Synthetic Aperture Microscopy

by

Michael Stephen Mermelstein

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

Doctor of Philosopy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1999

Author .....
Department of
Electrical Engineering and Computer Science
April 30, 1999

Certified by .....
Thomas F Knight Jr.
Senior Research Scientist
Thesis Supervisor

Certified by .....
Dennis M Freeman
Associate Professor
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Synthetic Aperture Microscopy

by

Michael Stephen Mermelstein

Submitted to the Department of
Electrical Engineering and Computer Science
on April 30, 1999, in partial fulfillment of the
requirements for the degree of
Doctor of Philosopy

Abstract

In the late 1800's, Ernst Abbe, research director of the Carl Zeiss Optical Works, wrote down the rules for a lens to form a sharp image. Advances in communications theory, signal processing, and computers have allowed us finally to break those rules.

Our "Synthetic Aperture Microscope" floods a large region with a richly complex, finely structured pattern of light—the interference pattern of a ring of n coherent sources.

A target within the volume of the interference fluoresces (or scatters or transmits) an amount of light that reveals correspondences with this "probing illumination." Modulating the phases and amplitudes of the n beams with carefully chosen modulation signals causes the probe illumination to step through a predetermined or measured family of patterns. A sensor records the target's response in a time-sequence. This time-sequence contains each of order n^2 complex Fourier coefficients of the target. Each of these coefficients is encrypted by a unique spread-spectrum key embedded in the amplitude and phase modulation signals. Signal processing picks out these coefficients to reconstruct an image of the target.

Low resolution conventional imaging maps an array of "targets" (actually portions of a larger target) to a CCD array, thus allowing this sensing process to be done in parallel over a large region. **The end result is to boost the resolution of a conventional imager by hundreds to thousands of sub-pixels per physical pixel.**

Both theoretical and experimental work on the engineering to make the concept practical are reported.

Thesis Supervisor: Thomas F Knight Jr.
Title: Senior Research Scientist

Thesis Supervisor: Dennis M Freeman
Title: Associate Professor

Acknowledgments

For John G. King, Dennis M. Freeman, Thomas F. Knight, Jr., and Berthold K.P. Horn, my teachers; and for my family.

This project was a collaboration between four major laboratories at MIT: the Research Laboratory for Electronics, Lincoln Laboratory, the Artificial Intelligence Laboratory, and the Media Laboratory. Contributors from each of these facilities, thank you. Notably, contributors included, from RLE, AJ Aranyosi, Stan S. Hong, and Abraham R. McAllister; from Lincoln, Lyle G. Shirley, John M. Fini, Daniel L. Feldkhun, Gregory Hallerman, Harold C. Payson, Emory Ariel, John F. Orthmann, Lee Zamir, John A. Tabaczynski, Kent R. Edwards, and Wade M. Kornegay; from the AI Lab, Nicholas R. Papadakis; and from the Media Lab, Michael Klug and Stephen Benton.

Elmer S. Hung kindly contributed MEMS mirror arrays for the M3 prototype.

Daniel L. Feldkhun co-invented the acousto-optic light projector technique used for the M4 prototype.

I will forever appreciate the many forms of help supplied by my dear family, John M. Fini, Ana Echániz, Oliver Thomas, Wilfredo Sánchez, and Chris Barnhart during my graduate years.

I am grateful to the Fannie and John Hertz Foundation for major support for my graduate education. Their generous fellowship afforded me the freedom to pursue this project. I am further grateful to DARPA (ETO/MEMS) for major support for this project under contract number F30602-97-2-0106.

Contents

List of Figures

List of Tables

Chapter 1

Perspective

1.1 The value of microscopes and of this thesis

Understanding starts small. Pictures help. See Figure 1-1.

Microscopes give us pictures of the small stuff from which everything is built. Our curiosity about the great variety of small stuff around us has inspired a variety of microscopes. There are two major types of practical value: lensed and scanning-probe. This thesis enables a third category for practical use: synthetic aperture.

1.2 Why build another type of microscope?

Synthetic aperture microscopes replace precision mechanical components with computation. In contrast with the economics of precision physical artifacts, computation is expected to continue its long-held scaling trends of ever-increasing complexity and power for ever-decreasing costs. Therefore, replacing the physical with computation is

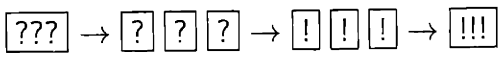

Figure 1-1: The value of microscopes. An important aspect of science is its quest to understand the complicated physical systems around us. The path to understanding a complicated system frequently involves examination of the subsystems and their interactions. Microscopes are often good for this type of examination.

a strategy that makes sense for the future. Ever-increasing demands on micro-imaging system performance for research, metrology, and inspection applications further motivate this paradigm-shift.

Synthetic Aperture Microscopy (SAM) is not just a single type of microscope. Rather, it is a system for making or upgrading microscopes that implements the shift to computation to a variable degree and in a scalable fashion. SAM ideas can apply to existing microscope types, maintaining many of the features and advantages demonstrated with conventional methods while adding its own *unique benefits.* As presented here, it is a way of using a combination of animated, structured illumination and computers to upgrade the performance of an imaging system or imaging concept, pre-existing or otherwise.

Therefore, SAM is a way of

1. **improving existing** imaging systems as well as

2. **making new kinds** of microscopes.

First, we will discuss the aspects of existing imaging systems we might want to improve, and then we will consider what types of new imagers are desirable.

1.2.1 Four fundamental improvements are possible

Four key parameters of an imager can benefit from the SAM ideas in this thesis. They are listed below with examples intended to connect with the reader's intuition.

1. **Working distance:**

   Working distance is the distance from the target (the object under examination) to the instrument. Often, working distance matters when interaction with the target is necessary. For example, an eye surgeon may need to operate on part of an eye while viewing it with a microscope. If the microscope has an insufficient working distance, it will get in the way of the surgery. Traditionally, longer working distances correspond with lower resolution. This correspondence makes sense from everyday life: you can't see far away things as sharply as things that

are up close. With SAM, however, working distance can be **improved with no sacrifice in resolution.**

2. **Field of view:**

   Field of view is the amount of the world that can be viewed at once, in a single picture. If you have a camera, chances are that you have a zoom lens or multiple lenses for it. Why not just use a single wide-angle lens for all of your photographs? After shooting, you could choose regions of your pictures for blow-up—saving money by buying fewer lenses or options and saving work by not bothering to aim your camera as precisely. Traditionally, you can't take this simplified approach because use of too small a portion of a photographic negative results in a blurry enlargement. With SAM, however, field of view can be **improved with no sacrifice in resolution.**

3. **Depth of field:**

   Depth of field is the range of distances from an instrument that are considered to be in focus. Focusing a camera, for example, is the act of aligning this range with the subject of the photograph. "Focus-free" cameras (such as disposal cameras and the lowest quality point-and-shoot cameras) provide a depth of field that typically includes everything from 6 feet to the horizon. Because of fundamental properties of lenses, the resolution of these cameras necessarily suffers accordingly. With SAM, however, depth of field can be **improved with no sacrifice in resolution.**

4. **Resolution:**

   In the preceding statements and in the technical jargon of the microscope world, resolution refers to a minimum separation distance of adjacent, distinguishable details in a picture. In common usage, however, resolution refers to the number of independent elements—how many dots, pixels, dabs of paint, grains of film, or details in general—in a picture. The definitions are linked by the field of view: common-resolution is the number of jargon-resolution-sized elements in

the field of view. Typically, jargon-resolution comes at the expense of field of view, keeping common-resolution unchanged. This is the case, for example, with the zoom lens on a video camera. Zooming in allows you to make out closer features in the scene, giving up your view of the surrounding features and keeping the total number of picture elements (pixels) the same. SAM, however, can **improve both types of resolution simultaneously.**

Anything from a hand-held camera to a research-grade microscope can benefit.

Putting it all together

Combining the best-cases of these four parameters, one imagines a camera that can capture every vein on every blade of grass in a vast meadow, every transistor in every chip on a production wafer, or every synapse on every neuron in a mouse's brain *in a single picture.* SAM enables such a camera.

The formula for taking the picture you want is simple:

1. Find an instrument that could take the low-resolution version of the picture, and

2. without reducing the working distance, field of view, or depth of field, **boost its resolution using SAM** until the details of interest are resolved.

1.2.2 SAM enables new and valuable microscopes

SAM provides a new flexibility for designing image sensors by decoupling the four previously entangled fundamental parameters listed above. It can not, however, achieve a resolution beyond the famous "diffraction limit" related to the wavelength of the light it uses[1]. [1]

[1]SAM resolution and this limit are discussed in Section 2.4.

Wavelength and resolution go hand-in-hand

Wavelength is a characteristic distance associated with colors of light and with particles of matter.

Light, as we normally think of it, has a wavelength in the range of half of a micron—one two-millionth of a meter. Roughly speaking, then, a microscope based on "normal light" can not resolve details much smaller than this characteristic size. Therefore, light microscopes can see the DNA in our cells, but not the genes in the DNA. They can see the wiring of our brains, but not the structure of the proteins that connect those wires.

Microscopes that basically touch the target, either with a probe (a pointy stick that scans across the target like a blind man's cane) or with a narrow flowing stream of particles, use the wavelength of *matter*, not normal light, to achieve a resolution in the range of a nanometer—a *billionth* of a meter. It is not always practical, however, to touch target features with one's instrument or to put the target in a vacuum, where particle streams flow.

Different wavelengths offer different ways to see the world

Many flowers that appear plain to us in fact have complex and beautiful patterns in the ultra-violet (UV)—*an alternate rainbow of colors* that bees and other pollinating insects can naturally appreciate far better than we can. In general, UV wavelengths and others such as Infra-red and X ray can provide important, interesting, and beautiful insights into our physical surroundings. Not all of these wavelengths are well suited for conventional approaches to microscopy, though.

X ray and electron wavelengths: an unfulfilled potential

X rays have wavelengths in the sub-nanometer range and could theoretically provide sub-nanometer resolutions from a reasonable working distance. However, no one has ever succeeded in building an X ray lens with resolution approaching X ray wavelengths. The difficulty is in the unprecedented degree of perfection a conventional

X ray lens would require.

Electron microscopes provide some of the best resolution pictures anyone has ever taken. As with X rays, though, the potential implicit in the electron's wavelength to resolve features as much as ten or a hundred times smaller has never been fulfilled.

Non-contact imaging with these wavelengths is enhanced or enabled

In the cases of X rays and electrons, as well as with some colors of UV light, the difficulty of achieving wavelength-scale resolution derives from the challenges of creating a good enough lens. Traditionally, such a lens would have a large, astoundingly perfect *continuous aperture*—a continuous space through which X ray, UV, or electron flow is channeled and controlled according to mathematical equations with a requirement for complete order and abidance. This daunting requirement calls to mind numbers obediently flowing in a computer.

SAM transcends the problem of making the perfect lens. Through its ideas, X ray, electron, UV, infra-red, and "normal light" microscopes that maximize resolution are enabled using *im*perfect lenses working together with the perfect order possible in mathematical ideas and computer programs.

1.3 A first glimpse at SAM

One of the prototype microscopes built for this thesis was centered on a standard camera lens. Figure 1-2 shows the lens augmented by a synthetic aperture. The SAM upgrade improved the (areal) resolution performance of the camera lens *a thousand fold* without reducing the working distance, field of view, or depth of field.

Not shown in Figure 1-2 is the second aspect of SAM technology: mathematical ideas and computer programs that make the lens and large ring shown in the figure as good as (and in some ways even better than) a mathematically perfect lens as big as the large ring.

Figure 1-2: In one of the microscope prototypes for this thesis, a standard camera lens (lower right) was augmented by a ring of diffraction gratings—using Synthetic Aperture Microscopy ideas to boost its resolution performance a thousand fold. The working distance, field of view, and depth of field were not reduced by the upgrade. Shown approximately life-size, the combination of lens and large ring (which is mostly cut off) provides the performance of a mathematically perfect lens as big as the large ring.

1.4 How does it work?

The theory of SAM is described fully in Chapter 2. In the way of introduction, the present section offers an overview including accurate explanations of the key phenomena by metaphor.

SAM has two parts: one physical, and one computational.

1.4.1 The physical part: an animated pattern projector

The physical part of SAM is a special illuminator—a light source of finely detailed patterns. The grating ring in Figure 1-2 is an example. Laser beams from each grating (the larger rectangular parts) in the ring overlap in the space where the lens is focussed and *interfere* with each other in that space.

Interference

Interference is a dance done by every two rays of light of similar orientation (polarization) that cross paths. Rays of very different (frequency) light make unwilling dance partners. They beat against each other at a violent rate, typically too fast for us to observe, and we see only a blur. Closely related (mutually coherent) light rays, on the other hand, like rays born in the same laser at nearly the same time, have similar values, so to speak. They slow-dance and form beautiful patterns that we can easily detect and appreciate.

The light rays from the grating ring all come from the same laser. Where they overlap, every *pair* of beams fulfills an obligation to dance. The resulting interference pattern is far more complex and interesting than we might first guess. That is because there are so many pairs. Each and every beam must partner with each and every of the remaining beams, leading to a squared number of partnerships starting with a given number of beams. In the prototype shown, about forty beams form about a thousand beam-pairs. The aggregate pattern is a combination of these thousand dances; it has an astoundingly intricate and fine level of detail. It fills the three-dimensional space where the beams overlap.

Modulation

A certain degree of control over the behavior of the pattern is achieved by (phase) *modulating* the source beams. Like whispering a secret into a dancer's ear, *modulating a beam endows it with information that influences how it will interact with any other beam.* Information placed in each beam, then, interacts with the information placed in each other beam, influencing the components of the interference pattern in a complicated but ordered way.

In SAM, each beam is modulated with a unique special code, designed to interact with the other codes. **Modulation animates the interference pattern.** Like the choice of music at a party, the choice of modulation codes determines the motions, speed, and orderliness of the participating light pairs.

Target objects

Generally, a target object is viewed by the instrument. Targets can be transmissive (mostly clear) or opaque. In each case, **there is an interaction between the target and the animated interference pattern.**

A mostly clear target is like a Jello dessert with strawberries suspended in it. Many biologically interesting structures such as the memory circuits in a brain are like this. With this type of target, the interference pattern will enter and fill the target's volume.

A mostly opaque target is like a Jello pudding dessert with swirls of chocolate and vanilla. Many technologically interesting structures such as the memory circuits in a computer fall into this category. In the case of a mostly opaque target, the surface of the target cuts through the three-dimensional interference pattern, which will light up the surface with a projected pattern—like a slide projector lighting up a projection screen.

Mapping contrast: a microscope's goal

The job of a microscope, camera, or image sensor in general is to locate the features of *contrast* in a scene. Contrast would be either the strawberries in the volume or the vanilla swirls on the surface respectively in the two dessert examples above. Knowing where the fruit or swirls are allows you to form an accurate picture of the target.

Suppose your camera lens is too far away, or the contrast features are too small and close together to make out. In the first example, the camera sees only a blurry bowl, perhaps, but not the strawberry locations. SAM boosts the camera lens resolution in this example in the following way.

1. Fine features in the animated interference pattern can be much smaller and closer together than features that the conventional lens can make out.

2. As bright parts of the pattern dance across a contrast feature (i.e. strawberry) the feature reflects more light into the camera than when only dark parts of the pattern overlap with the feature. Each feature receives an intermittent light signal, and a region of features returns a total amount of light that changes in time according to similarities of the illumination pattern with the arrangement of features.

3. The blurry bowl seen by the camera seems to twinkle and blink as a result.

4. The sequence of twinkles and blinks are *like a Morse code message* from the strawberries. If they were in different locations, but with the same animated pattern, a subtly different series of twinkles and blinks—*a different message*—would result.

5. A computer program can interpret the message with knowledge of the modulation codes that animated the pattern. The program determines where the contrast features must have been to give rise to the message it received.

6. Knowing where the features are allows the computer to form an accurate picture of the target, strawberry locations and all, even though the lens could see only

a blurry bowl.

1.4.2 Computer programs

Numbers and pictures go hand-in-hand. Appreciation for the connections between (1) numbers and pictures and (2) pictures and understanding inspires scientists to use instruments with an ever-increasing capacity for accurately collecting numbers. This trend has culminated in today's inseparability of science and computers.

Unlike "old-fashioned" light microscopes with only an eyepiece for viewing target specimens, more and more cameras and microscopes used by scientists these days are equipped with electronic detectors, often connected to computers. This connection is central to SAM.

These sensor-computers are usually only used to measure, manipulate, or store picture information made by the physical sensor. SAM recognizes the power of computers to play a role in sensing the picture as well. The approach shifts complexity and precision from the physical world, where it is often expensive or unavailable, to the world of numbers in a computer, where complexity and perfect order call home.

In SAM, a computer is connected to

1. **a detector** of the (blurry) picture from the traditional lens as well as to
2. **the modulators** that place codes of information in each beam of the illuminator.

The computer interviews the target. It sends information through the beams that are like questions. The resulting interference pattern interrogates different combinations of target locations with bright and dark regions. At the same time, the computer observes the twinkling (but blurry) picture. The twinkling message is like an answer to the questions: when the pattern corresponds well with the unknown feature locations, more light is detected, and more is known about the feature locations. After enough questions and answers, the computer can deduce the locations of the contrast features *within every blurry blob in the picture.* Doing this will produce a

new picture, potentially with **thousands of times more detail than the physical lens or detector could otherwise capture.**

The physics of SAM provides a *linear solution* to the linear imaging problem. The SAM computation amounts to a matrix multiplication with a precomputable matrix inverse. See Sections 2.2 and 2.3 for the details. Lenses, on the other hand, offer *trigonometric physics*—Snell's Law and the equal angles law of reflection—that can never exactly provide a linear transformation. Conclusion: computers are fundamentally better at the mathematics of imaging than glass lenses and mirrors can ever be.

1.5 Background

1.5.1 Today's Microscopes

The popular mind set in today's scientific community divides the world of microscopes into two major types: (1) optical microscopes, and (2) those which, to be straightforward, I will call "fancy" microscopes.

Optical microscopes

By far the most familiar, useful, convenient, practical, and widespread type of microscope is the optical microscope. Features and customizations aside, these instruments are very much like the hallmark magnifying glass used by Sherlock Holmes to solve mysteries. Faced with a mysterious material, a researcher will often take a close look with an optical microscope before employing more specialized equipment.

Optical microscopes haven't changed much in a long time. In fact, the resolution of a standard optical microscope and the depth of field at a given resolution **have not improved in over a hundred years.** Those parameters have been optimized to their physical limit (given the traditional approach to forming images). Improvements have been in the two other parameters listed in Subsection 1.2.1: working distance and field of view. Figure 1-3 offers a view of the trends in these parameters that could

indicate practical limits, if not fundamental ones, that these last two parameters could be facing.

Possibly the most important development in optical microscopy in recent decades was its marriage to electronic detection. The advent of video and its connection to microscopy—called "video microscopy"—advanced the *quantitative analysis* of microscope imagery while also bringing a world of motion into focus. [10] Today's video microscopes are typically connected to computers.

Fancy microscopes

There has been a recent boom in **scanning probe microscopes.** In these instruments, a fantastically pointy stick (the "probe") scans a target surface like a blind man's cane exhaustively exploring the mosaic layout of a tile floor. Most like the cane are instruments that use the attractive or repelling forces of atoms, electric fields, or magnetic fields to or from each other. A second category uses the probe as a type of funnel for light or electrons. [29]

While only a decade ago these microscopes were rare, cumbersome, and fussy, today's commercial instruments are increasingly convenient and simple—made to resemble a toaster-oven both in appearance and in operation. Furthermore, they offer a resolution that approaches the size of the atom—*about ten thousand times smaller than optical microscopes can resolve.* Aside from maintenance and expense, the only significant drawback of scanning probe technology is its requirement to physically contact a substantially flat target.

When targets have a three-dimensional shape on the scale of features that need to be resolved[2], an **electron microscope** is the tool of choice. Electron microscopes can provide resolutions nearly as good as the scanning probe microscopes discussed above. Even in situations for which optical microscopes provide sufficient resolution, electron microscopes are often used when a large depth of field is required.

[2]An example to explain what it means to have a three-dimensional shape on a given size-scale: the Earth has a (round) three-dimensional shape on the scale of continents but effectively a (flat) two-dimensional shape on the scale of a football field.

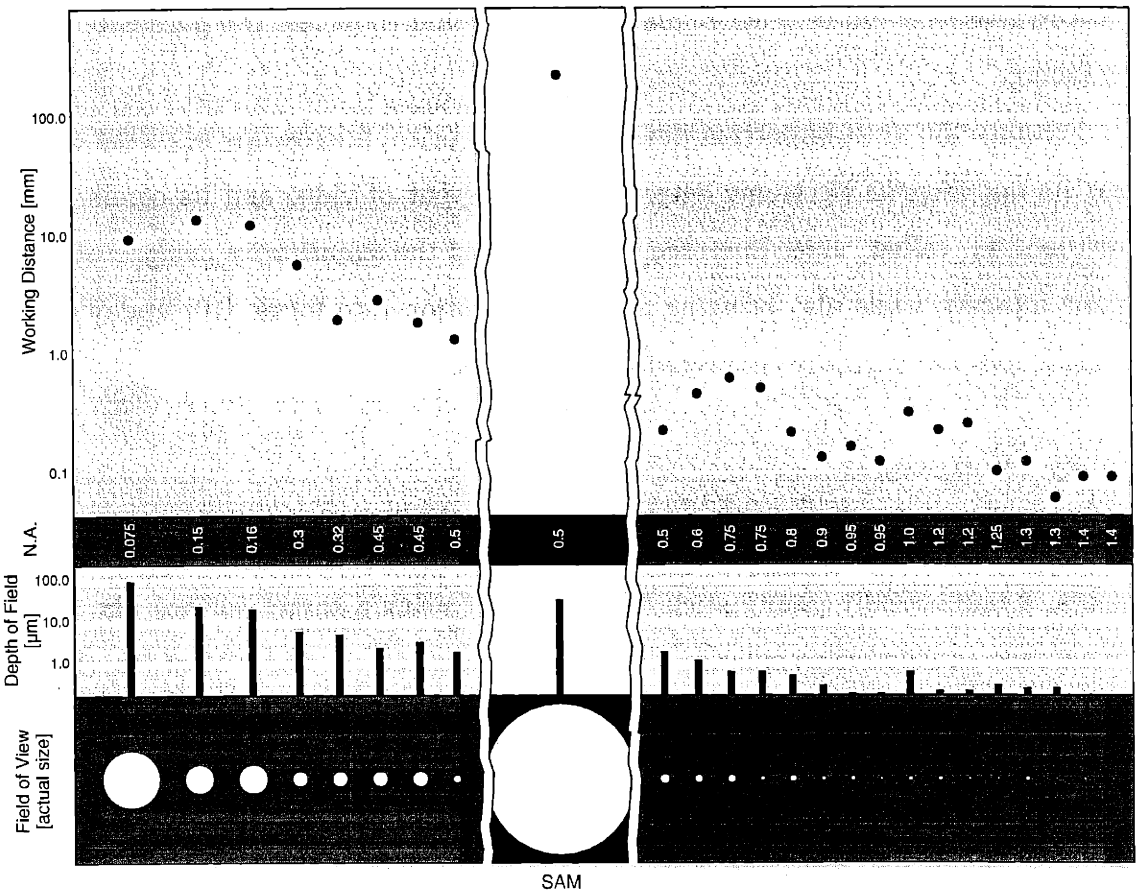

Figure 1-3: Traditional techniques in optical microscopy may be approaching physical or practical limits after the century of development on which today's lenses are based.

This chart summarizes four key performance parameters of microscope lenses from the leading manufacturer and of an example scaling of SAM that was worked out and simulated in detail. Imagers are sorted left to right by numerical aperture (N.A.), a convenient gauge of resolution. Each circle in the bottom section of the chart shows, in actual size, the imager's field of view—the area of the target seen by the imager. Above each circle, the depth of field for the corresponding imager is rendered as a bar showing the range of depths in the target that will be in focus. Symbols in the top section of the chart show the working distances of each imager. Working distance is the distance between the target and the instrument.

Working distance tends to decrease with increasing N.A. in conventional imagers. The lenses used for this plot conform to this trend spanning a range of over two orders of magnitude. The hundreds of millimeters of working distance afforded by the SAM approach departs significantly from the trend and potentially enables new capabilities in micro-surgery and remote inspection. Simultaneous departures from parallel trends in field of view and in depth of field potentially enable imaging of entire wafers or brains, as explained in the text. With visible light, the large available field of view, at the given numerical aperture, corresponds to a **billion-pixel image!** Such images can be realized because the technique is not limited by the pixel-count of electronic image detectors (such as CCD's).

These microscopes use a *lens made of electric and magnetic fields* (instead of glass or computer programs) to focus a stream of electrons to a tiny point from a narrow angle, like a needle. A detector senses how the electron point interacts with a spot on the target. Using additional, "deflection" fields, the microscope automatically scans the point around to probe an area of spots on the target surface as in the contacting scanning probe microscopes described above. However, since the probe in this case is a flowing stream of electrons, there is nothing that will "crash" into the target if it is bumpy.

The approach is like scanning a dinner plate under a fine stream of water from a faucet. Listening or feeling for the water that splashes back makes it easy to distinguish a portion of rice from the porcelain plate. Moving a pencil across a graph paper in step with the motion of the plate, and filling in each square on the graph paper in proportion with the degree of splash-back for the corresponding square of the dinner plate, will produce a map—a picture of sorts—of the dinner plate's contents. Because the water stream is not rigid, moving from the an empty region of the plate to a tall pepper will not break anything. Furthermore, because the stream is long and skinny, the spot where the water hits the top of the pepper will be nearly the same size as the spot hitting the plate. The water's likely tendency to disturb the arrangement of rice and to pool up in the plate are analogous to problems that complicate the interpretation of electron microscope images.

Using only fields and not matter, only *positive* lens elements can be made, and the resulting focus point is spread out typically a hundred times larger than the electron's wavelength. Therefore, the design limitation to use fields (and not matter) and to create a continuous aperture (and not a synthetic aperture) has prevented electron microscopes from achieving resolutions comparable to electron wavelengths. Early experiments using thin films of matter to create *negative* electron lenses to solve this problem have not been followed up sufficiently for general use. [18]

The major drawback of electron microscopes is a requirement to operate in vacuum. Air is made up of moving particles, generally much bigger than electrons, that would disrupt the electron stream if present. To continue the analogy from above,

operating an electron microscope in the presence of air would be like having a machine gun shooting bullets across the stream of flowing water. So that the electron stream is not dispersed, all of the air in a vessel enclosing the electrons and the target is pumped out. Changing targets usually involves repeating the inconvenient pumping process.

Another drawback is the difficulty of preparing many types of targets to provide good, meaningful contrast with electron streams. The most common step involves coating the target with metal to prevent the pooling up of electrons.

Several **other fancy microscopes** such as acoustic and X ray have advanced niche application areas in science, but on the whole, they have not yet made a broad impact. This is because, so far, they have lacked the combination of abilities needed for widespread acceptance. Those abilities include

1. to be broadly useful,
2. to be affordable to a research group or facility,
3. to create pictures quickly, and
4. to offer a justifiable advantage over the mainstay of microscopy: conventional optical microscopes.

1.5.2 Previous work on synthetic aperture microscopy

There have been two types of synthetic aperture microscopes previously investigated by other researchers:

1. microscopes that use **one laser beam** at a time, and
2. microscopes that use **two laser beams** at a time.

Figure 1-4 illustrates how these types are fundamentally different from each other and how the SAM approach in this thesis relates to them.

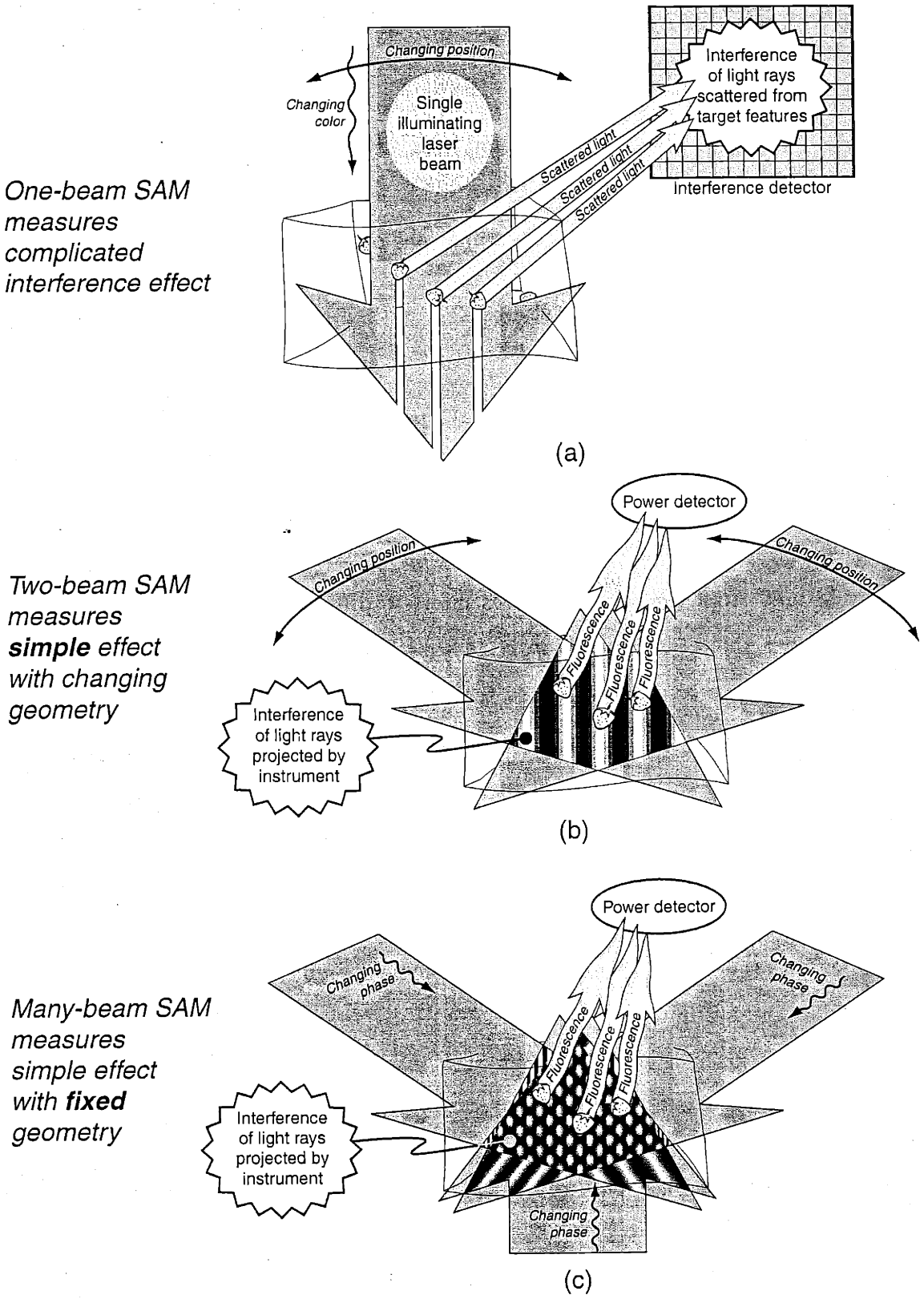

Figure 1-4: The three types of synthetic aperture microscopes: (a) one-beam SAM, (b) two-beam SAM, and (c) the new, many-beam SAM concept of this thesis.

One-beam synthetic aperture microscopes

In the first category, the "synthetic aperture" is formed with one beam and (usually) an array of detectors.

These microscopes work by analyzing how the light rays reflected from target features interfere with each other. As with SAM ideas in this thesis, these instruments "interview" the target over time. Combinations of these three interrogation tactics describe the existing devices:

1. the laser beam color is changed,

2. the laser beam position is changed, and/or

3. the detector position or set of positions is changed.

The simplest technique changes the color and uses a fixed arrangement of parts. Light scattered from the target features interfere differently depending on the distances between them *measured in wavelengths.* Therefore, as the wavelength (a distance related to the color) changes, so too will the interference patterns landing on the detectors.

Imagers of this type have not yet been built as microscopes per se. Instead, they have been more like macro-cameras for three-dimensionally measuring targets in the centimeter to meter size-range. [13, 16, 22] Scaling the concept to the microscopic scale would require such large changes in color as to introduce many complicated physical effects that would generally confuse the measurement, even if an appropriate light source were available.

Another approach is more like Synthetic Aperture Radar (SAR). In SAR, an airplane traces out a long arc as it flies over a target. The arc of space, traced out over time, behaves like a huge lens. The related one-beam SAM approach has a changing geometry, but typically a fixed color. In mimicry of an airplane flying over a field, a light source and detectors trace out paths in the space above and around the target.

One company (Essex Corporation, Columbia, Maryland) is attempting to develop a commercial SAM product of the SAR variety. Their "Virtual Lens Microscope" is based on rapidly changing the beam position with a large array of detectors. [26, 27]

One-beam SAM offers several useful **advantages over mainstream microscope techniques.** Most important among them are listed here.

1. The decoupling of working distance from resolution
2. Providing quantitative phase contrast
3. Shifting precision optics to computation
4. Offering possible imaging solutions in wavelength regimes for which high resolution optics are not feasible

Providing three-dimensional, volumetric imagery is often cited as an advantage of one-beam SAM by its proponents, but the capability is also offered by a mainstream variant of the conventional optical microscope, called a "confocal" optical microscope.

The principal disadvantage of one-beam SAM the large amount of computation required to make pictures. In time, as computers get faster, one would expect this problem to vanish. On the other hand, the size of the problem can quickly overrun any generation of computer given sufficient acquisitiveness for field of view.

The problem is fundamental: global interactions require a global inverse. In the one-beam SAM approaches, all of the many features within the target communicate with each other through light. Observing the outcome of this communication with electronic detectors is a simple matter. Connecting such observations with an image of the object, on the other hand, requires a daunting calculation. Computer programs based on three-dimensional Fourier transforms are more common, but equally intimidating as the so-called "inverse problem" approach to making pictures from one-beam SAM data. [11, 13, 14, 25, 26]

The proposed product (from Essex Corporation) would come with a special *optical computer* (rather than an electronic one) in order to make pictures at a reasonable rate. [28]

Perhaps because the technology is so new, one-beam SAM micrographs to date have been of visibly poor quality relative to today's norm for microscope pictures.

Since it relies on light from points on the target to interfere with each other observably, one-beam SAM requires the light from the feature points to be *mutually coherent*—like laser-light from the same laser. This requirement translates to certain qualities of the laser beam needed to illuminate the target and to the types of contrast available from the technique. Specifically, a powerful type of contrast that can map and identify chemicals and proteins in biological samples, namely "fluorescence contrast," is not available to one-beam SAM. The viability of fluorescence contrast is your key for identifying whether an imager is based on a one-beam concept or on the multiple beam concepts described below.

Two-beam synthetic aperture microscopes

Rather than relying on target features to produce interfering beams of light, the two-beam concept projects an interference pattern over the target and collects light returned from the target *incoherently.* In other words, the effect on which one-beam SAM is based—interference of the *mutually coherent* scattered light sources in the target—is suppressed so that **a different physical effect** can shine through.

