(21)

A System Level Design for Object Location and Identification In Unstructured Environments

Rolland L. Doubleday, Jr.

Bachelor of Engineering Science, Systems and Control
Montana College of Mineral Science and Technology, 1994

Master of Science in Mechanical Engineering
Massachusetts Institute of Technology, 1995

Submitted to the Department of Mechanical Engineering in Partial Fulfillment of the Requirements for the Degree of

Mechanical Engineer

at the
Massachusetts Institute of Technology

June 1998

Signature of Author: _____
Department of Mechanical Engineering
May 5, 1998

Certified by: _____
Haruhiko H. Asada
Ford Professor of Mechanical Engineering
Thesis Supervisor

Accepted by: _____
Ain A. Sonin
Professor of Mechanical Engineering
Chairman, Committee for Graduate Students

A System Level Design for Object Location and Identification In Unstructured Environments

by

Rolland L. Doubleday, Jr.

Submitted to the Department of Mechanical Engineering
on May 22, 1998 in Partial Fulfillment of the
Requirements for the Degree of Mechanical Engineer in
Mechanical Engineering

Abstract

A system was developed that uses an IR laser diode to fluoress an ink that absorbs at 785 nm and has emission at 830 nm, and a CCD camera is used to capture this low strength ink emission. There are many things that have emission in this band, and for this reason, the signal-to-noise ratio is very small. Robust methods were developed to deal with the noise in the viewable frame. The Perturbation/Correlation method is the most noteworthy method, and uses a sine wave superimposed on a DC value for laser intensity, and then monitors the change in intensity of the returned signal over this perturbation in laser intensity. These changes in viewable frame intensity are then correlated to the known perturbation in laser intensity, giving a large enhancement in the signal to noise ratio.

With the development of noise reduction methods that would result in a greater S/N ratio at any given laser power, the next step was to classify the noise in the viewable frame. The characteristics of the noise were investigated to determine when and how the signal processing techniques should be applied, as well as to give a general classification of these sources of noise for later intelligent scan planning.

Thesis Supervisor: Haruhiko H. Asada
Title: Ford Professor of Mechanical Engineering

Acknowledgements

After spending four years in graduate school at MIT, it is really hard to begin this section. There are so many people that feel I should say thank you to, both within the MIT community as well as outside of it. Let me start with the people outside of the community, since they are very much responsible for getting me here.

Again, I will start with my Dad. My Dad is the sharpest man I know, hands down. Of course, I don't get out much (I'm joking). He is a self-taught man, and I imagine he has learned most of what he knows from a genuine eagerness to learn. He is the kind of guy that can enter a room, solve almost any problem you may be struggling with, tell a few jokes, and then leave again before you have had a chance to feel stupid. I really try to pattern myself after him; of course, I only have the joke part down so far. If I posses any good qualities, I would like to think that I get them from him. I must also thank him for always accepting those collect calls from me when I had too many beers and just wanted to call and wake him up and ask him "you weren't asleep, were you?" Orneriness is another quality that I probably got from him.

Another person that is very much like my Dad in his humble treatment of his own excellence is Neil Wahl, my undergraduate advisor at Montana Tech. Like my Dad, he will dive into a troubleshooting nightmare and not surface until things are working properly. Neil is a very soft-spoken guy who will go the extra mile to actually "teach" somebody something. One day in Montana, after spending about six hours helping me troubleshoot a controller of some sort, he mentioned that he was going to get something to eat, and in fact it was his first food in several days. I asked him if he had been ill, but he said no, he just hadn't had time to eat. I thought maybe he was kidding at the time, but I quickly realized as I got to know him better that his dedication to his work ran deep. One of my deepest regrets as I write this is that my Dad and Neil never met when I was still in Montana.

Since I don't want to make my acknowledgements section larger than the body of my thesis, I will greatly reduce my mention of the important people in my life; this in no way reduces their importance to me. I would like to thank Kari A. Kulaszewicz, this blonde-Polish-Finnish bombshell that I may be marrying in a few weeks. Thank you for loving this engineer, even though he works too much and doesn't spend enough time with you, and probably almost always says the wrong thing at any particular moment in time. You are probably the only person that I know that may be more stubborn and hardheaded than I am. Your sense of humor is probably your most valuable asset, and never fails to make me laugh, even when life is tough; and life can be tough at MIT. Your spontaneity is sometimes overwhelming for this stick-in-the-mud, conservative engineer, but if it wasn't for this, I probably wouldn't have done anything but work while at MIT. I am not going to say anything truly mushy about you, since this document will be read by thousands of people, but believe me, I have many mushy things to say about you if so many people were not going to read this document. I have also learned a bit of cynicism during my stay here at MIT.

I would like to thank Tim Hogan and Mark James for being true friends, even though we haven't been around one another for almost four years. In particular, I would like to thank Tim for

always letting me know when I should "cut it away like a bad chute." This philosophy can be applied to almost any aspect of one's life. I would like to thank Mark for staying in school for so long; I don't feel quite so bad for spending nine years in school, now. As Tim would say about both Mark and I, there is a name for people like us that spend this much time in school: they are called neurosurgeons.

I would also like to thank Conrad Hilpert, Joe Kujawa, David Westine, Baily Hines, Dana Day, and Uncle Donnie (Don Day) for a combined effort of encouragement to me. Uncle Donnie is one of the funniest guys in our family, but that doesn't take away from the fact that he takes extreme pride in his work and was born with a big heart. Uncle Donnie, I must apologize for some of the things that Dana and I did over the years while in the presence of your company. Some things come to mind, such as the boat horn being honked at 4 am, the margaritas that Dana and I blended right above your bedroom as you desperately tried to sleep, and accidentally running into your brand new truck with an off-road vehicle. These are things that would make me feel guilty if they didn't make me laugh so much.

Conrad Hilpert, who is an old professor of mine back in Montana, was always very encouraging to his students. The fact is that he fought in WWII as an army officer, and when I say that he was always very encouraging, that was after he got to know you. You had to earn whatever you got in his class in the way of respect, but once you did, there wasn't enough good things he could say about you or do for you. After suffering a heart attack last year, he made a tape to apologize to his class for not being able to show up and teach. Even though he is nearing eighty, he still travels all over the country with his wife on their BMW motorcycles. He has a truly impressive resume of his fifty or so years in engineering, and is one of the only true design guys that I have met.

Baily Hines has always been somebody I looked up to. I guess you would have to know him to truly appreciate him. He always kept me working during my summers, which gave me a chance to work a set of tools and make friends with people that normally didn't like engineers all that much. I remember one time after a couple of us had finished a very tough rigging job, an engineer came out to inspect the installation of his design. We were all fairly sweaty and grimy, and pretty pumped from manhandling the rigging of this heavy piece of pipe (it was several thousand pounds). The engineer was fairly out of shape and was wearing a spotlessly clean white dress shirt, which of course was stuffed with pencils. He stood around awkwardly for several minutes, looking at this and looking at that, and then he did something very profound: He left us to do our job. I guess from that point on I realized the type of engineer that I did not want to be. These and other lessons I owe to Baily.

Joe Kujawa has been a friend for many years. We met in Montana during one of the many control classes that I took. Joe has always had natural talent in electronic design. If I were to make a comparison between Joe and a TV character, I guess I would have to say that he most closely resembles McGyver; I mean, as far as his ingenuity goes. Joe is a quiet, unassuming guy to most people, but when you get to know him, he is a true prince. He has the mentality of a true inventor in that he is always working on some new design and in the end, as far as I know has a 100% success rate at making it work, but will probably never have any patents since these things do not interest him. One of the most interesting things about Joe is that he has a liberal

arts background (I believe he was an English major), and most of his knowledge of electronics was self-taught. As he is reading this, he is probably looking at my extremely long sentences and shaking his head; Joe I can't help it, I'm an engineer.

Dave Westine is another professor of mine back at Montana Tech. He is an extremely nice guy, and would never turn me away when I had a question in one of his classes. We received some really bad news about Dr. Westine recently, but I hope by the time he reads this thesis he will be all better. Of course, this thesis may push anybody over the edge.

Alright, now that I have gone through the people out side of the MIT community, I will talk about the ones within MIT. I would like to thank Professor Asada for being my advisor for several years. Under his serious demeanor there is a very nice guy with a great sense of humor, although sometimes you forget these details when you have serious business to talk about. He has always encouraged lab unity in a number of ways, and always spends money so that the lab as a whole may go out and do fun things together. This is not always easy, considering the wide range of student personalities that exist in lab. He has always been a good sport, as can be seen in the pictures we have of him astride a horse on one of our outings in Vermont.

When I came to this lab, I was deathly afraid of giving presentations. Professor Asada broke me of this almost entirely, and now I am able to speak with a certain degree of confidence on almost any subject I am familiar with. Sometimes, for example, he would take a guy that knew nothing of electronics but new quite a bit of dynamics, strengths of materials, etc., and give this new guy an electronics project to complete before starting his mechanic's research. Another guy who knew quite a bit of electronics but less about some other aspect of engineering would be given a short project that dealt with his weakness. This approach worked well in two ways: First, the person would gain some experience in the area of his weakness; Second, lab communication would be enhanced when the person who knew nothing of the particular subject would seek out the person in the lab that did know something about this, and make friends with him.

The one thing that I can definitely say about Professor Asada is that no matter what happens during the weekly meetings, whether they are good or bad, he always backs his students in his representation of them to the people outside the lab. Some of the nice things he has said about me to the Mechanical Engineering faculty on two separate occasions have gotten back to me, and I would honestly like to thank him again for this. Professor Asada, what do these people know?

Also within this community, I must really thank Dr. Colin Brenan. Colin is a post-doc working in another lab, and is probably one of the sharpest guys I have met at MIT when it comes to optics, lasers, spectroscopy, image processing and vision based systems. Colin should sit on top of a very tall mountain and make seekers of knowledge climb that mountain to obtain this knowledge. Really, he should be one of the very first few that I acknowledge in this section since he has contributed so many good ideas in this research work. On top of all of this, Colin is a very nice guy and is always willing to openly discuss his research with you or vice versa, and on top of that, there is this rumor that he liked to drink beer when he was an undergraduate, so I guess that makes him alright in my book.

Within my own lab, I would like to thank some of the guys: Brandon Gordon, Bei Gu, Shrikant Savant, Muralhi Ravuri, Masa Wada, and Joe Spano, as well as Bernardo Aumond, Bill Finger, Elie Awad and Dan Schnidman. You guys have been a humorous influence in a serious place. In fact, I shouldn't leave out the other guys, but hey, you guys were more serious.

Jim Thompson has been an exceptionally nice guy to me as well as to most of the people he deals with in this department. He has a drive that is very high even by MIT standards, and has bared up well under the tough times of his recent past. I hope things will go very well for him in the future, since if anyone deserves this, he does.

Professor Sonin, if by chance you read this, I would like to thank you for your words of encouragement over the years. I enjoyed both semesters of Advanced Fluid Mechanics (a little joke), and I regret that I didn't come by your office more often. You are an extremely nice guy, and in case you didn't know it, everybody likes you.

I have saved the best for last. I would like to thank Leslie (Ms. Regan) and Joan of the Mechanical Engineering graduate office. Over my four years in grad school here, I have experienced almost every possible experience one may experience while in grad school here. Whether it was good or bad, they always were very willing to listen. They have always watched out for me, as best they could, and always lightened my moods when they were dark. I hope the many of thousands of people that read this thesis do not realize what a gold mine there is the graduate office, since they might not be able to handle thousands of people at once. One of the things I will miss most about MIT is being able to come into your office and just talk and listen to your stories and jokes.

Looking back at this acknowledgement section, I realize I should have probably used a table of contents and broken it into several chapters. Anyhow, it might be my last time to say some nice things about the people around me, and I didn't want to lose the opportunity. Take it easy.

Table of Contents

List of Figures

List of Figures (continued)

Chapter 1: Introduction

One of the main obstacles to home automation, such as home robotics, is the inability to locate and identify thousands of objects in this highly unstructured environment. In a well-defined environment, such as a laboratory, pattern recognition techniques have been used to identify and locate a few items within a small area of search. However, in the home environment, this stringent definition of environment does not exist. For this reason, a totally non-intrusive system was developed to locate and identify objects without any requirements placed on item orientation or search area.

1.1 Background and Motivation

Since there would be many thousands of items to identify within a large search area, emphasis was placed on computational efficiency for all applicable image processing. Since it was to be used in the home, the system had to be non-intrusive as well as meet ANSI standards on eye and skin safety. Also, steps were taken to insure that a good item identification had taken place, since mistakes at this level could prove costly to home occupants.

To illustrate the motivation of this project, Figure 1.10 depicts a typical kitchen table. The items on the table have a variety of shapes, sizes, and orientations. Some have clear surfaces, and others have opaque surfaces. The point of grip is different on each item. It would be very desirable to have a system that could easily identify the objects and locate them in three space, and from a database, pull information such as point of grip and special handling instructions.

With consideration of the home environment, it can be seen that the system that was to be used had to be able to locate and identify thousands of items quickly and accurately, and do this in a safe, non-intrusive manner. The objects would have to be identified at distances up to the maximum dimension of the room of placement. Many different types of systems were considered that might be able to locate and identify objects. A Pough chart was used to rate these different approaches based on our criteria (See Figure A1 in Appendix A).

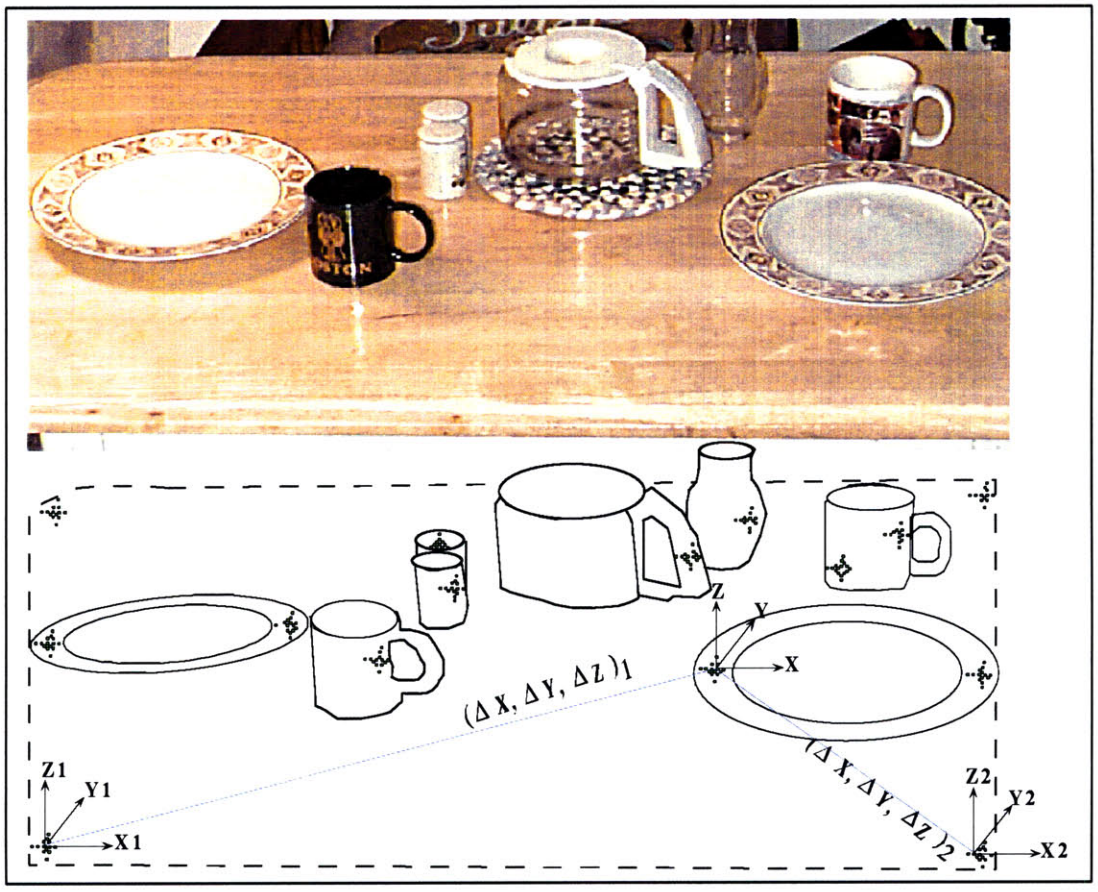


Figure 1.10: A Difficult Problem for Most Recognition Systems

Figure 1.11 is a summary of four different identification methods that were seriously considered, with the main pros and cons of each listed.

Object Location and Indentification (The Possibilities)

Active Elements: Transponders and Element Interrogation + Accurate Identification - Minimum Size and Power Requirement	Bar Codes + Unlimited Object ID + Accurate Identification - Noticeable
Structured Lightning and Shape Identification + Provides Contour Information - Computationally Intensive - Limited Object ID	Pattern Recognition - Computationally Intensive - Limited Object ID

Design Concepts Rated in a Selection Matrix with the following Criteria:

- SAFETY (ANSI Standards, Common Sense)
- COST (Standard Components, Modularity, Scaling, Complexity)
- ROBUSTNESS (Tracking Rate, Noise Considerations)
- AESTHETICS (Non-Intrusive?)

Selection Matrix Indication: Develop a Bar Code Based Object Location and Identification System.

Figure 1.11: The Four Most Popular Methods of Object Identification

As can be seen, transponders would be a desirable option for their accurate identification, but would fall short due to their minimum power and size requirements. Bar codes would also be desirable for their unlimited object identification capability, but would violate the "non-intrusive requirement." Structured lighting and pattern recognition would be fairly undesirable based on their computational intensity and the limitations on the number of objects that could be identified. From these considerations and the construction of a Pough chart, it was decided that a bar code reading system would be used, but due to the visibility of the bar code, steps would be taken to make them invisible so as to fulfill the non-intrusive requirement.

A system was developed that could read bar codes printed in laser dye, which is normally used in tunable cavity lasers to change the wavelength of emission. The system developed is composed of a laser diode that emits at 785 nm and fluoresces a laser dye that has peak absorption at this wavelength. The ink emission is at 833 nm, and a CCD camera that is sensitive in this near IR region is used to capture this emission signal. Since both the laser and ink emission are at wavelengths above what the human eye can see, both the activation and acquisition are invisible to the human, meeting the non-intrusive requirement.

The laser dyes that have absorption and emission in the near IR also have low quantum efficiencies associated with them, which is just the ratio of photons returned in emission to the total photons that are used to fluoress the ink. Due to this low returned signal strength, very robust signal processing methods had to be developed to enhance the signal-to-noise ratio (S/N). With every increase in S/N ratio, the laser power could be reduced, making the system even more eye and skin safe. However, with too great a reduction in laser power, the S/N ratio would drop and errors in the read would occur.

1.2 Previous Work

In the development of an invisible coding system with vision-based signal acquisition, there were many good pieces of research to reference over the past years. With regard to the development of a readable bar code in three space, the design of the bar code geometry was of particular interest in that particular visual cues may be used in the determination of shape [1][2]. Much work has been done with regard to light modulation and developed knowledge of reflection characteristics in object identification [3][5][6][11][15][16]. Some work has also been done in the other direction, using shadow information to determine surface characteristics [8].

As far as the methods that have been traditionally used in the determination of object identity, many good methods were found. One method was a straight correlation of static object images to references sets of object images [18][30][31][32][33][34][35]. A second method used object contours for object identification. A last method used non-rigid models of object motion for object identification, which is more of a dynamic analysis of a scene [43][45].

Many comparisons have been made between the amount of information that can be obtained from successive acquisition of single perspective views [21][22][23][26][54][57] and simultaneous, different perspective or stereo views of an object [12][13][14][29][49] with regard to depth and location accuracy. This information then leads into the application of object tracking and motion estimation, in which many research works have done [36][44][46][47][48][50][51][52][53][55][56][58].

One interesting direction that can be taken when a system exists that can track objects is the recognition of human intention. By analyzing the motion, human intention may be inferred, as well as some idea of environment makeup, which can be found when looking at sudden changes in trajectory which indicate an obstruction. In this analysis, if correct intention can be determined then the motion itself, if reproduced mechanically, may be cleaned up and spurious motion removed. Several references were found that dealt with these ideas [7][9][10][20][38][39][40][41][42].

One problem that was found in object bar coding was the fact that the bar code may be occluded at times, either from other objects or it's own body. Two very good references were found that dealt with this problem [4][36].

Several books were found for general reference in image processing, signal detection and signal design for ease in acquisition [59][60][66]. For consideration of general imaging challenges, an article that was an overview of the basic concepts in vision based system's was found [68].

On the hardware side, many references were found. The problem of image variance between difference acquisition systems had to be addressed [25]. Understanding of color information gathered from a CCD camera with regard to the signal wavelength was needed [17]. To identify many thousands of objects, a well-structured object database is needed [19]. In search of an efficient coding method, many references were found dealing with digital coding and error checking [62][63][64]. Also, for future application of this research, a need exists for parallel processing and code refinement for computational efficiency [27][28].

1.3 Scope of this Research Work

This research will address such issues as the system architecture, the signal processing algorithms used to deal with the low signal strength, and some work in the description of the noise in the viewable frame. Modeling of the proposed system was used to find component interaction relationships to be used at the design level. Some alternate designs will be presented with regard to light diffusion and modulation, and some ideas for future applications of this work will be discussed.

Some possible applications of this technology is in object location and identification of parts in parts feeders, counterfeit checking, spotlight tracking of performers, surgical tool preparation, stamp and money counting, and any inventory or logistics system. A provisional patent exists on this system.

Chapter 2: An Invisible Code System

With the basic class of design solution set, requirements could be set for this approach. The requirements for this bar code reading system are simple and will be stated briefly. This system must be able to locate and identify, in three dimensions, all objects within a home environment that have been bar coded, and do so in an efficient manner. It is desirable to have an overall system that is totally invisible to humans, from the fluoresced ink to the light activation system. The reader must be able to read a bar code at a distance equal to the largest dimension of the largest room of the house, and the bar code must be of a size that can be mounted on small curved surfaces, such as on the handle of a coffee mug. The bar code must be able to be read correctly each time with some redundancy in error checking provided. The power flux provided by the laser and diffusion\expansion elements must be high enough for ink detectability at a given distance, but low enough to maintain eye safe requirements as set forth in the ANSI standards. The ink\solvent combination must be non-toxic and non-flammable, with the required viscosity a function of the printing method employed.

In addition, Figure 2.00 shows the six main categories of design issues that were of concern in the development of this system.

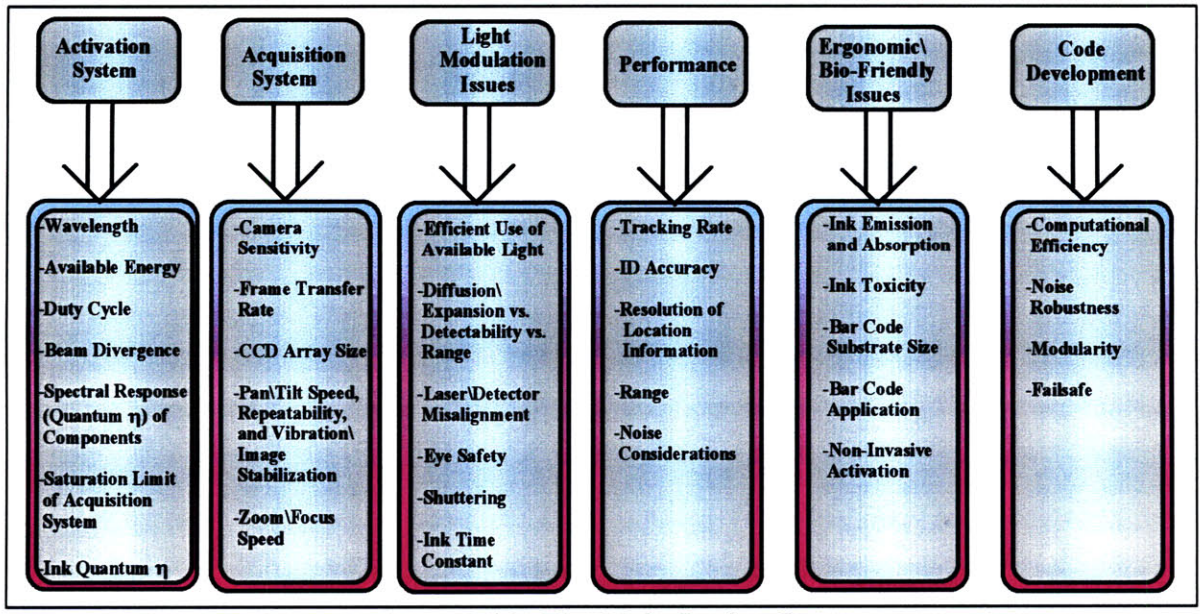


Figure 2.00: The Main Design Issues

2.1 Principle of Invisible Code Systems

A system was desired that could be totally invisible to the humans within the home environment, as well as provide a signal that could be easily differentiated from noise. To accomplish this, a detection method had to be found that would detect at wavelengths above or below what a human could detect. A silicon CCD chip was found to have a spectral response that started at around the same wavelength as the human eye, but extends roughly 200 nm above the top end of the human eye. Figure 2.10 gives a graphical illustration of this feature. The red area denotes the desirable design range for picking an ink and its corresponding detection system.

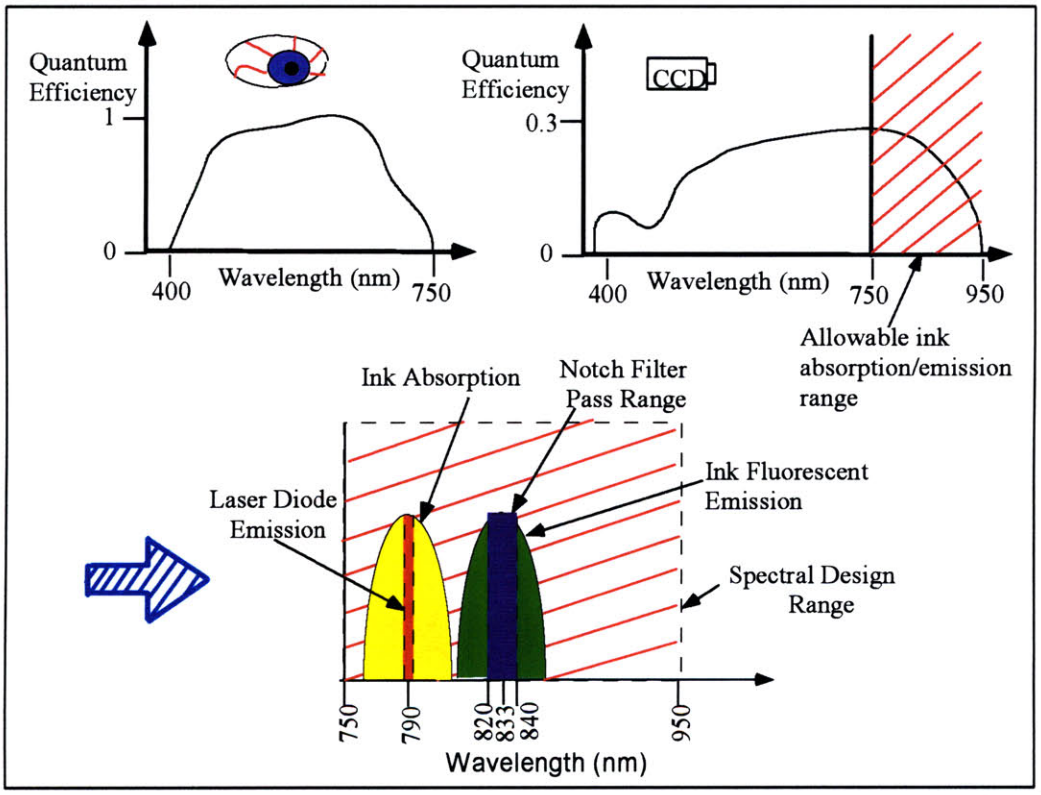


Figure 2.10: The Spectral Responses for the Human Eye and a Typical CCD Chip

Some basic design rules were developed at this point to constrain the design direction for the laser dye activation system. First, the bottom end of the ink absorption curve must be higher than the top end of the human eye spectral response. Specifying this rule leads to a search of all inks/dyes that have absorption at just above 750 nm, the top end of human spectral response. Second, the peak of the ink emission curve must be below the top end of the CCD spectral response. This rule is intended to keep the cost of detection relatively cheap, since cameras

designed specifically for IR applications are upwards of $10,000, and a monochrome CCD detector can be obtained for a few hundred dollars. Third, the top of the laser emission curve must be below the bottom end of the notch filter response curve. If this rule were not obeyed, a very intense noise source, the laser, would be introduced, making detection of the low strength ink emission very difficult. This also gives motivation for finding a laser source with a fairly narrow wavelength operating range. The fourth rule that was used to select an ink\activation system was that the product of the quantum efficiencies for the camera, camera optics, ink, and filter must be such that at a given laser power flux, the ink is detectable. This is the most limiting rule, and does not take into consideration such issues as light diffusion and search time. Finally, the "nominal" power flux provided by the laser source/diffuser combination must still be eye safe. A strong coupling between laser light diffusion and signal detectability should be noted. Figure 2.10 also shows the values of the ink absorption and emission, laser emission and the notch filter pass range for the given system.

