MEMS Spatial Light Modulator for Holographic Displays

Elroy L Pearson
B.S.E.C.E, USU, 1999
A.S.E.E., Ricks College, 1996

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
In Partial Fulfillment of the Requirements for the Degree of

Master of Science

All Rights Reserved

Program in Media Arts and Sciences

Stephen A. Benton
E. Rudge ('48) and Nancy Allen Professor of
Media Arts and Sciences, MIT Media Laboratory

Dr. Andrew B. Lippman
Chair, Departmental Committee on Graduate Students
Program in Media Arts and Sciences
MEMS Spatial Light Modulator for Holographic Displays

Elroy L. Pearson

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in Partial Fulfillment of the Requirements for the Degree of Master of Science

Abstract

This thesis presents new approaches to building holographic displays. The approaches use a diffractive spatial light modulator (SLM) built using microelectromechanical system (MEMS) technology. Several related MEMS SLMs are reviewed. In particular, the capabilities of Silicon Light Machines' Grating Light Valve (GLV™) and the Polychromator by Honeywell, Sandia, and MIT are examined because they are both diffractive SLMs. A holographic display design is presented that is based on a diffractive SLM that combines the spatial resolution of the GLV with the addressability of the Polychromator. The holographic display system design and the MEMS fabrication process are described. Although the MEMS device was not fully fabricated, much of the process was developed. The successful process steps and suggestions for process improvements are laid out. Finally an improved MEMS SLM process and a scalable MEMS holographic display are proposed.
MEMS Spatial Light Modulator for Holographic Displays

Elroy L. Pearson

Stephen A. Benton
E. Rudge ('48) and Nancy Allen Professor of Media Arts and Sciences, MIT Media Laboratory

Scott Manalis
Professor of Media Arts and Sciences
MIT Media Laboratory

Pierre St-Hilaire
Research Scientist
Templex Corp.
Acknowledgements

More people have helped make this thesis possible than can adequately be addressed here. Another thesis could probably be written on that topic alone. I would like to thank all of those who contributed in one way or another to this effort.

In particular I would like to thank Shelly for her unconditional support and love. I am not whole without you. My children Kaili, Louis, and LeRoyce have been the spice of my life. There is nothing like having them all run up and jump into my arms for a hug.

Stephen Benton has been an inspiration and guide. He has supported even my craziest ideas and corrected my errors.

Scott Manalis has introduced me to a whole new world of microfabrication. What fun!

Thanks to the MEMS fabrication team from the fall 2000 class. Diana Young, Yael Macguire, James White, Matt Reynolds, and Guy Rasmuth combined some incredible skills to design a great project.

Many thanks go to Aditya Prabhakar, Aaron Weber, and Wendy Plesniak – the other grad students in the group. Aditya endured many sleepless nights of thesis writing with me and helped keep me sane. Aaron is an amazing optical engineer and more often than not a great cook. Wendy is a talented engineer, scientist, and artist – a rare find. Her timely advice has helped me keep everything in perspective. Thanks to Ollie for proofreading my thesis and making me formulate my thoughts more clearly.

Thanks to Steve Smith and Thomas Nwodoh. Steve has terrific practical experience to go with his scientific understanding. I would like to thank him for letting me build his shed with him. Thanks for the hammer. It has been fun to compare stories about the kids with Tom and Steve.

I have much appreciation for the many talented UROPs that have contributed to group. I have always been able to count on Daniel, Rob, Julie, Altay, and the others to lighten my spirits and ask really good questions. Thanks to Daniel for being a good friend that I can talk to for much too long.
Thanks to the previous SPIers for inspiration and advice. Pierre St-Hilaire continues to amaze me with his technical prowess and scientific understanding. Bill and Julie Parker been friends since we met them two years ago. I look forward to working with you.

Thanks to the faculty and staff at MTL. In particular Paul Tierney has been a great support.

Thanks to Richard Rallison for giving me my first real introduction to holography and research. I learned laser safety from the guy that talked me into jumping from a cherry picker over a pool attached to a bungee cord.

I would like to thank Mom and my brothers and sisters for helping me get to where I am today. Mom taught me to really think about things and not take anything for granted.

Finally I would like to thank Dad for teaching me the value of integrity and hard work and the importance of family. You'll never see me graduate from college – on this Earth. Thanks for watching from the other side.
## Contents

**ABSTRACT** .................................................................................................................3

**ACKNOWLEDGEMENTS** .............................................................................................9

1. **INTRODUCTION** .................................................................................................13
   1.1 Motivation ...........................................................................................................13
      1.1.1 History of Image Production ...........................................................................13
      1.1.2 The Human Visual System ...........................................................................14
      1.1.3 Three-dimensional Displays .........................................................................15
      1.1.4 The Acousto-optic Spatial Light Modulator .................................................16
      1.1.5 MEMS-based SLMs .....................................................................................18
   1.2 Content of Thesis .................................................................................................18

2. **SPATIAL LIGHT MODULATORS** .........................................................................21
   2.1 Structure and Operation of Three MEMS SLMs .................................................21
      2.1.1 Texas Instruments’ DMD™ ............................................................................21
      2.1.2 Diffractive SLMs .........................................................................................22
      2.1.3 Silicon Light Machines’ GLV .........................................................................22
      2.1.3 The Polychromator .......................................................................................24
   2.2 SLM Comparisons ...............................................................................................25
      2.2.1 Diffraction Efficiency ....................................................................................25
      2.2.2 Bandwidth .....................................................................................................25
   2.3 Feasibility of Holographic Displays Based on These MEMS-SLMs .......................27

3. **SYSTEM DESIGN** .................................................................................................29
   3.1 SLM Design .........................................................................................................29
      3.1.1 SLM Mechanical and Electrical Design .........................................................29
      3.1.2 SLM Pinout ...................................................................................................35
      3.1.3 SLM Microfabrication Process Design .........................................................37
      3.1.3 SLM Optical Analysis ...................................................................................43
   3.2 Optical Architecture Design ................................................................................46
   3.3 Summary .............................................................................................................49

