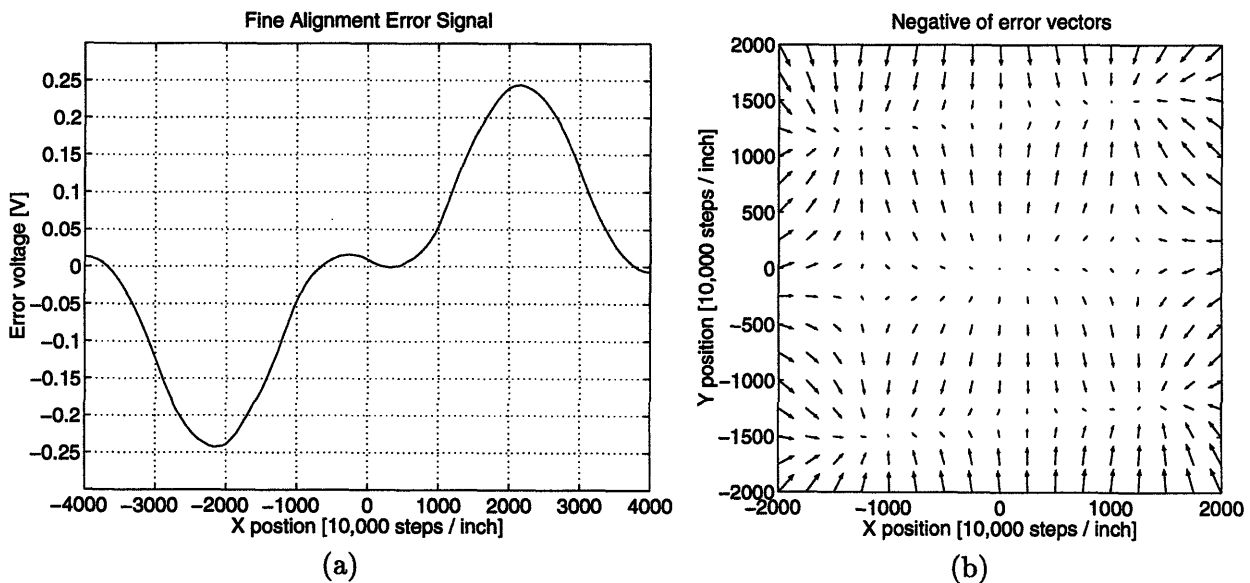


FIGURE 4-10: Fine alignment error signals after the mark is rotated. (a) The x -direction error signal for a 5.6 degree rotation just past the critical angle. Note the reduced signal amplitude, broadened capture region, and increased offset. (b) The error vector field for a seven degree rotation showing the multiple equilibrium points.

occurred at 5.5 degrees, slightly larger than the approximate analysis predicted. Generally, the critical angle was observed at 5.4 ± 0.1 degrees for rotations that bring the white quadrants over the edge detector and at 5.2 ± 0.1 degrees for rotations in the opposite direction. The difference between the experiment and predicted values could be due to the simple theoretical model which omitted the effects of imperfect contrast, the systematic mismatch γ , and the blurring of the edge. Also note the reduced amplitude of the error signal, the broadening of the capture region, and the offset at the origin. The offset arises from the illumination mismatch of the two edge detectors which normally cancels for no rotation but adds when the mark is rotated to the critical angle. Part (b) of the figure shows the vector field for a seven degree rotation where the four equilibrium points are evident. The agreement with the calculated field in Figure 3-7 is very good.

4.4.4 Alignment Trajectories

The mock alignment system was used to perform complete closed loop alignments in order to demonstrate the capabilities of the sensor. A simple discrete proportional control loop was implemented as follows. A program running on the computer measured the error signal voltages via the HP 4145. The resulting values were multiplied by a fixed gain to convert the voltage into distance. The motor controllers were then instructed to move the table by that amount and the resulting position was recorded. This process continued until the error signal voltages dropped below a predetermined threshold. Initial alignment used the coarse alignment error signal. When that was minimized, the fine alignment error signal was used to complete the alignment. Because of the large amount of time spent communicating over the interfaces, each position update took about 1.3 seconds to complete and the entire alignment process took a few tens of seconds. Because of the communication and measurement delays this system could not demonstrate the potential speed of the alignment but it did show the trajectories taken by a simple proportional control system using the sensor and the potential repeatability of the alignment.

Coarse alignment trajectories for several initial positions are shown in Figure 4-11. The measured error vector field is also plotted to show how the trajectories follow the error vectors. The coarse alignment gain was set to 1000 steps/V for this experiment. It took between 25 and 38 iterations to complete the alignment depending upon the initial position. A low gain resulting in many intermediate positions was used so that the trajectories could be seen clearly. Gains up to 6500 steps/V can be used before the trajectories begin to overshoot. At high gain the alignments were completed in 2 to 5 iterations. The trajectories taking the greatest number of iterations start in the upper right and lower left corners where the error signal is the weakest. The coarse alignment was terminated when the error signal for both axes dropped below 50 mV. Figure 4-5 shows that this threshold guarantees a position well within the ± 2000 step capture region of the fine alignment.

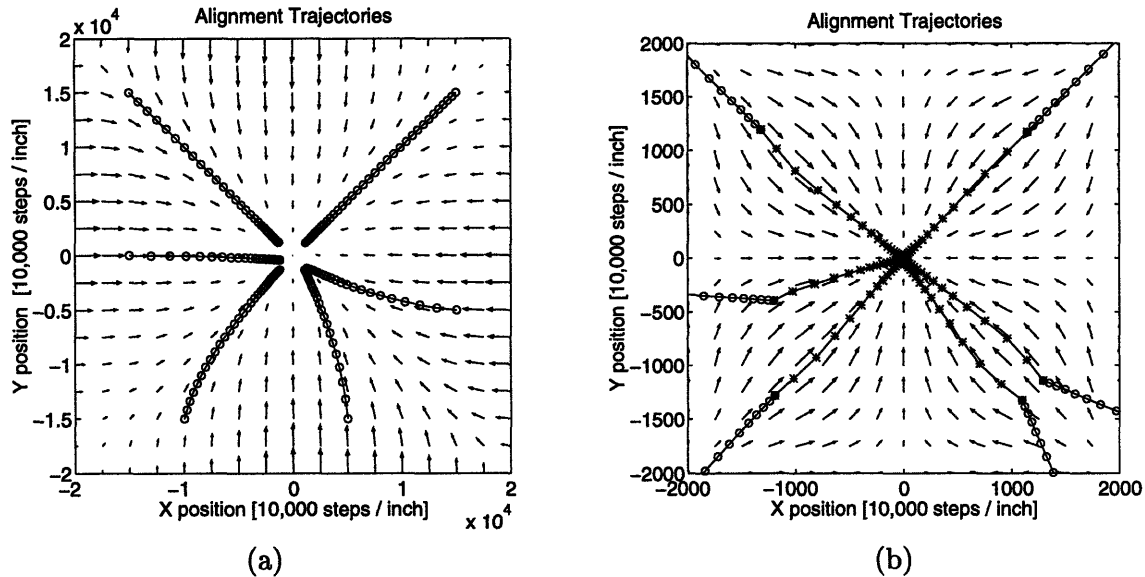


FIGURE 4-11: Trajectories using (a) the coarse alignment error signal (o's) and (b) the fine alignment error signal (*'s) with the corresponding error vector fields superposed.

After completion of the coarse alignment the system was switched over to the fine alignment signal. The fine alignment gain was also set to a low value of 200 steps/V in order to show many points along the trajectory. These trajectories are plotted in Figure 4-11(b) with asterisks and the previous coarse alignment trajectories are shown with circles. The fine alignment error vector field is also plotted. Slope discontinuities in the trajectories occur at the transition point between the coarse and fine alignment. The fine alignment trajectories took between 23 and 26 iterations to complete. The gain can be set as high as 900 steps/V before overshoot occurs, in which case the alignments were completed in 5 to 8 iterations.

4.4.5 Alignment Repeatability

The repeatability of the mock alignment system was also investigated. The null point position measured in motor steps was found to vary by a small amount depending from which side it was approached, indicating some hysteresis in the positioning system. The y -axis showed only one or two steps of hysteresis but the x -axis had almost five steps. This appears to be due to the mechanical drive of the table. It was minimized by overshooting and backtracking so that each new position was approached from the same direction.

Several sequences of 100 fine alignments were done. The algorithm updated the target position until the error voltages were less than 1 mV as measured by an HP4145. Note that these measurements take a few milliseconds and therefore substantially reduce the sensor noise. The final x and y positions in motor steps were recorded and compiled into histograms such as that shown in Figure 4-12. The repeatability of the alignment for both axes had 3σ values of 2.1 steps or 53 ppm of the field of view. The positioning table was specified to have about

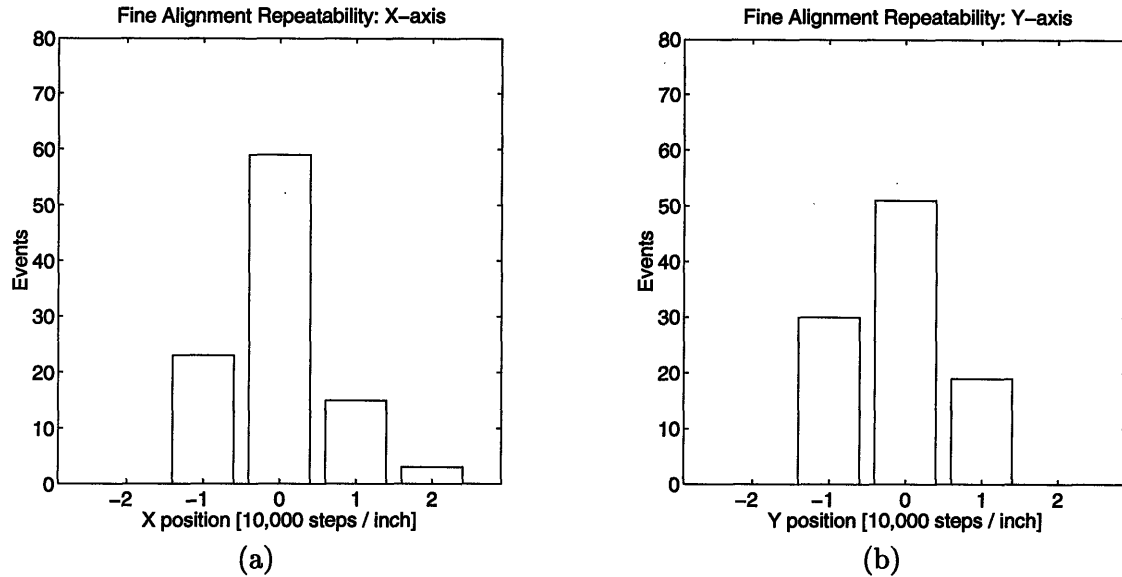


FIGURE 4-12: Histograms of the final alignment position along the (a) x -axis ($3\sigma = 2.13$) and (b) y -axis ($3\sigma = 2.07$) for a sequence of 100 alignments.

30 ppm of repeatability. An additional 30 ppm of drift can be caused by only a one degree Celsius rise in the leadscrew temperature during the alignment sequence. Furthermore, the HP4145 had only a 1 mV resolution corresponding to an additional 25 ppm of uncertainty in the measurement. Thus it seems likely that the sensor itself was not the limiting factor in these repeatability measurements.

A repeatability of 2.1 steps corresponds to a repeatability of $5.3 \mu\text{m}$ in the target plane across the four inch (10 cm) range of motion. The image of the target on the sensor is related to this by the system magnification

$$\delta x_i = M \Delta x_o. \quad (4.4)$$

The mock alignment system used a magnification of 5.9×10^{-2} yielding a repeatability of $0.31 \mu\text{m}$ at the image plane. If one were to use an alignment target of $200 \mu\text{m}$ on a side with $M=60$, the target repeatability would be 5.3 nm assuming the system repeatability scaled with the magnification. At such small dimensions the alignment repeatability would likely be limited by other factors such as the mechanical system, lighting, and quality of the alignment mark rather than by the sensor itself.

4.5 Bandwidth

Measuring the sensor bandwidth presents some difficulties. In order to test the current amplifier by itself requires generating rapidly time varying input currents in the range of 500 pA to 500 nA at frequencies up to 100 kHz. This did not seem to be practical considering the large resistances required to generate such small currents combined with the parasitic capacitances

from the package pins and bonding pads. In order to measure the bandwidth of the sensor with an imaged mark would require moving the target rapidly while monitoring the output error signal. However, controlled motions fast enough to test the limits of the sensor were not possible with the equipment available. Thus, an indirect approach was taken to measure the sensor bandwidth. Rather than rapidly moving the target, flashing LED's were used in place of the target. The resulting time varying photocurrents generated on the sensor formed the amplifier input used to measure the bandwidth.

Bandwidths were measured for both the coarse and fine alignment amplifiers. For the fine alignment measurements two LED's were placed on the table in place of the target such that the image of one appeared on photodiode *B* and the other on photodiode *H*. A DC bias current was applied to each photodiode and adjusted so that the resulting differential error signal from the sensor was zero. A small time varying current was added to one of the LED's to create a differential output signal with a peak-to-peak amplitude of 100 mV. The sensor bandwidth was then measured using an oscilloscope and an HP 3585A spectrum analyzer. After a bandwidth measurement was made, the sensor was removed from the camera and the total irradiance at the focal plane was measured with the Newport³ power meter. This power was divided by the area of the illuminated photodiodes to yield the equivalent irradiance of the sensor per unit area. The technique was the same for the measurements of the coarse alignment amplifier with the modification that four LED's were used in order to provide illumination to each of the quadrant photodiodes.

A series of these measurements were made at different illumination levels by placing neutral density filters in front of the lens. The active cascode circuit could be bypassed and thus measurements were made both with and without it. Figure 4-13 shows the resulting bandwidths plotted versus the sensor irradiance. Measurements made of the fine alignment signal are shown as *'s and the o's are from the coarse alignment signal. The lower data set was measured without the active cascode circuit and the data points lie close to the solid line given by Equation 3.61. The bandwidth is limited by the pole formed from the photodiode capacitance coupled with the amplifier input resistance. Because the input resistance is inversely proportional to the photocurrent, the bandwidth is proportional to sensor irradiance.

The upper set of data was taken with the active cascode circuit operating and the bandwidth is improved substantially. The solid line is given by the equation

$$f_h = \frac{(A + 1)RE_0}{2\pi n \frac{kT}{q} C_{da}} \quad (4.5)$$

which models the dominant pole as determined by the photodiode capacitance coupled with the resistance looking into the active cascode. For this calculation C_{da} increases to 21 nF/cm² because of the reduced bias across the photodiode. The data points deviate from this result for

³See Section 4.2 for model number.

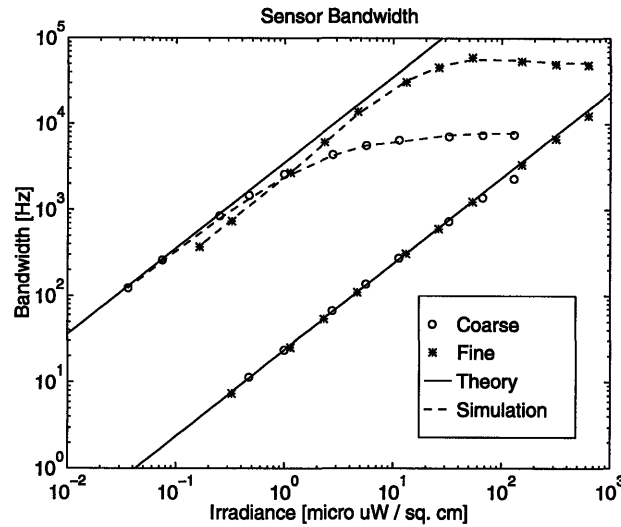


FIGURE 4-13: Sensor bandwidth plotted versus irradiance. The upper data set used the active cascode, the lower set did not. Coarse alignment measurements are shown with o's and fine alignment measurements are shown with *'s. The solid lines show the bandwidth limit from the photodiode capacitance and the dashed lines show values obtained using HSPICE simulation.

a variety of reasons. At high levels of irradiance, the coarse alignment measurements flatten out around 7.5 kHz. Pass transistors in series with the photodiodes are the cause of this limiting pole. These transistors were included for testing purposes and due to their small size present a resistance of about 5.7 k Ω apiece. They form a pole with the photodiode capacitance which places an upper bound on the achievable bandwidth. This behavior is also reflected in the simulations of the circuit shown with the dashed line.

Because the edge detecting photodiodes have substantially less capacitance, the pole due to the pass transistors does not appear until about 55 kHz. Furthermore, the effect of the amplifier poles on the bandwidth is stronger. At lower levels of irradiance the measurements fall along, but slightly below, the solid line modeling the dominant input pole. This occurs because the input pole has been pushed so far out by the active cascode that it is no longer completely dominant. The frequency limitations of the amplifier itself now affect the bandwidth of the sensor.

These measurements demonstrate that the sensor is capable of operation substantially faster than video rate. The actual speed achieved is directly proportional to the amount of light energy collected by the sensor. Using the active cascode circuit dramatically improves the bandwidth by reducing the voltage fluctuation across the photodiode. The poorly sized pass transistors limited the bandwidth to about 7.5 kHz, but simulations using the active cascode and omitting the pass transistors show that bandwidths up to 100 kHz are achievable.

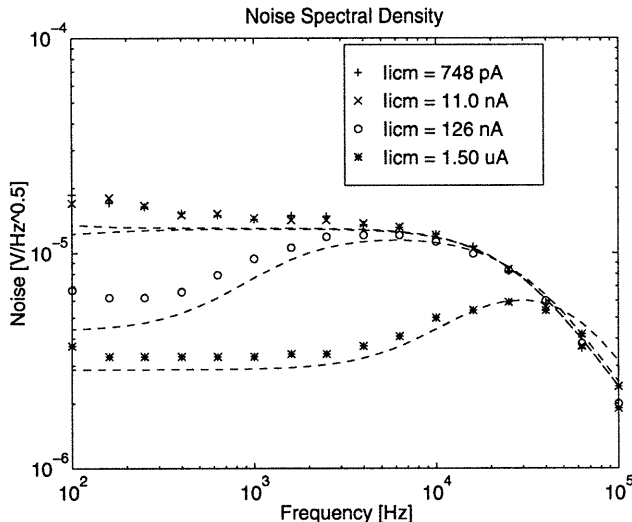


FIGURE 4-14: Output noise spectrum from the fine alignment portion of the sensor with the active cascode circuit disabled. From top to bottom the common-mode input currents were 748 pA, 11.0 nA, 126 nA, and 1.5 μ A corresponding to sensor irradiances of 0.36, 5.3, 60, and 720 μ W/cm². The dashed lines show values obtained using HSPICE simulation.

4.6 Noise

The test set-up for noise measurements was very similar to that used above to measure bandwidth. This time, however, no time varying signal was added to the light emitting diodes. An HP 3585A spectrum analyzer was used to measure the noise voltage at the output. The noise was measured at five frequencies per decade over the range from 100 Hz to 100 kHz. At frequencies below 100 Hz the measurements were found to be corrupted by additional noise arising from mechanical vibrations in the system. Data was taken at four different levels of illumination obtained by placing neutral density filters in front of the camera lens. The noise measurements were only made for the fine alignment portion of the sensor since it determines the ultimate accuracy.

The first set of measurements was made with the active cascode circuit bypassed and is shown in Figure 4-14. As noted in the previous chapter, the noise spectrum divides into two regions. At low frequencies the noise is primarily shot noise from the input devices in the amplifier plus the shot noise from the photodiodes and is given by Equation 3.69. This equation predicts $v_{on} = 3.0 \mu\text{V}/\sqrt{\text{Hz}}$ for $I_{icm} = 1.5 \mu\text{A}$ (the lowest data set). The measured value was $3.3 \mu\text{V}/\sqrt{\text{Hz}}$, somewhat higher than predicted. The simulation plotted in the figure shows v_{on} even lower because it does not include the noise from the photodiodes themselves.

The noise spectral density at the intermediate frequencies is given by Equation 3.70 and was expected to be about $14 \mu\text{V}/\sqrt{\text{Hz}}$. This agrees well with the experimentally obtained values of around $15 \mu\text{V}/\sqrt{\text{Hz}}$. The high frequency cutoff near 20 kHz also appears as predicted. The RMS noise at the output ranged from 1.7 to 2.4 mV. Peak-to-peak measurements of the noise

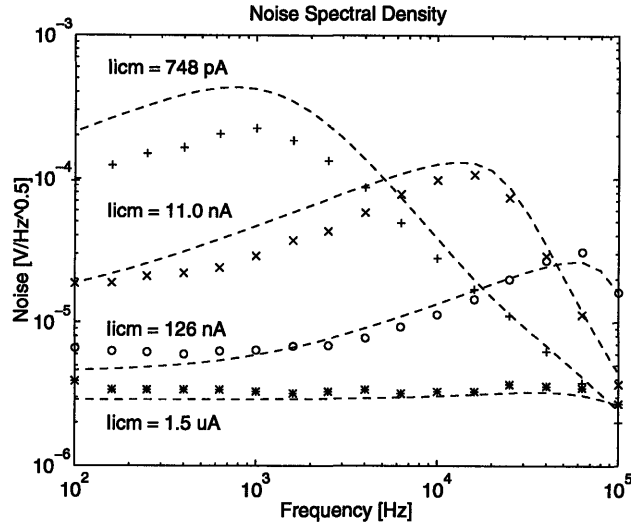


FIGURE 4-15: Output noise spectrum from the fine alignment portion of the sensor taken with the active cascode circuit enabled. Common-mode currents and corresponding irradiances are the same as in Figure 4-14. The dashed lines show values obtained using HSPICE simulation.

waveform generally ranged from 11 to 15 mV which is in good agreement with a $\pm 3\sigma$ extension of this value.

The maximum signal from the fine alignment sensor is about 1.1 V which yields a signal-to-noise ratio of 53 dB. The 3σ noise level of 7.2 mV can be used to estimate the effect of the noise on alignment accuracy. A 7.2 mV error in the fine alignment signal corresponds to a position error of 160 ppm of the sensor field of view. The experiments made earlier achieved better precision than this because the measurements were made with much lower bandwidth, thereby reducing the effect of the sensor noise.

Figure 4-15 shows the noise measurements made with the active cascode circuit included. At low frequencies the dominant noise sources are the photodiode junctions and the input devices of the current amplifier as noted before. At the intermediate frequencies the noise is increased substantially due to the additional noise from the active cascode devices. The expected hump in the spectrum appears at the point where the current amplifier cuts off. However, the data show only a qualitative agreement with the simulations which predict substantially more noise at the low light levels than is observed. There is considerable uncertainty in the estimated flicker noise coefficients used for the simulations. However, changing their values would only move the simulation curves up or down in magnitude and would not explain the deviation seen between the different light levels. A satisfactory explanation for this discrepancy has not been found.

The RMS noise voltage obtained by integrating the noise spectrum ranged from 1.2 to 16 mV. The worst case RMS noise voltage occurred at the intermediate illumination levels for common-mode currents around 10 nA. This corresponds to a signal-to-noise ratio of 59 to 37 dB

for the fine alignment signal. The corresponding 3σ accuracy ranged from 80 to 1100 ppm. At high light levels, the noise performance was superior with the active cascode but at moderate to levels of illumination the active cascode contributed substantially more noise.

4.7 Summary

Tests with the surveyor's mark sensor yielded the expected coarse and fine alignment error signals. They had signal slopes of $140 \mu\text{V}/\text{step}$ and $1.1 \text{ mV}/\text{step}$, respectively, near the aligned position. Alignment repeatability, at low speeds, was 2.1 steps (3σ) and was limited by the resolution of the positioning system and the resolution of the analog-to-digital conversion. This precision is 53 ppm of the total sensor field of view which corresponds to a target displacement of $5.3 \mu\text{m}$ and an image displacement of $0.31 \mu\text{m}$. The dominant pole determining the sensor bandwidth was due to the photodiode depletion capacitance reacting with the amplifier input impedance. The ratio of the sensor irradiance to this capacitance per unit area formed the constraint on the bandwidth. The use of an active cascode moved this pole out by the gain its amplifier and demonstrated a maximum bandwidth of 7.5 kHz. This was limited by pass transistors included for testing purposes and simulations indicate that bandwidths of 100 kHz can be achieved if these devices are removed. Measurements of the sensor noise indicate that it would limit repeatability to between 80 and 1100 ppm depending upon illumination level when using the full bandwidth of the sensor.

Square Grating Mark Alignment Sensor

5.1 Introduction

Many alignment systems use gratings as the alignment marks and detect the moiré fringe signal created when two such marks are superposed. The chief advantage of using a grating is the large signal-to-noise improvement compared to detecting a single edge. The resulting alignment signal is periodic with the grating spacing, but the signal strength is determined by the grating size. Thus, small displacements, a fraction of a pitch spacing, can be detected with a relatively large signal accumulated over the entire target area. Grating alignment techniques are also insensitive to local defects within the grating, although they are very sensitive to distortions of the entire grating.

This chapter describes an alignment sensor designed to take advantage of the strengths of gratings. It begins by introducing the square grating mark used for the alignment and outlining how misalignment can be detected with this mark. An edge detection scheme is presented that can initially position the mark to within one pitch spacing. A fine alignment method is then discussed that detects the moiré fringe pattern formed between the target and sensor gratings. This alignment technique is very similar to that used by King and Berry with circular gratings (see Section 2.4.5). An interpretation of the sensor as a spatial phase detector is given. Following the discussion of the algorithm, the circuits and other implementation issues are presented. The main focus is on an absolute value circuit used to perform the edge detection in the coarse alignment mode. The signal amplifiers are otherwise similar to those used in the surveyor's mark sensor discussed previously.

5.2 Alignment Method

The square grating mark is shown in Figure 5-1. It consists of a concentric set of alternating light and dark squares whose sides are spaced equally apart. When two such patterns are overlaid, as shown in part (b) of the figure, a large displacement is indicated by the shape and size of the resulting crosshatched regions. Detecting these crossings forms the basis of the initial coarse positioning of the target mark. If the two gratings have slightly different pitch spacings, then a geometric moiré fringe will appear when the two gratings are near alignment as shown in Figure 5-2. The moiré fringe is very sensitive to the relative displacement of the underlying gratings and is used to generate the fine alignment error signal.

5.2.1 Coarse Alignment

In order to detect the crosshatched regions shown in Figure 5-1(b) the sensor itself forms a square grating pattern. The photodiodes are laid out as a nested set of squares which mates with the image of the alignment mark as shown in Figure 5-3. The photodiodes are grouped into quadrants labeled with the cardinal directions N, S, E, and W referring to the side of the squares on which they lie. In the figure an image of a square grating target, displaced to the upper right, is shown in grey. The resulting light and dark bars lie across the rows of photodiodes in the north and east quadrants indicating a displacement in that direction. By detecting the number of these edges in each quadrant one can form an error signal related to the image's misalignment.

Edge Counting

In order to generate the coarse alignment signal, some method of detecting and measuring the edges is required. To understand how this is done, we examine the problem in one dimension as shown in Figure 5-4. The irradiance along one row of photodiodes will be modeled as a brightness function $E(x)$. Each photodiode samples this function by integrating it over the area of the diode to yield a sequence of photocurrents

$$I[n] = 2RP_d \int_{nP_d}^{(n+1)P_d} E(x) dx. \quad (5.1)$$

From this sequence a count of the number of edges must be obtained. One way to accomplish this is to choose a function $I_f(I[n], I[n-1])$ that measures, in some sense, the “amount” of an edge between photodiodes n and $n-1$. By summing the output of this function over all the photodiodes in a quadrant one can obtain a measure of the total number of edges crossing the array. The function should have a positive output that is a symmetric function of the two photocurrents. One possibility is the absolute value function

$$I_f(I_1, I_2) = |I_1 - I_2|. \quad (5.2)$$

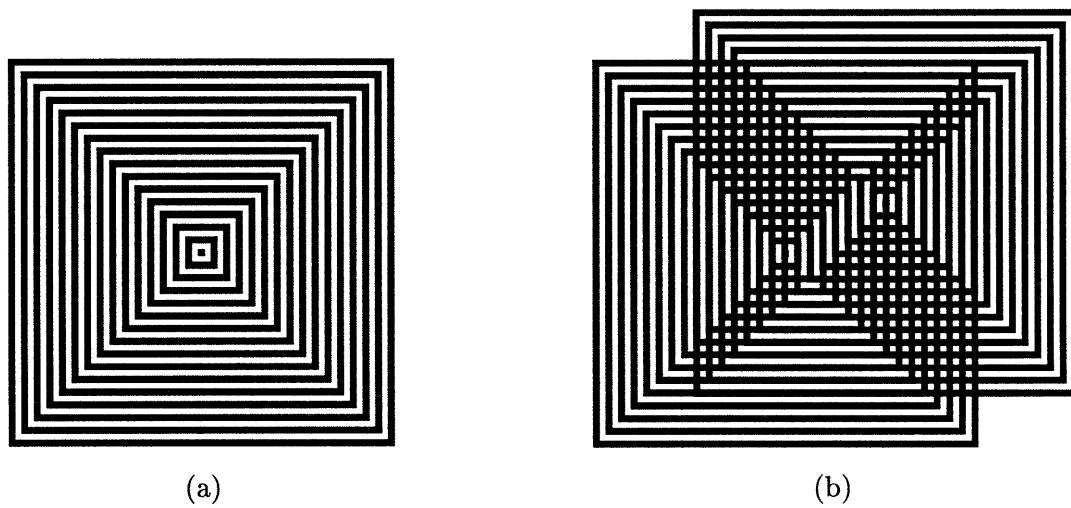


FIGURE 5-1: (a) A square grating and (b) the resulting pattern when two such gratings are overlaid and misaligned. The crosshatched pattern is detected to generate a coarse alignment signal.