The most straightforward way to suppress the effect of observable interference of points in the target is to work with *fluoresced* light. In fluorescence imaging, laser light is absorbed by dye molecules in the target. The laser power energizes the molecules which emit their own, new light rays, typically with a different, longer-wavelength color. Therefore, light rays from different regions in the target are "born again" in their respective regions, and no pair of rays from different regions will share the values required for "slow dance," observable interference. As you will understand soon, two-beam SAM delivers the resolution of the (finer) illuminating wavelength and not the (coarser) fluoresced wavelength.

Fluorescence imaging offers a fantastically powerful way to see the world. Fluorescent dye molecules can be *chemically programmed* to preferentially stick to certain structures or chemicals in a target specimen so that, for example, with a brain as

a target, only the wires are dyed. Different dyes can be used in combination, each with a different fluoresced or exciting wavelength and each with a different chemical program. To continue the brain imaging example, wires (axons) could appear red, connectors (synapses) could appear green, and brain signals (in the form of chemical ionic concentrations) could appear orange to the instrument. Many dyes are commercially available and many researchers make their own as well. Dyeing a biological sample is accomplished by applying a drop of the dye with a dropper and letting it soak in.

When fluorescence is not an option, and when observable interference of target points does occur, two-beam SAM can *ignore*, rather than suppress the effect. This is possible through *spatial averaging*. Because the interference pattern generated by the target can be very fine, spatial averaging can be easy to arrange. The trick is to provide many details of the interference pattern to your detector so that its one-number summary of how much light it collects at a given time is relatively insensitive to changes in the pattern details. In one-beam SAM, those detail changes are the *signal* while in two-beam SAM (not using fluorescence), they are the *noise*. By analogy, if you wanted to gauge the excitement-level of a crowd at a football game, you could listen to one fan's display of enthusiasm, which, typically, would come in spurts. A method that would provide a more steady measure of the excitement-level would be to provide many details of the sound pattern to your ear—to listen (perhaps from the inevitable blimp) the average of thousands of fans: a constant roar.

Two-beam SAM illuminates a target with a particular series of interference patterns: the set of "sinusoidal fringe" patterns that arise with a single beam-pair. Fringe patterns are made of alternating, equally spaced, parallel sheets of light and dark. The "spatial frequency" (the pitch or spacing between the sheets), the "spatial phase" (where the first light sheet starts), and the angular orientation (which way the stack of sheets aims) are all controlled by the directions of the beams and "phase information" carried in the two beams. In accordance with a famous mathematical theorem, "Fourier's Theorem," [23] the amount of light seen by a detector looking at the target with fringes hitting it (the integrated product of the three-dimensional sinusoidal

fringe pattern with the unknown distribution of scatterers in a target) sampled with a sufficient number and arrangement of fringe parameters (phase, frequency, and orientation) supply a computer with the data needed to make a picture.

The two-beam concept has an **important advantage over the one-beam** concept: because it relies on interference with known beams, and because it suppresses or ignores the communication (through interference) of target features with each other, two-beam SAM does not require a global inverse. Functionally, this means that **regions of the target can be treated independently**, both in the detecting and in the computer reconstruction steps of sensing an image. It also means that **lenses can be used** for part of the job, to divide the target into small regions. Each region would require a far easier computation to get to the corresponding final high-resolution segment of the picture than the equivalent fraction of the global problem.

Therefore, two-beam SAM solves the main problem with the one-beam technique by relying on a different physical principle.

The two-beam synthetic aperture microscope has been reported *conceptually* twice before [2, 9]. My best efforts did not turn up evidence of experimental work on the idea to date.

The principal drawback of the two-beam SAM concept (as described before), is a limitation in the **speed of data collection.** Whereas computing the picture had the potential to be relatively quick, acquiring the data to feed the computer program was limited to mechanical speeds—the speeds of motorized or robotic devices to position and re-position things like mirrors, stepping through a long sequence of fringe parameter settings.

As this problem is not a fundamental one, it is merely work for an engineer...

1.6 The scope(s) of my thesis

Two-beam SAM contains the physical concept on which this thesis is based. Confirming the previously reported *concept* was the starting point of the *experimental* work for

this thesis. Advancing a many-beam extension of the technique, *both conceptually and experimentally*, comprised the important remainder of the project. Many-beam SAM addresses the two-beam limitation to mechanical speeds with a **fixed architecture containing no moving parts.**

The following two chapters chronicle my theoretical (Chapter 2) and experimental (Chapter 3) work on SAM. The theory is a generalized view of two-beam SAM that applies to any number of beams with any choice of modulation strategy. The experimental work began with a confirmation of the physical principles of two-beam SAM and continued to work out, completely, the principal engineering challenge of many-beam SAM: the controllable many-beam illuminator.

The result is the basis for a microscope that overcomes the problems of previous synthetic aperture microscopes along with many of the limitations of other conventional techniques. Because of its compatibility with lens technology, it can be implemented as a *lens enhancement device* offering "the best of both worlds" in a wide range of application areas.

Chapter 2

Theory

The imaging concept explored in this thesis is designed to form an image of a target object. In the concept, image information is computed from the target's measured response to a sequence of pattern exposures. The method links these three steps:

1. **Illumination:** A target is illuminated with a known sequence of high-resolution patterns. (The high-resolution patterns projected on the target are the interference patterns that form at the intersection of coherent laser beams.)

2. **Target interaction:** The target sends energy to a detector in proportion with correspondences to these patterns. (The target can respond by scattering relatively more light at a detector when bright areas of the pattern fall on target locations containing relatively high densities of scatterers.)

3. **Signal interpretation:** Image information about the target is refined from the detector's output signal. (Digital signal processing methods are used to estimate a spatial Fourier transform component of this density of scatterers for each pair of laser beams.)

2.1 Many-beam illumination

Consider n laser beams, all passing through some volume of space for some period of time. Figure 2-1 shows three such beams overlapping in a labeled region. The

beams are of unit amplitude[1] and have direction vectors, $\vec{k}$, linear polarization in the $\hat{p}$ direction, optical frequency, ω, and optical phase as a function of time, $\phi(t)$.[2] We write the electric field, $\vec{E}$, of one of those beams as a function of time, t, and space, $\vec{r}$, as

$$\vec{E} = \cos(\vec{k} \cdot \vec{r} - \omega t + \phi(t))\, \hat{p}, \tag{2.1}$$

and the electric field in the space where the beams all overlap as the sum of all n of those fields:

$$\begin{aligned} \vec{E}_{overlap} &= \sum_{l=1}^{n} \vec{E}_l & (2.2) \\ &= \sum_{l=1}^{n} \cos(\vec{k}_l \cdot \vec{r} - \omega_l t + \phi_l(t))\, \hat{p}_l. & (2.3) \end{aligned}$$

The intensity in that space is given as

$$\begin{aligned} I &= \left|\vec{E}_{overlap}\right|^2 = \vec{E}_{overlap} \cdot \vec{E}_{overlap} \\ &= \left[\sum_{l=1}^{n} \cos(\vec{k}_l \cdot \vec{r} - \omega_l t + \phi_l(t))\, \hat{p}_l\right] \cdot \left[\sum_{m=1}^{n} \cos(\vec{k}_m \cdot \vec{r} - \omega_m t + \phi_m(t))\, \hat{p}_m\right] & (2.4) \\ &= \sum_{l=1}^{n}\sum_{m=1}^{n} \cos(\vec{k}_l \cdot \vec{r} - \omega_l t + \phi_l(t))\, \cos(\vec{k}_m \cdot \vec{r} - \omega_m t + \phi_m(t))\, \hat{p}_l \cdot \hat{p}_m & (2.5) \\ &= \sum_{l=1}^{n}\sum_{m=1}^{n} \frac{1}{2}\, [\, \cos((\vec{k}_l + \vec{k}_m) \cdot \vec{r} - (\omega_l + \omega_m)t + (\phi_l(t) + \phi_m(t))) + \\ &\qquad \cos((\vec{k}_l - \vec{k}_m) \cdot \vec{r} - (\omega_l - \omega_m)t + (\phi_l(t) - \phi_m(t)))\,]\, \hat{p}_l \cdot \hat{p}_m. & (2.6) \end{aligned}$$

Given that we plan on sensing intensity information with an electronic detector, the first cosine term in Equation 2.6 (which oscillates at the optical frequency $(\omega_l + \omega_m)$) contributes only its time-average: zero. To further simplify the expression for I, we

[1] In a real setup, the beam amplitudes could differ from each other. The present derivation neglects this factor for simplicity.

[2] For the sake of a few examples, the phase function, $\phi(t)$, can account for an optical phase modulation imposed on a beam by a phase-shifting device, a frequency-shift from ω, or the phase-jitter of a laser's output which gives rise to its optical linewidth.

can take all of the ω's to be the same.[3] Now we have:

$$I = \sum_{l=1}^{n} \sum_{m=1}^{n} \frac{1}{2} \left[0 + \cos((\vec{k}_l - \vec{k}_m) \cdot \vec{r} + (\phi_l(t) - \phi_m(t))) \right] \hat{p}_l \cdot \hat{p}_m. \tag{2.7}$$

Rearranging the double-sum into the case with $l = m$ and the case with $l \neq m$, and exchanging l and m in those cases with $l > m$ gives

$$\begin{aligned} I \;=\; & \sum_{l=1}^{n} \frac{1}{2} \cos(0) + \\ & \sum_{l=1}^{n-1} \sum_{m=l+1}^{n} \frac{1}{2} \left[\cos((\vec{k}_l - \vec{k}_m) \cdot \vec{r} + (\phi_l(t) - \phi_m(t))) + \right. \\ & \qquad \left. \cos((\vec{k}_m - \vec{k}_l) \cdot \vec{r} + (\phi_m(t) - \phi_l(t))) \right] \hat{p}_l \cdot \hat{p}_m. \end{aligned} \tag{2.8}$$

Finally, renaming $\vec{k}_{lm} \equiv (\vec{k}_l - \vec{k}_m)$, and $\phi_{lm}(t) \equiv (\phi_l(t) - \phi_m(t))$, and simplifying, we are left with

$$I = \frac{n}{2} + \sum_{l=1}^{n-1} \sum_{m=l+1}^{n} \cos(\vec{k}_{lm} \cdot \vec{r} \; + \; \phi_{lm}(t)) \, \hat{p}_l \cdot \hat{p}_m. \tag{2.9}$$

This speckle intensity pattern is the combination of a constant term and $\frac{n(n-1)}{2}$ spatial fringe interference patterns whose spatial phase is given by the optical phase difference of the pair of beams that produced it. Figure 2-2 illustrates this combination in a case with $n = 3$.

You can see from Equation 2.9 that the contrast of these fringes is influenced by polarization; namely, the modulation of a given fringe pattern is $\hat{p}_l \cdot \hat{p}_m$. From here on, let's assume, for simplicity, that the optical system is engineered so that all n beams have substantially the same polarization—that is, take $\hat{p}_l \cdot \hat{p}_m$ to equal 1.

Figure 2-3 shows, as a preview of the class of patterns used in the experimental setups, an intensity pattern from a circular array of 41 laser sources, and, as a hint of the parallelism described in Subsection 2.4.5, a pattern from a square array of 49 sources. These patterns comprise $\frac{41 \times 40}{2} = 820$ and $\frac{49 \times 48}{2} = 1176$ fringe patterns, respectively. The size scale and overall structure of the patterns depends on the

[3] This operation sacrifices no generality, since we can put any optical frequency differences into the phase functions.

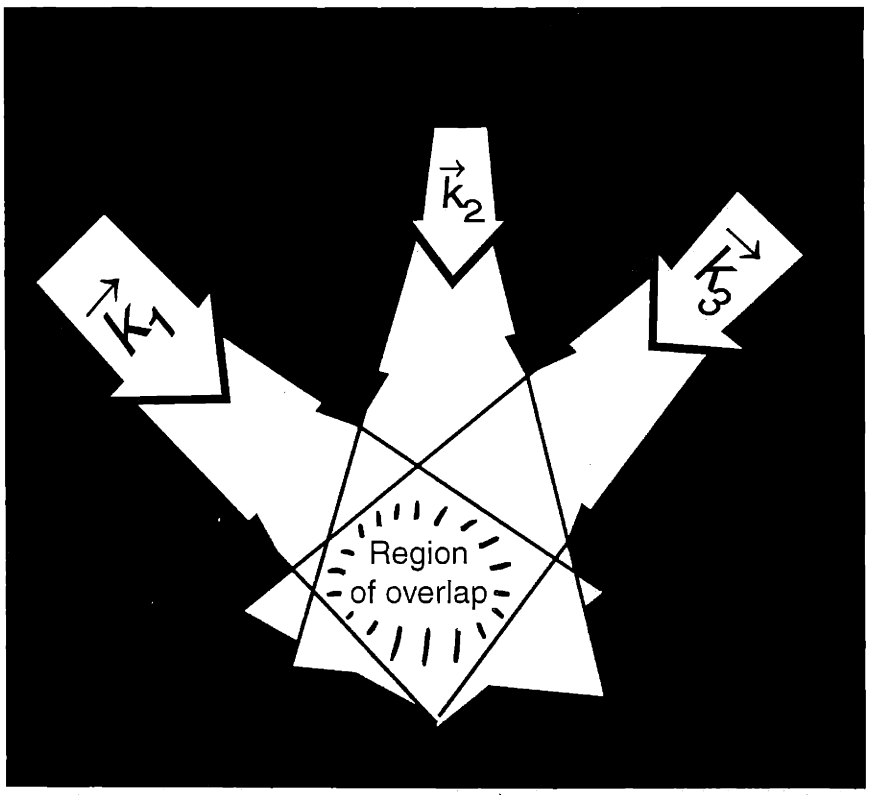

Figure 2-1: Each of three beams fills a region with coherent laser light.

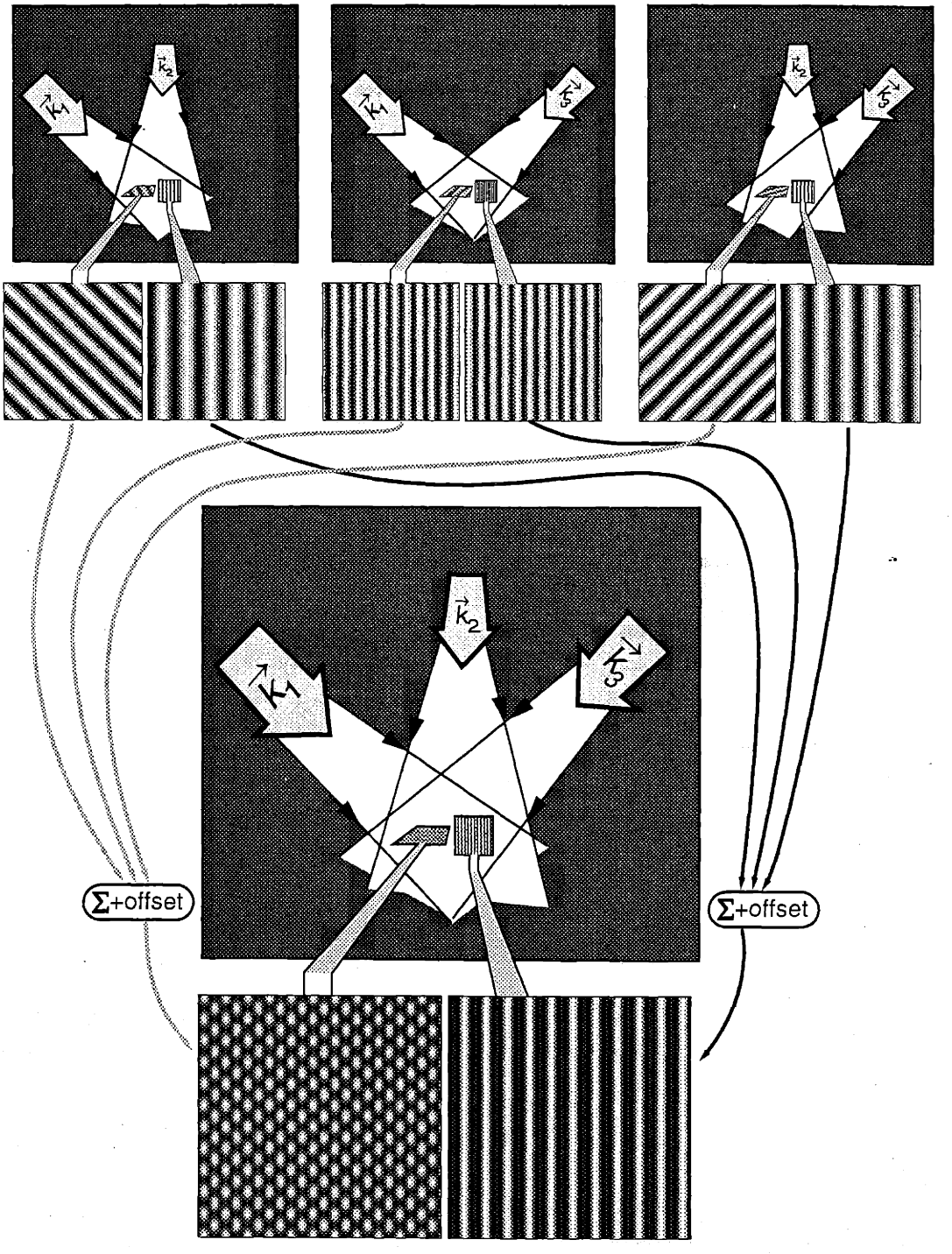

Figure 2-2: The intensity pattern where three beams overlap is synthesized by the patterns of each pair of beams. The interference pattern exists throughout the (three dimensional) region of overlap; each case in the figure is shown with a horizontal and a vertical "slice" through that 3D intensity pattern.

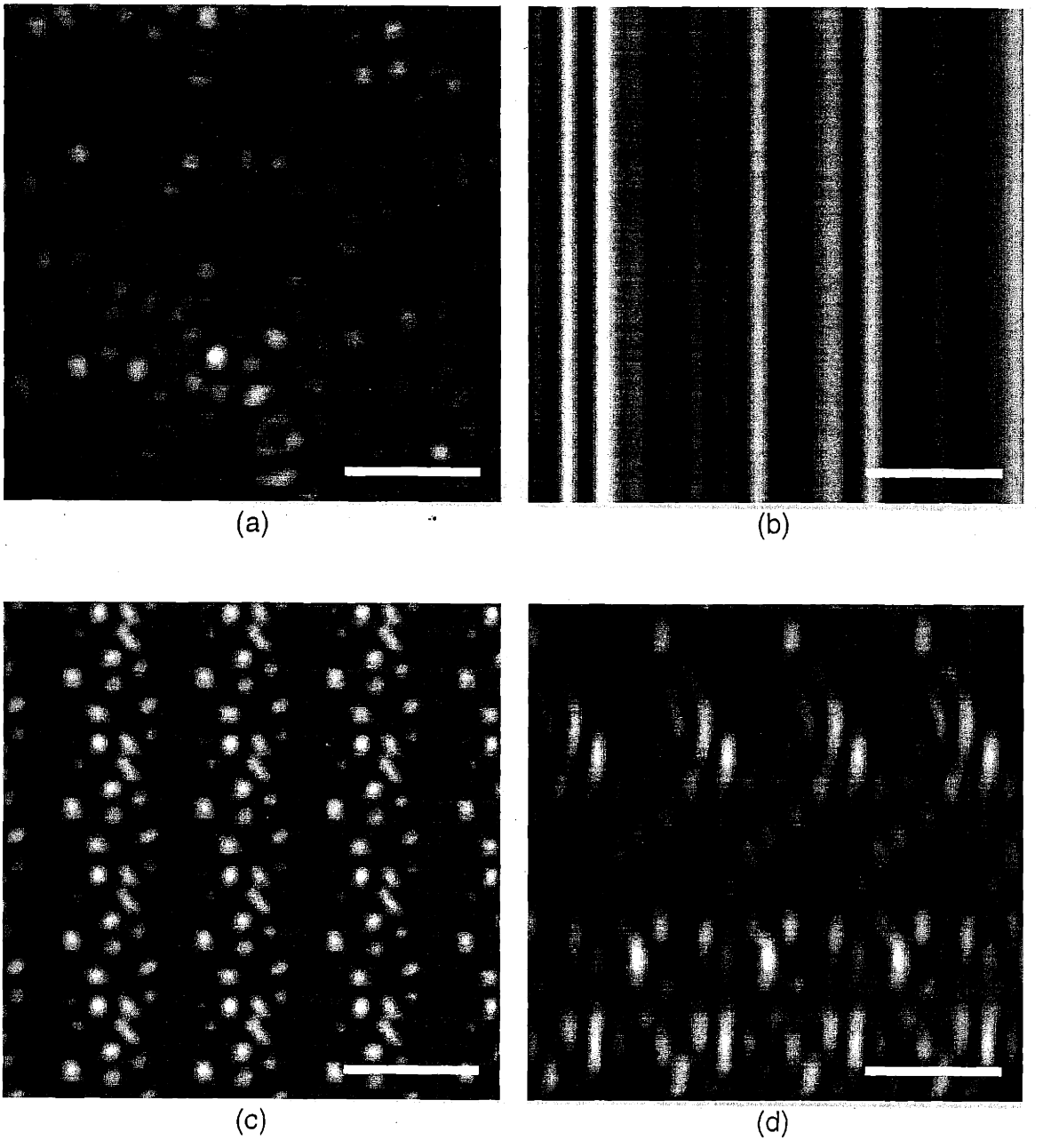

Figure 2-3: The top pair of images shows (a) a horizontal slice and (b) a vertical slice through the illumination pattern from a ring of $n = 41$ beams. Parts (c) and (d) are horizontal and vertical slices through the interferogram formed with a 7 by 7 regular grid of beams ($n = 49$). All four scale-bars represent 1 micrometer for the particular optical wavelength and geometry used to synthesize the images.

relative angles of the input beams, but the specifics of these images depend on the relative phases of the beams (chosen arbitrarily for the figure).

Significance

Whereas projecting an arbitrary high-resolution pattern generally requires a sophisticated system of lenses, a class of richly complex, high-resolution, three-dimensional patterns can be created by intersecting surprisingly few laser beams.

2.2 Target Interaction

Now, let's introduce a target into the region of space containing the intensity pattern I. We can represent that target by a contrast value as a function of position, $C(\vec{r})$, which is the target's response to some linear sensing process.[4] Then, the signal, T, of a detector positioned to receive light which has interacted with a region of the target is the integral over that region[5] of the product of C and I:

$$T = \int C(\vec{r})\, I(\vec{r}) d\vec{r}. \tag{2.10}$$

[4]Three examples of processes in which a detector's output can be linear in I and C are the detection of transmitted, reflected, or fluoresced light from the target.

[5]Assuming incoherent summation as in fluorescence. In scattering or transmission, the coherence of light rays from different features within the target adds *speckle noise* to the signal. See Section 2.5.1 for more on speckle noise.

Plugging Equation 2.9 in for I and expanding strategically, we find

$$T = \int C(\vec{r}) \left[\frac{n}{2} + \sum_{l=1}^{n-1} \sum_{m=l+1}^{n} \cos(\vec{k}_{lm} \cdot \vec{r} + \phi_{lm}(t)) \right] d\vec{r} \tag{2.11}$$

$$= \frac{n}{2} \int C(\vec{r})\, d\vec{r} + \int C(\vec{r}) \left[\sum_{l=1}^{n-1} \sum_{m=l+1}^{n} \cos(\vec{k}_{lm} \cdot \vec{r} + \phi_{lm}(t)) \right] d\vec{r} \tag{2.12}$$

$$= \underbrace{\frac{n}{2} \int C(\vec{r})\, d\vec{r}}_{\mathcal{A}} + \underbrace{\sum_{l=1}^{n-1} \sum_{m=l+1}^{n} \left[\overbrace{\cos(\phi_{lm}(t))}^{\mathcal{C}} \overbrace{\int C(\vec{r}) \cos(\vec{k}_{lm} \cdot \vec{r})\, d\vec{r}}^{\mathcal{D}} - \overbrace{\sin(\phi_{lm}(t))}^{\mathcal{E}} \overbrace{\int C(\vec{r}) \sin(\vec{k}_{lm} \cdot \vec{r})\, d\vec{r}}^{\mathcal{F}} \right]}_{\mathcal{B}}. \tag{2.13}$$

We can see from Equation 2.13 that the detector signal is the sum of a constant term (part $\mathcal{A}$) and a time-dependent double-sum (part $\mathcal{B}$). The double-sum in part $\mathcal{B}$ enumerates every pair of laser beams: $\{l, m\}$. For each of those pairs, a contribution to the total signal is made by the cosine transform coefficient of $C(\vec{r})$ at the vector spatial frequency synthesized by $\cos(\vec{k}_{lm} \cdot \vec{r})$ (part $\mathcal{D}$) modulated by a time-dependent encoder (part $\mathcal{C}$), and the negative of the sine transform coefficient of $C(\vec{r})$ at the vector spatial frequency synthesized by $\sin(\vec{k}_{lm} \cdot \vec{r})$ (part $\mathcal{F}$) modulated by a different encoder (part $\mathcal{E}$).

Significance

If we demodulate all of the $\mathcal{D}$'s and $\mathcal{F}$'s from the detector signal (by using our knowledge of the carriers $\mathcal{C}$ and $\mathcal{E}$), we will have measured order n^2 spatial Fourier transform coefficients of the target's contrast.

2.3 Signal multiplexing and separation

The time-dependent encoders (parts $\mathcal{C}$ and $\mathcal{E}$ of Equation 2.13) modulating the sine and cosine transform coefficients are both derived from the $\phi(t)$'s, a set of n phase

sequences potentially under our control. To separate the signals well, linear algebra requires the time-dependent encoders to be pairwise orthogonal or nearly so. The two approaches discussed below meet this orthogonality requirement in different ways that seem natural in light of communication schemes in common use today. Namely, the coefficients can be transmitted either on a family of sinusoids with different frequencies (exactly as in FM radio) or with a spread-spectrum approach employing a family of independent "random" seeds.

2.3.1 Individual carrier frequency technique

Subsection 2.3.2 sets up the basis for processing the detector signal for any choice of $\phi_{lm}(t)$'s, including the sinusoids discussed here. The following example is a primer, meant to connect with the reader on an intuitive level.

The individual carrier frequency technique boils down to modulating the n beams with different frequency-shifts from a reference laser.

Consider a single pair of beams, $\{l, m\}$, with frequencies $\{\omega_l, (\omega_l - \Delta\omega_{lm})\}$. That pair contributes to the total intensity pattern (see Equation 2.9) the term $\cos(\vec{k}_{lm}\cdot\vec{r} + \phi_{lm}(t))$ where $\phi_{lm}(t) = -\Delta\omega_{lm} \times t$. This is a fringe pattern whose spatial phase rolls across the target at the rate of a full fringe spacing every $\frac{2\pi}{\Delta\omega_{lm}}$ seconds. Figure 2-4 illustrates how a stationary target illuminated by this pattern-sequence "transmits" to a detector the amplitude and phase of the individual Fourier component probed by the pair $\{l, m\}$ at the carrier frequency $\Delta\omega_{lm}$.

In a system of n beams, the corresponding n frequency shifts would be selected so that the frequency differences of all pairs of beams are unique. Then, the total illumination pattern sequence consists of order n^2 fringe patterns traveling in their various directions, each at a different speed relative to its own period. Measuring the amplitude and phase (relative to some point) of a certain probed Fourier component of the target is exactly accomplished by measuring the amplitude and phase (relative, perhaps, to a phase-reference signal[6]) of the "radio station" associated with the beam-

[6]Such a signal could be derived from the drive voltage to the laser beam phase modulators or from a phase-reference detector located in the region of overlap.

pair that probes that Fourier component.

2.3.2 Spread-spectrum encoding

The idea of spread-spectrum encoding is to arrange that each of the illuminating beams has an optical phase (relative to some chosen beam), $\phi(t)$, that we know (by controlling or measuring it), that is uniformly randomly distributed between $-\pi$ and π, and whose value at any given time is statistically independent of the optical phase of any other beam. If those criteria are satisfied, then all of the $\phi_{lm}(t)$'s are statistically independent of each other, and moreover, the time-dependent encoders derived (parts $\mathcal{C}$ and $\mathcal{E}$ of Equation 2.13) serve their purpose of synthesizing a signal which can be parsed for recovery of all sine and cosine transform coefficients probed by the illumination pattern.

Data reduction

Rewriting Equation 2.13 as a matrix equation is straightforward:

$$\underbrace{T}_{b} = \underbrace{\begin{bmatrix} 1 & \mathcal{C}_{12}(1) & -\mathcal{E}_{12}(1) & \mathcal{C}_{13}(1) & -\mathcal{E}_{13}(1) & \dots & \mathcal{C}_{(n-1)\,n}(1) & -\mathcal{E}_{(n-1)\,n}(1) \\ 1 & \mathcal{C}_{12}(2) & -\mathcal{E}_{12}(2) & \mathcal{C}_{13}(2) & -\mathcal{E}_{13}(2) & \dots & \mathcal{C}_{(n-1)\,n}(2) & -\mathcal{E}_{(n-1)\,n}(2) \\ & & & & \vdots & & & \\ 1 & \mathcal{C}_{12}(m) & -\mathcal{E}_{12}(m) & \mathcal{C}_{13}(m) & -\mathcal{E}_{13}(m) & \dots & \mathcal{C}_{(n-1)\,n}(m) & -\mathcal{E}_{(n-1)\,n}(m) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathcal{A} \\ \mathcal{D}_{12} \\ \mathcal{F}_{12} \\ \mathcal{D}_{13} \\ \mathcal{F}_{13} \\ \vdots \\ \mathcal{D}_{(n-1)\,n} \\ \mathcal{F}_{(n-1)\,n} \end{bmatrix}}_{x} \tag{2.14}$$

Here, m is the number of time samples taken during a measurement.

This is a standard linear system with the form $b = Ax$. We know b from the target detector and A from measuring or controlling the $\phi(t)$'s. The information contained

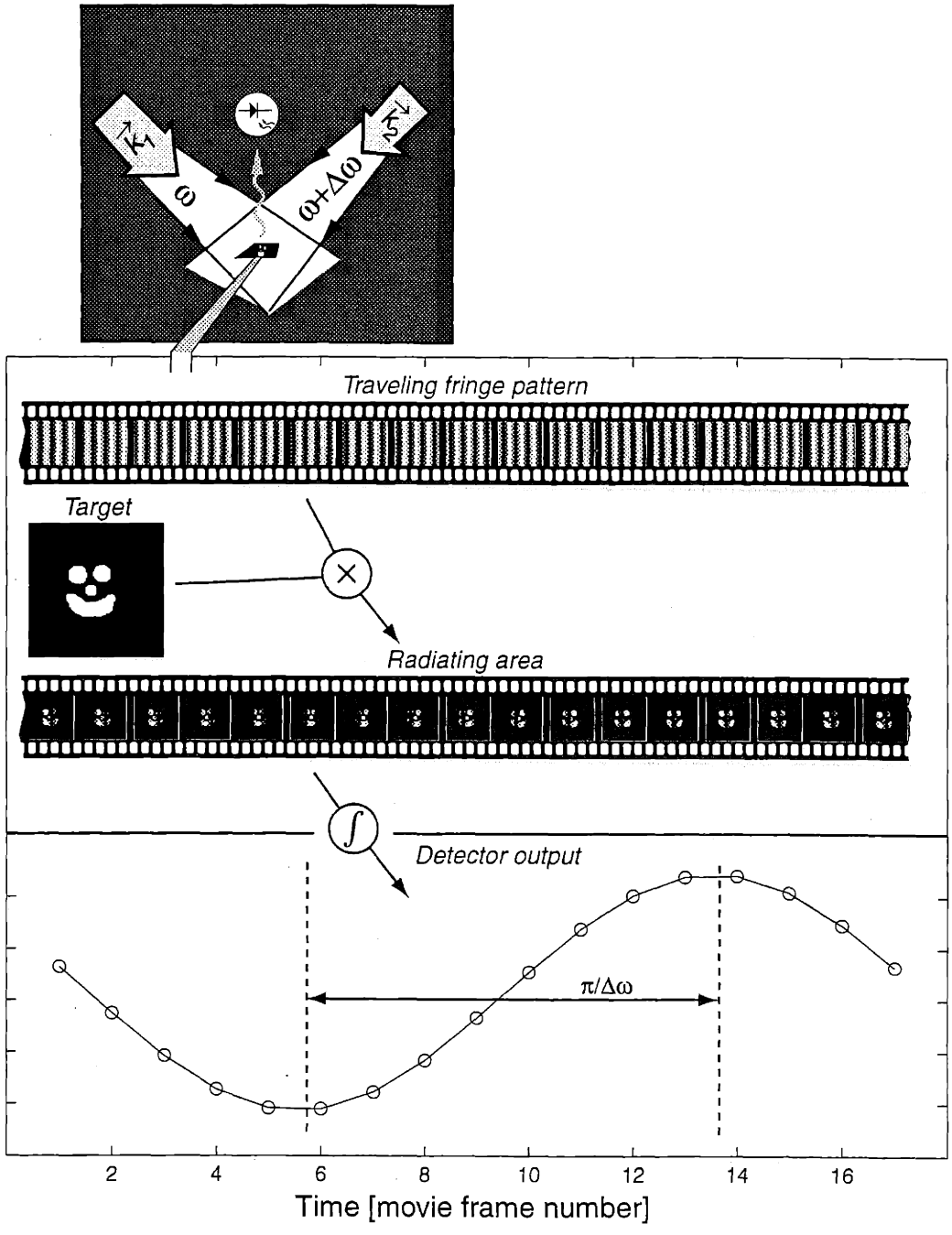

Figure 2-4: Two laser beams with optical frequency difference $\Delta\omega$ overlap in the region containing a target. The traveling fringe pattern projected by the pair sends more light to bright regions of the target at some times than at other times. The signal from a single photodiode positioned to receive some of the scattered light reveals at which spatial phase and to what degree the target has the most correspondence with the fringe pattern.

in x represents the target in Fourier space and can be extracted[7] using the formula

$$x = A^{+}b \tag{2.15}$$

where A^{+} is the Moore-Penrose pseudoinverse of A [15]. The pseudoinverse solves the system in a linear-least-squares sense, and computer code that implements it is widely available [19]. A good choice of experimental parameters is necessary to ensure that the linear-least-squares solution is correctly constrained [24].[8]

Significance

Now we have the mathematical means to convert a time-sequence of intensities from a single photo-detector into order n^2 spatial Fourier transform coefficients of the unknown target.

2.4 Implementing the concept

This section introduces aspects of a physical setup that must be considered to plan a useful measurement using the abstract framework outlined above.

2.4.1 Positioning the sources

This imaging concept can be implemented to serve many purposes. The specific purpose of a given setup dictates which transform coefficients are important to measure. To wit, a system needs to collect different information for acquiring two-dimensional images of an unknown target as opposed to for acquiring three-dimensional images

[7] Conventional spread-spectrum communications theory predicts a certain signal-to-noise ratio (SNR) given a frequency-spread transmitter in a hostile environment of many other frequency-spread jammers [4]. In this setup, however, we have the advantage of knowing all of the "jamming" carriers since they are all derived from their own, known $\phi_{lm}(t)$'s.