4. **PROCESSING** ......................................................................................................51
   4.1 Process Overview ................................................................................................51
      4.1.1 Chapter Contents ........................................................................................51
   4.2 Thin Film Application ........................................................................................51
      4.2.1 Thermal Oxide Growth .................................................................................51
      4.2.2 Polysilicon Deposition ................................................................................52
      4.2.3 Nitride Deposition .......................................................................................53
   4.3 Lithography ..........................................................................................................54
      4.3.1 Photoresist ..................................................................................................54
4.3.2 Stepper ........................................................................................................ 54
4.4 PLASMA ETCHING ......................................................................................... 55
  4.4.1 Nitride Etch .............................................................................................. 55
  4.4.2 Oxide Etch ................................................................................................. 55
  4.4.3 Polysilicon Etch .......................................................................................... 55
4.5.1 KOH Etch of Polysilicon ........................................................................... 56
  4.5.2 Stepper Lithography for Lift-off ................................................................. 57
  4.5.3 Metal Deposition for Electrodes and Traces ............................................... 57
  4.5.4 Metal Lift-off for Traces ............................................................................. 57
4.6 RELEASE ETCH ............................................................................................... 58
  4.6.1 Release Etch ................................................................................................ 58
4.7 PACKAGING .................................................................................................... 58
  4.7.1 Packaging .................................................................................................. 58
4.7 CONCLUSION .................................................................................................. 59

5. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK .............. 61
  5.1 REVIEW ........................................................................................................... 61
    5.1.1 Thesis Review ........................................................................................... 61
  5.2 SUGGESTIONS FOR FUTURE WORK ......................................................... 61
    5.2.1 SLM Fabrication ...................................................................................... 61
  5.3 CONCLUSIONS ............................................................................................... 67
    5.3.1 Conclusions .............................................................................................. 67

REFERENCES ....................................................................................................... 68
1. Introduction

1.1 Motivation

1.1.1 History of Image Production

For millennia people have imagined illusions so realistic that the viewer would not be able to determine whether the observed scene is real or fabricated. At first, it was likely assumed that the most probable way to create such an illusion would be through magic. The first realizations of convincing images were probably achieved through art forms such as sculpture and painting. Artists such as Michelangelo created amazing likenesses of people and other subjects using these media. Michelangelo's ability to recreate accurate details went virtually unmatched for centuries.

Figure 1.1 Michelangelo's David (left) and Sistine Chapel fresco (detail of "The Creation of Adam" at right) demonstrate the amazing realism that Michelangelo achieved through sculpture and painting.

During the last few hundred years, technology advanced enough to enable great breakthroughs in recording and displaying our visual world. Still cameras enabled us to record moments in incredible detail. Movies have let us capture and replay the complexity of movement. These media changed the way we communicated visually and empowered those who wielded them masterfully.

The closer our visual communications come to matching the abilities of the human visual system, the more our potential for communicating
expands. The paintings and pictures are able to show two-dimensional images. Sculptures are able to replicate the three-dimensionality of a scene. Movies are able to make two-dimensional images move. However, none of these media alone are capable of reproducing all of the visual cues that most humans are able to see. Conspicuously lacking in the tool set of visual communications is the ability to generate three-dimensional moving pictures.

1.1.2 The Human Visual System

People have invented various approaches to creating three-dimensional moving images. The great majority of these efforts have used techniques to send different images to each eye to take advantage of stereopsis.

Stereopsis is the ability to discern depth through the slightly differing images that each eye sees because they are physically located a short distance from each other. The brain interprets the differences in these images along with other depth cues to determine the depth of objects in a scene. Using color or polarization filtering glasses or head-mounted displays can produce stereo vision. However, stereopsis is only one depth cue among many that are important to three-dimensional imaging. Most humans actually combine many visual cues to see depth in scenes. These cues include stereopsis, occlusion, ocular accommodation, shading, convergence, and motion parallax. Motion parallax comes from seeing new perspectives of a scene as the viewer’s head (and thus eyes) move relative to the scene. Ocular accommodation is the ability of the eyes to focus at different depths. Accommodation is a monocular cue, or one that can be seen by using a single eye. This can be observed by looking through a window at a distant object. As the eye focuses on the object the window will seem mostly transparent (even if it is fairly dirty). Then, when the eye focuses on the surface of the window, the distant object will blur out and mostly disappear from the viewer’s vision. Convergence is the effect of one’s eyes turning slightly inward to view an object at close range. Occlusion is the disappearance of a distant object as it passes behind an opaque near object. Most modern displays (as of the writing of this thesis) do not attempt to provide binocular cues, but are able to convey the concept of depth fairly well through the use of pictorial and motion cues (not from the viewer moving but from moving cameras) cues. A good three-dimensional display will make use of not only these cues, but binocular cues as well.
1.1.3 Three-dimensional Displays

Most attempts at three-dimensional displays have used primarily stereopsis, convergence, and two-dimensional depth cues to display depth at the expense of accommodation. The result is that many people get headaches after viewing these displays for even short periods of time. This problem can be kept to a minimum if measures are taken not to make convergence-accommodation conflicts too great.

Holographic displays promise to deliver all of the depth cues needed to make truly three-dimensional moving images. Notable holographic displays that have been built to date include several holographic video displays built in the MIT Media Lab by the Spatial Imaging Group, a display built in Japan, and a holographic display built by DERA in England.

The MIT\textsuperscript{1,2} and Japanese displays are based on a Scophony-style scanned spatial light modulator (SLM) approach. A Scophony system using a scanned SLM was used in a large screen TV in 1939\textsuperscript{3,4}. The general layout of a Scophony-style holographic display is shown in figure 1.2. In these systems, the light is modulated as it passes through a crystal SLM. The crystal modulates the light as sound waves are induced by a piezoelectric transducer to travel through the crystal. As these waves pass through the crystal, the crystalline structure is microscopically stretched and compressed to create areas of increased and decreased index of refraction. Light passes through the areas of increased index

![Figure 1.2 Scophony-style Scanned SLM Holographic Display](image-url)
slower than in other areas. This induces a phase delay into the light in that area.

The DERA system uses a combination of SLMs to make holographic images. Light is projected from an electrically addressable spatial light modulator (EASLM) onto a larger optically addressable spatial light modulator (OASLM). Though the two styles of display are very different, SLMs are at the heart of both. This thesis focuses primarily on the SLM used in the Scophony-style holographic display.

1.1.4 The Acousto-optic Spatial Light Modulator

The capabilities of the SLM determine in large part the ability of the display to show three-dimensional images. The speed at which the SLM can be updated is important. Its modulation efficiency and the maximum spatial frequency are also key attributes. The smaller the area that an SLM can address the larger an angle it can diffract light. If the modulation area is small enough, significant diffraction effects can be observed. The resulting angle of diffraction can be determined directly from the grating equation:

\[
\lambda \cdot f_m = \sin(\theta_{\text{in}}) - \sin(\theta_{\text{out}})
\]

where \(\lambda\) is the wavelength of light measured in meters, \(f_m\) is the maximum spatial frequency measured in meters\(^{-1}\), \(\theta_{\text{in}}\) is the angle the light is incident to the modulating medium with respect its normal, and \(\theta_{\text{out}}\) is the angle that light leaves the medium. From this relation it becomes obvious that for an SLM to diffract light at higher angles, it must be able to address higher spatial frequencies.