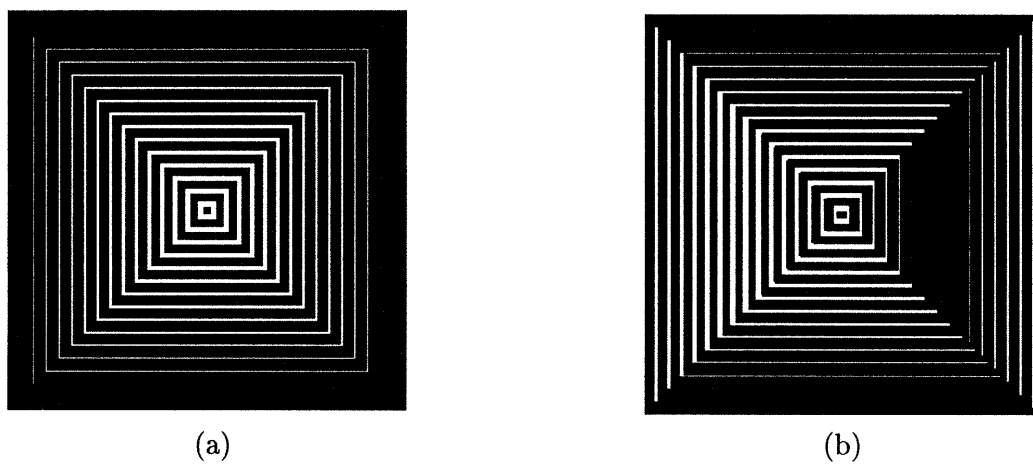


FIGURE 5-2: Moiré patterns from overlaid square gratings whose pitch spacings differ by 3.3 percent. Part (a) shows the beat fringe which appears in the aligned position and part (b) shows the fringe when the larger pitch grating is shifted one quarter period to the right. The fringe position is detected to generate a fine alignment signal.

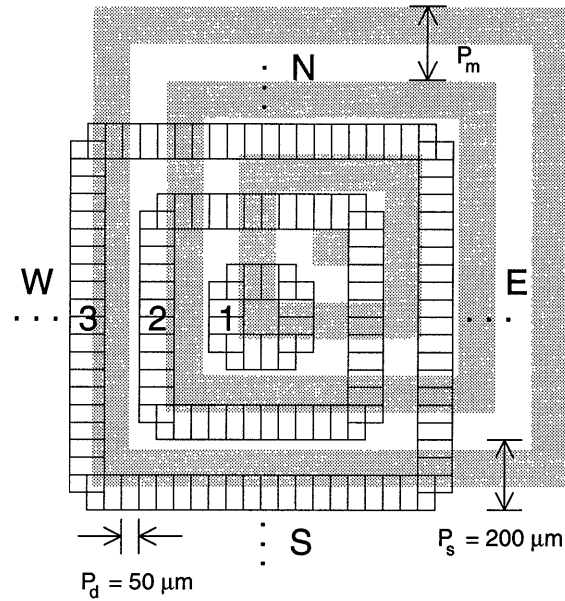


FIGURE 5-3: Rows of photodiodes arranged in concentric squares form the square grating sensor. The shaded regions represent the image of an offset mark. Edges of the mark crossing the rows in the north and east quadrants indicate a displacement in that direction and are detected to generate the coarse alignment error signal.

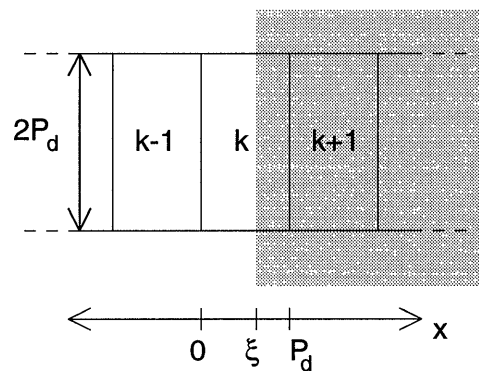


FIGURE 5-4: A single edge falling across the k -th photodiode in a row. The edge signal is obtained by summing the absolute value of the difference between adjacent photocurrents over the entire row.

Other possible functions are the square of the difference or a constant provided the difference exceeds some threshold. The total edge signal is obtained by summing the function over all photocurrents from each row in a quadrant

$$I_E = \sum_{m=1}^{N_s} \sum_{n=2}^{N_r[m]} I_f(I_E[n, m], I_E[n-1, m]). \quad (5.3)$$

Here N_s is the number of squares in the sensor array, $N_r[m]$ is the number of photodiodes in row m , and $I_E[n, m]$ is the photocurrent from photodiode n in row m of the east quadrant.

The absolute value function has the nice property that the summed signal remains constant regardless of the position of the edge. This is not the case for functions like the squared difference or a threshold function. To demonstrate this, consider the situation depicted in Figure 5-4 where the illuminated photodiodes receive irradiance E_0 and the shaded photodiodes receive no irradiance. The resulting photocurrents will be

$$I[n] = \begin{cases} I_0 & n < k \\ I_0\xi & n = k \\ 0 & n > k \end{cases} \quad (5.4)$$

where $I_0 \triangleq 2E_0RP_d^2$ is the photocurrent of a fully illuminated photodiode. For $E_0 = 10 \mu\text{W}/\text{cm}^2$ and $R = 0.13 \text{ A/W}$ this is only 65 pA. The total signal current from this single edge after summing across the entire row is

$$I_e = I_f(I[k], I[k-1]) + I_f(I[k+1], I[k]) \quad (5.5)$$

$$I_e = I[k-1] - I[k] + I[k] - I[k+1] \quad (5.6)$$

$$I_e = I_0. \quad (5.7)$$

This turns out to be a constant equal to I_0 regardless of the edge position ξ . If the function $I_f(I_1, I_2)$ were a squared difference or threshold function, then the summed signal would vary periodically with the edge position ξ with a period of P_d . When multiple edges are present, as in the grating, the summed current over a row of photodiodes will be $N_e I_0$, where N_e is the number of edges falling on the array, provided that the edges are greater than $2P_d$ apart. If the edges are closer than this spacing, then the signal will be reduced and fluctuate periodically with the image position. To ensure this is not the case, the photodiodes are sized with $P_d = P_s/4$ since the grating has two edges per period. The pitch of the mark is actually slightly greater than the sensor pitch in order to generate the moiré fringe, as will be discussed later.

Using the absolute value function to measure edges has some practical difficulties. For instance, consider the effects of some fixed pattern noise in the image by modeling each photocurrent as the sum of an ideal value $V[n]$ with some deviation $D[n]$

$$I[n] = V[n] + D[n]. \quad (5.8)$$

The deviation can arise from a variety of sources such as mismatch between the photodiodes, spatial noise in the image of the target grating, as well as temporal noise such as shot noise in the photodiodes. The total signal is then

$$I_{\Sigma} = \sum_n |\Delta V[n] + \Delta D[n]| \quad (5.9)$$

where

$$\Delta V[n] \triangleq V[n] - V[n - 1] \quad (5.10)$$

$$\Delta D[n] \triangleq D[n] - D[n - 1]. \quad (5.11)$$

The terms $\Delta V[n]$ are the desired edge signals and the $\Delta D[n]$ are noise.

Now consider the fact that most of the photodiodes in a quadrant will have no edge present when the target image is displaced. The signal $\Delta V[n]$ will be zero for most of the terms in the summation and one can write

$$I_{\Sigma} = \sum_{n \in M} |\Delta V[n] + \Delta D[n]| + \sum_{n \in \bar{M}} |\Delta D[n]| \quad (5.12)$$

where the set M denotes those absolute value circuits with an edge signal. Because the set M is usually small compared to its complement, the summed signal can easily become small relative to the total amount of noise. Furthermore, the $D[n]$ are generally proportional to the brightness of the image at that position. This brightness varies dramatically as the light and dark bars of the grating pattern move across the sensor and will cause strong periodic variation in the summation of the $\Delta D[n]$ terms.

A similar problem arises due to imperfections in the implementation of the absolute value function. For example, if the circuit which computes the absolute value has a gain mismatch such as

$$I_f(I_1, I_2) = |I_1 - (1 + \delta)I_2| \quad (5.13)$$

then the total signal will have the form

$$I_{\Sigma} = \sum_{n \in M} |\Delta I[n] - \delta I[n - 1]| + \sum_{n \in \bar{M}} |\delta I[n - 1]|. \quad (5.14)$$

Again, note the potentially large number of terms, each contributing a small error, which could overwhelm the actual signal.

A solution to this difficulty is to implement a slightly modified absolute value function with a cutoff threshold. For signals below the cutoff point, the output is zero. If the cutoff is set larger than the expected size of the noise deviations, then the modified function can greatly reduce the total noise coming from the many locations without any edge signal. One way to define such a modified absolute value is to build in a gain mismatch δ ($\delta > 0$) and define the

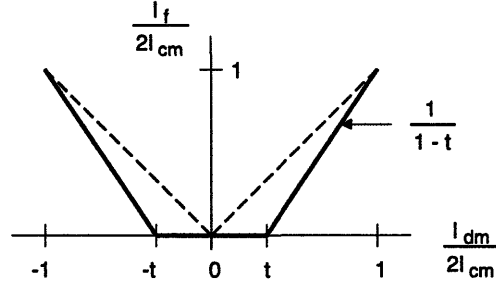


FIGURE 5-5: Plot of the modified absolute value input/output characteristic normalized to the common-mode input currents.

function as

$$I_f(I_1, I_2) = \begin{cases} I_1 - (1 + \delta)I_2 & \frac{I_1}{I_2} > (1 + \delta) \\ I_2 - (1 + \delta)I_1 & \frac{I_2}{I_1} > (1 + \delta) \\ 0 & \frac{1}{1 + \delta} < \frac{I_1}{I_2} < (1 + \delta) \end{cases} \quad (5.15)$$

It is informative to look at this function in terms of the differential-mode and common-mode currents (as defined in Equation 3.28)

$$I_f(I_{dm}, I_{cm}) = \begin{cases} \frac{2I_{cm}}{1-t} \left(\frac{I_{dm}}{2I_{cm}} - t \right) & \frac{I_{dm}}{2I_{cm}} > t \\ \frac{2I_{cm}}{1-t} \left(-\frac{I_{dm}}{2I_{cm}} - t \right) & \frac{I_{dm}}{2I_{cm}} < -t \\ 0 & -t < \frac{I_{dm}}{2I_{cm}} < t \end{cases} \quad (5.16)$$

where the cutoff threshold t is defined as

$$t \triangleq \frac{1}{2} \frac{\delta}{1 + \frac{1}{2}\delta} \quad (5.17)$$

Figure 5-5 shows this function plotted on axes normalized to $2I_{cm}$. The dashed line shows what an ideal absolute value looks like. The essential feature of this circuit is that the output is zero when the input currents are close together. The meaning of “close” is determined by the ratio of the input currents rather than their difference. The output is zero when

$$\frac{1-t}{1+t} < \frac{I_1}{I_2} < \frac{1+t}{1-t} \quad (5.18)$$

For example, if a cutoff threshold of $t = 1/3$ were chosen, the output would be zero unless one current exceeded twice the value of the other. There are other ways one could choose to define an absolute value with some kind of cutoff. This particular choice has the advantage that the cutoff point for zero output depends upon the ratio of the input currents rather than on their

absolute difference. Thus the cutoff point does not have to be adjusted relative to the level of illumination.

The modified absolute value circuit yields an edge signal that is no longer constant as the edge moves across the photodiode array. Returning to the situation illustrated in Figure 5-4, it is useful to examine this fluctuation more closely. The effect of imperfect contrast in the image will be accounted for by modeling the dark level of illumination as having brightness ϵE_0 . The resulting photocurrents are then

$$I[n] = \begin{cases} I_0 & n < k \\ I_0\xi + I_0(1-\xi)\epsilon & n = k \\ \epsilon I_0 & n > k \end{cases} . \quad (5.19)$$

The summed current measuring this edge is found by evaluating Equation 5.7 which yields a function of the edge position ξ that is periodic with period P_d

$$I_e = \begin{cases} \frac{I_0}{1-t} \left\{ 1-t-\epsilon-t\epsilon - (1+t)(1-\epsilon)\frac{\xi}{P_d} \right\} & 0 < \xi < \xi_1 \\ \frac{I_0}{1-t} \left\{ 1-t-\epsilon-3t\epsilon-2t(1-\epsilon)\frac{\xi}{P_d} \right\} & \xi_1 < \xi < \xi_2 \\ \frac{I_0}{1-t} \left\{ -2t\epsilon + (1-t)(1-\epsilon)\frac{\xi}{P_d} \right\} & \xi_2 < \xi < P_d \end{cases} \quad (5.20)$$

where

$$\xi_1 \triangleq \frac{2t\epsilon}{(1-t)(1-\epsilon)}P_d \quad (5.21)$$

$$\xi_2 \triangleq \frac{1-t-\epsilon-t\epsilon}{(1+t)(1-\epsilon)}P_d \quad (5.22)$$

and ϵ must satisfy

$$\epsilon < \left(\frac{1-t}{1+t} \right)^2 . \quad (5.23)$$

A plot of this function is shown in Figure 5-6. Note the substantial variation in the signal as the edge position changes. There is a dramatic drop in the signal even for small values of ϵ which emphasizes the importance of good image contrast. If ϵ becomes too large

$$\epsilon > \left(\frac{1-t}{1+t} \right)^2 \quad (5.24)$$

then the total output current actually goes to zero for some positions of the edge

$$I_e = \begin{cases} \frac{I_0}{1-t} \left\{ 1-t-\epsilon-t\epsilon - (1+t)(1-\epsilon)\frac{\xi}{P_d} \right\} & 0 < \xi < \xi_2 \\ 0 & \xi_2 < \xi < \xi_1 \\ \frac{I_0}{1-t} \left\{ -2t\epsilon + (1-t)(1-\epsilon)\frac{\xi}{P_d} \right\} & \xi_1 < \xi < P_d \end{cases} . \quad (5.25)$$

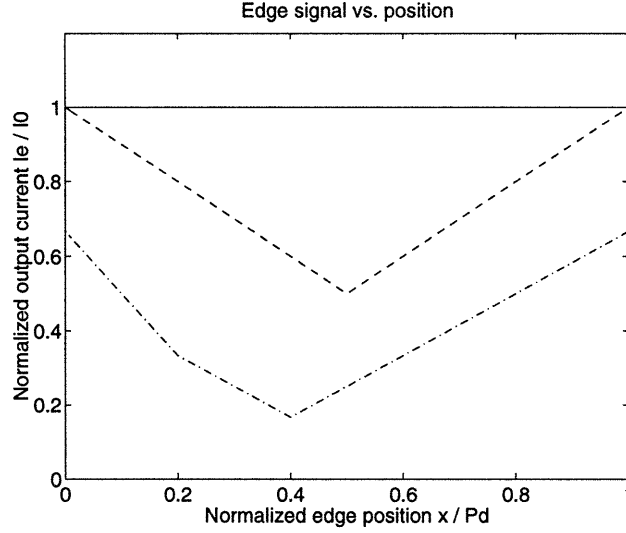


FIGURE 5-6: Plots of the normalized edge signal I_e/I_0 versus the normalized edge position ξ/P_d . The top plot shows the case with no cutoff and perfect contrast ($t = 0$ and $\epsilon = 0$), the middle plot shows a cutoff threshold $t = 1/3$ with perfect contrast, and the bottom plot shows the same cutoff point with contrast $\epsilon = 1/6$.

The average value of this function over one period is given by

$$\bar{I}_e = I_0 \frac{(1 - t - \epsilon - t\epsilon)^2}{(1 - \epsilon)(1 + t)(1 - t)^2}. \quad (5.26)$$

For the example $t = 1/3$ and $\epsilon = 1/6$ the average value of the signal is $\bar{I}_e = (2/5)I_0$ which is a sixty percent reduction from the maximum signal. A high contrast image ($\epsilon \ll 1$) will naturally yield a stronger edge detection signal. Little can be done about this in the design other than to assure a high contrast target that is well illuminated. There is a trade-off between the cutoff threshold t and the average signal value \bar{I}_e . Choosing a small cutoff point will increase the signal strength, but at the same time allow more noise into the total signal. The choice for a cutoff point that is in some sense optimal is not clear and depends on more specific knowledge of the particular characteristics of the mismatch and noise that one hopes to reduce.

Approximate Error Signal

The error signals are obtained by summing the edge detection currents over each quadrant as shown in Equation 5.3. These currents are then subtracted to obtain the x - and y -direction error signals as

$$e_x = I_E - I_W \quad (5.27)$$

$$e_y = I_N - I_S. \quad (5.28)$$

It is difficult to derive an exact expression for these currents as a function of the target mark position. However, an approximate analysis which yields reasonable results will be given here.

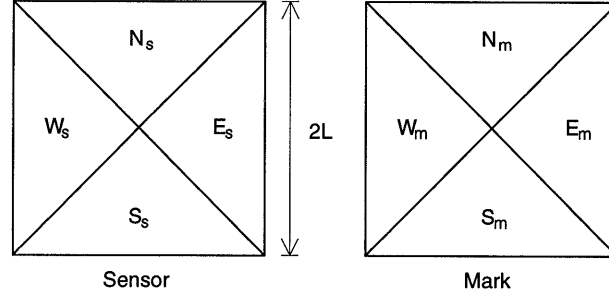


FIGURE 5-7: Dividing the sensor and mark gratings into their respective quadrants.

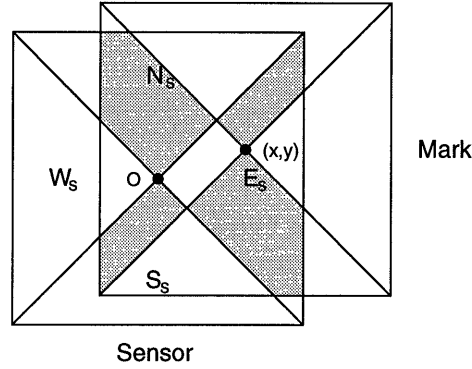


FIGURE 5-8: When the two patterns are overlaid, edge crossings will occur where the horizontal grating of one pattern overlaps the vertical grating of another (shaded regions).

In order to determine the strength of the error signal, one needs to describe the number of edge crossings in each quadrant as a function of position. This can be done approximately by measuring the areas of the regions in which edge crossings can occur and then multiplying by the number of edge crossings per unit area expected in those regions. Figure 5-7 shows abstracted representations of the sensor and target gratings divided into quadrants. The north and south quadrants contain horizontal gratings whereas the east and west quadrants contain vertical gratings. When the two patterns are overlaid, crossings will occur only in those regions where the horizontal and vertical gratings overlap. These are the shaded regions depicted in Figure 5-8. The quadrant error signals will be proportional to the amount of this overlap in each quadrant of the sensor. Using set notation ($\vee \triangleq$ or, $\wedge \triangleq$ and) these areas will be

$$a_N = (E_m \vee W_m) \wedge N_s \quad (5.29)$$

$$a_S = (E_m \vee W_m) \wedge S_s \quad (5.30)$$

$$a_E = (N_m \vee S_m) \wedge E_s \quad (5.31)$$

$$a_W = (N_m \vee S_m) \wedge W_s. \quad (5.32)$$

The areas of these regions can be solved for as functions of the mark position. The resulting equations depend upon the quadrant in which the center of the mark lies and are given below.

When the mark is in the east quadrant ($0 < x < L, |y| < x$)

$$a_N = L(x + y) - \frac{3}{4}x^2 - \frac{1}{2}xy - \frac{1}{4}y^2 \operatorname{sgn} y \quad (5.33)$$

$$a_S = L(x - y) - \frac{3}{4}x^2 + \frac{1}{2}xy + \frac{1}{4}y^2 \operatorname{sgn} y \quad (5.34)$$

$$a_E = 2Lx - \frac{3}{2}x^2 \quad (5.35)$$

$$a_W = 0. \quad (5.36)$$

When the mark is in the west quadrant ($-L < x < 0, |y| < -x$)

$$a_N = L(-x + y) - \frac{3}{4}x^2 + \frac{1}{2}xy - \frac{1}{4}y^2 \operatorname{sgn} y \quad (5.37)$$

$$a_S = L(-x - y) - \frac{3}{4}x^2 - \frac{1}{2}xy + \frac{1}{4}y^2 \operatorname{sgn} y \quad (5.38)$$

$$a_E = 0 \quad (5.39)$$

$$a_W = -2Lx - \frac{3}{2}x^2. \quad (5.40)$$

When the mark is in the north quadrant ($0 < y < L, |x| < y$)

$$a_N = 2Ly - \frac{3}{2}y^2 \quad (5.41)$$

$$a_S = 0 \quad (5.42)$$

$$a_E = L(x + y) - \frac{3}{4}y^2 - \frac{1}{2}xy - \frac{1}{4}x^2 \operatorname{sgn} x \quad (5.43)$$

$$a_W = L(-x + y) - \frac{3}{4}y^2 + \frac{1}{2}xy + \frac{1}{4}x^2 \operatorname{sgn} x. \quad (5.44)$$

When the mark is in the south quadrant ($-L < y < 0, |x| < -y$)

$$a_N = 0 \quad (5.45)$$

$$a_S = -2Ly - \frac{3}{2}y^2 \quad (5.46)$$

$$a_E = L(x - y) - \frac{3}{4}y^2 + \frac{1}{2}xy - \frac{1}{4}x^2 \operatorname{sgn} x \quad (5.47)$$

$$a_W = L(-x - y) - \frac{3}{4}y^2 - \frac{1}{2}xy + \frac{1}{4}x^2 \operatorname{sgn} x. \quad (5.48)$$

The error signals for the x and y directions are formed by subtracting the quadrant signals

$$a_x = a_E - a_W \quad (5.49)$$

$$a_y = a_N - a_S \quad (5.50)$$

and can be written as

$$a_x = \begin{cases} 2Lx - \frac{3}{2}x^2 \operatorname{sgn} x & |y| < |x| < L \\ 2Lx - xy \operatorname{sgn} y - \frac{1}{2}x^2 \operatorname{sgn} x & |x| < |y| < L \end{cases} \quad (5.51)$$

$$a_y = \begin{cases} 2Ly - xy \operatorname{sgn} x - \frac{1}{2}y^2 \operatorname{sgn} y & |y| < |x| < L \\ 2Ly - \frac{3}{2}y^2 \operatorname{sgn} y & |x| < |y| < L \end{cases}. \quad (5.52)$$

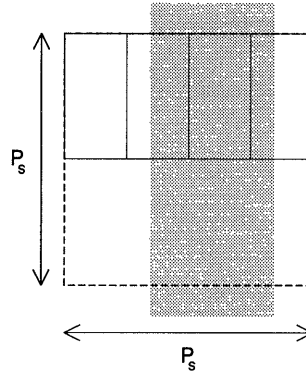


FIGURE 5-9: A small area element of width P_s will contain four photodiodes detecting two edges in the regions where edge crossings occur.

The equations derived so far are measures of the area in which edge crossings can occur. Next, one must estimate the amount of edge signal that will be obtained per unit area. Consider a small area of the sensor array one pitch spacing across as illustrated in Figure 5-9. Within this area will be four photodiodes and roughly two edge transitions in a region with edge crossings. The amount of current signal due to these two edges will be a complex function of the edge position as shown by Equations 5.20 and 5.25. Generally, there will be many such edges except in the case of a small displacement. Because the sensor and mark pitches differ slightly, the edges will occur at different positions relative to the photodiodes. Thus the function $I_e(\xi)$ in Figure 5-6 will be sampled at many points. For these reasons one can assign the average value \bar{I}_e as the average signal expected from each edge. Thus, in an area P_s^2 one would expect a signal of about $2\bar{I}_e$ on average. With this relationship established between area and the signal current it is possible to write down approximate equations for the coarse alignment error currents

$$e_x = 2 \frac{\bar{I}_e}{P_s^2} a_x \quad (5.53)$$

$$e_y = 2 \frac{\bar{I}_e}{P_s^2} a_y \quad (5.54)$$

with similar equations holding for I_N, I_S, I_E , and I_W .

Figure 5-10 shows the error signals expected from the sensor using this model and assuming the parameters were $I_0 = 65$ pA, $\epsilon = 0.173$, and $t = 0.33$. Part (a) shows the x -direction error signal along the x -axis across the sensor field of view and part (b) shows the negative error vector field. As was the case for surveyor's mark sensor, the vector field is not radially directed into the origin. Thus, the trajectories will take slightly curving paths to the center during alignment.

This analysis has only addressed the case when the center of the target mark falls within the domain of the sensor ($|x| < L, |y| < L$). However, error signals will still be generated as long as any portion of the target overlaps any portion of the sensor ($|x| < 2L, |y| < 2L$). It is a conjecture that error signals over this area are still sufficient to guide an alignment system. This

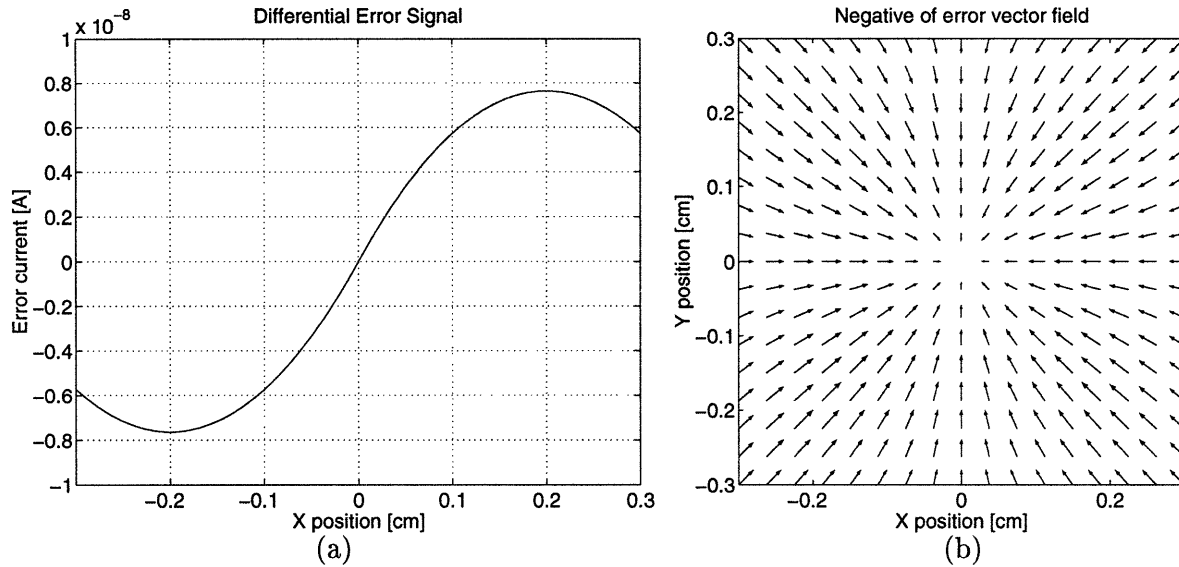


FIGURE 5-10: Predicted coarse alignment (a) x -direction error signal currents along the x -axis and (b) negative error vectors in sensor coordinates.

has not been determined analytically. If it is true, then the capture region of this technique is effectively twice as large in each dimension as the sensor itself.

Error Sources

Some possible error sources to consider are the effects of photodiode mismatch, illumination variations, and rotations of the target. Mismatch in the responsivity between the photodiodes will effect the shape of the absolute value input/output function. The result will be to asymmetrically shift the cutoff points and modify the slopes of the function shown in Figure 5-5. It is important to ensure that the cutoff threshold does not become too small. This is done by choosing the built-in gain error δ large enough to accommodate the expected variation in photodiode matching. In practice, the variation in the cutoff point due to photodiode mismatch is negligible compared to the effect of device mismatch in the absolute value circuit.

Variation in the level of illumination across the target will affect the coarse alignment signal because the edge detection signal is proportional to the sensor irradiance. If the irradiance is larger on the right side of the sensor than on the left, then the total edge detection signal will also be larger on the right. This will cause an asymmetry in the alignment signal. However, it will not cause an offset in the aligned position because there are no edges to be detected when the patterns are in alignment. Nonuniform illumination and the resulting asymmetry in the alignment signal will not cause a position offset in the null point of the signal.