2.2 System Configuration

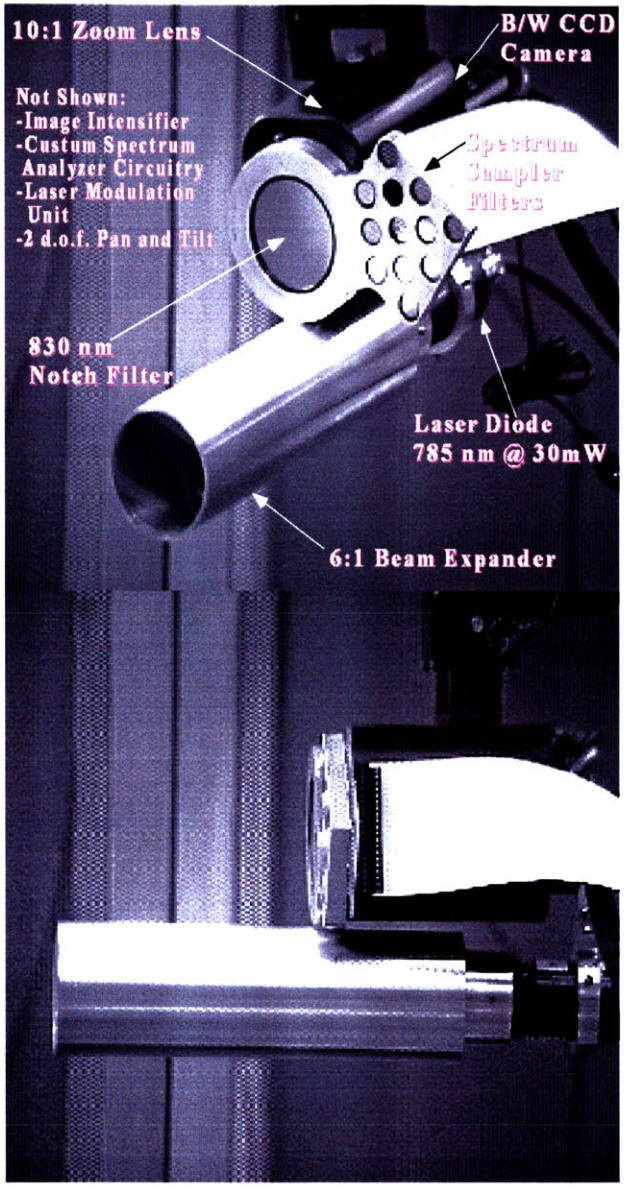
With great consideration of the above issues, as well as some trial and error with different inks and their respective absorption and emission wavelengths, a combination was found that seemed to meet the design requirements. Figure 2.20 gives a functional block diagram of the developed system.

The system consists of a laser diode that emits at 785 nm with a nominal power of 20mW. This emission wavelength is the peak absorption frequency of a special formulation of IR 125 laser dye and solvent. This solvent/dye combination has a peak fluorescent emission at 833 nm. This laser light is taken through a spatial filter to reduce the effects of spatial aberrations in the laser source optics and provide a nice gaussian distribution in power. This light is then taken though a double convex and concave lens combination that is used to expand the 6.96 mm elliptical beam to 42 mm (1:6).

On the side of image acquisition, the image is taken through a notch filter with a pass range of 825-835 nm, with the center wavelength roughly the wavelength of the peak ink fluorescent emission. This is passed through a 10:1 motorized zoom lens. Three channels of D/A are used to control the iris, zoom, and focus of the lens, and one channel of A/D is used to bring in the zoom position information.

Since the ink emission signal is fairly weak due to a low quantum efficiency (measured experimentally to be less than 15%), an image intensifier was used to amplify the light signal by a factor of 30,000. A monochrome CCD camera is then used to measure this conditioned image intensity signal, with a PCI bus frame grabber used to bring this information into the computer.

A system for the Location and Measurement of Fluorescent Emission for the Application of Object Location and ID



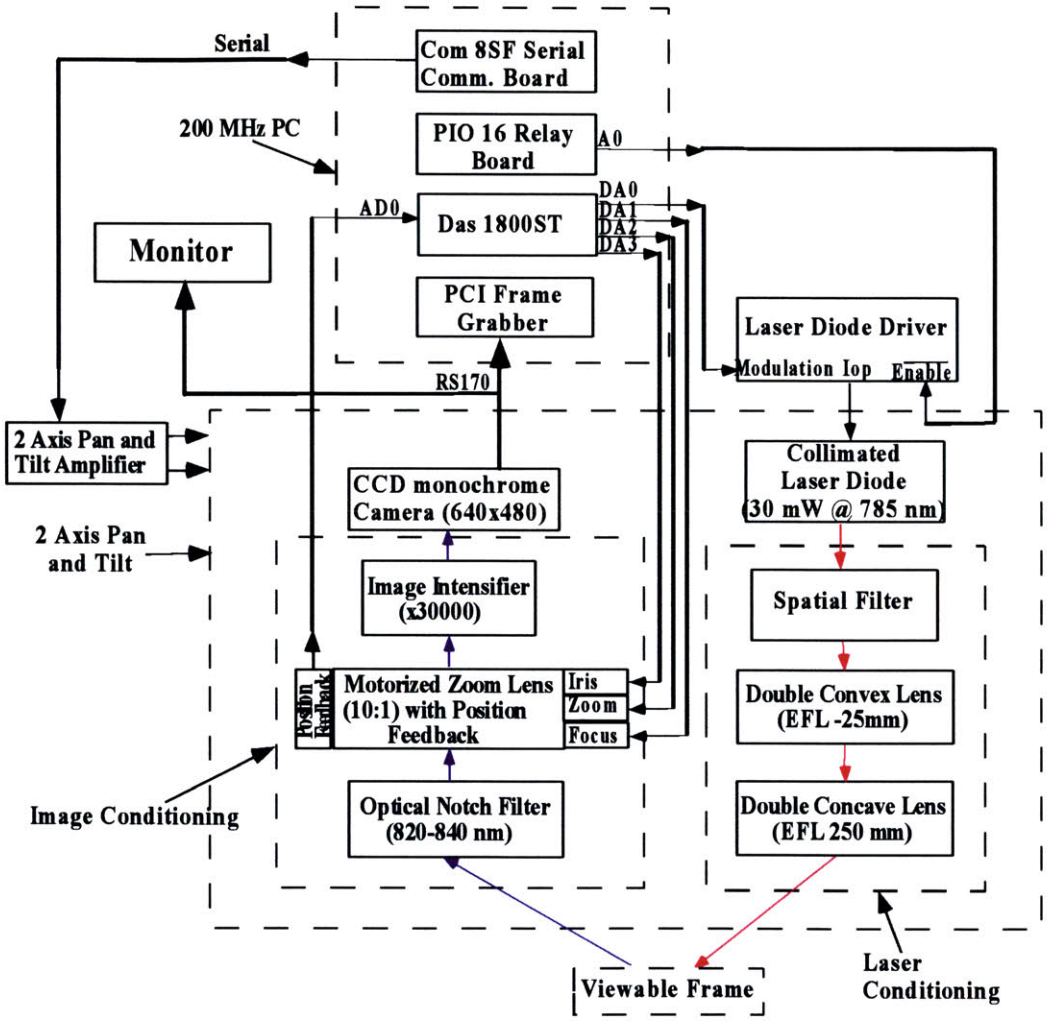


Figure 2.20: A Functional Block Diagram of the Present Day Bar Code Identifier

2.30 The Signal: Evolution of the Bar Code

Up until now, all of the focus has been on system development and methods to deal with the presence of noise in the viewable frame. However, there has been some clear changes in the bar code from the original 1D bar code.

Figure 2.30 depicts the original 2D bar code that was used to replace the 1D version. It consists of three dots that are used for orientation information, and a four by four matrix of dots that are used to give 16 bit identification.

This bar code is read by first identifying the reference points which are at the extreme edges. Then, the distances between the reference points are measured in frame pixels, and used to give a spacing value for the pixel reads. Angles are computed that relate the orientation of the bar code to the frame array, and these angles are used in the coordinate transformation to the bar code coordinate frame. Also shown in Figure 2.30 are the important angles and lengths that are measured. With the pixel read spacing calculated, the center of each bit of the identification matrix can be checked for fluorescence (a 1) or no fluorescence (a 0). However, since nothing is perfect, some error is introduced due to integer roundoff\discretization in the measurement of the angles and reference distances, as well as from aberrations or inconsistencies in the bar code print.

Early on, there were many problems maintaining an accurate bar code due to ever changing lighting conditions, changes in background reflections, and an ever changing magnitude of noise. A need was present to implement a parity check, and one that could be used on not only the rows but the columns as well. For this reason, a fifth column and row were added to the bar code and serve as nothing more than an even (or odd) parity check (also shown in Figure 2.30). The point where this column and this row meet is termed the "redundant parity check bit," in that a change in any row or column parity bit will cause a corresponding change in this bit.

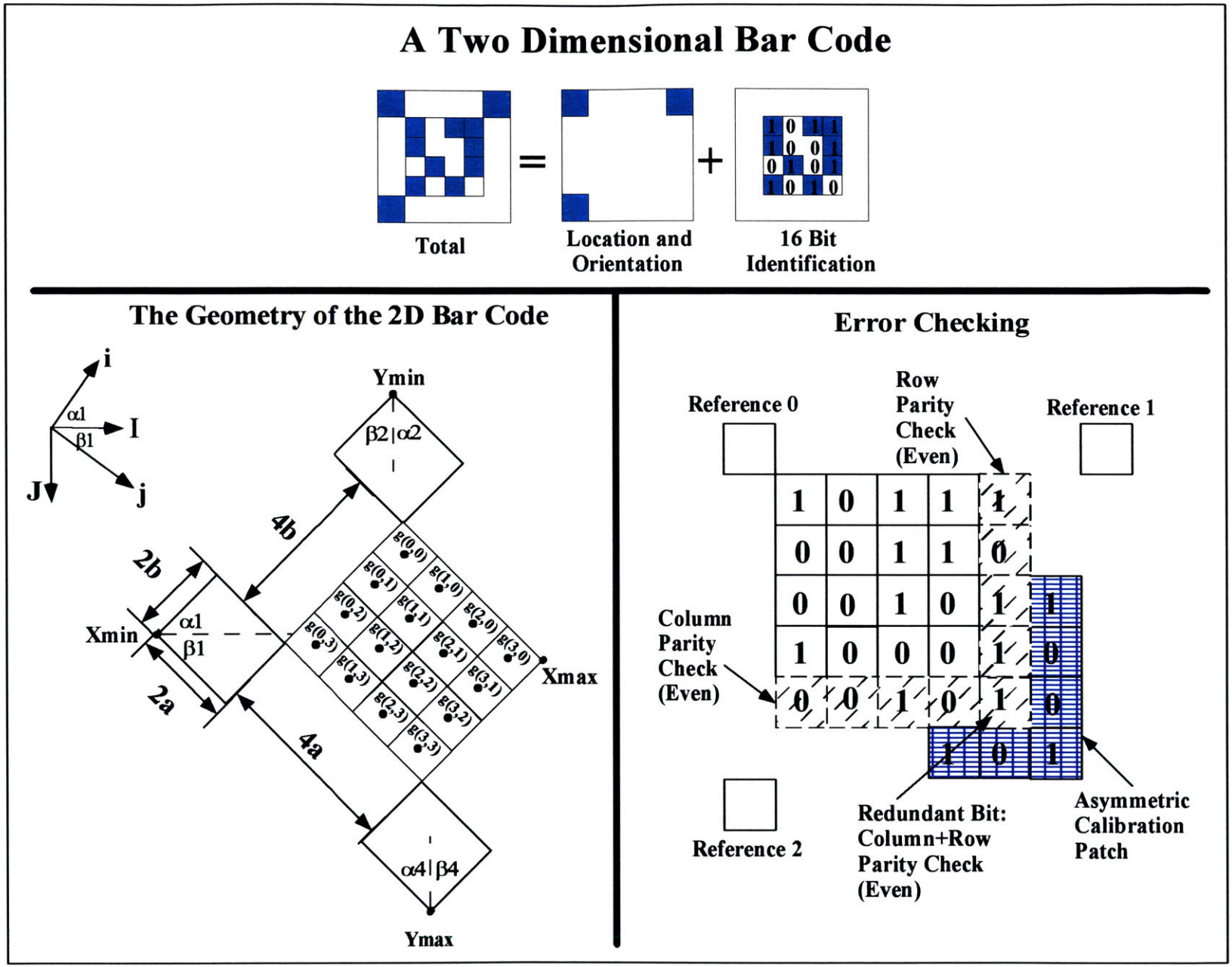


Figure 2.30: The 2D Bar Code

One of the more difficult challenges in this project was to know where to set the threshold to capture the bar code position and identification information. From experience it was observed that the magnitude of the noise and the magnitude of the signal are of the same order. This means that, without any noise removal methods, the threshold must be very precisely set to avoid bogus information. To reduce the probability of an incorrect threshold selection, the " Asymmetric Calibration Patch" was added. Since the positional errors propagate outwards from reference point 0 due to roundoff error, the patch was placed at the extreme end across from this reference point. This works in such a way as to allow the iterative selection of a threshold based on a measured read of the expected values for this patch. If the expected "1's" are being read as "0's", the threshold needs to be lowered. If the expected "0's" are being read as "1's", the threshold needs to be raised. From this, a range of suitable thresholds can be found, which gives insight into the signal to noise ratio.

At this point, all of the bar codes have been printed using a stencil printer. With the given stencil dimensions, the resulting bar code is a 0.325" (8.25 mm) square without the parity bits and calibration patch, and 12 mm with these features included. The ink dot drops were on the order of 15 mils. In the future, it is hoped that this technology will catch on, and manufacturers will apply these bar codes to their product.

2.4 An Alternative Approach to Light Modulation

Up until now, the light modulation has been through motion of a pan and tilt, which, because of it's weight, is very slow. The light diffusion is accomplished through a galilean beam expander, but other methods of diffusion were tried. With the current setup, the more the laser is diffused, the less signal that will be received at the CCD. To measure the emission of the ink at varying radial distances, a motorized zoom lens is used, but there is definitely time associated with this process when changing fields of view. The frame rate of the frame grabber is the limiting factor to search velocity, and for these reasons, an alternate design was made for future implementation of this technology.

Figure 2.40 depicts a two-axis galvo scan system. The laser diode beam is taken through a collimator and is reflected off the notch filter (this works since the filter blocks this wavelength). This light is then taken through the two axis's of the galvos, and when it returns, the image returns on the same axis of the outgoing beam. The part of the image that corresponds to ink emission is then passes through the notch filter. This is taken to a mirror-type beam splitter that breaks into two perpendicular beams of 50% power each. These beams are then taken to another set of beam splitters which splits them again. The overall beam (image) intensity is now 25% of the original beam intensity. At the output of each beam splitter a double convex and double concave lens are used to give a fixed magnification zoom, and another double convex lens is used to make a radiometer to focus this image energy on to an avalanche photo-diode. Four distinctly different magnification settings are used to a system that can, almost instantly, look at four different radial ranges. This removes the slow zoom and focus of the motorized zoom lens. The avalanche photo-diode was selected because it produces a current proportional to intensity but with a very large gain. The sample time for these modules can be up to 10^{12} Hertz, which is compared to the 30 Hz frame rate of the frame grabber.

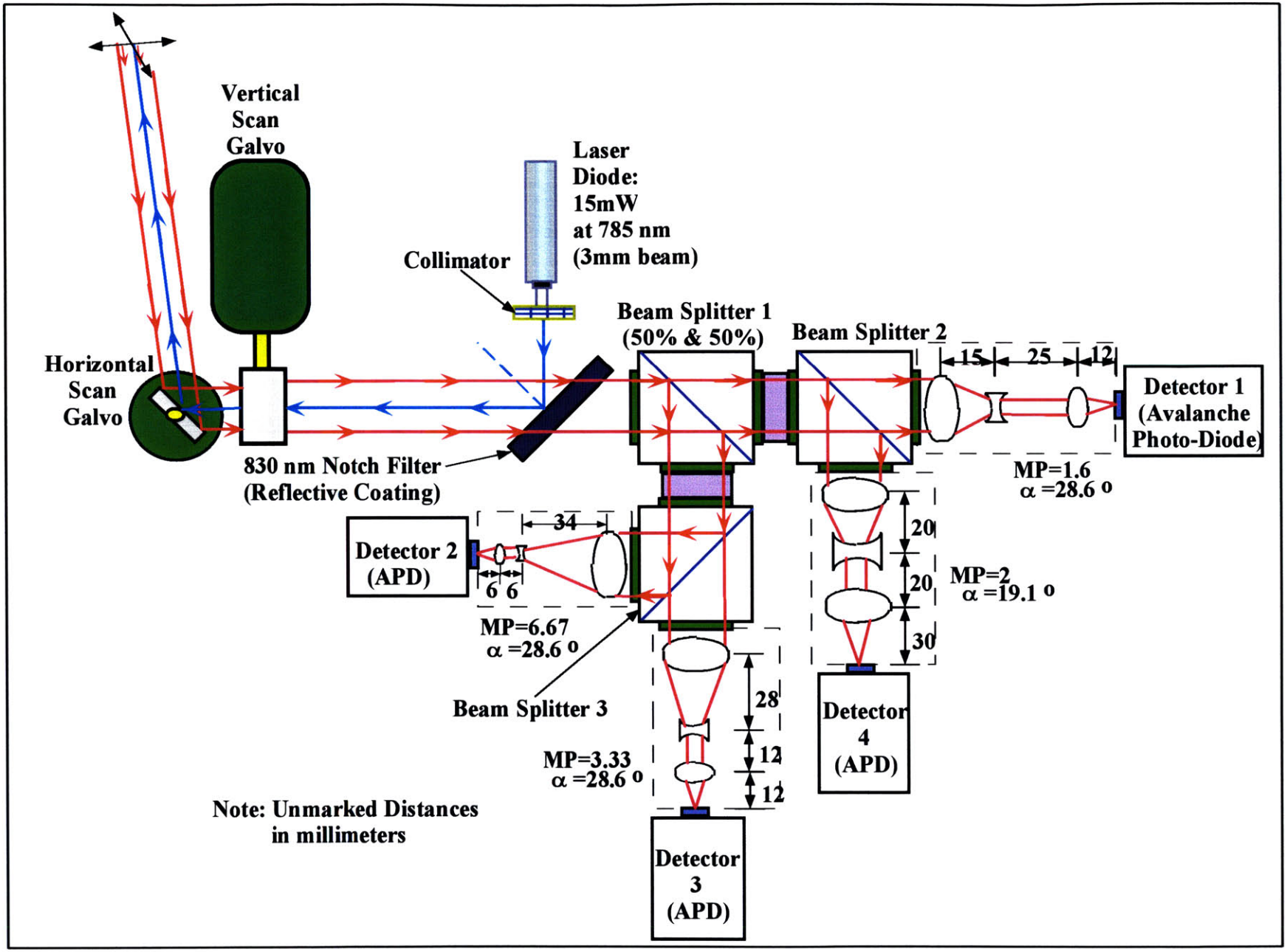


Figure 2.40: An Alternative Design for Light Modulation

Chapter 3: System Modeling

With the hardware selected, it was time to take a look at the processes involved starting from the image activation to the image acquisition, and everything in between. This chapter deals with the modeling that was performed to get an idea of the interaction of the components within the design, as well as to give insight into design changes that could be made to give more weight to the signal with regard to the noise.

3.1 Modeling of the Ink Fluorescent Emission and Capture

With the selection made for the ink, laser, and notch filter, it was possible to find some relations between the laser power, quantum efficiencies of the camera, optics, filters, and ink, and the CCD frame dimensions and to relate this to a signal to noise ratio. For this derivation, some basic assumptions had to be made:

-Assume the laser is perpendicular to the fluoresced material to give a circular and not an elliptical coverage area. This also assumes that the laser has a circular beam and not an elliptical beam.

-Although the laser is gaussian in power and wavelength, assume an RMS value at a given wavelength, and assume it is constant over the coverage area. This makes it possible to say all of the laser energy is at a given wavelength λ_a.

-Assume that the laser beam diameter is less than the diameter of the ink substrate so that all of the laser photons are used in the fluorescence process.

-Assume that the full CCD array sees the emission, and that there is no dead space between pixels.

-Assume other sources of light at ink emission wavelength are negligible for contribution to measured intensity.

-Assume no blockage of incoming light due to improper use of the iris.

-Assume that the CCD area is much less than the surface area of the virtual sphere of radius L from the laser source to the fluoresced ink.

-Assumes the number of photons incident on the CCD is less than some saturation value.

-Assumes there is no divergence to the laser beam (perfectly collimated).

By knowing the laser emission wavelength, λ_a, it is possible to calculate the energy of one light photon at this wavelength:

$$E = \frac{hc}{\lambda_a} \tag{3.101}$$

where **h** is plank's constant ($\mathbf{6.626x10^{-34}}$ Js), **c** is the speed of light in vacuum ($\mathbf{3x10^{8}}$ m/s), and E is in Joules. By knowing the power at which the laser is operating, it is possible to calculate the number of photons per second, $\dot{\xi}_t$, that are emitted by the laser:

$$\dot{\xi}_t = \frac{P_s \lambda_a}{hc} \tag{3.102}$$

where P_s is the operating power of the laser source. The number of photons returned ($\dot{\xi}_R$), per second, from the fluorescent emission is related to the number of photons into the process through the quantum efficiency of the ink, η_i, evaluated at the absorption wavelength:

$$\dot{\xi}_R = \frac{\eta_i|_{\lambda_a} P_s \lambda_a}{hc} \tag{3.103}$$

These photons are returned to the environment in a "half-sphere" manner, as shown in Figure 3.10:

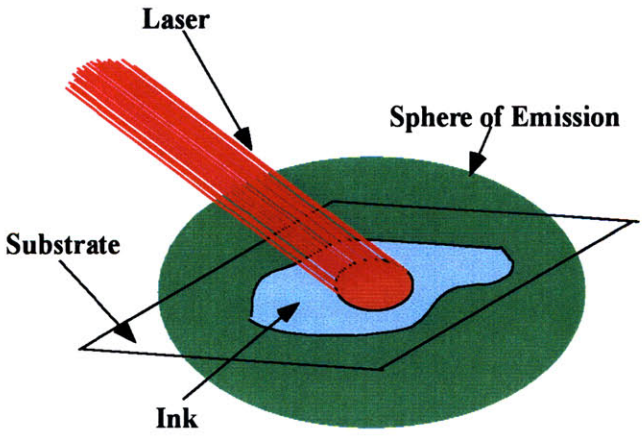


Figure 3.10: The Sphere of Photon Emission

The photon flux is related to the surface area of this half sphere and the total number of photons returned per second:

$$\dot{\xi}'' = \frac{\dot{\xi}_R}{2\pi L^2} \tag{3.104}$$

where L is the distance from the fluoresced ink to the CCD detector. If each pixel of the CCD camera detects these photons, the number of photons per second to the CCD may be determined:

$$\dot{\xi}_{CCD} = \dot{\xi}'' x_L y_L \eta_z|_{\lambda_e} \eta_c|_{\lambda_e} \eta_f|_{\lambda_e} \tag{3.105}$$

where x_L and y_L are the horizontal and vertical dimensions of the CCD array, η_z is the quantum efficiency of the zoom lens, η_c is the quantum efficiency of the camera optics, and η_f is the quantum efficiency of the notch filter. These efficiencies are evaluated at the fluorescent emission wavelength, λ_e.

$$\dot{\xi}_{CCD} = \dot{\xi}'' x_L y_L \eta_z|_{\lambda_e} \eta_c|_{\lambda_e} \eta_f|_{\lambda_e} \tag{3.106}$$

Given the shutter time of the camera, t_s, it is possible to determine the number of photons that hit the CCD in one shuttered frame. It is also possible to determine the number of photons hitting an individual pixel over this same time if the horizontal and vertical pixel resolution values, P_H and P_V, are known:

$$\xi_{pixel} = \frac{\dot{\xi}_{CCD} t_s}{P_H P_V} \tag{3.107}$$

Combining the above equations gives:

$$\xi_{pixel} = \frac{P_s \lambda_a t_s x_L y_L \eta_z|_{\lambda_e} \eta_c|_{\lambda_e} \eta_f|_{\lambda_e} \eta_i|_{\lambda_a}}{2\pi L^2 hc P_H P_V} \tag{3.108}$$

At the CCD detector, a current is produced that is proportional to the measured intensity through a gain g_{pi} (mA/photon), which relates the number of photons hitting the pixel to a current. This current is converted to a voltage through gain g_{iv} (volts/mA). The voltage range of the A/D can be denoted as R_{AD}, and is normally 5V. The digital value range can be denoted D_{AD}. For an 8 bit A/D, this value is 255. Using this information, a rough idea can be formed as to the number of photons hitting a pixel based on the measured digital value, N_m, representing intensity:

$$N_m = \text{int}\left[\frac{P_s \lambda_a t_s x_L y_L \eta_z|_{\lambda_e} \eta_c|_{\lambda_e} \eta_f|_{\lambda_e} \eta_i|_{\lambda_a} g_{pi} g_{iv} D_{AD}}{2\pi L^2 hc P_H P_V R_{AD}}\right] \tag{3.109}$$

It should be noted that there is a saturation value to the number of photons that may enter and still be detected:

$$\xi_{SAT} = \frac{R_{AD}}{g_{pi} g_{iv}} \tag{3.110}$$

Due to the discrete nature of the measured digital value, a range of incident photons may give the same measured digital value. Now that a relation has been developed to relate the measured digital value of a CCD to the laser power and wavelength, and the characteristics of the optics in between, a more general form may be developed. A summation may be made for the power source, over all wavelengths, of the energy due to a light photon at that particular wavelength, and noting that each element's quantum efficiency is also a function of wavelength:

$$N_m = \text{int}\left[\frac{(\sum_{\lambda=0}^{\infty} P_s(\lambda)\eta_z(\lambda)\eta_c(\lambda)\eta_f(\lambda)\eta_i(\lambda)\lambda) t_s x_L y_L g_{pi} g_{iv} D_{AD}}{2\pi L^2 hc P_H P_V R_{AD}}\right] \tag{3.111}$$

At this point, it is possible to consider laser emitters that are slightly divergent, and are not circular but elliptical. Let the major and minor axis diameters be d_1 and d_2, respectively, and the divergence angles from the major and minor axis θ and ϕ, respectively. Figure 3.11 shows the beam divergence variables. The new diameter, larger due to the divergence of the beam over a distance L, is given by:

$$d_1' = 2L\tan\theta + d_1 \tag{3.112}$$

$$d_2' = 2L\tan\phi + d_2 \tag{3.113}$$

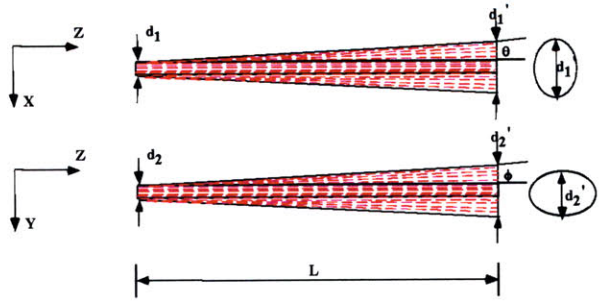


Figure 3.11: The Laser Beam Divergence

The area of the resulting ellipse, at the fluoresced object, is just the major and minor axis multiplied by a constant:

$$A_1 = \frac{\pi}{4} d_1' d_2' \tag{3.114}$$

or

$$A_1 = \frac{\pi}{4}[2L\tan\theta + d_1][2L\tan\phi + d_2] \tag{3.115}$$

At this point, it would be appropriate to include a beam expander term. Assuming that there is no change in the divergence angles, and given a magnification factor L_{MF}, the area A_1 is increased by a factor of L_{MF}^2:

$$A_1' = A_1 L_{MF}^2 \tag{3.116}$$

and the passed photons are affected by the quantum efficiency of the expander, η_e:

$$N_m = \text{int}\left[\frac{(\sum_{\lambda=0}^{\infty} P_s(\lambda)\eta_z(\lambda)\eta_c(\lambda)\eta_f(\lambda)\eta_i(\lambda)\eta_e(\lambda)\lambda)t_s x_L y_L g_{pi} g_{iv} D_{AD}}{2\pi L^2 hc P_H P_V R_{AD}}\right] \tag{3.117}$$

If the laser emission totally covers the bar code substrate, but there isn't ink everyplace that is covered by the laser ellipse, then a simple area ratio adjustment must be made. If there are **n** ink points of diameter d_3 composing the bar code, then the following is the adjustment that can be made to equation 3.108 to give the actual number of photons returned due to the ink dots:

$$\xi'_{pixel} = \xi_{pixel} \frac{A_{bc}}{A_1'} \tag{3.118}$$

where the area of the bar code, A_{bc}, is:

$$A_{bc} = \frac{n\pi d_3^2}{4} \tag{3.119}$$

which leads to:

$$\xi'_{pixel} = \frac{n d_3^2 \xi_{pixel}}{L_{MF}^2[2L\tan\theta + d_1][2L\tan\phi + d_2]} \tag{3.120}$$

Modifying equation 3.117 appropriately gives:

$$N_m = \text{int}\left[\frac{(\sum_{\lambda=0}^{\infty} P_s(\lambda)\eta_z(\lambda)\eta_c(\lambda)\eta_f(\lambda)\eta_i(\lambda)\eta_e(\lambda)\lambda)t_s x_L y_L g_{pi} g_{iv} n d_3^2 D_{AD}}{2\pi L^2 hc P_H P_V R_{AD} L_{MF}^2 [2L\tan\theta + d_1][2L\tan\phi + d_2]}\right] \quad (3.121)$$

Most of the earlier assumptions have been eliminated. However, the above equation still assumes that the full CCD array sees the emission, there is no dead space between the pixels, and the number of photons incident on the CCD is less than some CCD saturation value. There is also the assumption that other sources of light at ink emission wavelengths are negligible for their contribution to the measured intensity, that there is no blockage of incoming light due to improper use of the iris. Also assumed is that the CCD area is much less than the surface area of the virtual sphere of radius L. These all seem to be reasonable assumptions, except possibly the assumption about a total absence of noise in the form of other sources of light photons.