An AOM must be driven with higher frequency acoustic waves if higher spatial frequencies are to be achieved. The current MIT holographic video display uses acoustic frequencies up to 100 MHz to produce spatial frequencies up to 81 line pairs per mm. (See the Calculation of spatial frequency and fringe wavelength in AOMs below.) This enables a maximum deflection of light of 2.9 degrees. This is a much smaller angle than is practical to use.

A holographic display needs to have a viewing angle at least wide enough to permit a viewer to fit both eyes into the view zone at a reasonable viewing distance. The view zone should be wide enough to
allow the user to also move his or her head side to side. For a viewing distance of 600 mm, given that the average human interpupillary distance is 65 mm, the minimum view angle needs to be 6.2 degrees. The holographic image can be greatly improved by increasing this angle to 30 degrees. This widening of the view angle allows the viewer to move side to side and observe motion parallax. The deflection angle of the SLM must be magnified by a series of lenses for this to happen.

Each component that is added to the display increases the complexity of the design. The design must be carefully balanced so as not to unnecessarily limit the bandwidth by a poorly chosen part. In the style of display that MIT uses, the speed of the horizontal scan mirror is dictated by the speed of sound in the crystal and the distance between the scanner and the AOM. The sound velocity and the drive frequency of the acoustic transducer determine the maximum spatial frequency and, hence, the maximum diffraction angle the AOM can produce. This angle in turn sets one half of the demagnification specification with the view zone requirement setting the other half. In fact, setting the specifications for a holographic display is a bit like smoothing wallpaper - when you fix one area a bubble pops up in another. All of the components are intimately linked.

### Calculation of spatial frequency and fringe wavelength in AOMs

The maximum spatial frequency produced by the AOM can be determined from the speed of sound in the crystal and the temporal frequency at which it is driven. The sampling theorem dictates that the sampling frequency must be twice the spatial frequency.

\[
f_m = \frac{f_s}{2} = \frac{F}{2v} = \frac{81 \text{ lines}}{\text{mm}}
\]

\[
\Lambda = \frac{1}{f} = \frac{0.00617 \text{ mm}}{\text{line}} = 6.17 \text{ \mu m/line}
\]
Greater holographic display performance may be achieved by investigating new components. Acousto-optics have improved in the years since the second generation MIT display was built. Likewise, faster and smaller scanning mirrors are now available. Much stands to be gained from improvements in computation. Currently, a reconfigurable computing board called Holo-Chidi is being built in the Media Lab. A dedicated team of engineers and scientists could probably take advantage of these advances to build a much improved AOM-based Scophony-style holographic display.

1.1.5. MEMS-based SLMs
During the last few years many new spatial light modulators have been introduced to the market. Many of these have been MEMS-based devices. MEMS, or "microelectromechanical systems," are devices fabricated on a microscopic scale usually on a silicon wafer. These systems often perform electrical, mechanical, chemical, heat, and/or optical transducing. Because their size is about the same order of magnitude as the wavelength of visible light they have begun to be explored for their optical transducing potential. Probably the most well-known MEMS SLM is the Texas Instruments Digital Micromirror Device (DMD). The DMD uses a two dimensional array of tilting mirrors to reflect light into and away from an output image. Each mirror is 16 microns across. A MEMS SLM by Silicon Light Machines shrinks the size of its optical transducers to one micron and uses diffraction rather than plane tilting to direct light into an output image. This device is based on the Deformable Grating Optical Modulator (DGM) created at Stanford by Solgaard, et al.6

1.2 Content of Thesis
Much of this thesis is based on fabricating a modified version of the DGM. The details of these SLMs are explained in chapter 2.

Chapter 3 covers the design of a MEMS-based holographic display. This work began as part of a MEMS fabrication class. The team of students in the class (with the consultation of many experts in the various fields) designed and began implementation of a MEMS SLM, an array of MEMS scanning mirrors, a holographic display optical layout, and drive electronics. After the class ended, the author continued to develop the MEMS SLM process and it became his thesis project.
Chapter 4 describes the implementation of the MEMS SLM. This chapter discusses the changes that were made in the original process and the reasons for these changes. Also suggestions are made to successfully fabricate a MEMS SLM. Most of the steps in the MEMS SLM process were developed successfully, however, a few steps remain to be completed before the device will work.

Chapter 5 gives the conclusions of this thesis and suggestions for future work.
2. Spatial Light Modulators

MEMS spatial light modulators (SLMs) have taken many forms and been used for many applications. MEMS SLMs have been used as optical switches in optical networks, in two-dimensional image projectors, and in holographic printers.

Chapter 2 introduces three MEMS SLMs: Texas Instruments' DMD™, Silicon Light Machines’ GLV™, and Honeywell/Sandia/MIT's Polychromator. The physical structure and operating mechanisms of each device are presented. Then they are compared quantitatively with the AOM used in the Mark II Holographic Video display.

2.1 Structure and Operation of Three MEMS SLMs

2.1.1 Texas Instruments' DMD™

Probably the most well known MEMS SLM is the Digital Mirror Device (DMD)™ by Texas Instruments. It has recently gained fame for replacing film projectors in some theaters viewings of movies such as "Star Wars: The Phantom Menace" and "Toy Story II." The DMD is a two-dimensional array of mirrors (see figure 2.1). The mirrors are tilted...
such that each mirror reflects light into or away from an output image. Because each pixel can be either 'on' or 'off', pixel brightness is determined by the ratio of 'on' time to 'off' time. Because human eyes can only detect changing light levels at a certain speed (around 30 changes per second) any changes faster than that tend to get averaged together. During that averaging time, the longer a mirror is switched to the 'on' position the brighter the corresponding pixel will appear. To create a dimmer pixel the mirror turns on fewer times during the same period, thus delivering less energy to the eye.

The DMD is able to do optical pulse width modulation because its mirrors switch on and off many times faster than the averaging time of the eye. TI claims a 15 µs mechanical switching time and a 2 µs optical switching time.

The optical efficiency of the DMD is very good. The brightness of a fully 'on' pixel is limited by the reflectivity of the surface aluminum and the mirror's fill factor.

The DMD is a very complex MEMS device. The substrate upon which the mirrors are fabricated actually contains CMOS SRAM cells. These cells connect to address electrodes which electrostatically tilt the mirrors. The mirrors rotate on flexure hinges which support a yoke. The mirrors are connected to the yokes through support posts. TI has spent much effort and money on developing this complex MEMS SLM.

2.1.2 Diffractive SLMS
Other MEMS SLMs operate on different optical principals. Two notable MEMS SLMs use diffraction to modulate light. These SLMs are Silicon Light Machines’ Grating Light Valve (GLV™) and the Polychromator, a MEMS-based optical correlation spectrometer system developed jointly by Honeywell, Sandia National Laboratories, and MIT.