Rotation of the alignment mark has an interesting effect on the error signal. It is useful to consider the extreme case of a 45 degree rotation. When the mark is aligned there will be edge crossings in all four quadrants of the sensor and the edges will cross the rows of photodiodes

at 45 degree angles. Contrast this with the case of no rotation where the edge crossings are perpendicular to the photodiode rows and there are no edge crossings to detect when the patterns are aligned. As the rotated mark is translated, the number of edge crossings in each quadrant will vary in such a way that a useful alignment signal is still generated. For example, a translation to the right will increase the number of edges in the east quadrant while reducing the number in the west. In fact, the sensor will generate a usable alignment signal for any rotation. The major disadvantage of a rotation is a reduction in the strength of the edge signals. It was noted earlier that in order to obtain maximum signal from an edge crossing, the edges in the grating pattern must have at least one fully illuminated or fully darkened photodiode between each partially illuminated photodiode. The mark pitch must satisfy the constraint

$$P_m \geq 4P_d \quad (5.55)$$

The implemented design uses $P_m = 4P_d(1 + \epsilon_p)$ where ϵ_p is a small number to generate the moiré fringe. Thus, a very small rotation will cause adjacent photodiodes to be partially occluded by edges and reduce the differential signal between them. This will reduce the strength of the coarse alignment signal when the mark is rotated.

5.2.2 Fine Alignment

The fine alignment error signal is obtained by detecting the moiré pattern formed between the sensor grating and the image of the target grating. The moiré fringe is created by slightly mismatching the pitches these two gratings and its position, or spatial phase, is very sensitive to the alignment of the two patterns. The position is detected using the sensor array to implement a spatial phase detector. The alignment system can be viewed as a spatial phase-locked-loop that minimizes the phase error in order to bring the two grating patterns into alignment.

Moiré Fringe Generation

Geometric moiré patterns are usually understood to occur when two gratings are overlaid. For example, take the case of two one-dimensional sinusoidal gratings with transmission functions

$$r(x) = \frac{1}{2}(1 + \cos k_s x) \quad (5.56)$$

$$m(x) = \frac{1}{2}(1 + \cos k_m(x - \xi)) \quad (5.57)$$

where ξ denotes a shift, the misalignment, in the position of the $m(x)$ grating. The pitches of the gratings will be assumed to be very close with

$$P_s = (1 + \epsilon_p)P_m \quad (5.58)$$

$$k_s = \frac{2\pi}{P_s} \quad (5.59)$$

$$k_m = \frac{1}{1 + \epsilon_p} \frac{2\pi}{P_s} \quad (5.60)$$

where $\epsilon_p \ll 1$ denotes the pitch mismatch. Multiplying these two functions yields the transmission of the overlaid gratings

$$t(x) = r(x)m(x). \quad (5.61)$$

In the frequency domain this function has components at k_m and k_s as well as at the sum and difference frequencies

$$k_\Sigma = k_s + k_m = \frac{2 + \epsilon_p}{1 + \epsilon_p} \frac{2\pi}{P_s} \quad (5.62)$$

$$k_\Delta = k_s - k_m = \frac{\epsilon_p}{1 + \epsilon_p} \frac{2\pi}{P_s}. \quad (5.63)$$

The moiré fringe is the lowest frequency component. Mathematically it can be defined by either averaging $t(x)$ over one period P_s or by ideal low pass filtering $t(x)$ with a cutoff frequency at $k_s/2$. The low pass filtered version of $t(x)$ in this case is

$$t_L(x) = \frac{1}{4} + \frac{1}{8} \cos k_\Delta \left(x + \frac{\xi}{\epsilon_p}\right). \quad (5.64)$$

The spatially varying portion of $t_L(x)$ is the moiré fringe. There are several things to notice about it. First, it has a period

$$P_\Delta = \frac{1 + \epsilon_p}{\epsilon_p} P_s \quad (5.65)$$

which is much larger than the pitches of the two gratings which formed it. The fringe is the aliased low frequency (or beat frequency) component that appears when two signals with similar frequencies are multiplied. Also note that the position of the fringe is $-1/\epsilon_p$ times the position of the grating. Thus, a small movement of the grating $m(x)$ yields a much larger movement, in the opposite direction, of the moiré fringe.

The generation of the moiré fringe in the alignment sensor will be modeled somewhat differently. Rather than being the result of a multiplication of two grating functions, the fringe arises from the aliasing which occurs when the imaged target grating is sampled by the rows of photodiodes in the sensor. Consider the one-dimensional case where the mark is modeled as a sinusoidal variation in irradiance

$$m(x) = E_0 \frac{1}{2} (1 + \cos k_m(x - \xi)) \quad (5.66)$$

and the sensor is modeled by a square wave responsivity function $r(x)$ that represents a grating of photodiodes as shown in Figure 5-11.

The photodiodes in the sensor can be viewed as transducing the irradiance function $m(x)$ into a sequence of photocurrents by integrating the light across each photodiode row

$$f[n] = \int_{(n-1/2)P_s}^{(n+1/2)P_s} r(x)m(x) dx. \quad (5.67)$$

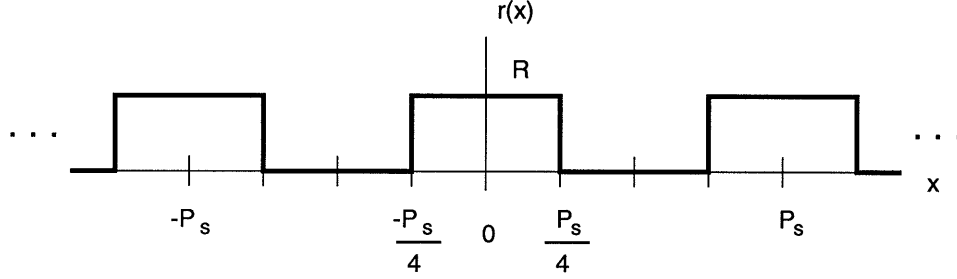


FIGURE 5-11: One-dimensional sensor responsivity function modeling a grating of photodiodes with responsivity R .

The result of this integration yields a sequence of photocurrents that forms a moiré fringe. There is an interesting way to derive this sequence in the frequency domain. Note that the sequence can be expressed as samples of a box-car averaged function

$$m_a(x) = \int_{-\infty}^{\infty} m(u)b(x-u) du \quad (5.68)$$

$$f[n] = m_a(nP_s) \quad (5.69)$$

where

$$b(x) = \begin{cases} R & -\frac{P_s}{4} < x < \frac{P_s}{4} \\ 0 & \text{otherwise} \end{cases} . \quad (5.70)$$

In the frequency domain the convolution becomes multiplication

$$M_a(k) = M(k)B(k) \quad (5.71)$$

where

$$B(k) = \frac{RP_s}{2} \operatorname{sinc} \frac{k}{2k_s} \quad (5.72)$$

$$M(k) = \pi E_0 \left\{ \frac{1}{2} e^{jk_m \xi} \delta(k + k_m) + \delta(0) + \frac{1}{2} e^{-jk_m \xi} \delta(k - k_m) \right\} . \quad (5.73)$$

The spectrum of $M_a(k)$ is shown in Figure 5-12(a) where

$$a_1 = \frac{\pi}{4} E_0 RP_s e^{-jk_m \xi} \operatorname{sinc} \frac{k_m}{2k_s} \quad (5.74)$$

$$a_0 = \frac{\pi}{2} E_0 RP_s \quad (5.75)$$

$$a_{-1} = \frac{\pi}{4} E_0 RP_s e^{jk_m \xi} \operatorname{sinc} \frac{k_m}{2k_s} . \quad (5.76)$$

The sampling process is modeled by multiplying the averaged function by an ideal impulse train ([108] ch. 3)

$$f_p(x) = p(x)m_a(x) \quad (5.77)$$

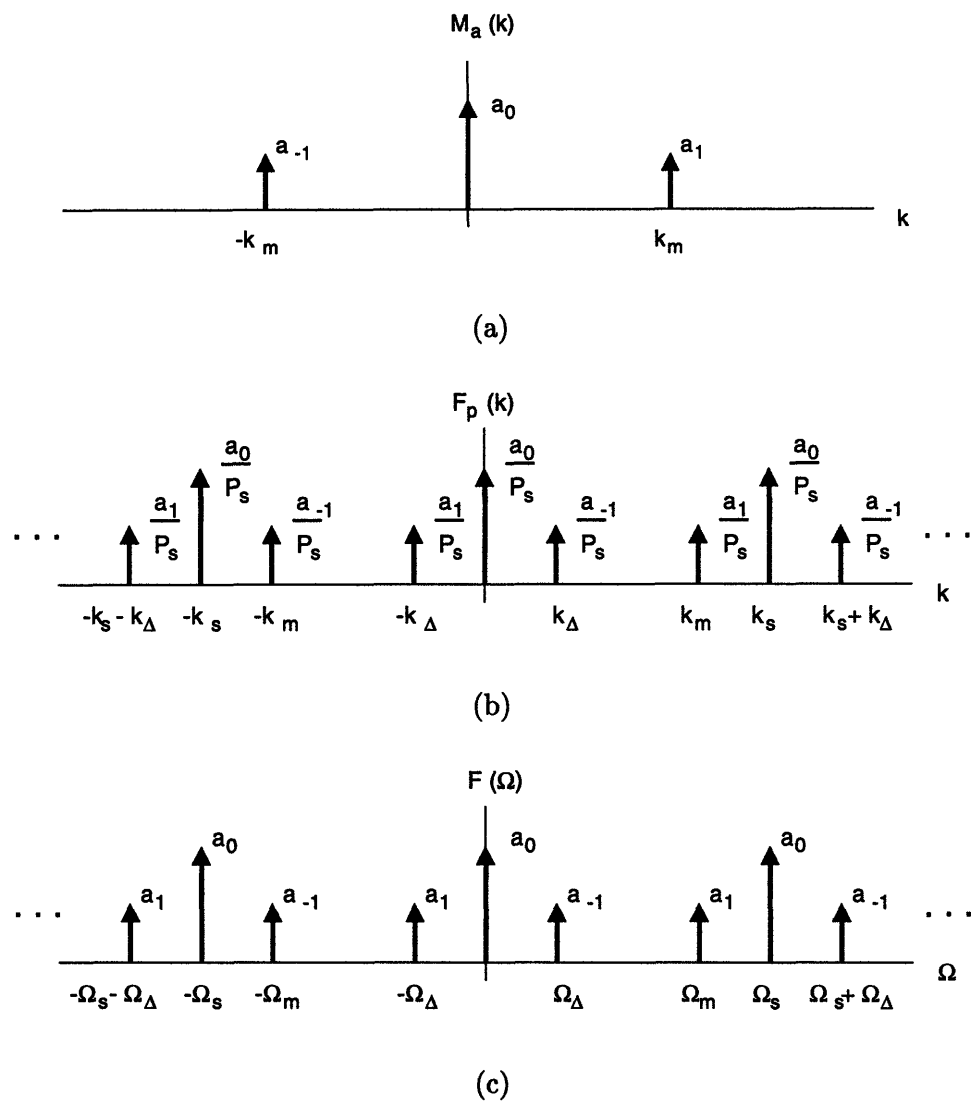


FIGURE 5-12: Fourier transforms of (a) $M_a(k)$ averaged function, (b) $F_p(k)$ after multiplication with the impulse train, and (c) $F(\Omega)$ after conversion to a discrete Fourier transform.

where

$$p(x) = \sum_{n=-\infty}^{\infty} \delta(x - nP_s). \quad (5.78)$$

In the Fourier domain this multiplication becomes a convolution

$$F_p(k) = \frac{1}{2\pi} M_a(k) * P(k) \quad (5.79)$$

where $P(k)$ is also an impulse train with spacing k_s

$$P(k) = k_s \sum_{n=-\infty}^{\infty} \delta(k - nk_s). \quad (5.80)$$

The transform of the sampled function is a sum of shifted replicas of the original function $M_a(k)$ scaled by $1/P_s$ as illustrated in Figure 5-12(b).

Finally, to convert the continuous Fourier transform function $F_p(k)$ to a discrete Fourier transform $F(\Omega)$ one must scale the frequency axis by $\Omega = P_s k$. This also scales the areas of the delta functions by the factor P_s . The resulting discrete time Fourier transform is depicted in Figure 5-12(c) where

$$\Omega_{\Delta} = 2\pi \frac{\epsilon_p}{1 + \epsilon_p} \quad (5.81)$$

$$\Omega_m = 2\pi \frac{1}{1 + \epsilon_p}. \quad (5.82)$$

The inverse discrete Fourier transform yields the sequence

$$f[n] = \frac{E_0 R P_s}{4} \left\{ 1 + A \cos \Omega_{\Delta} \left(n + \frac{\xi}{\epsilon_p P_s} \right) \right\} \quad (5.83)$$

where

$$A = \text{sinc} \frac{k_m}{2k_s} \approx \frac{2}{\pi}. \quad (5.84)$$

This sequence contains the moiré fringe with discrete frequency Ω_{Δ} . The moiré fringe can perhaps be seen more clearly in the band limited reconstruction of a continuous function from this sequence¹

$$f_L(x) = \frac{E_0 R}{4} \left\{ 1 + A \cos k_{\Delta} \left(x + \frac{\xi}{\epsilon_p} \right) \right\}. \quad (5.85)$$

This moiré fringe function is similar to that of Equation 5.64 with the same period and phase shift. The above result could be obtained directly by integrating Equation 5.67. However, viewing this as a sampling problem in the frequency domain shows how the moiré fringe can be interpreted as an aliased signal that arises from undersampling by the sensor array.

¹Reconstruction obtained with a unity gain low pass filter with cutoff at $k_s/2$.

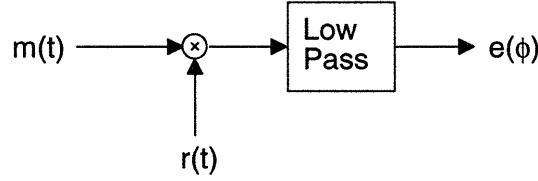


FIGURE 5-13: Signal flow diagram for a phase detector.

Spatial Phase Detection

Referring to Figure 5-2 one can see that a displacement of the target image towards the right reduces the brightness in the east quadrant and increases it in the west quadrant. Thus, an error signal can be formed by subtracting the total signal in the west quadrant from that in the east. In this case the quadrant signal currents are simply the summed currents from all of the photodiodes in the respective quadrants. For example, the east quadrant current is now

$$I_E = \sum_{m=1}^{N_s} \sum_{n=1}^{N_r[m]} I_E[n, m]. \quad (5.86)$$

The fine alignment error signal is formed by subtracting these currents as

$$e_x = I_W - I_E \quad (5.87)$$

$$e_y = I_S - I_N. \quad (5.88)$$

This description of the fine alignment error signal has an interesting interpretation as an implementation of a spatial phase detector. Phase detection of time varying waveforms is often modeled with the basic system shown in Figure 5-13. Here an input signal $m(t)$ is compared to a reference waveform $r(t)$ by multiplying the two signals together and low pass filtering the result to obtain the average value

$$e(\phi) = \frac{1}{T} \int_T r(t)m(t)dt. \quad (5.89)$$

For example, in the case where the signal and reference waveforms are functions of the form

$$r(t) = \sin(\omega t) \quad (5.90)$$

$$m(t) = \cos(\omega t + \phi) \quad (5.91)$$

the error signal is

$$e(\phi) = \frac{1}{2} \sin \phi \quad (5.92)$$

which provides a measure of the phase shift between the signal and reference. When the phase error is zero, the two waveforms are orthogonal. Phase detectors of this type are used to generate the error signal for phase-locked-loops. These systems adjust the phase shift in order to maintain zero phase error and synchronize the signal with the reference. For the alignment

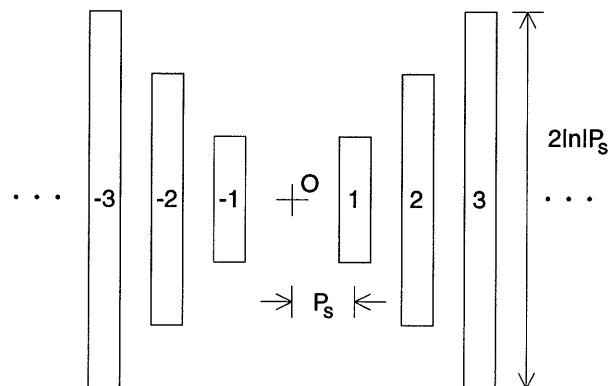


FIGURE 5-14: Simplified two-dimensional responsivity function for the portion of the square grating sensor used to generate the x -direction fine alignment error signal. Rectangles have responsivity R and everywhere else has responsivity zero.

sensor the input waveform is the moiré fringe generated by the sensor array sampling the image of the mark and the reference waveform is the spatial response of the sensor itself.

Several approximations will be made in the following analysis in order to show the analogy between this sensor and a phase detector without becoming burdened by relatively unimportant mathematical details. The first approximation made is to model the responsivity of the sensor array with the two-dimensional function $r(x, y)$ shown in Figure 5-14. In this figure the rectangles represent regions where photodiodes of responsivity R are located. All other regions have zero responsivity. There are three important features to note. First, the horizontal rows of photodiodes in the north and south quadrants are omitted because they do not concern the detection of the x -direction error signal. This discussion will consider the x -direction error signal with the understanding that the y -direction signal is detected with the same technique. Second, the center of this sensor has been widened by an amount P_s . Finally, the vertical extent of the photodiode arrays has been modified slightly. These changes, when compared to the actual layout in Figure 5-3, do not significantly affect the operation of the sensor but simplify the notation to be used in the analysis.

As done previously, the mark is modeled as a sinusoidal grating which varies only in the x -direction as defined by Equation 5.66. In reality, the mark function $m(x)$ should also be modeled as a square wave. This would result in a trapezoidal waveform for the moiré fringe as derived in [71] (section 1.1.3) and can be treated analytically by expanding $m(x)$ in a Fourier series to represent the square wave. Because of the complexity this adds to the analysis, the one term sinusoidal model for $m(x)$ will be used in this discussion. The quadrant error currents are then given by the area integrals

$$I_E = \int_{x>0} \int r(x, y) m(x) dy dx \quad (5.93)$$

$$I_W = \int \int_{x < 0} r(x, y) m(x) dy dx. \quad (5.94)$$

The error signal can be written as a summation

$$e_x = \sum_{-N_s}^{-1} s[n] - \sum_0^{N_s} s[n] \quad (5.95)$$

where $s[n]$ is the sequence of photocurrents collected by each vertical photodiode

$$s[n] = \int_{(n-1/2)P_s}^{(n+1/2)P_s} \int_{-\infty}^{\infty} r(x, y) m(x) dy dx. \quad (5.96)$$

Since each stripe contains a photodiode row of width $P_s/2$, height $2|n|P_s$, and constant responsivity R , the integral over y can be carried out yielding

$$s[n] = 2|n|P_s \int_{(n-1/2)P_s}^{(n+1/2)P_s} r(x) m(x) dy dx \quad (5.97)$$

where $r(x)$ is the function shown in Figure 5-11.² The remaining portion of this integral is the moiré fringe sequence $f[n]$ defined in Equation 5.67 and solved for in Equation 5.83. Thus, the sequence of photocurrents is given by

$$s[n] = 2|n|P_s f[n] \quad (5.98)$$

and the error signal can be written as the sum

$$e_x = \sum_{-N_s}^{N_s} d[n] f[n] \quad (5.99)$$

where the new sequence $d[n]$ is defined as

$$d[n] \triangleq -2nP_s. \quad (5.100)$$

This sequence is the the detector weighting of the moiré fringe sequence $f[n]$.

Summing the sequence to get the error current e_x is difficult so an additional approximation is made by turning it into an integral. This is accomplished by defining

$$x \triangleq nP_s \quad (5.101)$$

$$\Delta x \triangleq P_s \quad (5.102)$$

$$N \triangleq \frac{L}{P_s} \quad (5.103)$$

and approximating

$$e_x = \sum_{-N_s}^{N_s} d\left[\frac{1}{P_s}(nP_s)\right] f\left[\frac{1}{P_s}(nP_s)\right] \frac{1}{P_s} P_s \quad (5.104)$$

²To treat the actual sensor layout would require integrating over stripes from $(n-1)P_s$ to nP_s with photodiode heights of $(|2n+1|-1/2)P_s$.

with

$$e_{cx} = \int_{-L}^L d(x) f_L(x) dx \quad (5.105)$$

where

$$d(x) \triangleq d\left[\frac{x}{P_s}\right] = -2x \quad (5.106)$$

$$f_L(x) \triangleq \frac{1}{P_s} f\left[\frac{x}{P_s}\right] = \frac{E_0 R}{4} \left\{ 1 + A \cos k_\Delta \left(x + \frac{\xi}{\epsilon_p}\right) \right\} \quad (5.107)$$

with A as defined previously and $f_L(x)$ is identical to Equation 5.85. Equation 5.105 is in the same form as the phase error Equation 5.89. The input signal is the moiré fringe $f(x)$ and the reference is a ramp function $d(x)$ that models the spatial response of the sensor.

With the previous approximations it is a simple matter to solve the integral to find the error signal for this spatial phase detector

$$e_{cx}(\xi) = I_{max} a(k_\Delta) \sin k_\Delta \frac{\xi}{\epsilon_p} \quad (5.108)$$

where

$$I_{max} \triangleq \frac{E_0 R L^2}{2} \quad (5.109)$$

$$a(k_\Delta) \triangleq \frac{2A}{k_\Delta L} \left(\frac{1}{k_\Delta L} \sin k_\Delta L - \cos k_\Delta L \right). \quad (5.110)$$

This can be written in terms of the mark position as

$$e_{cx}(\xi) = I_{max} a(k_\Delta) \sin k_m \xi. \quad (5.111)$$

The resulting error signal is a sinusoidal function of the mark position with a period equal to the pitch of the mark grating. The term I_{max} in front represents the maximum possible current from one quadrant of the sensor. The second term $a(k_\Delta)$ determines the amplitude of the error signal and is a function of the moiré fringe spatial frequency.

It is important to look closely at the amplitude factor. It is really a function of the pitch mismatch ϵ_p since both A and k_Δ depend upon this parameter. Figure 5-15 shows a plot of $a(k_\Delta)$ as ϵ_p is varied. There are several interesting things to note in this plot. First, for $\epsilon_p = 0$ the amplitude factor is zero. This occurs because k_Δ is also zero and the moiré fringe has infinite extent when the grating pitches are matched. Zeros also occur at other special spatial frequencies. These are akin to the orthogonality that exists between two sinusoids when one has a frequency that is an integral multiple of the other. In this case rather than two sinusoids, the functions are a sinusoid $f_L(x)$ and the detector ramp function $d(x)$. The amplitude factor reaches its highest and lowest extrema for small ϵ_p on either side of the origin. This occurs when the period of the moiré fringe is comparable to the length of the sensor. Larger pitch mismatch results in lower amplitude extrema because more periods of the moiré fringe fit in the sensor dimensions. This reduces the potential disparity in the quadrant signals.

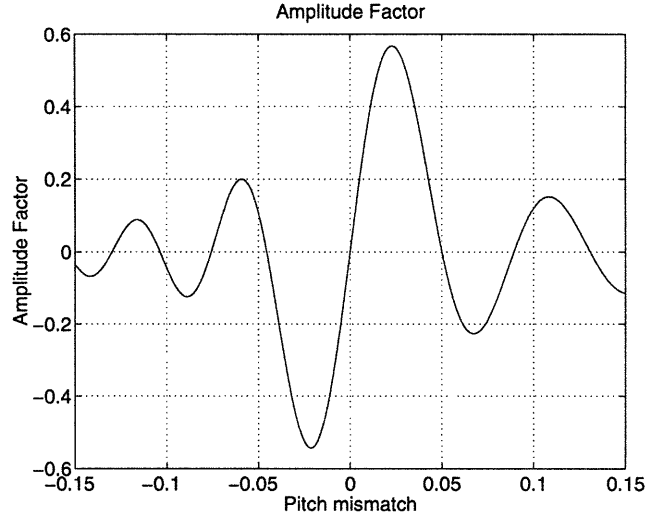


FIGURE 5-15: Plot of the amplitude factor $a(k_\Delta)$ over a range of pitch mismatch ϵ_p near zero. The maximum signal is obtained for $\epsilon_p = 0.023$ where the amplitude factor is 0.57.

Note that for small ϵ_p one can treat the factor A as a constant equal to approximately $2/\pi$. In this case the amplitude factor is only a function of the spatial frequency of the moiré fringe k_Δ . The maximum of 0.57 occurs when $k_\Delta = 2.082$ ($\epsilon_p = 0.023$). This implies that the moiré fringe bears a fixed relationship to the sensor width which can be written as

$$\frac{P_\Delta}{L} = \frac{2\pi}{2.082} \approx 3.0. \quad (5.112)$$

The moiré fringe must be about three times the quadrant length to obtain the strongest error signal. Another way to write this is that

$$\frac{P_s}{\epsilon_p L} \approx 3.0. \quad (5.113)$$

The optimal pitch mismatch is determined by the sensor size and the grating pitch. If the pitch mismatch is made smaller than this, then the moiré fringe becomes too large and the amplitude factor becomes small. Even though a small pitch mismatch means the moiré fringe is more sensitive to grating displacements, it is of no use unless the sensor is large enough to encompass approximately one period of the fringe.

Another way to see the benefit and tradeoffs involved with this technique is to examine the slope of the error signal at the origin

$$\frac{de_{cx}}{d\xi} = \frac{E_0 R L^2}{2} \frac{a(k_\Delta)}{1 + \epsilon_p} \frac{2\pi}{P_s} \quad (5.114)$$

Note that the signal is proportional to the sensor area L^2 and inversely proportional to the sensor pitch P_s . The benefit of using a grating for the alignment mark is obtained by making a fine pitch grating which covers a large area. Once these parameters have been fixed, the optimal

pitch mismatch between the mark and sensor gratings is determined such that approximately one moiré fringe covers the sensor. One must also keep in mind that the pitch spacing cannot be made arbitrarily small without endangering the overlap between the coarse alignment final position and the fine alignment capture region. The fine alignment capture region is given by the ratio P_s/L and must be large enough to accommodate the precision of the coarse alignment.

Error Sources

The spatial phase detection for fine alignment is reminiscent of the large area photodiodes used in the surveyor's mark sensor. Both gather light from across the entire image to generate the error signals. A similar sensitivity to illumination variation can be expected. To verify this a simple linear variation in the irradiance is assumed, as was done in Chapter 3, by multiplying the moiré fringe function $f_L(x)$ by the factor $(1 + bx)$. After recalculating the integral of Equation 5.105 one can solve for the approximate offset

$$\xi \approx \frac{bL(1 + \epsilon_p)P_s}{2\pi a(k_\Delta)} \left\{ \frac{2}{3} + \frac{2A}{k_\Delta L} \left[\left(1 - \frac{2}{k_\Delta^2 L^2} \right) \sin k_\Delta L + \frac{2}{k_\Delta L} \cos k_\Delta L \right] \right\}. \quad (5.115)$$

Plugging in the parameters for the implemented design results in

$$\xi \approx \frac{bL}{5.3} P_s. \quad (5.116)$$

Thus, for example, a nonuniformity of $b = 0.05/(3000 \mu\text{m})$ would result in an offset of about $0.01P_s$, one percent of the sensor pitch and 333 ppm of the sensor size.

Rotation of the mark will reduce the amplitude of the detected signal by causing the moiré fringe to rotate away from vertical. This reduces the differential signal in the x -direction. Just as a small displacement of the grating results in a large displacement of the moire fringe, a small rotation of the grating results in a large rotation of the fringe pattern. It can be shown³ that the angle of the fringe pattern is related to the rotation angle of the grating by [71]

$$\theta_f = \arctan \left(\frac{1}{\frac{1 + \epsilon_p}{\sin \theta} - \cos \theta} \right). \quad (5.117)$$

With no rotation the fringe is vertical. For a small rotation and a small pitch mismatch the equation can be approximated by

$$\theta_f \approx \arctan \left(\frac{\epsilon_p}{\theta} \right). \quad (5.118)$$

Thus, a 45 degree rotation of the fringes occurs when $\theta \approx \pm \epsilon_p$. For $\epsilon_p = 0.023$ the corresponding grating rotation is ± 1.3 degrees. The exact effect of the rotated fringe upon the spatial phase

³I have corrected for Patorski's unconventional choice of clockwise θ .

detection was not studied. However, it is clear that a horizontal fringe cannot result in any x -direction differential signal. As the fringe rotates towards horizontal the phase detector signal must diminish. In order to retain a reasonable signal level for the spatial phase detector, a small pitch mismatch requires tight control on the rotation of the target.

Magnification is a serious concern for the fine alignment signal because it determines the pitch mismatch between the gratings. This in turn controls the amplitude of the alignment signal. Magnification is defined as the ratio of the image height to the object height. Suppose the magnification required to match the image and sensor pitches ($P_s = P_m$) is defined as M_0 . Then the required magnification for optimal pitch mismatch is $M_0(1 + \epsilon_p)$ because P_m is proportional to magnification. The sensitivity between pitch mismatch and magnification is

$$\frac{M}{\epsilon_p} \frac{\partial \epsilon_p}{\partial M} = \frac{1 + \epsilon_p}{\epsilon_p}. \quad (5.119)$$

Thus, a small pitch mismatch results in a large sensitivity to magnification error. This can be related to the optical distances using the Gaussian lens law for a single lens system [96]. Assuming that the distance between the lens and image is fixed, the magnification depends upon the focal length f and the distance between the lens and the object s_o . From the lens law this relationship is

$$M = -\frac{f}{s_o - f}. \quad (5.120)$$

The sensitivity of the magnification to object distance is

$$\frac{s_o}{M} \frac{\partial M}{\partial s_o} = -\frac{s_o}{s_o - f} \quad (5.121)$$

which has a magnitude greater than one for $s_o > f$ and approaching one for $s_o \gg f$. For example, if $\epsilon_p = 0.023$ one must control the object distance to significantly better than two percent. Small pitch mismatch requires tight control of both the magnification and the object distance.