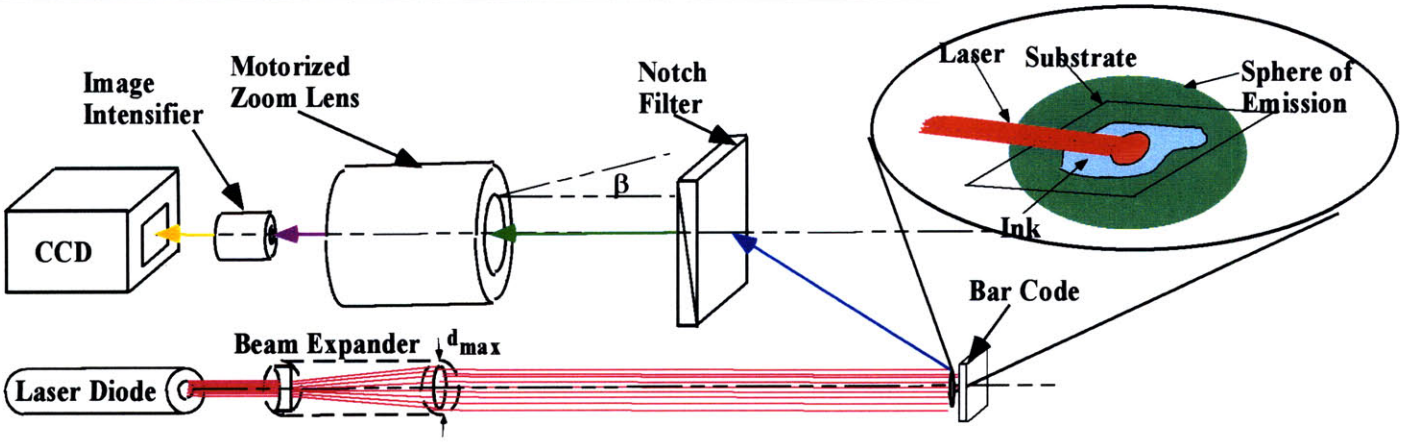
If equation 3.121 is said to be the measured intensity due to signal only, it is possible to modify this equation with a noise term in the following manner, where ξ_{Noise} is the photons, per pixel, due to noise:

$$N_M = \text{int}\left[\left[\frac{t_s x_L y_L g_{pi} g_{iv} D_{AD}}{2\pi L^2 hc P_H P_V R_{AD}}\right]\left[\frac{(\sum_{\lambda=0}^{\infty} P_s(\lambda)\eta_z(\lambda)\eta_c(\lambda)\eta_f(\lambda)\eta_i(\lambda)\eta_e(\lambda)\lambda) n d_3^2}{L_{MF}^2 [2L\tan\theta + d_1][2L\tan\phi + d_2]} + \xi_{Noise}(\lambda)\right]\right] \quad (3.122)$$

As shown before in Figure 2.20, an image intensifier is used to amplify the viewed light signal. This is inserted between the camera zoom and the camera. This intensifier amplifies the light signal by a gain L_g, and also shifts the wavelength by λ_S. This only affects the quantum efficiency term associated with the camera. There is also a quantum efficiency associated with this element, η_{ap}:

$$N_M = \text{int}\left[\left[\frac{t_s x_L y_L g_{pi} g_{iv} D_{AD} L_g}{2\pi L^2 hc P_H P_V R_{AD}}\right]\left[\frac{(\sum_{\lambda=0}^{\infty} P_s(\lambda)\eta_z(\lambda)\eta_c(\lambda-\lambda_S)\eta_f(\lambda)\eta_i(\lambda)\eta_e(\lambda)\eta_{ap}(\lambda)\lambda) n d_3^2}{L_{MF}^2 [2L\tan\theta + d_1][2L\tan\phi + d_2]} + \xi_{Noise}(\lambda-\lambda_S)\right]\right] \quad (3.123)$$

Figure 3.12 summarizes the elements in the design with respect to the component variables.



Element	Laser Diode	Beam Expander	Bar Code	Notch Filter	Zoom Lens	Image Intensifier	CCD Camera	Medium (Air)
Associated Parameters	$P_s(\lambda)$ d1 d2 θ,ϕ	L_{MF} $\eta_e(\lambda)$ d_{max}	$\eta_i(\lambda)$ n d_3 λe, λa	$\eta_f(\lambda)$	Z $\eta_z(\lambda)$ β	$\eta_{ii}(\lambda)$ λs	X_L, Y_L P_v, P_h $\eta_c(\lambda)$, t_s g_{pi}, g_{iv}	L c $\zeta(\lambda)$

Figure 3.12: Photon Counting

It should be noted that the image intensifier amplifies not only the signal, but the noise as well. With this in mind, a signal to noise ratio can be defined, which is the ratio of the signal variance to the variance of the noise:

$$S/N = \frac{\sigma^2{}_{\xi\,pixel}}{\sigma^2{}_{\xi\,Noise}} \tag{3.124}$$

This ratio may be determined experimentally by taking a frame with the laser on and with it off, positioned to point to a bar code. With the laser off, the spatial deviation of the resulting frame can be found and used for the noise component value. By taking a filtered version of the "laser on frame" and calculating the square of this deviation, the signal deviation may be approximated. This method is only approximate, since the true signal deviation may change slightly due to the filtering.

3.2 The Optics

Up until this point, the assumption that every pixel sees the bar code photons equally has been used. However, relations can be developed to find the number of pixels in the CCD frame that actually "see" the photons, as shown in Figure 3.23. The following is a list of variables that are used to find these relations:

Z = zoom setting
LFL = lens focal length
L = distance to the bar code
cf = camera format
fov_H = field of view, horizontal
fov_V = field of view, vertical
P_V = vertical pixels in frame
P_H = horizontal pixels in frame
x_L = horizontal dimension of CCD chip
y_L = vertical dimension of CCD chip
x_{BC} = horizontal dimension of bar code
y_{BC} = vertical dimension of bar code
$S_{H,BC}$ = number of horizontal pixels that see the bar code
$S_{V,BC}$ = number of vertical pixels that see the bar code
α = horizontal rotation angle of bar code respect to the camera plane
β = vertical rotation angle of bar code respect to the camera plane

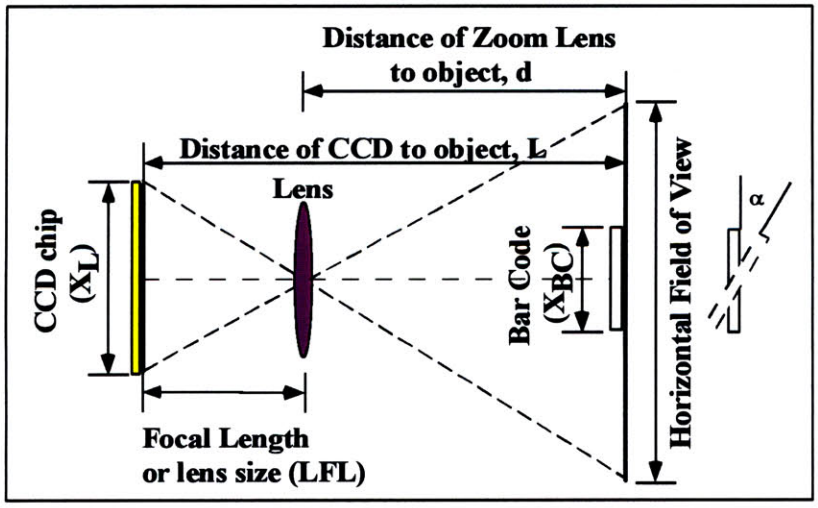


Figure 3.23: The Camera Optics

The distance, d, from the camera lens to the viewed object (bar code) can be found by multiplying the zoom setting by the lens focal length and the horizontal field of view, and dividing by the camera format:

$$d = \frac{Z(LFL)(fov_H)}{cf} \tag{3.225}$$

The horizontal and vertical field of views are related by:

$$fov_H = \frac{1}{r} fov_V \tag{3.226}$$

where **r** is the ratio of the number of vertical pixels to the horizontal pixels for a given CCD chip:

$$r = \frac{P_V}{P_H} \tag{3.227}$$

This gives the following two expressions for the horizontal and vertical field of views, respectively:

$$f_{OV,H} = \frac{(c_f)(d)}{(Z)(LFL)} \tag{3.228}$$

$$f_{OV,V} = \frac{r(c_f)(d)}{(Z)(LFL)} \tag{3.229}$$

The distance **d** varies according to the following:

$$d = L - LFL \tag{3.230}$$

The ratio of the projected horizontal bar code side ($\mathbf{x_{BC}}\cos\alpha$) to the horizontal field of view is equal to the horizontal pixels of the CCD that see the bar code ($\mathbf{S_{H,BC}}$) to the total horizontal pixels ($\mathbf{P_H}$)of the CCD:

$$\frac{x_{BC}\cos\alpha}{f_{OV,H}} = \frac{S_{H,BC}}{P_H} \tag{3.231}$$

A similar relationship can be found for the vertical direction:

$$\frac{y_{BC}\cos\beta}{f_{OV,V}} = \frac{S_{V,BC}}{P_V} \tag{3.232}$$

Solving for the number of pixels seeing the bar code in the above two expressions, and substituting what is known about the horizontal and vertical fields of view, the following two expressions can be given:

$$S_{H,BC} = \frac{(Z)(LFL)x_{BC}P_H\cos\alpha}{(L-LFL)(c_f)} \tag{3.233}$$

$$S_{H,BC} = \frac{(Z)(LFL)y_{BC}P_H \cos\beta}{(L-LFL)(c_f)} \tag{3.234}$$

Multiplying the above two quantities, an area, in pixels, that see the bar code can be found. To find the actual number of pixels that are "lit" from the bar code presence (S_{lit}) it is possible to ratio the area of the bar code dots to the total area of the bar code, and multiply by the pixel area that sees the bar code:

$$S_{lit} = S_{H,BC}S_{V,BC}\frac{(n\pi d_3^2/4)}{x_{BC}y_{BC}} \tag{3.235}$$

It should be noted that dividing both sides of equation 3.235 by **n** gives the number of pixels lit per bar code dot, which can be used at the design stage to specify the minimum number of pixels required to give a measured "high." When specified, this information can be used to eliminate spatial noise in the viewable frame.

Simplifying the above equation and solving for the pixels lit per bar code dot, as well as a conversion between square feet and square inches:

$$\frac{S_{lit}}{n} = \frac{(\pi d_3^2)((Z)(LFL)(P_H))^2}{576((L-LFL)(c_f))^2}\cos\alpha\cos\beta \tag{3.236}$$

Assuming that the Length **L** to the object is much greater than the Lens Focal Length (**L>>LFL**):

$$\frac{S_{lit}}{n} = \frac{(\pi d_3^2)((Z)(LFL)(P_H))^2}{576((L)(c_f))^2}\cos\alpha\cos\beta \tag{3.237}$$

The above equation can be used to set a length **L** in feet and maximum bar code rotation angles, α and β, and based on the camera zoom optics, calculate the number of pixels lit per bar code dot. It should be noted that the dot diameter, $\mathbf{d_3}$, is in inches, and the lens focal length, **LFL**, and camera format, $\mathbf{c_f}$, are in millimeters. If it is desired to calculate the maximum distance $\mathbf{L_{max}}$ that the object can be at and still provide a given number of pixels lit, the following applies:

$$L_{\max} = \sqrt{\frac{(\pi d_3^2)((Z)(LFL)(P_H))^2}{576(c_f)^2(S_{lit}/n)}\cos\alpha\cos\beta} \tag{3.238}$$

A relationship exists between the bar code dot diameter and the actual size of the bar code. A ratio must be given for the dot diameter to the "space" length for a given template, since the dots will not be exactly next to one another. Assuming a square bar code and a ratio $\mathbf{R_{d3,s}}$, as well as **m** array columns for the signal and parity checking, the following relationship exists:

$$X_{bc} = Y_{bc} = mR_{d3,s}d_3 \tag{3.239}$$

Plugging in a modified version of equation 3.237 into the above equation, it is possible to see the effect of changing the system parameters on the overall bar code size:

$$X_{bx} = Y_{bc} = mR_{d3,s}\sqrt{\frac{576(S_{lit}/n)((L_{\max})(c_f))^2}{(\pi)((Z)(LFL)(P_H))^2\cos\alpha\cos\beta}} \tag{3.240}$$

It should be noted that the above equation is meaningless for bar codes that are perpendicular to the camera plane. Also, the accuracy of the 2D location information is dependent on the pixel resolution of the camera, which is dependent on the zoom setting and view length. The resolution $\mathbf{R_{p,h}}$ is in inches/pixel and **L** is in feet The horizontal resolution is used, since this value is larger than the vertical resolution:

$$R_{p,h} = \frac{12(c_f)(L)}{(Z)(LFL)(P_H)} \tag{3.241}$$

3.3 Design Considerations from Modeled System

As can be seen from the previous sections, much design insight can be obtained through proper tracking of photon movement through the system as well as some consideration of the optics and bar code geometry. By analyzing equation 3.123, it can be seen that to increase the actual signal part of the measured photon emission, many design adjustments can be made. The laser power could be increased, which would give more returned photons from the ink emission, or the quantum efficiencies of the notch filter, beam expander, zoom or image intensifier could be increased within the wavelength of ink emission and reduced in the other wavelengths. The bar code dot size could also be increased, increasing the number of returned photons for the same laser power. On the reception side, the CCD surface area could be increased, which basically increases the view factor from the bar code to the CCD. The sample time could be increased to "count" photons for a longer time, which is "time averaging." The distance from the bar code to the camera could be reduced, which would give a disproportional increase since the returned signal would quadruple for a halving of this distance. The same idea applies to the beam expander; for a one half reduction in beam expansion, the signal would be four times as large. Of course, one of the more important considerations is the reduction of the noise around the ink emission wavelength. Since this is very hard to do with pure hardware solutions, other methods were found to get past this problem. This is the subject of chapter 4.

Chapter 4: Methods to Increase the Signal to Noise Ratio

In the first system which used UV activated ink, the problem of a small signal to noise ratio was addressed in depth. A technique dubbed the "Perturbation\Correlation Method" was developed that worked fairly well to reduce the effect of noise in the viewable frame.

4.1 The Perturbation\Correlation Method

The Perturbation\Correlation Method is designed specifically to increase the signal to noise ratio by actively changing the parameters that are known to bring the most change in the signal, and then performing a correlation to the known perturbation input. For this application, the ink was known to fluoress at a given wavelength, and the fluorescent emission strength was a function of the power of the light source at this wavelength.

Due to a lack of sensitivity to this parameter change, the magnitude of the noise remains fairly constant over this perturbation. When a correlation is made between the noise+signal measurement and the perturbation, the noise falls away due to a weak correlation and only the signal, the ink spot, remains. It should be mentioned that this method need not be restricted to a power variance of the light source, but could be applied to any measurable parameter that is sensitive to perturbation.

For the case of this application, a sinusoid perturbation in light power is applied, and a certain number of frame array (intensity) are sampled over the sine period, and for each frame, a spatial gradient is performed. This gives an idea of the areas of the frame that are most sensitive to the given perturbation. A sequence is then formed that is composed of pixel intensity values for the same pixel over the many frames, and a discrete time correlation is made of this sequence to the perturbation input. Figure 4.10 gives an illustration of this concept.

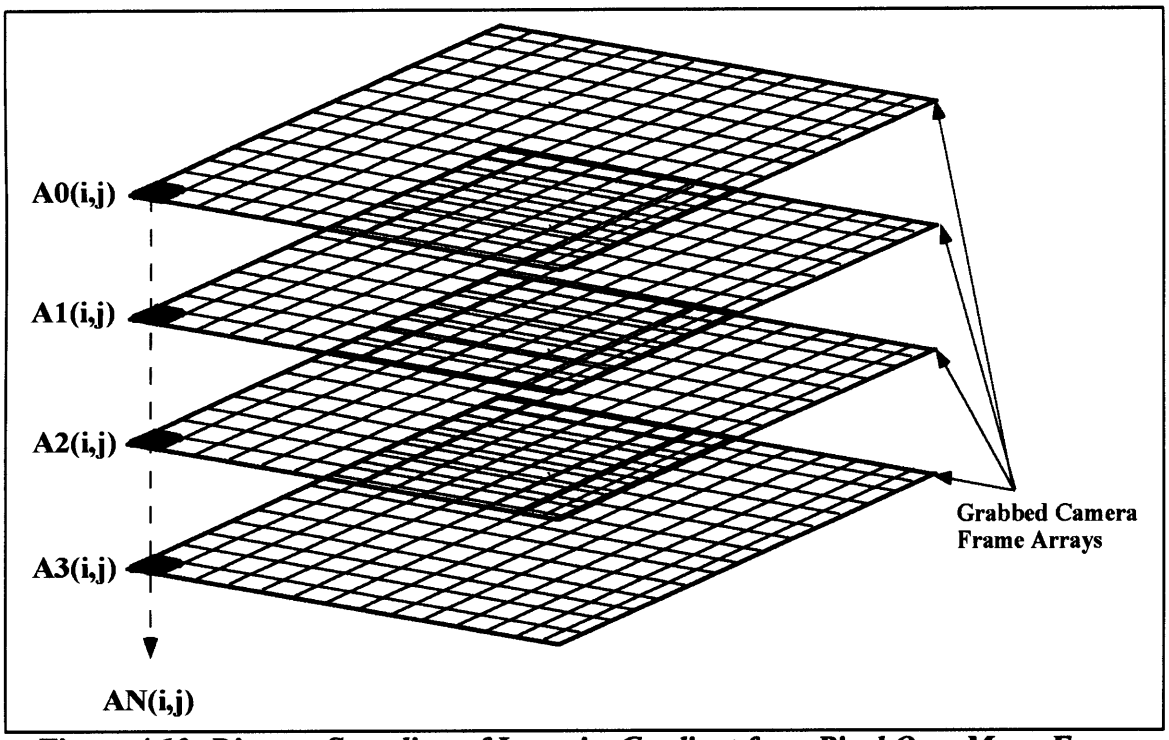


Figure 4.10: Discrete Sampling of Intensity Gradient for a Pixel Over Many Frames

The sequence that would be made would be the following:

$$Y[n,i,j] = [A_0(i,j), A_1(i,j), A_2(i,j), A_3(i,j),.,.,A_N(i,j)] \tag{4.101}$$

This sequence would then be correlated to the power perturbation sequence:

$$X[n,i,j] = [B_0 + B_1 \sin(2\pi n / T_k)] \tag{4.102}$$

At this point, an argument can be made that each pixel is not independent from one another, and in fact, there is a "smearing" between pixels when creating a composite signal. Knowing that the brightness of one pixel is the weighted superposition of the brightness of each of the neighboring pixels plus a nominal brightness, a new matrix can be found that accounts for this neighboring effect form the correlation matrix.

A square Laplacian matrix can be composed to "weight" the value of each of the neighboring pixels contribution to the overall pixel brightness. The number of elements of one row should be made proportional to the distance from the detector to the fluorescent emission. This can be interpreted to mean that the further away the object is from the detector, the less area of the CCD actually sees this emitted light from the object, and the less the contribution of neighboring

pixels to the pixel intensity. If the object is very close to the detector, then most all of the CCD pixels see the same light, and the more of a contribution that the neighbor pixels make.

For example, consider a 3x3 chunk of the correlation array, $f(i,j)$, and a Laplacian of weights, $W(i,j)$:

$$f(i,j)=\begin{bmatrix} f_{00} & f_{10} & f_{20} \\ f_{01} & f_{11} & f_{21} \\ f_{02} & f_{12} & f_{22} \end{bmatrix} \tag{4.103}$$

$$w(i,j)=\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{4.104}$$

A new value for the center correlation coefficient, f_{11}, can be found:

$$f'_{11} = \sum_{m=-1}^{m=+1} \sum_{n=-1}^{n=+1} f(m+1,n+1)W(m+1,n+1) \tag{4.105}$$

This new factor can then be normalized by the factor ε :

$$\varepsilon = \sum_{m=-1}^{m=+1} \sum_{n=-1}^{n=+1} W(m+1,n+1) \tag{4.106}$$

This operation is then performed for each element of the correlation matrix, with the result a matrix that accounts for the brightness of the neighboring pixels.

Figure 4.11 is a block diagram of the Perturbation/Correlation method.

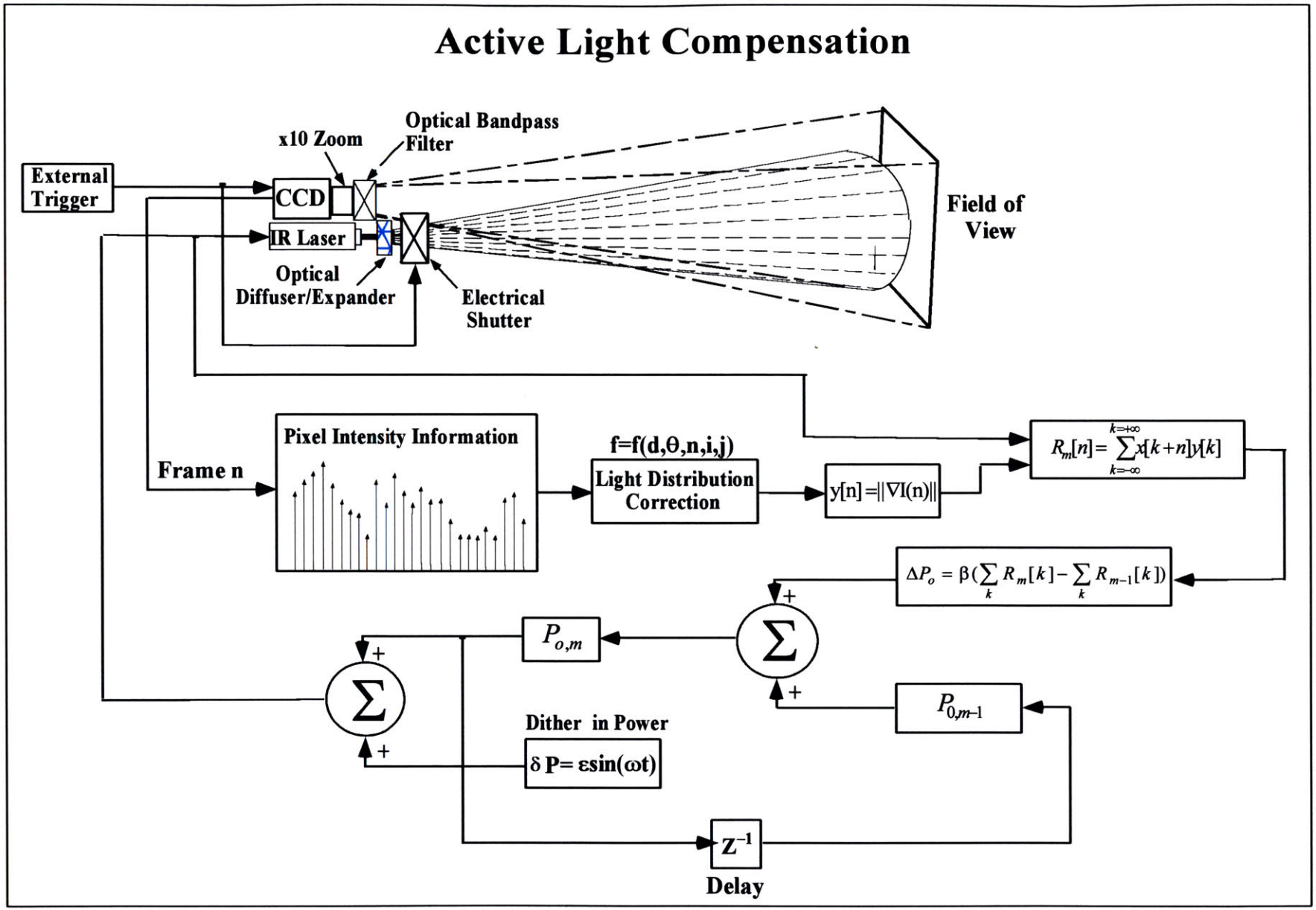


Figure 4.11: The Perturbation/Correlation Method

4.2 Verification of the Perturbation\Correlation Method

Within the software developed for this identification system, the perturbation/correlation was a method that was applied to "pre-process" a region of interest (ROI) when the noise in the ROI was of the same order as that of the signal. The signal, in this case, was a 2D bar code that used a four by four matrix of ones (an ink dot) and zeros (no dot), with three reference dots. The bar code was a 0.325" (8.25 mm) square without the parity bits and calibration patch, and 12 mm with these features included. The ink dot drops were on the order of 15 mils.

It should be noted that the light source was a 30 mW laser diode emitting at 785 nm, but for the use of this type of active image processing technique, any light source could be used that could be modulated (not with the given ink, though).

The laser power perturbation was produced by feeding a sinusoid in current to the laser diode driver. This AC current was superimposed on a nominal DC current that was used to forward bias the diode. The resulting power from the diode in the form of photon emission at 785 nm was proportional to this current.

$$I = I_o + A \sin(2\pi n / s) \tag{4.201}$$

I_o is the DC current value, A is the magnitude of the sinusoid perturbation in current, n is the discrete sample number, and s is nothing more than a frequency divider. For the lab prototype, a value of 20 was used for s.

It should be noted that for very long camera-to-object distances, a sinusoidal variation in laser power will still work, but the intensity magnitudes are so tiny that it is impossible to detect the variation with the CCD camera without some conditioning optics such as a zoom lens. Also, the software searches for an appropriate value of I_o by looking for the largest changes in correlation magnitudes from one frame process event to another. This, in effect, adapts to changing distances to the viewable object, as well as changes in the characteristics of the forward bias level of the diode.

Twenty CCD frames were taken from a trial run of this method in the prototype software, and are included in appendix A. Frame 5 corresponds to the frame grabbed with the least laser power out in the perturbation sequence, and frame 15 corresponds to the frame that was grabbed with the most laser power out. Figures 4.21 and 4.22 depict various views of the sensed intensity for frame 5 and frame 15. These figures should be compared carefully to appreciate the "noise" problem that results. Noise in this setup can come from reflective surfaces, objects that contain a similar color makeup spectrum, or light sources. Since the ink that was used for this experiment fluoresces at about 830 nm, which is near IR, there are many items that can fall within the same bandwidth and confuse the read. Tungsten filament lamps and sunlight reflections are the most common sources, since they have a wideband spectral emission.

The actual bar code "signal" is the sharp peaks that can be seen clearly in the middle of the frame, with the outlying edges containing the noise. The side view of Figure 4.21 clearly shows the problem that is faced in reading the bar codes in distinguishing the bar code from the surrounding reflective, since the signal and the noise are of the same order.

In the beginning, a correlation was made between the perturbation sequence and the intensity information; a sample of this can be found in Figure 4.23. As can be seen, the signal to noise ration is greatly enhanced, but noise is still a problem when setting thresholds. For this reason, in the final version of the prototype software, the change in intensity was used in the correlation.