2.1.3 Silicon Light Machines’ GLV
The GLV consists of an array of parallel beams that can be pulled down. This height difference creates a phase delay in the light that is reflected off the lower beam. When two adjacent beams are adjusted to heights that differ by a quarter wavelength of light they strongly diffract light. See figure 2.2. The GLV uses groups of beams to form pixels. These pixels are usually made groups of even numbered beams where every
Alternate ribbons are deflected creating a square-well diffraction grating.

Ribbons are longer than the width of the pixel creating 100% diffraction region in center.

Figure 2.3 Silicon Light Machines' GLV is constructed from a raised set of beams. Pixels are turned on or off by deflecting alternating beams or ribbons down toward the substrate electrode.

other beam moves. This means that beams can only be controlled only in groups and not individually.

The GLV beams are actuated electrostatically. The beams, or ribbons as they are sometimes referred to, are normally suspended above the bottom electrode. This rest state is due to spring restoring forces in the beams. The ribbons are pulled down when a voltage is applied between the electrode that runs across the top of the beam and bottom electrode (figure 2.3).

The size and geometry of the beams determines the required actuation voltage, switching time, and diffraction angle of the device. Many different sizes of GLV are possible. A GLV with beams 5 microns wide by 20 microns long was reported to have a switching time of 20 nanoseconds. More typically the ribbons are driven at 250 KHz. A typical beam size is 3 microns wide by 100 microns long. One pixel may be made up of up to six ribbons.

Some of the proposed applications for the GLV include high-definition projection displays, handheld displays, and automobile heads-up displays. The GLV may be of limited use in a holographic display because adjacent beams are not individually controllable.
2.1.3 The Polychromator

The Polychromator is a diffractive spatial light modulator that was created from a collaboration between MIT, Honeywell, and Sandia National Labs. It enjoys a few advantages over the GLV while performing less well in other areas. The Polychromator allows individual beams to be controlled. It also has a much larger optically active diffracting area than the GLV due to a complex double tiered beam structure that it employs (see figure 2.4). While the GLV uses only about 20 microns of a 100 micron long beam to diffract light the Polychromator is able to diffract light over the entire length of its 1 cm long beams! This beam structure also helps create a more controllable bending. Normally, when a doubly-clamped beam is pulled one third of the distance between the beam and the bottom electrode it will reach a an unstable state in which the beam will spontaneously collapse to the bottom of the substrate without any added charge. This effect is known as ‘pull-in’. If a beam is pulled from close to its ends rather than near the center, it can be controlled over much more of the gap. This is known as leveraged bending.

The Polychromator was designed for use in the infrared region to create artificial spectra for identifying the chemical compositions of specimens. Because the Polychromator was designed to manipulate infrared light, its spatial frequency is quite low. Each beam is 10 microns wide with a 2 micron gap in between. The Polychromatic is actually not intended to be used as an image-producing SLM.

![Figure 2.4](image)

**Figure 2.4** The Polychromator is built with a double tiered beam structure that allows the beams stay optically flat along their entire length. The bottom beams are bent by electrostatic forces near the beam ends and not in the middle. This leveraged bending helps prevent ‘pull-in’ and enables the beams to travel nearly the entire gap.
2.2 SLM Comparisons

2.2.1 Diffraction Efficiency

Each of the aforementioned diffractive SLMs is subject to similar principles of diffraction efficiency. The fill-factor or ratio of the area of the beams to the area of the gaps between the beams plays a large role in this. Publications by Silicon Light Machines refer to the diffraction of the structures as square-well diffraction. They give the intensity of the 1st order diffraction lobes as:

\[ I_{1st} = I_{max} \sin^2 \left( \frac{2\pi d}{\lambda} \right) \]

where \( I_{max} \) is the maximum 1st order diffracted intensity (at \( d = \lambda/4 \)), \( d \) is the grating depth, and \( \lambda \) is the wavelength of the incident light.

They also claim about 70% optical efficiency for the GLV using 0.6 μm microfabrication design rules. This is a product of diffraction efficiency (81%), aluminum reflectivity (91%) and ribbon/gap efficiency (95%). The ribbon/gap efficiency seems dubious, however, because a .6 micron gap would require a 12 micron wide beam to achieve a 95% fill factor. Beams that wide were not reported.

2.2.2 Bandwidth

An SLM must be able to process an enormous amount of information to be considered for a holographic display. Just to match the processing capabilities of the Mark II Holographic Video display, an SLM or array of SLMs must temporally modulate over 37 million areas of light 30 times a second. This equates to a temporal bandwidth of over 1.1 billion samples per second. The spatial bandwidth product is also impressive with over 260 thousand samples being squeezed into a single holo-line just 150 mm wide. That is an effective spatial frequency of 1747 lines per millimeter or sample sizes of .57 microns wide.

To compare the bandwidths of various SLMs one needs to consider many characteristics. The switching speed of the modulator, the number of modulating channels, and ability of these channels to be tiled to modulate in parallel are important considerations. Silicon Light Machines claim a 20 ns switching time with certain GLVs. This is about 1000 times faster than the Texas Instruments DMD. This speed translates into added shades of gray in 2D images. GLVs are currently being built in parallel channels of 1080 pixels. The DMD is commonly
built with 1280*1024 parallel modulating channels. The GLV may have an inherent speed advantage in that it can be operated in analog mode so that it need not sacrifice switching cycles to obtain grayscale samples.

It is interesting to note how these MEMS SLMs stack up against acousto-optic modulators. AOMs are capable of modulating light at very high bandwidths. The AOM used in the Mark II Holovideo system has an eight-bit modulation bandwidth of 55 MHz or 55*2*10^6 samples/second. (The eight-bit restriction is actually a limitation of the framebuffer system and not the AOM itself.) Thus the 18 channel system has a time bandwidth product of 1980 Msamples/sec. However the scanned nature of the AOM Holographic display system limits the effective bandwidth to 1130 Msamples/sec due to turn-around times during which no modulation is performed.

The MEMS SLMs are made up of more channels of lower bandwidth. The 1080 pixel GLV has an eight-bit modulation bandwidth of 270 Msamples/sec when the beams are driven in analog mode at one-eighth their natural frequency. If the beams are driven at one-half their natural frequency of 2 MHz then that figure jumps to 1080 Msamples/sec. If a GLV could be constructed that could utilize a 20 nanosec switching time in analog mode the bandwidth would be 54,000 Msamples/sec. The TI DMD is capable of achieving eight-bit modulation at 341 Msamples/sec. Because the GLV is a scanned modulator it also takes a performance hit during the retrace period. These results are compiled in table 2.1.