5.3 Absolute Value Circuit

The absolute value circuits compute the modified absolute value of the difference between currents from adjacent photodiodes along a row. These circuits are located in the spaces between the rows of photodiodes. The circuitry next to one photodiode is shown in Figure 5-16. It is comprised of four current mirrors and a pair of transistors acting as switches. The current $I[n]$ from the adjacent photodiode first encounters a switch formed by devices MP and MN that is controlled by the signal V_s . In the fine alignment mode V_s is low, thereby directing the current to the output of the cell and bypassing the current mirrors. All of the outputs from the circuits in one quadrant of the sensor are connected together in order to sum the signals. With V_s high the photocurrent enters the current mirror formed by devices M1A through M1D. Device

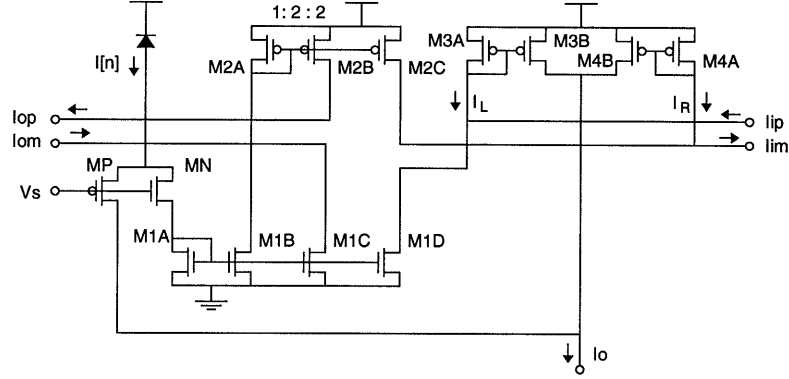


FIGURE 5-16: The circuitry next to each photodiode in the sensor which computes the modified absolute value of the difference between adjacent photocurrents.

M1C provides a sink current $I_{om} = I[n]$ to the next cell to the left. Similarly, device M1B with the mirror M2A and M2B provides a source current $I_{op} = I[n]$. Source and sink copies of the photocurrent to the right $I[n+1]$ enter as I_{ip} and I_{im} . These currents are compared to copies of $I[n]$ from devices M1D and M2C to form the differences

$$I_L = I[n] - I[n+1] \quad (5.122)$$

$$I_R = I[n+1] - I[n]. \quad (5.123)$$

Note that these currents are equal and opposite. The current mirrors formed by devices M3A, M3B, M4A, and M4B rectify these currents so that only the positive portion appears as the output current

$$I_o = I_L u(I_L) + I_R u(I_R) \quad (5.124)$$

where $u(I)$ is the unit step function. The resulting output is the absolute value of the difference of the two photocurrents

$$I_o = |I[n] - I[n+1]| \quad (5.125)$$

assuming the current mirrors are perfect.

In order to generate the modified absolute value function of Equation 5.15, the transistors in the second current mirror are resized to add some gain. By making devices M2B and M2C slightly wider than device M2A, the current gain through this mirror can be changed from unity to $(1 + \delta)$. This results in currents

$$I_L = I[n] - (1 + \delta)I[n+1] \quad (5.126)$$

$$I_R = I[n+1] - (1 + \delta)I[n] \quad (5.127)$$

and an input/output characteristic as given earlier by Equations 5.15 and 5.16.

In practice, the current mirrors will not provide the desired gains for two reasons. First, the drain-to-source voltages of the input and output devices in each current mirror differ significantly. The finite output resistance of the transistors results in a gain error which can be

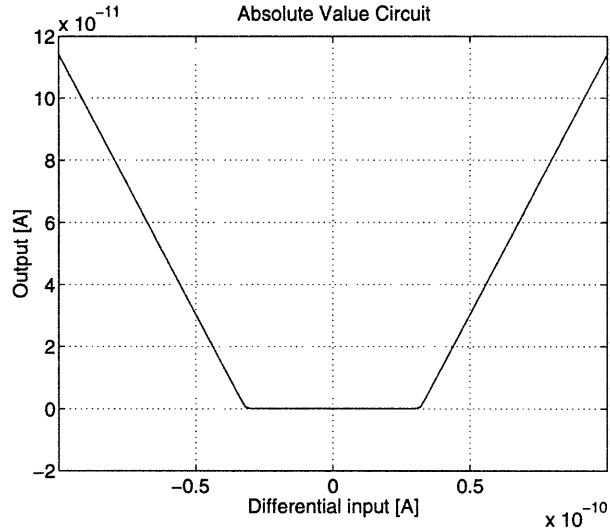


FIGURE 5-17: Simulation of the absolute value output as the input currents are differentially varied about a common-mode level of 50 pA.

anticipated to some extent in the choice of the transistor size. Second, and more importantly, the mismatch between devices causes a substantial mismatch in the mirror currents. Because the photocurrents are quite small, on the order of 1 to 100 pA, the devices operate in sub-threshold where the effect of mismatch is exacerbated. For example, current mismatch of 9% for NMOS devices and 40% for PMOS devices has been reported for $3 \mu\text{m}/8 \mu\text{m}$ devices in the Orbit 2.0 μm p-well process [109]. These numbers vary between processes and depend on device size, but the important point is that the mismatch can be quite large at low currents. This causes asymmetric gain and offset errors in the input/output characteristic of the absolute value circuit.

In the fabricated chip the second current mirror was sized with M2A $8 \mu\text{m}/9 \mu\text{m}$ and M2B and M2C $14 \mu\text{m}/9 \mu\text{m}$. This results in an ideal δ of 0.75. Simulations accounting for the variation between actual and drawn size as well as the effect of finite output resistance showed that the effective δ was about 1.0. The corresponding cutoff point for the modified absolute value function is 0.33. Such a large cutoff threshold was an intentionally conservative choice to ensure zero output from the absolute value when the two photocurrents are within twenty percent of each other. This requires a minimum current gain of $(1 + \delta) = 1.2$. Assuming that output current can vary by up to forty percent due to device mismatch implies that $\delta = 1.0$ to ensure that the current gain is always greater than 1.2. In addition to choosing the proper W/L ratio, the mirror devices were sized with as much area as could be squeezed into the allowed space in order to reduce the device mismatch [109]. All of the M1 devices were sized $8 \mu\text{m}/7 \mu\text{m}$ and the M3 and M4 devices were sized $8 \mu\text{m}/10 \mu\text{m}$. Figure 5-17 shows the simulated input/output characteristic for the absolute value circuit when the two input currents were varied differentially about a fixed common-mode level of 50 pA.

A special situation arises at the ends of the rows of photodiodes which form the corners of the square pattern. These photodiodes were split into two smaller photodiodes (see Figure 5-3) so that the inner one could be shared between rows in two different quadrants of the sensor. An extra current mirror was used to duplicate these photocurrents for use in the absolute value circuits of each adjoining row. This is important when the target image displacement is less than P_s because only these photodiodes will detect an edge that can contribute to the coarse alignment error signal.

5.3.1 Absolute Value Circuit Dynamics

One might not expect the absolute value circuit to operate very quickly because it is comprised of current mirrors operating in subthreshold. The small size of the the photodiode pixel will result in a very small photocurrent. For example, a focal plane irradiance on the order of $10 \mu\text{W}/\text{cm}^2$ with a photodiode size of $100 \times 46 \mu\text{m}^2$ will collect only 60 pA of photocurrent. Simulations reveal that the time response of this circuit is dominated by the input capacitance from the photodiode. HSPICE calculates the photodiode depletion capacitance to be 580 fF with a total gate and drain-to-substrate capacitance at this node of 160 fF. Thus the first current mirror is loaded with about 740 fF of input capacitance, far more than the subsequent mirrors.

For small signals the resulting bandwidth is determined by this capacitance reacting with the impedance presented by the diode connected device M1A yielding

$$f_h = \frac{I_p}{2\pi n \frac{kT}{q} C_i}. \quad (5.128)$$

Note that this is essentially the same formula as seen previously in Equation 3.61. Here the subthreshold slope is given by $n = 1.7$ with $C_i = 740$ fF and I_p being the photocurrent. Setting aside the device capacitance, the bandwidth is again determined largely by the irradiance divided by the capacitance per unit area of the photodiode. Thus, small-signal bandwidths comparable to the surveyor's mark sensor (without the active cascode) are achievable. Figure 5-18(a) shows the transfer function for the circuit for several different photocurrents (with I_{im} and I_{ip} set to zero). The above equation accurately predicts these bandwidths which range from 5 Hz to 5 kHz. For $I_p = 60$ pA the equation predicts a bandwidth of 290 Hz which compares well with the surveyor's mark sensor under similar illumination.

In actual operation this circuit will undergo large scale transients as the grating image moves across the sensor. A large signal transient simulation is shown in Figure 5-18(b). In this case one photocurrent was stepped from 10 pA to 100 pA at $t = 0$ and then back to 10 pA at $t = 6$ ms. This models a darkened pixel (contrast $\epsilon = 0.1$) that becomes fully illuminated and is then returned to darkness. The other photocurrent was set to zero. Note the relatively rapid rise time compared to the slower fall time. This occurs because the the photocurrent charging up the input capacitance is a faster process than the discharge through device M1A.

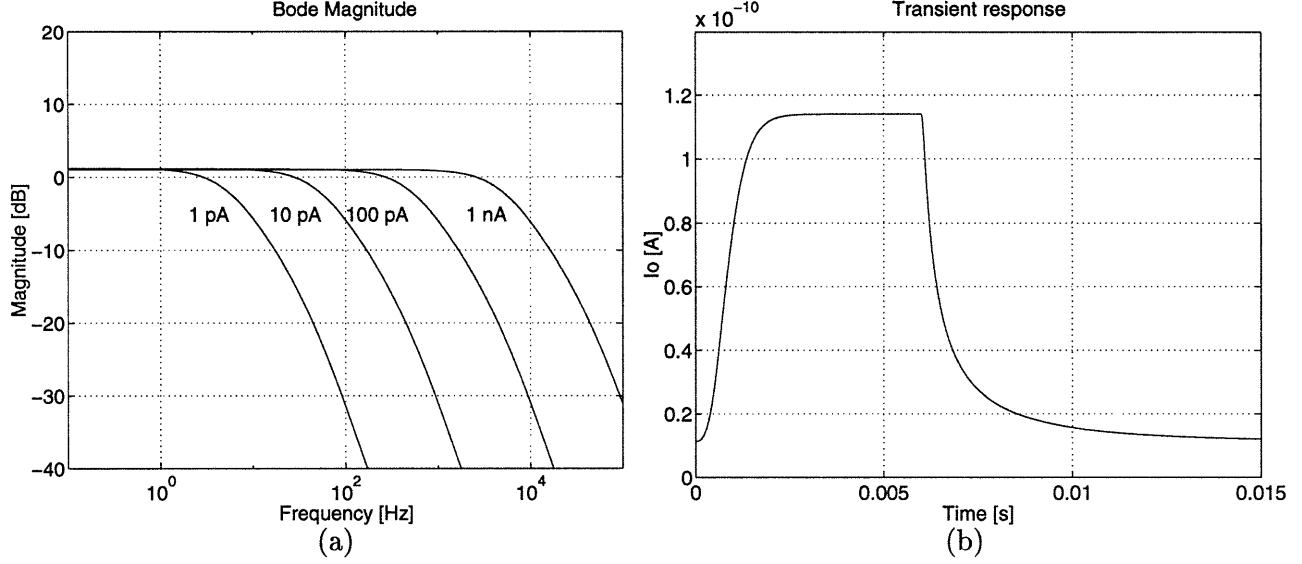


FIGURE 5-18: Dynamics of the absolute value circuit (a) small-signal frequency response with various photocurrents and (b) large-signal transient response for a 10 pA to 100 pA step input.

As before, it will be assumed that the dynamics are dominated by the first current mirror. A differential equation relating the photocurrent to the current I_m flowing through device M1A can be written as

$$I_p(t) = n \frac{kT}{q} C_i \frac{1}{I_m(t)} \frac{dI_m(t)}{dt} + I_m(t) \quad (5.129)$$

where the subthreshold current is modeled as

$$I_m(t) = I_s \exp\left(\frac{V_{gs}}{n \frac{kT}{q}}\right) \quad (5.130)$$

with $I_s = 2$ fA. This differential equation can be solved for the time required for I_m to reach a fraction kI_p ($k < 1$) of the peak value

$$t_r = \frac{n \frac{kT}{q} C_i}{I_p} \ln \frac{\frac{1}{\epsilon} - 1}{\frac{1}{k} - 1} \quad (5.131)$$

where I_p is the peak input current (100 pA in this case) and $\epsilon = 0.1$ denotes the starting current of $\epsilon I_p = 10$ pA. From this the 10% to 90% rise time can be computed by noting that it is the time required for I_m to go from $0.19I_p$ to $0.91I_p$

$$T_{10-90} = \frac{n \frac{kT}{q} C_i}{I_p} (3.76). \quad (5.132)$$

The equation predicts a rise time of 1.23 ms which compares well with the 1.05 ms obtained in the simulation. A similar equation can be found for the fall time from I_p to kI_p

$$t_f = \frac{n \frac{kT}{q} C_i}{\epsilon I_p} \ln \frac{1 - \epsilon}{1 - \frac{\epsilon}{k}}. \quad (5.133)$$

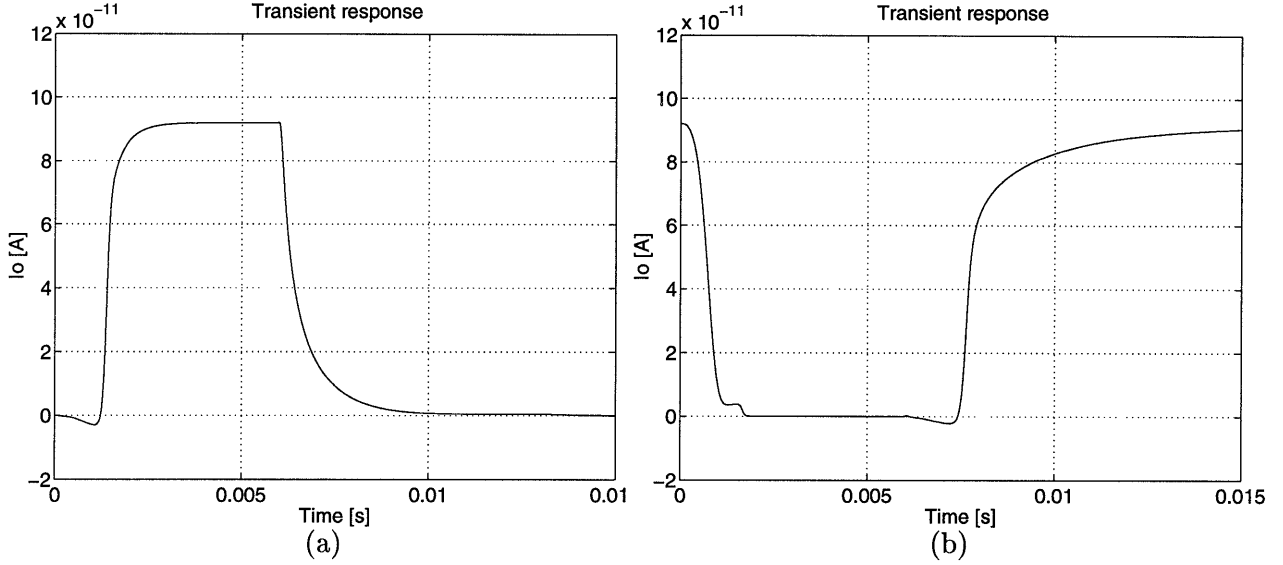


FIGURE 5-19: Large signal transients of the absolute value circuit with one input stepped from 10 pA to 100 pA and back with (a) the other input held at 10 pA and (b) held at 100 pA.

The corresponding 90% to 10% fall time is

$$T_{90-10} = \frac{n \frac{kT}{q} C_i}{I_p} (6.31). \quad (5.134)$$

This equation predicts a fall time of 2.06 ms compared to a simulated value of 2.18 ms.

A more realistic pair of transient simulations is shown in Figure 5-19. In this case the photocurrent $I[n]$ was stepped from 10 pA to 100 pA as before, but the photocurrent $I[n + 1]$ from the next pixel to the right was held at a constant value. Part (a) of the figure shows the result for $I[n + 1] = 10$ pA. In this case the transient simulation models both pixels as covered by a dark region with contrast $\epsilon = 0.1$. At $t = 0$ the left pixel becomes fully illuminated ($I_p = 100$ pA) while the right pixel remains darkened. At $t = 6$ ms both pixels return to the darkened state. The important feature of this simulation is the large delay in the output between the onset of the photocurrent step and when the output starts to rise. This occurs because both photocurrents are initially equal. Thus, the absolute value circuit output is zero and the gates of the current mirrors M3 and M4 are close to V_{dd} . When the photocurrent steps to 100 pA, the gates of mirror M3 must drop down by about 0.5 V to turn on the devices. While this node is charging the output remains close to zero with a slight negative excursion due to the gate/drain overlap capacitance of device M3B. Once this mirror turns on there is a quick rise in the output followed by settling that is largely determined by the input capacitance. The length of the delay is determined by the current through M1D discharging the approximately 100 fF capacitance on the gates of mirror M3. Solving for this time explicitly would require integrating the solution to Equation 5.129. The total delay to reach 90% of the final value can be estimated from Equation 5.131 assuming the input capacitance dominates the final settling

as I_m rises from $0.1I_p$ to $0.91I_p$

$$T_{0-90} = \frac{n \frac{kT}{q} C}{I_p} (4.51). \quad (5.135)$$

This predicts a total rise time, including the delay, of 1.48 ms compared to a simulated value of 1.76 ms. The fall time can be calculated again from Equation 5.133 keeping in mind that the output is zero when $I_m = 0.2I_p$. The the 90% to 10% fall time at the output is given approximately by the 92% to 28% fall time for I_m

$$T_{90-10} = \frac{n \frac{kT}{q} C}{I_p} (3.27). \quad (5.136)$$

which yields a fall time of 1.07 ms compared to the simulated value of 1.4 ms.

Part (b) of Figure 5-19 shows the same transient in $I[n]$ as part (a) except that now $I[n+1]$ is held at 100 pA to model a fully illuminated neighboring pixel. The output transient shows an initial fall to zero as the two photocurrents become equal followed by a return to the full output as $I[n]$ drops back down to 10 pA. There is a substantial delay from $t = 6$ ms until the waveform begins rising which is caused by the time taken to turn on the gates of mirror M4. The indentation at the base of the falling waveform is caused by the gate/drain overlap capacitance of transistor M4A as it turns off. This device also causes the slight negative ramp at the base of the rising waveform as it turns on again. The 90% to 10% fall time corresponds to the rise time of I_m from $0.14I_p$ to $0.46I_p$ given by Equation 5.131

$$T_{90-10} = \frac{n \frac{kT}{q} C}{I_p} (1.65) \quad (5.137)$$

which yields a fall time of 540 μs compared to the simulated value of 590 μs . The rise time, including delay, to 90% of the final value is also accurately modeled by the fall time I_m back to $0.14I_p$

$$T_{0-90} = \frac{n \frac{kT}{q} C}{I_p} (11.5). \quad (5.138)$$

This predicts a rise time of 3.8 ms which equals the simulated value.

All of these transient situations yield rise and fall times of the form

$$T = \frac{n \frac{kT}{q} C}{I_p} K \quad (5.139)$$

where K is a factor roughly between 1 and 12. The agreement between these equations and the simulations holds up for a wide range of input currents as long as the devices are operating in subthreshold. The speed of the absolute value circuit can be approximated by taking the largest value of K as the maximum settling time. For the typical case of $I_p = 60$ pA the resulting value is 6.5 ms. If one considers these settling times as approximately due to the response of a first order system, then a corresponding bandwidth can be obtained using the model ([110] Sec. 3.5)

$$f_h = \frac{2.2}{2\pi T}. \quad (5.140)$$

Obviously these large scale signal transients are not due to a first order system response, but this allows one to compute a “bandwidth” for comparison purposes. The settling time of 6.5 ms corresponds to a bandwidth of 54 Hz. Thus, these large signal transients will limit the speed of the coarse alignment to something on the order of video rate. The speed can be improved to an extent by increasing the illumination. However, the circuit is really limited by the slow process of turning on and off the current mirrors. To substantially increase the speed without inordinate amounts of illumination would require substantially redesigning or replacing this circuit.

5.4 Current Amplifiers

Both the coarse and fine alignment error signals consist of four currents, one from each quadrant of the sensor. These currents form two differential pairs related to the x and y misalignment. As was the case for the surveyor’s mark sensor, these currents are also proportional to the overall illumination level. Differential current amplifiers are used to normalize and amplify them. The amplifiers are identical to that shown in Figure 3-8 except that the active cascode devices M7 through M12 were not included. Adding the active cascode would have required a refabrication of the sensor. This was deemed not necessary since the benefit of this circuit had already been verified in the surveyor’s mark sensor. Only two amplifiers are used on this sensor, one each for the x and y error signals. The same amplifiers, and summing wires, are used for both the coarse and fine alignment error signals. A detailed discussion of these amplifiers was presented earlier in Section 3.3. Essentially all of the power dissipated in this sensor is consumed by the two current amplifiers. The output currents alone total $200\ \mu\text{A}$ and use 1.0 mW with a 5 V supply. The remaining amplifier stages, bias currents, and photocurrents use an additional 0.22 mW.

5.5 Summary

This chapter has presented a sensor design which uses a square grating array of photodiodes in order to detect a square grating alignment target. A coarse alignment signal is generated by detecting and comparing the number of edge crossings in the four quadrants of the sensor. Absolute value circuits integrated alongside the photodiodes detect the edges. The fine alignment signal is obtained by using the sensor as a spatial phase detector to locate a geometric moiré fringe. This fringe is formed when the target grating is imaged upon the sensor grating and is a sensitive indicator of their relative positions. The resulting error currents for both coarse and fine alignment are normalized and amplified using circuits similar to those used in the surveyor’s mark sensor. The coarse alignment signal is expected to be tolerant of nonuniform illumination, rotations of the target, and magnification error. On the other hand, the fine alignment signal is very sensitive to all three of these error sources.

Test Results: Square Grating Mark Alignment Sensor

6.1 Introduction

Testing of the square grating sensor was very similar to that done for the surveyor's mark sensor. Two versions of the chip were fabricated in the Orbit 2 μm n -well BiCMOS process provided by MOSIS [107]. Both had imaging areas $6 \times 6 \text{ mm}^2$ in size. The first sensor used an absolute value circuit without any cutoff threshold and exhibited a periodic artifact in the coarse alignment signal. The second version added a cutoff point to the absolute value circuit to eliminate this problem.

The test results from the photodiodes and the current amplifiers were essentially the same as those already presented in Chapter 4 and will not be repeated here. This chapter focuses on those aspects of the square grating sensor that differ from the surveyor's mark sensor. The absolute value circuit is discussed first with measurements showing the built-in cutoff threshold. This is followed by various measurements characterizing the alignment error signals. Alignment trajectories are shown along with discussions of the repeatability and the effect of a rotated target. The chapter concludes with measurements of the sensor bandwidth, transient response, and noise.

6.2 Absolute Value Circuit

The modified absolute value circuit was discussed earlier in Section 5.3 and shown in Figure 5-16. It was tested by fixing the common mode of the input current pair and then varying the currents differentially. The resulting input/output plots are shown in Figure 6-1. Part (a) of the figure shows the characteristic with $I_{icm} = 500 \text{ pA}$. There is significant asymmetry which arises from the mismatch in the current mirrors. Because the devices are operating at sub-

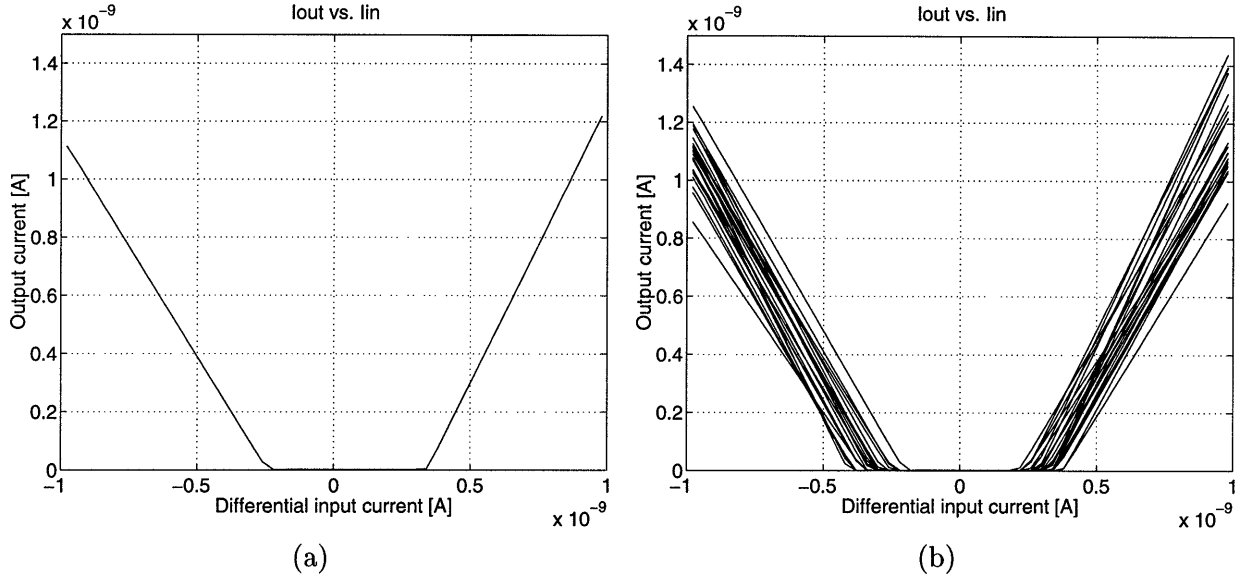


FIGURE 6-1: Input/output characteristic from (a) a typical absolute circuit and (b) a superposition of characteristics from twenty-four different circuits showing the effects of device mismatch. The cutoff threshold is generally between 0.2 and 0.4 of $2I_{icm}$.

threshold current levels, this mismatch is large and causes asymmetry in both the gain as well as the cutoff threshold in the plot. Part (b) of this figure shows a superposition of the characteristics from 24 different circuits on six different chips. The cutoff point, designed to be about 0.33, ranges anywhere from 0.2 to 0.4. Even greater variation was anticipated which led to the choice of such a large cutoff threshold. The data indicates that a lower cutoff point, about 0.15 to 0.2, would still maintain zero output for zero input in the face of device mismatch. Using a smaller cutoff point has two benefits. First, it would increase the average current signal \bar{I}_e of a detected edge which would raise the overall signal level. For example, lowering the cutoff threshold from $1/3$ to $1/6$ would increase the average edge current by about fifty percent. Second, it would also reduce the amplitude of the ripple on this signal. Because the choice of the cutoff point is determined by device mismatch, it will depend on process variations. Thus, it should be kept in mind that the data shown here are for a single run of a particular process.

6.3 Alignment Error Signals

Characterization of the alignment error signal was done using the same mock alignment system discussed in Section 4.4.1. The only difference was the alignment target used. In this case the mark was a square grating, as illustrated in Figure 5-1(a), that was 10 cm across consisting of fifteen nested squares with a period of 3.33 mm. This pattern was laser-printed onto white bond paper with a 600 dpi printer and then placed on the movable table for the experiments.

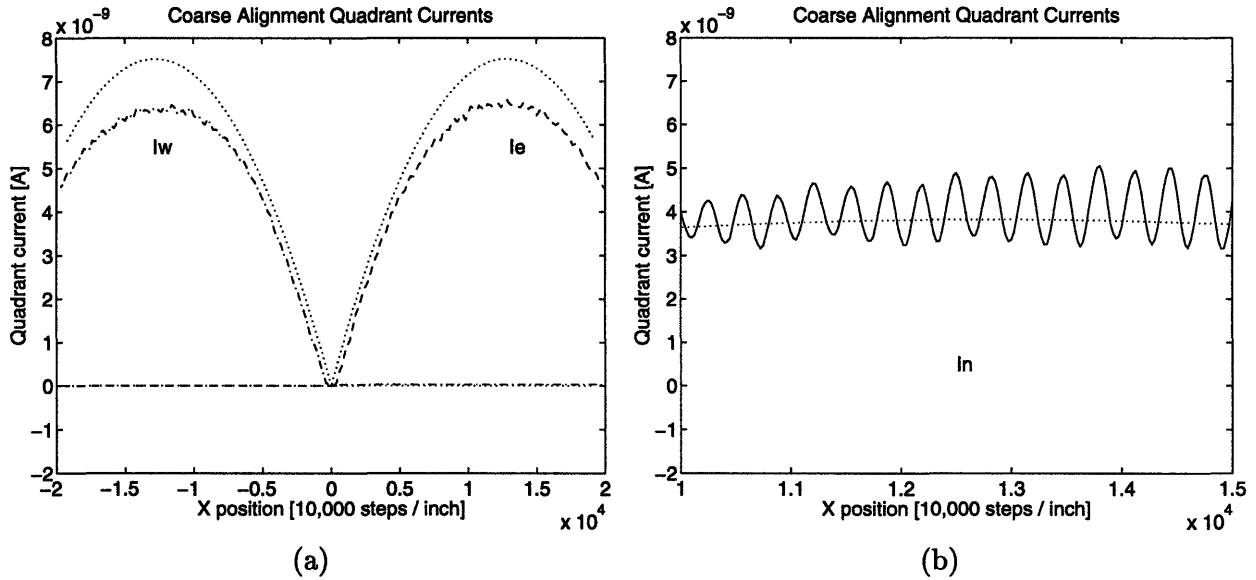


FIGURE 6-2: Coarse alignment quadrant currents as the target is moved along the x -axis. (a) The east and west quadrant currents and (b) the ripples in the north quadrant current as the edges move along the rows of photodiodes. Dotted lines show the theoretical results from the approximate model.