This spatial intensity gradient information can provide a better signal to noise ratio since it basically finds the edges with the frame. Figures 4.24 and 4.25 show various views of the spatial gradient of intensity for frame 5 and frame 15, respectively. The gradient of the intensity function can be written as follows:

$$\nabla I = \sqrt{(\partial I/\partial x)^2 + (\partial I/\partial y)^2} \angle \tan^{-1}((\partial I/\partial y)/(\partial I/\partial x)) \qquad (4.202)$$

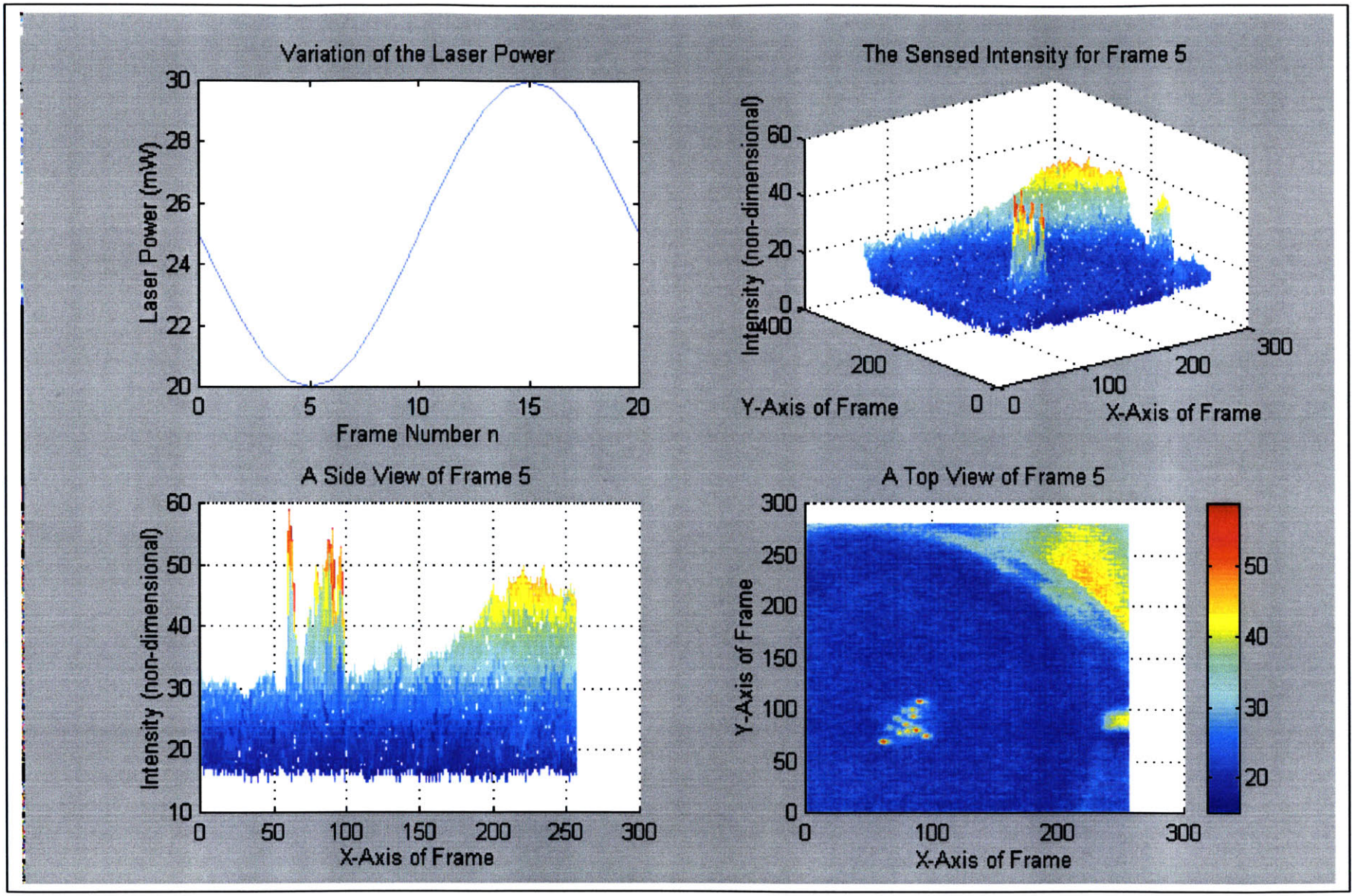


Figure 4.21: The Sensed Intensity for Frame 5

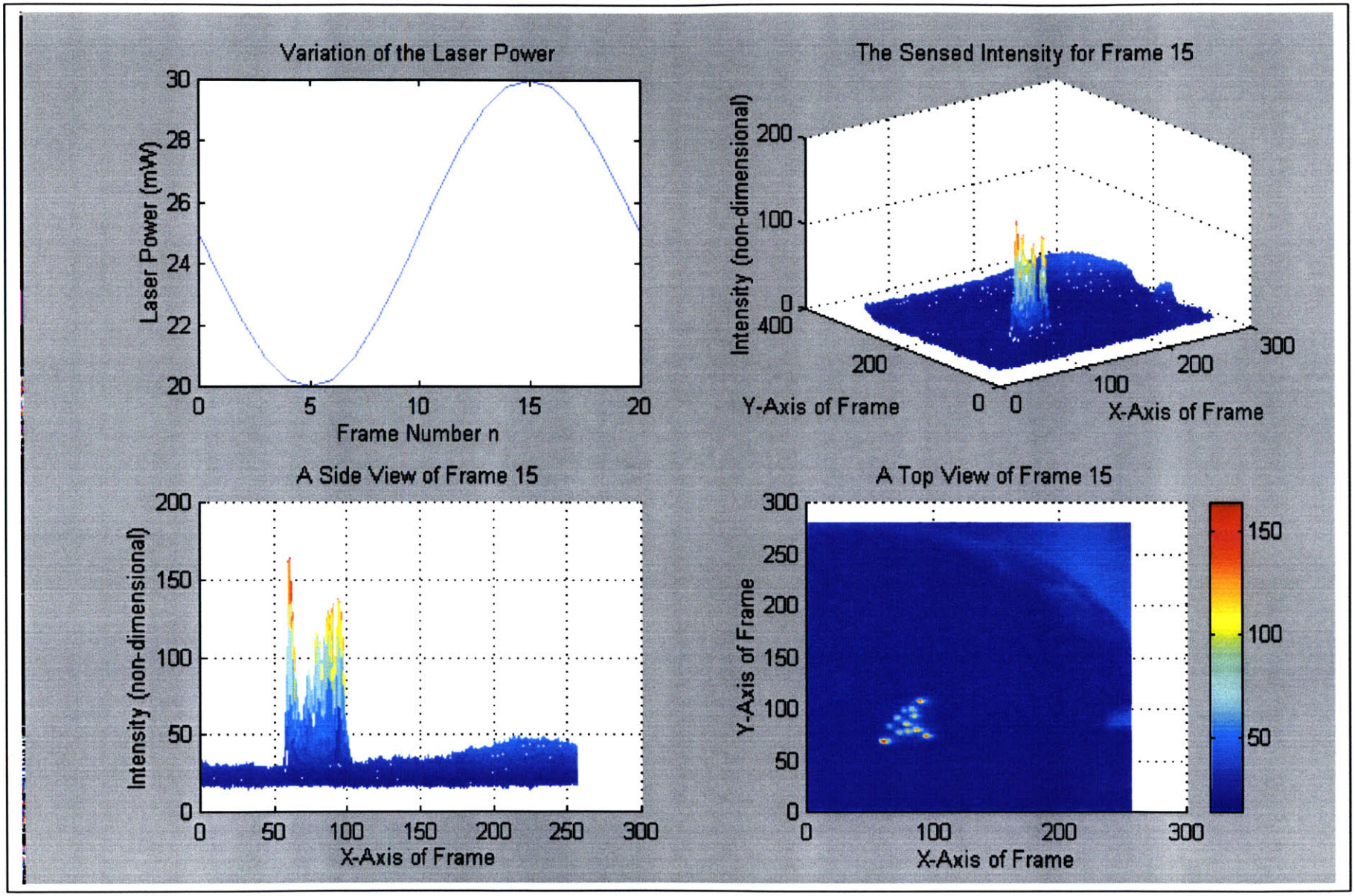


Figure 4.22: The Sensed Intensity for Frame 15

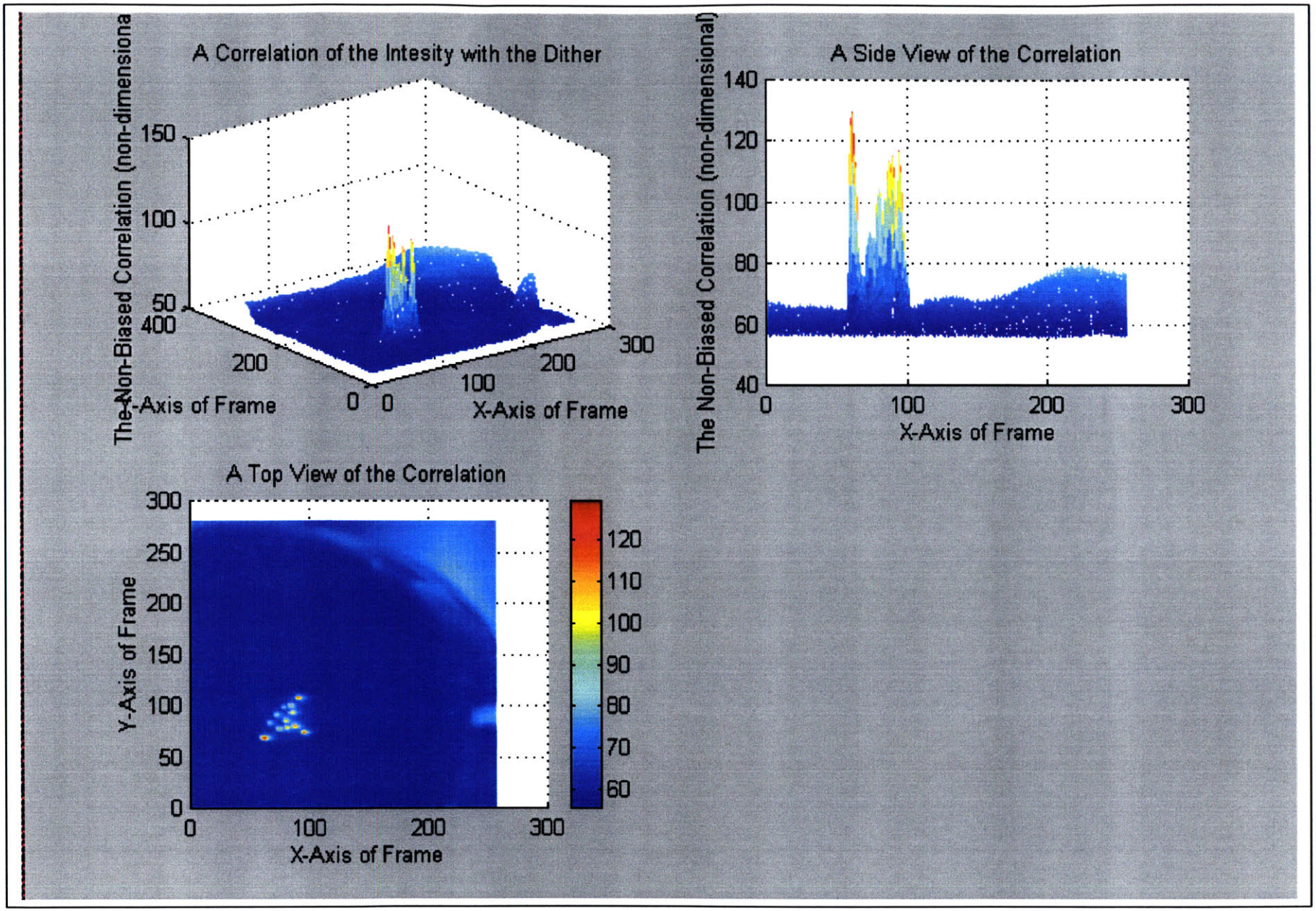


Figure 4.23: The Perturbation to Sensed Intensity Correlation

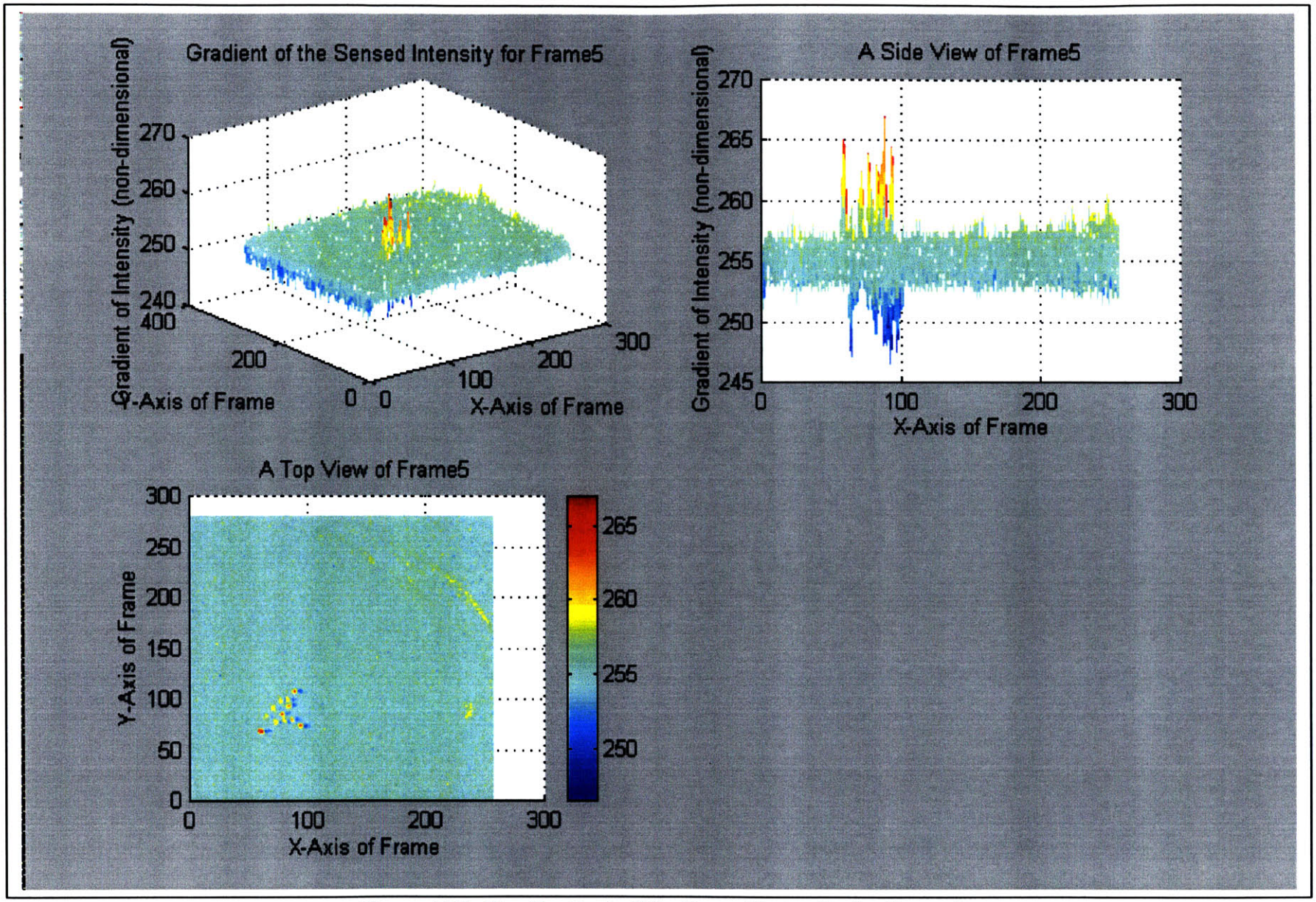


Figure 4.24: The Gradient of the Sensed Intensity for Frame 5

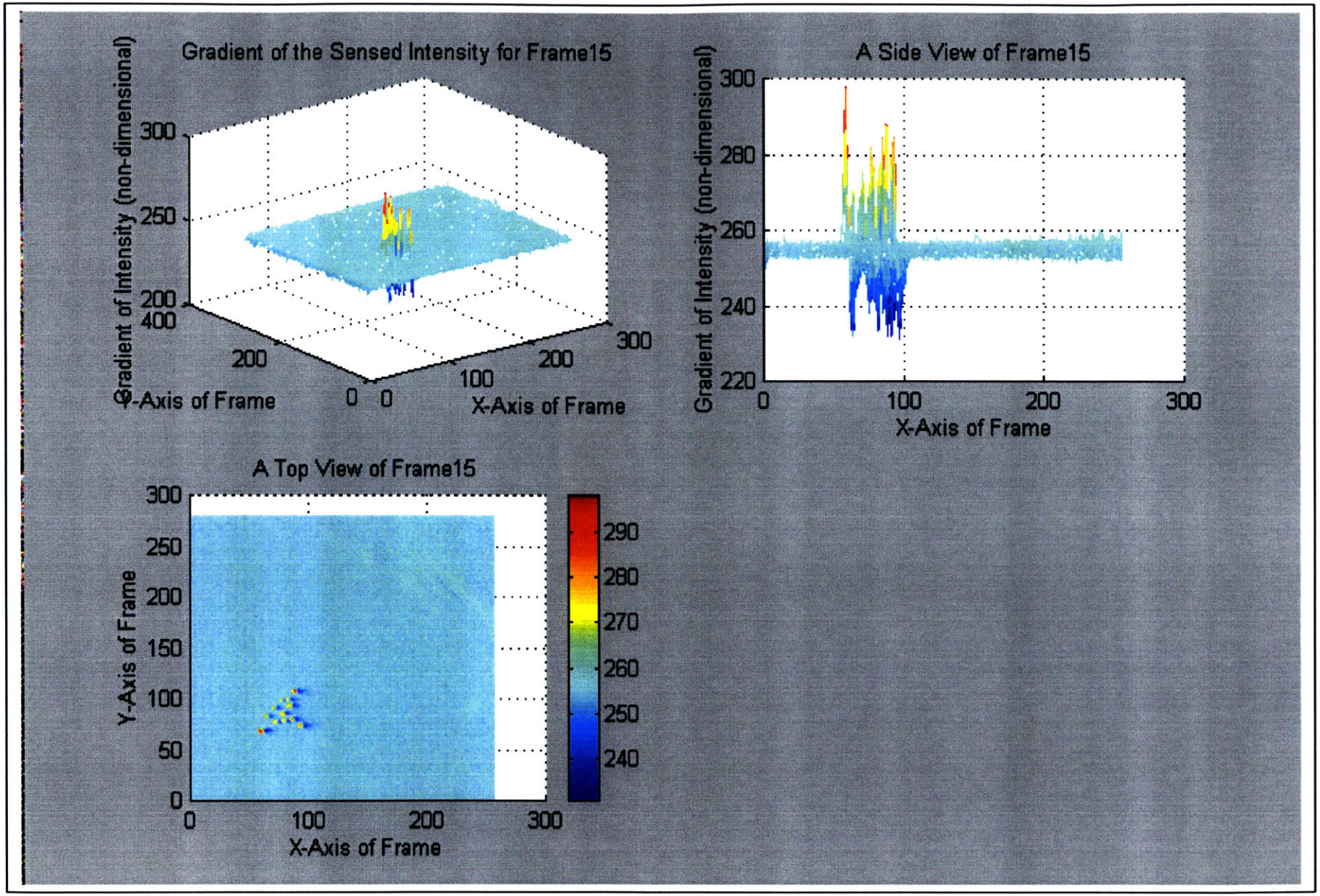


Figure 4.25: The Gradient of the Sensed Intensity for Frame 15

$$\frac{\partial I}{\partial x} = \frac{I(x+h) - I(x)}{h} - \frac{I(x) - I(x-h)}{h} \tag{4.203}$$

$$\frac{\partial I}{\partial y} = \frac{I(y+m) - I(m)}{m} - \frac{I(y) - I(y-m)}{m} \tag{4.204}$$

This provides some insightful information in that the biggest changes in the frame should occur within the bar code ink area, since this ink is "selectively" stimulated by using the laser wavelength (785 nm) that is known to fluoresce the ink. The correlation used in the prototype software was between the laser perturbation and the change in sensed intensity at the CCD camera. The correlation function [61], in discrete time, can be written as:

$$\phi[n] = \sum_{k=-\infty}^{+\infty} x[n+k]y[k] \tag{4.205}$$

In this case, $x[n]$ is the laser power, and $y[n]$ is the change in sensed intensity (gradient) function for each pixel over the discrete sample period. This correlation was performed for each pixel over 20 frames, and the results can be seen in the various correlation views depicted in Figure 4.26. As can be seen, the noise is greatly reduced, giving a significant increase in the signal to noise ratio.

As a side note, a monochrome CCD camera will give a "white" color corresponding to an area of great intensity ($\frac{R+G+B}{3}$), such as what is found at the ink surface, and a "black" color to areas of low intensity. These facts should be kept in mind when interpreting gradient information with regard to intensity change and the direction of change. The bar code ink corresponds to the "peaks" of the gradient function, since the ink normally has a higher intensity than that of its local environment, although globally it may not.

It should be noted that for this run of the prototype software, the bar code and the CCD camera frame of view were made parallel to one another. The CCD camera and bar coded object were kept at a fixed distance from each other, and the camera was focused to the bar code depth. Other than that, no attempts were made to improve the quality of frame information that was to be taken.

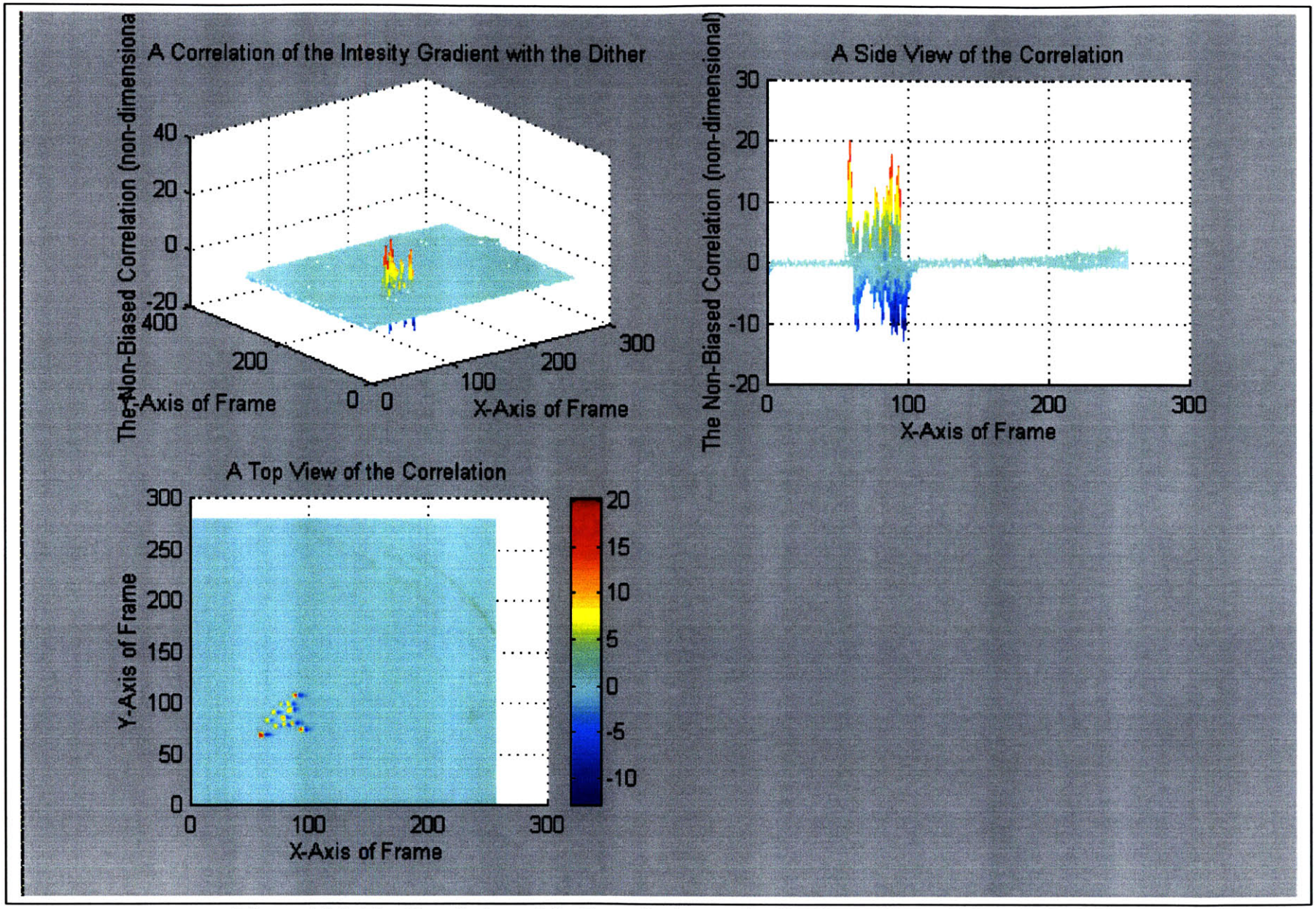


Figure 4.26: The Correlation of the Perturbation to the Gradient of Sensed Intensity

A little information should be given about the software that was developed and used in the above mentioned bar code read. Since the correlation is a computationally intensive operation, a basic methodology has been developed for giving a bar code read of reduced computation. Step one is to compare images when the lamp is on and when it is off through intensity subtraction. Figure 4.27 give various views of this on/off subtraction. As can be seen, the area that contains the bar code can be seen clearly, although the side view shows that there is still a lot of noise in the frame. A window is then set about this bar code area, reducing the number of pixels to be considered.

Step two consists of the perturbation in the laser power, and the corresponding correlation. This step "pre-processes" the frame by removing as much of the noise as possible before any attempt is made to read the bar code. The third step is the read of the "Asymmetric Calibration Patch" bits. The value of the bits in this patch are known already, and are placed at the extreme distance from reference 0, since the spatial error in calculating the read points propagates outward from this point. This is an intuitive fix in that if the extreme edge bits can be read correctly, then the probability of a correct read for the internal four by four matrix bits is increased.

At this point, the software threshold is raised or lowered until a correct read of the calibration patch is found. A range of thresholds are found that give a 100% correct calibration patch read, and the middle threshold value is picked. With this complete, the geometric information about the bar code is used to read the internal four by four matrix, and this information is used in the row and parity checks. If the read is error free, the read is complete; if not, the DC value of the laser current is adjusted and the same process is reapplied to the same area. It should be noted that the first software threshold (the base threshold) is picked by setting a threshold on a histogram made of the intensity gradient to perturbation correlation information. Figure 4.28 gives a software map of the prototype software.