[8] Notice that Equation 2.14 is the same as Equation 2.13 and makes no assumption about random encoding seeds versus any other choice of encoder. In the individual carrier frequency technique (of Subsection 2.3.1), the matrix A^{+} of Equation 2.15 computes the discrete Fourier transform (DFT) of the detector signal and thus "tunes in" to each "radio station" in a row-wise manner.

of an unknown target, or for measuring parameters of a target with some known information. This subsection explains just which Fourier transform coefficients will be measured.

Our ability to position the n k-vectors as we choose constitutes our (incomplete) freedom to position the Fourier transform coefficients. The freedom is incomplete insofar as n free parameters determine order n^2 outcomes.

For each pair of beam k-vectors, a Fourier sample arises at the difference-vector. The spatial frequency, F, is determined by the angle, θ, between the beams, and the refractive index, n_r, of the medium in which the fringes are made:

$$F \approx \frac{2\, n_r \sin(\frac{\theta}{2})}{\lambda} \tag{2.16}$$

in the approximation that both beams have essentially the same wavelength, λ, during the period of observation.[9]

Figure 2-5 shows example geometries of the source-beams and the probe patterns that result. Each part on the left side of that figure represents a collection of laser beams aiming at a region of overlap where a target's Fourier space is probed; (a) is a single pair of beams, (c) is a ring of 41 beams, and (e) is a 7×7 grid of beams. The three points in (b) represent the single fringe pattern created by the beam-pair of (a) and correspond to a positive spatial frequency, the negative of that frequency, and DC. Three points are necessary to represent a physical fringe pattern, which is both real-valued and positive-biased. Notice that the samples in (d) arising from the ring in (c) are actually 3D Fourier samples which happen to lie in the DC plane of the vertical dimension. Thus, no vertical information is encoded by these samples, and the depth of field of the instrument is limited only by the beam diameter. The cloud of points in (f) resulting from the grid of sources in (e) can represent a sample in 3D, but with less resolution in the vertical than in the remaining dimensions.

[9]Notice that if we define numerical aperture (*N.A.*) as $n_r \sin(\frac{\theta}{2})$, Equation 2.16 exactly predicts the famous $\frac{\lambda}{2\,N.A.}$ limit [7, p. 498] for the resolving power of an optical microscope in critical illumination.

Dimensionality

Notice that a collection of source-vectors all with the same value in a given dimension can have no resolution in that dimension. This fact explains how a ring of sources all having the same elevation angle above the target constitutes an ideal two-dimensional scanner—a scanner that projects all target contrast into a single plane of range. (See Figure 2-5 (c) and (d) and Figure 2-3 (a) and (b).) This geometry would make a microscope convenient for visualizing irregular, bumpy targets, obviating the need to choose a single plane of focus or, in fact, to focus at all.

To collect 3D information, two or more view-angles can be scanned with a 2D instrument and combined using a tomographic algorithm. Alternatively, a 3D instrument can be made by ensuring that source beams enter the target from a variety of azimuth and elevation angles. (See Figure 2-5 (e) and (f) and Figure 2-3 (c) and (d).)

2.4.2 Field of view

Transform domain aliasing limits the spatial extent of an image unambiguously contained in a discrete Fourier representation of a certain bandwidth and number of samples [23]. This unambiguous field of view in a given direction is usually predicted by the period of the lowest spatial frequency probed in that direction. The implication for this technique as an image-scanner is a dependence of field of view on the number of beams, n, and their corresponding k-vectors.

2.4.3 Abberations

A fringe pattern created at the intersection of two collimated beams is sensitive to wave-front non-uniformities in one beam that are not balanced by complementary non-uniformities in the other. A fringe pattern, then, can be distorted by imperfections in the illumination optics. Such distortion would constitute a departure from an ideal Fourier basis vector in this technique. On the other hand, the distorted fringe pattern can carry information in a non-Fourier basis.

A photographic process can capture a set of distorted fringe patterns from beam-

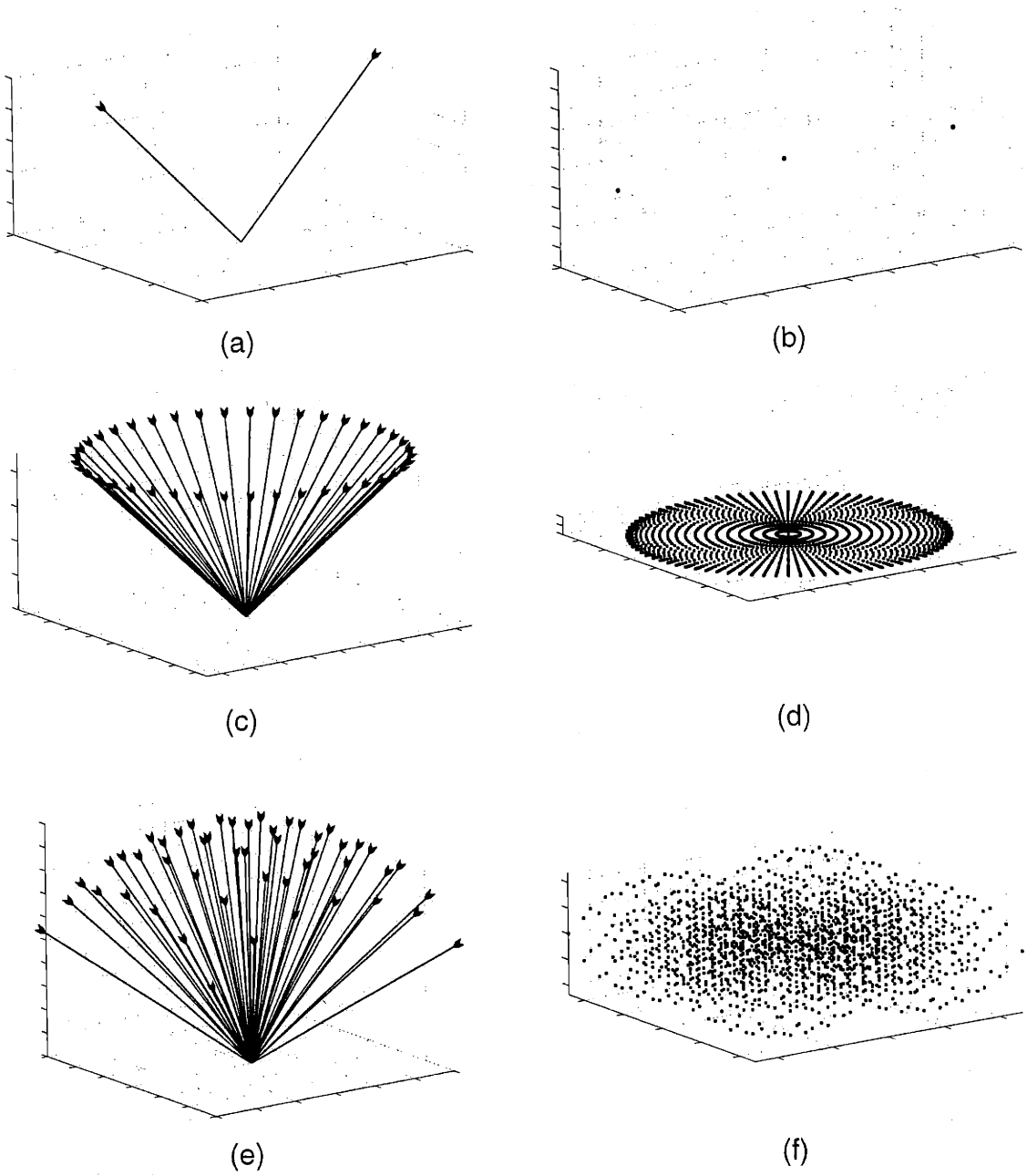

Figure 2-5: Parts (a), (c), and (e) show three arrangements of source k-vectors. Parts (b), (d), and (f) show the corresponding sets of points in the target's Fourier space that will be probed.

pairs of a given system. These data can be employed in the reconstruction of images produced with the same (imperfect) optical system. And so, a savvy image reconstruction technique can undo some abberations using computation.[10]

2.4.4 Diverging beams

Our tolerance of deviations from perfect sinusoidal fringe patterns constitutes a certain amount of flexibility with the optics. One way to exploit that flexibility is to broaden the volume of beam overlap by using diverging source beams—thus decoupling the size of the target from the size of any lens or mirror in the setup. Diverging beams interfere in pairs to form fringe patterns that can appear systematically distorted on one size scale but indistinguishable from purely sinusoidal fringe patterns on a smaller size scale.

2.4.5 Imaging lens

Piecing the last four subsections together, we have the means to fill a large region with a fine, predictable pattern but a limitation on the size of the largest target probed by such a pattern resulting in a correct image.

Here, we integrate yet more parallelism into the imaging concept. The idea is to use a relatively low-resolution imaging lens to segment the target into a number of sub-regions. Each of these sub-regions should be thought of as a separate "target" in terms of the above descriptions. Detectors can be placed at different point-spread-function-sized spots in the back focal-plane, each corresponding to an independent reconstructed image. If those detectors occupy adjacent regions of the focal-plane, an algorithm for stitching the images together may be employed to eliminate any seam.[11]

[10] In the case of imaging at the focal plane of a lens, information needed to perform this sort of correction is not present in the image. And so, imperfections in the optics of conventional imagers lead to abberations that irreversibly degrade resolution—even with a completely characterized lens system. (These abberations are to be distinguished from distortions and other effects that can be undone with computation.)

[11] The blur associated with a typical point-spread function ensures that abutting detectors see some overlapping information, even though the detectors don't overlap themselves. This enables a powerful class of stitching algorithms.

The resolution of the lens needs to be no better than the unambiguous field of view of the multi-beam sensing process (see Subsection 2.4.2). From the point of view of setting up an instrument, one would choose the best lens afforded by cost, working distance, depth of field, and image-circle requirements and provide resolution enhancement with a number and arrangement of beams sufficient to subdivide each resolution-cell of that lens to satisfy the system's resolution requirement.

Significance

Picking the angles of laser beams (or locations of point sources), setting the size of the target, and possibly selecting an imaging lens constitute the first steps of engineering a multi-beam image scanner. This section described how to relate those choices to image content.

2.5 Sources of error

Likely and fundamental sources of error in measuring the target contrast fall into three categories: optical, electrical, and numerical.

Use of a greater number of distinct illumination patterns can mitigate any of these errors when the setup is forced to allow them. Specifically, Equation 2.15 has the automatic behavior of "averaging" away non-systematic errors.

2.5.1 Optical response

The nature of light provides several mechanisms for deviations from Equation 2.10. These deviations, if not modeled or accounted for when interpreting the T signal, constitute a noise source that will corrupt our estimate of the target contrast function.

Stray light

No amount of baffling and anti-reflection coatings can completely eliminate stray light from some optical systems. If (at least) the stray light is unchanging during

the measurement, then, regardless of its structure, it will influence only our estimate of the DC Fourier component of the target contrast and can be subtracted trivially from that single estimate.

Laser intensity fluctuations

Any laser's output carries intensity fluctuations characterized by a finite noise spectral density. Depending on the relationship between this density function and the timescale of a measurement, data normalization using a monitor of the laser output power might be helpful.

Speckle noise

Coherent light remains coherent after it is reflected or transmitted by a typical object.[12] Rays of coherent light from different scatterers in the target interfere to form a pattern of dark regions and bright regions called speckles. Intuitively, speckle noise arises from an uncertainty in the number of speckles that might land in a collector's aperture [3].[13]

The fraction of uncertainty, f, in a radiometric measurement due to speckle noise depends on the wavelength, λ, the diameter of the illuminated patch of surface being measured, d, and the numerical aperture of the collecting optic, $N.A.$:[14]

$$f = \frac{\lambda}{2\,d\,N.A.}. \tag{2.17}$$

2.5.2 Electrical response

Assuming we use an electronic detection system and computation engine, we are stuck with an array of associated problems. Most troublesome among them are these:

[12] Generally, coherence is lost in a fluorescent excitation and so there is *no* speckle noise in a fluorescence image collected by a multi-beam scanner.

[13] This "speckle noise" can actually carry valuable information about the target. No effort has been made here to model this image speckle, but in other setups, a target's entire 3D structure has been extracted from speckle information-content [13, 16, 17].

[14] A good description of speckle geometry and behavior can be found in [21].

Dark noise

Dark noise is a double curse—first using up some of the sensor's dynamic range and then reducing the remaining SNR with its associated square root of uncertainty [12]. Generally, reducing the exposure time and lowering the detector's temperature are the paths to a lower dark count, providing linear and exponential benefit, respectively.

Digitization round-off

Assuming the T signal is digitized, it is recorded as one of a set number of codes. The difference between the actual value and some representative value of the code can be a systematic error unless sampled values are uniformly distributed within the digital bins associated with the codes.

2.5.3 Numerical response

Some pattern-sequences are better at probing an arbitrary target than others. Fortunately, linear algebra provides a gauge of how well a target can be represented by the sequence (contained in the matrix A of Equation 2.14): the singular value decomposition (SVD) of A. For a brief but efficient introduction to interpreting the SVD of a matrix, see [19].

Engineering a pattern-sequence that is sufficient to represent the target well even in the presence of optical and electrical sources of noise generally comes down to having enough distinct patterns.

Significance

With an understanding of the system's basic mode of operation as well as its fundamental sources of error, we are ready to discuss the prototype microscopes of the next chapter at an engineering level.

Chapter 3

Experiment

Four Synthetic Aperture Microscopes of the type described in Chapter 2 were constructed. Each apparatus was an improvement on the previous one. Table 3.1 tells the story.

M1, the first apparatus, proved the concept of spatial heterodyning with a moving beam system. M2, the second generation, was a first stab at engineering the complex optical system for speeding up M1 with parallelism. The system was a proving ground for the theory of many beam illumination, a two-camera approach, and a suite of algorithms for analyzing reference imagery. M3 sought to correct the optical efficiency problem of the M2 approach for generating many beams. This involved considering and experimenting with different ideas for achieving the ring source geometry with non-ring, but highly efficient beam-splitting mechanisms. Successful ideas from the ill-fated M3 were inherited by M4, a solid-state, stable, fast, optically efficient, and positively controllable synthetic aperture microscope.

Table 3.2 summarizes the key features and specifications of the four machines.

3.1 M1: First-generation, moving-beam microscope

M1 was intended as a proof of concept of the spatial Fourier transform heterodyne microscope concept. It traced out a synthetic aperture over time. Figure 3-1 shows the apparatus.

M1	M2	M3	M4
Provides: Proof of physical principles with a moving beam system.			
Needs: Speed. Freedom from precision motorized stages.	Provides: Fixed beam geometry, divorcing the microscope from precision motorized stages. Many beam and many pixel parallelism concepts demonstrated.		
	Needs: Better light efficiency. More light-efficient method for generating many beams with easier alignment to an array of modulators.	Provides: Efficient beam splitting and alignment technology for an array of modulators. Array of beams generated with digital optics automatically aligned to a MEMS modulator array.	
		Needs: Better modulators with greater power-handling and surface quality than available MEMS.	Provides: Auto-aligned "modulators" of good quality and efficiency. Acousto-optic device used both as a highly-efficient beam-splitter and also as a modulator of each beam.

Table 3.1: The story of the four generations of apparatus.

	M1	M2	M3	M4
Number of beams	2	41	15	15
Beam-splitter	Beam-splitting cube	Hologram	Digital Optic	Acousto-optic modulator
Beam phase modulator	Translating wedge	Piezo-driven mirrors	Electrostatically driven MEMS mirror array	
Beam amplitude modulator	none	none	none	
Beam delivery to sample	Mirrors	Gratings, mirrors, and one spherical lens	Single-mode polarization-preserving fiber and spherical lenses	Mirrors and cylindrical lenses
Spatial filtering	Pinhole-type	Custom-etched copper slit array	Fiber optic as pinhole	none
Phase reference	Photodiode with slit in magnified path	0.5 NA microscope objective + CCD in magnified path	Far-field CCD	1.25 NA microscope objective + CCD in unmagnified path
Laser type	He-Ne	Argon ion	Argon ion	Diode
Wavelength Spectral purity	633 nm (red) multiple cavity modes	488 nm (blue) single-line and single-frequency	514 nm (green) single-line and single-frequency	685 nm (red) Single longitudinal mode

Table 3.2: Key specifications of the four prototypes.

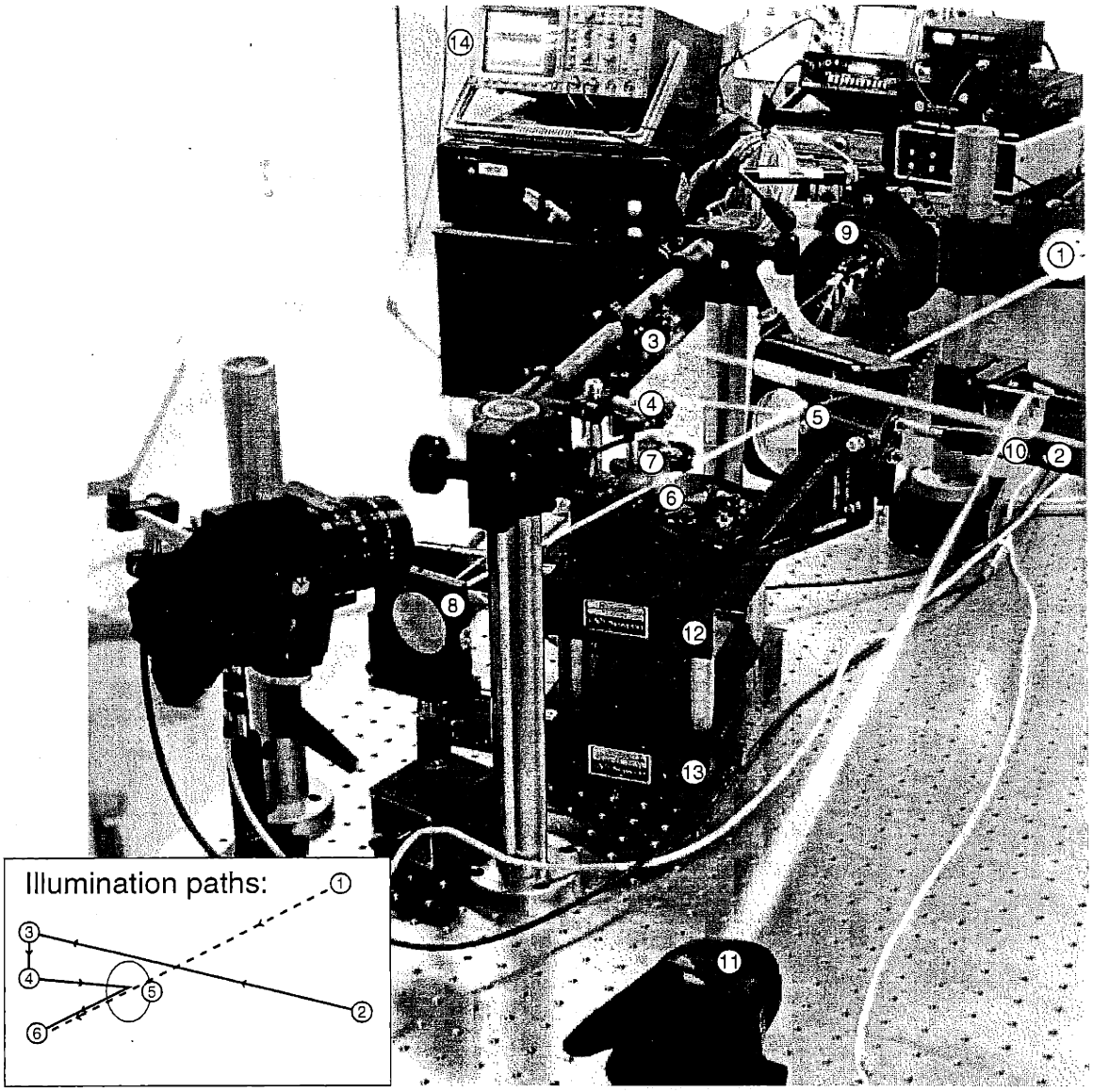

Figure 3-1: A photograph of M1, the first microscope prototype. The labels indicate (1) a mirror delivering a stationary beam to the target area; (2) the second beam; (3) the top, stationary mirror of a periscope for the second beam; (4) the bottom periscope mirror that rotates on the top rotation stage; (5) a beam-splitter that also rotates on the top rotation stage; (6) the target area; (7) the target scattered light detector; (8) a fold mirror, directing transmitted light to (9) the target transmitted light detector; (10) an off-axis parabolic mirror receiving light from both beams via the beam-splitter; (11) a reference detector comprising a slit and a photodiode under a felt shroud; (12) the top rotation stage, that rotates the moving beam; (13) the bottom rotation stage, that rotates the target at half the rate of the top rotation stage; and (14) a digitizing oscilloscope for capturing target and reference sinusoids.

3.1.1 Apparatus overview: three motors, two beams, one oscilloscope

Figure 3-2 is the most informative starting point for understanding M1.

The apparatus projected traveling sinusoidal intensity fringe patterns on a target and collected the back-scattered light with a large-area photodiode. Each sinusoidal intensity fringe pattern sampled a corresponding point in the 2D Fourier transform space of the target. The photodiode current oscillated sinusoidally as a function of time with an amplitude and phase corresponding to the complex value of the probed 2D Fourier transform component of the target.

The fringe pattern was formed by the interference of two beams from a 633 nm (red) helium neon laser. The two beams, generated by splitting the output of a single laser with a non-polarizing beam-splitting cube, were directed into a region of overlap where the target was placed. One beam was frequency-shifted relative to the other with a wedge-shaped piece of glass actuated by a motorized linear translation stage. This frequency shift caused the fringe interference pattern to "roll" perpendicular to the fringes as a function of time. The pitch and angular orientation (relative to the sample) of the fringe pattern were dictated by the angle of the beams relative to the sample, set by motorized rotation stages. These angular positions were computer-controlled to step through a sequence of Fourier transform sample locations.

Measurement time proportional to the number of resolution elements

Generally, the number of resolution elements in an inverse Fourier transform corresponds to the number resolution elements in the original Fourier transform data.[1] Because of the mechanical positioning time associated with each Fourier transform sample point, the number of resolution elements was kept low. Runs of M1 were limited to 1D measurements to this end.

[1] The correspondence is one-to-one when samples are spaced on a regular grid.

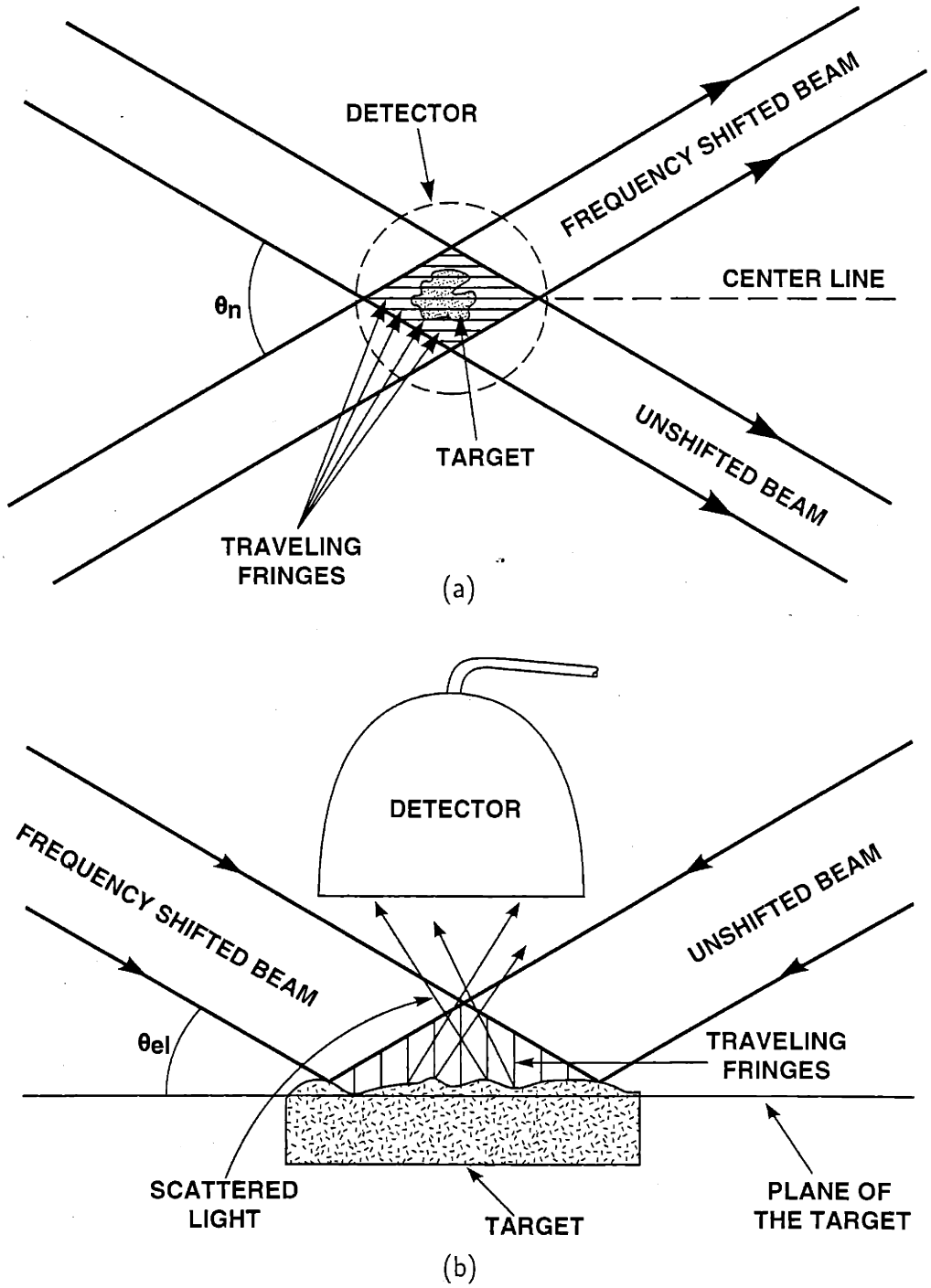

Figure 3-2: Two laser beams of slightly differing frequencies form a traveling sinusoidal fringe pattern that interacts with a target. In response, the target transmits, reflects, or fluoresces a time-varying intensity signal which carries a complex spatial Fourier transform coefficient to the detector. Part (a) shows an overhead schematic of the concept. Part (b) shows a side view for reflection-mode imaging.

Reference detector gives meaning to the phase measurement

Because of mechanical uncertainties, the spatial phase of the fringe pattern had to be measured by a "reference detector." A magnified copy of the fringe pattern projected on the target was produced with a beam splitter and a parabolic mirror. A slit, fabricated by cementing two razor blades to a frame under a dissection microscope, was positioned parallel to the magnified fringe pattern. The slit was much thinner than the finest fringe pattern's magnified pitch, and therefore a second photodiode behind the slit received an oscillating intensity as the pattern rolled by the slit.

A digitizing oscilloscope collected and buffered the target- and reference-detector signals simultaneously for eventual computer analysis.

3.1.2 Detailed description of the apparatus

A 30 mW (rated) 633 nm (red) Helium-Neon multiple-cavity-mode laser (Spectra-Physics) was split with a non-polarizing beam-splitting cube and directed through different paths at a target.

Path lengths matched to within 10 centimeters

Five prominent cavity modes spaced by approximately 480 MHz were observed with a (Burleigh) super-invar Fabry-Perot interferometer fitted with $\lambda/200$ multi-layer dielectric mirrors. The response was measured with a (Motorola MRD500 PIN) photodiode, a sensitive transimpedance amplifier, an analog Tektronix oscilloscope, a Polaroid camera, and a ruler. The Polaroid is shown in Figure 3-3. The measurement suggested the path lengths in the system should be matched to within about 10 cm, the useful coherence length of the primary coherence region of the source.

Phase shifted with a translating glass wedge

After abandoning a 40 MHz frequency-shifting Bragg cell because of bandwidth limitations of convenient detectors, a method was devised to frequency-shift one of the two beams approximately 7.5 Hz relative to the other. Figure 3-4 shows the arrange-

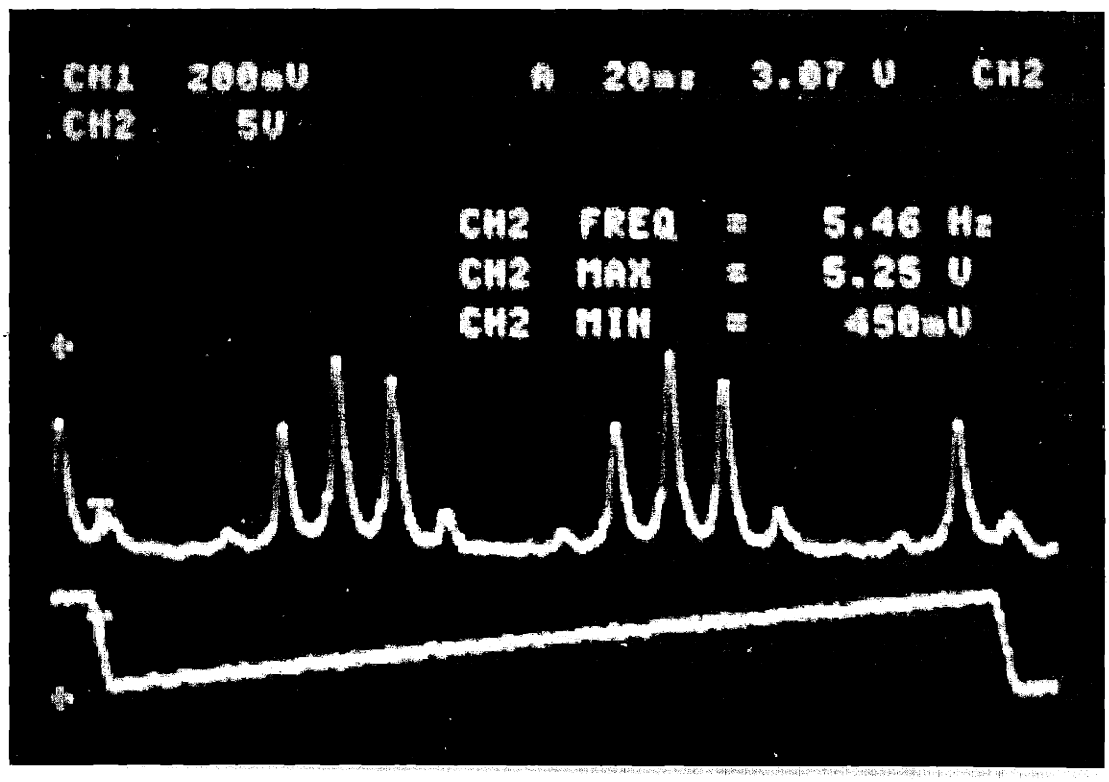

Figure 3-3: M1 laser coherence measurement. Polaroid snapshot of central intensity (top trace) as a function of a ramping mirror phase command signal (bottom trace) with a ten centimeter Fabry-Perot cavity. The primary coherence length was deduced to be 15.5 cm with a useful sub-range of about 10 cm. For an efficient treatment of Fabry-Perot measurements, see [6, pp. 111-115].

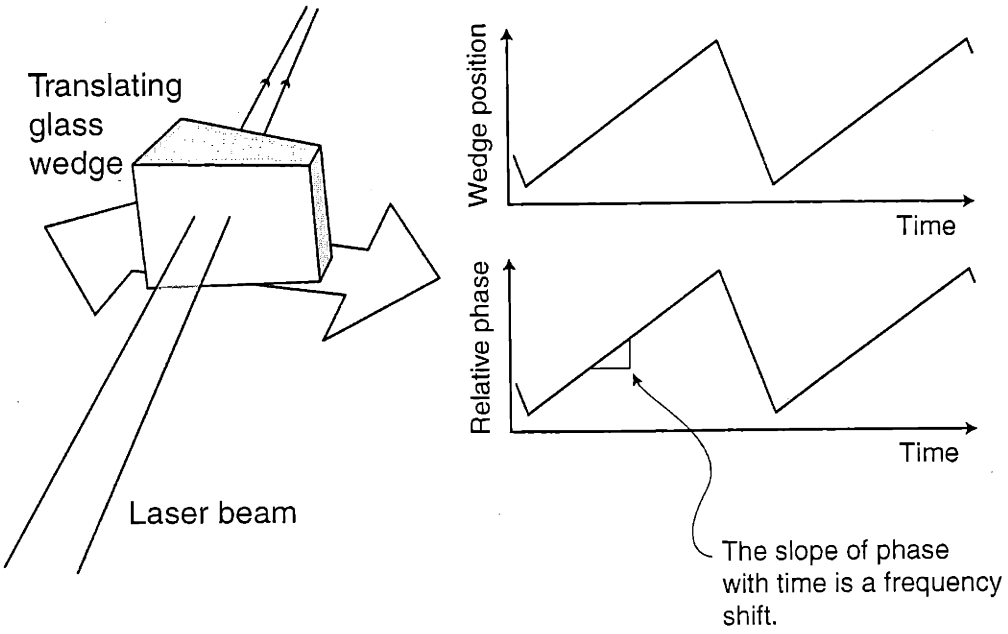

Figure 3-4: A glass wedge on a constant-velocity translation stage imparts a linearly ramping phase to one beam equivalent to a relative frequency-shift.

ment. Translating a mildly-sloped glass wedge 5 mm retards the transmitted beam about 10 waves. A (Newport PM500) motorized micrometer actuated the wedge with a linear velocity for the one second data acquisition window ($\approx$7.5 waves) and then reset to the original position. Stage accelerations and decelerations were programmed for smooth operation to avoid excessive vibration of the optical table.

Beam geometry controlled with precision rotation stages

Instead of controlling the directions of two beams and leaving the target stationary, only one beam was dynamically positioned, and the target was rotated. Mirrors controlling the moving beam and the target were mounted on (Newport PM500) precision rotation stages under computer control. Figure 3-5 shows the arrangement. For the 1D measurements, the target was rotated one angular increment for every two angular increments of the moving beam. Thus, the fringe pattern maintained its angular relationship with features on the sample while changing pitch alone.

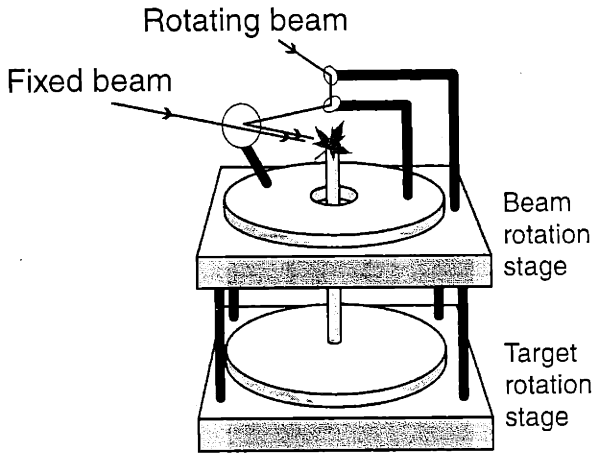

Figure 3-5: It was more convenient to rotate the target and one beam than to rotate both beams over a stationary target. Because of the resolutions of interest for this proof of concept, the required mutual angles between beams were very small. Therefore, the moving beam was directed at the target after reflection from a beam splitter through which the fixed beam passed. This allowed the mutual angle to be brought all the way down to zero.