6.3.1 Coarse Alignment Signal

Moving the target along the x -axis resulted in the coarse alignment signal currents shown in Figure 6-2(a). These are quadrant currents measured directly from the sensor. Reasonable agreement is seen between the data and the model expressed in Equation 5.53. In this case the value of I_0 was obtained by measuring the total photocurrent in a quadrant with a white sheet as the target and then dividing by the number of photodiodes. The resulting value was 64 pA, corresponding to a sensor irradiance of about $10 \mu\text{W}/\text{cm}^2$. The model used a cutoff point $t = 0.33$ and assumed a contrast parameter $\epsilon = 0.173$, the same value measured from the surveyor's mark target. The agreement is good considering the uncertainty in these parameters and the approximations made in deriving the model. The error currents I_E and I_W also show a small ripple with period $P_s/2$ that arises from the cutoff in the absolute value circuit. Ripples due to this effect are more evident in the currents I_N and I_S as shown in part (b) of the figure. In the north and south quadrants the mark edges are moving along the array of photodiodes as shown in Figure 5-4. The resulting ripples are large in amplitude and have a period of $P_d = P_s/4$ due to the variation in the edge current signal shown in Figure 5-6.

Note that the current I_W is zero for displacements to the left of the origin and I_E is zero for displacements to the right. This shows that essentially no "distortion" currents are being summed in the quadrants without any edge crossings (see Equation 5.12). The modified absolute value circuit successfully eliminates these currents. Unfortunately this also points out a flaw in this implementation. Because both error currents go to zero for small displacements, the

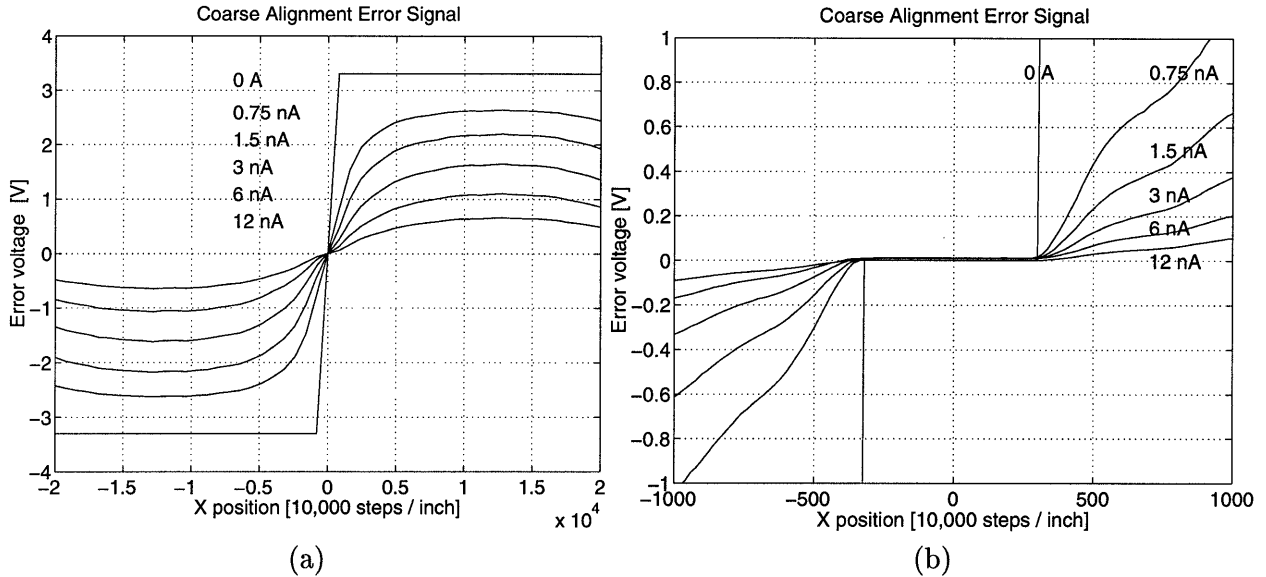


FIGURE 6-3: Error signal voltage for the x -direction as the target is moved along the x -axis (a) over the full field of view and (b) near the origin. Multiple curves show the effect of adding bias currents of various values to the signal current.

current amplifier becomes starved for current. This leads to very slow settling behavior for target positions near the origin.

The fact that one of the quadrant currents is always zero also leads to a saturated output error signal after the normalization performed by the current amplifier. This is seen in the top curve of Figure 6-3. The signal is saturated at ± 3.3 V over most of the range of motion with a steep transition region as the mark crosses the origin. This is not a good error signal for a control system. A signal that was linear over the whole range of motion would be preferable. The signal can be improved by adding a DC bias current, from off-chip, to each quadrant current as shown by the remaining curves in the figure. As the bias current is increased, the curves become lower and begin to resemble the shape of the unamplified current signal of Figure 6-2(a). The bias currents prevent the amplifier from being starved for current near the aligned position, decrease the signal settling time, and broaden the high gain transition. The price paid is a reduction in the signal amplitude and the addition of circuitry that would be needed to provide the currents. The size of the bias currents is an important choice. In the figure, currents of 6 nA result in a reasonable error signal with a slope through the origin that is $200 \mu\text{V}/\text{step}$, comparable to that from the surveyor's mark sensor. This bias current level is nearly equal to the 6.5 nA peak current signal seen in Figure 6-2(a). Thus, it appears that making the bias currents equal to this peak current level would be a good choice. The peak signal is proportional to the illumination level. Therefore, the bias currents would also need to be proportional to the illumination level in order to obtain a normalized error signal that is independent of illumination changes. Such a current could be generated, for example, by summing all of the photocurrents from the sensor

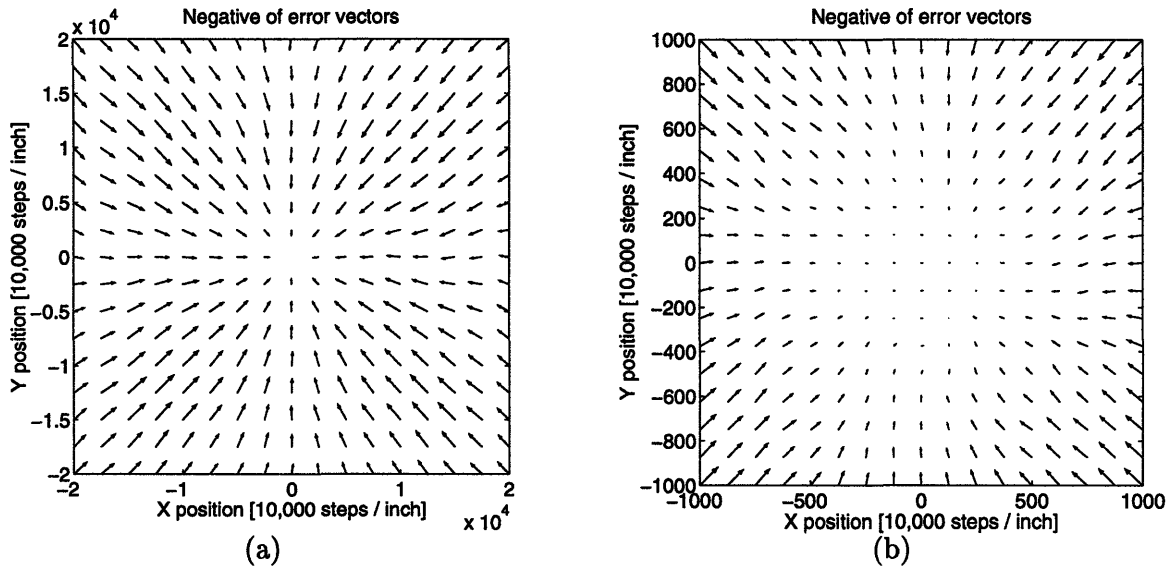


FIGURE 6-4: Negative error vector field obtained from the unamplified quadrant currents (a) over the full field of view and (b) showing the dead zone near the origin.

and dividing them by an appropriate constant.

Part (b) of the figure shows a close-up view of the origin revealing a small dead zone in the signal. This arises from the cutoff threshold in the modified absolute value circuit. For small displacements around the origin, only a fraction of an edge will be detected at the corners of the square photodiode arrays. If this signal does not exceed the cutoff point, then the output of the corresponding circuit will be zero. For a small region about the origin the outputs of all of the absolute value circuits are zero which causes the dead zone.

A plot of the negative error vectors made from the unamplified error currents is shown in Figure 6-4. These error vectors compare well with the expected result plotted in Figure 5-10(a). Part (b) of this figure shows the dead zone near the origin.

The effect of rotations of the mark was investigated experimentally. Figure 6-5(a) shows the unamplified error currents at a few different rotations. The curve with no rotation compares well in shape to the theoretical result shown in Figure 5-10(a). As the target is rotated, edges appear in all the quadrants, although the detected edge signals weaken as they are rotated away from perpendicular to the photodiode arrays. A 45 degree rotation causes all four quadrants to detect edges oriented 45 degrees to the photodiode arrays. At this point the individual edge signals are very weak, however this is compensated to an extent by their large number. As the mark is translated, the number of these edges in each quadrant varies and still provides a usable alignment signal. A strong ripple appears on the signal in this case as well as a substantial offset. The source of this offset is not known for sure, although it is likely due to nonuniformity in the illumination. A smaller cutoff threshold in the modified absolute value should increase the signal when the edges are rotated although no investigation of this was done. Part (b) of

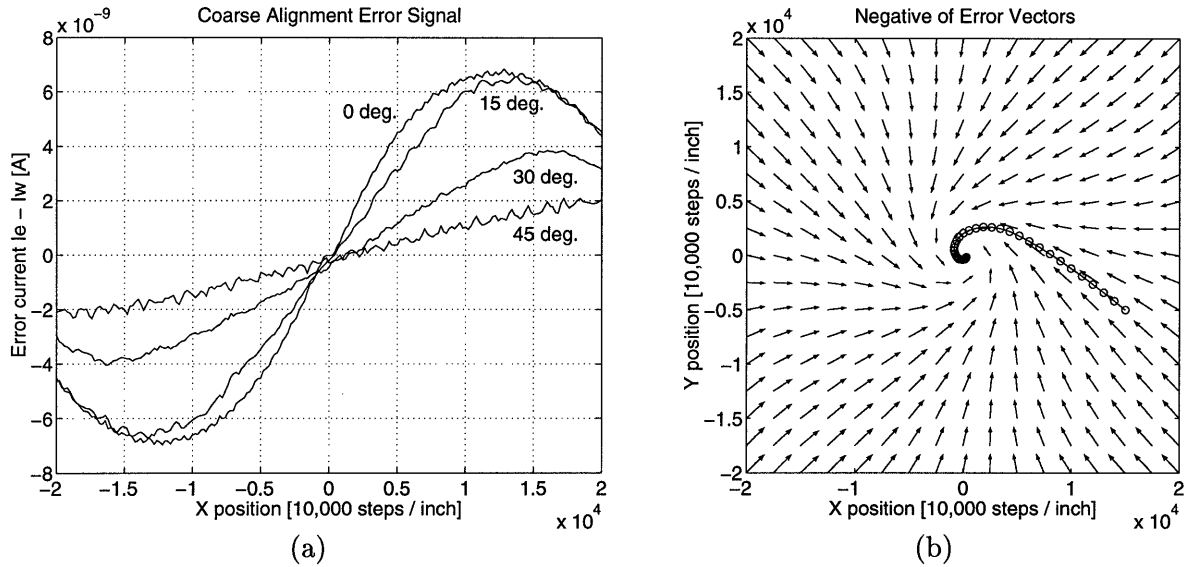


FIGURE 6-5: The effect of rotations of the target. (a) Rotations attenuate the unamplified error signal currents. (b) Output error vectors for a 15 degree rotation result in a coarse alignment trajectory that spirals into the origin.

the figure shows the output error vector map for a 15 degree rotation. An alignment trajectory is also shown spiraling into the origin as it follows the error signal.

6.3.2 Fine Alignment Signal

Measurements of the fine alignment signal clearly showed the spatially phase detected signal from the moiré fringe. The x -direction error signal for displacements along the x -axis is plotted in Figure 6-6(a). The signal appears approximately sinusoidal in shape with a capture region of ± 660 steps. The capture region is 3.3% of the sensor field of view and corresponds to one period of the mark grating (ideally 1333 steps). The peak signal is ± 1.7 V. The unamplified differential current signal had an amplitude of 30.8 nA which was 55% of the maximum possible signal ($I_{max} = 55.6$ nA was measured with a white sheet as the target). This corresponds surprisingly well with the maximum amplitude factor of 0.57 obtained using the simple sinusoid model of the image grating (see Figure 5-15 and Equation 5.108). This analysis did not model the square wave nature of the target grating, which would tend to increase the amplitude factor, nor the imperfect image contrast, which would tend to decrease it. For example, the fundamental component of a unit height square wave has an amplitude of $4/\pi = 1.27$ whereas the attenuation due to the contrast is $1 - 0.173 = 0.827$. Multiplying these two factors together yields a factor of 1.05, demonstrating the coincidental cancelling of the effects of these omissions in the model. The signal slope through the origin is 6.5 mV/step which easily allows a voltmeter with 1 mV resolution to detect a single step displacement of the table. Part (b) of the figure shows the measured fine alignment error vectors for displacements within the capture region.

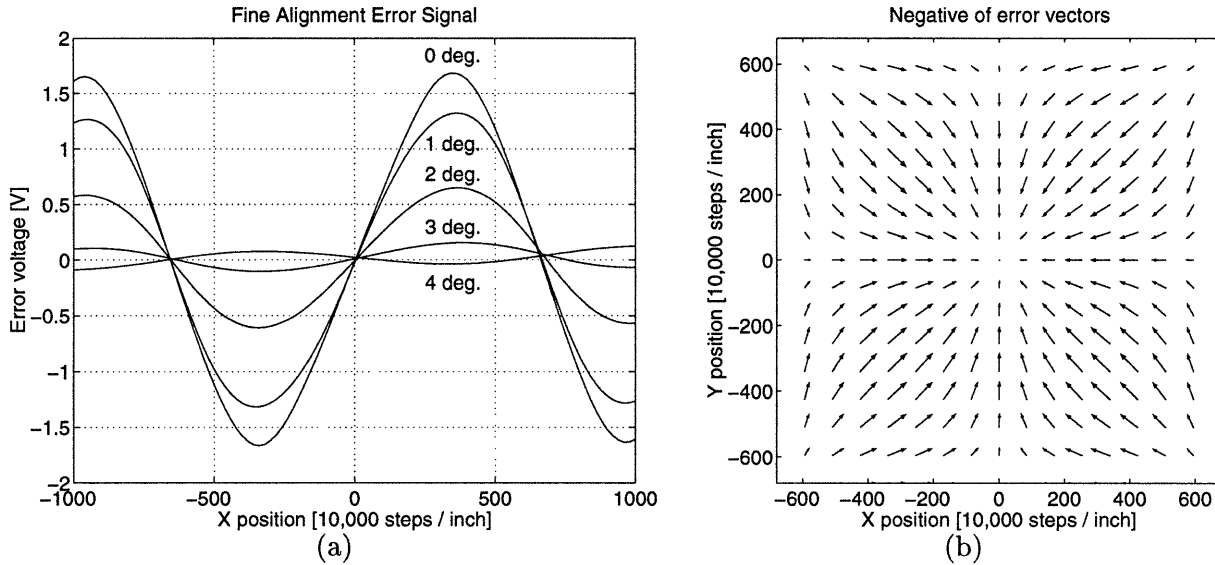


FIGURE 6-6: Fine alignment error signals: (a) x -direction error signal along the x -axis at various rotations and (b) an error vector plot over the capture region (no rotation).

The effect of rotations is also illustrated in Figure 6-6(a). As the target is rotated relative to the sensor, the error signal amplitude is reduced as expected. Equation 5.118 predicted the signal would degrade for rotations of about 1.3 degrees and beyond. This corresponds well with the significant attenuation in the signal seen between rotations of 1 and 2 degrees.

6.3.3 Alignment Trajectories

Alignment trajectories were obtained using the mock alignment system as described previously (Section 4.4.4). The coarse alignment trajectories from several different initial positions are shown in Figure 6-7(a) along with the measured error vector field. The iteration gain was set to a low value of 200 steps/V in order to prevent overshoot oscillations at the high slope transition in the error signal near the origin. Coarse alignment was terminated when the error signal dropped to within ± 1 V which assured a position within the capture region of the fine alignment signal. A large threshold was required to detect the sharp transition at the origin. The trajectories for this sensor take more direct paths to the origin than was the case for the surveyor's mark (see Figure 4-11). All of the coarse alignment trajectories shown took 25 iterations to complete.

Trajectories for the fine alignment are shown in Figure 6-7(b). These trajectories also used a relatively low gain of 30 steps/V in order to show a detailed path. They took between 17 and 20 iterations to complete. Alignment was terminated when the error voltage dropped below ± 35 mV.

A series of alignments was also done with the 6 nA bias currents added to the coarse alignment signal. This allowed a higher coarse alignment gain of 600 steps/V to be used and

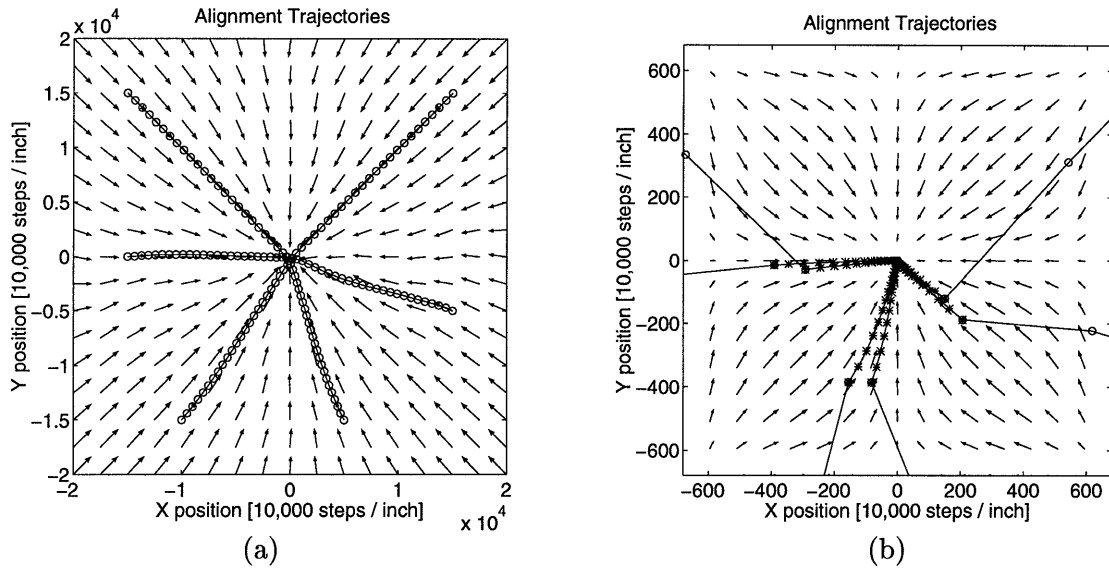


FIGURE 6-7: Trajectories and error vector fields for (a) the coarse alignment and (b) the fine alignment.

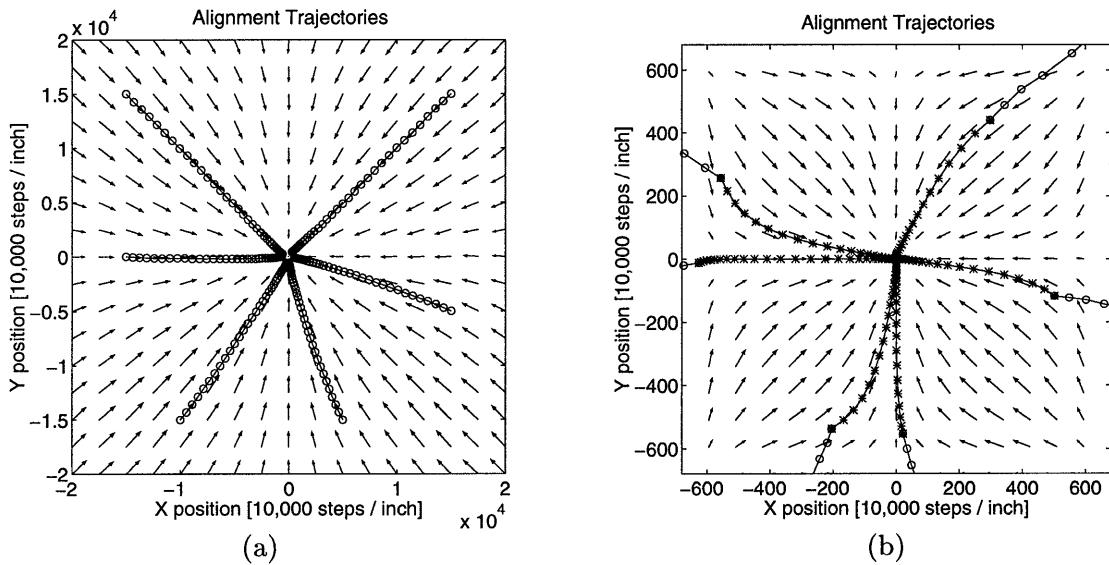


FIGURE 6-8: Trajectories and error vector fields for (a) the coarse alignment with added bias currents and (b) the fine alignment.

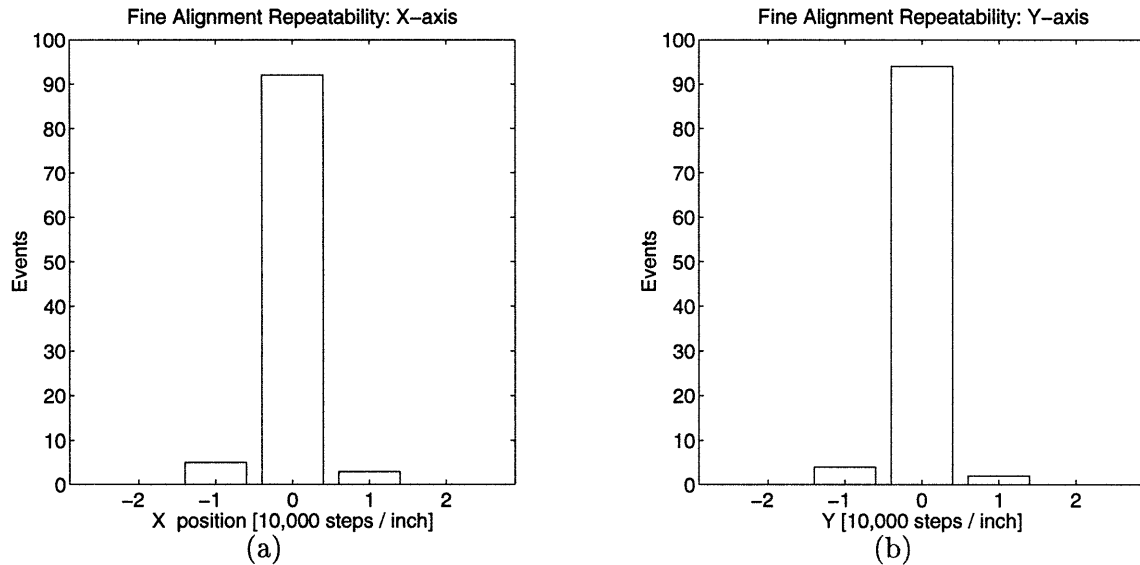


FIGURE 6-9: Histograms of the final alignment (a) x position ($3\sigma = 0.85$) and (b) y position ($3\sigma = 0.74$) for a sequence of 100 alignments.

to see the effect of this modification to the sensor. Because of the reduced signal slope at the origin, the threshold was reduced to ± 70 mV for termination of the coarse alignment. The fine alignment operation was unaltered for these experiments. The resulting trajectories are shown in Figure 6-8. The paths are very similar to those seen without the bias currents, even though the convergence properties and error signals are changed. The coarse alignments took 35 iterations to complete with this modified error signal and gain setting. Using a high gain of 5000 steps/V allowed the alignments to complete in three or four iterations. The fine alignment trajectories took between 22 and 31 iterations to complete using the low gain. This gain could be set as high as 150 steps/V without overshooting the aligned position. In this case the fine alignment was completed in two to five iterations.

6.3.4 Alignment Repeatability

Alignment repeatability in the mock alignment system was investigated as described earlier in Section 6-9. A series of 100 fine alignments were done from different starting points in the capture region. An iteration gain of 145 steps/V was used and the alignment was stopped when the error signal was within ± 4 mV. Histograms of the final positions in motor steps are shown in Figure 6-9. The repeatability for both axes had 3σ values of better than one step which corresponds to 25 ppm of the sensor field of view. Note that the statistics become somewhat misleading when the 3σ point becomes comparable to the bin size of the histogram. The repeatability in this case was clearly limited by the positioning system. A repeatability of 25 ppm of the sensor field of view corresponds to $0.15 \mu\text{m}$ at the image plane and to 2.5 nm for a target $200 \mu\text{m}$ across. In practice the repeatability with such a small target would likely

be limited by factors other than the sensor. The repeatability demonstrated here was achieved using the sensor with very low bandwidth measurements (16 ms integration time). This removes most of the sensor noise that becomes an issue at higher speeds. It is interesting to recall that the repeatability with the surveyor's mark was 53 ppm. The additional uncertainty in that case was likely due to the 1 mV quantization limit of the HP4145 voltmeter. One millivolt of error signal corresponds to one step (25 ppm) in position for the surveyor's mark sensor, whereas for the square grating sensor it corresponds to only 0.15 step (4 ppm).

6.4 Bandwidth

6.4.1 Fine Alignment

Bandwidth measurements for the fine alignment portion of the sensor were made in a similar fashion to those for the surveyor's mark sensor. A pair of LED's were placed in the target plane with one each in the east and west quadrants. Their positions were adjusted until the differential output was zero. A small-signal sinusoidal variation was then added to one of the LED's and its frequency was adjusted to measure the bandwidth. Then the DC quadrant photocurrents were measured with 2.75 V of reverse bias on the photodiodes. The equivalent irradiance of the sensor could not be measured directly with the Newport power meter because the grating pattern only collects a portion of the incident light. An equivalent uniform irradiance can be calculated by assuming that the grating collects half of the light incident upon it. Thus,

$$E_0 = \frac{2I_{icm}}{A_q R} \quad (6.1)$$

where $A_q = 0.09 \text{ cm}^2$ is the quadrant area and $R = 0.13 \text{ A/W}$ is the responsivity.

Measurements were made at different illumination levels by placing neutral density filters in front of the lens. Figure 6-10 shows the resulting bandwidths plotted versus the equivalent irradiance calculated from the measured photocurrents. As was the case for the surveyor's mark sensor, the bandwidth is determined by the pole formed from the input capacitance and the amplifier input impedance

$$f_h = \frac{A_q R E_0}{4\pi n \frac{kT}{q} C_d}. \quad (6.2)$$

Here $C_d = 700 \text{ pF}$ was measured from the photodiode array and includes both the photodiode and wiring capacitance. This equation is plotted as the solid line in the figure and agrees well with the measured data. The plot is very similar to that shown for the surveyor's mark sensor in Figure 4-13. This occurs because the bandwidth limit in both cases is determined by the irradiance divided by the photodiode depletion capacitance per unit area. Adding an active cascode to this sensor would improve the bandwidth substantially as was shown for the surveyor's mark sensor.

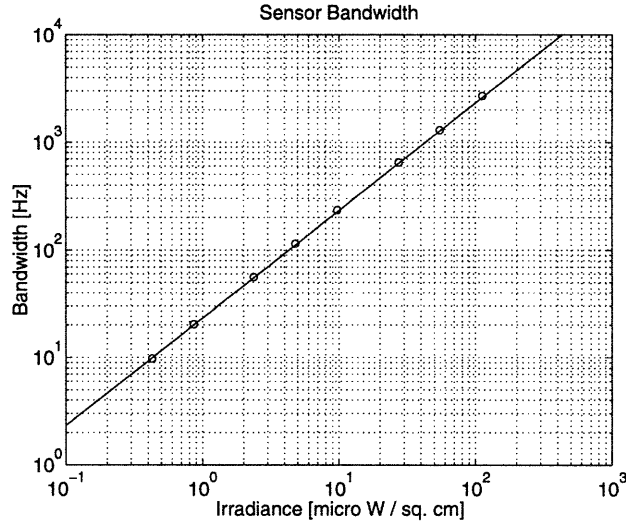


FIGURE 6-10: Bandwidth of the fine alignment portion of the sensor plotted versus equivalent irradiance. The solid line shows the limit imposed by the input capacitance.