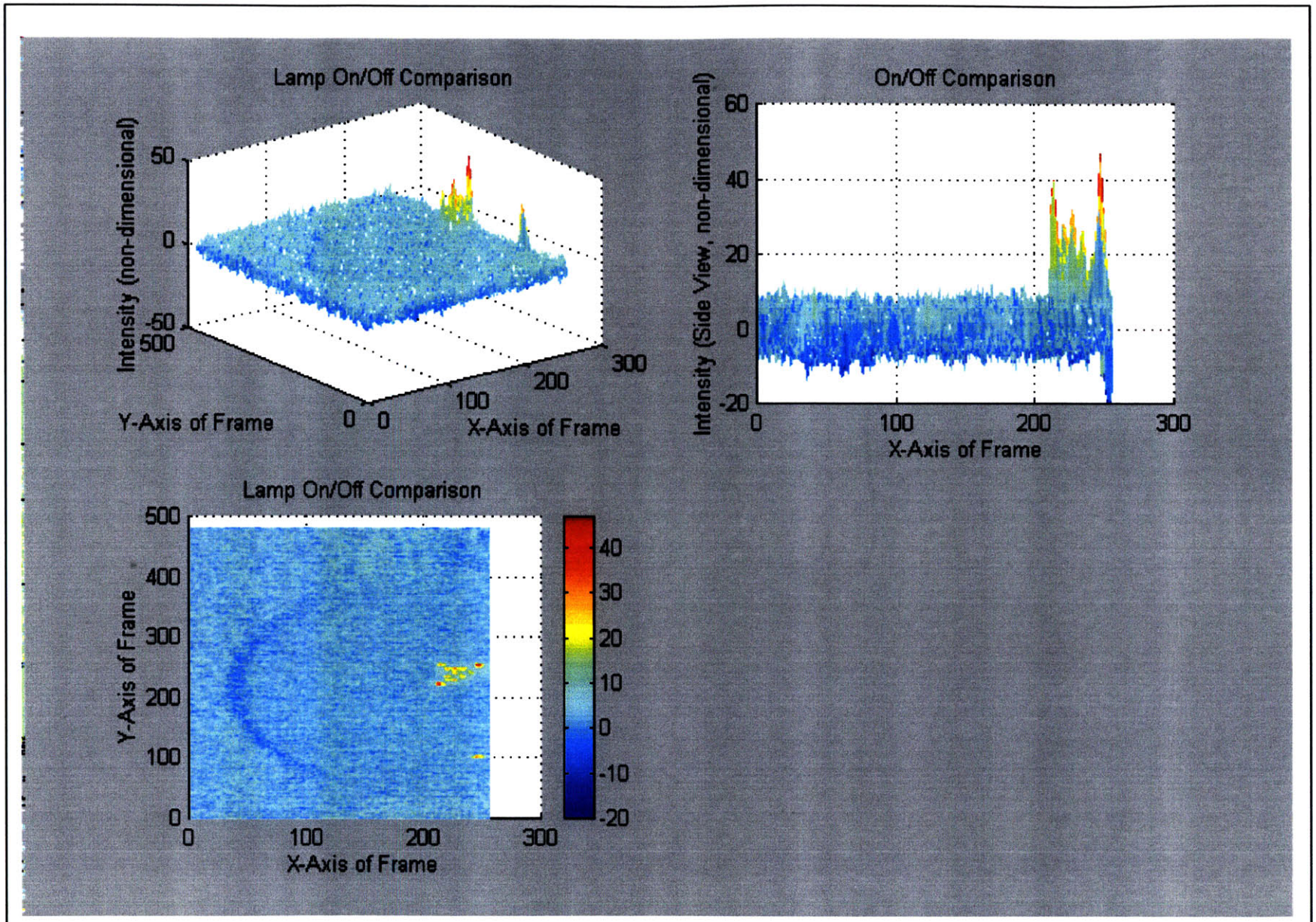


Figure 4.17: An On/Off Comparison for ROI Determination

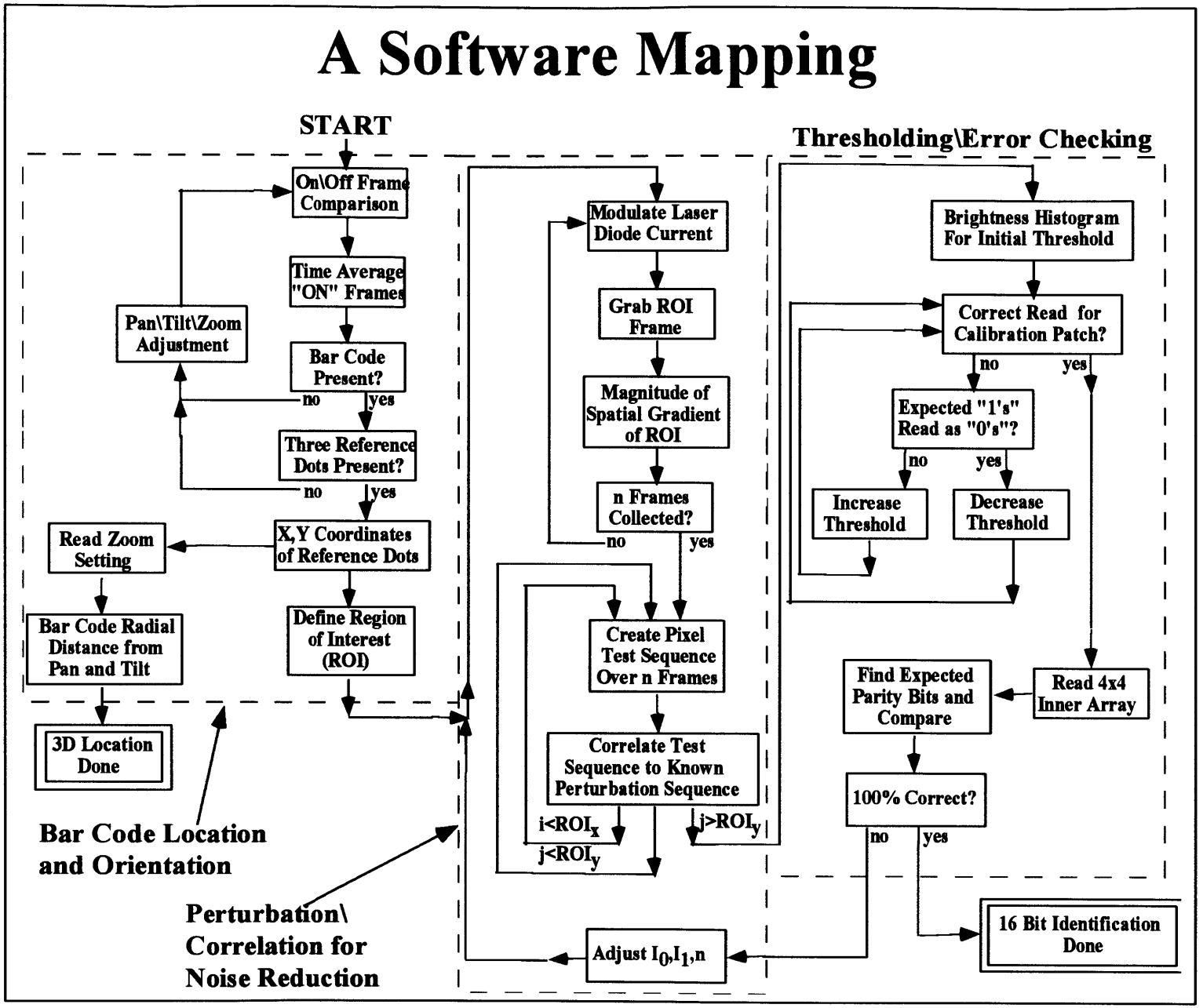


Figure 4.18: A Software Mapping

4.3 The Perturbation/Correlation Method and Radiation Heat Transfer

A closer look is taken at the information that is actually measured by the CCD camera in determining the bar code number. Since light and heat are often analogous quantities, a simple analogy is drawn from the experimental setup measuring reflected light intensity to a heat radiation problem. The input perturbation in lamp intensity and the corresponding sensed frame intensity of the CCD detector can be analyzed in the following manner (See Figure 4.30: An Analogy to Radiation Exchange). The lamp intensity is on the order of the lamp power:

$$P \sim I \tag{4.306}$$

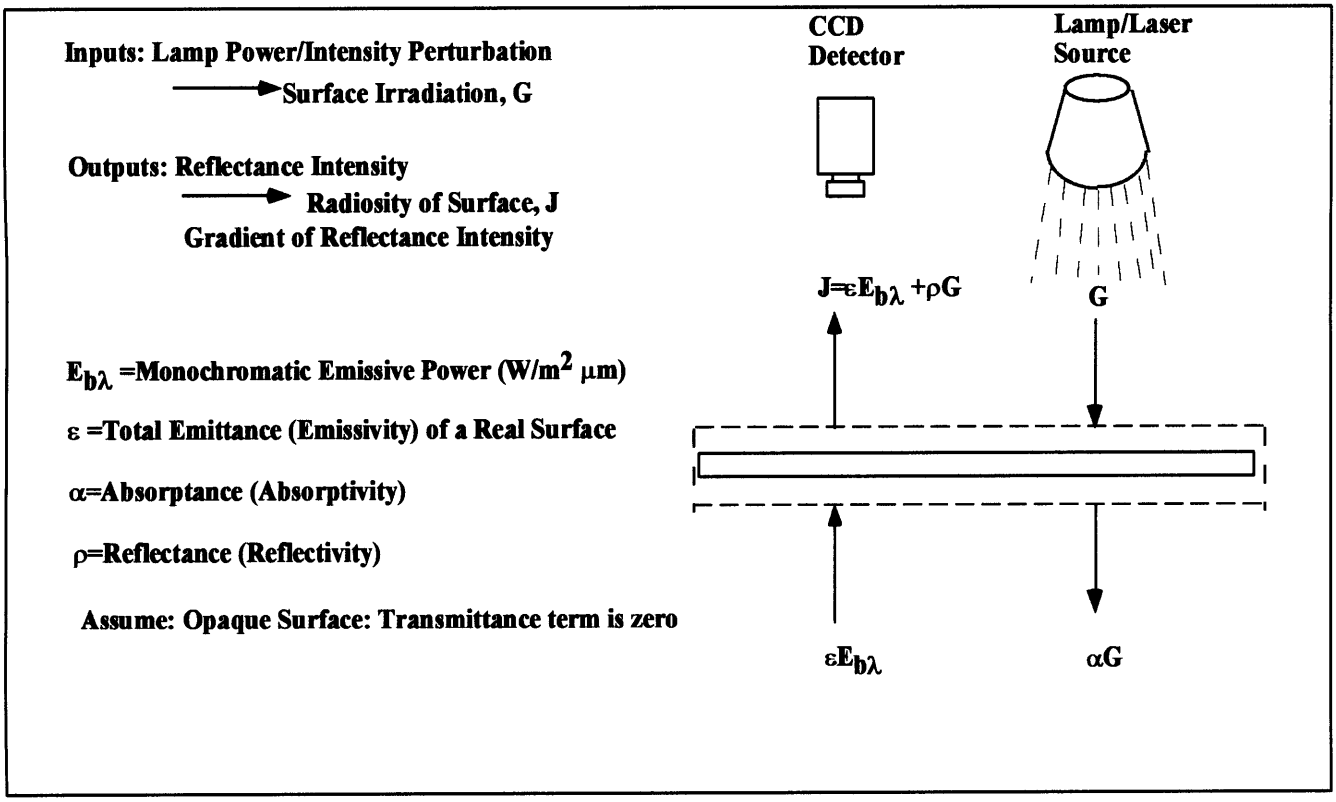


Figure 4.30: An Analogy to Radiation Exchange

If G is defined as the surface irradiation flux, then it can be seen that, for small irradiated areas, that the lamp intensity is on the order of the surface irradiation:

$$G \sim I \tag{4.307}$$

The fraction of this irradiation that is absorbed in the object surface is just the material absorptivity, α, multiplied by the irradiation flux.

The material has a certain radiant emissive energy, $E_{b\lambda}$, which is a function of wavelength and temperature, and is usually referenced about the "black body" emissive power through the emissivity constant, ε, which is specific to the material. The portion of this energy that is transferred to the atmosphere is just εE_b, where E_b is the black body emissivity.

In addition, some of the irradiation energy is reflected back from the surface, and the reflectivity, ρ, is used to quantify this energy amount. The sum of the reflected irradiant energy flux and the material specific emissive energy flux is the surface radiosity flux, J:

$$J = \varepsilon E_b + \rho G \tag{4.308}$$

Since a perturbation in lamp intensity is given, which is proportional to the surface irradiation at a given wavelength, a perturbation in surface irradiation can be defined:

$$G' = G + \delta G \tag{4.309}$$

From this relation, a radiosity flux as a function of this perturbation can be given:

$$J' = (\varepsilon E_b)' + \rho(G + \delta G) \tag{4.310}$$

The wavelength of the perturbation source remains fairly constant, so:

$$(\varepsilon E_b)' = (\varepsilon E_b) \tag{4.311}$$

Substituting equation 4.311 into equation 4.310, and subtracting equation 4.308 from it, we obtain:

$$J' - J = \rho(\delta G) \tag{4.312}$$

The perturbation in irradiation of the surface is known, and a dimensionless measure of the relative radiosity of the surface can be made. From this, and a dimensional scaling factor, *a*, the reflectivity of the surface can be evaluated at a given lamp source wavelength:

$$\rho = \frac{a(J' - J)}{\delta G}\Big|_{\lambda=\lambda_L} \tag{4.313}$$

This reflectivity parameter is fairly important in that there is a reflection component to the bar code signal as well as to the overall noise. In correlating the gradient of the intensity to the input

perturbation in lamp intensity, the largest magnitude reflectivity coefficients for the frame are preserved over time.

Now, from the work in section 2.4, it is possible to make the following equalities:

$$\delta G = \xi_{t2} - \xi_{t1} \tag{4.314}$$

$$J' - J = \xi_{R2} - \xi_{R1} \tag{4.315}$$

The interpretation of this is that the change in surface irradiation is simply the difference in transmitted photons over two frames, and the change in radiosity of the surface over two grabbed frames is the difference in photons returned from the surface. Dividing these two quantities, a quantum efficiency for the surface, $\eta_S(\lambda, x_n, y_n)$, may be obtained, where x_n and y_n are the horizontal and vertical pixel frame coordinates:

$$\eta_S(\lambda, x_n, y_n) = \frac{\xi_{R2} - \xi_{R1}}{\xi_{t2} - \xi_{t1}} \tag{4.316}$$

It can be seen that if this is then evaluated at the known absorption wavelength of the ink for each pixel, then the ink is present at that particular pixel when:

$$\eta_S(\lambda_a, x, y) \approx \eta_i(\lambda_a) \tag{4.317}$$

It should be noted that the constant **a** in equation 4.313 accounts for the quantum efficiency of the zoom lens, camera optics, filter, beam expander, and the light amplifier. It also accounts for the distance **L** from the target to the detector, the ratio of areas of the beam to the bar code dot area, the beam divergence, the light amplification factor, the laser power, and the CCD characteristics. It does not account in any way for the presence of noise. The following is the equivalent value for **a**:

$$a = \left[\frac{t_s x_L y_L g_{pi} g_{iv} D_{AD} L_g}{2\pi L^2 hc P_H P_V R_{AD}}\right]\left[\frac{(\sum_{\lambda=0}^{\infty} P_s(\lambda)\eta_z(\lambda)\eta_c(\lambda-\lambda_S)\eta_f(\lambda)\eta_e(\lambda)\eta_{ap}(\lambda)\lambda) n d_3^2}{L_{MF}^2[2L\tan\theta + d_1][2L\tan\phi + d_2]}\right] \tag{4.318}$$

This approach uses information about a measured surface quantum efficiency and then attempts to match this to the known quantum efficiency of the given ink, evaluated at the emission wavelength of the ink. Since the ink is more sensitive to changes in lamp power at the absorption frequency, this approach works well.

Chapter 5: Noise Modeling and Measurement

As mentioned in section 4, a lot of work has been done with regard to signal to noise enhancement through the above mentioned Perturbation/Correlation method and signal time averaging, and very remarkable gains in S/N ratio have been realized. However, up until now, these methods have been applied blindly to both the large amplitude noise case as well as the low amplitude noise case, without any real means of capturing the characteristics of the noise in the viewable frame at any instant in time. With some knowledge of the noise makeup, it is possible to apply these image processing algorithms in a smart manner, as well as to find an optimal search path based on the probability of signal detection at any spatial location (if the signal were present).

5.1 Hardware Setup for Noise Measurement

Figure 5.10 shows a simple spectrum sampler that was constructed by using a bank of photo-diodes that were placed behind notch filters at different center pass wavelengths (refer to Figure 5.11 for specific filter information). The photo-diodes were chosen on the basis of the high gain requirement as well as the linearity of the spectral response for the photo-diode for wavelengths between 430 nm and 880 nm (the visible to the near IR range). Some signal conditioning was performed to change the small amplitude current, produced from incident radiation on the detector's surface, to a 0-5 V signal that could be read by the DAS 1800 A/D. Op-amps were used to provide isolation to the converted current to voltage signal, as well as to provide a "zero and span" for the calibration of the individual detection system. A requirement was found for this calibration since each diode had a different base voltage signal (voltage when there is no radiation on the detector surface). This problem was taken care of by using the zero portion of the circuit, and the different filter pass gains could be nulled by appropriate settings of the span circuit.

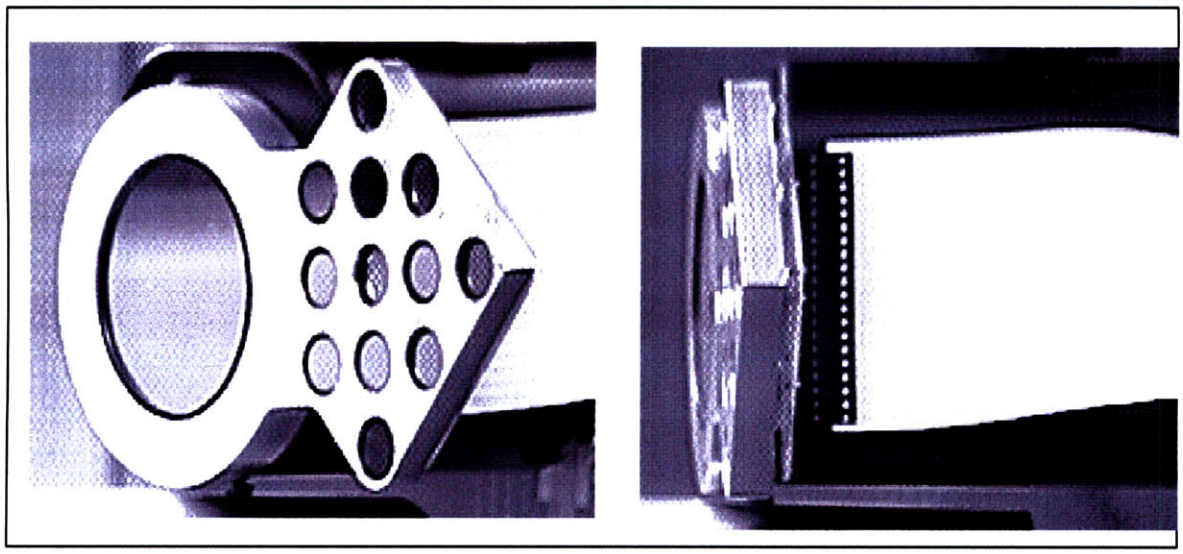


Figure 5.10:The Noise Measurement System

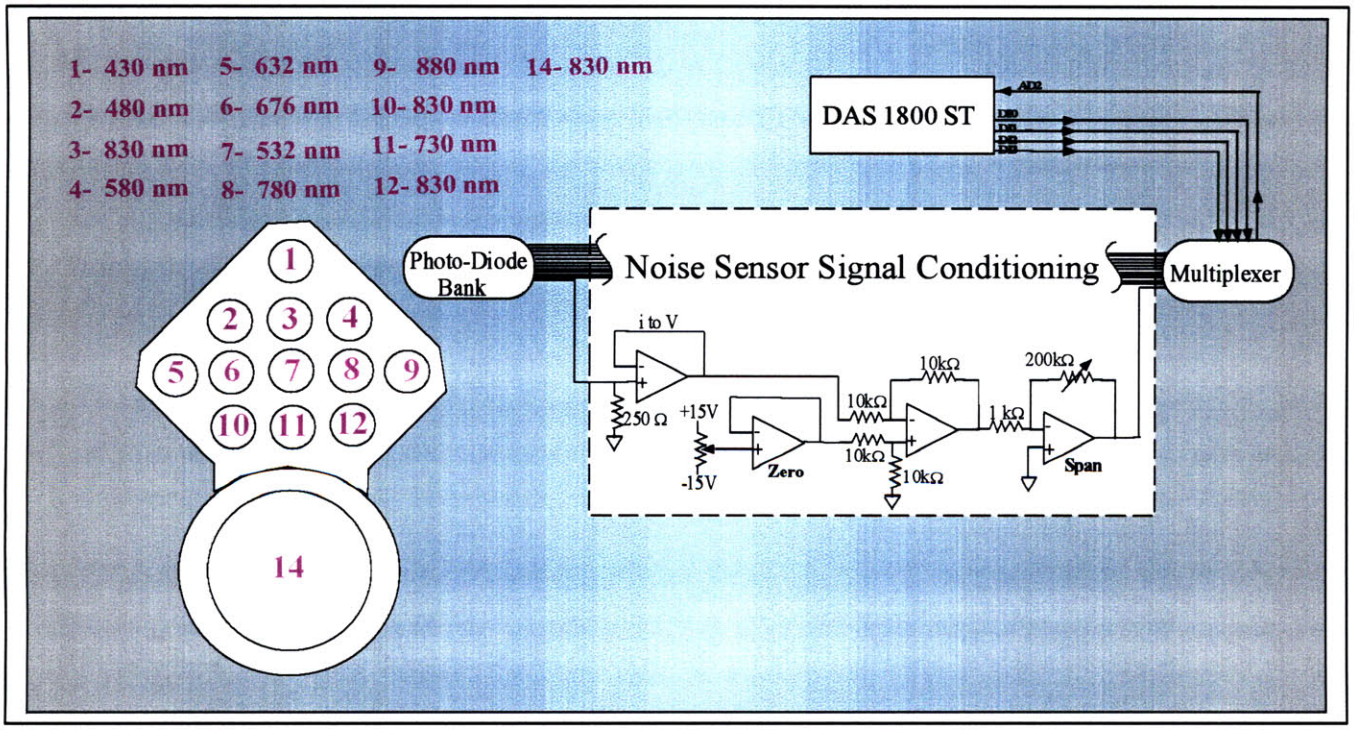


Figure 5.11: Signal Conditioning and Pass Ranges for the Spectrum Sampler

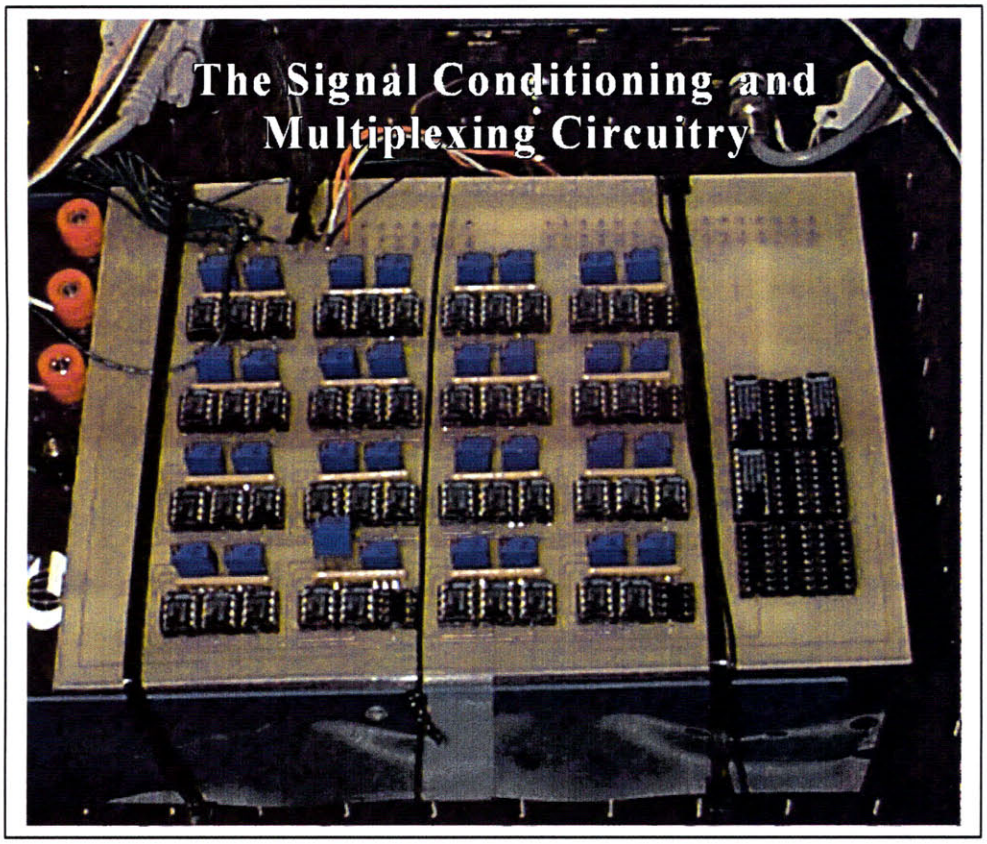


Figure 5.12: The Signal Conditioning and Multiplexing Circuitry

5.2 Calibration Procedure for the Noise Measurement System

The circuitry developed for the collection of the incident radiation at a particular wavelength was calibrated by using a narrow band light source of known intensity. Before the filter was placed in front of the photo-diodes, each diode was covered with black electricians tape and the zero circuit potentiometer was adjusted to give zero volts out for each photo-diode, measured at the output of the signal conditioning circuitry. Then, with the tape removed, a hollow tube was placed over the diode and the light source was shined upon the detector surface. The span potentiometer was then adjusted to give 5V at the output of the signal conditioning circuitry. This was done for each of the twelve diodes.

At this point, the calibration became rather intuitive. The feedback resistance of each span element was measured at the given calibration. Then, after checking the pass gain found from the spectral response curve for each filter, this gain was divided into 1 and multiplied by the span resistance measured. This gave a new resistance to which the span potentiometer was then set. This, in effect, compensated for the attenuation of the signal as it was passed through the

notch filter. By doing this, it was possible to equate a 5V signal to the wattage of the light source used in the calibration, but now at a given wavelength. From this calibration procedure, it was possible to develop curves that represented the noise in the viewable frame and find the power of this noise signal over a given wavelength range.

5.3 One Motivation for this Noise Power Estimation

Often, a simple, black body emission model is used to describe noise sources such as tungsten lamps and other broad band emitters. For instance, a good approximation for the tungsten element in an incandescent light bulb is a 1mm black body emitter at 3000K. For these particular black body approximations, a relationship for the monochromatic emissive power for a black surface is used:

$$E_{b\lambda} = \frac{C_1 \lambda^{-5}}{e^{C_2 (\lambda T)^{-1}} - 1} \tag{5.30}$$

Where C_1=3.742x10^8 Wμm^4/m^2 and C_2=1.4389x10^4 μm K [67], and $E_{b\lambda}$ is in [W/(m^2μm)]. Figure 5.30 gives a plot of the black body emissive power as a function of wavelength and at five discrete temperatures.