<table>
<thead>
<tr>
<th>Modulator</th>
<th>Number of Channels</th>
<th>Bandwidth per Channel (8-bit samples/sec)</th>
<th>Total Bandwidth at 8-bits/sample (Msamples/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shear mode TeO₂ AOM in Mark II</td>
<td>18</td>
<td>55*10^6</td>
<td>1980</td>
</tr>
<tr>
<td>Silicon Light Machines’ GLV (driven at 1/8 natural frequency)</td>
<td>1080</td>
<td>250*10^3</td>
<td>270</td>
</tr>
<tr>
<td>Silicon Light Machine’s GLV (driven at 1/2 natural frequency)</td>
<td>1080</td>
<td>1*10^6</td>
<td>1080</td>
</tr>
<tr>
<td>Silicon Light Machine’s GLV (driven at 50 MHz)</td>
<td>1080</td>
<td>50*10^6</td>
<td>54000</td>
</tr>
<tr>
<td>Texas Instruments’ DMD</td>
<td>1280*1024</td>
<td>260</td>
<td>341</td>
</tr>
</tbody>
</table>

Table 2.1 Comparison of time bandwidth products of several SLMs.

Each of these modulators has the ability to improve given additional development. The AOM approach could be improved in a number of ways. The channels could be placed closer together on the crystal allowing more total channels. This would allow more horizontal lines
and also give room to trade-off scan time to get longer lines. However, this approach would require the horizontal scan mirrors to sweep through a larger angle unless a longer focal length lens is used after the AOM. This angle is already constrained by the diameter of the output lens and the abilities of the scanning galvonometers. This approach would work if a stronger output lens could be used and faster scanners found. Higher bandwidth acousto-optic materials are also available if an appropriate balance of system components can be managed.

2.3 Feasibility of Holographic Displays based on these MEMS-SLMs

In the previous section the MEMS SLMs have been compared to the TeO₂ AOM used in the Holovideo system without considering whether they are suitable for use in a practical holographic display. Given the nature of the DMD pixels it may be difficult to build a truly holographic display based on that device. Even if the pixels could be made to diffract light in a useful manner, they would need to be demagnified significantly to provide a sufficient viewing angle.

The GLV is not much more suitable for driving a holographic display than the DMD. It does, however, have higher spatial frequencies and time bandwidth products. But this is of limited use without individually controllable beams.

The Polychromator was conspicuously dropped from the discussion of temporal bandwidth because no figures relating to switching speed could be found. Also, because of pin-out issues only 128 of the 1024 total beams can be individually controlled. Thus between the unknown temporal bandwidth, limited individual control, and limited spatial bandwidth the Polychromator appears to be an unlikely candidate for holographic displays also.

A MEMS SLM that combines the spatial resolution and switching speed of the GLV and the addressability of the Polychromator could be a potential contender in the quest to build the next generation of holographic display. The next chapter describes a prototype holographic display designed to use just such a device. In chapter 5 improvements on extending the MEMS-based holographic display to a larger scale are proposed.
3. System Design

The design of the MEMS diffractive spatial light modulator began in a Media Arts and Sciences MEMS project class co-taught by Professors Scott Manalis and Marty Schmidt. The intent of the class was to introduce the members to MEMS design and fabrication by creating a MEMS-based holographic display. The display would consist of a MEMS spatial light modulator, a MEMS array of scanning mirrors, imaging optics, and drive electronics. The project team was loosely assigned as follows: James White and Guy Rasmuth took primary responsibility for the MEMS scanning mirror array, Matt Reynolds designed the drive electronics, Yael Maguire, Diana Young, and Elroy Pearson took care of the MEMS SLM, and Elroy Pearson designed the optical layout.

The design process began with a system level design. It was decided that the system would follow the basic optical principles used in the MIT Mark I Holographic Video system with a few exceptions. The light modulator would be constructed following the plans of the Grating Light Modulator built by Solgaard, Sandejas, and Bloom⁶. Previously an AOM had been used as the SLM. The spinning polygonal mirror would be replaced by an array of scanning mirrors as was done in the Mark II display. These scanning mirrors would also be MEMS devices. Custom electronics would be built to drive the SLM, the mirror array, and a vertical scanning mirror.

3.1 SLM Design

3.1.1 SLM Mechanical and Electrical Design

The capabilities of the SLM and scanning mirror array determined many of the properties of the optical subsystem and the drive electronics and vice versa. In the Solgaard paper that we followed, several sizes of beam arrays were constructed. We decided to make the beam width just slightly larger than the feature size allowed by the lithography equipment available in the Microsystems Technology Lab (MTL) at MIT. Since the achievable feature size was estimated to be around one micron we decided to make the beams 1.5 microns wide. The length of the beams was chosen to be 100 microns long. Solgaard measured the natural frequency of beams of this size to be 1.5 MHz. Because tension is the
major restoring force in the beams, they can be approximately modeled as vibrating strings. Thus the resonant frequency can be found from

\[ F_r = \frac{1}{2L} \sqrt{\frac{\sigma_o}{\rho}} \]  

where \( L \) is the length of the beam, \( \sigma_o \) is the tensile stress, and \( \rho \) is the density. According to Kovacs\(^{11}\) low-stress nitride has a density of 2.9 to 3.1 g/cm\(^3\) and a stress of 1000 MPa. He notes, however, that low-stress nitride can have stresses approximately equal to 0 Pa. Commercial thin film deposition services list nitride stresses between 110 and 160 MPa. A quick calculation with 1000 Mpa stress shows a resonant frequency of 2.9 MHz. A stress of 100 MPa would yield a resonant frequency of .928 MHz. One lesson here is that processes can vary widely and numerical analysis of a MEMS part is only valid if the figures used have been well characterized. A value in a table in a book cannot blindly be used to give accurate predictions of performance unless the process available to the user has well known results. However, numerical analysis can help one understand generally the effects of a process change on the finished product even when specific data are not yet available. So a prudent MEMS builder will not assume a process will work until experience has shown her it will.