6.4.2 Coarse Alignment

The bandwidth of the coarse alignment portion of the sensor is difficult to characterize adequately using the LED's as optical inputs. The LED's are about 3 mm wide and thus generate bright spots at the image plane about $180 \mu\text{m}$ in diameter. This spot size then illuminates a small number of photodiodes, perhaps between four and eight. The resulting signal current, that can be measured, is the summed absolute value of differences from all of the illuminated photodiodes in a quadrant. It is thus impossible to determine the individual diode photocurrents with this experimental set-up. This in turn prevents quantitatively verifying by experiment the bandwidth and time response equations developed in Section 5.3.1. To truly verify the time response would require the ability to target several spots of light on individual photodiode pixels and then vary the brightness of some them dynamically. Alternately, one could actually image the target grating onto the sensor and then move the image swiftly by either moving the target or using a scanning mirror to move the image. Unfortunately the mock alignment system was not capable of performing these more sophisticated tests. Instead, the flashing LED's were used to qualitatively verify the expected performance.

In this case the LED's were placed such that two each were in the east and west quadrants and one apiece were in the north and south quadrants. This ensured that both of the amplifiers were operating without being starved for current. All of the LED's were held at DC except for one in the west quadrant. This particular LED was flashed with both small and large signals to examine the sensor time response. These experiments were repeated for different light levels by placing ND filters in front of the lens. The total DC currents from the quadrants were also measured to get an idea of the signal level.

The sensor bandwidth was measured for a small amplitude signal and is plotted in Fig-

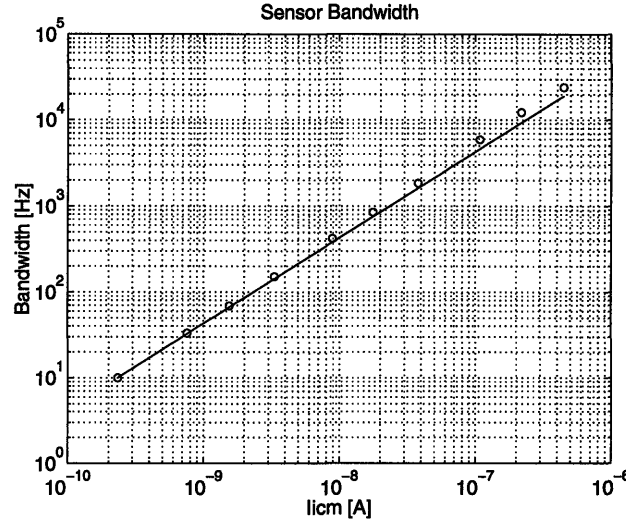


FIGURE 6-11: Small-signal bandwidth of the coarse alignment portion of the sensor plotted versus the amplifier common-mode input current. The solid line shows the current amplifier dominant pole. The bandwidth is limited by the amplifier rather than the absolute value circuit in this experiment.

Figure 6-11 versus the common-mode input current to the amplifier. It is important to realize this is not the actual photocurrent in the absolute value circuits, however it is proportional to it. The solid line in this plot shows the bandwidth limit imposed by the amplifier

$$f_h = \frac{I_{icm}}{2\pi(1+n)\frac{kT}{q}C_w} \quad (6.3)$$

where $C_w = 65$ pF is the capacitance of the summing wire which loads the input to the amplifier and forms the dominant pole. Since the measurements are accurately predicted by this equation, one can conclude that bandwidth was limited by the current amplifier rather than the absolute value circuit in this experiment. This is not surprising due to the large currents flowing in the few absolute value circuits that are illuminated by the two light spots. For example, assume that the total current I_{icm} arises from ten absolute value circuits. Then the photodiode currents are about one tenth of I_{icm} . The input impedance of both the current amplifier and the absolute value circuit is approximately set by the $1/g_m$ of the input devices. Thus, the impedance is inversely proportional to the current. Since the photodiode currents are about one tenth of I_{icm} , the impedance seen by the photodiode capacitance is about ten times that seen by the wiring capacitance. However, since the photodiode capacitance is almost 100 times smaller than the wiring capacitance, the photodiode capacitance pole is roughly ten times higher than the pole from the wiring capacitance in this experiment.

Another experiment was done in order to observe the large signal time response of the absolute value circuit. In this case the flashing LED was turned completely off and on in order to induce large signal transients in the illuminated photodiodes and associated absolute value

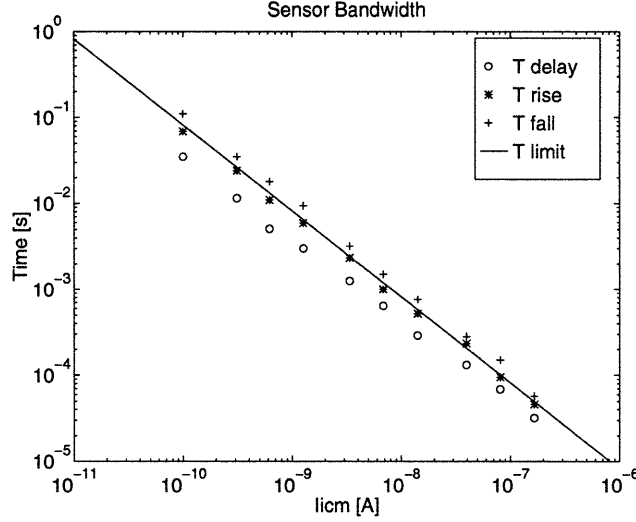


FIGURE 6-12: Transient delay time T_d , rise time T_r , and fall time T_f plotted versus the amplifier common-mode input current.

circuits. Although these transients were large for the absolute value circuits, they were kept small as a fraction of the total quadrant current. This prevented the current amplifier from also undergoing large signal transients.

The resulting waveform observed on an oscilloscope showed the same basic features seen in the Figure 5-19(a) simulation. After the onset of illumination there was a distinct delay time followed by the rise. The subsequent fall was about twice as long as the rise time. The delay, rise, and fall times were measured at various illumination levels and are plotted in Figure 6-12 versus the total measured common-mode input current to the amplifier. The solid line in the plot shows the settling time limit imposed by the current amplifier dominant pole. This is the rise/fall time calculated from Equation 6.3 assuming a single pole model

$$T = \frac{(2.2)(1+n) \frac{kT}{q} C_w}{I_{icm}} = \frac{8.18 \times 10^{-12} [\text{A s}]}{I_{icm}}. \quad (6.4)$$

Absolute value transients faster than this will be limited by the current amplifier transient response.

In order to apply the equations of Section 5.3.1 the pixel photodiode current is needed. This can be crudely estimated from the total signal current as follows. First, the amount of current due to the flashing photodiode can be obtained from the amplitude of the resulting output waveform and the measurement of I_{icm}

$$\Delta I = I_{icm} \frac{\Delta V}{V_{ocm}} \quad (6.5)$$

where $\Delta V \approx 200$ mV was the typical peak-to-peak amplitude of the waveform and

$$V_{ocm} = (0.82)(39.1 \text{ k}\Omega)(50 \mu\text{A}) = 1.6 \text{ V} \quad (6.6)$$

is the corresponding common-mode level of the output voltage. Thus one has $\Delta I \approx I_{icm}/8$. Because the flashing LED creates a spot about $180 \mu\text{m}$ across, it should illuminate between four and eight photodiodes. Of these about two to four should have enough edge signal present to generate an output. One can then estimate the pixel photocurrent in terms of the total quadrant signal as $I_p \approx \Delta I/3 \approx I_{icm}/24$.

The delay time can now be estimated as the time it takes the photocurrent to charge up the input capacitance and turn on device M1A

$$T_d = \frac{V_{on} C_i}{I_p} = \frac{8.88 \times 10^{-12} [\text{A s}]}{I_{icm}} \quad (6.7)$$

where $C_i = 740 \text{ fF}$ and $V_{on} \approx 0.5 \text{ V}$. The data shows a trend lying closer to $4 \times 10^{-12} [\text{A s}]/I_{icm}$ which is in reasonable agreement considering the numerous approximations involved. Rise time (10% to 90%) can be estimated from Equation 5.131 yielding

$$T_r = \frac{n \frac{kT}{q} C_i}{I_p} (4.39) = \frac{3.45 \times 10^{-12} [\text{A s}]}{I_{icm}}. \quad (6.8)$$

This is about two times faster than the limit imposed by the current amplifier. Thus, it is not surprising that the rise time data is found to lie along the limit in the figure. Similarly, the fall time (90% to 10%) can be obtained from Equation 5.133 resulting in

$$T_f = \frac{n \frac{kT}{q} C_i}{I_p} (8.89) = \frac{6.98 \times 10^{-12} [\text{A s}]}{I_{icm}}. \quad (6.9)$$

This is slightly below the limit imposed by the current amplifier whereas the data appears slightly above it. Note that this analysis has underestimated the fall time and overestimated the delay time. This might be explained by the fact that with several circuits being illuminated the observed delay time will be the shortest of the group and the fall time will be that of the slowest of the group. Considerable variation in the speed of the individual circuits would be expected given that they will receive different amounts of illumination from the spot. These measurements are consistent, given the experimental limitations, with the models and simulations discussed in Section 5.3.1 and illustrate the speed limitations of the coarse alignment operation of the sensor.

The shortest fall time measured was $60 \mu\text{s}$ which corresponds roughly to a bandwidth of 5.8 kHz. This is quite high because the current signal is roughly $I_{icm} = 170 \text{ nA}$. The estimated photodiode currents are about 7.1 nA which implies an equivalent irradiance of around 1.2 mW/cm^2 . This is very large compared to the maximum irradiance level achieved for the fine alignment measurements and the surveyor measurements. This occurs because a similar amount of light power is concentrated into a much smaller photodiode in this situation. An irradiance level of $100 \mu\text{W/cm}^2$ is more comparable with the earlier measurements which is twelve times less than the estimated maximum illumination level obtained in these experiments. At this irradiance the resulting bandwidth would be about twelve times less or 480 Hz. One

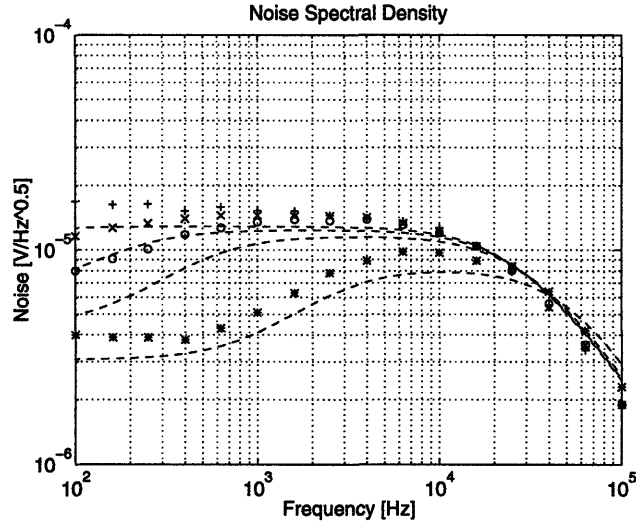


FIGURE 6-13: Output noise spectrum from the fine alignment portion of the sensor. From top to bottom the common-mode input currents were 11.5, 45.5, 120, and 703 nA corresponding to irradiances of 2.0, 7.8, 21, and 120 $\mu\text{W}/\text{cm}^2$, respectively. The dashed lines show values obtained using HSPICE simulation.

can also estimate a bandwidth from Equation 5.140. Using a conservative value $K = 12$ for the fall time results in a bandwidth estimate of about 540 Hz. This is in good agreement with the value extrapolated from the measurements.

6.5 Noise

The noise from the fine alignment portion of the square grating sensor is essentially the same as that from the surveyor's mark sensor. In this case there was no active cascode circuit and thus the noise is primarily due to the shot noise from the photodiodes and the bipolar devices at the input of the current amplifier (see Section 3.3.8). Noise measurements were made as described in Section 4.6 using photodiodes to illuminate each quadrant of the sensor.

Measurements of the noise spectral density are shown in Figure 6-13 at four different light levels. The data is very similar to Figure 4-14 from the surveyor's mark sensor. As seen there, the noise spectrum divides into two regions. The shot noise from the photodiodes and amplifier input devices dominates at low frequencies. At mid-frequencies the noise from the amplifier replaces the input shot noise resulting in a plateau around $15 \mu\text{V}/\sqrt{\text{Hz}}$. This agrees well with the expected value of $14 \mu\text{V}/\sqrt{\text{Hz}}$ from Equation 3.70. The agreement between the simulations and the data at low frequency is not particularly good and the reason for the discrepancy is not known.

The RMS noise at the output ranged from 1.9 to 2.4 mV as the light level was varied. Since the sensor yields a maximum signal of 1.7 V, the signal-to-noise ratio is 57 dB. A 3σ noise level

of 7.2 mV implies an alignment repeatability of 1.1 steps. This corresponds to a repeatability of 30 ppm of the sensor field of view.

6.6 Summary

The test results from the square grating sensor are similar to those from the surveyor's mark sensor. The absolute value circuit was found to be affected by substantial mismatch in the devices, although this effect was not as large as had been anticipated. The coarse alignment signal saturated with a sharp transition at the aligned position due to the single-sided nature of the error signal currents. By adding appropriate bias currents a more reasonable signal could be obtained with a slope of about $200 \mu\text{V}/\text{step}$ near the aligned position. The fine alignment signal was successfully sensed with the spatial phase detector and had a slope of $6.5 \text{ mV}/\text{step}$. Alignment repeatability was limited by the x - y table to about one positioning step which corresponds to 25 ppm of the sensor field of view. Bandwidth was found to be proportional to sensor irradiance and the fine alignment signal demonstrated a bandwidth of 3 kHz. However, the real limitation on sensor speed is determined by the settling time of the absolute value circuits used for the coarse alignment. For bright levels of illumination this bandwidth was estimated at about 480 Hz. The worst case measured noise corresponded to a position uncertainty of 30 ppm of the sensor field of view. Because the square grating sensor did not incorporate active cascodes, it had both lower bandwidth and lower noise than the surveyor's mark sensor.

Comparison

7.1 Introduction

The two sensor designs presented in this thesis embody different philosophies in the alignment algorithm. In this chapter a brief discussion will be presented emphasizing the similarities and differences between the sensors. The most obvious difference is in the alignment marks they are designed to detect. These two marks reflect a division that can be made between the various alignment techniques that were summarized in the second chapter. The major similarities between the two sensors lie in the details of the implementation.

7.2 Alignment Methods

A review of the numerous papers published on alignment systems reveals a wide variety of alignment techniques. The choice of target marks and detection schemes differs greatly between manufacturers and even between systems from the same manufacturer. There seems to be little agreement on an optimal alignment mark, detector, or algorithm. In the second chapter the various alignment methods were divided into categories such as correlation, edge detection, moments, and so forth. In this spectrum of alignment techniques one can make a rough classification that serves to separate the methods into two groups. The distinction concerns the number of features, or amount of information in the alignment mark.

One group of methods depends upon a single feature, or small number of features, detected in an alignment mark. For example, edge detection techniques are concerned only with the position of edges. Similarly, correlation systems search for a single image feature or shape that matches the template pattern. Centroid methods might also be considered in this class since the centroid of a target is the single important feature. The strength of these single feature techniques lies in their insensitivity to defects in the target across most of the sensor field of view. For example, a correlation system will not usually lock onto extraneous features or defects

in the image unless they correspond almost exactly to the template. However, in this strength also lies a weakness. Any defect in the specific feature these systems search for can cause them to fail. For example, in a correlation system a defect in the target mark can cause it to no longer adequately match the template.

In contrast to single feature methods are techniques that make use of a large number of features, or a pattern, in the alignment mark. The chief examples of these techniques are geometric moiré and diffraction methods. These systems use the information present in a large pattern to generate a position signal and are, in a sense, the opposite of the single feature methods. Pattern methods are generally insensitive to local defects in the pattern. However, broad distortions in the pattern can cause large errors. Nonuniform illumination, asymmetrically blazed diffraction gratings, and magnification error are examples of this type of distortion.

The two alignment sensors presented in this work can also be distinguished along these lines. This is most clear in how they perform the fine alignment. The surveyor's mark sensor generates the fine alignment error signal by detecting the four edges in the mark. It can be considered as a single feature method. The error signal is relatively insensitive to nonuniform illumination, a broad distortion, but is sensitive to defects that occur at the edges such as photodiode mismatch and errors in the target itself. In contrast, the spatial phase detection used by the square grating sensor is clearly a pattern method. The moiré fringe results from the superposition of the two grating patterns. This sensor is less sensitive to local defects in the gratings, but it is very sensitive to nonuniformity in the illumination and to magnification error.

The coarse alignment methods used by the two sensors do not divide as clearly between single feature and pattern methods. The surveyor's mark coarse alignment can be viewed as an edge detection method or as a system which detects total irradiance in the quadrants as the features. However, since the photodiodes encompass the entire sensor field of view, broad effects like nonuniformity in illumination can cause significant errors in alignment. Similarly, it is difficult to classify along these lines the coarse alignment algorithm performed by the square grating sensor. This sensor detects edges as features and sums them to obtain the error signal. At large displacements it resembles a pattern method due to the large number of edges. It is sensitive to nonuniform illumination and is insensitive to any defect that might cause a spurious edge to be detected. However, near the aligned position the number of detected edges becomes small. In this case nonuniform illumination is less of a problem, whereas any incorrectly detected edges will cause significant alignment error.

7.3 Errors

The sensors differ in their vulnerability to several error sources. Magnification error is one example. For the surveyor's mark sensor it was assumed that the image of the mark was twice

the size of the sensor. In fact, the mark can be considerably larger or smaller than this and the alignment sensor will still function. There will be a loss of signal strength for marks smaller than the ideal but no penalty occurs for marks that are too large. This sensor has considerable insensitivity to magnification errors.

This is not the case for the square grating sensor. Although the coarse alignment signal is also insensitive to magnification errors, the fine alignment signal is very dependent upon the proper magnification. The moiré fringe is created by the slight mismatch in the pitches of the sensor and mark gratings which is determined by the magnification. A small error in magnification will cause a distortion of the grating that can easily destroy the fringe. The amplitude factor dependence on pitch mismatch, plotted in Figure 5-15, shows this sensitivity. The square grating sensor has an optimal pitch mismatch of $\epsilon_p = 0.023$ which means a magnification error of 2.3% is enough to erase the moire fringe. If the number of grating lines is increased, then not only is the displacement sensitivity increased, but so is the sensitivity to magnification error. Unlike the surveyor's mark sensor, the square grating sensor requires tight control of the magnification.

Another source of error arises from rotations of the mark relative to the sensor. Both designs show a similar intolerance for rotations which corrupt the fine alignment error signal. The surveyor's mark sensor can tolerate rotations up to approximately (from Equation 3.27)

$$\theta_c \approx \frac{2W_e}{L_q} \quad (7.1)$$

whereas the square grating sensor can tolerate (from Equations 5.113 and 5.118)

$$\theta_c \approx \frac{P_s}{3L}. \quad (7.2)$$

These equations are both of the form

$$\theta_c \approx k \frac{W_c}{L_s} \quad (7.3)$$

where W_c is the width of the fine alignment capture region, L_s is the sensor length, and the factor k is a number on the order of unity. The surveyor's mark sensor can tolerate rotations up to 5.2 degrees whereas the square grating sensor can only tolerate rotations of about 1.2 degrees. This difference is not due to any inherent advantage of the surveyor's mark, but rather to the fact that the surveyor's mark sensor was designed with a larger capture region relative to the sensor length.

The coarse alignment error signal from both sensors shows more tolerance to rotation error than the fine alignment error signal. However, in this case the square grating sensor is definitely superior. The surveyor's mark sensor can only tolerate rotations up to 45 degrees, whereas the square grating sensor is able to provide alignment signals at any rotation. For rotations near 45 degrees this signal becomes weak and has an offset beyond the capture region of the fine alignment. No effort was spent attempting to improve the rotation tolerance of this sensor.

It seems likely that better performance is possible if one redesigned the photodiode layout to more easily detect rotated edges.

Another source of error that considerably affects the alignment is nonuniform illumination. It is assumed that this is a slowly varying function across the target in which case it was shown that the edge detection method employed by the surveyor's mark fine alignment was much less sensitive to this error than the coarse alignment. This occurs because the fine alignment signal is derived from only a thin slice of the image over which the illumination varies little. The reverse is true for the square grating sensor. The fine alignment signal is obtained by integrating light over the entire image and thus it is very sensitive to illumination nonuniformity.

7.4 Capture Regions

A significant difference between the sensors is the size of the coarse alignment capture regions relative to the target mark size. For the square grating sensor there is a good signal over the entire field of view and it is conjectured that this is the case as long as any portion of the square grating is within the sensor field of view. Thus a mark of length $2L$ can be captured over distances up to $2L$ from the aligned position (L_∞ norm). This is much better than the surveyor's mark sensor which requires a mark of size $4L$ that can only be captured over distances up to L away from the aligned position.

The two sensors also have different fine alignment capture regions, although this difference is not as dramatic as for coarse alignment and is the result of different design choices. The capture region for the surveyor's mark sensor was sized to be ten percent of the field of view due to concerns about the effect of nonuniformity of illumination. Because this is less of a concern for the square grating sensor, it was designed with a capture region of only 3.3%. Even with the moderate illumination uniformity in the mock alignment system, both of these capture regions could have been reduced further without causing difficulties.

7.5 Implementation

Although the alignment sensors differ greatly in their alignment methods, they have many similarities in their implementation. Both use the same photodiode structure to perform the light sensing. They also use similar circuits to amplify the detected photocurrents. The same differential normalizing amplifier is used for the fine alignment signal of the surveyor's mark sensor and for both the coarse and fine alignment signals in the square grating sensor. The coarse alignment signal of the surveyor's mark sensor requires an additional four input normalizing amplifier that doubles the amplification circuitry on this chip. These amplifiers are located at the edge of the imaging area on both sensors.

One implementation difference is that the square grating sensor has circuits throughout

the imaging area whereas the surveyor's mark sensor has none. These circuits are necessary to locally compute the absolute value function. They reduce the fill factor of the sensor, and hence the signal strength, by a factor of two. The complexity of the absolute value circuit is also restricted by the need to make it compact enough to fit inside the pixel area. This is a definite disadvantage of the square grating sensor because there is little room to add features such as a buffering amplifier like the active cascode that would increase the sensor bandwidth.

7.6 Performance

7.6.1 Alignment Signals

The alignment error signals from both sensors are fairly similar overall, although there are some minor differences. For coarse alignment the normalized signals were comparable (depending upon the bias current used for the square grating sensor), although the error vectors from the square grating sensor yielded more direct paths to the origin. The fine alignment signal from the surveyor's mark sensor was a triangle wave compared to a sinusoid from the square grating sensor. The error vector fields and the resulting fine alignment trajectories from both sensors looked very similar.

The square grating sensor had a significantly higher signal slope at the aligned position than did the surveyor's mark sensor. This was primarily due to its smaller capture region. This led to a somewhat higher repeatability due to the quantization error from the analog-to-digital conversion. In practice the repeatability for both sensors was limited by the resolution of the positioning system and the analog-to-digital conversion.

It is also interesting to compare the magnitude of the detected photocurrents for the two sensors. This magnitude sets the upper limit on the signal-to-noise ratio determined by shot noise. The fine alignment signal of the surveyor's mark sensor has relatively small photocurrents due to the small size of the photodiodes that detect the edges of the mark. In contrast, the error signal currents are much larger from the square grating sensor because the photocurrents are summed over the entire sensor area. This situation is reversed when one examines the coarse alignment signals. In this case the surveyor's mark sensor has large signal currents due to the large size of the four quadrant photodiodes. On the other hand, the square grating sensor has relatively smaller currents because signal is only gathered from pixels where there is an edge and then only if the contrast exceeds the cutoff threshold built into the absolute value circuit.

7.6.2 Bandwidth

The difference in signal strength has an effect on the potential circuit speed. The bandwidth of the current amplifier is proportional to the magnitude of the input signal currents. All other things being equal, the small fine alignment signal of the surveyor's mark sensor will be slower than the larger signal from the square grating sensor. The reverse will hold for the coarse

alignment signal. This is true only if the current amplifier is limiting the sensor dynamics. In the chip designs presented in this thesis this was not the case. The dominant pole was always due to the photodiodes, although that need not be the case if the active cascode is designed more aggressively.

The measured sensor bandwidth was clearly superior for the surveyor's mark sensor. However, this was due to the active cascode circuit which was omitted from the square grating sensor design. Both sensors have essentially the same bandwidth limitations imposed by the photodiode depletion capacitance and the amplifier input impedance. The exception is the absolute value circuit used for coarse alignment in the square grating sensor which is not as amenable to simple modifications that could improve its speed.

7.6.3 Noise

The difference in signal strengths also affects the ideal noise performance. The fundamental noise limit of these sensors comes from the shot noise of the photocurrents. Because shot noise is proportional to the square root of the current, the larger the total photocurrent, the larger the signal-to-noise ratio. The noise in the fine alignment signal ultimately determines the repeatability of the alignment signal. Because the square grating sensor collects a much larger photocurrent for this signal, it should have a superior repeatability. For coarse alignment, the surveyor's mark sensor has the largest signal. Setting other limitations aside, this implies that a smaller capture region could be used which would tend to counteract the relatively lower signal-to-noise ratio of the fine alignment signal. In practice, however, the capture region will likely be determined by other factors such as diode mismatch and illumination uniformity that will exceed the noise limit.

The noise limits set by the shot noise were not reached in the implemented designs. Noise from the current amplifier was generally larger than the photodiode noise and mostly determined the RMS noise of the signal. Both sensors had essentially identical noise spectra because they used identical current amplifiers (ignoring the active cascode) and photodiodes. If lower noise amplifiers were designed so that the photodiode noise became dominant, then the square grating sensor would offer superior performance.

7.6.4 Power

Almost all of the power dissipated by the sensors is consumed by the current amplifiers and thus the sensor power requirements are similar. The surveyor's mark sensor dissipates more power because it requires the additional coarse alignment amplifier and also has the active cascode circuits for improving the sensor speed.

Characteristic	Surveyor		Square Grating	
	Coarse	Fine	Coarse	Fine
Method	Brightness comparison	Edge detection	Edge detection	Moiré pattern
Signal	± 2.7 V	± 1.1 V	± 1.1 V	± 1.7 V
Slope	140 μ V/step	1.1 mV/step	200 μ V/step	6.5 mV/step
Capture region	0.5	0.05	2	0.033
Rotation limit	45 deg.	5.2 deg.	none	1.3 deg.
Bandwidth	7.5 kHz	55 kHz	480 Hz	2.5 kHz
RMS noise (σ)	-	1.2 mV	-	2.4 mV
Repeatability	-	80 ppm	-	30 ppm
Power dissipation	4.0 mW		1.2 mW	

Table 7-1: Summary of measured performance. Bandwidths for the square grating sensor are lower because it did not contain the active cascode circuit.

7.7 Summary

Table 7-1 summarizes various aspects of the two sensor designs. The signal strength values state the maximum range of the measured alignment signals. The signal slopes are for the x -direction error signal at the aligned position. The maximum signal and signal slope for the coarse alignment of the square grating sensor are stated with the 6 nA bias current added to the error signal. The capture region number is a dimensionless figure expressing the ratio of the size of the capture region to the size of the target mark's image. The bandwidths are for an irradiance of 100 μ W/cm², approximately the brightest level achievable in the test set-up. The bandwidths reported for the surveyor's mark sensor use the active cascode circuit. The bandwidth reported for the coarse alignment of the square grating sensor is the estimate from the transient fall time measurements. The rotation limits express the maximum rotation that can be tolerated. Practical rotation limits will be less than these. The RMS noise is the lowest noise level observed on the fine alignment signal with full illumination. This limits the possible alignment accuracy at high bandwidth and is reported as the repeatability expressed as a fraction of the sensor size.

The contrasting strengths and weaknesses between the two sensors prompts the question: Could both sensors be used together in such a way that the combination would be superior to using either sensor alone? The answer to this question is probably no. The two alignment sensors embody the fundamental differences between the single and multiple feature alignment techniques. They have complementary strengths that are not easily blended. For example, consider a situation in which the image suffered from both nonuniformity in the illumination

and small, evenly distributed spot defects. If both sensors were used they would each yield alignment information with offsets due to the image defects. A single feature method would see errors due to the spot defects and a multiple feature method would have difficulty with the nonuniform illumination. Without knowing which image noise source is dominant, one cannot decide which of the sensor signals is more accurate. The choice of which sensor technique to use depends upon knowledge of the expected image noise.

Potential Improvements

8.1 Introduction

Although the sensor designs in this thesis work well, testing them and further reflection upon their operation revealed a variety of ways in which they could be improved. This chapter will present several ideas covering the alignment algorithm down to the circuits which should yield superior performance in a revised sensor design. Some of the proposed changes are simple to implement whereas others would require additional design effort. Some tasks that were left incomplete or could have been done better in this work will also be pointed out.