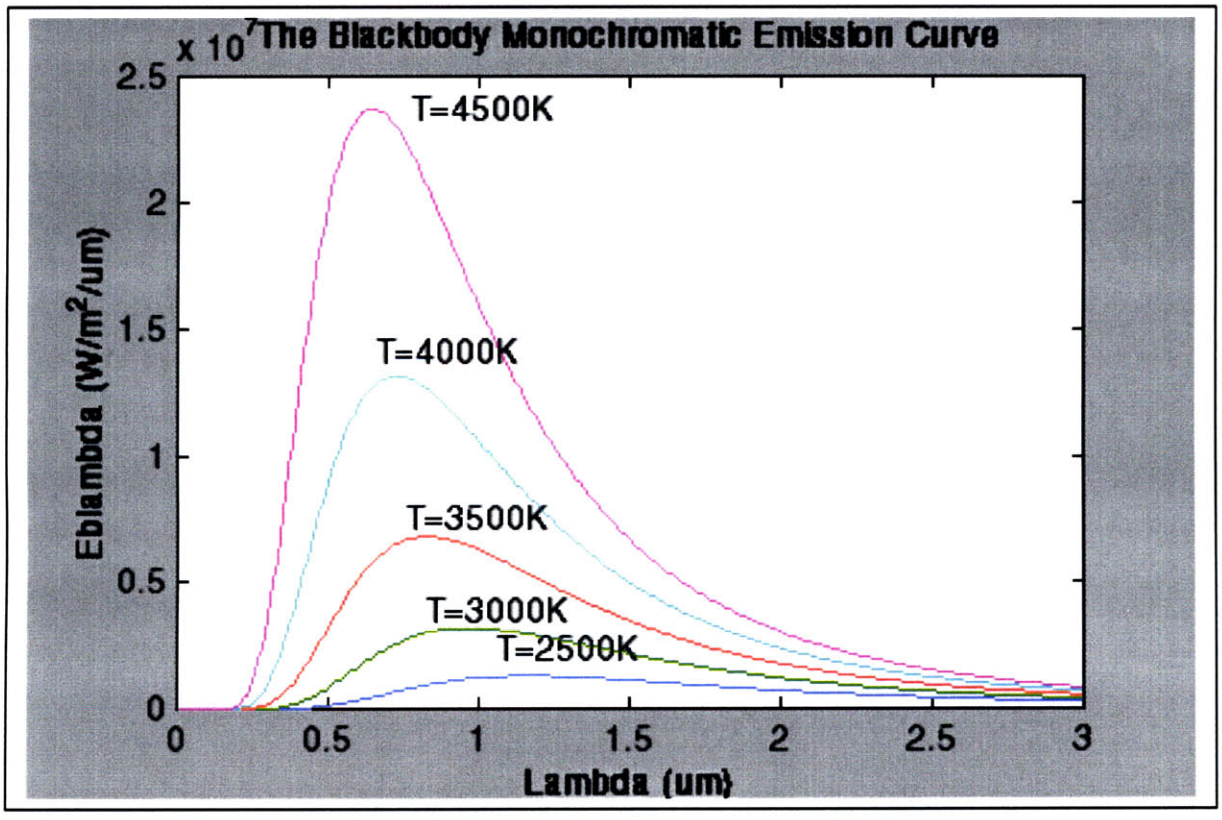


Figure 5.30: The Black Body Emission Curve for Varying Body Temperatures

The total emission of a black body over all wavelengths is:

$$E_{bb} = \sigma T^4 \tag{5.31}$$

The amount of power contained between zero and some wavelength, λ, can be found by integrating equation 5.30 with respect to lamda, and the percentage of total power found in this range can be found by dividing by the total black body emissive power. This new quantity is referred to as the external fractional function (Mills):

$$f_e(\lambda,T) = \frac{\int_0^{\lambda} E_{b\lambda}(T,\lambda)\,d\lambda}{\sigma T^4} \tag{5.32}$$

This fractional function is shown in Figure 5.31

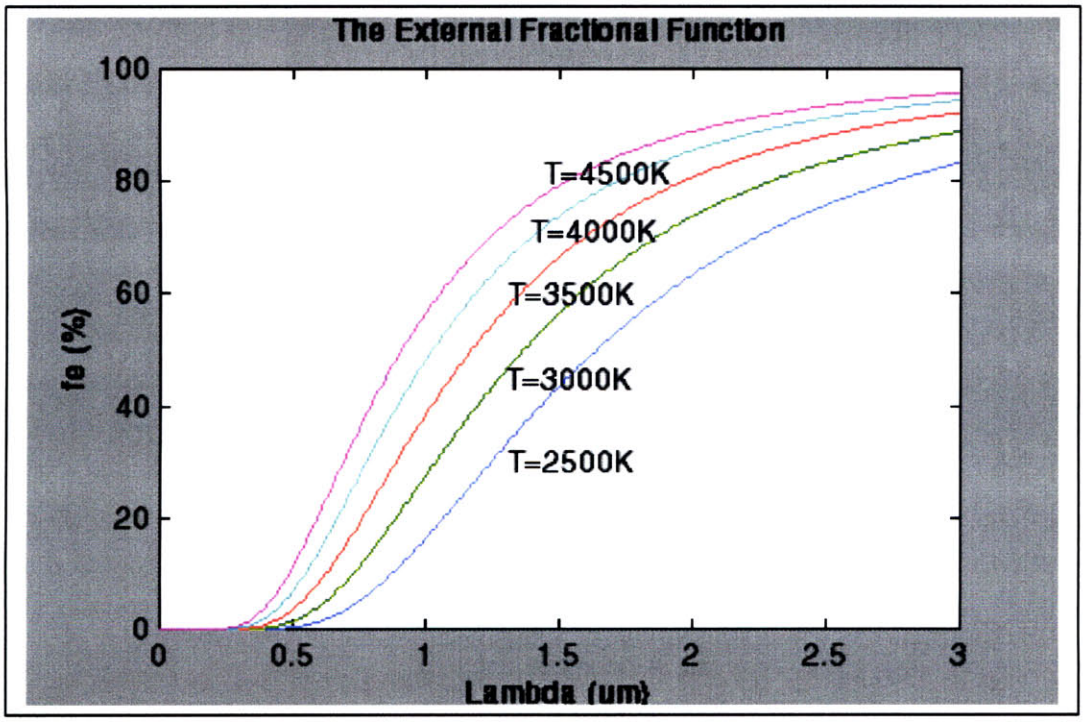


Figure 5.31: A Plot of the External Fractional Function For Varying Body Temperatures

In the case of noise measurement in a specific band, what is required is the emissive power between two wavelengths, which is obtained by first subtracting the fractional function at the bottom wavelength from the fractional function of the top wavelength:

$$\Delta f_e = f_e(\lambda,T)_{top} - f_e(\lambda,T)_{bottom} \tag{5.33}$$

This quantity is then multiplied by the total blackbody emissive power:

$$\Delta f_e \sigma T^4 = \int_0^{\lambda_{top}} E_{b\lambda}(T,\lambda)\,d\lambda - \int_0^{\lambda_{bottom}} E_{b\lambda}(T,\lambda)\,d\lambda \tag{5.34}$$

which can also be written as:

$$\Delta f_e \sigma T^4 = \int_{\lambda_{bottom}}^{\lambda_{top}} E_{b\lambda}(T,\lambda)\,d\lambda \tag{5.35}$$

This quantity now represents the power contained between two wavelengths for a noise source modeled as a black surface and at a given temperature. The two wavelengths, in this case, are the bottom and top cutoff wavelength for the notch filters. However, this does not give the amount of radiation incident at the detector, but rather the total power flux emitted into the environment between two wavelengths.

To find the actual radiation that is incident upon the surface of the photo-diode detector, the total power flux from the noise source must be multiplied by a view factor. By modeling the noise source to detector combination as two parallel coaxial disks (See Figure 5.32), a view factor may be found.

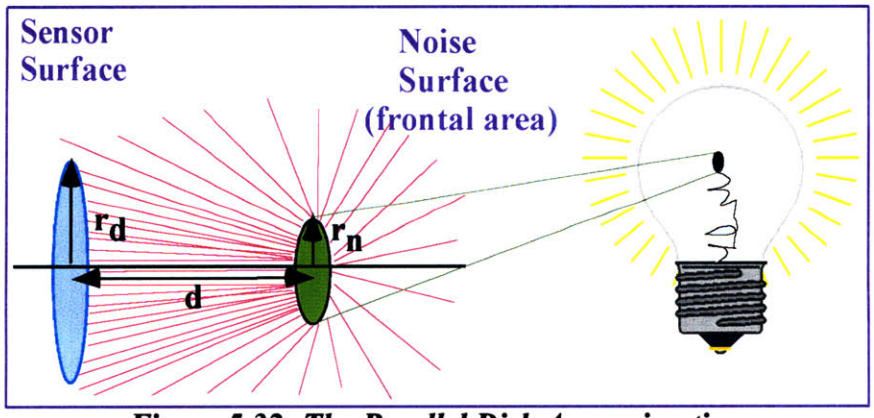


Figure 5.32: The Parallel Disk Approximation

$$F_{nd} = \frac{1}{2}\left[\left[1+\left[\frac{1+\left(\frac{r_n}{d}\right)^2}{\left(\frac{r_d}{d}\right)^2}\right]\right] - \left[\left[1+\left[\frac{1+\left(\frac{r_n}{d}\right)^2}{\left(\frac{r_d}{d}\right)^2}\right]\right]^2 - 4\left(\frac{r_n}{r_d}\right)^2\right]^{\frac{1}{2}}\right] \tag{5.36}$$

It should be noted that $\mathbf{r_n}$ is the radius of the noise source, $\mathbf{r_d}$ is the radius of the photo-diode detector surface, and **d** is the distance from the noise source to the detector. Although the tungsten lamp filament is spherical in nature, only the frontal area of the filament is seen by the detector, so the coaxial disk approximation is justified. The main point to be made here is that the view factor is a strong function of the distance from the sensor surface to the noise source surface. Figure 5.33 illustrates the view factor as a function of **d** for varying ratios of the sensor radius to the noise source radius.

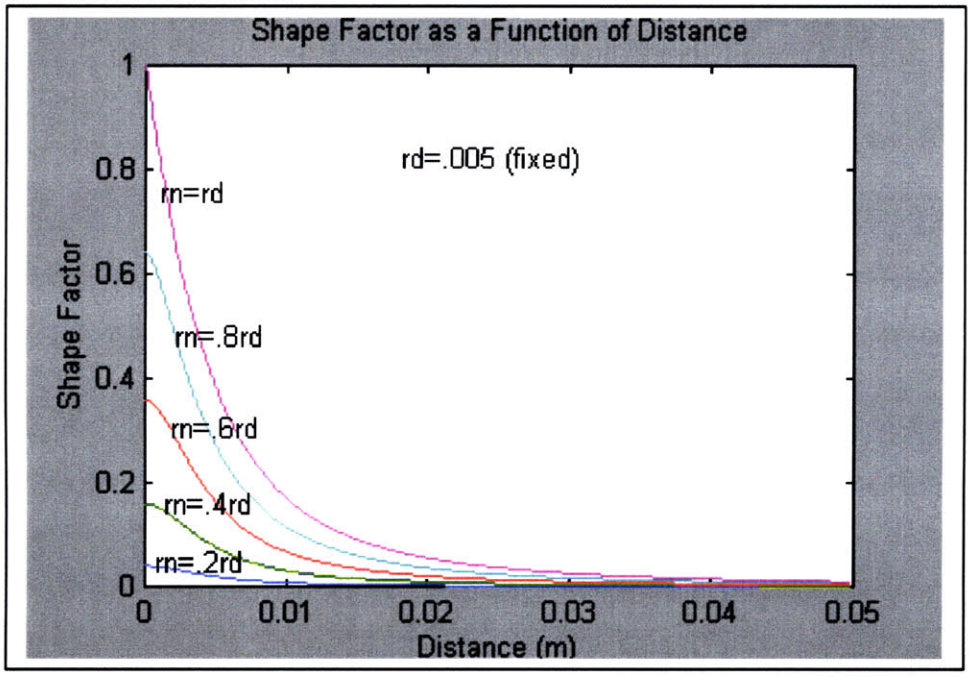


Figure 5.33: The View Factor as a Function of Surface to Surface Distance

Assuming a general form for the transfer function for the photo-diode as $G_d = G_d(\lambda)$ and for the filter as $G_f = G_f(\lambda)$, then the total power flux incident upon the photo-diode detector is:

$$P_{measured} = \frac{1}{2} G_d(\lambda) G_f(\lambda) \left[\left[1 + \left[\frac{1+\left(\frac{r_n}{d}\right)^2}{\left(\frac{r_d}{d}\right)^2} \right] \right] - \left[\left[1 + \left[\frac{1+\left(\frac{r_n}{d}\right)^2}{\left(\frac{r_d}{d}\right)^2} \right] \right]^2 - 4\left(\frac{r_n}{r_d}\right)^2 \right]^{\frac{1}{2}} \right] \int_{\lambda_{bottom}}^{\lambda_{top}} E_{b\lambda}(T,\lambda)\, d\lambda \quad (5.37)$$

This quantity is actually a power flux. To get the actual power received by the detector, then the area of the detector must be multiplied by the flux. It should be noted that the measured power is very sensitive to the distance **d** from the noise source to the detector. However, since each detector is mounted very close to one another on the mount, the radiation must travel

approximately the same distance to each detector. This approximation breaks down at extremely low distances in which the distances are not of the same order for each detector, as well as large angles of rotation of the detector mount. Also, at very long distances, the view factor is very tiny and the measured radiation is very small in magnitude, which gives a breakdown in this noise measurement system.

The motivation of this exacting measurement of noise is that it makes it possible to get an idea of the overall wavelength makeup of the noise in the viewable frame. If the noise characteristics follow that of a blackbody surface at a given temperature, then the noise is of the repeatable type, and the location of this noise source can be determined as well as a basic classification for the noise. Also, if the noise is repeatable in nature, the Perturbation/Correlation method works very well to remove this noise. If it doesn't follow the black body surface characteristics, then it is random, dark current noise, and can be effectively removed by just time averaging the frames at this given spatial location. In addition, with some information about the distance from the sensor surface to the noise source, which could be found from a ranging sensor (such as a sonar based device), the actual power of the noise source may be determined and compared to the known power of various noise sources.

5.4 Explanation of Measurements Made by the Noise Detection System

Figures 5.40 and 5.41 are actual plots of the noise measured by this detection system. The measurements are of power flux incident on the detector, with the 0-255 representing a 0 to P_{MAX} of the calibration light source. Figure 5.42 then plots the noise, as a function of wavelength, for two different spatial locations; one with a noise source present and one without. This basically gives the spectral response curve for two different spatial locations; one that contains large amplitude, repeatable noise; and the other, low amplitude dark current or random noise. Figure 5.43 is a weighted spectral response for the viewable frame. More weight is given to the noise that falls in or near the band of the notch filter (825-835nm), and less weight is given to the noise well outside the notch filter range. A gaussian function is used in picking the weights, with the center of this corresponding to the center of the notch pass range. This gaussian multiplier determines the net weight of each noise measurement at a given wavelength in the overall accumulation, much like what the notch filter does in hardware. This is in effect is a "penalty schedule" for the viewable frame, and can be used in the design of search patterns for a particular bar code. The probability of detection, given the presence of a bar code, goes down with an increase in elevation of the penalty schedule, so the intuitive search would be the shortest path to connect all areas of equal and least elevation. The following is the penalty schedule, where N(λ, x, y) is the actual noise measurement:

$$PS = \sum_{\lambda_{bottom}}^{\lambda_{top}} N(\lambda, x, y)(.6e^{-((\lambda-830)^2/20000)}) \qquad (5.40)$$

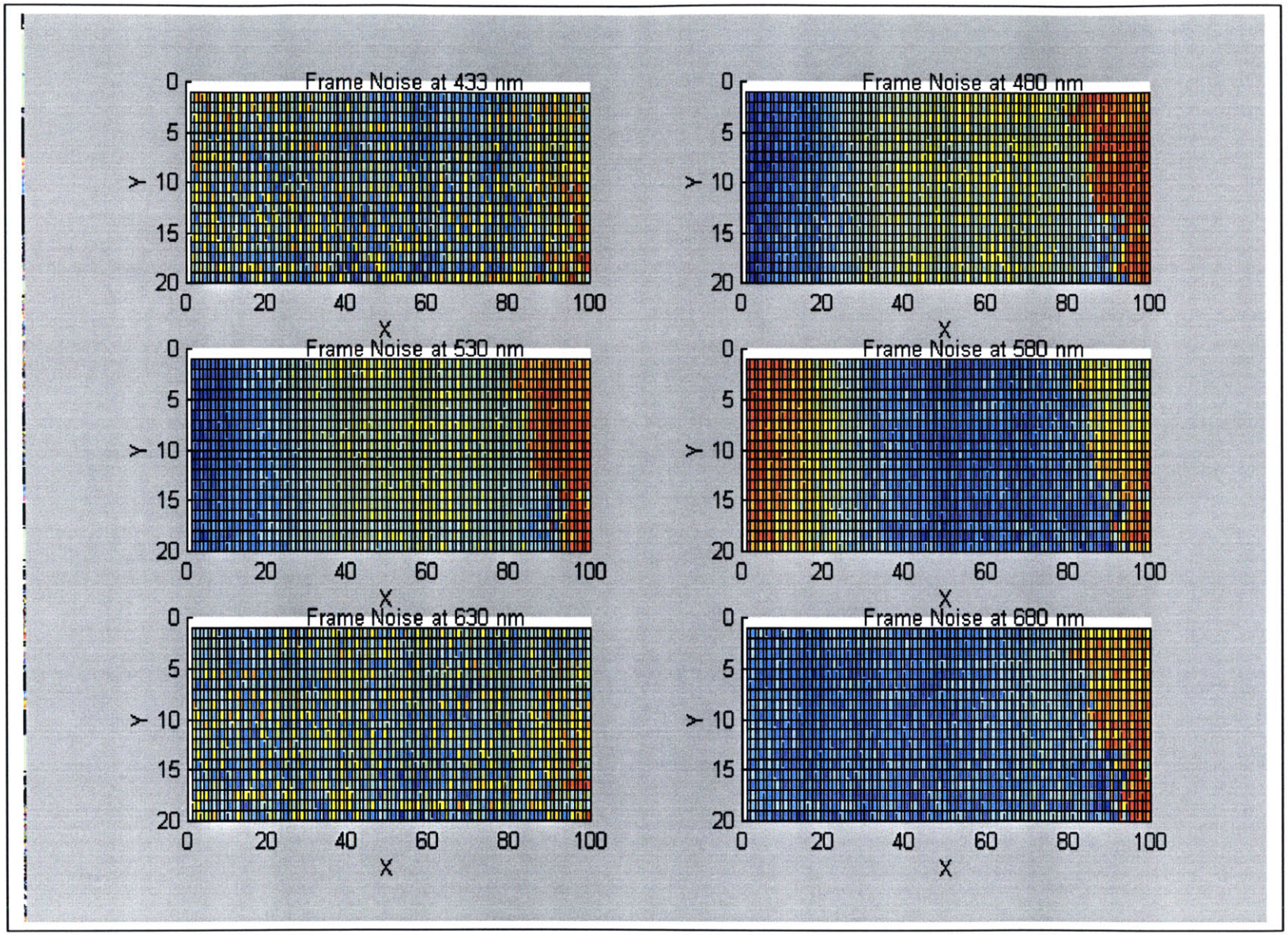


Figure 5.40: Frame Noise at 433, 480, 530, 580, 630, and 680 nm

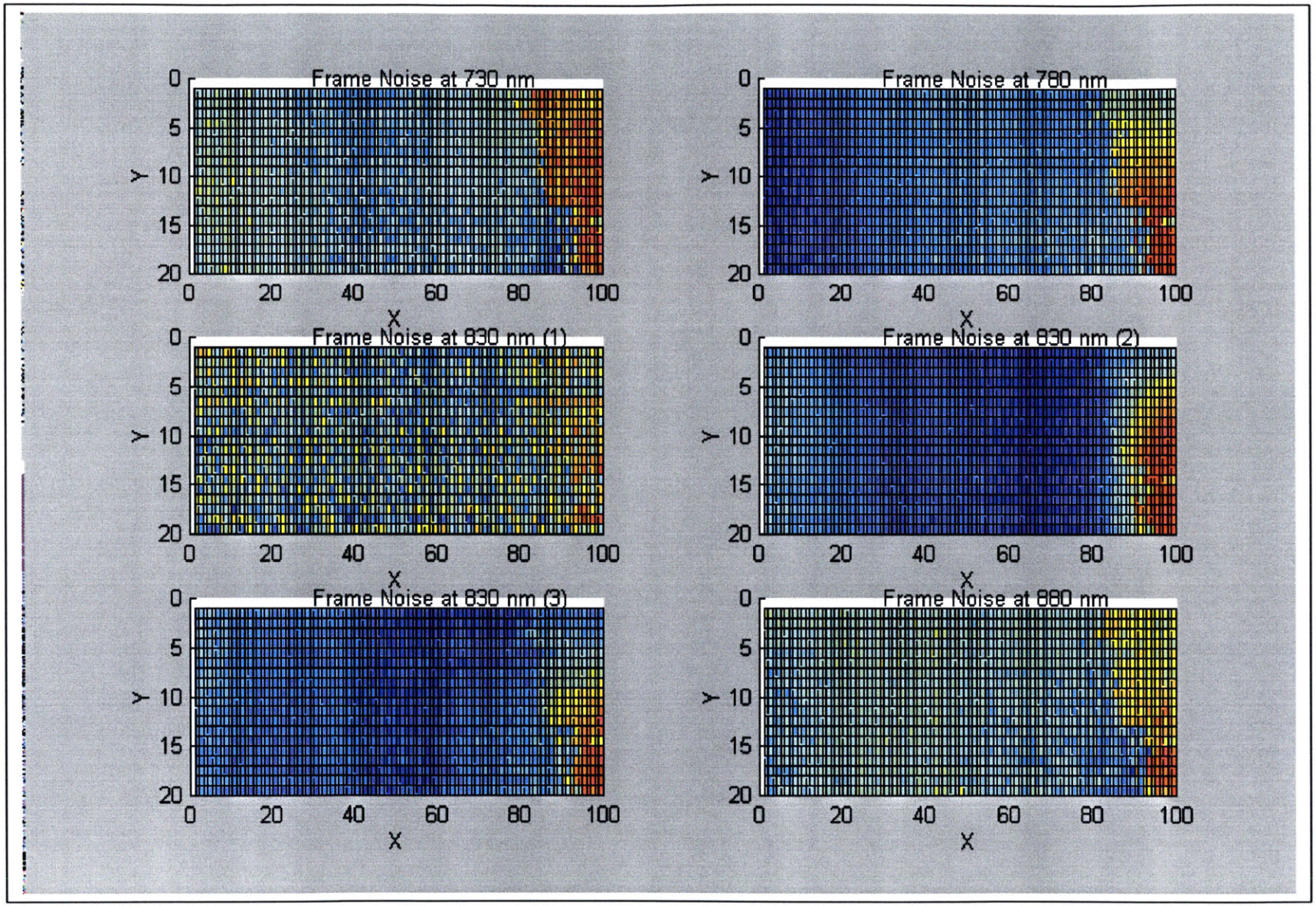


Figure 5.41: Frame Noise at 730, 780, 830, 830, 830, and 880 nm

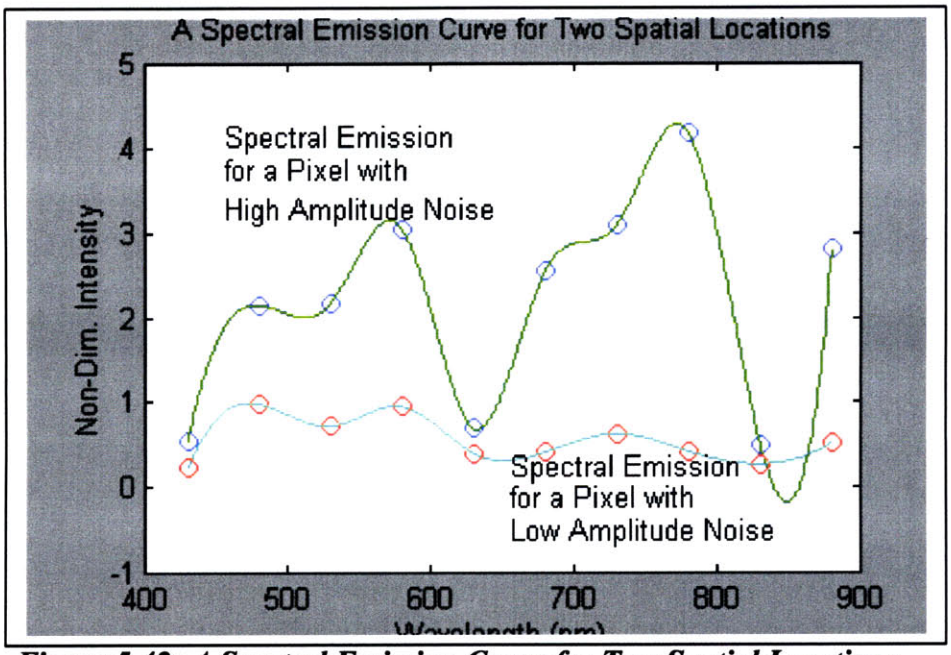


Figure 5.42: A Spectral Emission Curve for Two Spatial Locations

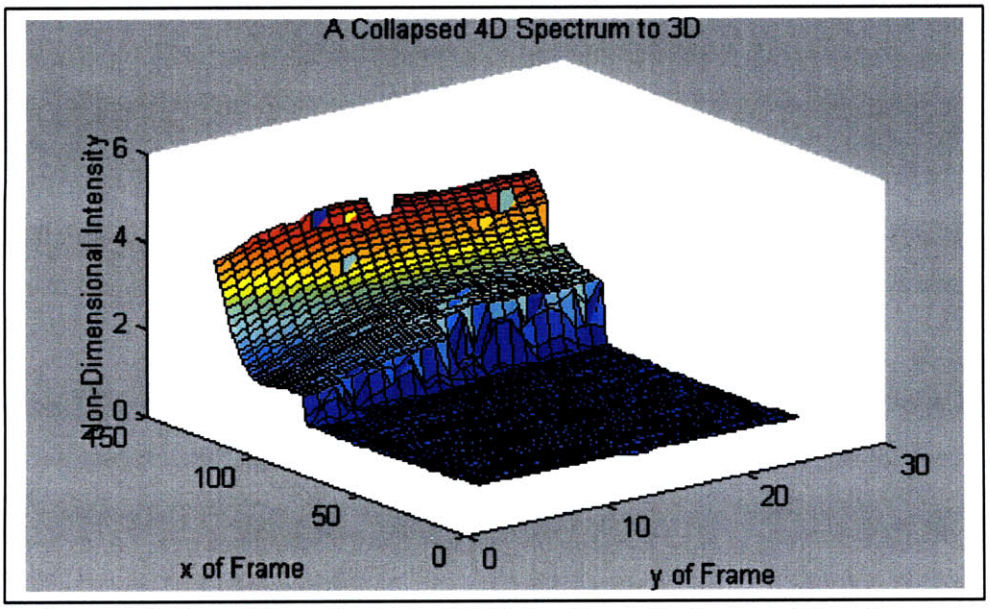


Figure 5.43: A "Penalty Schedule" for Scan Trajectory Planning

Chapter 6: Applications of this Technology

6.1 Application: Monitoring Human Hand Motion in a Task Manipulation

The monitoring of human motion has been a topic of research for many years. A person's "gait" has been studied in the Newman Laboratory at MIT []. The motion of dots mounted at various places on the leg were monitored to develop ideas about the mechanics of walking. However, these dots were red, and for this reason, visible. The main difference between the Newman lab's research and this application is that the dots that are used in this application are invisible, as well as having the capability of being uniquely identified if bar codes were used to replace the dots. However, this would require the use of a high resolution camera as well as a tracking motion that could be obtained from a pan and tilt.

For this application, the same approach was taken, but this time the dots were placed at various points on a surgical glove. The dots were roughly 4 mm in diameter, and were placed in a triangle configuration as shown in Figure 6.10. The ink is invisible, but was printed on white paper so that the configuration could be easily seen in print.

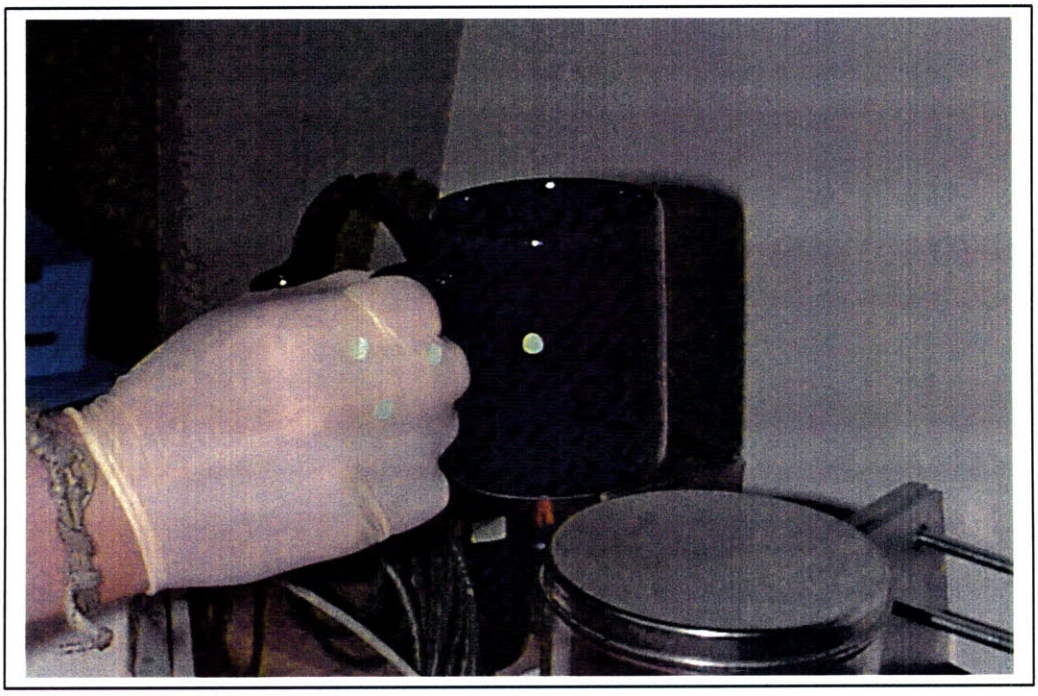


Figure 6.10: The Task

The hardware was modified slightly for this application. The galilean beam expander was replaced by a hologram diffuser which gave a 10 degree cone of diffusion. Figure 6.11 shows this new diffusion setup. The desire was to have a 2 foot diameter of coverage in which the motion could be tracked.

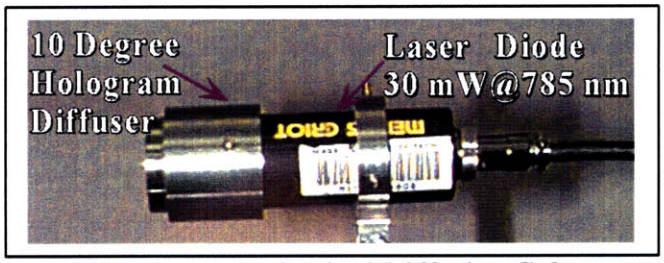


Figure 6.11: A Conical Diffusion Scheme

The acquisition hardware was also modified slightly. Figure 6.12 shows this new setup; the main differences being that the setup is fixed on a tripod and the image intensifier is added. Great care was also take to remove all sources of noise from the area that would be amplified when passed through the image intensifier. It should be noted that due to different focal lengths needed between the image activation and image acquisition, the diode was mounted separately from the camera and intensifier.