Large-area detector output analyzed by computer

Targets were simple, high-contrast patterns generated with a laser printer on laser printer paper for back-scatter measurements, or on laser printer transparencies for transmission measurements. Some of the light scattered back or transmitted by the simple targets was collected by a large-area (1 cm^2, round collection aperture) PIN photodiode approximately 2 cm from the target. The collection numerical aperture, therefore, was about 0.3. Amplified by a transimpedance amplifier and digitized by a (Tektronix TDS 460) oscilloscope, the signal was analyzed by computer to extract the amplitude and phase of the target's Fourier transform coefficient at the addressed Fourier sample.

3.1.3 1D measurements

Figure 3-6 shows an example trace of the reference and target detector outputs during a one second measurement of a single Fourier transform coefficient of a target. The amplitudes and relative phases of hundreds of such pairs of traces, collected with

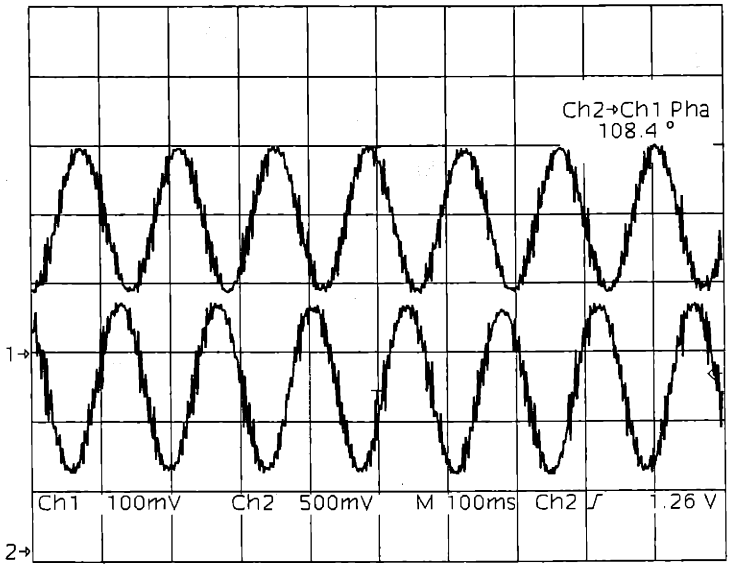

Figure 3-6: M1 sinusoids. The upper trace shows the response of the phase reference detector, and lower trace shows the response of the detector positioned to receive light reflected from the sample.

different angles between the beams for a given target, were estimated by curve-fitting each trace. Figure 3-7 (a) shows a plot of the amplitudes as a function of the angular difference between the pair of beams with a fine grating as the target. Figure 3-7 (b) shows a similar plot with a (four times) coarser grating as the target. Both plots clearly show the expected symmetry and DC and fundamental terms of the grating's Fourier magnitude spectrum. The angular differences between DC and first harmonic can be read off of the plots and converted to an expected period. The first is 0.102 degrees and the second 0.026 degrees corresponding to periods of 0.353 mm and 1.39 mm, respectively. The gratings were measured with a Leitz measuring tool-maker's microscope indeed to have these periods within the tolerance of the measuring microscope. The excellent agreement not only between the appearance of the plots and expectation, but also between the quantitative features of the data and an independent measurement prove the microscope concept.

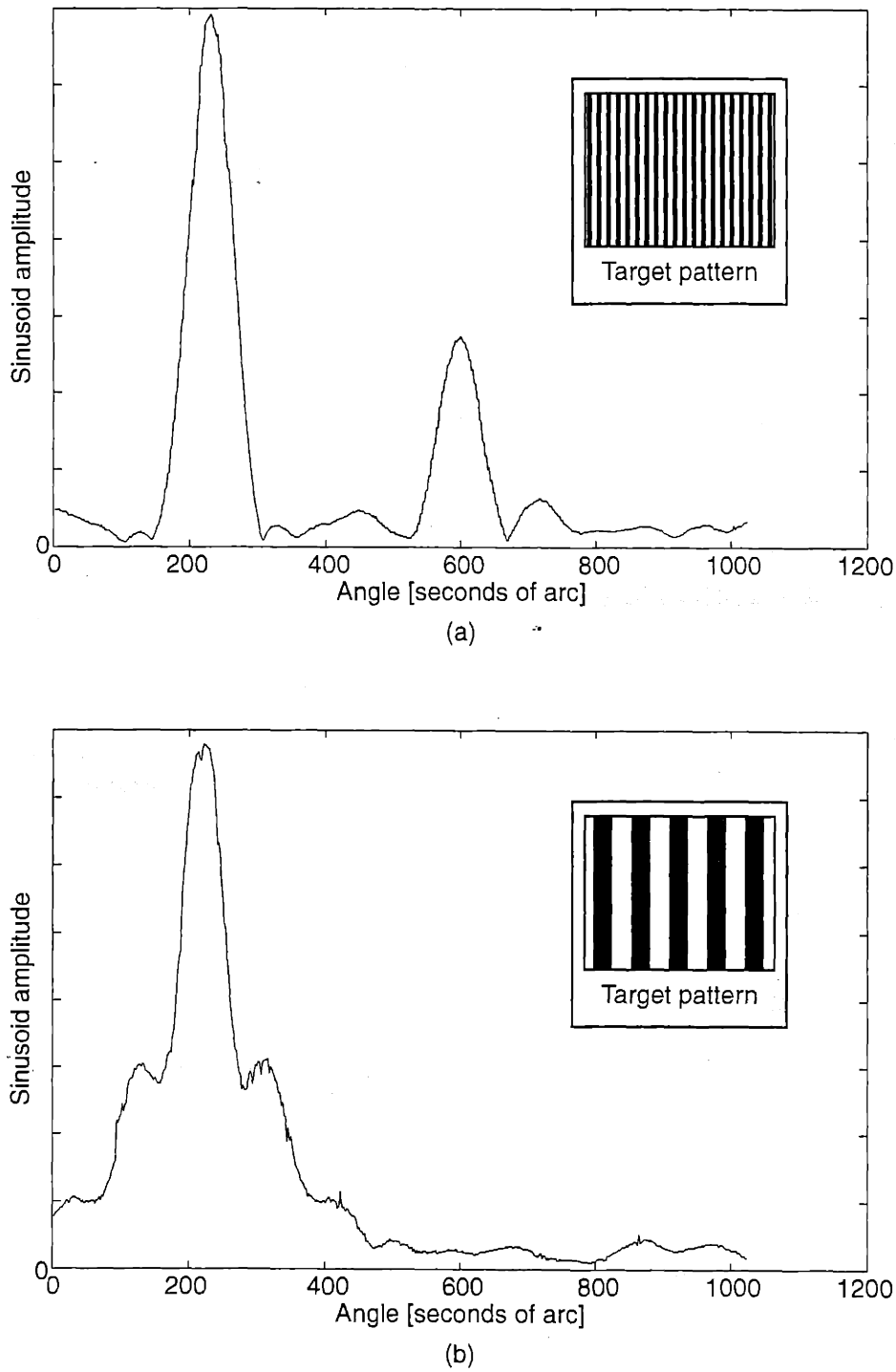

Figure 3-7: Fourier magnitude spectra of two samples measured by M1. Plot (a) shows the magnitude of 512 sinusoids like the lower trace of Figure 3-6 for a grating target, as shown in the inset. The beams pass through zero angular difference (and the DC Fourier component of the target) at the high point of the curve. Plot (b) shows the same type of un-filtered data for a coarser grating, as shown in its inset.

3.1.4 Conclusion: A useful demonstration of principles, but too slow for general microscopy

The major feature of M1 was that it demonstrated the physics of synthetic aperture microscopy. The major drawback was that even 1D measurements took hours, owing to the serial, mechanical method of addressing Fourier transform samples.

3.2 M2: The first synthetic aperture microscope with fixed beam geometry

M2 tested novel concepts for a fixed beam geometry synthetic aperture. These concepts fundamentally overcame the speed problem of M1. Most significant among them are

- the discrete-source ring geometry for providing a good sampling of the 2D spatial Fourier transform of a target scene.
- spread-spectrum communication of spatial Fourier transform coefficients from an array of sample regions to an array of detector elements.
- reference image measurement and associated amplitude and phase estimation algorithms.

Figure 3-8 shows the apparatus.

3.2.1 Speedup from parallelism

M2 embodied a twofold approach to speeding up measurement of complex Fourier transform coefficients of a target. Both tactics are based on parallelism and constitute a solution to the major practical problems of previous state-of-the-art synthetic aperture microscopes.

1. First, **many beams**, forty-one (41) to be exact, were used in parallel instead of just two. Since the number of beam-*pairs* determines the number of simple

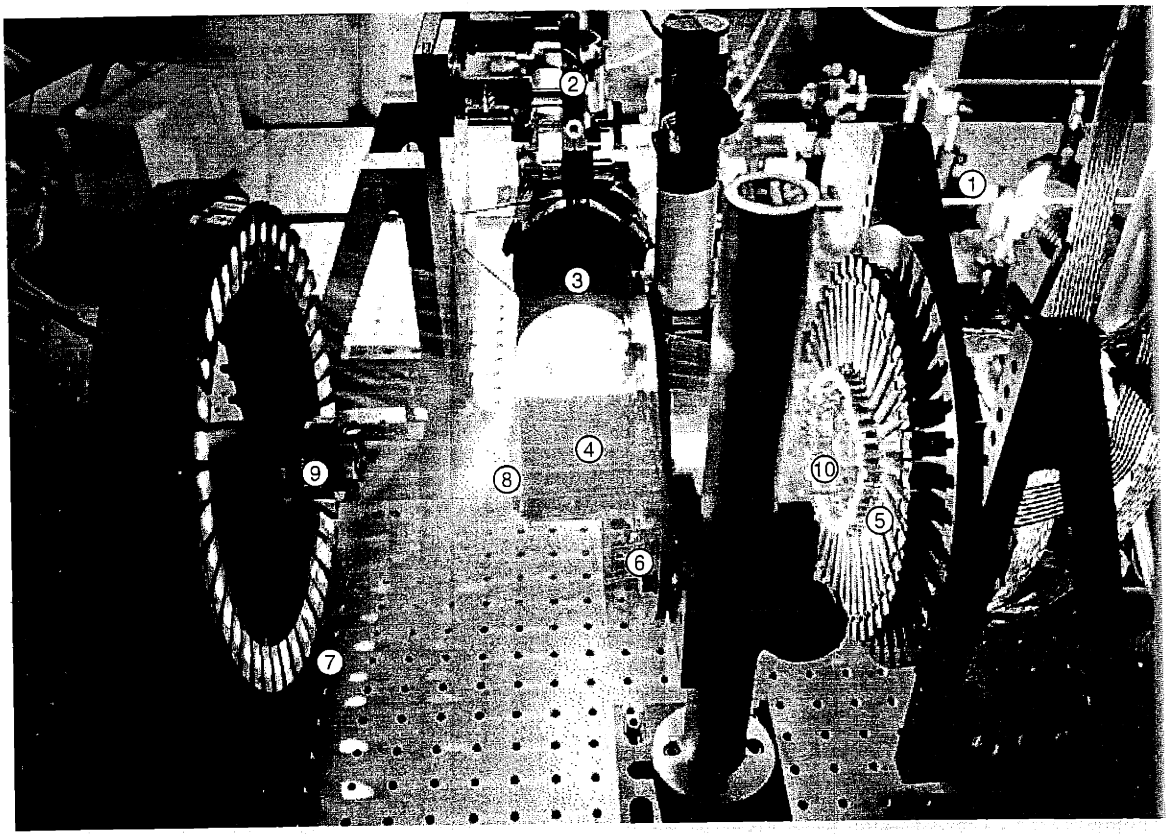

Figure 3-8: M2, the second microscope prototype. The labels indicate (1) the source beam from the argon ion laser (laser head not shown) which passes through a spatial filter before entering (2) a periscope assembly which directs the expanding beam into (3) a large aperture copy lens. The now converging beam enters (4) the hologram housing which contains an elliptical fold mirror and the beam-splitting hologram. A diverging cone of 41 converging beams emerges from the right side. The beams strike (5) the ring of path-length modulating mirrors, one beam per mirror. Each mirror is glued to a piezo which is glued to a steel drill-rod post. Each post is fastened to a miniature tip-tilt mount to facilitate alignment with (6) the "slit array," an array of matched spatial filters at the unfolded focus point of the copy lens. The now diverging beams strike gratings in (7) the grating wheel from which they diffract. Second-order diffracted beams are the strongest (because of the blaze angle of the gratings), and a converging cone of diverging second-order beams land on (8) the target. Light scattered or fluoresced by the target is collected by (9) the target camera macro lens. First-order diffracted beams from the grating wheel pass through angled cuts in the slit-wheel mounting plate and interfere in the focal plane of (10) the reference camera microscope objective lens.

fringe patterns simultaneously projected, the number of complex Fourier transform sample points varies as the square of the number of beams. So, instead of two pieces of information (an amplitude and a phase) as in M1, 1,641 independent pieces of information were probed by the expanded arrangement of beams.

2. Second, in conjunction with a relatively low-resolution imaging lens, **an array of detectors**—the 256×256 pixels on a CCD detector—measured the response of as many regions on the target in parallel.

Thus, 1,641 numbers for each of 65,536 regions are probed by the fixed geometry system providing a degree of parallelism of 107,544,576.

3.2.2 Apparatus overview: one hologram, two cameras, forty-one of everything else

M2 projected a complex, finely structured pattern of light at a target. An imaging lens collected some of the light scattered from the target and directed it to a CCD array.

Changing probe pattern formed by modulated interference

The "probe pattern" was formed by interfering forty-one (41) beams split from a single argon ion laser beam by a holographic optical element. The path lengths, and thus the relative phases, of the 41 beams were modulated between CCD frames by piezoelectrically actuated mirrors. The result of this phase modulation was to change the probe pattern landing on the target between each of a sequence of CCD frames.

A picture in every pixel

Each pixel in the CCD mapped to a region of the target many times larger than the characteristic size of probe pattern features. Over time, reacting with a sequence of probe patterns, each region would transmit its probed spatial Fourier transform

coefficients—shape information on the size-scale of the probe pattern features—to its CCD pixel.

Reference images help parse Fourier transform coefficients from target detector signals

The transmission of Fourier coefficients was encrypted by the phase modulation sequence used. Because of uncertainties associated with piezoelectrically controlling phase, the phase modulation sequence had to be measured. A magnified copy of a region of the probe pattern was directed at a second CCD camera. Unlike the target detector, this "reference detector" resolved the pattern. A computer algorithm was used to extract the phase modulation sequence from a sequence of frames from the reference detector. Because target detector frames were collected simultaneously with the frames from the reference detector, this measured phase modulation sequence could be used to decrypt the spatial Fourier transform coefficients of regions of the sample.

3.2.3 Detailed description of the apparatus

Figure 3-9 shows the apparatus schematically.

A Coherent Innova 304 argon ion laser supplied one (1.00) Watt of 488 nm (blue) light to M2. The laser was fitted with an inter-cavity prism as well as an inter-cavity etalon and adjusted to operate in single-line, single-frequency mode. In this mode, the useful coherence length was many meters. The current to the laser tube dynamically adjusted for the conditions of the laser to maintain a substantially constant output optical power.

A spatial filter (a 25 μm pinhole, mounted in Newport model 910), a periscope (Newport model 670), and a large-aperture copy lens (1:1 ILEX 210 mm f4.5, Bruning P/N 308668) expanded the beam and directed it into the center of the apparatus.

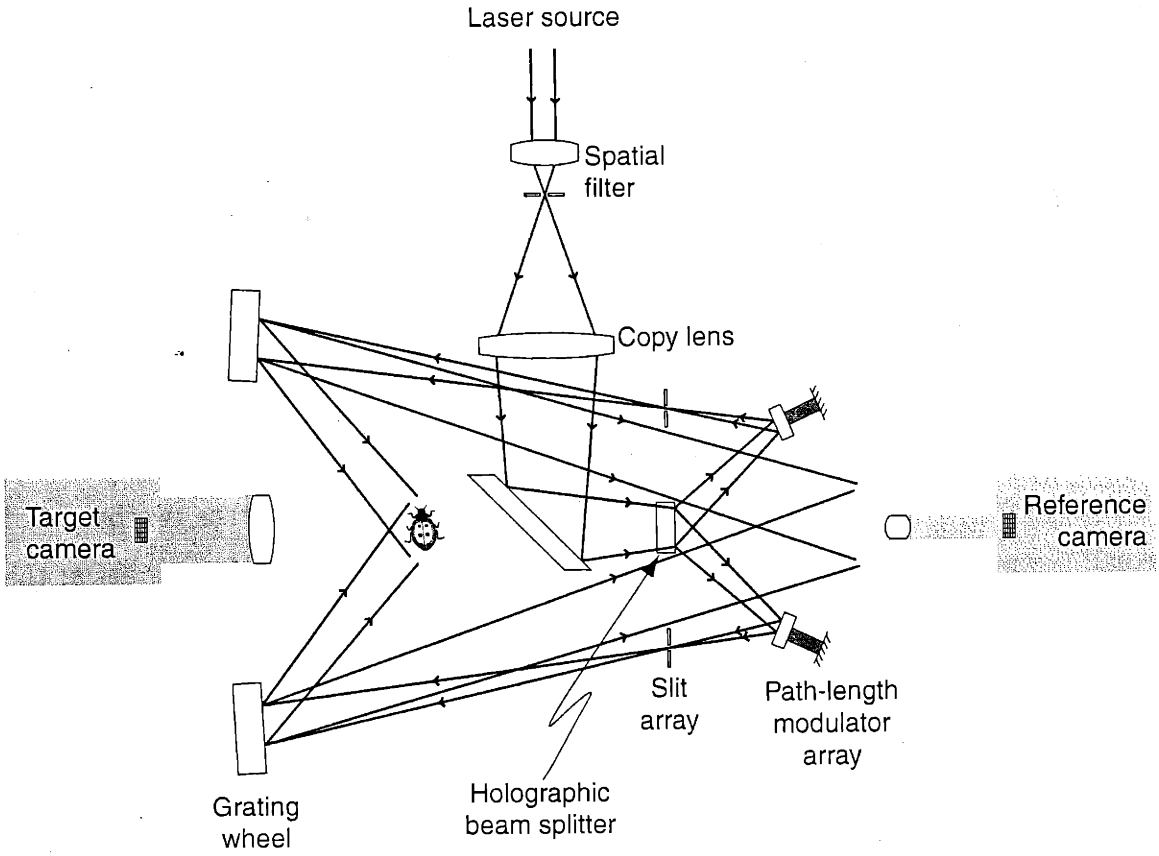

Figure 3-9: The optical path of M2 splits twice: once from the holographic beam splitter (into 41 used beams) and again from the grating wheel (into 82 used beams). The axially symmetric design was folded to provide a compact structure and to allow for high-efficiency reflective blazed gratings.

Custom digital optics too expensive as beam splitter

To split the single laser beam into a cone of forty-one beams, several approaches were considered, including the fabrication of a custom digital optic. The fine pitch required to turn our relatively short wavelength an appreciable angle combined with our requirement for a large number of beams relative to the 1997 state-of-the-art would have made digital optics an expensive choice. It was more cost-effective to fabricate a conventional holographic optical element and compensate for the low optical efficiency with a more powerful laser.

Holographic beam splitter creates the right number and arrangement of beams

Figure 3-10 shows the setup used to fabricate the custom holographic optical element for M2.[2] The 532 nm (green) beam of a YAG laser (Coherent DPSS532) was expanded to a two inch diameter with a 65x microscope objective and a 127 mm plano convex lens. The expanded beam was split with a dielectric beam splitter into two beams of equal power. A mirror re-joined the beams with a mutual angle of 22.7 degrees at the surface of a silver-halide-based photo-sensitive holographic plate (Illford SPG72T) mounted on a sheet of glass affixed to a rotation stage. The holographic plate was stuck to the sheet of glass using block printing ink ("Speedball Black Water Soluble Ink for Block Printing," purchased at an art supply store) applied in an even coat with a heavy roller. The beams encountered the emulsion side (the photo sensitive side) of the plate first, and the ink served jointly as a temporary glue and as an anti-reflection coating for the back side of the plate. The beams, coherent for 10 km (rated), interfered in the volume of the emulsion to expose one component grating of the hologram.

A series of forty-one flashes, controlled by a shutter (Newport model 845), were exposed on the holographic plate with a $2\pi/41$ rotation between each flash. Neutral

[2]The setup was constructed in the MIT Media Laboratory's holography facility with assistance from Michael Klug and Stephen Benton.

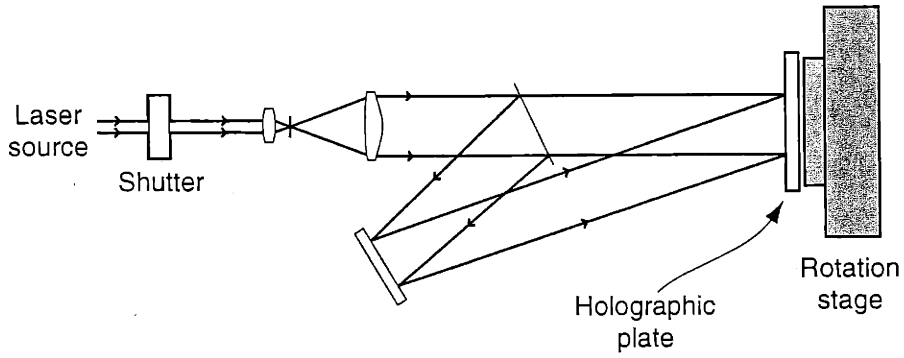

Figure 3-10: The optical setup for fabricating the holographic beam splitter for M2. A shutter exposed a holographic plate with flashes of varying duration for each of 41 angular orientations of the plate.

density filters were used to bring the optical power into the range of the shutter's best accuracy given the rated film sensitivity of 60 μJ/cm^2.

When all exposures were of equal duration, the holographic efficiency of the forty-one component gratings declined monotonically. That is, the early exposures "used up" the film sensitivity. One solution would have been to give tiny exposures to the component gratings in a sequence that addressed all the components and then repeated many times to provide the full exposure dose. However, this solution required that the rotation stage address the same angle more than once to an accuracy well beyond that of the available stage. Instead, a nonlinear model of the film's hysteresis was estimated from test hologram efficiencies measured with a power meter. After several iterations, the computer model predicted a sequence of forty-one increasing exposures commanded to the shutter for the fabrication of a holographic optical element providing forty-one beams matched in power.

After exposure, the holographic plate was treated with standard chemical processing followed by a bleaching step—changing the function of the plate from *amplitude* spatial light modulation to *phase* spatial light modulation.

Measured at 488 nm, the holographic optical element delivered more than three percent of the input beam to the forty-one transmitted output beams. The power of these beams was measured to match within ten percent. Negative orders and higher

orders were well suppressed (by the volume-nature of the hologram), but, as expected with a silver-halide holographic plate, scatter was noticeable. The zero transmitted order contained almost 42 percent of the energy. The cone angle defined by the output beams was approximately 40 degrees.

Piezo stacks push mirrors, control phase

The 41 beams were received by as many 6×6 mm first surface mirrors glued (with Devcon five-minute epoxy: Devcon Corporation, Danvers, Massachusetts) to corresponding piezoelectric stacks (AE0203D08: NEC Corporation, Japan). With a multi-channel analog controller (UVA 128 Channel Controller: Digital Designs and Systems, Arlington, Massachusetts), the piezo voltages were computer-commanded in the range of 0-40 V. The resulting length change in the reflected beam's path corresponded to a phase shift relative to its previous phase or to the other beams.

A measurement of the frequency response and hysteresis effects of this phase modulator was performed with a laser Doppler vibrometer (OFV3001 vibrometer controller with OFV511 fiber interferometer: Polytec Corporation, Germany). Figure 3-11 shows characteristic data. With the final mounting, mirror, and drive signals, the modulator was found to have a settling time in the range of less than 200 μs to within 30 nm. The acceptable ringing tolerance was determined by the impact of a given phase uncertainty on the probe pattern and the radiometric sensitivity of the cameras (see below) to that impact. Drift and hysteresis were noticeable, but less so with a DC bias to the piezos.

Slit array as spatial filter and alignment mask

The 41 phase-modulated beams came to sagittally astigmatic foci after the copy lens and holographic optical element. A ring of matched spatial filters—slits of the right dimensions instead of the usual pinholes—was etched from a 381 μm thick ETP half-hard copper sheet. The slit array was 143 mm in diameter; the copper sheet was just larger. After etching, the sheet was black-oxidized. Tip-tilt adjustment of the phase-modulators placed each beam's focus in its corresponding slit.

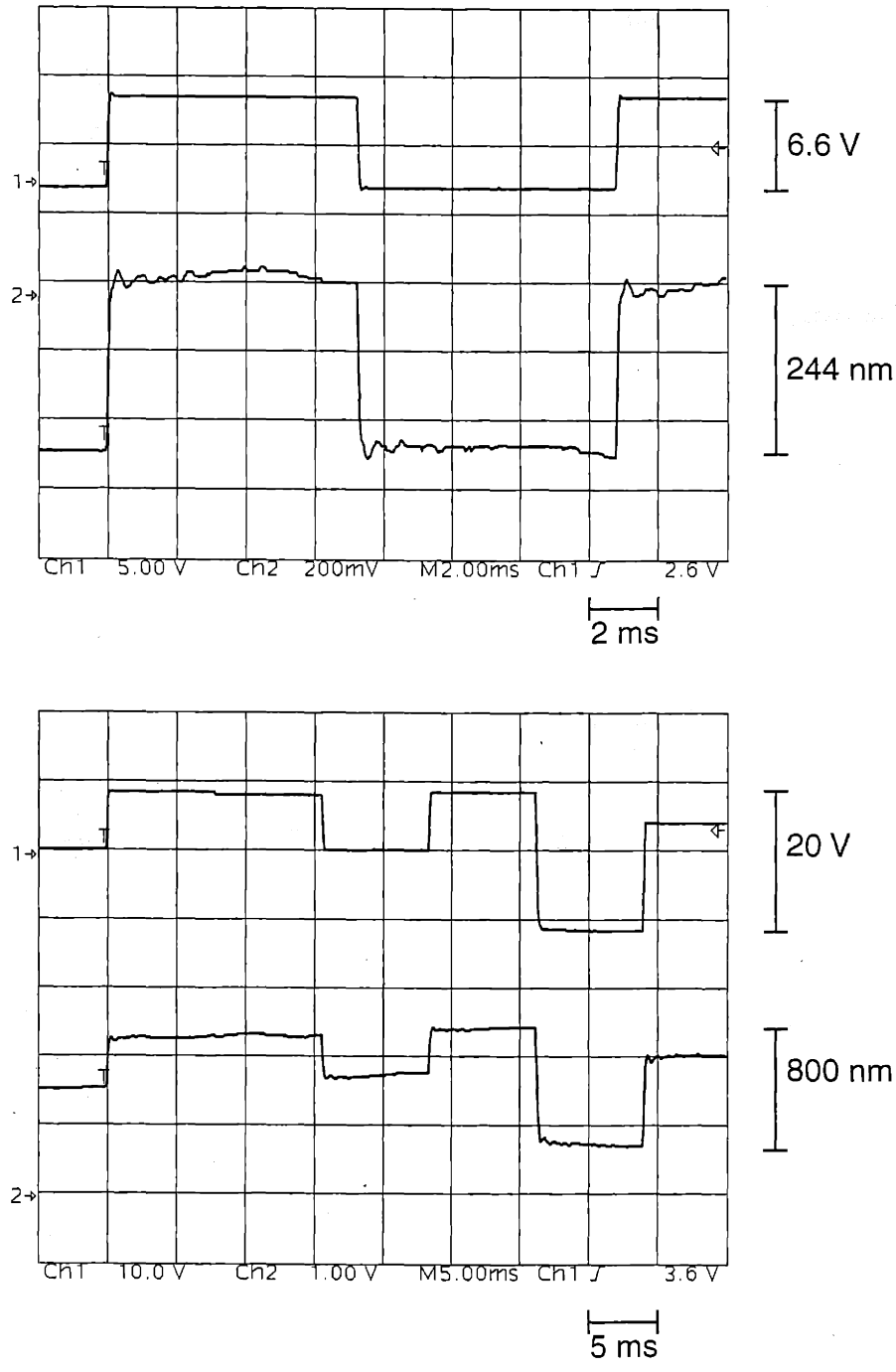

Figure 3-11: Measured responses of a single M2 phase modulator. In each plot, the top trace is the command voltage, and a bottom trace is a measure of position. The command voltage rode on a DC offset of +11.1 V, not shown, intended to bias the piezo away from a hysteretic region of operation. The position signal is nominally 0.5 μm/V. The ringing in the top measurement falls within a 30 nm worst-case envelope after no more than 200 μs of settling time.

The slit array served to spatially filter away the scattered light background from the hologram and also to provide accurate and sure alignment reference points for the complicated optical system.

M2 front element: a ring of gratings

After passing through the slits, the beams landed on a 241 mm diameter ring of 41 diffraction gratings. This ring, the "front element" of M2, defined the physical synthetic aperture of the microscope. The aperture is on the order of 100 times larger than that of a conventional microscope lens of the same effective N.A., allowing for a far greater working distance than possible conventionally.

The gratings (12.5×25 mm 600 mm^{-1} 17 degree blazed gratings, catalog number p43,206: Edmund Scientific, Barrington, New Jersey) were chosen to send most of the incident light to their second order—the location of the target—and some light to their first order—the location of the reference detector. Better than 25 percent efficiency in sending incident power to these two orders was observed.

Second-order beams from gratings illuminate target

The cone of second-order beams overlapped at the location of a target. Their interference formed a complex, finely structured probe pattern of light. With a cone-angle of 76.8 degrees, the effective numerical aperture comes out to 0.62 with a working distance of 152.4 mm. The cutoff spatial period, $\lambda/2$ N.A., was 394 nm/cycle.

Camera images target response

Light fluoresced or scattered from the target was collected by a 1:2 macro lens (a reverse-mounted Apo-Rodagon D 2x 75 mm f4.5 enlarger lens, catalog number 273 0075 002: Rodenstock, Germany) and focussed on a CCD camera imager (200 fps with 256×256 16×16 μm pixels and unity fill-factor, model number CAD1-0256A: Dalsa Inc., Ontario, Canada). The lens was used at its optimized magnification with no tilts, so each CCD camera pixel received energy from an 8×8 μm region of the sample.

The target camera viewed the target from the center of the grating wheel.

First-order beams from gratings provide a reference pattern

The cone of first-order beams interfered to form a reference pattern. Because of the unchanging phase relationship between the first and second-order beams, this reference pattern had an unchanging relationship to the probe pattern at the target. Because the first-order cone angle was more shallow than the second-order cone angle, the reference pattern was magnified relative to the probe pattern.

A 50x microscope objective (LD Epiplan 50x 0.50 N.A. $\infty/0$, part number 44 28 50: Zeiss, Germany) imaged this reference pattern onto a second CCD camera, identical to the first.

Cameras cooled, dark noise reduced

The cameras were augmented with a thermoelectric/water refrigeration system that used the surplus capacity of the argon laser heat exchanger and four matched thermoelectric cooler elements (catalog number CP0.8-31-06L-1: MELCOR Corporation, Trenton, NJ) per camera. Dark noise of the sensors was significantly reduced by the chilling system, which lowered the free-running temperature of the high-speed CCD's from approximately 30 degrees to -2 degrees Celsius. No condensation on the detectors was observed.

Cameras synchronized to maximize value of reference images

For the reference pattern to predict the probe pattern as well as possible, it was measured substantially at the same time as the target response. To this end, the two cameras were controlled by the same, external master clock and control signals and thus were synchronized, pixel-for-pixel. The two 8-bit outputs were received simultaneously, as if from a single 16-bit camera, by a digital frame grabber (Road Runner model 44: BitFlow, Inc., Woburn, Massachusetts).

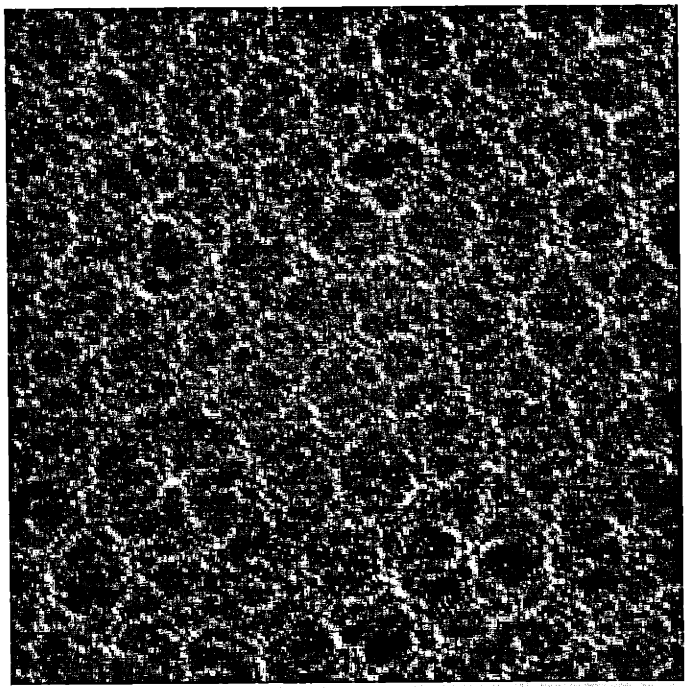

Figure 3-12: Example reference data as captured by the reference camera.

Turbulence tent improves stability

At first, unwanted phase fluctuations between the target and reference paths caused by room turbulence plagued the system. Therefore, the apparatus was enclosed in a tent made of building insulation (1 inch thick Owens Corning Foamular 250) to minimize turbulence in the beam-paths.

The further addition of fans to equalize thermal gradients within the turbulence tent reduced the overall stability of observed beam phases and was subsequently undone.

3.2.4 Reference measurements

Figure 3-12 shows a typical frame from the reference camera. The image therefore also shows the type of probe pattern present at the target.

A spectral analysis of the image reveals how the data are useful. Figure 3-13 shows the 2D Fourier transform of Figure 3-12. The image contains 1,641 hot spots—two

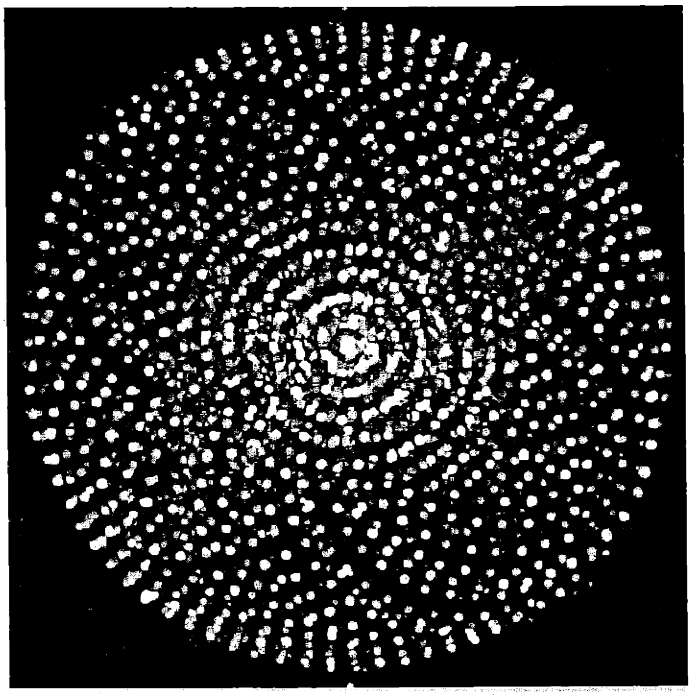

Figure 3-13: 2D power spectrum of Figure 3-12. The grey codes in the plot are the log of the Fourier magnitude of the reference image windowed by a raised cosine to eliminate edge effects in the spectrum. The image contains 1,641 hot spots.

for each beam pair (symmetric about DC) and one at DC (in the center).