8.2 Photodiodes

The photodiodes used in these sensors have plenty of room for improvement. The original design did not consider what would be the best reverse bias to place across them. Experiments revealed that increasing the reverse bias improved the quantum efficiency somewhat, presumably by enlarging the depletion region. This has the added benefit of reducing the depletion capacitance. Both effects would serve to improve the sensor bandwidth. On the other hand, an enlarged depletion region leads to more dark current which might be a concern.

A better way to improve the photodiodes might be to choose a different pn junction in the process. Both sensor designs used p -type source/drain diffusions in n -wells to create photodiodes that have a quantum efficiency of 27%. This junction was chosen because it was expected to collect more light due to its relatively shallow depth. Furthermore, the wells provide isolation from the substrate. However, other people have achieved much better performance by using the junction between the well and substrate for the photodiodes. Aswell et al. [111] used an n -well to p -substrate junction to form photodiodes with a quantum efficiency of 85% for 660 nm illumination. Similarly, Aubert et al. [112, 113] characterized a number of junctions in a p -well CMOS process and found that the best was the p -well to n -substrate junction. This

yielded a quantum efficiency of about 90% for 620 nm illumination. The worst junction was the n^+ -diffusion in the p -well which had a quantum efficiency of only 27% and is analogous to the junction used in the alignment sensors. In general they found that the deeper and more graded junctions were better at collecting the photo-generated carriers than were the shallower abrupt junctions. They further improved the quantum efficiency by using special processing to remove the protective nitride and to thin the oxide above the photodiodes in order to reduce interference effects. Another advantage of the well-to-substrate junction is its substantially lower capacitance. In Aubert's process this was lower by about a factor of six.

It would be a simple change to use the well junction rather than the source/drain junction to form the photodiodes in the Orbit process. If an improvement is seen similar to that reported in these papers, then one could potentially raise the quantum efficiency by about a factor of three while lowering the capacitance by a factor of six. Thus, at a given light level the photodiode pole would be pushed out by a factor of eighteen compared to the present implementation. Furthermore, the additional photocurrent would increase the signal relative to the shot noise. These benefits could be realized without any special processing, but they do come at a small price. The deeper and more lightly doped well diffusion also needs greater lateral spacing to isolate it. Thus, the gaps between photodiodes would become larger. The Orbit process design rules would require that the gaps be increased from 3 μm to 9 μm . Another potential problem is that these junctions are no longer isolated from wandering minority carriers in the substrate. This might increase the dark current and allow minority carriers generated under one photodiode to be more easily collected by an adjacent photodiode.

8.3 Current Amplifier

The surveyor's mark sensor incorporated an active cascode circuit to buffer the photodiode depletion capacitance from the amplifier input impedance. This moved the dominant pole out to higher frequency and significantly improved the sensor bandwidth. It also added some new noise sources to the signal. A variety of possible improvements could be made to this circuit to improve its bandwidth and lower its noise. The primary source of noise appears to be flicker noise arising from the NMOS transistor active load (M8 and M12 in Figure 3-8). This device could be replaced with an NPN bipolar transistor which would eliminate the flicker and thermal noise and replace it with a much lower shot noise source. Furthermore, the area saved could be used to make the PMOS gain transistor even wider to further reduce its noise. These changes should lead to a significant reduction in the noise from the active cascode.

Another step up in complexity would be to use a high gain voltage amplifier, such as an operational amplifier, in place of the inverter in the active cascode. This was also done by Aswell et al. in their design. By increasing the gain of this stage the photodiode pole can be pushed so far out as to be irrelevant. The stability of such an amplifier would have to be

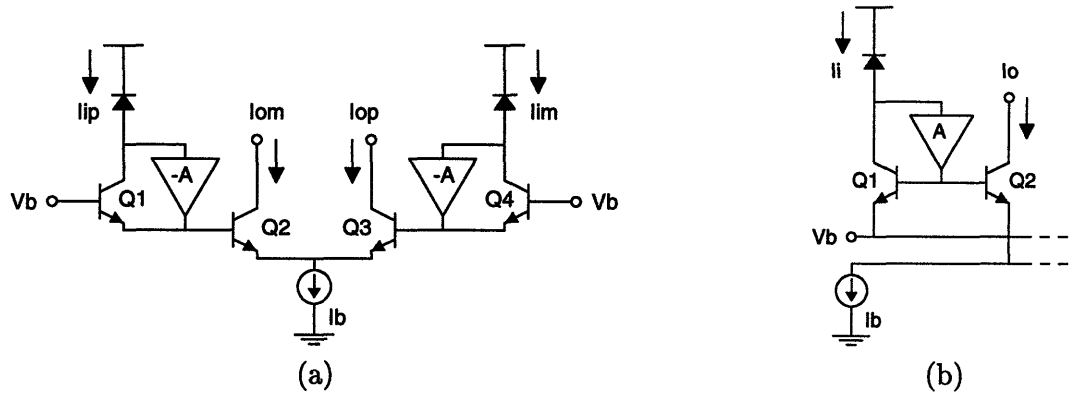


FIGURE 8-1: Adding gain around the bipolar transistors is a way to improve the circuit bandwidth for the (a) inverting multiplier (after Aubert [113]) and (b) one stage of a multiple input normalizing current amplifier. The high gain amplifiers fix the diode reverse voltage and, for circuit (a), the collector-base voltage.

carefully considered as well as the noise generated by it.

Another possibility is to replace the active cascode with a high gain amplifier around the bipolar transistors. For example, Aubert et al. constructed a differential normalizing amplifier out of the inverting multiplier as shown in Figure 8-1(a). Here the gain of the amplifier is placed between the collector and emitter of the bipolar devices. This serves to hold both the reverse voltage across the diode and the collector-base voltage constant. Thus, the photodiode pole is eliminated as well as the nonlinearity caused by the base width modulation in the input bipolar transistor. Aubert et al. used inverters for this amplification, but one could use higher gain amplifiers as well. They reported rise times of better than $1 \mu\text{s}$ for current steps of 0 to 3 nA from this circuit.

The inverting multiplier could be used in the alignment sensors for the two input normalization. However, the four input normalization used by the coarse alignment error signal in the surveyor's mark sensor requires a different multiplier topology. This normalization was accomplished using the circuit shown in Figure 3-15 and it could be improved by replacing the active cascode with an amplifier as shown in Figure 8-1(b). The amplifier again serves to hold the voltage across the photodiode constant. In this case the collector-base voltage will still vary. Placing high gain around bipolar transistors in this way raises stability issues that would need to be addressed before proceeding with such a design.

One could also replace the active cascode with a circuit to bootstrap a constant voltage around the photodiode. Figure 8-2 shows how this could be done using a source follower. As long as the photocurrent is small compared to the source follower bias current, the photodiode voltage will be held constant at the V_{gs} of the follower device. This requires a reverse biased floating diode which necessitates the use of the junction between the source/drain diffusion and the well. Unfortunately, this is the junction that appears to have the lowest quantum efficiency. Furthermore, the wells for individual photodiodes would have to be isolated which

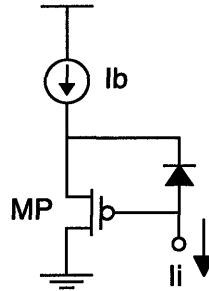


FIGURE 8-2: A source follower can be used to bootstrap the voltage around the photodiode. By holding this voltage fixed the bandwidth can be improved.

would enlarge the photodiode gaps. Finally, in order to substantially improve the bandwidth, the gain of the source follower would need to be made very close to unity which is a difficult task. Despite these disadvantages, this circuit might have compensating benefits in terms of noise that should be investigated.

Assuming the above techniques are successful in removing the photodiode limit on the bandwidth, then one would have to focus on improving the bandwidth of the current amplifier itself. One possibility would be to run the initial normalizing gain stage at a higher current level. The original design used both normalizing and inverting current gain stages to lower the gain per stage. This was done in order to minimize the diffusion component of the base capacitance by running the input stage at low current. However, the current levels could be raised considerably before the diffusion capacitance would become comparable to the already present base-emitter depletion capacitance. By doing so one could improve the speed of the normalizing multiplier and eliminate the second gain stage. This was already done for the four input current amplifier of Figure 3-15.

The current amplifier speed could also be improved by altering its layout. All of the bipolar transistors in this circuit were placed in a common centroid layout that used four parallel devices to realize each circuit device. This reduces offset problems related to device mismatch. However, the extra devices and interconnect increase the parasitic capacitances. Reducing the symmetry of the layout could lower these capacitances by a factor of four or more. The penalty would be an uncertain degradation in the device matching which would affect the amplifier offset.

8.4 Square Grating Sensor

8.4.1 Analysis

The fine alignment portion of the square grating sensor was modeled as a spatial phase detection of the position of a moiré fringe. Several approximations were made in the mathematical analysis that should be reviewed. Some of these were made to reduce the detail which

threatened to cloud the meaning behind the equations. Others were made in order to obtain equations that could be solved. The end result was an interpretation of the sensor as a spatial phase detector which revealed the underlying compromises between the sensor size, grating pitch, pitch mismatch, and sensitivity. However, the equation for the error signal was not a quantitative model that could be directly compared to the measured data.

The first, and most important, approximation was to model the mark as a sinusoidal function in Equation 5.66. In fact, the mark is better modeled by a square wave in irradiance that alternates between full illumination E_0 and darkness ϵE_0 with ϵ being the image contrast parameter. The omission of this parameter, which is not insignificant, is a major deficiency in the quantitative accuracy of the model. A more appropriate representation of the mark would be a Fourier series expansion of the square wave function

$$m(x) = b_0 + \sum_{q=1}^{\infty} b_q \cos qk_m(x - \xi). \quad (8.1)$$

The sampling process of detection by the photodiode grating would then result in a more complex equation for the sequence of photocurrents than Equation 5.83. This would also have a series form

$$f[n] = c_0 + \sum_{q=1}^{\infty} c_q \cos q\Omega_{\Delta}(n + \frac{\xi}{\epsilon_p P_s}). \quad (8.2)$$

This new sequence can probably be modeled as samples of a continuous function $f_L(x)$ that represents the moiré fringe formed by two square gratings. This fringe function has been shown to be a trapezoidal waveform [71]. The analysis completed in Chapter 5 essentially only treated the fundamental component of this series for simplicity and clarity.

The analysis of the spatial phase detection made some additional approximations. First, the model assumed a spatial layout of the photodiodes that was slightly altered from reality in order to simplify the mathematical notation in the model. Adhering more closely to the actual layout could be done, but it is probably not worth the effort. A great deal of notational complexity would be added that should not materially affect the accuracy of the result. This approximation resulted in a simple discrete model of the spatial phase detector described by Equation 5.99. It is probably possible, although difficult, to sum this sequence with either the original approximate fringe sequence $f[n]$ or with the more accurate version from above. Approximating this sum with an integral would again simplify the analysis greatly without compromising the accuracy of the model to a great extent. Modeling the moiré fringe function with a Fourier series

$$f_L(x) = d_0 + \sum_{q=1}^{\infty} d_q \cos qk_{\Delta}(x + \frac{\xi}{\epsilon_p}) \quad (8.3)$$

and integrating according to Equation 5.105 will result in an error signal of the form

$$e_{cx}(\xi) = \sum_{q=1}^{\infty} g_q \cos qk_{\Delta}(x + \frac{\xi}{\epsilon_p}). \quad (8.4)$$

The error signal is no longer a simple sinusoid with a single amplitude factor. Instead, it is a Fourier series that describes a more complex waveform. Furthermore, the series coefficients depend upon the pitch mismatch ϵ_p and corresponding spatial frequency k_Δ . This implies that the shape of the error signal will vary slightly with the pitch mismatch. Thus, the relationship between the error signal amplitude and pitch mismatch will be more complex than the earlier model which contained only a single amplitude factor for the fundamental component. However, this equation should also yield a more quantitative prediction of the measured data because the coefficients g_q will incorporate the effect of the imperfect image contrast ϵ on the error signal.

8.4.2 Implementation

A number of changes could be made to the square grating sensor in order to improve its performance. The most important of these is to repair the coarse alignment signal which saturates for even small displacements from the aligned position. This arises from the algorithm and not in any implementation flaw. A displacement of the square grating image along the positive x -axis results in edge crossings in the east, north, and south quadrants and no crossings in the west quadrant. This results in a zero error current from the west quadrant and nonzero currents from the others. The normalizing amplifier output for the x -direction error signal will be saturated as long one input is zero and other input is nonzero, regardless of its value. This flaw causes two major problems. First, the saturating characteristic is awkward to use as an error signal for alignment. The high gain transition in the signal at the aligned position requires a low gain control loop to avoid chatter about this position. Another problem is that all of the quadrant signal currents become small near the aligned position which causes the current amplifier to become very slow.

It was shown in Chapter 6 that these flaws could be eliminated by adding a constant bias current to all of the quadrant error currents. The value of this bias current should be comparable in magnitude to the maximum possible quadrant current. Furthermore, it should be proportional to the sensor irradiance, as are the error currents. Such a bias current could be generated in several ways. One could sum a copy of all of the pixel photodiode currents across the imager and then copy the sum to create the bias currents. This would require additional wiring and circuitry to be added to the pixels in the sensor array. A less complex solution would be to designate a small portion of the sensor area to measure the light level. A photodiode the same size as the pitch spacing located in the middle of each quadrant would be adequate. The photocurrent from these diodes could be scaled using a fixed, or perhaps controllable, gain and then copied to generate the bias currents. The copies would need to be accurate in order to avoid creating an offset in the error signal currents.

Another small improvement could be made by exploring more deeply the best cutoff threshold to build into the absolute value circuit. It could be seen in Figure 6-1 that the cutoff point used in the design was too conservative. Lowering this cutoff threshold would increase

the average edge detection signal. This would in general increase the bandwidth and lower the signal-to-noise ratio for a given illumination level. The choice of the cutoff point depends greatly upon the matching characteristics of the transistors in the process. A study of these matching properties would have to be done to determine the best choice for the cutoff point.

The square grating sensor generally had lower bandwidth than the surveyor's mark sensor because it did not contain any extra circuitry to buffer the photodiode capacitance. At a minimum the active cascodes used in the surveyor's mark sensor should be added to improve the bandwidth. In addition, the techniques discussed above to improve the speed of the current amplifier could be incorporated. The fine alignment portion of this sensor consists of a large number of photodiodes wired together into quadrants that connect to the current amplifier inputs. This is similar to the coarse alignment of the surveyor's mark sensor which has large photodiodes connected to the amplifier. Thus, the fine alignment signal of the square grating sensor should be capable of achieving similar bandwidths.

Increasing the bandwidth of the coarse alignment signal is more difficult. Active cascodes or similar remedies can be used to buffer the current amplifier inputs from the significant capacitance of the summing wire. However, the chief limitation on the bandwidth in this case is not the current amplifier. Instead, the absolute value circuits that perform the local current subtractions are the bottleneck. Because of the small size of these currents (1 to 100 pA), this circuit is quite slow. There are several ways one can attempt to surmount this difficulty. One possible solution is to incorporate the active cascode into the absolute value circuit. Only a very simple version of the active cascode could be squeezed into the small area available in the pixel. An inverter might be acceptable but a high gain operational amplifier is likely to be too large. Furthermore, such a circuit would only address the photodiode capacitance pole and not the current mirror poles. Another possible solution is to place some current gain in front of the absolute value circuit by using parasitic phototransistors in place of the photodiodes [114]. This would increase the current signal by the gain of the transistor and correspondingly speed up the absolute value circuit. Finally, one could try to avoid computing the absolute value with the very small photocurrents. Instead, the photocurrents could be converted to a voltage signal and a folding amplifier or similar circuit running at higher currents could be used to compute the absolute value.

8.5 Rotation Alignment

The square grating sensor was designed to detect translational misalignment of the target mark. When considering the effects of rotation it was realized that the mark could possibly still align a rotated target. Figure 8-3(a) shows the situation when the mark is rotated 45 degrees relative to the sensor at the aligned position. In this case the edge crossings occur in equal number in all four quadrants. When the mark is translated to the right, as shown in part (b)

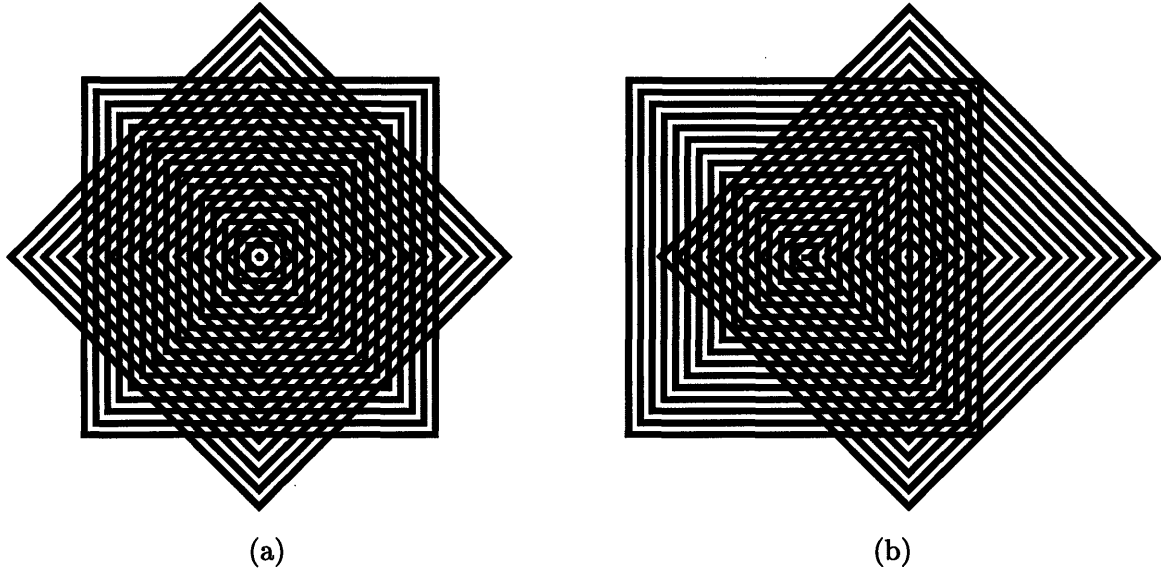


FIGURE 8-3: The pattern formed by (a) rotating one of the square gratings at the aligned position and (b) then translating it to the right. The number of edge crossings in each quadrant can still be used as an alignment signal.

of the figure, the number of edge crossings is reduced in the west quadrant and increased in the east quadrant. Thus, it appears that the sensor can be used to align a target which is rotated. Furthermore, at the aligned position all of the quadrant signal currents are nonzero and equal. Thus, there is no need for bias currents. On the other hand, nonuniform illumination will cause an offset in this case. The alignment signal and its capture region as a function of rotation could be studied more deeply.

Some experiments were done measuring the alignment signal at a few rotations. It was observed that the strength of the edge detected signal weakened as the rotation angle increased. This is explained by the fact that between every photodiode that is partially occluded by an edge there must be at least one fully illuminated, or fully darkened, photodiode in order to obtain the maximum difference signal across the edges. This condition was used to determine the maximum photodiode spacing in terms of the grating pitch

$$P_d \leq \frac{1}{4}P_s \quad (8.5)$$

and is illustrated in Figure 8-4(a). When the edges are rotated, as shown in part (b) of the figure, partially occluded photodiodes are adjacent to one another and the difference signal between them is reduced. In order to strengthen the detected signals one needs to reduce the photodiode pitch P_d . To ensure that an unoccluded photodiode lays between each edge for a 45 degree rotation would require a photodiode pitch of

$$P_d \leq \frac{1}{4}(\sqrt{2} - 1)P_s. \quad (8.6)$$

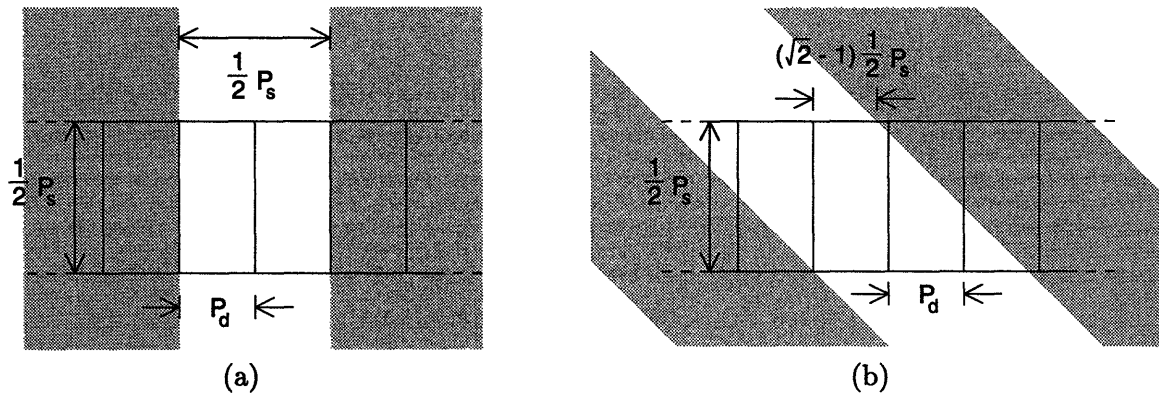


FIGURE 8-4: (a) The unrotated image edges need to be spaced at least $2P_d$ apart to obtain the maximum signal difference between adjacent photodiodes. (b) When the edges are rotated by 45 degrees, adjacent photodiodes are occluded and the difference signal between them is reduced. This effect can be minimized by decreasing the photodiode pitch relative to the grating pitch.

This is about a 60% reduction of the pitch implemented in the current design. With this spacing edges rotated by up to 45 degrees would still be detected at full strength.

It is also interesting to note that the sensor could be used to remove rotational alignment errors in addition to translational errors. This could be done as follows. A rotated target would first be translated to the aligned position by minimizing the differences between the quadrant error currents. At the aligned position all four currents would be equal. If they are not also zero, then the target is rotated in the direction to reduce them to zero. If the sensor was divided into octants, by splitting the quadrants in half, one could even detect the necessary direction of rotation. Further work needs to be done studying this possibility, but it appears that the sensor could be used to align a target with both rotational and translational errors.

Similar changes could be made to the surveyor's mark sensor to detect rotations. Again the strength of the error signals for both the coarse and fine alignment depends upon rotation. For example, a large rotation of the surveyor's mark about the aligned position (see Figure 3-1) would reduce the irradiance in quadrants 1 and 3. Thus, one could rotate the target about the aligned position to maximize these photocurrents. Small rotations could be discerned with the edge detection photodiodes. For example, a small rotation about the aligned position would cause a difference in the irradiance between photodiodes B and E. Minimizing this difference would remove the rotation error. The sensor could be modified to detect this rotation error by rewiring the photodiodes to one of the differential current amplifiers. Photodiodes A, B, C, G, H, and I could be connected to one input and the remaining edge detecting photodiodes to the other. This would form a simple detector for small rotations of the target about the aligned position.

8.6 Focus

Both alignment sensors can also be used for adjustment of focus, or equivalently z height to the target. In fact, the alignment sensor signals were used to adjust the focus in the mock alignment system. For example, to focus the image on the surveyor's mark sensor the target was aligned using an initial guess at the correct focus. The target was then translated to the point where the fine alignment edge detection signal was at a maximum. This occurs when the imaged edge falls between the outer edge detection photodiodes and the center ones. The height of this signal maximum depends upon the sharpness of the image. The focus is adjusted to maximize the fine alignment signal at this position and thereby bring the image into focus.

The square grating sensor can be focussed in a similar way. In this case the coarse alignment signal is based upon edge detection. The target is initially misaligned such that there are a significant number of edge crossings in a quadrant. The focus is then adjusted to maximize this error current. The fine alignment signal from the square grating sensor can also be used as a sensitive indicator of height. It was shown in Section 7.3 that the magnitude of this signal, determined by the amplitude factor, is very sensitive to the optical magnification. This, in turn, depends on the distances between the object, lens, and imager. One can position the target in z by translating the target to a maximum in the fine alignment signal and then adjusting the height to maximize this signal.

8.7 Summary

In this chapter a variety of ways were presented that could improve the sensor designs and the understanding of their operation. Several modifications to the circuitry in the sensors were suggested. The photodiodes could be improved in a number of ways to increase their responsivity and reduce their capacitance. Alterations to the current amplifiers were also presented that would increase their speed and reduce their nonlinearity. Similarly, potential changes to the absolute value circuitry in the square grating sensor were covered in the hopes of increasing its speed as well. The shortcomings in the analysis of the spatial phase detection used by the square grating sensor were discussed. A more thorough analysis was outlined and its expected results were presented. Some methods were examined to extend the operation of the sensors to detect rotation and focusing (or height) errors in the position of the target. The modifications suggested in this chapter could dramatically increase the sensors' bandwidth, lower their noise, and extend their functionality.

Conclusion

9.1 Summary

Image processing at high speeds is a difficult task for machine vision systems. However, for certain problems that do not require substantial amounts of computation one can implement the processing in the imager itself. Imagers which use signal processing in the focal plane can be substantially faster and consume less power than systems which separate the imager from the signal processor. This thesis investigated the design and implementation of focal plane processors for the particular machine vision problem of automatic alignment.

Automatic alignment of parts or products is an important task in many manufacturing processes. The sensors presented in this work were designed to detect the translational position error of a target constrained to move in a plane. They are for use in mark-to-sensor alignment systems where the target is aligned directly to the sensor. This type of alignment has applications in photolithography, die bonding, and color printing. Much of the published literature concentrates on the photolithography alignment problem in the manufacture of semiconductors. A review of this work reveals a wide variety of alignment techniques including methods based on image moments, edge detection, correlation, diffraction gratings, and moiré fringe patterns. The signal processing for alignment often takes on the order of 10 to 1000 ms to complete and can be an important factor limiting the throughput in these applications.

Two different sensors were designed to explore the potential of focal plane processing for automatic alignment. The first detects the position of a target in the shape of a surveyor's mark. A coarse alignment signal is obtained using quadrant photodiodes. Additional photodiodes perform edge detection to provide a more sensitive fine alignment signal. Alignment is accomplished by using the coarse signal for initial positioning with a large capture region and the fine signal to complete the alignment.

The second sensor design uses a photodiode array arranged to form a square grating in order to detect a square grating alignment target. Alignment is again accomplished by a

combination of coarse initial positioning followed by sensitive final alignment. In this case the coarse alignment signal is generated using absolute value circuits to perform local edge detection. The fine alignment signal is obtained by operating the sensor as a spatial phase detector to locate a geometric moiré fringe. This fringe is formed when the target grating is imaged upon the sensor grating and is a sensitive indicator of their relative positions.

The error signals from both sensors are generated by appropriately summing the currents from the photodiodes. The resulting signal currents are quite small, on the order of nanoamperes, and depend on both the position of the target mark and on the brightness of its illumination. In order to amplify this signal and remove the illumination dependence, normalizing differential current amplifiers were used to boost the signal currents up to a level suitable for off-chip use. These BiCMOS amplifiers contain translinear multipliers for normalization and an active cascode to improve the bandwidth. Test results from the current amplifiers showed that they could amplify differential input currents with gains ranging between 8.2 and 820,000 in order to provide an output common-mode signal of 50 μA . This was accomplished with offsets and nonlinearity of 0.2% and 0.3% of the full scale output signal, respectively.

The two alignment sensors were tested in a mock alignment system. They were placed inside a video camera body which was mounted above a computer controlled x-y table. The table had a four inch travel with a resolution of 10,000 steps/inch. Using this system measurements of the alignment error signal could be made for various positions of the target. Furthermore, closed loop alignments could be accomplished with a simple proportional control algorithm implemented in software.

Tests with the surveyor's mark sensor yielded a coarse alignment signal of 140 $\mu\text{V}/\text{step}$ and a fine alignment signal of 1.1 mV/step . Alignment repeatability, at low speeds, was 2.1 steps (3σ) and was limited by the resolution of the positioning system and the analog-to-digital conversion. This precision is 53 ppm of the total sensor field of view which corresponds to a displacement of 5.3 μm relative to the 10 cm target and a displacement of 0.31 μm across the 6 mm imager. The dominant pole determining the sensor bandwidth was due to the photodiode depletion capacitance reacting with the amplifier input impedance. Because this impedance is inversely proportional to the total photocurrent, the ratio of the sensor irradiance to the depletion capacitance, per unit area, formed the constraint on the bandwidth. The use of an active cascode moved this pole to higher frequency and increased the bandwidth to 7.5 kHz. The higher bandwidth limit comes from some pass transistors included for testing purposes and simulations indicate that removal of these devices would allow bandwidths of 100 kHz to be achieved. Measurements of the sensor noise indicate that it would limit repeatability to between 80 and 1100 ppm, depending upon illumination level, when using the full bandwidth of the sensor.