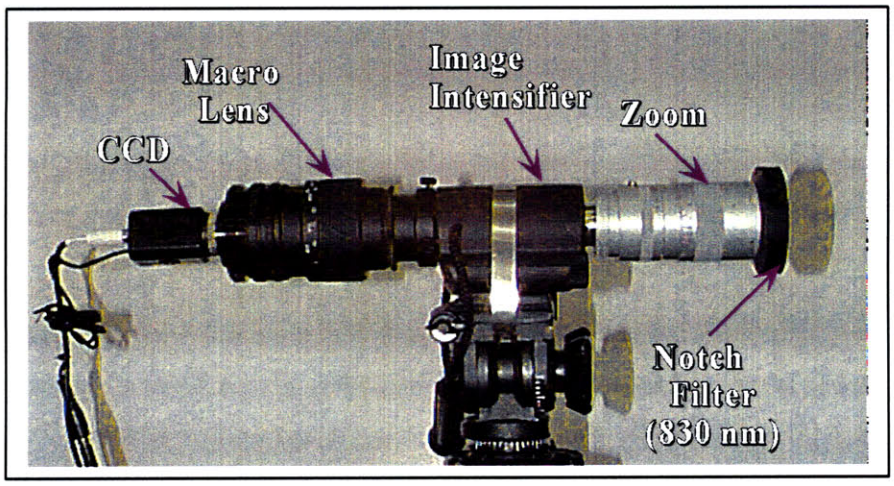


Figure 6.12: Hardware Setup for Human Motion Monitoring

The task that was undertaken for monitoring was the placement of a coffee mug on a tin canister. This task seemed to be appropriate since inaccurate motion could result in either damage to the mug or to the tin can.

Figure 6.13 shows the motion of the dots in Cartesian coordinates. The spatial scaling is 40 pixels to 1 inch, and the time scaling is 0.2 seconds between frames. Since only one camera is used in this task monitoring, the actual depth of motion is not evident. However, two pieces of information may be extracted from Figure 6.13 that give insight into the depth of motion. Since the camera field of view is known, as well as the size of the dots (4 mm), the distance from the camera to the dot may be determined. This feature is referred to as "pose estimation [ ]" in which multiple points on the same object are monitored. In this case, the multiple points are the reference dots at known size and location on the hand and mug. Additionally, if the ink print is uniform and the laser power is constant, then the intensity of the return signal varies as the square of the distance from the dot to the laser diode, which is mounted at a different position than that of the camera. This gives inherent triangulation.

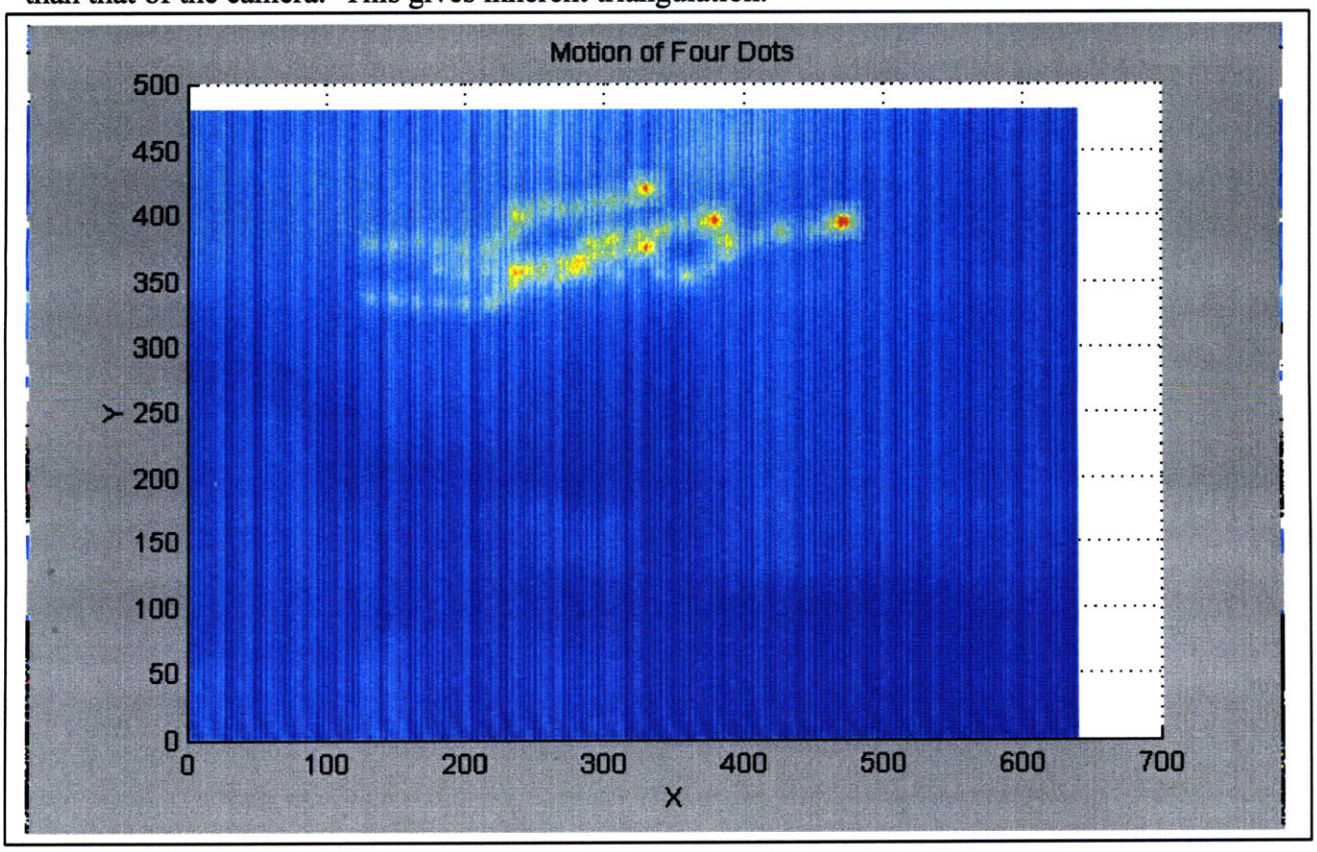


Figure 6.13: The Motion of Four Discrete Dots (on the Hand) During Task Manipulation

From this information, the centroid of each dot may be calculated for each frame. In the case of this task, since unique identity of each dot is not established (since dots are used and not bar codes), the initial position is established prior to the task, and "temporal consistency" is used to identify the dots as they progress through the task. When a series of images are acquired at a frame rate that is much higher than the inverse time for task completion, then there will be only a slight change in the appearance of a small region from frame to frame. This method is used to identify the dots as they transition through the area of the grabbed frames during the performance of the task. Figure 6.14 gives a plot of the calculated dot centroid as it moves through the motion of the task. It should be noted that when calculating the dot centroids, the center calculated is very sensitive to software thresholds as well as inconsistencies in ink application on the dots.

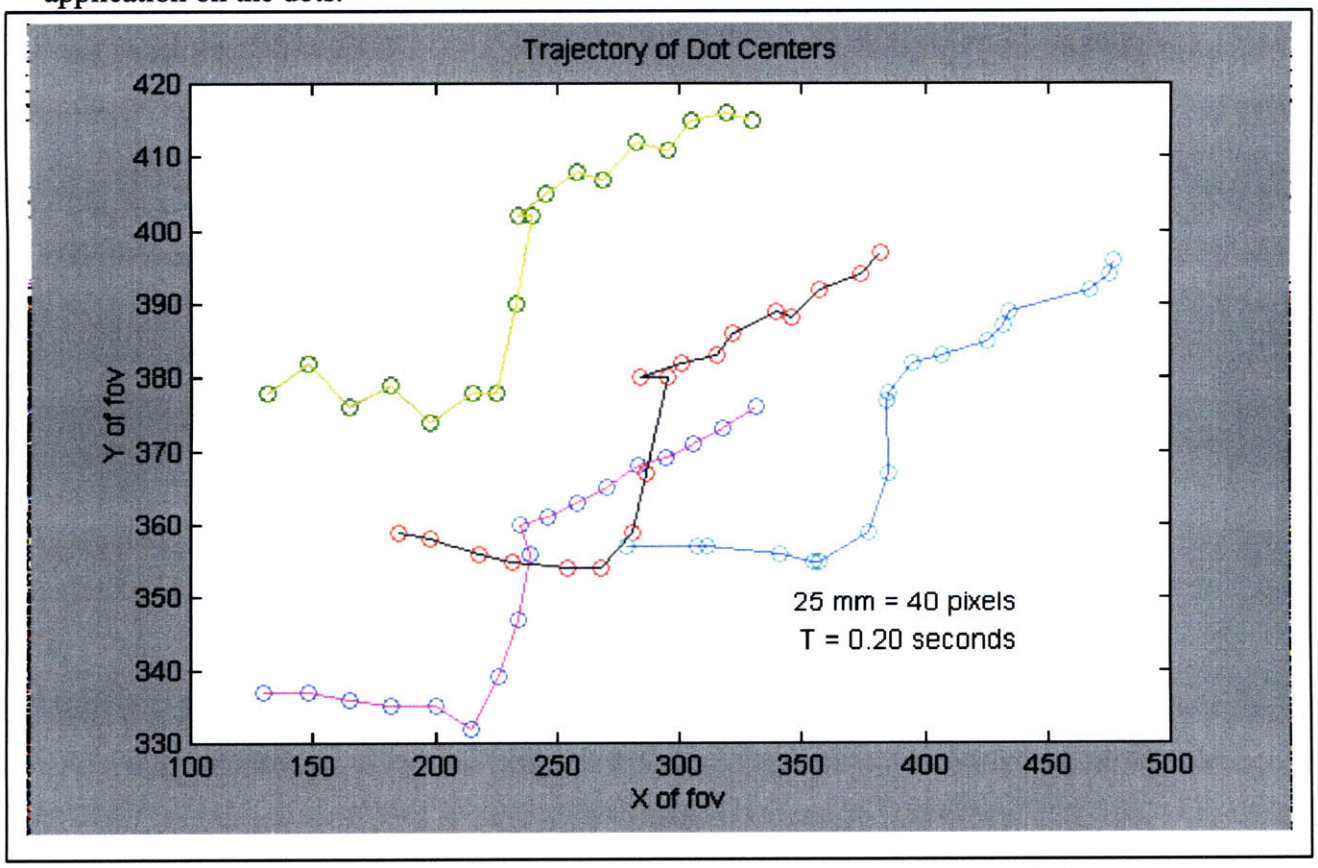


Figure 6.14: The Trajectory of the Dots in the Task Manipulation

6.2 Application: Tracking

The prototype has a simple tracking algorithm included, although limited success has actually been found due to the extremely small area of diffusion with the galilean beam expander as well as the need for a change in zoom setting with a change in depth. A camera basically gives a 2D representation of cylindrical motion, with an associated acceptance angle. The tracking routine given in this section uses setpoints in the form of desired spatial Cartesian locations of dot centroids in the viewable frame. The region of interest is set about the given beam diffusion area, and this is a function of zoom setting.

In this tracking routine, the zoom is servoed to give roughly the same area taken up in pixels for a given dot, and an autofocus routine is used set focus at varying levels of depth of zoom. This autofocus was developed from the observation that when an item is in focus, such as a bar code, the relative area that the object takes up in the frame is at a minimum when the object is in focus. In a given range of focus, the area in pixels taken up by an object will be parabolic in shape, with the minimum of the parabola the "in-focus" point. Figure 6.20 illustrates this autofocus logic. The minimum is found by taking the derivative of pixels taken up to focus percentage and finding where this quantity is zero. Of course, this whole process is very sensitive to correct software thresholds. It should also be noted that this process is the limiting factor in the speed of track, since it takes much more time to stroke the zoom than it does to pan or tilt.

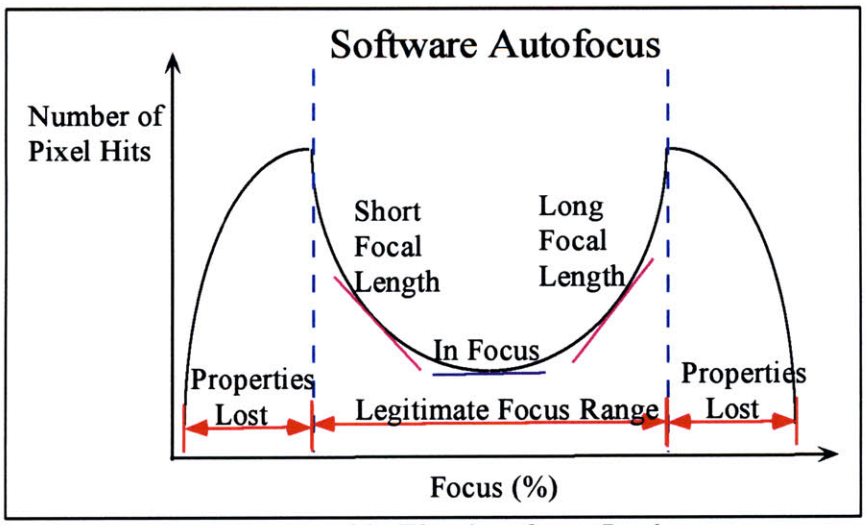


Figure 6.20: The Autofocus Logic

Given the quantities found in section 3.2 for the horizontal and vertical fields of view, and if the bar code has x and y dimension of X_{BC} and Y_{BC}, respectively, but basically gives a circular emission such that the diameter of this emission is:

$$D_{BC} = X_{BC} = Y_{BC} \tag{6.201}$$

This gives the number of horizontal and vertical pixels that see this emission:

$$S_{H,BC} = \frac{X_{BC}\cos\alpha(P_H)(Z)(efl)}{(d)(C_f)} = \frac{D_{BC}\cos\alpha(P_H)(Z)(efl)}{(d)(C_f)} \tag{6.202}$$

$$S_{V,BC} = \frac{Y_{BC}\cos\beta(P_V)(Z)(efl)}{(d)(C_f)} = \frac{D_{BC}\cos\beta(P_V)(Z)(efl)}{(d)(C_f)} \tag{6.203}$$

Simplifying the approach by assuming small angles of rotation, α and β, then the distance d from the camera to the bar code is:

$$d = \frac{Z(efl)(D_{BC})(P_H)}{(C_f)(S_{H,BC})} \tag{6.204}$$

If an origin is assigned to the camera frame, for instance at the dead center of the frame, and measurements from this origin in pixels are denoted X and Y, then the physical location of the bar code in Cartesian coordinates from this origin may be denoted X_P and Y_P:

$$X_p = \frac{X(C_f)(d)}{Z(efl)(P_H)} \tag{6.205}$$

$$Y_p = \frac{Y(C_f)(d)}{Z(efl)(P_H)} \tag{6.206}$$

It should be noted that this method does not worry about strict definition of the origin or coordinate transformations, but is concerned with the relative motion from one frame to the next. Figure 6.21 gives an illustration of this. The change in pixel location in the X and Y directions is denoted δ_X and δ_Y , respectively. The relative change in physical position is denoted δ_{PX} and δ_{PY}, and represents the actual change in position from one frame to the next.

$$\delta_{px} = \delta_x D_{BC}\cos\alpha\left(\frac{1}{S_{H,BC_{i+1}}} - \frac{1}{S_{H,BC_i}}\right) \tag{6.207}$$

$$\delta_{py} = \delta_y D_{BC}\cos\beta\left(\frac{1}{S_{V,BC_{i+1}}} - \frac{1}{S_{V,BC_i}}\right) \tag{6.208}$$

$$\delta_d = \frac{Z(efl)(P_H)}{C_f} D_{BC} \cos\alpha \left(\frac{1}{S_{H,BC_{i+1}}} - \frac{1}{S_{H,BC_i}} \right) \tag{6.209}$$

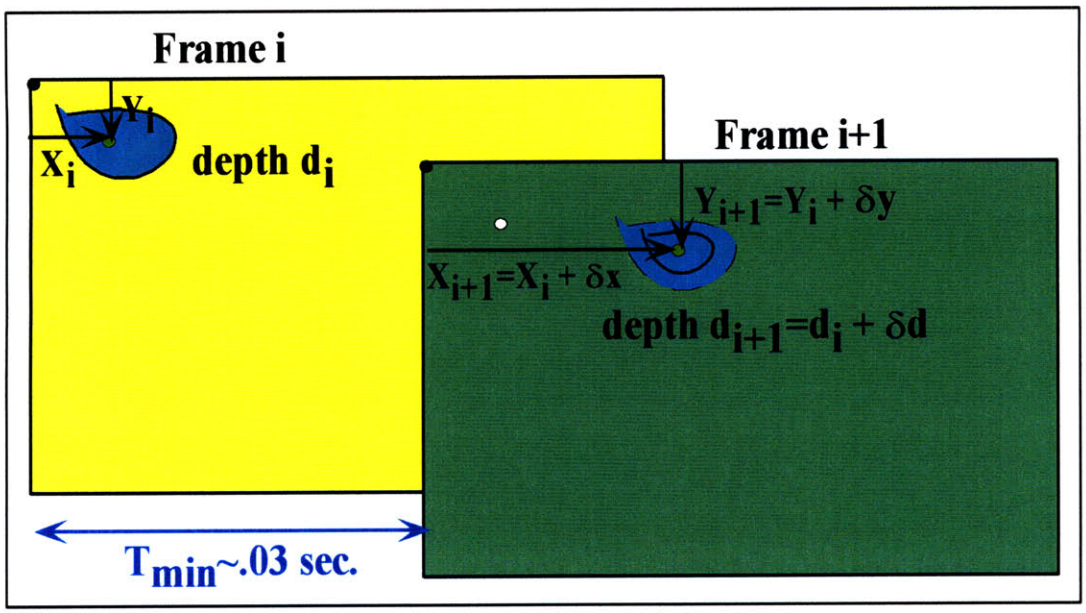


Figure 6.21: Frame to Frame Motion

The relative velocity may be determined knowing the frame rate of the acquisition, f_s:

$$V_{px} = (\delta_{px}) f_s \tag{6.210}$$

$$V_{py} = (\delta_{py}) f_s \tag{6.211}$$

$$V_{pd} = (\delta_d) f_s \tag{6.212}$$

Since the pan and tilt is servoed with respect to the angle of pan or tilt, it makes sense to convert the Cartesian measurements to an angular system of measurement (See Figure 6.22). Denoting the angle of pan rotation as θ_X and the angle of tilt rotation as θ_Y, and noting that the following relationships can convert the X and Y pixel positions to angles:

$$\phi_x' = \tan^{-1}\left(\frac{X(fov_H)}{P_H/2} \right) \tag{6.213}$$

$$\phi_y' = \tan^{-1}\left(\frac{Y(fov_V)}{P_V/2} \right) \tag{6.214}$$

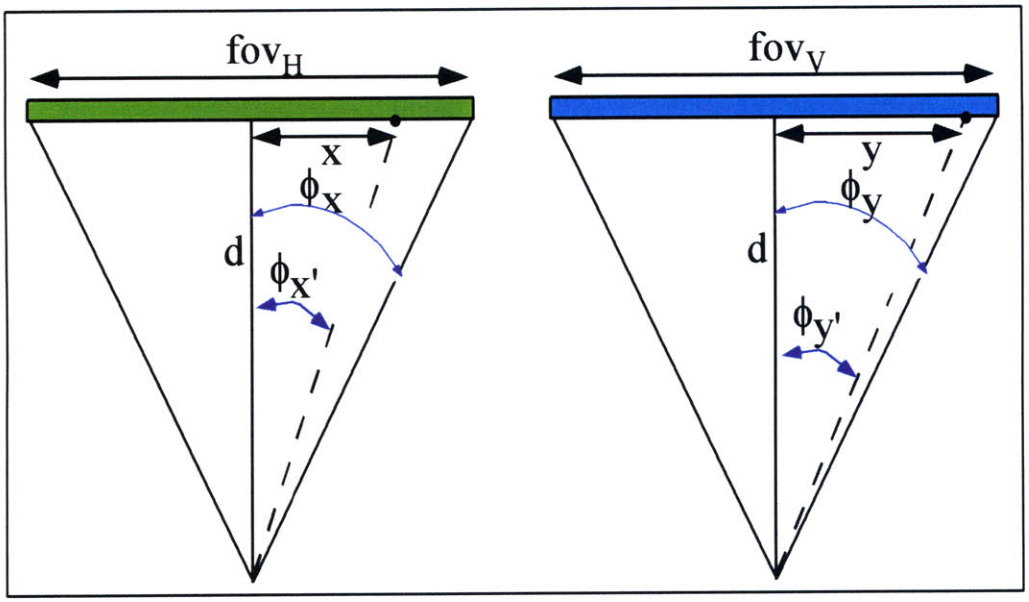


Figure 6.22: Cartesian to Spherical Coordinates

With these definitions, then the absolute angles to the tracked bar code centroid (φ_x and φ_y, respectively) are:

$$\varphi_x = \theta_x + \phi_x' = \theta_x + \tan^{-1}\left(\frac{X(fov_H)}{P_H/2}\right) \tag{6.215}$$

$$\varphi_y = \theta_y + \phi_y' = \theta_y + \tan^{-1}\left(\frac{Y(fov_V)}{P_V/2}\right) \tag{6.216}$$

A spatial setpoint or desired spatial location may be set in Cartesian coordinates, and then later converted to spherical coordinates. The spatial error is defined as the difference between this setpoint and the actual spatial location of the bar code:

$$error_x = SP_x - x \tag{6.217}$$

$$error_y = SP_y - y \tag{6.218}$$

The zoom error is defined a little differently. Since it is desired to keep the same amount of frame area occupied by the bar code as it changes depth, the setpoint is a desired pixel area (SP_{dA}) and the measurement is the number of pixels taken up by the bar code ($S_{H,BC} \bullet S_{V,BC}$). It should be noted that due to limits on the stroke of the zoom, some setpoints may not be possible.

$$error_{dA} = SP_{dA} - (S_{H,BC})(S_{V,BC}) \tag{6.219}$$

The derivatives of the error are as follows:

$$\frac{d}{dt}(error_x) = ((SP_{x_{i+1}} - x_{i+1}) - (SP_{x_i} - x_i))(f_s) \tag{6.220}$$

$$\frac{d}{dt}(error_y) = ((SP_{y_{i+1}} - y_{i+1}) - (SP_{y_i} - y_i))(f_s) \tag{6.221}$$

$$\frac{d}{dt}(error_{dA}) = ((SP_{dA_{i+1}} - (S_{H,BC_{i+1}})(S_{V,BC_{i+1}})) - (SP_{dA_i} - (S_{H,BC_i})(S_{V,BC_i}))(f_s) \tag{6.222}$$

Under simple PD control, the controller output is of the form:

$$r_x = K_{p_x} error_x + K_{d_x} \frac{d}{dt}(error_x)$$

It should be noted that the controller output r_x is a speed command to the pan axis of the pan and tilt. When the error and the change in error go to zero, this command goes to zero, and the pan "servos there," meaning it stops. If pure position control were desired, then an integral term would be added to the control action and the controller output would be in pan position. The controller terms for the tilt and zoom axis are similar in form.

A version of this control is found in the prototype software. However, since the diffusion area of the beam is small (with the galilean beam expander) and the speed at which the zoom may be changed is fairly slow, tracking becomes difficult in anything but relatively slow, fixed depth motion. However, the galvo design of section 2.6 offers a solution to these inadequacies in light modulation and zoom speed.

Chapter 7: Conclusions and Future Work

A system was presented that can be used to measure the fluorescent emission of a bar code printed in laser dye. The dye absorption was at 785 nm, and the peak ink emission was at 833 nm. The system architecture that was used in the prototype in the "read" of this bar code was also presented.

Some of the methods that were used to enhance the signal to noise ratio (S/N) were presented, with emphasis on the Perturbation/Correlation method. This method used a perturbation in laser intensity to give a corresponding change in ink emission magnitude, and these changes were correlated to separate the noise in the viewable frame from the signal.

With noise reduction methods outlined, research in the description of the noise was also presented. A system that could be used to determine the wavelength content of the noise in the viewable frame was presented. The calibration of this system was also discussed, with the basic architecture of the system described. Some motivation for this type of analysis was shown through black body model approximations to measure the relative intensity of the power flux at a given wavelength for a particular noise source. Examples of this noise measurement were presented, and some uses of this information were discussed.

For future work, a system was presented that would be one alternative to the prototype of the bar code reader that exists in the lab. It utilizes galvos to modulate the outgoing laser beam, as well as to bring in the incoming image. It also reduces the time required to sample the emission at different radial distances from the detector, which was found to be quite large with the motorized zoom lens. Finally, it uses avalanche photo-diodes in the emission detection since they have quite a large gain, which is useful for low-signal strength radiation measurement. In addition, they have a very high sample rate, much higher than the frame rate of the frame grabber used in the current system.

The emphasis at the design level was to produce a system for high bandwidth tracking applications, such as "Teaching by Showing," in which a human demonstrator performs a task, to be later reproduced in robot motion. The first step for this application has been completed in that human hand motion has been monitored with discrete points of the human hand tracked through the completion of a task. Also, a simple tracking routine has been included within the prototype software. With these steps taken, it is a small step to implement this technology in a "Teaching by Showing" application.

References

1. M. Oren and S. Nayar, "A Theory of Specular Surface Geometry," in Proceedings of Int'l Conference on Computer Vision, 1995, pp. 740-747.
2. R. Zhang et. al., "Analysis of Shape from Shading Techniques," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 377-383.
3. T. Wada, H. Ukida, and T. Matsuyama, "Shape from Shading with Interreflections under Proximal Light Source," in Proceedings of Int'l Conference on Computer Vision, 1995, pp. 66-71.
4. K. Ikeuchi and B. K. P. Horn, "Numerical Shape from Shading and Occluding Boundaries," Shape from Shading, ed. B. K. P. Horn and M. Brooks, MIT Press, Cambridge, 1989.
5. D. Forsyth, and A. Zisserman, "Reflections on Shading," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 13, no. 7, July 1991, pp. 671-9.
6. G. Healey, "Estimating spectral reflectance using highlights," Image and Vision Computing, vol. 9, no. 5, October 1991, pp. 335-9.
7. B. A. Maxwell and S. A. Shafer, "A Framework for Segmentation using Physical Models of Image Formation," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 361-8.
8. Funka-Lea and R. Bajcsy, "Combining color and geometry for the active, visual recognition of shadows," in Proceedings of Int'l Conference on Computer Vision, 1995, pp. 203--9.
9. Y. Ohta, T. Kanade, and T. Sakai, "Color Information for Region Segmentation," Computer Graphics and Image Processing, vol. 13, 1980, pp. 222-241.
10. G. J. Klinker, S. A. Shafer, and T. Kanade, "A Physical Approach to Image Understanding," Int'l Journal of Computer Vision, vol. 4, no. 1, Jan. 1990, pp. 7-38.
11. Nayar, K. Ikeuchi, and T. Kanade, "Shape from Interreflections," Int'l Journal of Computer Vision, vol. 6, no. 3, 1991, pp. 73-195.
12. L. B. Wolff, "Spectral and Polarization Stereo Methods using a Single Light Source," in Proceedings of Int'l Conference on Computer Vision, 1987, pp. 708-715.
13. R. J. Woodham, "Surface Curvature from Photometric Stereo," Technical Report, University of British Columbia, Computer Science TR 90-29, October 1990.
14. P. H. Christensen, and L. G. Shapiro, "Three-Dimensional Shape from Color Photometric Stereo," Int'l Journal of Computer Vision, vol. 13, no. 2, 1994, pp. 213-227.
15. Y. Sato and K. Ikeuchi, "Temporal-color space analysis of reflection," Journal of the Optical Society of America, vol. 11, no. 11, Nov. 1994, pp. 2990-3002.
16. S. Nayar, K. Ikeuchi, and T. Kanade, "Shape from Interreflections," Int'l Journal of Computer Vision, vol. 6, no. 3, 1991, pp. 73-195.
17. M. Swain, D. Ballard, "Color Indexing," Int'l Journal of Computer Vision, vol. 7, no. 1, 1991, pp. 11-32.
18. H. Murase, and S. K. Nayar, "Visual Learning and Recognition of 3-D Objects from Appearance," Int'l Journal of Computer Vision, vol. 14, 1995, pp. 5-24.