With an estimate of the source locations, we should be able to tap into each of the hot spots for amplitude and phase information about the probe pattern. In M2, the process of finding the source locations was done using three sub-arrays, as shown in Figure 3-14. Finding the right bright regions in a matched Hough transform of data from each sub-array was the first step. The estimation of source locations from reference imagery, including a description of the matched Hough transform used, is explained in detail in Subsection 3.4.4.

Now, with knowledge of the source locations most consistent with the reference image, we should be able to predict the hot spot coordinates in the transform domain. Figure 3-15 shows the same spectrum as in Figure 3-13 tagged with markers. The good agreement of measured hot spot locations with the coordinates predicted from the source locations in k-space confirms the theory of many-beam illumination, derived

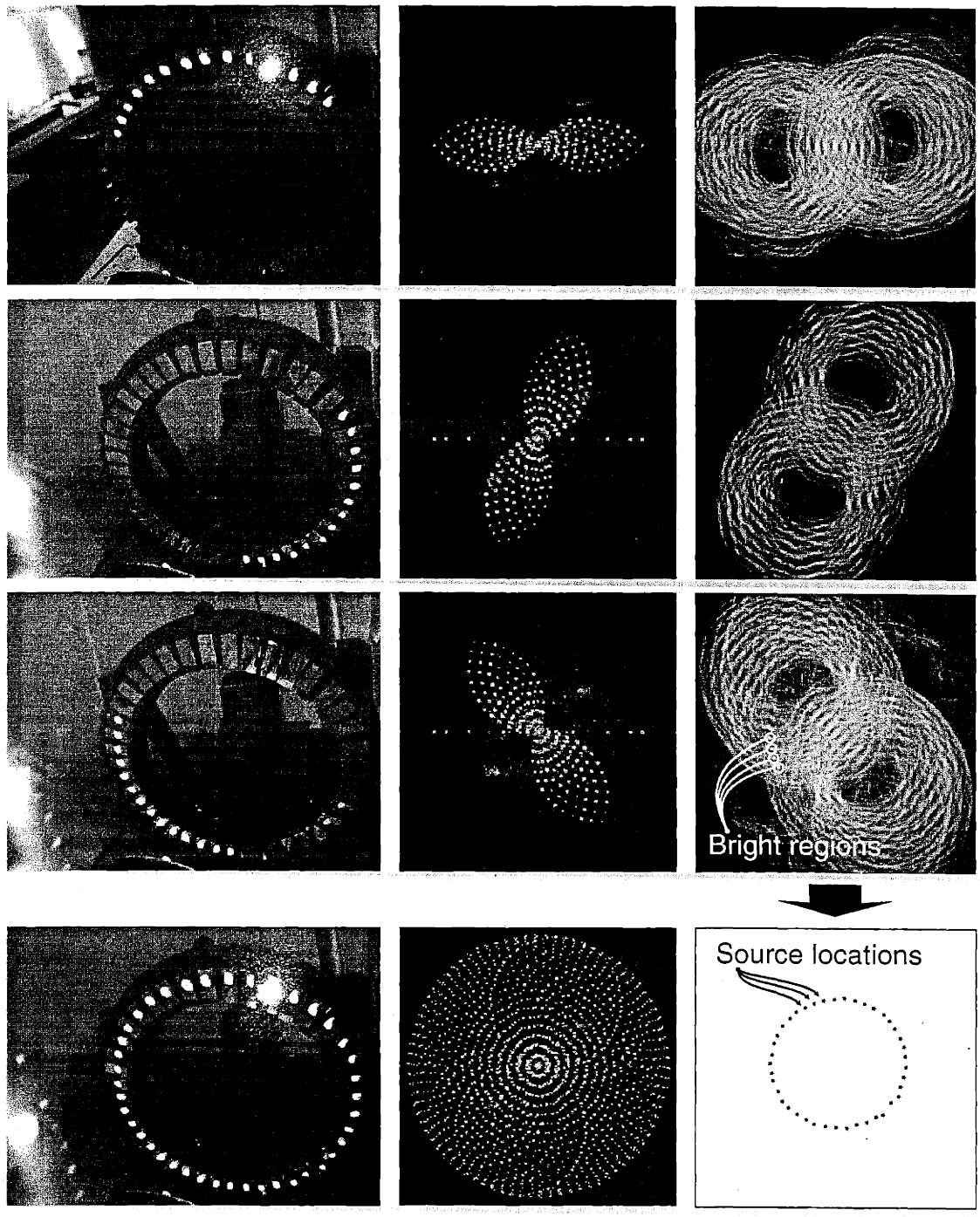

Figure 3-14: Covering about 2/3 of the beams in each of 3 reference images made recovering the source beam locations a lot easier. The left column shows which beams were on for each power spectrum in the center column. The right column shows the matched Hough transform (used to find the sources in k-space) from each sub-array (top three right panels), and the deduced source locations for the whole array (bottom right panel).

in Chapter 2.

Amplitude and phase values for the source beams were estimated from these reference images. (See Appendix B.) Figure 3-16 shows these values plotted in the complex plane for each beam over a one hour period of measurement. During the measurement period, the phase modulators were set not to change.

The plausibility and interpretability of the data further support the model of the many-beam interferograms developed in Chapter 2. Specifically, the data agree with the theory that the optical table deformed during the measurement, tipping the top of the grating wheel away from the top of the modulator array for a clockwise phase rotation of top beams and a counterclockwise phase rotation of bottom beams as plotted in Figure 3-16. The clustering of nearby points in time and the clear trend over time in the fully independent measurements indicate that the phase estimation process is valid. This validity itself testifies that the amplitudes and phases, and not just the locations, of the 1,641 hot spots in the transform have the expected relationship to each other. The illumination model is correct.

On the other hand, the data betray an unacceptable mechanical instability over the time periods needed to acquire the requisite thousands of frames with the low light levels finally reaching the target.

3.2.5 Conclusion: parallelism and reference concepts sound, optical system too large and too inefficient

M2 showed that phase sequences are measurable from reference data and that the theory of many-beam illumination is correct.

The apparatus was electronically capable of completing the measurement in tens of seconds. Beam splitting and steering methods used, however, provided less than one percent optical efficiency necessitating long runs (approximately 30 minutes) of the apparatus for measuring a target. Even so, because of the high degree of parallelism of the system, this constituted a five order of magnitude boost in information rate over M1. Unfortunately, though, this timescale proved incompatible with the tight

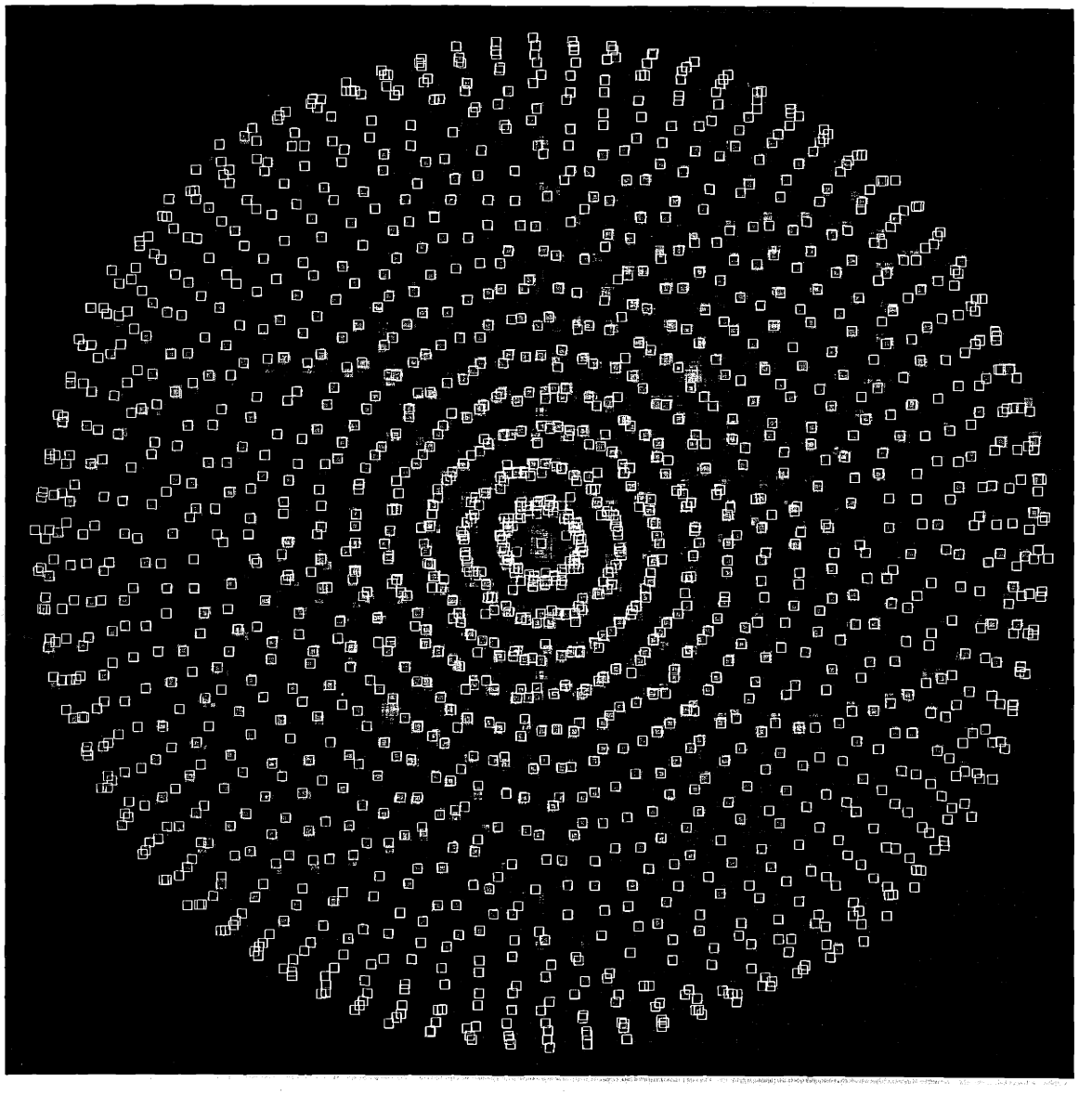

Figure 3-15: A major success of M2. Figure 3-13 with the hot spots tagged with square markers. The marker positions were constrained to predict consistent locations for the 41 source beams according to the theory derived in Chapter 2.

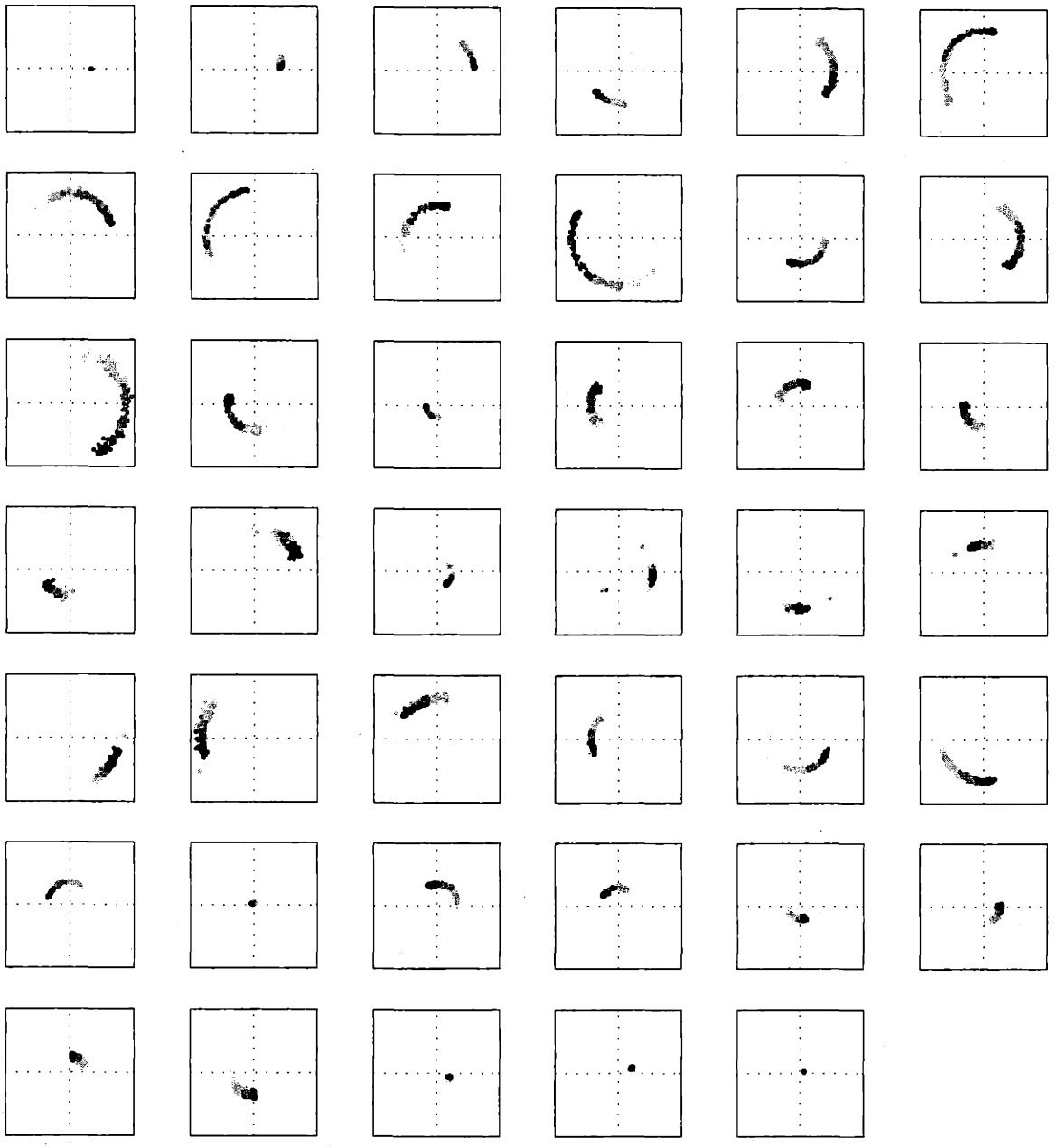

Figure 3-16: Another major success of M2, but also its major shortcoming. Amplitudes and phases of each beam were extracted from a sequence of reference patterns over a one hour period and plotted in the complex plane in each panel above. The darkness of the points encodes the time, with darker points later in the measurement. Reading the panels first left to right and then top to bottom, like the words on this page, reveals a trend in the phase angle motion of each next source around the ring. Even though the command voltages were not modulated during the measurement, the plots show a measured phase drift. The drift is consistent with an approximately eighth-wavelength (61 nm) change in positions of the rings.

limits on mechanical drifts and instability of the interferometric measurement.

3.3 M3: Micro Electro Mechanical System as experimental phasing element

M3 sought to fix the efficiency and drift problems of M2.

The efficiency was improved principally by sacrificing the direct connection between the geometry of the modulated beams with the geometry of the source beams. In other words, lifting the requirement for a ring-type beam splitter with a large, odd number of output beams opened up many commercially available options for high efficiency beam splitting.

Drift was essentially eliminated with a more compact, more rigid mechanical design that also kept beam paths very close to each other, minimizing the relative phase shifting effect of turbulence.

M3 made up the summer project of Stan Hong, an undergraduate member of Professor Freeman's research group. Figure 3-17 shows the apparatus.

3.3.1 Apparatus overview: MEMS, lenses, and fibers

A commercially available digital optic beam splitter diffracted, from a single input beam, 16 output beams in a 4×4 grid arrangement with high efficiency. A lens system brought the beams to a grid of focused points and allowed the points to be aligned with the microscopic structures of a Micro Electro Mechanical System (MEMS). The MEMS comprised an array of electrostatically actuated mirrors intended to replace the piezoelectrically actuated mirrors of M2 as path length modulating phase shifters.

Because the modulated beams were in a grid arrangement, some mechanism was needed to bring the source beams into a ring geometry. Single mode, polarization preserving optical fibers were chosen for the job.

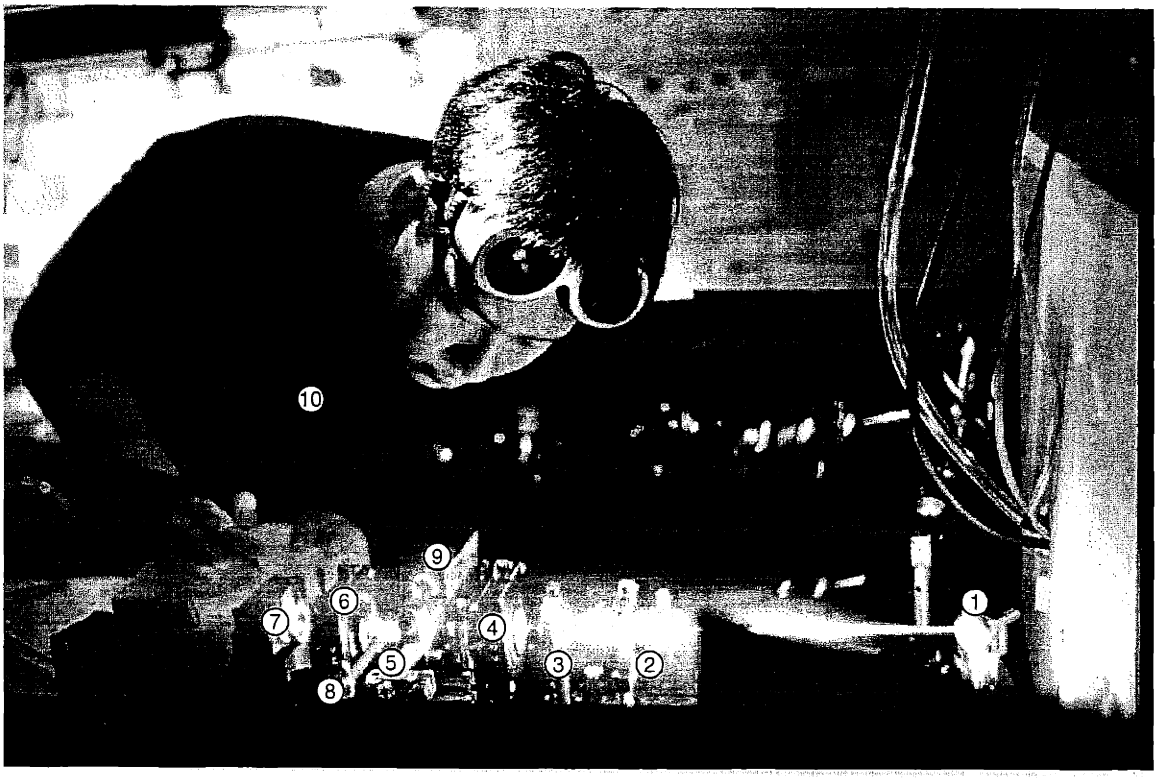

Figure 3-17: A photograph of M3, the third microscope prototype. Indicated are (1) a fold mirror bringing laser power from the argon laser (not shown); (2) a spatial filter assembly; (3) the diffractive beam splitter in a rotating mount; (4) a field lens on a five-axis gimbal mount; (5) a non-polarizing beam splitting cube; (6) microscope objective lens; (7) MEMS mirror array; (8) fold mirror to fiber grid array chuck (not shown); (9) a beam stop for the unwanted reflection from the beam splitting cube; and (10) Stan Hong.

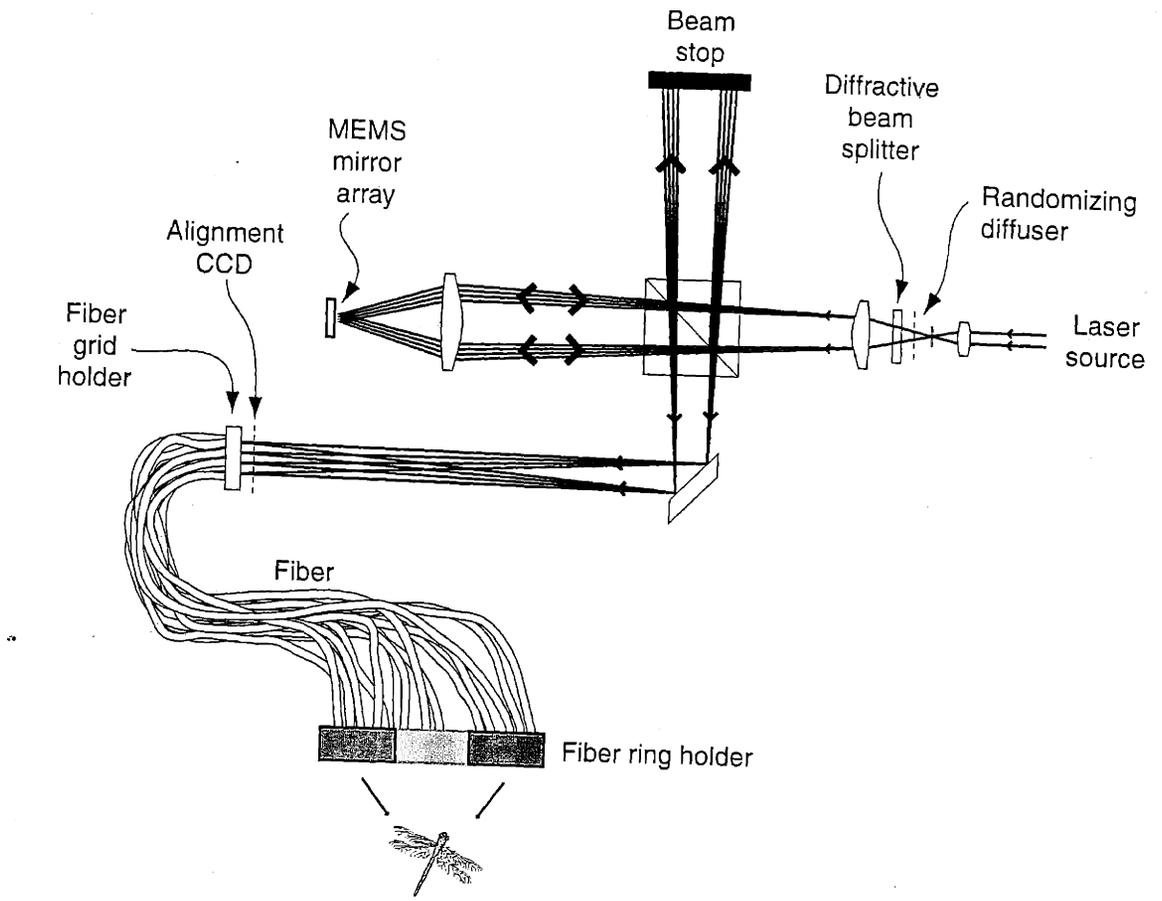

Figure 3-18: A schematic representation of the M3 illumination system. The two elements drawn as dotted lines were only used for alignment.

3.3.2 Detailed description of the apparatus

Figure 3-18 shows the apparatus schematically.

Highly efficient commercial diffractive beam splitter

Off-the-shelf diffractive beam splitters are available in a variety of regular patterns. There are two varieties: plastic stamped optics and glass etched optics. The latter type is better in all ways but price. A glass 4x4 spot array diffractive pattern generator optimized for 633 nm light was chosen (catalog number D53,191: Edmund Scientific Company, Barrington, NJ) and was found to have good performance at 488 nm (blue) and 514 nm (green). Figure 3-19 shows the pattern measured with a lens to focus the

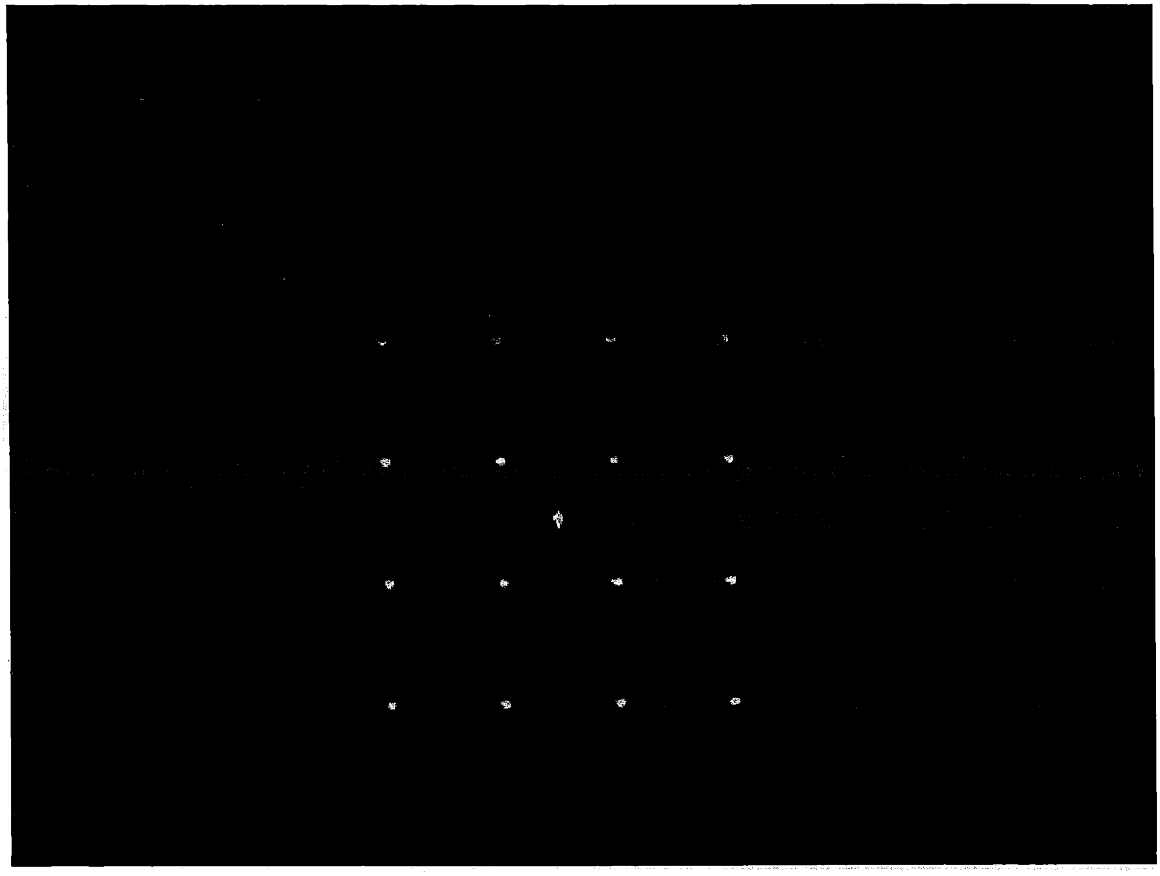

Figure 3-19: The grid of focused spots generated by the commercial digital optic diffractive beam splitter with focusing optics. The center dot in the pattern comes from the zero-order (straight through) beam.

grid of beams from the beam splitter to a grid of small spots on a CCD.

The beam powers were observed (with a power meter) to match with a 5 percent standard deviation at 488 nm and with a 7 percent standard deviation at 514 nm. The efficiency (the fraction of the input energy diffracted into the 16 output beams) was measured to be 59 percent at 488 nm, and 57 percent at 514 nm. Though some scatter was visible, most of the remaining energy was simply passed through the optic into zero the order. Efficiency strongly depended on filling a large fraction of the optic with an expanded, Gaussian-profile beam.

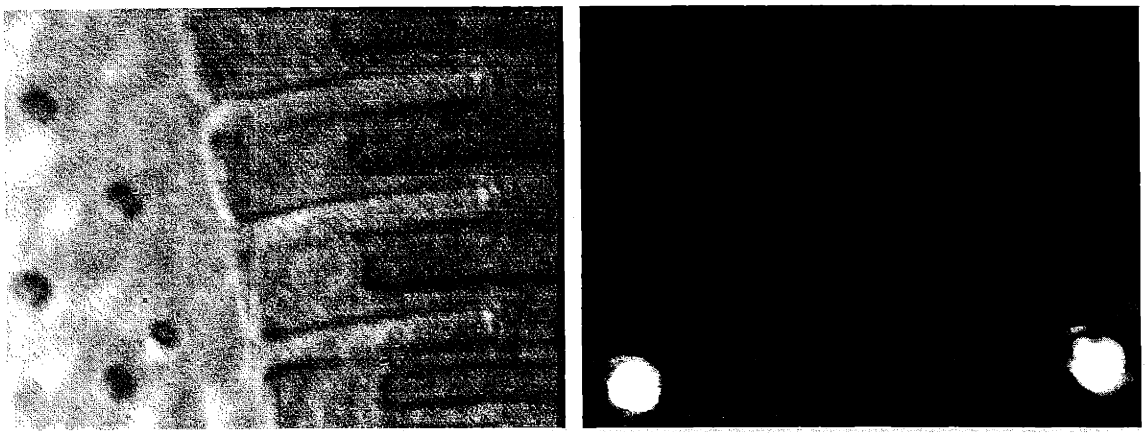

Figure 3-20: M3 alignment technology. The left image shows microscopic structures of a MEMS torsional comb drive illuminated with the bobbing randomizing diffuser in place. The right image has the randomizing diffuser removed, but otherwise no other changes in the setup were made. Two of the points from the 4×4 grid are clearly visible. Setting the diffuser in a larger bobbing motion gives an alternating view that makes aligning the focussed points of light to microscopic structures relatively easy.

High precision micro-alignment technology: the randomizing diffuser

Figure 3-18 has two elements drawn with dotted lines. These are used for system set up and alignment, but not during operation.

A randomizing diffuser, adjacent to the diffractive beam splitter, was supported with a steel spring. Set in a bobbing motion with a gentle tap, the optic added a random phase to the carefully engineered wave front exiting the beam splitter—undoing the highly ordered grid pattern in the far field. Thus the MEMS was illuminated uniformly.

A CCD imager was placed just in front of the fiber grid holder. With the randomizing diffuser in place but unmoving, the CCD would receive a speckled image of the MEMS. Setting the diffuser in motion randomized the speckle, and the integrated image was free of speckle and other coherent artifacts. With the diffuser removed, only the grid of focused points illuminates the MEMS, and only a grid of focused points lands on the CCD. Figure 3-20 shows a pair of these CCD images of a MEMS electrostatic comb drive, with and without the diffuser.

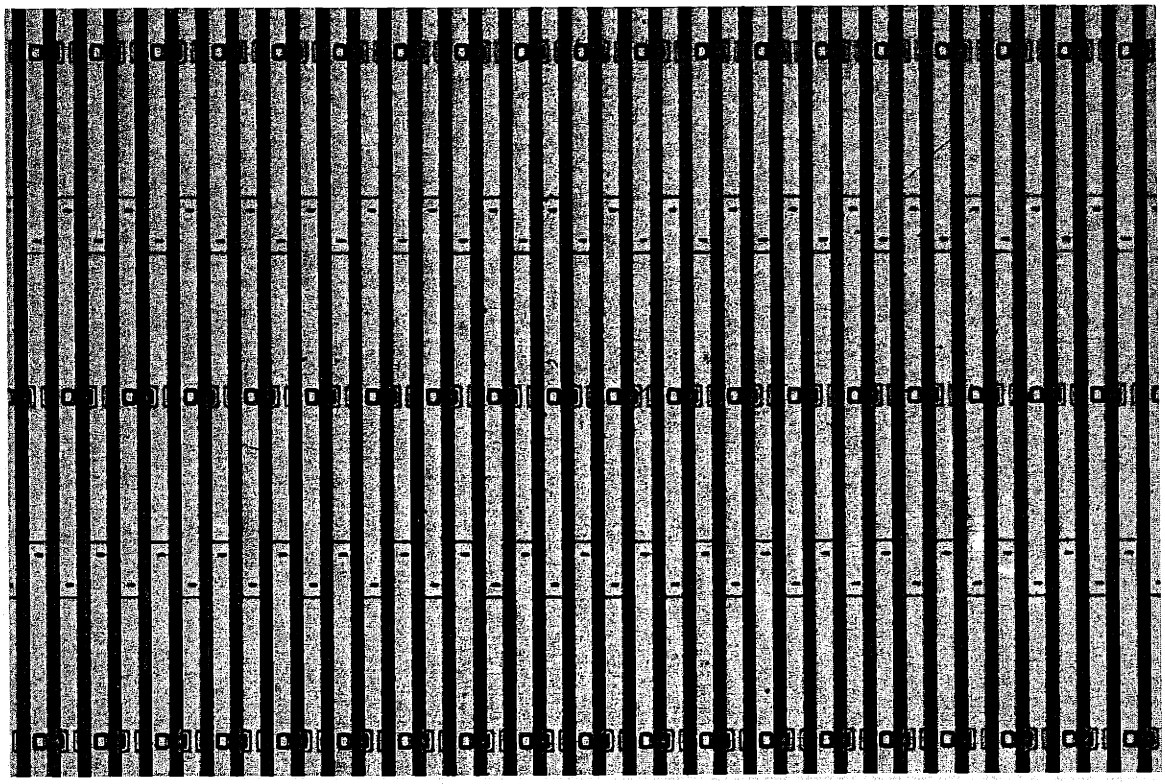

Figure 3-21: An optical micrograph of the MEMS mirror array used to phase modulate the source beams in M3. The structure has a nominal period of 12 μm.

Magnification of the spot array along with translation and rotation of the MEMS, all accomplished with manual, micrometer-driven stages, allowed the grid to be aligned precisely and confirmably to MEMS structures.

MEMS mirror array modulates beam phases

We were unable to get the MEMS originally promised to us for this project. The device contained a 4×4 array of electrostatically driven mirrors. Instead, we used 16 mirrors of the grating mirror array designed by Elmer S. Hung for the "Polychromator" project described in his Ph.D. Thesis [8]. Figure 3-21 shows an optical micrograph of the array, fabricated by the MCNC MUMPS process. The width of the mirrors was 10 μm, and the spacing was 2 μm (nominal).