Unfortunately for the design of the MEMS SLM we had to make the best guess we could on several material values. Because we didn't have solid data on the stress level of nitride available through the MTL, we had to pick a value and hope that we could adjust the processing or drive voltage parameters until we achieved the desired results. The beams needed to be deflected at least \( \frac{1}{4} \) wavelength of light (158 nm for 633 nm light). The thickness of the beams was set at 158 nm to follow the Solgaard design. Next the initial gap distance \( g_o \) between the beam and bottom electrode had to be set. Here we varied from the Solgaard design so that the beams could travel the full 158 nm without experiencing "pull-in". Pull-in is a state in which the electro-mechanical forces become unstable and the structure spontaneously collapses to the bottom of the substrate without increased voltage. The regions of travel can be seen in figure 3.1. According to Senturia\(^{12}\) the pull-in distance for a beam that is approximately one-half as thick as the gap \( g_o \) is one-third \( g_o \). For a beam that is thinner the stable travel distance increases slightly because of "strain-stiffening". This is because as the beam bends the
Figure 3.1 Pull-in of a doubly clamped beam occurs when the beam is deflected past a certain point and spontaneously collapses to the bottom of the gap. The stable deflection region is usually one-third the gap when the beam is approximately one-half as thick as the gap. In this figure the vertical deflection is greatly exaggerated.

increased strain in the beam increases the tensile stress. The strain-stiffening effect is reduced, however, if a large initial stress exists.

A case study in the Senturia book examines the mechanical characteristics of fixed-fixed fully clamped beam structures with respect to the GLV. The beams’ pull-in voltage \( V_{pl} \) can be found with the equation

\[
V_{pl} = \sqrt{\frac{8k'_{eff} g_o^3}{27}} \tag{3.2}
\]

where \( g_o \) is the original gap and \( k'_{eff} \) is a modified spring constant. \( k'_{eff} \) is given by

\[
k'_{eff} = \frac{P_c}{w_{max}} = \frac{8N}{WL^2} \left[ \frac{1}{1 + 2 \frac{1 - \cosh u}{u \sinh u}} \right] \tag{3.3}
\]

where \( u = k_c L/2 \), \( N \) is the tension in the beam, \( W \) is the width of the beam, \( H \) is the thickness of the beam, and \( L \) is the length of the beam. \( P_c \) is the
effective pressure on the beam and \( w_{\text{max}} \) is the deflection of the beam at the point of greatest travel – the center. The previous variables are found through the following equations:

\[
N = \sigma_o W H
\]

3.4

\[
P_e = \frac{\varepsilon V^2}{2(g_o - w_{\text{max}})^2}
\]

3.5

\[
w_{\text{max}} = \frac{qL}{4N} \left( \frac{L}{2} - 2 \frac{\cosh(k_o L/2) - 1}{k_o \sinh(k_o L/2)} \right)
\]

3.6

\[
k_o = \frac{\sqrt{12N}}{EWH^3}
\]

3.7

where \( E \) is the Young's modulus, and \( V \) is the voltage between the beam and the electrode that forms the bottom of the gap. \( q \) is the distributed load on the beam or force per unit length so it is just the effective pressure \( P_e \) times the beam width \( W \). These equations use a parallel plate approximation that can be improved somewhat using more sophisticated non-linear numerical simulations. The pull-in voltage in particular is off by about 20%, but is close enough to give one an idea of the effects of changes in geometry and drive voltage on the system.

Using these equations we can see what region of operation our choices of beam structure and geometry places us in. Because the residual stress \( \sigma_o \) is a parameter that has to be determined empirically it is important to know what effect it has on the final performance. Figure 3.2 shows the change in pull-in voltage versus residual stress for a nitride beam 100 microns long, 1.5 microns wide, and .15 microns thick with an initial gap \( g_o \) of 1 micron.
Figure 3.2 These graphs show the pull-in voltage versus the residual stress in the beam. Because residual stress in nitride can vary greatly the left figure shows the pull-in voltage for stresses between 0 and 2 MPa and the right figure shows the pull-in voltage for stresses between 1 and 200 MPa.

From these simulations it is apparent that the range of pull-in voltages for low-stress nitride beams could vary from about 3 to 29 volts. Typical nitride stress values available commercially through mems-exchange.org are between 110 and 160 MPa. This leads to pull-in voltages between about 17 to 26 volts.

The operational voltage, or voltage required to deflect the beams $\lambda/4$, determines the range required of the digital-to-analog converters that will drive the beams. It can be found by solving equations 3.5 and 3.6 for V.
For this design the pull-in voltage exceeds the operational voltage as seen in figure 3.3. A good rule of thumb when designing beams is that the original gap $g_o$ should be 3 times the deflection voltage because the stable travel range is usually one-third the initial gap.

\[ V = \sqrt{\frac{8N(g_o - w_{\text{max}})^2 w_{\text{max}}}{\varepsilon WL \left( \frac{L}{2} - 2 \frac{\cosh(k_o L/2)}{k_o \sinh(k_o L/2)} - 1 \right)}} \quad 3.8 \]

Figure 3.3 The operational voltage plotted with the pull-in voltage. Because the operational voltage does not exceed the pull-in voltage, the beams of this design will be able to travel the full $\lambda/4$ distance.

Once $\sigma_o$ has been determined the voltage/deflection curve can be plotted from equation 3.9.
The first root of 3.9 gives the deflection distance where $z=w_{\text{max}}$. The voltage/deflection curve for a beam with the dimensions as used previously and residual stress of 110 MPa can be seen in figure 3.4.

![Figure 3.4 The voltage/deflection curve for a beam with residual stress of 110 MPa.](image)

### 3.1.2 SLM Pinout

Once the operational voltage level is determined a digital-to-analog converter (DAC) can be chosen. The DAC must have a high enough voltage rating to be able to deflect the beams through the entire $\lambda/4$ distance. Also important is the resolution of the DAC. To control the
beams to 256 discrete levels a DAC with 8 bits of precision is required. However, more bits are likely to be required in order to make equal deflections throughout the stable travel distance because equal voltage changes at distances close to the full gap and the pull-in level will move the beam unequal amounts.

Another important consideration in choosing the DAC and designing the SLM is how to connect the DAC to the beam. To make the chip as high-bandwidth as possible as many beams should be constructed on the chip as is feasible. For our design we decided to construct 256 adjacent beams. This corresponded nicely to the number of pins on a commercially available dip package and to the area available to form bond pads around the periphery of the die. The bond pad arrangement around the edge of the die is shown in figure 3.5.

Figure 3.5 Image of die layout with bond pads on the periphery.
Each bond pad is connected to a metal trace that conducts electricity to a single beam in the center of the die (figure 3.6). Traces to the electrodes on adjacent beams approach the beams from opposite side of the array. This allows a smaller tolerance for the traces during patterning. Each beam is 1.5 microns wide by 100 microns long with gaps 1.5 microns wide between pairs of beams. The traces are 3 microns wide with spaces 3 microns wide between them.

3.1.3 SLM Microfabrication Process Design

The physical size and composition of the beams were chosen to utilize equipment available in the MIT Microsystems Technology Laboratory (MTL). The processing of this devices involves applying various layers of materials to a single crystal silicon wafer and then patterning those layers with a variety of lithography techniques and etches. The original design called for nitride beams to be constructed on a silicon dioxide sacrificial layer with gold electrodes, traces, and bond pads. Due to processing difficulties this was changed to a process with nitride beams on a polysilicon sacrificial layer with gold electrodes, traces, and bond pads. The reasons for the switch are explained in chapter 4. This chapter discusses only the polysilicon process.