19. D. Lee, R. Barber, and W. Niblack, "Indexing for Complex Queries on a Query-By-Content Image Database," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 142-146.
20. A. Bobick and A. Wilson, "A State-based Technique for the Summarization and Recognition of Gesture," in Proceedings of Int'l Conference on Computer Vision, 1995, pp. 382-388.
21. C. Rothwell et. al., "Extracting Projective Structure from Single Perspective Views of 3D Point Sets," in Proceedings of Int'l Conference on Computer Vision, 1993, pp. 573-582.
22. J. L. Mundy and A. Zisserman, "Towards a New Framework for Vision," Geometrical Invariance in Computer Vision, ed. J. Mundy and A. Zisserman, MIT Press, Cambridge, 1992, chapter 1.
23. C. Coelho et. al., "An Experimental Evaluation of Projective Invariants," Geometrical Invariance in Computer Vision, ed. J. Mundy and A. Zisserman, MIT Press, Cambridge, 1992, chapter 4.
24. J. Ponce and D. Kriegman, "Toward 3D Curved Object Recognition from Image Contours," Geometrical Invariance in Computer Vision, ed. J. Mundy and A. Zisserman, MIT Press, Cambridge, 1992, chapter 21.
25. R. Mohr, L. Morin, and E. Grosso, "Relative Positioning with Uncalibrated Cameras," Geometrical Invariance in Computer Vision, ed. J. Mundy and A. Zisserman, MIT Press, Cambridge, 1992, chapter 22.
26. J. L. Mundy and A. Zisserman, "Projective Geometry for Machine Vision," Geometrical Invariance in Computer Vision, ed. J. Mundy and A. Zisserman, MIT Press, Cambridge, 1992, chapter 23.
27. J. A. Webb, "Latency and Bandwidth Considerations in Parallel Robotics Image Processing," in Proceedings of SUPERCOMPUTING `93, 1993, pp. 230-9.
28. G. Ahearn, "MaxVideo 200: A pipeline image processing architecture for performance-demanding applications," SPIE Vol. 2368, Image and Information Systems, 1994, pp. 225-8.
29. B. Ross, "A Practical Stereo Vision System," in Proceedings of Int'l Conference on Computer Vision, 1993, pp. 148-153.
30. F. Stein and G. Medioni, "Structural Indexing: Efficient 3-D Object Recognition," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 14, no. 2, Feb. 1992, pp. 125-145.
31. M. D. Wheeler and K. Ikeuchi, "Sensor Modeling, Probabilistic Hypothesis Generation, and Robust Localization for Object Recognition," IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 14, no. 2, March 1995, pp. 252-265.
32. R. J. Vayda and A. C. Kak, "A Robot Vision System for Recognition of Generic Shaped Objects," CVGIP: Image Understanding, vol. 54, no. 1, pp. 1-46.
33. D. P. Huttenlocker and S. Ullman, "Recognizing Solid Objects by Alignment with an Image," Int'l Journal of Computer Vision, vol. 5, no. 2, 1990, pp. 195-212.
34. R. Basri and S. Ullman, "The Alignment of Objects with Smooth Surfaces," CVGIP: Image Understanding, vol. 57, no. 3, May 1993, pp. 331-345.
35. P. Flynn and A. K. Jain, "3D Object Recognition Using Invariant Feature Indexing of Interpretation Tables," CVGIP: Image Understanding, vol. 55, no. 2, March 1992, pp. 119-129.

36. J. M. Rehg and T. Kanade, "Model-Based Tracking of Self-Occluding Articulated Objects," in Proceedings of Int'l Conference on Computer Vision, 1995, pp. 612-617.
37. F. K. H. Quek, T. Mysliwiec, and M. Zhao, "FingerMouse: A Freehand Pointing Interface," in Proceedings of Int'l Workshop on Automatic Face and Gesture Recognition, 1995, pp. 372-377.
38. M. J. Black and Y. Yacoob, "Tracking and Recognizing Facial Expressions in Image Sequences using Local parameterized Models of Image Motion, "Technical Report, University of Maryland Computer Science Dept., CS-TR-3401, January 1995.
39. T. Darrell and A. Pentland, "Space-Time Gestures," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1993, pp. 335-340.
40. J. Davis and M. Shah, "Gesture Recognition," Technical Report, Univ. of Central Florida, Computer Science Dept., CS-TR-93-11.
41. W. T. Freeman and M. Roth, "Orientation Histograms for Hand Gesture Recognition," in IEEE Workshop on Automatic Face and Gesture Recognition, Zurich, 1995.
42. W. T. Freeman, and C. D. Weissman, "Television control by hand gestures," in IEEE Workshop on Automatic Face and Gesture Recognition, Zurich, 1995.
43. D. Terzopoulos, A. Witkin, and M. Kass, "Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion," Artificial Intelligence, 36, 1988, pp. 91-123.
44. D. Metaxas and D. Terzopoulos, "Shape and Nonrigid Motion Estimation through Physics-Based Synthesis," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 6, June 1993, pp. 580-591.
45. H. Delingetter, "Simplex Meshes: A General Representation for 3D Shape Reconstructions," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 856-860.
46. J. Bergen et. al., "Hierarchical Model-Based Motion Estimation," in Proceedings of European Conference on Computer Vision, 1992, pp. 237-252.
47. J. Y. A. Wang and E. H. Adelson, "Layered Representation for Motion Analysis," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1993, pp. 361-366.
48. B. K. P. Horn and B. G. Schunck, "Determining optical flow," Artificial Intelligence, 17, 1981, pp. 185-203.
49. B. D. Lucas and T. Kanade, "An iterative image registration technique with an application in stereo vision," In Proceedings of Seventh International Joint Conference on Artificial Intelligence (IJCAI-81), 1981, pp. 674-679.
50. S. Uras, F. Girosi, A. Verri, and V. Torre, "A computational approach to motion perception," Biological Cybernetics, 60, 1988, pp. 79-87.
51. P. Anandan, "A computational framework and an algorithm for the measurement of visual motion.," International Journal of Computer Vision, vol. 2, no. 3, January 1989, pp. 283-310.
52. D. J. Heeger, "Optical flow using spatiotemporal filters," International Journal of Computer Vision," vol. 1, 1988, pp. 279-302.
53. D. Fleet and A. Jepson, "Computation of component image velocity from local phase information.," International Journal of Computer Vision, vol. 5, 1990, pp. 77-104.

54. H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," Nature, vol 293, 1981, pp. 133-135.
55. C. Tomasi and T. Kanade, "Shape and Motion from Image Streams under Orthography: a Factorization Method," Int'l Journal of Computer Vision, vol. 9, no. 2, 1992, pp. 137-154.
56. A. Azarbayejani and A. Pentland, "Recursive Estimation of Motion, Structure, and Focal Length," IEEE Trans. on Patterna Analysis and Machine Intelligence, vol. 17, no. 6, June 1995, pp. 562-575.
57. R. I. Hartley, "Projective Reconstruction and Invariants from Multiple Images," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 17, no. 6, June 1994, pp. 1036-1041.
58. B. K. P. Horn and W. J. Weldon, Jr., "Direct Methods of Recovering Motion," Int'l Journal of Computer Vision, vol. 2, 1988, pp. 51-76.
59. D.J. Creasey, "Advanced Signal Processing," Peter Peregrinus, Ltd., 1985, pp. 97-102.
60. C.L. Weber, "Elements of Detection and Signal Design," McGraw-Hill Book Company, 1968, pp. 74-76.
61. J.S. Bendat et al, "Engineering Applications of Correlation and Spectral Analysis," John-Wiley & Sons, Inc., 1993, pp. 43-76.
62. N.S. Jayant, "Digital Coding of Waveforms, Principles and Applications to Speech and Video," Prentice-Hall, Inc., 1984, pp. 86-110.
63. W.E. Hartnett, "Foundations of Coding Theory," D. Reidel Publishing Company, 1974, pp. 61-80.
64. I.F. Blake et al, "An Introduction to Algebraic and Combinatorial Coding Theory," Academic Press, Inc., 1976, pp. 95-100 & 161-165.
65. J.L. Miller et al, "Photonics Rules of Thumb," McGraw-Hill Book Company, 1996, pp. 1-409.
66. R. N. McDonough et al, "Detection of Signals in Noise," Academic Press, Inc., 1995, pp. 202-243.
67. A.F. Mills, "Basic Heat and Mass Transfer," R.D. Irwin, Inc., 1995, pp. 403-495.
68. S. Hutchinson et al, "A Tutorial on Visual Servo Control," IEEE Transactions on Robotics and Automation, vol. 12, No. 5, 1996, pp. 653-659.
69. L. Beiser, "Fundamental Architecture of Optical Scanning Systems," Applied Optics, vol. 34, No. 31, 1995, pp.7307-7316.

Appendix A

Concept / Selection Criteria	Vision\ Pattern Recognition	Invisible Coding (UV)	Invisible Coding (IR)	Active Objects (Transp.)	Traditional Bar Code Reader (Reference)	Laser Scan for Geometry (Shape)	Sonar or Acoustic (Shape)	Miniature Tuning Forks	Smell Emitters and Detectors
Safety	-	-	0	0	0	-	-	0	0
Cost	0	0	0	-	0	-	0	-	-
Location Accuracy	+	+	+	+	0	+	+	+	+
ID Accuracy	+	+	+	+	0	0	0	+	0
Robustness	0	+	+	0	0	+	+	0	0
Noise	0	+	+	+	0	+	+	+	0
Tracking Rate	-	+	+	0	0	0	0	-	-
Modularity	-	-	-	-	0	-	-	-	-
STD Components	-	0	0	-	0	-	-	-	-
Complexity	-	0	0	-	0	-	-	-	-
Ergonomics	+	-	+	-	0	0	0	-	-
Range	+	0	+	0	0	+	0	0	-
Sum of +'s	4	5	7	3	0	4	3	3	1
Sum of 0's	3	4	4	4	12	3	5	3	4
Sum of -'s	5	3	1	5	0	5	4	6	7
Net Score	-1	+2	+6	-2	0	-1	-1	-3	-6
Rank	4	2	1	7	3	4	4	8	9
Continue?	NO	YES	YES	NO	COMB	NO	NO	NO	NO

Figure A1: The Pough Chart Used in Method Selection

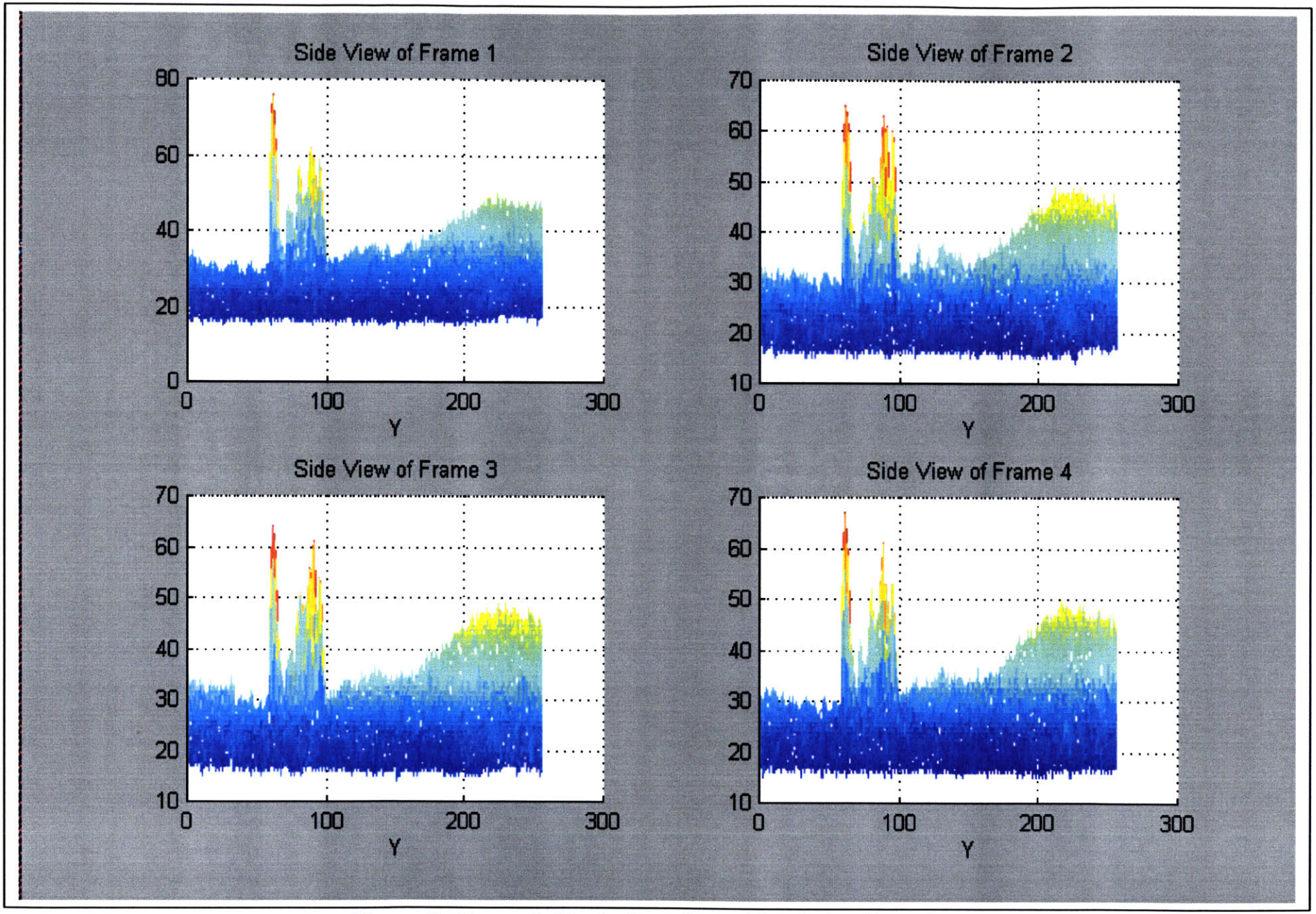


Figure A2: Frames 1-4 in the Perturbation/Correlation Example

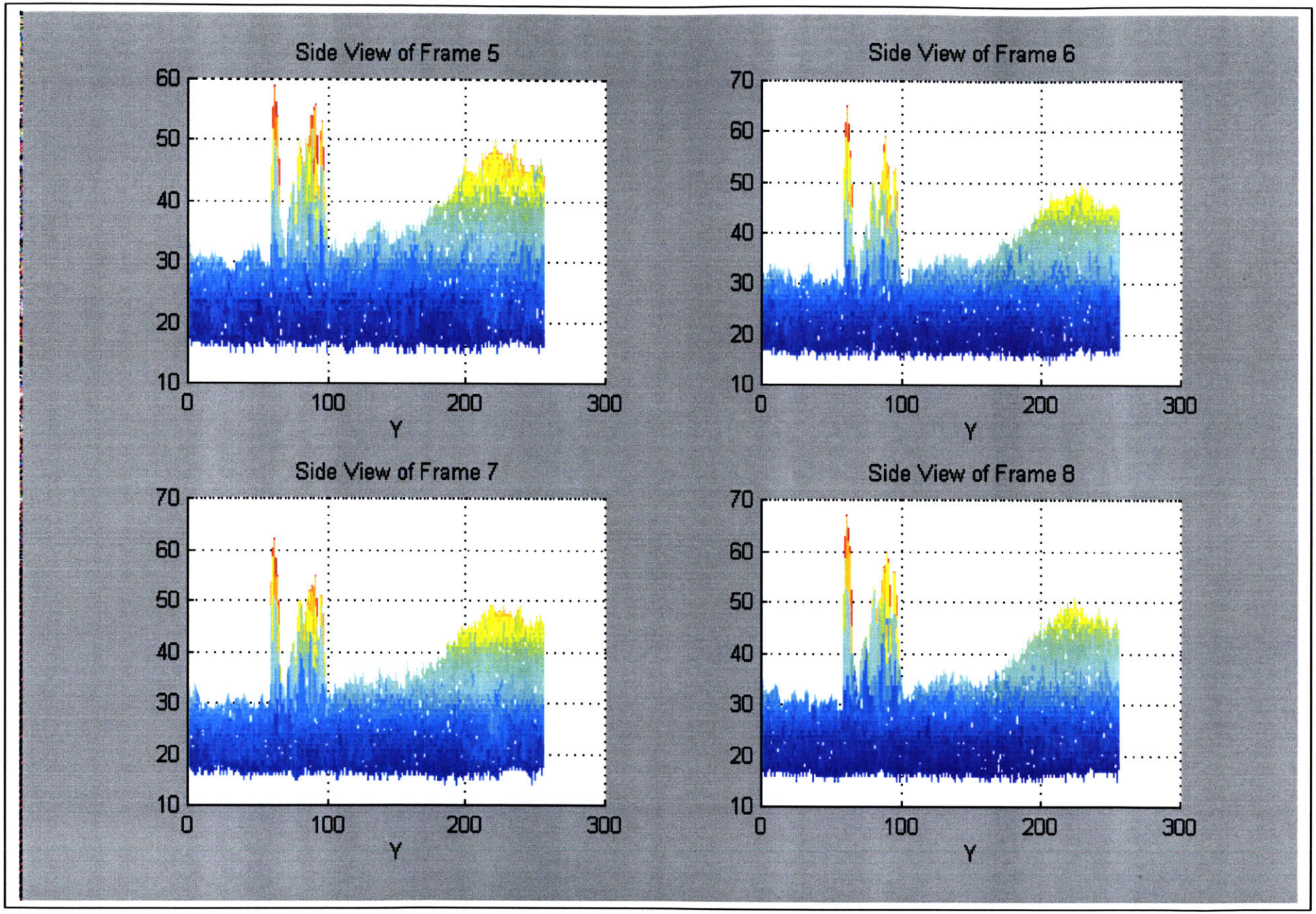


Figure A3: Frames 5-8 in the Perturbation/Correlation Example

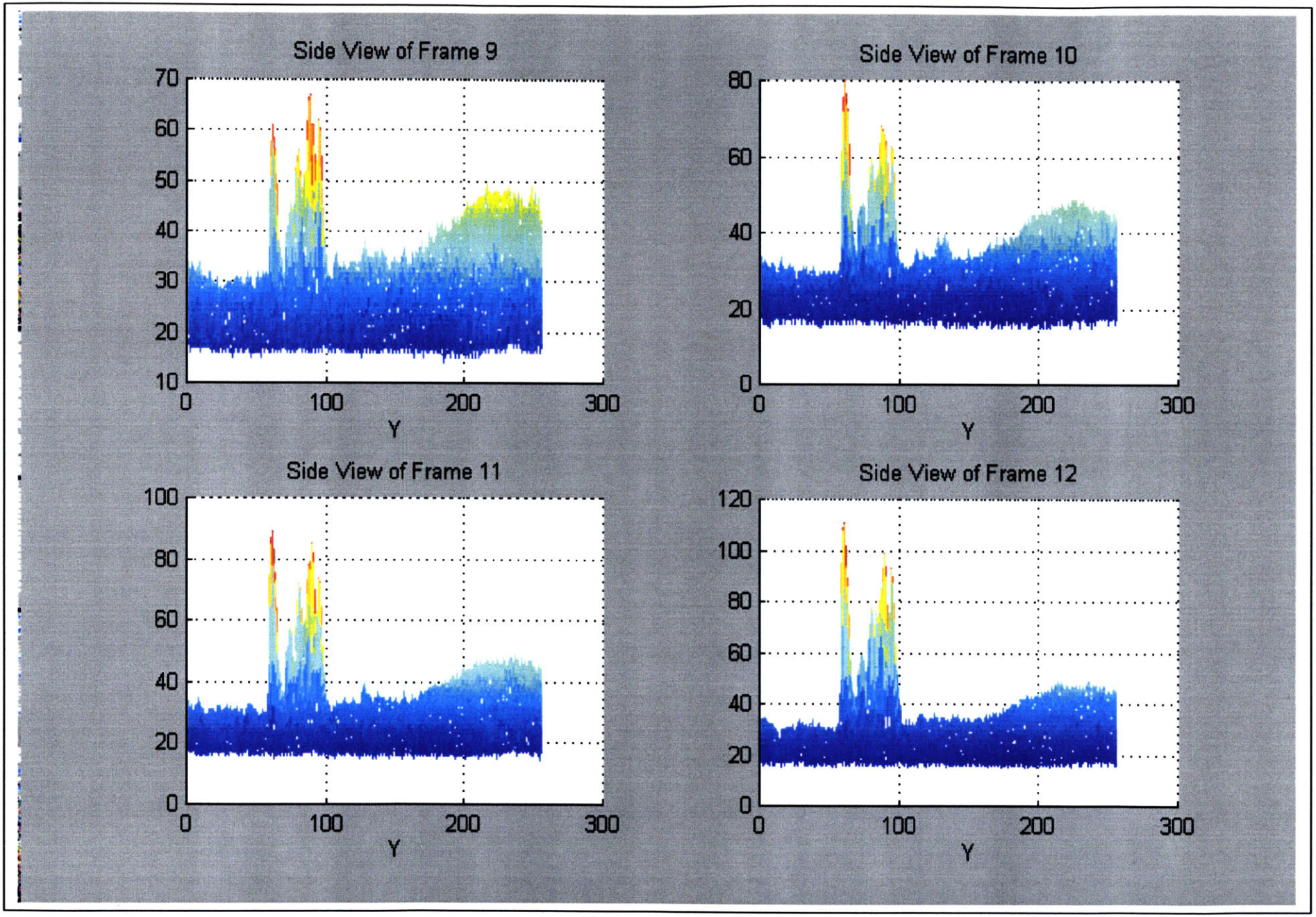


Figure A4: Frames 9-12 in the Perturbation/Correlation Example

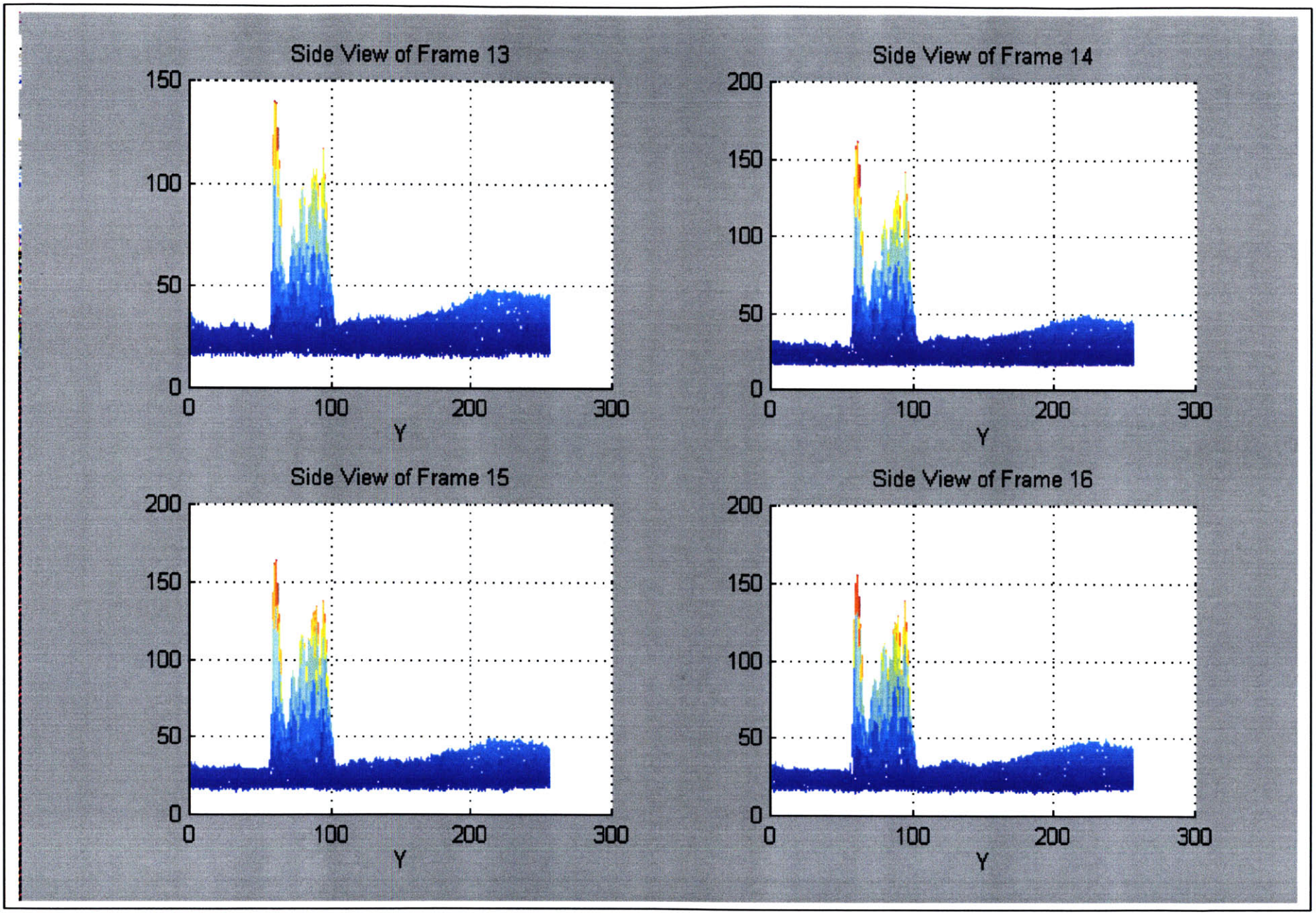


Figure A5: Frames 13-16 in the Perturbation/Correlation Example

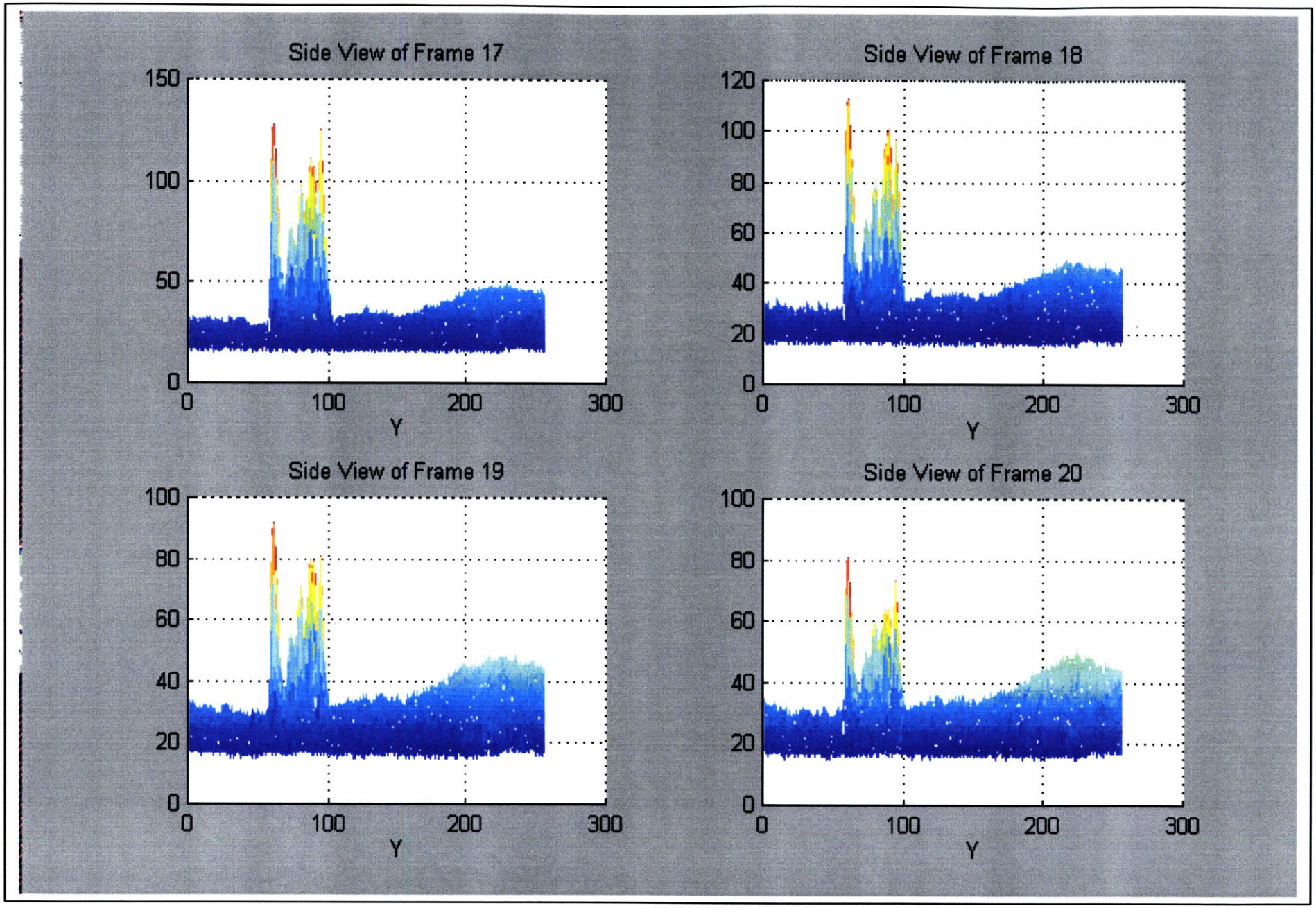


Figure A6: Frames 17-20 in the Perturbation/Correlation Example