Our strategy for using this structure with a 4×4 spot array is illustrated in Fig-

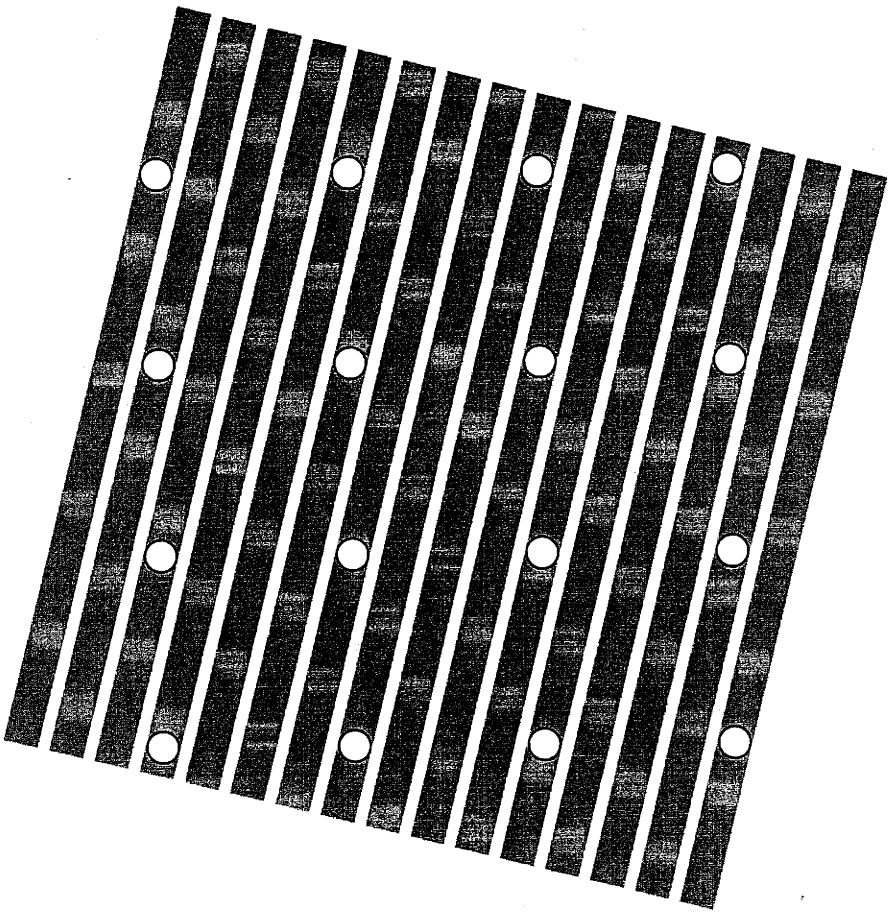

Figure 3-22: By using the MEMS device at an angle (of about 14 degrees) relative to the grid of focused spots, each beam could be independently modulated. The gray bars represent the MEMS mirrors and the white circles are the focused spots.

ure 3-22.

The frequency responses of individual mirrors were assessed with a laser Doppler vibrometer (OFV3001 vibrometer controller with OFV511 fiber interferometer: Polytec Corporation, Germany). A spectrum analyzer (Hewlett-Packard 3562A Dynamic Signal Analyzer) with a high voltage amplifier stimulated the mirror under inspection and recorded the response from the vibrometer. Typical data are shown in Figure 3-23. With a peak-to-peak excitation of 100 V, the mirror displacement was 800 nm peak-to-peak up to a few kHz.

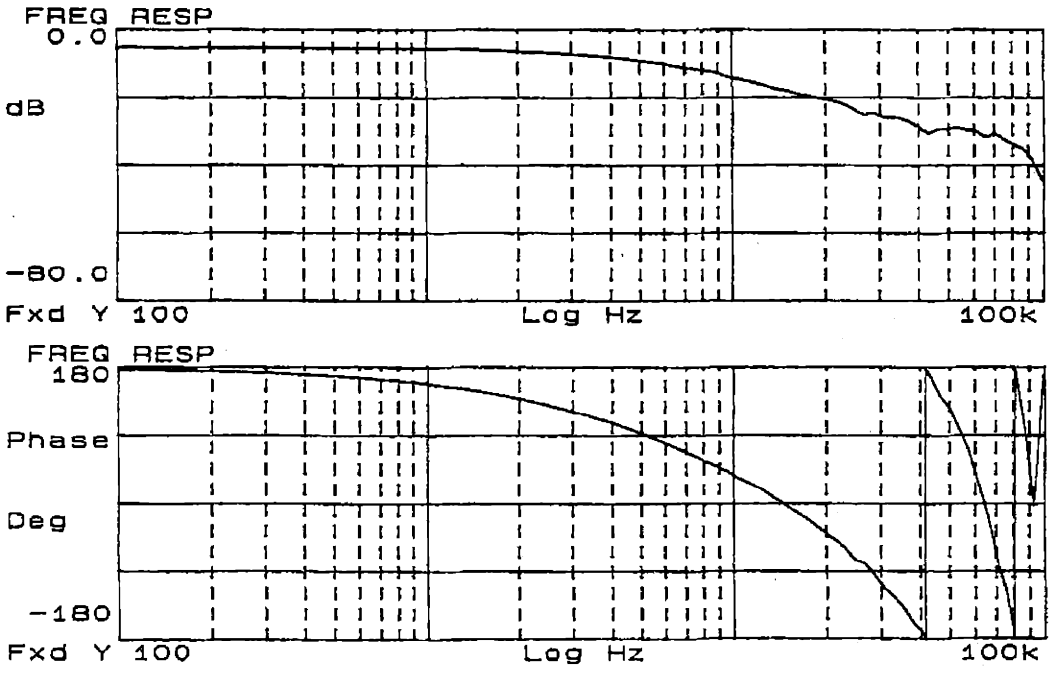

Figure 3-23: The measured frequency response of a single of the MEMS mirrors shown in Figure 3-21.

Fiber turns a square into a circle

Single mode polarization preserving optical fiber (catalog number F-SPA: Newport Corporation, Irvine, California) delivered the modulated beams from the grid geometry to a ring of source beams. Several generations of mounting schemes resulted in the plan sketched in Figure 3-24.

Better success in coupling the 514 nm (green) line into the fiber led us to standardize on that wavelength in the final alignment.

3.3.3 Conclusion: beam splitter efficiency a blessing, fiber a curse; MEMS setback stops work

Polarization control is necessary to ensure proper interference of the beams at the target. Polarization preserving fiber, in general, is single mode, and single mode fiber, in general, is tricky to align. Therefore, while fiber seems an obvious and simple solution for creating the ring of sources, our M3 fiber experience left us wanting greater coupling efficiency from the modulator to the source ring.

The MEMS modulator array worked well until, in an effort to couple more light into the fibers, too much laser power was applied to the MEMS, resulting in its total obliteration. Blame for this demise was shared by three factors:

1. the MEMS mirrors degraded the wave front quality of reflected beams because of surface roughness,

2. the single mode fibers act as spatial filters, passing only the fraction of the input beam with good wave front quality, and

3. the long, skinny MEMS mirrors were less capable of dissipating heat than the MEMS mirrors tested to determine the damage threshold.

With the summer ended, the undergraduate student assigned to M3 taking classes, the anticipated 4×4 MEMS mirror array still unaccounted for, and the apparently more promising M4 already underway, work on M3 never resumed.

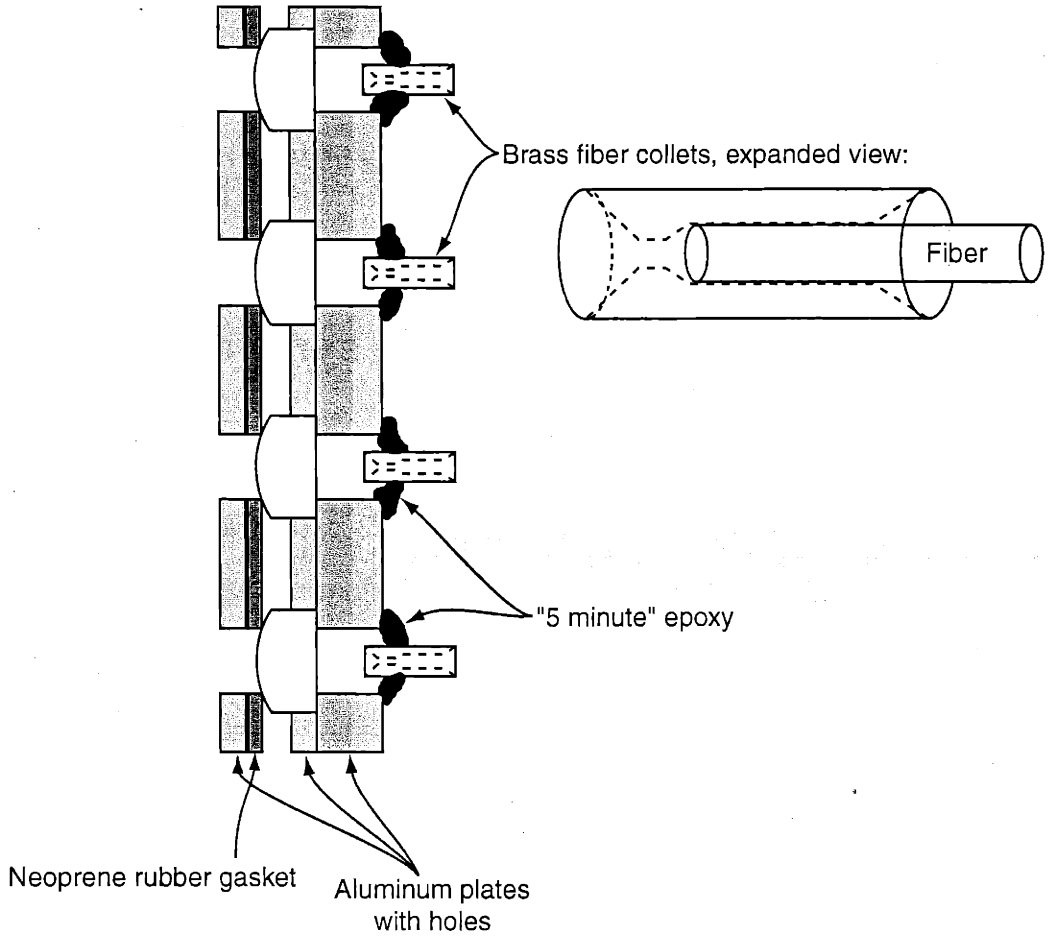

Figure 3-24: The fiber grid holder. An array of fiber collimators was fabricated to couple the grid of modulated beams into single mode, polarization preserving optical fibers. This simple design worked well and was the result of experimenting with many ideas for holding the array of lenses and fibers in place. Lenses dropped into holes in the middle aluminum plate, and were sandwiched in place by the rest of the assembly. The sandwich was held together with machine screws, not shown. Each fiber, press-fit into a brass cylindrical collet, was aligned to the focal point of its lens with a multi-axis stage. Once aligned, the collet was bonded in place with epoxy. The same mechanical technique was planned for the fiber ring holder.

MEMS mirror arrays with better surface quality and more compatible geometry exist. The approach taken in M3 continues to hold promise. Some successful strategies of M3 were appreciated early on and applied in M4—most notably

1. divorcing the beam splitter geometry from the source arrangement for improved optical efficiency, and
2. using bundles of beams (instead of wide fans as in M2) to minimize the effect of turbulence.

3.4 M4: An entirely solid-state synthetic aperture microscope

M4 answered the controllability issues of M2 and M3 with several innovative engineering solutions. The most important three were

- An acousto-optic modulator (AOM) strategy, providing
  - high efficiency beam-splitting
  - absolute linearity of phase control, and
  - modulation bandwidth in the MHz
- Reference camera viewing target area in transmission, allowing
  - near field coupling to the reference pattern for >1 N.A. reference information, and
  - alignment of reference imagery with the target camera
- Optimized periscope array design that is
  - mechanically compact,
  - based on flat mirrors and more optically efficient than fibers or gratings, and

– savvy of polarization and path length equalization.

Figure 3-25 shows the apparatus.

3.4.1 Apparatus overview: one AOM, fifteen radio channels, and thirty mirrors

The new modulating, beam steering, and camera arrangement concepts simplified and improved data acquisition speed, accuracy, and repeatability of M4 compared with previous generations. Figure 3-26 shows the apparatus schematically.

AOM splits and controls

The pulsing beam of a 685 nm (red) diode laser was diffracted into 15 first-order beams by a complex grating. The grating was comprised of 15 simple gratings superposed—formed in the volume of an acousto-optic material with sound waves of 15 different frequencies. The amplitudes and phases of these frequencies, set by a computer-controlled multi-frequency RF synthesizer, controlled the amplitudes and phases of the 15 diffracted beams. Subsection 3.4.2 explains how.

Periscope array turns a line of beams into a circle

A system of stationary mirrors and cylindrical lenses, assembled according to a computer optimization, directed the beams (fanned into a row by the AOM) to a region of overlap in a shallow cone from a ring of discrete locations.

Reference and target cameras view the same region of overlap

The target camera viewed this region of overlap in reflection at low magnification while the reference camera viewed in transmission at high magnification. A glass cover slide doubled as the target stage and as the front optical element of the reference lens system.

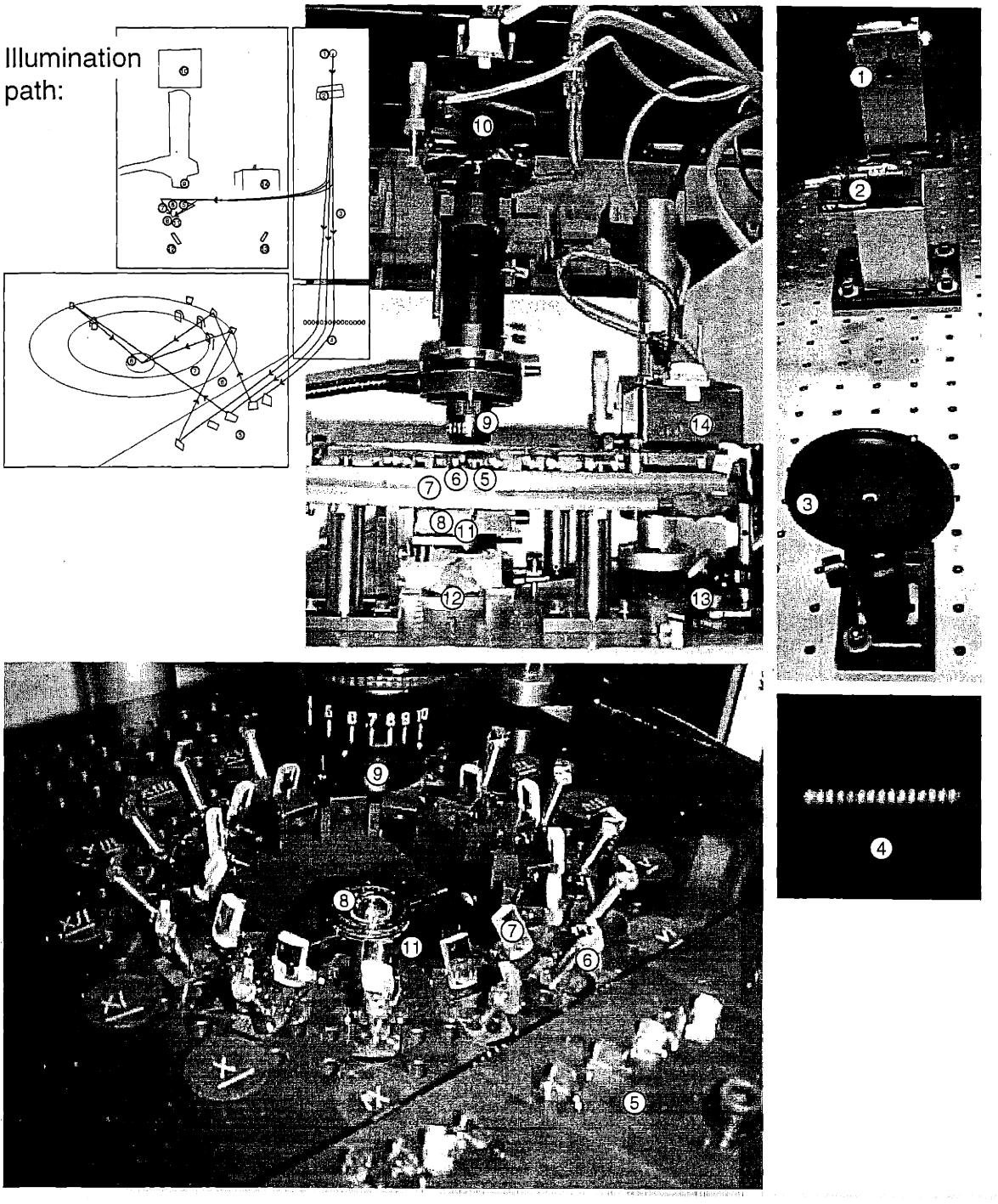

Figure 3-25: M4, the fourth microscope prototype. Labels indicate: (1) the diode laser; (2) the AOM; (3) an iris to block the zero-order beam from the AOM; (4) the row of 15 diffracted, controlled beams on a piece of paper just before they would enter (5) the fan mirror array; (6) the ring mirror array; (7) a ring of cylindrical lenses; (8) the cover-slip on which the target rests; (9) the video macro lens on (10) the target camera; (11) an oil-immersion microscope objective; (12) and (13) a periscope to direct reference images to (14) the reference camera.

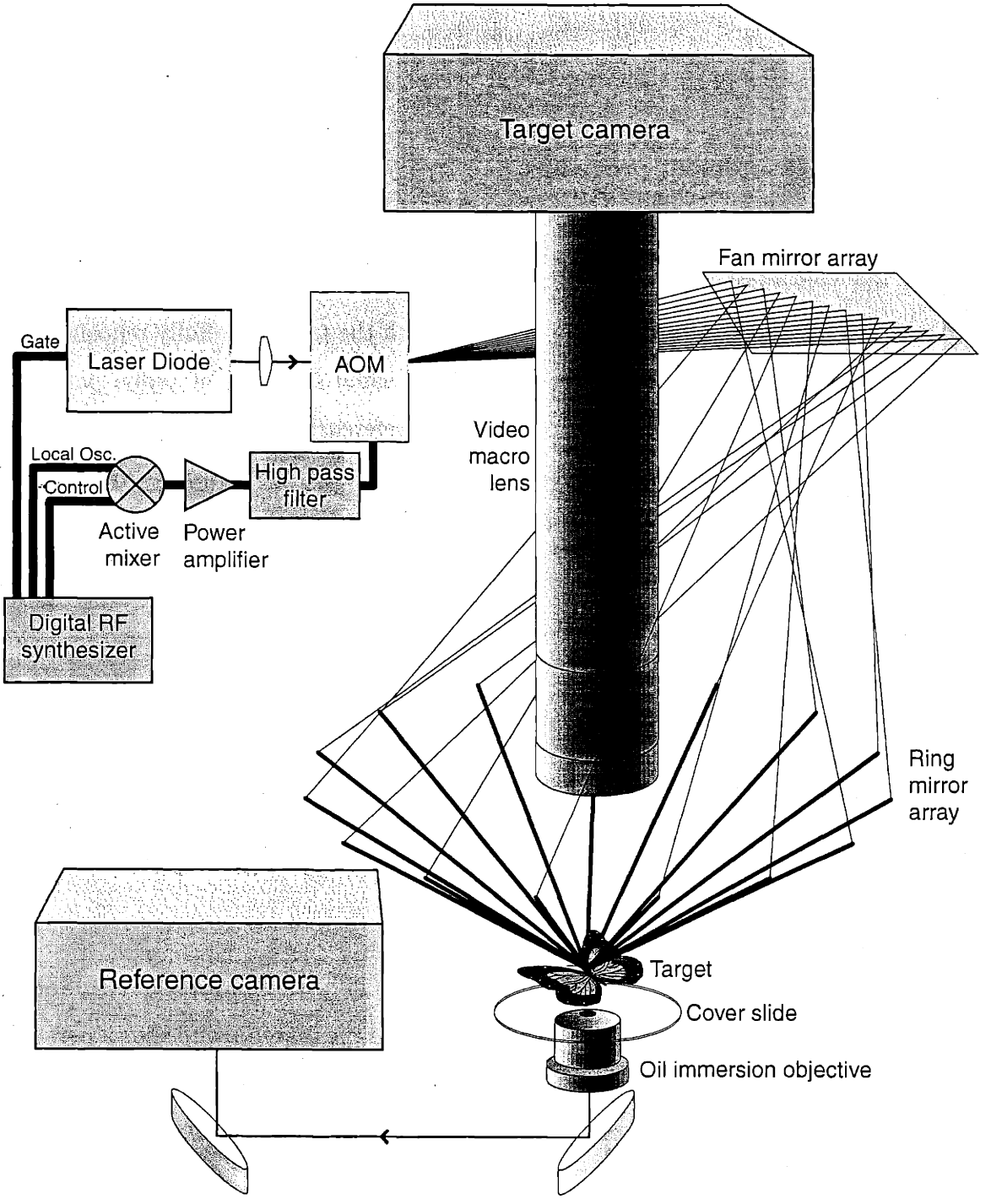

Figure 3-26: M4 schematic drawing.

System stable over many measurements

Reference and target data reproduced to acceptable tolerances over the time scale of tens of measurements. Therefore, the strategy of collecting reference and target data sequentially instead of simultaneously was acceptable.

Four-step measurement procedure

In a measurement, a thin target was brought to focus in both cameras. Next, the target was gently removed and reference data were collected in transmission with a sequence of hundreds of probe patterns. Then, with the target carefully repositioned, target data were collected in reflection with the same sequence of probe patterns. Finally, the target was gently removed again and reference data were collected again.

If the two reference data sets (collected before and after the target data) matched, the reference data was assumed to bear relevance optimally to the target data. Otherwise, the apparatus must have been perturbed accidentally at some point during the measurement, and new data were needed.

3.4.2 AOM beam control strategy

In M4, a single AOM generated multiple beams and also provided electronic (and fully independent) control of the phase and the amplitude of each beam. [5]

AOM background

An AOM converts electrical signals into sound waves in a transmissive crystal. Sound pressure influences the local index of refraction within the material and thus a periodic electrical signal generates a traveling phase grating capable of diffracting a laser beam. The idea is illustrated in Figure 3-27 (a). See [20].

High level overview

An AOM is driven by an electronic signal generator. The resulting acoustic waves propagating through the crystal form a traveling diffraction grating. Since the drive

signal can have several frequency components, the grating can be a superposition of several regular gratings. A laser beam incident on this aggregate grating diffracts differently for each acoustic wavelength passing through the crystal.

The resulting diffracted beams are combined (with additional optics) and interfere. The traveling grating imparts a different Doppler shift on each beam, resulting in a traveling interference pattern.

A central idea of the approach is that the traveling pattern can be "frozen" by amplitude modulating the laser power. One strategy is to supply laser power only when the traveling interference pattern corresponds to the desired pattern.

Controlling the amplitudes and phases of a discrete number of output beams

Generation of many beams: By driving the AOM with a family of distinct frequency components, an aggregate of regular diffraction gratings is formed within the crystal. A single beam traveling through the crystal diffracts off of the aggregate grating and forms a family of distinct beams in the far field. Since each beam diffracts from a grating of different pitch, each emerges from the AOM with a different first-order diffraction angle. Figure 3-27 (b) shows the idea with two beams.

In M4, an AOM is driven by the 15 frequencies from 56 to 98 MHz spaced by three megahertz. A single beam from a diode laser is split by the aggregate grating formed in the AOM into 15 first-order beams.

Independent control of each beam's amplitude: By varying the amplitude of a given "channel" of the drive frequency, the acoustic energy in the corresponding regular grating changes. Because the acousto-optic effect links the efficiency with the acoustic energy of a grating, the power in the beam corresponding to the selected grating changes as well.

As an example, in M4, if we doubled the power only at 59 MHz, the second beam (the one diffracted by the second smallest angle) would be brighter than the remaining fourteen. If we filtered away all of the energy at 59 MHz from the drive signal, that

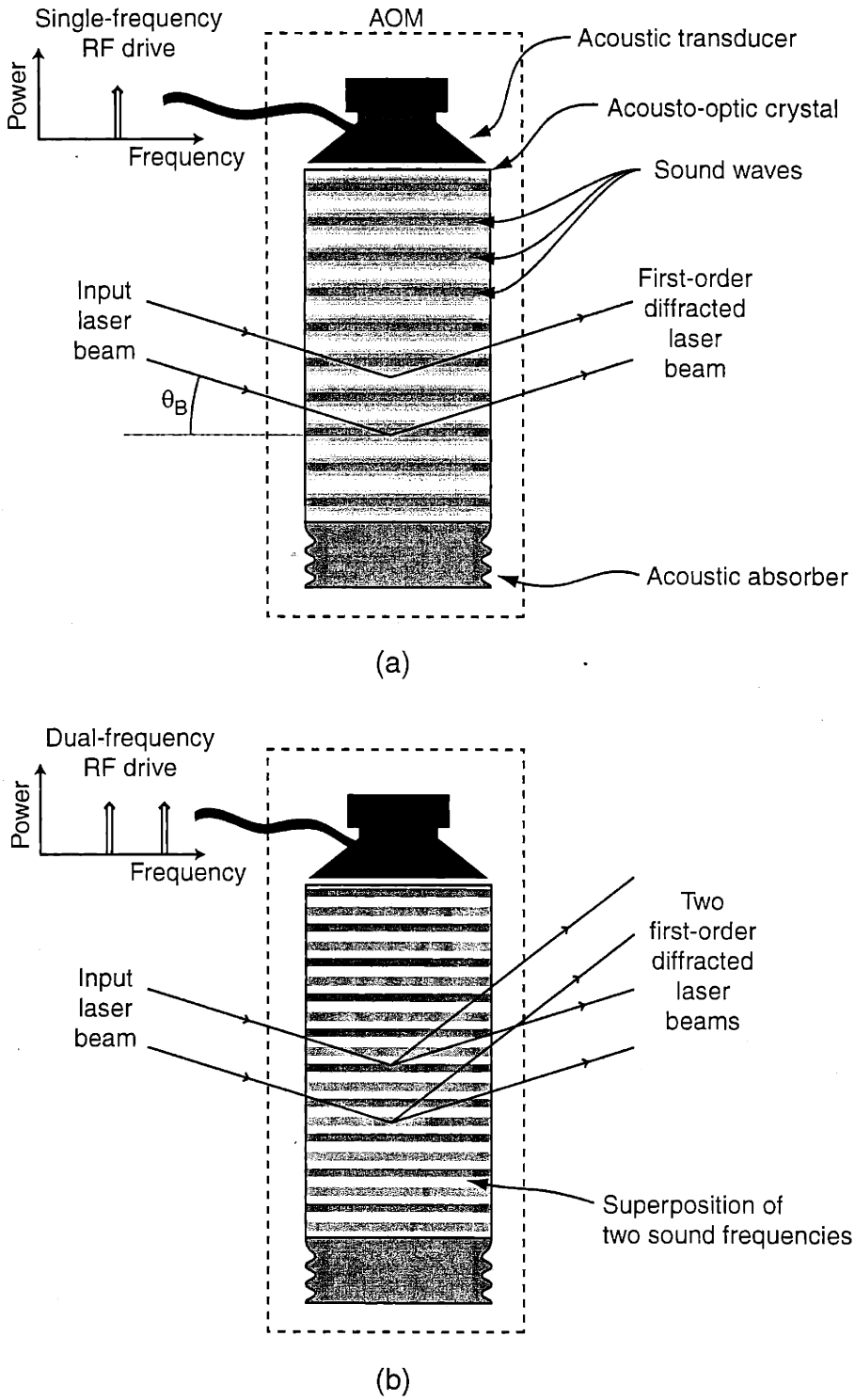

Figure 3-27: The basics of Acousto-optic modulators. Part (a) shows the component actions and actors with AOM's. Part (b) shows a superposition of acoustic gratings generating a pair of first-order diffracted beams.

beam would disappear.

Optical frequency-shift: An AOM shifts the optical frequency of a beam diffracted in the first order by the exact frequency corresponding to the drive-signal component leading to that beam.

Suppose the optical frequency of the diode laser is 438,000,000 MHz (red). The 15 first-order diffracted beams will have optical frequencies from 438,000,056 MHz to 438,000,098 MHz in three megahertz increments.

Combining the beams: A conventional optical system can redirect these beams into a region of overlap while controlling divergence angle, differences in path-length, and polarization.

M4 combines the beams with a pair of mirrors and a cylindrical lens in each beam path. The approach is described in more detail below.

The pattern in motion: The interferogram in the region of beam overlap is dynamic because of the optical frequency differences present in its constituent beams. The corresponding bandwidth sets the speed of the motion of the pattern.

Stroboscopic illumination: Suppose you wanted to take a still picture of the pattern with a camera. To eliminate motion-blur, you would choose a shutter speed that is fast on the timescale of the optical bandwidth. Supposing the pattern is in a dark room, you could put the shutter in front of the AOM or the laser instead of in front of the sensor. Equivalently, you could pulse the laser. With a sufficiently powerful laser, a single pulse may expose the sensor. Otherwise, multiple pulses are required. If those pulses occur when the phase relationships among the beams are the same as when the first pulse fired, the pattern will also be the same, and the image will be reinforced. If the pulses occur at different times, the integrated pattern will be the sum of several different patterns. Such a sum could be just as interesting or useful.

A phase relationship between 15 beams equally spaced in frequency repeats at the interval frequency.[3] *In M4, pulsing the diode laser with a four percent duty-cycle at 3 MHz (the interval frequency) provides sufficient contrast to apparently "freeze" the interferogram formed in the overlap region. Figure 3-28 shows the signaling strategy.*

Independent control of each beam's phase: Changing the phases of drive signal components relative to each other and to the laser modulation signal changes the effective phases of the beams and thus changes the pattern. Such changes could be informed by a measurement of the pattern produced with the previous set of phases.

3.4.3 Detailed description of the apparatus

The M4 apparatus is conveniently described by following electrical signals from the experiment control computer through RF electronics into the AOM, where it is converted to an acoustic signal and then to an optical signal, to the region of overlap, into the two cameras, and finally, back into the experiment control computer. Figure 3-26 is helpful for visualizing this path.

AOM signal requirements

The AOM chosen[4] (a lead molybdate acousto-optic deflector; model 1205C-2-804A: Isomet Corporation, Springfield, Virginia) operates efficiently in a frequency range from 50 MHz to 100 MHz (observed, the rated range was slightly smaller). Diffraction efficiency of the AOM was optimized with in the range of 1 to 1.5 W of RF power to its 50 Ω input.

[3]Notice that the phases of the beams do not necessarily repeat at the interval frequency (they do not with the M4 numbers given). What matters is that all of the the phase *differences* repeat at the interval frequency, which they in fact always will.

[4]The AOM was chosen for its low cost and ability, derived from its bandwidth, acoustic velocity, and aperture size, to project 15 isolated spots. The resolution (in "spots") can be calculated by considering the diffraction-limited divergence angle of a single beam in relation to the range of deflection angles provided by the modulator. A slow acoustic velocity enables a finer grating pitch and thus a greater deflection angle at a given acoustic frequency. A large aperture yields a smaller divergence angle of each beam. Modulators with large apertures, high bandwidth, and slow acoustic velocity provide the greatest resolution, therefore, but require the most expensive transducers and crystal materials.

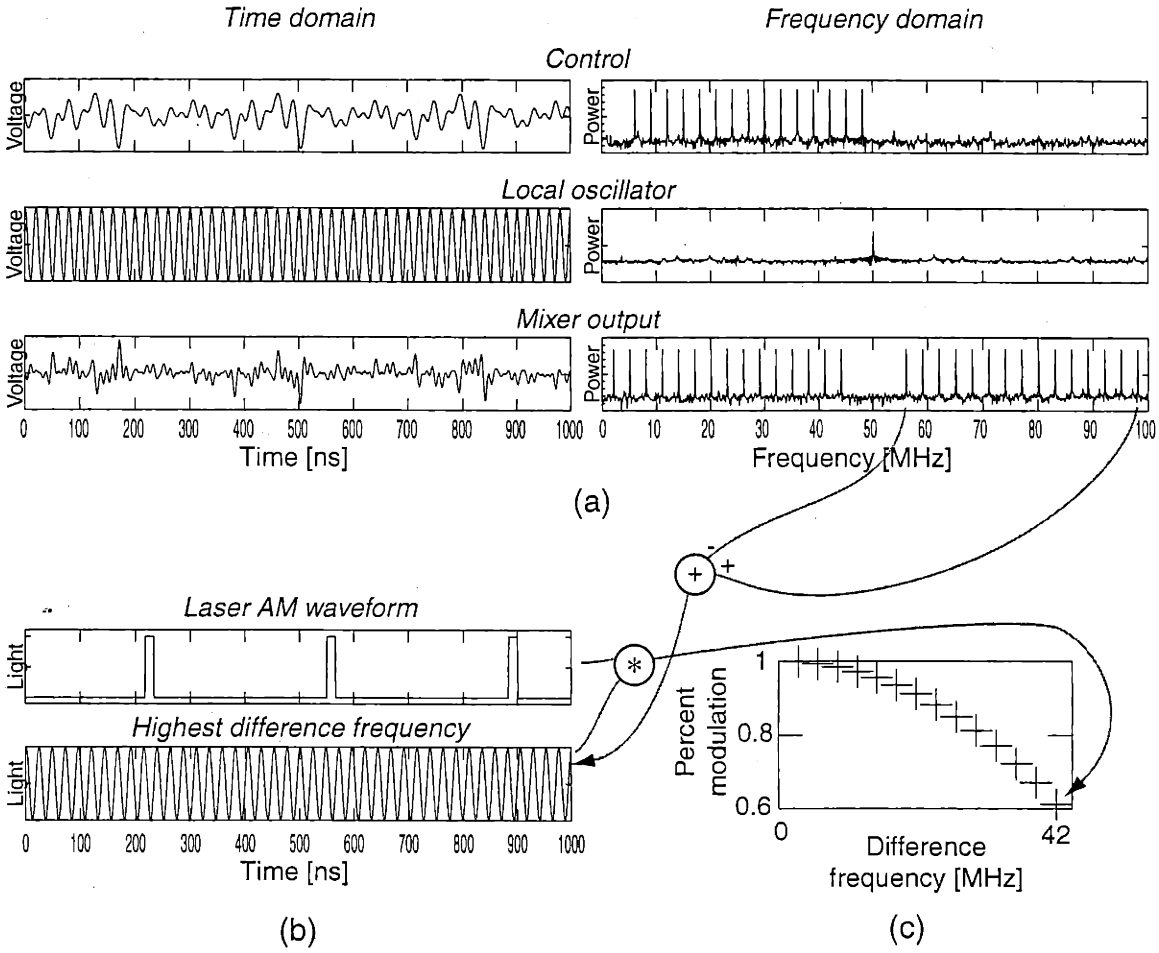

Figure 3-28: AOM and laser power signaling strategy for "freezing" the interferogram where diffracted beams overlap. Part (a) shows the control and local oscillator (LO) signals and their product. The spectral band from 50 to 100 MHz from the mixer output generates 15 acoustic pure tones in the AOM. Each of these tones diffracts a laser beam. Each pair of beams that interferes in a region of overlap forms a traveling fringe pattern. A point in this pattern would oscillate in intensity at the difference frequency of the tones corresponding to each beam in the pair. Pulsing the laser light as shown in the top panel of (b) at the spacing frequency of the tones stroboscopically "freezes" all of the traveling fringe patterns at a certain phase. The duty-cycle of the pulse train impacts the contrast of a frozen fringe pattern, depending on the difference frequency of the beam pair. The bottom panel of (b) shows the highest difference frequency on the same scale as the pulse train. Convolving the two signals produces the corresponding frozen fringe pattern. Part (c) shows the modulation contrast with the 4 % duty-cycle pulse train used in M4 as a function of difference frequency. Only one pair has the worst-case contrast, two pairs have the second worst, and so on.