The process starts with growing 100 nm of oxide on an n-type silicon wafer that is 500 microns thick. Then a layer of polysilicon is deposited
via low-pressure chemical vapor deposition (LPCVD) to a thickness of one micron. A 158 nm thick layer of low-stress silicon nitride is then deposited on top of the polysilicon. A cross-sectional view of these layers is shown in figure 3.7. All of the materials are also deposited to some degree on the backside of the wafer though they are not shown here.

To prepare the wafer for the first mask layer the nitride, polysilicon, and oxide layer are etched from the backside. Thus nothing need touch the frontside of the wafer after it has been patterned.

The first mask layer patterns trenches the length and width of the beams in the center of the die. To prepare the wafer for the etch it is coated in photoresist. Photoresist is a light sensitive material that can be patterned with light and then chemically developed such that the patterned areas wash away. Positive photoresist washes away in the places that have been exposed to light. Negative, or image-reversal, photoresist washes away in the places that haven’t been exposed to light. The remaining material is then baked to form a hardened mask. This mask is resistant to many chemicals, so places that are covered in resist are protected from the etchant while areas without resist are etched. This process uses positive photoresist.

Photoresist is deposited in liquid form onto a spinning wafer. This spreads the resist into a thin, uniform layer on the wafer. The wafer is then baked to pre-harden the resist before exposure. The wafer is then exposed to light that has been spatially modulated in the proper pattern. After exposure the wafer is developed and then post-baked to harden the resist mask. The post-baked photoresist pattern from the first mask layer is shown in figure 3.8.
After the resist is developed and hard-baked the nitride and polysilicon layers are etched in a plasma etcher. The plasma etcher uses etchant chemicals under vacuum in a high powered RF field to etch vertically oriented sidewalls. This is known as anisotropic etching because it favors removing material in one direction over another. Isotropic etching removes material in all directions at the same rate and thus makes dish-shaped etches rather than box shaped etches. The etched wafer in figure 3.9 shows the results of the anisotropic nitride and polysilicon etches.

After the trenches are etched in the nitride and polysilicon layers, the first layer of photoresist is removed. Then a second layer of photoresist
is applied and pre-baked. Into this layer of photoresist is exposed the second mask layer. This is the layer that defines the beam electrodes, leads, and bond pads. This layer is known as a "lift-off metal layer". The photoresist is patterned to stay where the gold is to be removed. Then gold is evaporated onto the photoresist and exposed areas of the wafer. The wafer is then soaked in an acetone bath to dissolve the photoresist and lift off any metal that was deposited on the resist. Figure 3.10 shows the photoresist pattern on the beams. Figure 3.11 shows the metal deposition and then liftoff steps. Before the 50 nm gold layer is deposited a 10 nm layer of Titanium is put down. This helps the gold adhere to the underlying nitride.

Figure 3.10 Wafer after lift-off photoresist has been developed, but before deposition of the metal layer.
Figure 3.10 The top image shows the wafer after metallization. The bottom picture shows the wafer after the wafer has been put in acetone, the photoresist dissolved, and the metal patterned by lift-off.
There are a couple of key assumptions in this liftoff process. The first is that the profile of the liftoff photoresist will be steep enough to make sure that the gold on top of the photoresist is not connected to the gold on the wafer. The second assumption is related. It is that sidewalls of the nitride-polysilicon trenches are straight enough that the gold on the top of the beams does not connect to the gold on the bottom of the trenches. The next chapter will discuss the impact of these design decisions.

The next fabrication step is to thicken up the bond pads. The traces are connected by gold wire bonding to the chip package. For the wire bonder to work a thicker layer of gold is required on the bond pads. A 200 to 300 nm thick layer of gold is needed to promote good adhesion by the gold connection wires. This thicker layer of gold is applied in a second liftoff process. In this process the only holes patterned into the photoresist are holes above the bond pads. Then a thick layer of gold is deposited and removed in an acetone bath as before (figure 3.11).

![Figure 3.11](image)

Figure 3.11 The 100 micron wide bond pads at the edges of the die are thickened up with a second layer of gold to help the gold wires to adhere to the pads better during wirebonding.

Once the second liftoff is completed successfully the beams can be released. The beams are released in a dry Xenon Difluoride (XeF₂) etch.
The XeF₂ etches the polysilicon sacrificial layer from under the nitride beams (figure 3.12).

![Diagram of etching process](image)

**Figure 3.12** The nitride beams are undercut by etching the sacrificial polysilicon layer with a Xenon Difluoride etch.

After the beams have been released, the wafer can be sawed into individual dies. The dies are then mounted into chip packages and wire bonded to the electrical connections. At that point the MEMS SLM is ready to be integrated into the full holographic display.

### 3.1.3 SLM Optical Analysis

The MEMS SLM modulates light through diffraction as does the Silicon Light Machines’ GLV. However, a more complex model of the diffraction is needed for analysis of the optical characteristics of this MEMS SLM because every beam is individually controllable. Also
hologram computation for the completed display will be based on the diffraction model used for analysis.

The MEMS SLM is essentially a one-dimensional dynamic surface relief grating. When the beams are all in the same plane light reflects off the front of the beams as it would off a mirror (except for the extraneous diffraction added by the spaces between the beams). When two adjacent beams are at different heights, the light bouncing off the lower beam is delayed in phase by

$$\phi = \frac{2\pi}{\lambda} 2w_{\text{max}}$$

where $\lambda$ is the wavelength of light. The resulting grating pattern can be modeled as a transmission phase grating (figure 3.13). A basic square-wave phase grating works by retarding light wavefronts as a function of position. Because a phase grating absorbs no light the diffracted efficiency can reach 100 percent when all of the diffracted orders are collected. Up to 40.5 percent of the light can be contained in each of the first orders. The diffraction efficiency $DE_m$ of the orders $m$ can be calculated as
DE_{m=0} = \cos^2 \Delta \phi \hspace{1cm} 3.11

DE_{m=\pm 1} = \left( \frac{2}{\pi} \sin \Delta \phi \right)^2 \hspace{1cm} 3.12

DE_{m=\text{even}} = 0 \hspace{1cm} 3.13

DE_{m=\text{odd}} = \frac{1}{m^2} DE_{+1} \hspace{1cm} 3.14

where $\Delta \phi$ is the phase delay. Of course these equations are for square waves only so a more complete model would take into account the movable beams and the immovable dead-space in between. Far-field diffraction effects can be determined by treating each beam or space as a one-dimensional aperture of width $W$ with a constant phase across its width and then summing the contribution from all of the apertures in the far-field. The light amplitude $A(x)$ from each aperture leaves the SLM plane as a square wave or rect function (figure 3.14). In the far-field the light amplitude from position $x$ is distributed as a sinc function or Fourier transform of the rect function due to diffraction and interference effects. This far-field amplitude at $x'$ which originated from position $x$ in the SLM plane is denoted $A'(x \rightarrow x')$. This distribution is given by

$$A'(x \rightarrow x') = A(x)W \text{sinc} \left( \frac{\pi \theta W}{\lambda} \right) \hspace{1cm} 3.15$$

where the angle between $x$ and $x'$ is $\theta$ which is equal to $\tan((x'-x)/(z+2w_{\text{max}}))$. The phase of the light is given by equation 3.16.