Test results from the square grating alignment sensor were similar to those from the surveyor's mark sensor. The coarse alignment signal was found to saturate with a sharp transition

at the aligned position due to the single-sided nature of the error signal currents. By adding appropriate bias currents a more reasonable signal could be obtained with a slope of about $200 \mu\text{V}/\text{step}$ near the aligned position. The fine alignment signal from the spatial phase detector had a slope of $6.5 \text{ mV}/\text{step}$. Alignment repeatability was limited by the positioning system to about one positioning step which corresponds to 25 ppm of the sensor field of view. Bandwidth was again found to be proportional to sensor irradiance as discussed above. The fine alignment signal demonstrated a bandwidth of 3 kHz. However, the limitation on sensor speed is determined by the settling time of the absolute value circuits used for the coarse alignment. The coarse alignment bandwidth was estimated to be about 480 Hz for reasonable illumination levels. The measured noise under bright illumination corresponded to a position uncertainty of 30 ppm of the sensor field of view. Because the square grating sensor did not incorporate active cascodes, it had both lower bandwidth and lower noise than the surveyor's mark sensor.

A comparison was made between the two sensors drawing a distinction between single feature and pattern alignment techniques. The contrast was most clearly seen in their fine alignment methods. The edge detection of the surveyor's mark sensor is a single feature method. As such it is sensitive to local defects affecting the edge and is insensitive to broader defects such as nonuniform illumination and magnification error. Its signal strength is relatively weak which implies a fundamentally lower signal-to-noise ratio and bandwidth. However, these limits were not achieved in the actual design. The spatial phase detection used by the square grating sensor is a pattern method. It is insensitive to local pattern defects but it is very sensitive to the broader problems of nonuniform illumination and magnification error. The signal currents are much larger for this sensor design which means that, all else being equal, it would have a higher signal-to-noise ratio and bandwidth. The fine alignment signals from the two sensors were similar in their error vector fields, sensitivity to rotation, and capture regions.

The coarse alignment signals also have some interesting differences. The square grating sensor was found to be able to operate at any rotation, although it works best with no rotation. It also has a capture region that, it is conjectured, is twice the size of the mark. It was only tested over a capture region equal to the mark's size. The surveyor's mark sensor can only tolerate rotations up to 45 degrees and has a capture region that is only half the size of the mark.

9.2 Future Work

There are a number of ways the sensor designs presented in this thesis could be improved. For example, better photodiodes could make a significant difference in performance. The photodiodes used in these sensors had a quantum efficiency of only 27%. By using a well-to-substrate photodiode this could be raised by a factor of three which would correspondingly increase the sensor bandwidth and signal-to-noise ratio at a fixed illumination level. This diode would also

lower the junction depletion capacitance which would improve the speed further.

The surveyor's mark sensor used an active cascode to reduce the effect of the photodiode capacitance and increase the bandwidth. This circuit was also found to add a substantial amount of additional noise. More effort could be spent here to design a higher gain amplifier with lower noise than the simple inverter used currently. Another possibility is to replace the active cascode circuit with a source follower to bootstrap the voltage across the photodiode. This might have advantages over the active cascode although it was not investigated very deeply.

The exact nature of the noise arising from the active cascode is a side issue that raises some questions. Simulations of the noise from this sensor indicated that flicker noise from the inverter would be a prominent component. Measurements of the noise spectrum qualitatively agreed with the simulations but showed significant discrepancies with the flicker noise model used in the simulations. Since this is the dominant noise source it is important to understand the exact nature of the noise in order to minimize it.

More sophisticated designs of the active cascode or its replacement should remove the bandwidth limit associated with the photodiode capacitance. In this case the other poles arising from the current amplifier will become dominant. These also could be pushed out to higher frequency with some design modifications. One simple change would be to modify the layout. The devices in the current amplifier were placed in a common centroid geometry that used four parallel devices to realize each actual circuit device. This reduces the problems associated with device mismatch, but it increases parasitic capacitance due to the extra devices and interconnect. Reducing the symmetry of the layout could reduce the parasitic capacitances by a factor of four or more. The price paid would be an uncertain degradation in the device matching.

Some other implementation improvements are required in the square grating sensor. A major problem was found in the coarse alignment error signal. The nature of the algorithm results in error currents which cause the current amplifier to have a saturated output over most of the range of target motion. Furthermore, near the aligned position the unamplified error currents are quite small which results in very slow circuit behavior. These problems can be easily fixed by adding bias currents to the signal currents. A way to provide these bias currents and to scale them with the overall image illumination level needs to be implemented.

The active cascode circuit was not used on the square grating sensor and should be added in order to realize the bandwidth improvements demonstrated in the surveyor's mark sensor. This will improve the speed of the fine alignment signal by regulating the voltage across the photodiodes. However, improving the speed of the coarse alignment signal is more difficult. In this case the modifications have to be made in the pixel where the absolute value of difference circuit is located. Possibilities include adding a simple inverter cascode to this circuitry or using a folding amplifier to add both gain and a current-to-voltage conversion in the pixel. In any case, the circuit used to implement the edge detection for the coarse alignment signal should be examined carefully for ways to improve its speed. As the design now stands the absolute

value circuit will limit the sensor speed to only a few times better than video rate.

Some additional features could be added to the sensors that might make them easier to use. For example, the chip output might be made voltage rather than current. Presently, the error currents are converted to voltages using off-chip circuitry. If voltage is the preferred output, then this conversion could be integrated onto the sensor. Similarly, one might also prefer a digital output for the position. One could integrate an analog-to-digital converter with this sensor if desired. Finally, for truly high speed applications it might be useful to incorporate circuitry onto the sensor that could drive positioners directly. This would restrict the use of the sensor to a particular positioning mechanism, but it might be worthwhile for some applications.

9.3 Contribution

This thesis has shown how focal plane processing could be used to build better sensors for automatic alignment. Certain kinds of computations lend themselves to implementation in the focal plane. The computation done in these sensors occurs partly in the physical layout of the photodiodes and how it interacts with the pattern of the alignment mark. The addition of signals from various regions of the image plane is then easily accomplished by summing photocurrents. The remaining computation is the normalization and amplification of the error signal by the current amplifiers on the periphery of the imager.

This work has also tried to contrast the various alignment techniques by implementing two sensors with different alignment philosophies. A rough division was made between alignment schemes that concentrate on a few features in a target mark and those that make use of a large number of features. Some of the trade-offs between these approaches are illustrated by the two sensor designs and their sensitivity to nonuniform illumination and defects in the alignment marks.

The two sensors that were designed and tested both operate effectively in the mock alignment system. They have demonstrated the capability to guide accurate alignments and can provide error signals at speeds substantially above video rate. They also consume relatively little power and are compact compared to solutions that separate the imager from the signal processing. Their only major drawbacks are the lack of flexibility in the alignment mark and algorithm as well as the time and expense required to design such special purpose sensors.

Focal plane processing is an appealing idea for realizing high performance machine vision systems. However, its successful application is limited to problems that satisfy some restrictive conditions. First, the required signal processing must be simple enough to be realized with compact circuits that can be integrated with the imager. Second, the desired image information must be required at high speed and low latency and/or with very little power consumption. These are the chief advantages of a focal plane processing system. A related constraint often required for high speed operation is that the signal processing on the imager reduce the amount

of information present in the image that must be transmitted off-chip. The problem of automatic alignment meets all of these conditions. It requires elementary signal processing at high speed and reduces the acquired image to a pair of position error signals. The sensor designs presented in this thesis demonstrate how focal plane processing can be applied to this problem in order to provide a better tool for industrial applications.

Bibliography

- [1] David Schatz. *Next-Generation Machine Vision Technology Enables Faster, More Accurate Automatic Wafer Alignment*. Cognex Corporation, Needham, MA. Reprinted from *Microelectronic Manufacturing and Testing*, July 1987.
- [2] Nikon Precision Inc. San Bruno, CA, Sales brochure.
- [3] J. T. M. Stevenson and J. R. Jordan. Use of gratings and periodic structures as alignment targets on wafer steppers. *Precision Engineering*, 11(2):63–69, April 1989.
- [4] C. B. Bose and I. Amir. Design of fiducials for accurate registration using machine vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(12):1196–1200, December 1990.
- [5] B. W. Crowley. Application of one-dimensional machine vision in the textile industry. *IEEE Transactions on Industry Applications*, 26(2):324–329, March/April 1990.
- [6] A. Hachicha and S. Simon. Subpixel edge detection for precise measurements by a vision system. In *Industrial Inspection*, volume 1010, pages 148–157. SPIE, 1988.
- [7] B. Kleinemeier. Measurement of CCD interpolation functions in the subpixel precision range. In *Industrial Inspection*, volume 1010, pages 158–165. SPIE, 1988.
- [8] Q. Tian and M. N. Huhns. Algorithms for subpixel registration. *Computer Vision, Graphics, and Image Processing*, 35:220–233, 1986.
- [9] S. C. Douglas. A frequency-domain subpixel position estimation algorithm for overlay measurement. In *Integrated Circuit Metrology, Inspection, and Process Control VII*, volume 1926, pages 402–411. SPIE, 1993.
- [10] J. L. Wyatt et al. Analog VLSI systems for early vision processing. In *Proceedings of the 1992 IEEE International Symposium on Circuits and Systems*, pages 1644–1647, 1992.
- [11] E. A. Vittoz. Future of analog in the VLSI environment. In *Proceedings of the 1990 IEEE International Symposium on Circuits and Systems*, volume 2, pages 1372–1375, 1990.
- [12] B. J. Hosticka. Performance comparison of analog and digital circuits. *Proceedings of the IEEE*, 73(1):25–29, January 1985.
- [13] P. Burggraaf. Overlay: Lithography’s big challenge. *Semiconductor International*, pages 58–61, February 1991.

- [14] J. M. Lavine, R. B. Fish, A. J. Weger, and C. Simpson. A direct-reticle-referenced alignment system. In *Optical Microlithography IV*, volume 538, pages 57–69. SPIE, 1985.
- [15] M. v. d. Brink, H. F. D. Linders, and S. Wittekoek. Direct-referencing automatic two-points reticle-to-wafer alignment using a projection column servo system. In *Optical Microlithography V*, volume 633, pages 60–71. SPIE, 1986.
- [16] N. Akiyama and Y. Yamauchi. Wafer alignment mark position-detecting technique for step-and-repeat projection aligners. *Bull. Japan Soc. of Prec. Eng.*, 23(3):200–204, September 1989.
- [17] S. McKaughan. Real-time automatic inspection of web materials. In *International Conference on Automatic Inspection and Measurement*, volume 557, pages 140–145. SPIE, 1985.
- [18] J. Holmes. Fast photonics at the film factory. *Photonics Spectra*, 28(12):137–141, December 1994.
- [19] B. K. P. Horn. A problem in computer vision: Orienting silicon integrated circuit chips for lead bonding. *Computer Graphics and Image Processing*, 4:294–303, 1975.
- [20] A. Casals, J. Amat, and A. B. Martinez. High speed processors for autonomous vehicles guidance. In *Proceedings of the IEEE International Workshop on Intelligent Motion Control*, pages IP-45 — IP-53, 1990.
- [21] J. T. M. Stevenson and J. R. Jordan. Absolute position measurement using optical detection of coded patterns. *J. Phys. E: Sci. Instrum.*, 21:1140–1145, 1988.
- [22] B. K. P. Horn. *Robot Vision*. MIT Press, Cambridge, MA, 1986.
- [23] D. J. W. Noorlag and S. Middelhoek. Two-dimensional position-sensitive photodetector with high linearity made with standard I.C.-technology. *Solid-State and Electron Devices*, 3(3):75–82, May 1979.
- [24] G. P. Petersson and L. Lindholm. Position sensitive light detectors with high linearity. *IEEE Journal of Solid-State Circuits*, SC-13(3):392–399, June 1978.
- [25] SELSPOT INC, 1233 Chicago Road, Troy, MI 48083. *Selspot II — A Complete System for Sophisticated Motion Analysis*. Sales brochure.
- [26] S. P. DeWeerth and C. A. Mead. A two-dimensional visual tracking array. In *Advanced Research in VLSI: Proceedings of the 1988 MIT Conference*, pages 259–275, Cambridge, Massachusetts, 1988. MIT Press.
- [27] D. L. Standley. An object position and orientation IC with embedded imager. *IEEE Journal of Solid-State Circuits*, 26(12):1853–1859, December 1991.
- [28] B. K. P. Horn. Parallel networks for machine vision. In P. H. Winston and S. A. Shellard, editors, *Artificial Intelligence at MIT: Expanding Frontiers*, volume 2, chapter 43, pages 530–573. MIT Press, Cambridge, MA, 1990.
- [29] R. E. Hughlett and K. A. Cooper. A video based alignment system for x-ray lithography. In *Electron-Beam, X-Ray, and Ion-Beam Submicrometer Lithographies for Manufacturing*, volume 1465, pages 100–110. SPIE, 1991.

- [30] E. Cullmann, K.A. Cooper, and W. Vach. An x-ray stepper for production lithography. In *Electron-Beam, X-Ray, and Ion-Beam Lithographies VI*, volume 773, pages 2–6. SPIE, 1987.
- [31] J. Lauria and T. Kerekes. Automatic wafer alignment technique for the 700SLR / 800SLR wafer steppers. In *Microcircuit Engineering 80*, pages 171–179. Delft University Press, 1981.
- [32] Cannon Inc., Kawasaki, Japan. *FPA-1550 M III W Reticle Design Guide*, 1989. Manual.
- [33] A. Suzuki. Double telecentric wafer stepper using laser scanning method. In *Optical Microlithography IV*, volume 538, pages 2–8. SPIE, 1985.
- [34] SVG Lithography Systems, Inc., 77 Danbury Road, Wilton CT 06897. *Micralign 600 HT Series Projection Mask Alignment Systems*, May 1985. Sales brochure.
- [35] H. E. Mayer and E. W. Loebach. An advanced wafer stepper for sub-micron fabrication. In *Optical Microlithographic Technology for Integrated Circuit Fabrication and Inspection*, volume 811, pages 149–159. SPIE, 1987.
- [36] H. E. Mayer and E. W. Loebach. Design and performance of a new step-and-repeat aligner. In *Microcircuit Engineering 80*, pages 191–209. Delft University Press, 1981.
- [37] T. A. Brunner and S. D. Smith. Moiré technique for overlay metrology. In *Integrated Circuit Metrology II*, volume 480, pages 164–170. SPIE, 1984.
- [38] D. R. Cote, R. H. Clayton, and J. E. Lazo-Wasem. Advanced optical overlay measurement system. In *Optical Microlithography VI*, volume 772, pages 124–133. SPIE, 1987.
- [39] C. P. Ausschnitt, T. A. Brunner, and D. J. Cronin. Overlay performance of the Perkin-Elmer Model 500. In *Optical Microlithography II: Technology for the 1980s*, volume 394, pages 64–72. SPIE, 1983.
- [40] J. Dey and K. Johannsmeier. An advanced step and repeat aligner for VLSI. In *Microcircuit Engineering 80*, pages 211–221. Delft University Press, 1981.
- [41] T. Soesanto and K. Harrison. Optimization of Optimetrix wafer steppers in very large scale integrated circuit (VLSI) process. In *Optical Microlithography III: Technology for the Next Decade*, volume 470, pages 92–103. SPIE, 1984.
- [42] S. M. Preston, D. W. Peters, and D. N. Tomes. Preliminary testing results for a new x-ray stepper. In *Electron-Beam, X-Ray, and Ion-Beam Technology: Submicrometer Lithographies VIII*, volume 1089, pages 164–177. SPIE, 1989.
- [43] D. J. Elliot, R. Fish, N. Gottier, M. Mason, and C. Sparkes. Advances in digital alignment strategies for step-and-repeat optical imaging. In *Microcircuit Engineering 85*, pages 261–265. North-Holland, 1985.
- [44] N. Bobroff, R. Tibbetts, J. Wilczynski, and A. Wilson. An optical alignment microscope for x-ray lithography. *J. Vac. Sci. Technol. B*, 4(1):285–289, Jan./Feb. 1986.
- [45] H. Sakou, T. Miyatake, S. Kashioka, and M. Ejiri. A position recognition algorithm for semiconductor alignment based on structural pattern matching. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(12):2148–2157, December 1989.

- [46] Cognex Co., 15 Crawford Street, Needham, MA 02194. *COGNEX Vision for Industry: Board-Level Machine Vision Systems Technical Description*, 1990. Sales literature.
- [47] LSI Logic Corp., 1551 McCarthy Blvd., Milpitas CA 95035. *Digital Signal Processing (DSP) Databook*, June 1990.
- [48] H. S. Ranganath. Hardware implementation of image registration algorithms. *Image and Vision Computing*, 4(3):151–158, August 1986.
- [49] R. S. Hershel. Autoalignment in step-and-repeat wafer printing. In *Developments in Semiconductor Microlithography IV*, volume 174, pages 54–62. SPIE, 1979.
- [50] J. S. Wilczynski. Optical step and repeat camera with dark field automatic alignment. *Journal of Vacuum Science and Technology*, 16(6):1929–1933, November/December 1979.
- [51] D. R. Beaulieu and P. P. Hellebrekers. Dark field technology — a practical approach to local alignment. In *Optical Microlithography VI*, volume 772, pages 142–149. SPIE, 1987.
- [52] M. A. Butler, K. B. Pfeifer, and S. J. Martin. Two-dimensional patterns for optical alignment. *Applied Optics*, 30(32):4600–4601, 1991.
- [53] S. Murakami, T. Matsuura, M. Ogawa, and M. Uehara. Laser step alignment for a wafer stepper. In *Optical Microlithography IV*, volume 538, pages 9–16. SPIE, 1985.
- [54] S. Slonaker et al. Enhanced global alignment for production optical lithography. In *Optical/Laser Microlithography*, volume 922, pages 73–81. SPIE, 1988.
- [55] N. Farrar and F. Sporon-Fiedler. Performance of through-the-lens/off-axis alignment systems and alignment algorithms on Nikon wafer steppers. In *Integrated Circuit Metrology, Inspection, and Process Control VI*, pages 369–380. SPIE, 1992.
- [56] Nikon Precision Inc., 1176 Cherry Avenue, San Bruno, CA 94066. *Nikon: The Next Generation NSR*, 1988. Sales brochure.
- [57] H. P. Kleinknecht. Diffraction gratings as keys for automatic alignment in proximity and projection printing. In *Developments in Semiconductor Microlithography IV*, volume 174, pages 63–69. SPIE, 1979.
- [58] G. Dubroeuq, M. Lacomat, and D. Sullerot. Double diffraction gratings as keys for high contrast alignment on wafer steppers. In *Microcircuit Engineering 80*, pages 181–190. Delft University Press, 1981.
- [59] G. Makosch and F. Prein. Phase-locked interferometry for automatic mask alignment in projection printers. *Applied Optics*, 26(14):2828–2835, July 15 1987.
- [60] B. Fay, J. Trotel, and A. Frichet. Optical alignment system for submicron x-ray lithography. *Journal of Vacuum Science and Technology*, 16(6):1954–1958, 1979.
- [61] B. Fay. Registration and distortion compensating techniques for a large field x-ray exposure system. In *Electron-Beam, X-Ray, and Ion-Beam Techniques for Submicron Lithographies II*, volume 393, pages 198–204. SPIE, 1983.

- [62] B. Fay, W. T. Novak, and I. Carlsson. Automatic x-ray alignment system for submicron VLSI printing. In *Electron-Beam, X-Ray, and Ion-Beam Techniques for Submicrometer Lithographies III*, volume 471, pages 90–95. SPIE, 1984.
- [63] B. S. Fay and W. T. Novak. Advanced x-ray alignment system. In *Electron-Beam, X-Ray, and Ion-Beam Techniques for Submicrometer Lithographies V*, volume 632, pages 146–155. SPIE, 1986.
- [64] R. B. McIntosh, G. P. Hughes, J. L. Kreuzer, and G. R. Conti Jr. X-ray step-and-repeat lithography system for submicron VLSI. In *Electron-Beam, X-Ray, and Ion-Beam Techniques for Submicrometer Lithographies V*, volume 632, pages 156–163. SPIE, 1986.
- [65] G. Chen et al. Implementation of two-state alignment system into CXrL aligner. In *Electron-Beam, X-ray, and Ion-Beam Submicrometer Lithographies for Manufacturing II*, volume 1671, pages 471–485. SPIE, 1992.
- [66] M. Feldman, A. D. White, and D. L. White. Application of zone plates to alignment in microlithography. *Journal of Vacuum Science and Technology*, 19(4):1224–1228, November/December 1981.
- [67] M. Feldman, G. G. Wong, and M. Cheng. High speed zone plate alignment systems. *J. Vac. Sci. Technol. B*, 5(1):241–243, Jan./Feb. 1987.
- [68] G. Oster and Y. Nishijima. Moiré patterns. *Scientific American*, 208(5):54–63, May 1963.
- [69] G. Oster, M. Wasserman, and C. Zwerling. Theoretical interpretation of moiré patterns. *Journal of the Optical Society of America*, 54(2):169–175, February 1964.
- [70] L. O. Vargady. Moiré fringes as visual position indicators. *Applied Optics*, 3(5):631–636, May 1964.
- [71] K. Paturski and M. Kujawińska. *The Moiré Fringe Technique*. Elsevier, New York, NY, 1993.
- [72] Y. Arai and T. Kurata. Analysis of moiré fringes produced by centrosymmetric gratings. *Journal of Modern Optics*, 34(8):1067–1075, 1987.
- [73] M. C. King and D. H. Berry. Photolithographic mask alignment using moiré techniques. *Applied Optics*, 11(11):2455–2459, November 1972.
- [74] T. M. Lyszczarz, D. C. Flanders, N. P. Economou, and P. D. DeGraff. Experimental evaluation of interferometric alignment techniques for multiple mask registration. *Journal of Vacuum Science and Technology*, 19(4):1214–1218, November/December 1981.
- [75] T. Allen, C. Mead, F. Faggin, and G. Gribble. Orientation-selective VLSI retina. In *Visual Communications and Image Processing '88*, volume 1001, pages 1040–1046. SPIE, 1988.
- [76] J. Guild. *The Interference Systems of Crossed Diffraction Gratings*. Oxford University Press, London, UK, 1956.
- [77] J. Guild. *Diffraction Gratings as Measuring Scales*. Oxford University Press, London, UK, 1960.

- [78] D. C. Flanders, H. I. Smith, and S. Austin. A new interferometric alignment technique. *Applied Physics Letters*, 31(7):426–428, Oct. 1977.
- [79] S. Ishihara. High precision positioning for submicron lithography. *Bull. Japan Soc. of Prec. Eng.*, 21(1):1–8, Mar. 1987.
- [80] J. Itoh and T. Kanayama. A new interferometric displacement-detection method for mask-to-wafer alignment using symmetrically-arranged three gratings. *Japanese Journal of Applied Physics*, 25(6):L487–L489, June 1986.
- [81] J. Itoh, T. Kanayama, N. Atoda, and K. Hoh. A new mask-to-wafer alignment technique for synchrotron radiation x-ray lithography. In *Electron-Beam, X-Ray, and Ion-Beam Lithographies VI*, volume 773, pages 7–14. SPIE, 1987.
- [82] T. Nomura, S. Kimura, Y. Uchida, and S. Hattori. Moiré alignment technique for the mix and match lithographic system. *J. Vac. Sci. Technol. B*, 6(1):394–398, Jan./Feb. 1988.
- [83] V. T. Chitnis, Y. Uchida, K. Hane, and S. Hattori. Moiré signals in reflection. *Optics Communications*, 54(4):207–211, June 15 1985.
- [84] Y. Uchida, S. Hattori, and T. Nomura. An automatic mask alignment technique using moiré interference. *J. Vac. Sci. Technol. B*, 5(1):244–247, Jan./Feb. 1987.
- [85] T. Nomura, K. Yamashita, Y. Yamada, M. Suzuki, and T. Takemoto. Interferometric nanometer alignment for a wafer stepper by two wave-front reconstruction onto deformed wafer gratings. *Japanese Journal of Applied Physics*, 26(6):959–964, June 1987.
- [86] N. Uchida, Y. Ishibashi, R. Hirano, N. Kikuri, and M. Tabata. A dual grating alignment method insensitive to mask-wafer gap variation. *Bull. Japan Soc. of Prec. Eng.*, 23(2):140–145, June 1989.
- [87] G. Bouwhuis and S. Wittekoek. Automatic alignment system for optical projection printing. *IEEE Transactions on Electron Devices*, ED-26(4):723–728, 1979.
- [88] S. Wittekoek, H. Linders, and H. Stover. Precision wafer-stepper alignment and metrology using diffraction gratings and laser interferometry. In *Micron and Submicron Integrated Circuit Metrology*, volume 565, pages 22–31. SPIE, 1985.
- [89] S. Wittekoek and J. v. d. Werf. Phase gratings as wafer stepper alignment marks for all process layers. In *Optical Microlithography IV*, volume 538, pages 24–31. SPIE, 1985.
- [90] M. A. v. d. Brink, S. Wittekoek, H. F. D. Linders, F. J. v. Hout, and R. A. George. Performance of a wafer stepper with automatic intra-die registration correction. In *Optical Microlithography VI*, volume 772, pages 100–117. SPIE, 1987.
- [91] M. v. d. Brink, H. Franken, S. Wittekoek, and T. Fahner. Automatic on-line wafer stepper calibration system. In *Integrated Circuit Metrology, Inspection, and Process Control IV*, volume 1261, pages 289–314. SPIE, 1990.
- [92] F. J. v. Hout, M. A. v. d. Brink, S. Wittekoek, and J. F. Coolen. Registration accuracy and critical dimension control for a 5x reduction stepper with magnification control. In *Optical Microlithographic Technology for Integrated Circuit Fabrication and Inspection*, volume 811, pages 118–129. SPIE, 1987.

- [93] W. R. Trutna and M. Chen. An improved alignment system for wafer steppers. In *Optical Microlithography III: Technology for the Next Decade*, volume 470, pages 62–69. SPIE, 1984.
- [94] J. T. M. Stevenson and J. R. Jordan. Dynamic position measurement technique for flash-on-the-fly wafer exposure. *Precision Engineering*, 11(3):127–133, July 1989.
- [95] A. Moel, E. E. Moon, R. D. Frankel, and H. I. Smith. Novel on-axis interferometric alignment method with sub-10 nm precision. *J. Vac. Sci. Technol. B.*, 11(6):2191–2194, Nov./Dec. 1993.
- [96] E. Hecht and A. Zajac. *Optics*. Addison-Wesley, Reading, MA, 1974.
- [97] B. Gilbert. A new wide-band amplifier technique. *IEEE Journal of Solid-State Circuits*, SC-3(4):353–365, December 1968.
- [98] B. Gilbert. A high-performance monolithic multiplier using active feedback. *IEEE Journal of Solid-State Circuits*, SC-9(6):364–373, December 1974.
- [99] B. Gilbert. Translinear circuits. Unpublished paper, May 1981.
- [100] C. Toumazou, F. J. Lidgley, and D. G. Haigh. *Analogue IC Design: The Current-Mode Approach*. Peter Peregrinus Ltd., Stevenage, United Kingdom, 1990.
- [101] B. Gilbert. A monolithic 16-channel analog array normalizer. *IEEE Journal of Solid-State Circuits*, SC-19(6):956–963, December 1984.
- [102] P. R. Gray and R. G. Meyer. *Analysis and Design of Analog Integrated Circuits*. John Wiley and Sons, New York, NY, third edition, 1993.
- [103] R. Unbehauen and A. Cichocki. *MOS Switched-Capacitor and Continuous-Time Integrated Circuits and Systems: Analysis and Design*. Springer-Verlag, New York, NY, 1989.
- [104] Z. Y. Chang and W. M. C. Sansen. *Low-Noise Wide-Band Amplifiers in Bipolar and CMOS Technologies*. Kluwer Academic Publishers, 1991.
- [105] Z. Y. Chang and W. Sansen. Test structure for evaluation of $1/f$ noise in CMOS technologies. In *ICMTS*, pages 143–146, 1989.
- [106] B. J. Gross. *$1/f$ Noise in MOSFETs with Ultrathin Gate Dielectrics*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, June 1992.
- [107] C. Tomovich. MOSIS — A gateway to silicon. *IEEE Circuits and Devices Magazine*, 4(2):22–23, March 1988.
- [108] A. V. Oppenheim, A. S. Willsky, and I. T. Young. *Signals and Systems*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1983.
- [109] F. Forti and M. E. Wright. Measurement of MOS current mirror mismatch in the weak inversion region. *IEEE Journal of Solid-State Circuits*, 29(2):138–142, February 1994.
- [110] J. K. Roberge. *Operational Amplifiers: Theory and Practice*. John Wiley and Sons, New York, NY, 1975.

- [111] C. J. Aswell, J. Berlien, E. Dierschke, and M. Hassan. A monolithic light-to-frequency converter with a scalable sensor array. In *Technical Digest: IEEE International Solid State Circuits Conference*, pages 158–159, 1994.
- [112] P. Aubert and H. Oguey. An application specific integrated circuit (ASIC) with CMOS-compatible light sensors for an optical position encoder. In *Proceedings of the IEEE 1987 Custom Integrated Circuits Conference*, pages 712–716, 1987.
- [113] P. Aubert, H. J. Oguey, and R. Vuilleumier. Monolithic optical position encoder with on-chip photodiodes. *IEEE Journal of Solid-State Circuits*, 23(2):465–473, April 1988.
- [114] C. A. Mead. *Analog VLSI and Neural Systems*. Addison–Wesley, Reading, Massachusetts, 1989.