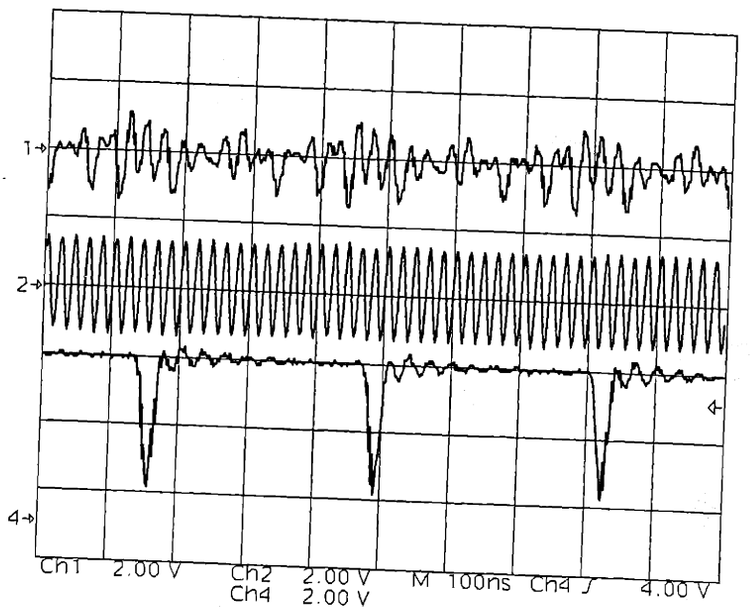

Figure 3-29: The three oscilloscope traces show the signals synthesized by the computer-controlled arbitrary waveform generator for M4. The top trace is the AOM control signal, the middle is the LO signal, and the bottom is the laser gate signal (negative true).

A 50 MHz bandwidth off-the-shelf synthesizer simplifies experiment control

A computer-controlled RF synthesizer (Sony/Tektronix model AWG2021 Arbitrary Waveform Generator) generated these three signals:

1. a **Control signal**, containing the sum of 15 sinusoids in the range of 6 MHz to 48 MHz spaced by 3 MHz,

2. an **Local Oscillator (LO)**: a 50 MHz 1.5 V peak-to-peak sinusoid, and

3. a **Laser gate signal**, a 3 MHz negative-true TTL pulse train with a duty cycle of four percent.

Since all three signals were derived from the same master clock (the synthesizer's internal clock running at 248.0 MHz), the signal relative phases were well-determined and steady. Figure 3-29 shows representative oscilloscope traces of the signals.

Active broadband mixer heterodynes control signal against LO

Although an RF synthesizer with 98 MHz bandwidth would have simplified the setup, a heterodyne approach [23] with the 50 MHz bandwidth synthesizer was chosen to save money.

To shift the control signal into the range of frequencies in which the AOM operated efficiently, an active multiplier circuit (based on the DC to 250 MHz AD835AN multiplier chip from Analog Devices) heterodyned the control signal against the LO. Passive mixers (model numbers ZP-3LH-S, ZP-3MH-S, ZLW-3H-S, and ZP-3H-S: Mini-Circuits, Brooklyn, New York) were found to provide insufficient dynamic range and freedom from distortion. Figure 3-30 shows the function and performance of the mixer.

Power amplifier boosts signal

A 5 W class A power amplifier (1 MHz to 500 MHz, +40 dB Gain, $\pm$1.5 dB worst-case flatness, +37 dBm maximum power into 50 Ω, model number ZHL-5W-1: Mini-Circuits, Brooklyn, New York) boosted the signal after 4 dB of attenuation (model SAT-4 4 dB attenuator: Mini-Circuits).

While only 1.5 W of RF power were actually required to optimally drive the AOM, this 5 W amplifier was chosen to avoid signal clipping when the 15 control signal frequencies conspired in phase to produce high input voltages. Figure 3-31 illustrates the potential problem. The fourth panel in Figure 3-30 shows the amplifier performance under typical conditions.

High pass filter improves AOM performance

Nonlinearities in the AOM yielded more favorable diffraction performance with only drive signals in the 50 to 100 MHz range present. A high pass filter (5 W 11-pole elliptical passive filter, model number H724-56M-50-720A: TTE Incorporated, Los Angeles, California) blocked the shifted control signal negative frequencies (from 2 MHz to 44 MHz) from the amplified signal with better than -50 dBc suppression

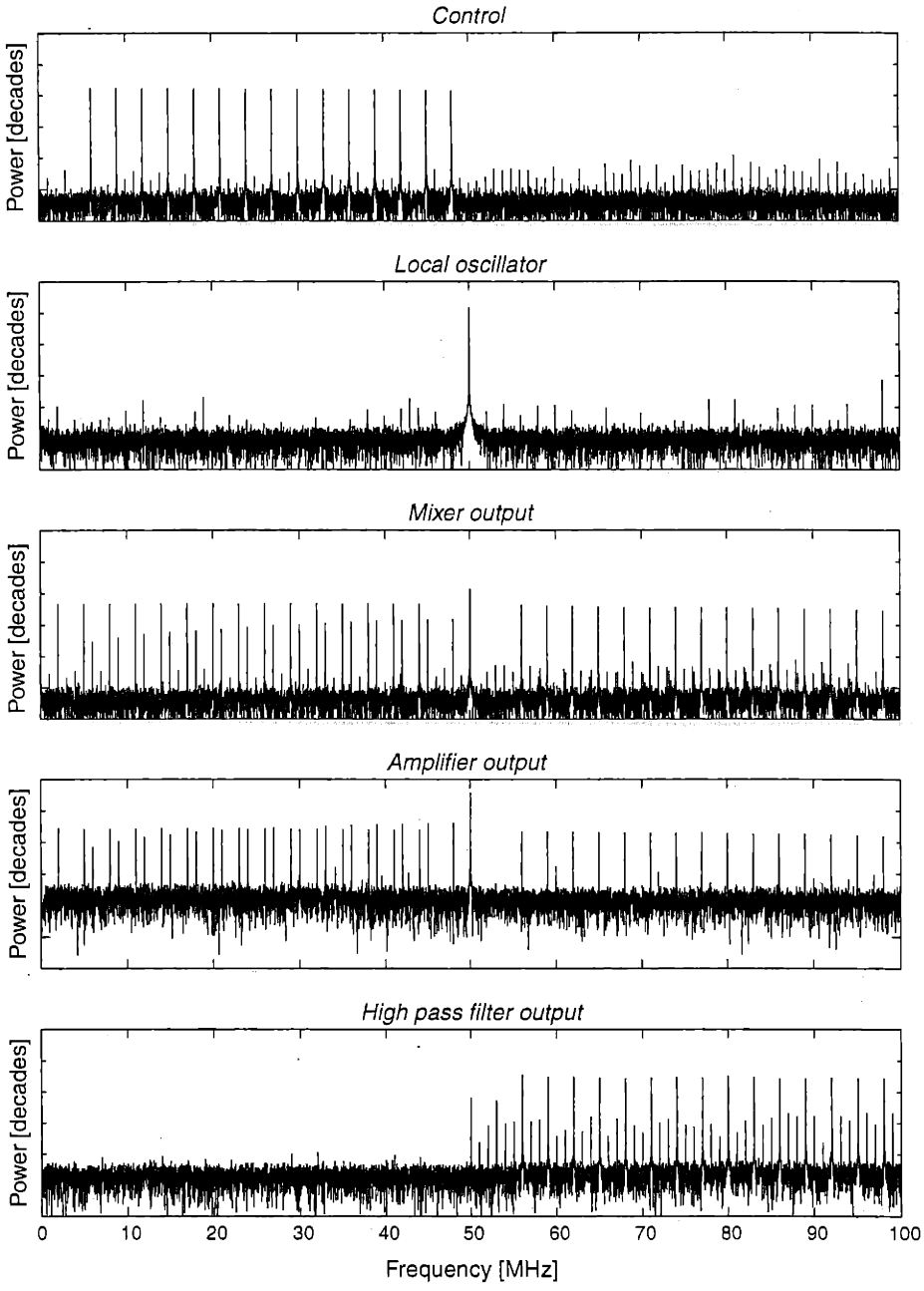

Figure 3-30: Measured power spectra of signals leading to the AOM

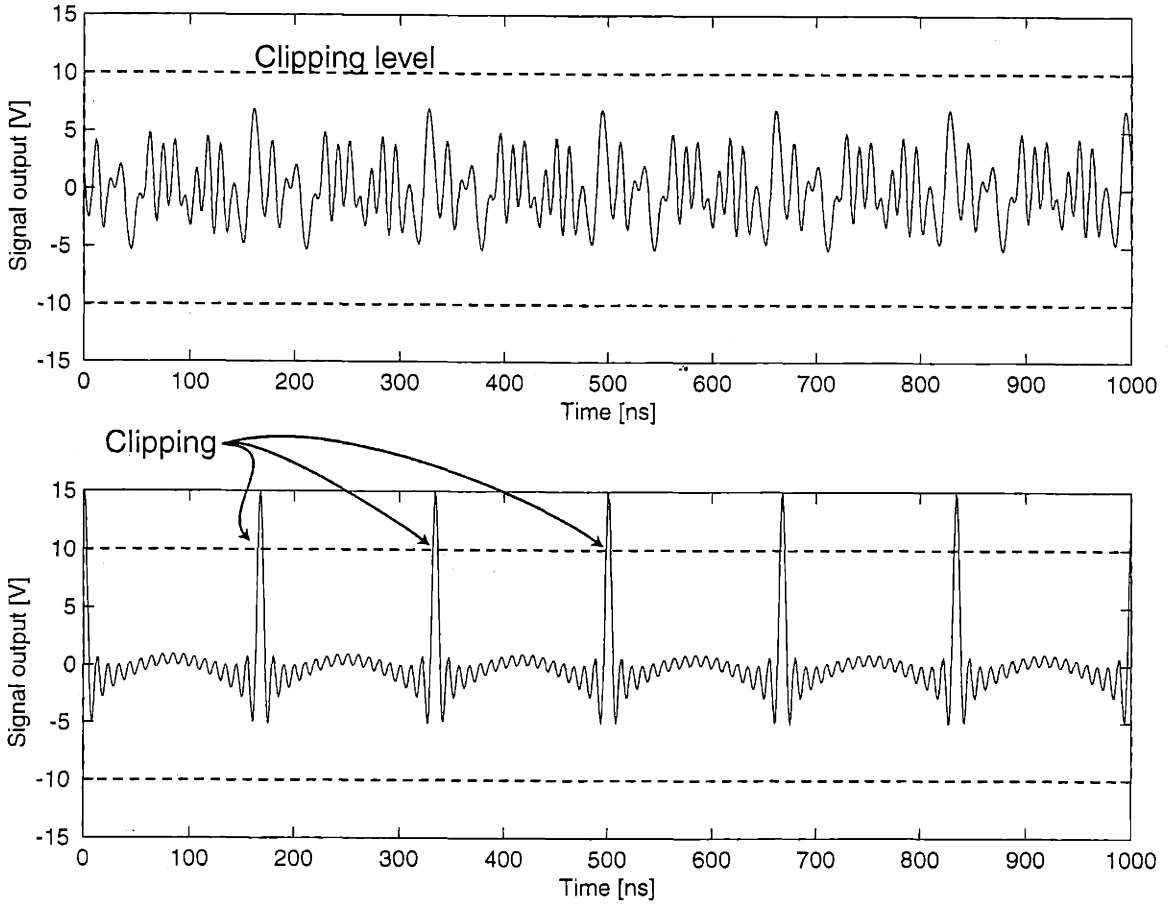

Figure 3-31: With random phases, the maximum AOM signal voltage is within the range of an normal power amplifier (top plot). Conspiring phases (bottom plot), on the other hand, require a much higher dynamic range and therefore, generally, a more powerful amplifier, even though the RF power is the same in both cases.

while passing the shifted positive frequency range (from 56 MHz to 98 MHz) with better than -2 dBc performance. The bottom panel in Figure 3-30 shows the filter performance under typical conditions.

Laser beam is strobed, split, and controlled

The TTL laser gate signal amplitude modulated a 685 nm (red) diode laser (single-longitudinal-mode, coherent for >20 cm, model number PMT50(685-60): Power Technology Incorporated, Mabelvale, Arkansas). The CW power of the laser was measured to be 46 mW (50 mW rated), and, when driven with a four percent duty cycle pulse train (13 ns negative-going pulses at 3 MHz), the laser produced 1.8 mW (measured CW equivalent) of light as expected.

The elliptical (virtually collimated: converging with a 2,286 mm focal point) beam had a horizontal major axis and a vertical linear polarization. The (2×6 mm) AOM window was horizontal, and the acoustic wavefronts were vertical. The angle of the input beam was adjusted to maximize the power diffracted by the AOM. In principle, this adjustment satisfied the Bragg condition as well as possible with the given range of deflection angles. The 15 diffracted beams fanned out with a beam-to-beam angle of approximately 0.03 degrees (measured).

Periscope array design

An array of mirrors redirected this fan of beams to a 145 mm diameter ring of locations in the plane of the fan. A mirror in each of these locations redirected its beam into a region of overlap 30 mm below the center of the ring. **The equivalent N.A. of this illumination cone is 0.92.**

The locations of all the mirrors as well as the permutation of which beams from the fan connect to which locations in the ring were computer-optimized with the following constraints:

- All of the path lengths must match as well as possible, and no length may differ from any other by more than 5 mm.

- Mirror edges must cross as little beam intensity as possible, and no edge may cross a region of intensity greater than one percent of the beam maximum intensity.

- The beam polarization angle after the compound-angle reflections must agree with the design angles shown in the upper left panel of Figure 3-32.

- No beam can pass through a mirror or through a mirror mount.

Path length matching: The path-length matching is important primarily for convergence reasons, and not coherence. The point is to ensure that the different beams spread to equal areas and apply equal energy densities to regions of the target.

Polarization matching: Best interference contrast is achieved when the polarizations of the beams mutually agree as well as possible. Linear polarization, however, is constrained to be perpendicular to a beam's propagation direction. This constraint gives rise to two regimes in which different strategies optimize polarization agreement and thus interference contrast. In systems with relatively pointy cones of beams (low N.A. systems), the best polarizations all point substantially perpendicular to the direction the cone points. In shallow cone systems like M4 (high N.A. systems), the best polarizations point to the middle of the cone. Figure 3-32 illustrates the idea.

Cylindrical lenses make round beam profile in the plane of the target

Lenses cut from a plastic lenticular array (catalog number H43026: Edmund Scientific Company, Barrington, New Jersey) with about a 9 mm focal length were glued 18 mm along the beam path from the ring mirrors. The lenses made the slice of the region of overlap occupied by a thin flat target into a 5 mm circle.

Mirror and cylindrical lens array assembly aided by automation

The positions for the first array of mirrors, calculated by the computer-optimization described above, were programmed into a numerically controlled milling machine (a

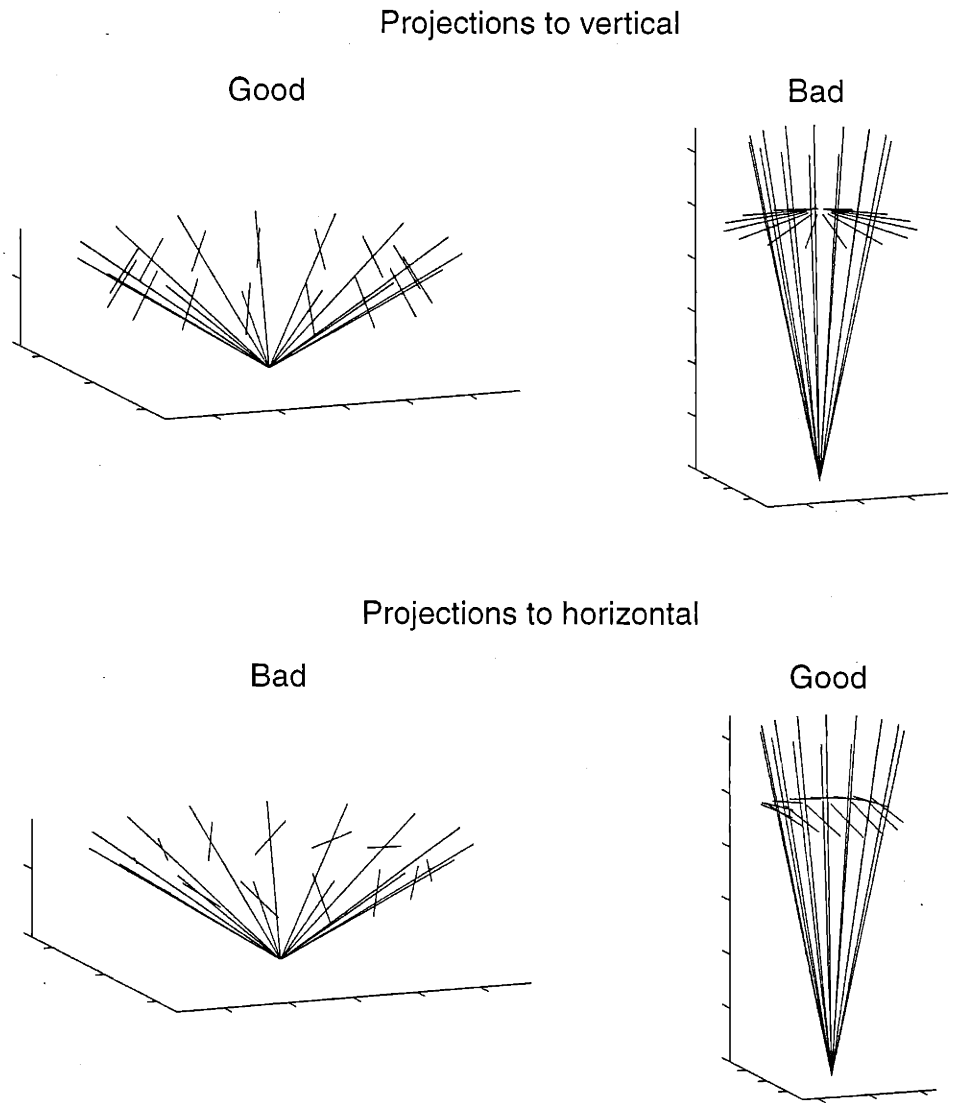

Figure 3-32: The best polarization strategy depends on N.A. Polarization vectors are drawn on the beam k-vectors in each plot. Beams with polarizations that generally point in the same direction will interfere with good contrast. Beams with nearly perpendicular polarizations will interfere poorly. Good arrangements (labeled) provide high interference contrast for every beam pair. The polarization strategy adopted for the high N.A. synthetic aperture of M4 is shown in the upper left.

Bridgeport mill controlled by TRAK A.G.E. 2 v2: Southwestern Industries, Inc., Rancho Dominguez, California), and guide points were milled into an aluminum substrate (strain-relieved, type 300 aluminum tooling plate, ground to half-inch thickness) with a nominal accuracy of ± 12.7 μm over the extent of the substrate. Each mirror in the first array was aligned with these guide points and held in place using a manually adjustable precision stage. A short length of fine wire, protruding from each of a ring of holes in a second plate, defined an aim point for each beam. When the light reflected straight back from the wire was optimized, two drops of five-minute epoxy (Devcon five-minute epoxy: Devcon Corporation, Danvers, Massachusetts) were applied between the mirror and the substrate. The stage was disengaged from the mirror once the epoxy completely hardened.

With the first array of mirrors glued in place, the array of wires was removed and the second array of mirrors (the ring) was assembled. The ring-mirrors required a more sophisticated mounting system comprising a large plate with 15 channels in a star burst pattern, aluminum blocks that slid in the channels, and hardened steel dowel pin posts that swiveled in the blocks. The star burst plate, milled on the same machine out of the same material as the substrate for the first array of mirrors, is shown in Figure 3-33. The arrangement constrained the mirrors to be substantially equally spaced in angle about a common center. A photo diode was placed in the center of the ring, 30 mm below the level of the beams with the help of a special jig. A manually adjustable precision stage and five-minute epoxy were used as before, to position and fix one mirror at a time, optimizing the power received by the photo diode from each mirror.

A second set of blocks and a similar strategy was used to mount a cylindrical lens in each channel on the star burst plate.

In the five months since the alignment was finalized, no pointing drift was observed.

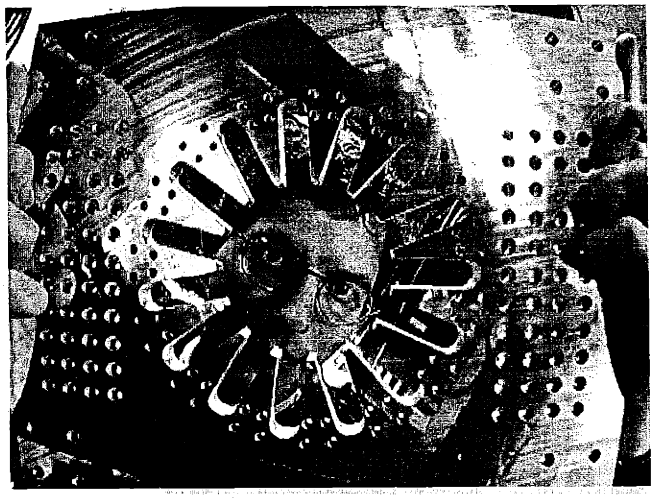

Figure 3-33: The channels in this plate simplified mirror and lens mounting and alignment. They were milled by computer control.

Parfocal cameras image region of overlap

The two CCD cameras from M2 (200 fps with 256×256 16×16 μm pixels and unity fill-factor, model number CAD1-0256A: Dalsa Inc., Ontario, Canada) were reused in M4.

Because the beams enter the region of overlap in a very shallow cone, a large collection aperture was required to image the reference pattern. This large aperture was achieved through physical contact (near field communication) of the reference imaging system with the pattern. The reference camera imaged the region of overlap in transmission at 175× magnification, sampling the scene on a 90 nm spaced grid. This high magnification (and at least 180 nm resolution) was achieved with an oil immersion microscope objective (100×, 1.25 N.A.: Nikon), a drop of immersion oil (nD = 1.515: Leitz, Germany), and a cover glass (22 mm circle, No. 1, selected micro cover glass, catalog number 48380 068: VWR Scientific). Manual micrometer adjustment of the microscope objective (five degrees of freedom) and of the cover glass (focus only) allowed a thin sample to be focused accurately on the reference camera sensor.

The target camera imaged light scattered from a thin target in the region of overlap

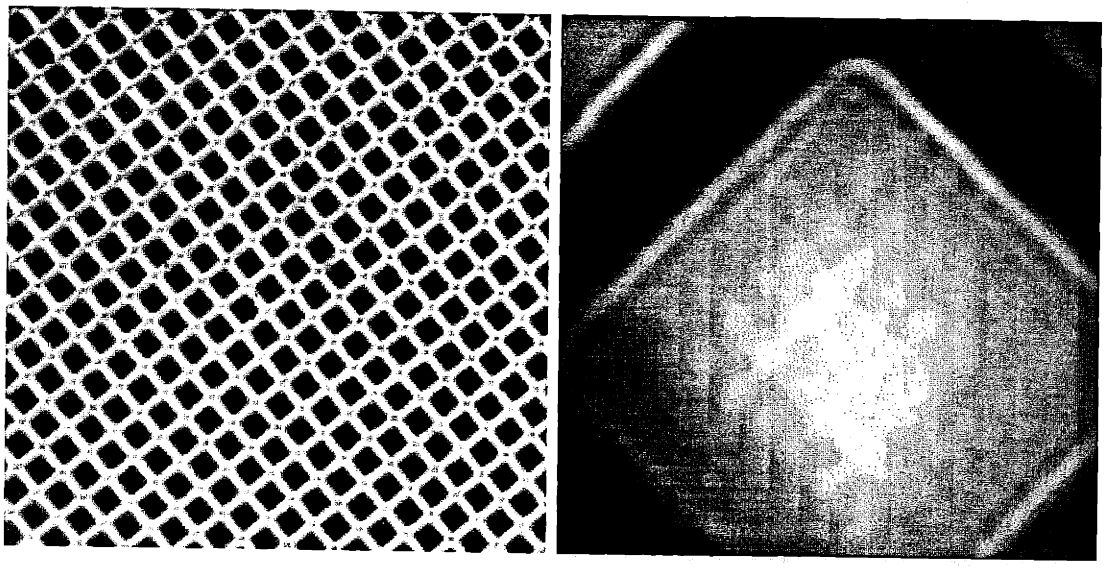

Figure 3-34: The relative magnifications and complementary contrast of the target and reference cameras can be appreciated from these images of a 1,000 mesh TEM grid. Grid spacing is 25.4 μm.

with a 10× magnification and greater than 30 mm of working distance through a zoom video macro lens (VZM model 1000 microscope head, catalog number F54396: Edmund Scientific Company).

The relative magnifications and complementary contrast of the target and reference cameras can be appreciated from Figure 3-34, which shows images collected from the two cameras with a 1,000 mesh (1,000 squares per inch) TEM grid (Gilder G1000HS copper TEM grid purchased from Ted Pella Incorporated, Redding, California) as a target with white-light illumination.

3.4.4 Phase data

Reference image Fourier transform agrees with theory

Figure 3-35 shows a typical reference image from M4. Its power spectrum, shown in Figure 3-36 shows a pattern of 211 "hot spots," two for each beam-pair, and one at DC (in the center of the pattern).

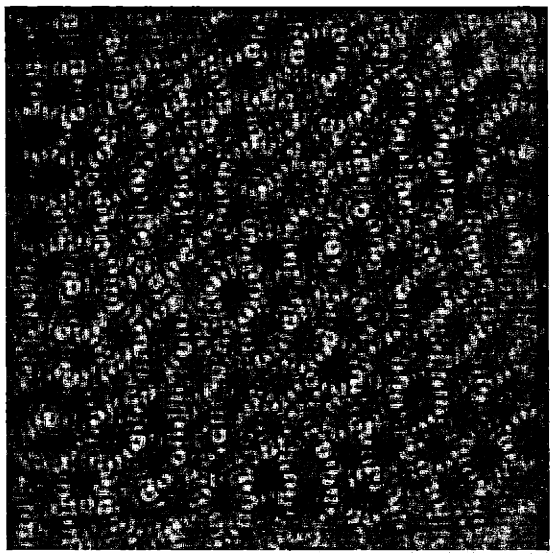

Figure 3-35: A typical reference image from M4.

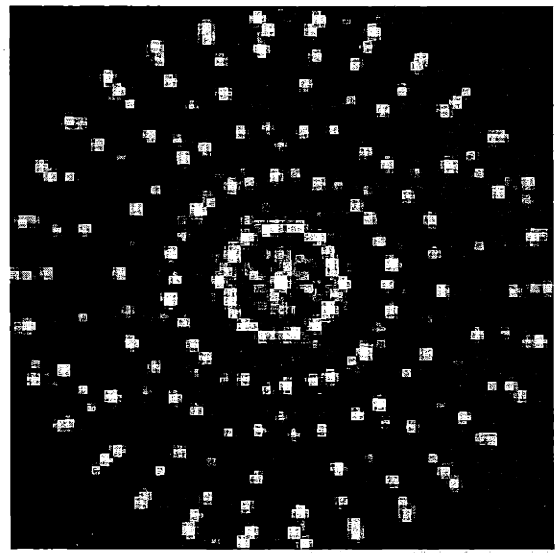

Figure 3-36: The power spectrum of the data in Figure 3-35.

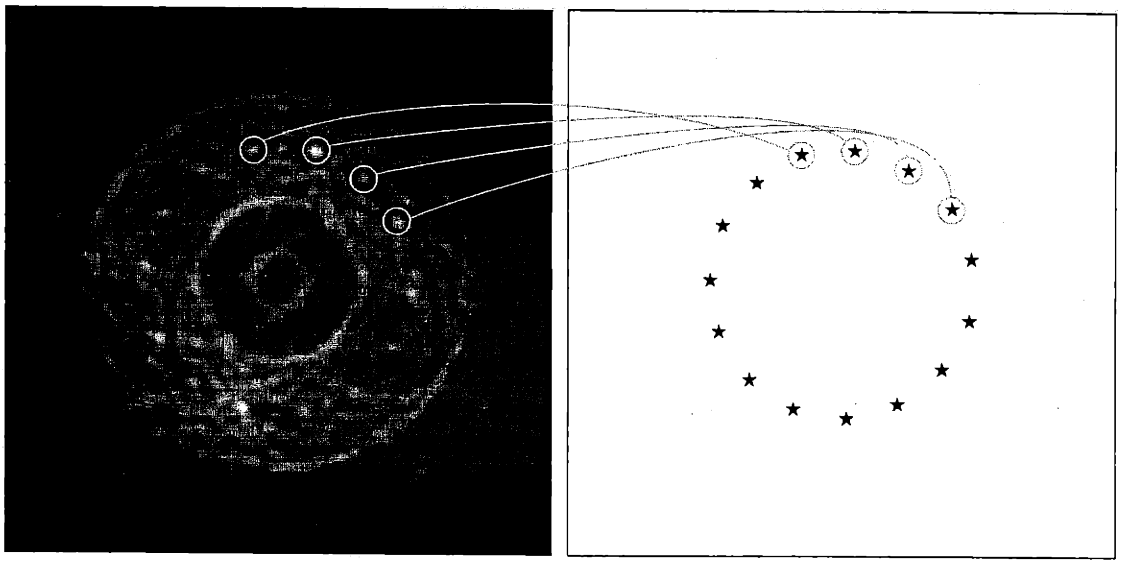

Figure 3-37: A matched Hough transform (left) aids in locating the beam k-vectors (right).

Hough transform locates beams roughly

According to the theory developed in Chapter 2, the hot spots should be distributed according to the auto correlation of the ring of sources. In general, converting 2D auto correlation data to its parent function is an inverse problem. However, given our special knowledge that the parent function is an array of 15 sources approximately equally spaced around a ring, the inverse problem is simplified.

A matched Hough transform algorithm was developed by Berthold Horn for this problem. The transform is computed by replacing each pixel in the power spectrum with the line integral around a circle centered on that pixel. The circle has the supposed radius of the ring of 15 sources, sufficiently estimated to be half of the radius of the pattern of hot spots. Figure 3-37 shows the result of the transform applied to Figure 3-36. The ring of bright regions (four are marked in the figure) arise where the integrating circles overlapped well with circles of hot spots in the power spectrum. Those regions indicate the most likely concentrations of energy in the parent distribution with the given model. To understand the transform further, see Appendix A.

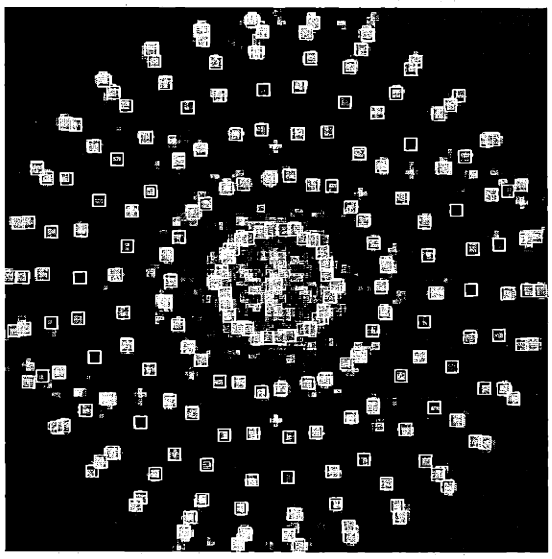

Figure 3-38: Markers overlayed on the spectrum from Figure 3-36 show the locations of "hot spots" predicted by the refined estimate of the beam k-vectors.

Hill climbing method refines beam location estimates

Centroids of the bright regions in the Hough image were extracted and used as an initial guess at the beam locations. A hill climbing algorithm (a Nelder-Mead type simplex search method) optimized the weighted sum of power spectrum values sampled by the auto correlation of the beam positions to produce a final estimate of the beam positions with sub-pixel resolution. The weights emphasized regions of the power spectrum where the hot spots were well isolated.[5]

Figure 3-38 shows the power spectrum from Figure 3-36 with markers showing the auto correlation of the final estimate of beam locations. The good agreement of the 211 marker positions with the hot spot positions supports the theory of many-beam illumination developed in Chapter 2.

[5]Power spectra contained seven rings of hot spots centered around DC, as in Figure 3-36. Hot spots in the second, third, fourth, and fifth rings from DC were weighted by a factor of three relative to the remaining hot spots for the hill climbing metric due to their better isolation from each other. This simple improvement was empirically found to yield favorable convergence speed and agreement of estimated source locations with the data.

Beam amplitudes and phases estimated from reference image

Because the beam geometry was never changed, the process of locating the beam positions in reference image Fourier transforms was done just once. Those locations were used with all subsequent reference images to estimate the amplitudes and phases of the source beams.

Discrete Fourier transform values of a given reference image at the spatial frequencies corresponding to marker positions in Figure 3-38 were estimated by 2D bicubic interpolation from an FFT of the image (windowed by a 2D raised cosine, to suppress edge artifacts). DC is disregarded, and only half of the remaining samples are needed because of the symmetry in the transform of a real function. The resulting vector of 105 complex numbers was used to estimate the 15 amplitudes and 14 relative phases of the 15 laser beams. Appendix B explains the theory behind the iterative algorithm for this data reduction.

Linear phase control with no cross-talk observed

The 15×15 array of subplots in Figure 3-39 shows linear phase control and freedom from crosstalk of the M4 probe illumination system. Each subplot contains 400 points, one for each reference image analyzed in a single measurement. The commanded phases in this measurement were selected to be uniformly randomly distributed from $-\pi$ to π. A given subplot is located in a column, i, and a row, j, corresponding to the measured phase of the i'th beam and the commanded phase of the j'th AOM RF channel that determined the coordinates of each data point for each reference image in each plot.

It is apparent from the plot that for each commanded phase channel, only one beam responds, and it responds linearly. The white-noise appearance of the remaining response phases for the commanded phase channel testifies to a freedom from cross-talk from one command channel to an unaddressed beam. The spread in each correlated plot (the width of the lines) corresponds to the relative power in the responding beam during the measurement and is interpreted as measurement error

Commanded RF channel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Responding beam	7	6	8	5	4	3	2	9	1	15	14	13	10	12	11

Table 3.3: The permutation of commanded RF channels to responding beam was set by the arrangement of mirrors, and easily visualized in Figure 3-39.

and not phase error in the actual beam. Commanded beam number 1 was chosen arbitrarily as the phase reference beam.

The permutation of commanded to measured phases, set by the arrangement of mirrors and read directly off of Figure 3-39 is summarized in Table 3.3. The responding beam indices correspond to a clockwise progression of beam locations (in the coordinate frame of the reference images) starting with beam 1 at three o'clock.

Phase offsets

The phase offsets evident in the correlated plots of Figure 3-39 were stable for hours, but changed from day to day because of thermal drifts and other slow effects. Having demonstrated (with plots like Figure 3-39) linear phase control, the only reference information needed anymore was a list of 14 numbers: the phase offsets.

In M4, for the first time, the ability to control the phases precisely exceeded the ability to measure any single set of phases accurately. Therefore, the control sequence and a sequence of phase measurements were used together. Accurate knowledge of the offsets was refined from phase estimates extracted from hundreds of frames of reference images and then used with knowledge of the phase commands to correct the phase estimates to their most likely values. The procedure is equivalent to replacing the fuzzy lines in Figure 3-39 with best-fit lines.

Point of light

Fresh phase offsets in hand, the phases of all of the beams could be commanded to equal at a location in the reference camera. This condition yielded positive interference of all 15 beams at the chosen location, resulting in a point of light of high intensity surrounded by relative darkness.

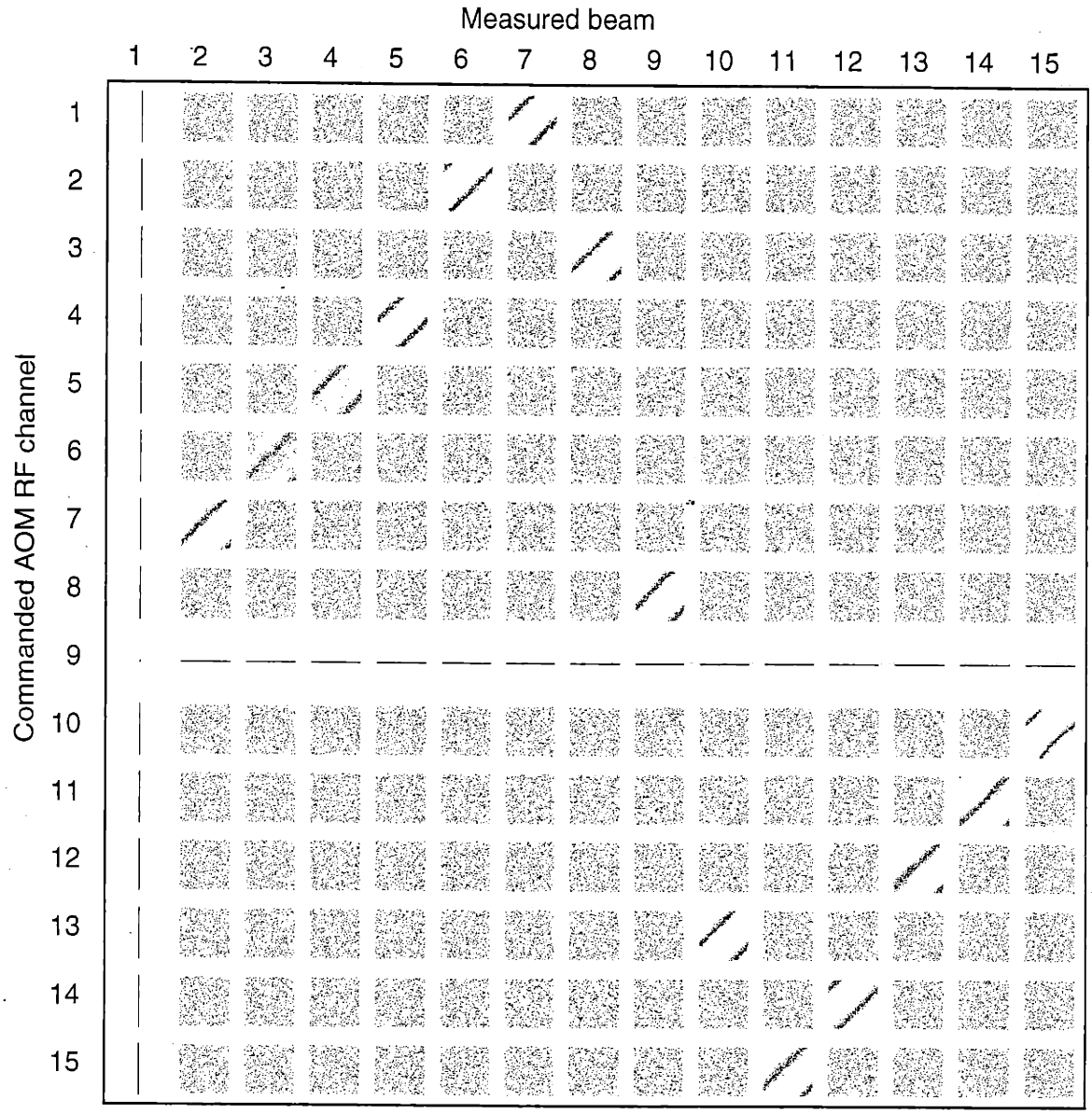

Figure 3-39: This 15×15 array of subplots shows linear phase control and freedom from crosstalk of the M4 probe illumination system. Each row of plots corresponds to a single phase command sequence, and each column corresponds to a single measured phase response. Horizontal (commanded phase) and vertical (measured phase) axes in each subplot range over 2π of phase.

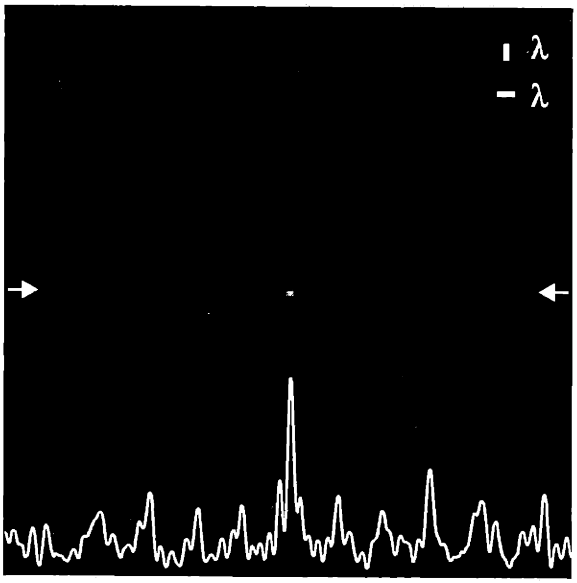

Figure 3-40: Reference image and cross-sectional intensity profile with the beam phases set to zero in the center of the image.

Figure 3-40 shows a reference image with the phases set to zero in the center. The point has a full-width-half-maximum diameter of $\lambda/2.5$ (274 nm) and was projected from a stand-off distance of greater than 30 mm.

Raster scanning was accomplished by commanding the phases to equal at each of an array of locations in a time sequence.

Point is actually a needle

In theory, if the ring were fabricated perfectly, the intensity would not vary out of the plane of the reference interferogram and the point would extend throughout the entire range extent of the region of overlap. It would be a diffraction-limited needle. In practice, variation of source elevation angles and non-complementary wavefront abberations of the source beams would limit the height of the needle.

The out-of-plane evolution of the probe pattern was observed in the real time video imagery from the reference camera by adjusting the height of the reference lens assembly with forceful pressure applied to the assembly mount by hand. Qualitative agreement with the needle theory was thus established.

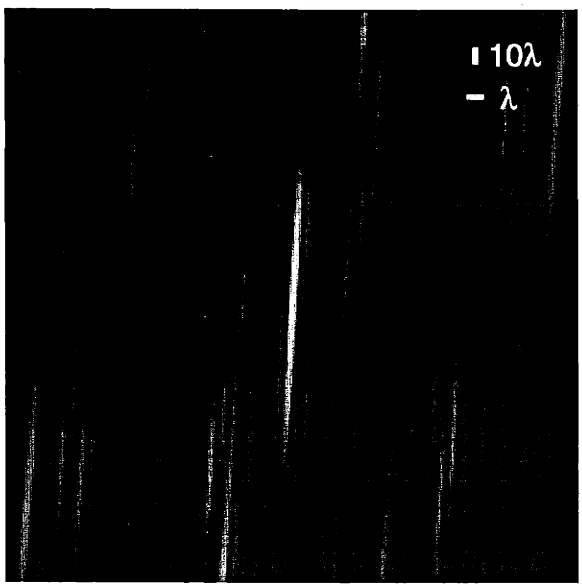

Figure 3-41: Depth intensity profile extrapolated from Figure 3-40 with a simple model.

Our measurement of the most probable source locations from 2D interferograms (see above) enabled the prediction of the most probable evolution of the intensity pattern outside the plane of the reference image. The calculation sums the electric fields of the beams with the measured amplitudes, phases, and source locations, assuming only the polarizations and beam divergences from the design. Figure 3-41 shows the extrapolation from the data in Figure 3-40. Variations in elevation angle were modeled, but wavefront abberations were not, and the extrapolation predicts the point stayed intact beyond the range over which the wavefront approximations used are expected to be valid. Considering only the trustworthy range, the extrapolation suggests the sub-wavelength point remained substantially intact for more than ten wavelengths in depth.

Point stabilized with feedback

The scaling of the probe illumination system to extremely short wavelengths and to application areas characterized by adverse conditions could require active feedback stabilization of the probe patterns. We experimented with a simple under-relaxed

control approach to prove the concept. With a loop-time of just less than 5 seconds, reference images were captured and analyzed to extract the phases of the beams. Deviations from the design phases (all to equal in the center of the reference image) were multiplied by an under-relaxation constant (chosen to be 0.9 after a small number of poorer choices) and added to the previous command phase to generate the next command phase.

With feedback enabled, the point did not exceed 5 % of its minimum diameter over a 24 hour period despite visible tremors of objects in the laboratory caused by the ongoing violent demolition of MIT Building 20 next door.

3.4.5 Conclusion: Synthetic Aperture Microscopy bears fruit

Long term stability justifies measurement procedure

The acquisition time for a sequence of 400 probe patterns was just under one minute. Phase offsets estimated from a sequence of such acquisitions were found to drift slowly, but were essentially unchanging over a two-hour timescale with no feedback and at night (with the Building 20 demolition crew safely elsewhere). Therefore, the measurement procedure adopted for M4 (namely, (1) to collect reference frames with a given probe pattern sequence, (2) to measure the target response to that sequence, and (3) to collect the reference frames again) was justified.

AOM strategy a big win

The AOM control strategy greatly simplified the illumination system, sped it up, and improved imaging accuracy and repeatability compared with previous laboratory methods. Because all of the beams were modulated using a single crystal, this technology was far less expensive to implement than other obvious solid-state approaches. Finally, by shifting from physically uncertain mechanical means to highly repeatable solid-state electrical means of control, the strategy enabled amplitude and phase precalibration. Thanks to this capability and the observed long-term mechanical stability, future microscope generations may be simplified, mechanically and optically,

by shifting from two cameras to one camera and a "reference target."

Fixed assembly successful

While M1, M2, and M3 contained arrays of position adjustments in each final apparatus, M4 only contained position adjustments to accommodate differences in targets. Sufficient aim-point stability over many months was observed (so far), hopefully justifying not only the architectural and material choices but also the lack of arrays of position adjustments.

Greater automation in assembly desired

A numerically controlled milling machine greatly simplified the assembly and alignment of M4. An opto-mechanical design and assembly technology not requiring human alignment of arrays of elements would be a certain improvement. MEMS technology might provide part of the solution.

Needle of light a powerful tool

The ability to create and electronically position an isolated needle of light is a powerful one. Stabilized and raster-scanned demonstrations of this ability inspired hope in several new technology concepts currently under investigation.

Active feedback could help scale the technology

The illumination system's tolerance for vibration and other environmental disturbances is a critical factor in the application of the technology to much shorter wavelengths and to certain application areas. A successful demonstration of one concept to immunize the probe illuminator from these disturbances elicited the expectation that synthetic aperture microscopy ideas could scale to these regimes.

Chapter 4

Conclusions

You have just learned about a new microscope technology with the potential to impact a wide range of application areas. It breaks the hundred-year-old rules of lensed microscopy by **replacing precision physical artifacts with computation.**

The technology is a many-beam extension of the previously reported two-beam SAM concept.[1] It eliminates the acquisition bottleneck of two-beam SAM by replacing the previous mechanically-limited architecture with a fixed-geometry implementation that relies on two forms of parallelism working together. The first form is the simultaneous interrogation of on the order of n^2 spatial Fourier transform coefficients of a target region illuminated by n coherent, phase-modulated beams. The second form of parallelism is the segmentation of the target into many regions with a low-resolution imaging lens and an array-type detector (CCD camera). Combining these two strategies yielded **a degree of parallelism of one hundred million** in one of the prototypes.

Much of the relevant physical theory of the new technology was worked out with an emphasis on critical engineering and implementation issues. Verification of the physical principles of operation was secured under the simplest possible experimental conditions with M1, the first generation of prototype apparatus. Therefore, with confidence in the physical basis of the image-sensing technology, the next step was

[1] Three separate US patents (currently pending) resulted directly from this thesis project.

to tackle the principal engineering challenge in implementing many-beam SAM: the construction of a controllable many-beam illuminator. Three subsequent prototype generations (M2, M3, and M4) addressed the entire spectrum of engineering issues in the construction of such an illuminator, culminating (with M4) with the invention and demonstration of a novel solid-state technique for generating and electronically controlling many laser beams. This invention enabled an optics tour de force: the synthetic aperture projection of an actively stabilized, isolated needle of light—*a deeply useful pattern that is impossible to achieve using traditional ideas in optics.* The final prototype **substantially solved the engineering problem of controllable many-beam illumination using visible light.** Because of the diffractive and reflective nature of the illuminator, and because of the (demonstrated) potential for active feedback control to immunize the illuminator against vibration and other environmental disturbances, the illuminator solution holds promise to scale to much shorter wavelengths.

The concept as a whole **decouples traditionally entangled imaging parameters.** Specifically, it renders working distance, field of view, and depth of field independent of resolution, and of each other. Furthermore, it enables images with a pixel-count orders of magnitude greater than the pixel-count of a given generation of array-type image detectors (such as CCD's). It accomplishes this by subdividing each physical pixel in the given detector into an array of smaller *virtual pixels.* Table 4.1 lists demonstrated values for these parameters for the final prototype and for the closest comparable state-of-the-art microscope objective lens along with functional relationships for each technology.

This thesis project constitutes the beginning of a research activity that will involve **future students and projects.** One important future task is the reduction to practice of the image synthesis concept disclosed herein. A second future task is a careful feasibility analysis for adapting the concept to electron or X ray wavelengths. A third future task is adapting the many-beam illuminator to a commercial research-grade optical microscope. This retro-fit will allow the convenient use of a single, very long working distance, low magnification objective lens with a large depth of field *for all*

	M4 SAM prototype (N.A. $\approx$ 0.92)	SAM in theory	Conventional 0.95 N.A. lens (63$\times$ Zeiss)	Lens technology
Working Distance	>30,000 μm	Set by low-N.A. lens	120 μm	Practical limit: decreases with N.A. according to a power law
Depth of field	>20λ	Set by low-N.A. lens	0.35λ	Theoretical limit: $\left\lceil \frac{\lambda\sqrt{n_r^2-(N.A.)^2}}{(N.A.)^2} \right\rceil$
Field of view	>5,000 μm available, 560 μm used	Set by low-N.A. lens or size of region of beam overlap	317 μm available	Practical limit: decreases with N.A. according to a power law
Resolution	>200 sub-pixels per pixel	Goes as the square of the number of beams	1 pixel per pixel	Always 1
Composition	No precision optics		Precision optics	

Table 4.1: A comparison of the final SAM prototype with state-of-the-art microscope lens performance and of SAM and lenses in general.

imaging scales—replacing the familiar turret of lenses with a controllable many-beam illuminator and a software "zoom tool" on the computer console attached to the pre-existing CCD camera. Operating in UV fluorescence, the instrument could provide resolution of a fraction of the UV-wavelength. It would not need a reference camera and could be occasionally phase-calibrated with a known reference structure such as a fluorescent polystyrene micro-bead. Such a microscope would constitute **a powerful and affordable tool to researchers in many fields.** It would advance our quantitative understanding of complex, intricate physical systems that can not easily be studied or visualized using existing technologies. It would simplify visualization of micro-structures, particularly those with three-dimensional extent. Moreover, it would maintain the look, feel, and functionality of a conventional optical microscope, thus enhancing its likelihood of widespread acceptance by the scientific community.

Appendix A

A matched Hough transform for locating beam sources

Equation 2.9 indicates that a pair of beam with k-vectors, $\vec{k}_l$ and $\vec{k}_m$, will synthesize a cosine intensity component at the vector spatial frequency $\vec{k}_{lm} \equiv (\vec{k}_l - \vec{k}_m)$. Figure A-1 shows a pair of k-vectors and their difference vector.

In Fourier space, a sample point arises at the vector spatial frequency $\vec{k}_{lm}$. In terms of the plot, this point would be located at the arrowhead of the $\vec{k}_{lm}$ vector with the other end at the origin. Figure A-2 shows the construction with the example vectors from Figure A-1 in the (F_x, F_y) plane.

Assuming the k-vectors, $\vec{k}_l$ and $\vec{k}_m$, are of equal magnitude and elevation angle,

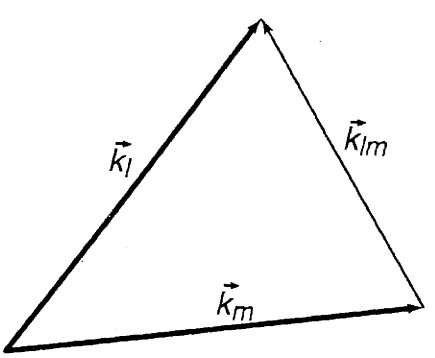

Figure A-1: A pair of k-vectors and their difference vector.

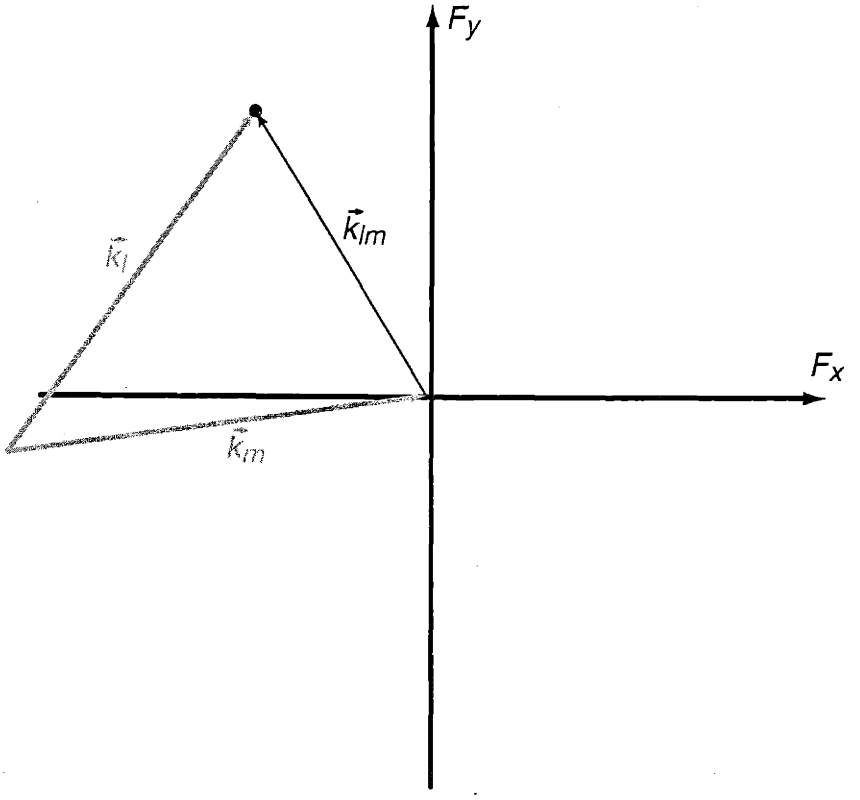

Figure A-2: The difference vector as a sample point in Fourier space.

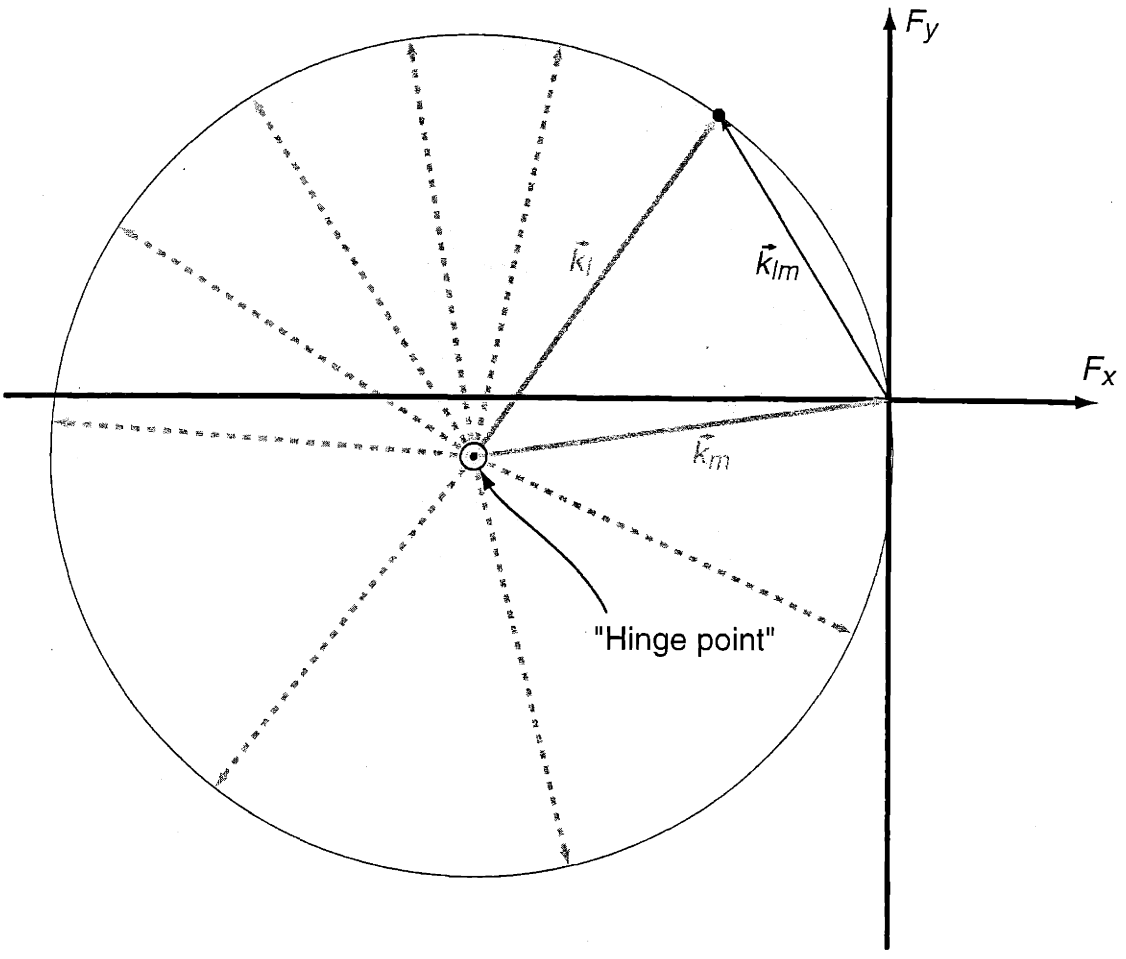

Figure A-3: All difference vectors from an ordered pair with the same elevation angle and magnitude fall in a circle in Fourier space.

the non-arrow side of those vectors in the figure could be above the plane of the paper, for example. The difference vector, $\vec{k}_{lm}$, and the Fourier sample point, however, would be in the plane of the paper nonetheless.

To understand at an intuitive level why a path-integral over a circle was used in the matched Hough transform, we can view the non-arrow side of $\vec{k}_m$ in Figure A-2 as a mechanical hinge, fixed in space. The partnering beam, $\vec{k}_l$, attaches to the hinge with its non-arrow side and points at the Fourier sample point in the (F_x, F_y) plane. Figure A-3 shows several other example candidates for $\vec{k}_l$ as dotted vectors. All of the difference vectors land in a circle, centered around the hinge at $-\vec{k}_m$.

We integrate around all circles of this radius in the power spectrum of the reference image. This power spectrum comprises the sum of $2n$ of these circles of spectral peaks (one at $+\vec{k}_i$ and one at $-\vec{k}_i$ for each of the n i's), and the integrals centered on those $2n$ circles are expected to yield a higher value than the integrals centered in other locations in the spectrum.

Therefore, locating those points in a power spectrum that are the centers of circles of this size tags likely positions of the source vectors.

Appendix B

Source phase recovery from reference images

This Appendix describes an algorithm for converting reference image information to a list of beam phases (and amplitudes). Berthold Horn developed the algorithm specifically for the problem at hand.[1]

B.1 Method for estimating the phase differences

There are n beams whose complex amplitude is represented by (x_i, y_i) for $i = 1 \ldots n$. The amplitude of the i'th beam is $\sqrt{x_i^2 + y_i^2}$. The phase of the beam is $\text{atan2}(y_i, x_i)$, measured at the origin (at the center of the reference image, typically) with respect to an arbitrary reference beam.

We cannot directly observe the complex amplitudes of the n beams; we have to infer them from the $n(n-1)/2$ peak responses in the transform of the reference image generated by their interference. Suppose that we locate the transform domain peak corresponding to interference between beams i and j and find that the value of the transform there is $(c_{i,j}, s_{i,j})$ (in terms of the cosine and sine coefficients, respectively). Since we are dealing with the transform of a real function (namely the reference

[1]The Appendix is self-contained in terms of notation. In other words, the notation has local scope—it is locally defined and locally valid.

image) we know that $c_{j,i} = c_{i,j}$ and $s_{j,i} = -s_{i,j}$.

The magnitude $(c_{i,j}, s_{i,j})$ should be the product of the magnitudes of the two beams, and the phase should be the phase difference of the two beams. If there were no noise or measurement error we would have

$$c_{i,j} = x_i x_j + y_i y_j \tag{B.1}$$

$$s_{i,j} = x_i y_j - x_j y_i \tag{B.2}$$

Our problem then is to estimate (x_i, y_i) for $i = 1 \ldots n$ from the observed $(c_{i,j}, s_{i,j})$ for $i = 1 \ldots n$ and $j = i + 1 \ldots n$. The problem is overdetermined since $n(n-1)/2$ is greater than n.

We can try to minimize the sum of squares of errors

$$E = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j}((c_{i,j} - x_i x_j - y_i y_j)^2 + (s_{i,j} - x_i y_j + x_j y_i)^2) \tag{B.3}$$

where $w_{i,j}$ is a weight that indicates the reliability of our measurement of the complex magnitude of the transform domain peak corresponding to interference between beams i and j. We let $w_{j,i} = w_{i,j}$. In the above sum we actually count each measurement twice, since the complex amplitude for (j, i) is just the complement of that for (i, j). The duplication is just for convenience of the following derivation.

Differentiating E with respect to x_k, we get

$$\begin{aligned} \frac{dE}{dx_k} = \; & 2\sum_{j=1}^{n} w_{k,j}(c_{k,j} - x_k x_j - y_k y_j)(-x_j) + \\ & 2\sum_{i=1}^{n} w_{i,k}(c_{i,k} - x_i x_k - y_i y_k)(-x_i) + \\ & 2\sum_{j=1}^{n} w_{k,j}(s_{k,j} - x_k y_j + x_j y_k)(-y_j) + \\ & 2\sum_{i=1}^{n} w_{i,k}(s_{i,k} - x_i y_k + x_k y_i)(y_i). \end{aligned} \tag{B.4}$$

the first and second terms have the same value, as do the third and fourth terms.

After some further simplification the above expression becomes

$$\frac{dE}{dx_k} = 4\sum_{i=1}^{n} w_{i,k}(-c_{i,k}x_i + s_{i,k}y_i + (x_i^2 + y_i^2)x_k) \tag{B.5}$$

Setting this equal to zero (to identify likely minima) yields:

$$x_k \sum_{i=1}^{n} w_{i,k}(x_i^2 + y_i^2) = \sum_{i=1}^{n} w_{i,k}(c_{i,k}x_i - s_{i,k}y_i) \tag{B.6}$$

Differentiating E with respect to y_k, we get, after some simplification

$$\frac{dE}{dy_k} = 4\sum_{i=1}^{n} w_{i,k}(-c_{i,k}y_i - s_{i,k}x_i + (x_i^2 + y_i^2)y_k) \tag{B.7}$$

Setting this equal to zero yields:

$$y_k \sum_{i=1}^{n} w_{i,k}(x_i^2 + y_i^2) = \sum_{i=1}^{n} w_{i,k}(c_{i,k}y_i + s_{i,k}x_i) \tag{B.8}$$

We need to solve this set of non-linear equations for the unknown (x_i, y_i) for $i = 1, \ldots n$, given the values of $(c_{i,j}, s_{i,j})$ from the transform of the reference pattern.

The problem would be easier if knew the value of

$$\sum_{i=1}^{n} w_{i,k}(x_i^2 + y_i^2) \tag{B.9}$$

This suggest an iterative scheme:

$$x_k^{(n+1)} = A\sum_{i=1}^{n} w_{i,k}(c_{i,k}x_i^{(n)} - s_{i,k}y_i^{(n)}) \tag{B.10}$$

$$y_k^{(n+1)} = A\sum_{i=1}^{n} w_{i,k}(c_{i,k}y_i^{(n)} + s_{i,k}x_i^{(n)}) \tag{B.11}$$

for some positive A. The problem is that we don't know A. However, since A is a single overall scale factor on the estimated amplitudes, we can perhaps optimize it in

a second step. We need to find the value of A that minimizes

$$\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(c_{i,j} - A(x_i x_j + y_i y_j))^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(s_{i,j} - A(x_i y_j - x_j y_i))^2 \qquad \text{(B.12)}$$

for fixed (x_i, y_i). Differentiating with respect to A and setting the result equal to zero leads to:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(c_{i,j} - A(x_i x_j + y_i y_j))^2 (x_i x_j + y_i y_j) \; + \\ &\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(s_{i,j} - A(x_i y_j - x_j y_i))^2 (x_i y_j - x_j y_i) \; = \; 0 \end{aligned} \qquad \text{(B.13)}$$

Or

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(c_{i,j}(x_i x_j + y_i y_j) + s_{i,j}(x_i y_j - x_j y_i)) \; = \\ &\qquad A\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}((x_i x_j + y_i y_j)^2 + (x_i y_j - x_j y_i)^2) \end{aligned} \qquad \text{(B.14)}$$

The sum on the right simplifies to

$$\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(x_i^2 + y_i^2)(x_j^2 + y_j^2). \qquad \text{(B.15)}$$

So finally

$$A = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(c_{i,j}(x_i x_j + y_i y_j) + s_{i,j}(x_i y_j - x_j y_i))}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(x_i^2 + y_i^2)(x_j^2 + y_j^2)} \qquad \text{(B.16)}$$

We adjust the estimates of (x_i, y_i) by multiplying them by $\sqrt{A}$ before going on to the next iteration. If we are very far from the solution, A as computed above may turn out to be negative. In that case we use its absolute value to be able to continue the iterative adjustment.

Bibliography

[1] G. Bekefi and A.H. Barrett. *Electromagnetic Vibrations, Waves, and Radiation.* MIT Press, Cambridge, Massachusetts, 1990.

[2] T.M. Crawford. Novel fringe scanning/fourier transform method of synthetic imaging, 1993. Report No. EGG-M-93041.

[3] J.C. Dainty. *Laser speckle and related phenomena.* Springer-Verlag, New York, 1984.

[4] R.C. Dixon. *Spread Spectrum Techniques.* IEEE Press, New York, 1976.

[5] D.L. Feldhun. Design and implementation of a high-speed solid-state acousto-optic interference pattern projector for three-dimensional imaging. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, Spring 1999.

[6] R.D. Guenther. *Modern Optics.* John Wiley & Sons, New York, 1990.

[7] A.C. Hardy and F.H. Perrin. *The Principles of Optics.* McGraw-Hill, New York, 1932.

[8] E.S. Hung. *Positioning, Control, and Dynamics of Electrostatic Actuators for Use in Optical and RF Systems.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, August 1998.

[9] R.A. Hutchin. Microscope for producing high resolution images without precision optics. U.S. Patent No. 4,584,484, 1986.

[10] S. Inoué. *Video Microscopy.* Plenum Press, 1986.

[11] R. Juskaitis and T. Wilson. Synthetic aperture microscopy. *10'th International conference on 3D image processing in microscopy*, 4(2):232–233, 1997.

[12] C. Kittel and H. Kroemer. *Thermal Physics.* W.H. Freeman and Company, New York, 1980.

[13] G.R. Hallerman L.G. Shirley. Applications of tunable lasers to laser radar and 3d imaging. Lincoln Laboratory Technical Report 1025, February 1996.

[14] F. Liu, S. Nie, C. Tao, and S. Bian. Theory and computer simulation of the synthetic aperture microscope. *Proceedings of the SPIE*, 2890:58–62, 1996.

[15] L. Ljung. *System Identification.* Prentice-Hall, New Jersey, 1987.

[16] J. Marron and K. Schroeder. Three-dimensional lensless imaging using laser frequency diversity. *Applied Optics*, 35(5):255–262, January 1992.

[17] M.S. Mermelstein. A large-scale three-dimensional imaging system based on laser speckle. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, May 1995.

[18] M.S. Osofsky. An electron transparent unipotential foil lens for the correction of spherical aberration in electron optical systems. Bachelor's thesis, Massachusetts Institute of Technology, Cambridge, MA, 1979.

[19] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vettering. *Numerical Recipes in C.* Cambridge University Press, Cambridge, second edition, 1992.

[20] C.V. Raman and N.S.N. Nath. The diffraction of light by high-frequency sound waves: Part 2. *Proc. Indian Sci.*, 2:413–420, 1935.

[21] L.G. Shirley, E.D. Ariel, G.R. Hallerman, H.C. Payson, and J.R. Vivilecchia. Advanced techniques for target discrimination using laser speckle. *The Lincoln Laboratory Journal*, 5:367–440, 1992.

[22] L.G. Shirley and G.R. Hallerman. Nonconventional 3d imaging using wavelength-dependent speckle. *The Lincoln Laboratory Journal*, 9:153–186, 1996.

[23] W.M. Siebert. *Circuits, Signals, and Systems.* MIT Press, Cambridge, Massachusetts, 1986.

[24] G. Strang. *Linear Algebra and its applications.* Harcourt Brace Jovanovich, Inc., 3rd edition, 1988.

[25] O. Takeshi and U. Norihiro. Synthetic aperture microscopy using external optical feedback to a laser diode. *Kogaku (Japanese Journal of Optics)*, 25(6):323–328, 1996.

[26] T. Turpin, L. Gesell, J. Lapides, and C. Price. Theory of the synthetic aperture microscope. *Proceedings of the SPIE*, 2566:230–240, July 1995.

[27] T.M. Turpin. Image synthesis using time sequential holography. U.S. Patent No. 5,751,243, May 1998.

[28] T.M. Turpin, P. Woodford, C. Price, J. Lafuse, S. Evanko, and L. Phillips. Imsyn: optoelectronic product for image synthesis and correlation. *Proceedings of the SPIE*, 3073:178–184, 1997.

[29] R. Wiesendanger. *Scanning probe microscopy and spectroscopy: methods and applications.* Cambridge University Press, Cambridge [England], 1994.