![Figure 3.14 The far field intensity distribution emanating from a rectangular aperture is a sinc function.](image)

**Figure 3.15** Light propagates from the SLM plane on the left to the far-field plane on the right.
\[ \Phi'(x \rightarrow x') = j \sin \left[ \frac{2\pi(z + 2W_{\text{max}})}{\lambda \cos(\theta)} \right] \] 3.16

The complex amplitude contribution \( E'(x \rightarrow x') \) from the aperture at position \( x \) in the SLM plane to the position \( x' \) in the far-field is a combination of equations 3.15 and 3.16.

\[ E'(x \rightarrow x') = A'(x \rightarrow x') + \Phi'(x \rightarrow x') \] 3.17

The total amplitude \( E'_{\text{total}}(x') \) at the point \( x' \) is found by summing the contributions \( E'(x \rightarrow x') \) from every point in the SLM plane \( x \).

\[ E'_{\text{total}}(x') = \sum_{x=-X/2}^{X/2} E'(x \rightarrow x') \] 3.18

This summation would then be convolved with the lower frequency sinc function that results from the combined aperture of the entire beam array.

3.2 Optical Architecture Design

The purpose of the optical architecture is to form the light modulated by the SLM into a three-dimensional holographic image. The hologram is created by increasing the effective aperture of the SLM by tiling its image in the output plane of the display. This is accomplished in much the same way as the MIT Mark II holographic display. A vertical scanning mirror scans the image vertically and a bank of horizontal mirrors scan the image horizontally. Lenses demagnify the image to create a sufficient viewing angle. The optical layout is shown in figure 3.15.

Most of the lens requirements can be reduced to geometric relationships. The laser illuminates the SLM at normal incidence. The highest angle that the SLM can diffract light is found from the grating equation, \( f_d = \sin \theta \). The highest frequency that this grating SLM can synthesize is one half the spatial frequency of beams or 166 lines per millimeter. This leads to the first order being diffracted at 6 degrees. By placing the SLM at the focal length of a lens \( L_1 \) the light will be nearly collimated with the spatial frequency transformed to a spatial offset along the horizontal axis. The focal length of \( L_1 \) dictates the minimum width of the vertical scanning mirror and the minimum combined width of the horizontal
scanning mirrors. Because the width of the scanning mirrors was set to 2 cm a lens of focal length 18.85 cm is needed.

The total size of the array of scanning mirrors is 2 cm wide by 3 cm tall. They are designed to scan 5 mechanical degrees so they deflect 10 optical degrees (figure 3.17). The final imaging lens needs to have an
aperture wide enough to capture all of the light deflected by the mirrors (figure 3.18).

The vertical optical system scans the image line through a cylindrical lens L2 which focuses the line at the center of the horizontal scanning array. The unfolded vertical system in figure 3.19 shows that the light leaving the horizontal scanning area will be collimated parallel to the optical axis. The height of the image is controlled by the scan angle of the vertical scanning mirror.

Figure 3.18 Top view of the last half of the holographic display optical layout. The horizontal scanning mirror array is placed at the focal length of lens L3 which images the SLM at the vertical diffusing plane.
3.3 Summary
A complete holographic display was designed in the MEMS project class. Designs were submitted for the drive electronics, horizontal scanning mirror array, the optical architecture, and the MEMS SLM. Much progress was made in developing the fabrication techniques required to build the MEMS devices.
4. Processing

4.1 Process Overview

4.1.1 Chapter Contents

Chapter 4 discusses the practical considerations of process development for the MEMS SLM. Both the oxide sacrificial layer process and the polysilicon sacrificial layer process are presented side-by-side. This allows a concise comparison of characteristics of the processes. During the development of this process most of the steps were performed repeatedly until they became well characterized. However, the polysilicon process' gold electrode application step still remains to be characterized. Most of this chapter describes how to perform process steps that are proven to work. At the end of the chapter the focus shifts to suggestions on perfecting the remaining steps.

The majority of the steps carried out in this chapter are performed in class 100 or class 10 clean rooms. The class number of the clean room specifies how many one micron particles per cubic foot of air are present. Clean conditions reduce the number of imperfections in the process by keeping dust particles off the wafers. A dust particle can wreak havoc with a microfabrication project because it can interfere with proper image patterning and etching.

Chapter 4 is divided into six sections. These sections cover thin film growth and deposition, photolithography, plasma etching, metal liftoff, release etching, and packaging.

4.2 Thin Film Application

4.2.1 Thermal Oxide Growth

The first step in both the oxide sacrificial layer process and polysilicon sacrificial layer process, hereinafter referred to as the oxide process and the poly process respectively, is to grow a layer of thermal oxide. The oxide process calls for a micron thick layer of silicon dioxide to act as the sacrificial layer. The poly process requires only 100 nm of silicon dioxide for an etch stop for the polysilicon layer.
Thermal oxide is grown in a thermal oxidation furnace. To prevent tube contamination the wafers are first cleaned using an RCA clean\textsuperscript{14} to remove any organics, native oxide, and ionic or heavy metal contaminants. The RCA clean was first developed by RCA Laboratories in the 1970's and has now become an industry standard\textsuperscript{15}. The RCA clean has three major steps. The organics are removed with a 5:1:1 \( \text{H}_2\text{O}:\text{H}_2\text{O}_2:\text{NH}_4\text{OH} \) solution. Then the native oxide is removed in a diluted 50:1 \( \text{H}_2\text{O}:\text{HF} \) etch. Finally ionic and heavy metal atomic contaminants are removed using a solution of 6:1:1 \( \text{H}_2\text{O}:\text{H}_2\text{O}_2:\text{HCL} \). When the wafers are properly cleaned they can be placed in the oxidation furnace.

The oxidation furnace envelops the wafers in a high temperature atmosphere of oxygen. The reaction model is

\[
\text{Si}(s) + \text{O}_2(g) \rightarrow \text{SiO}_2(s)
\]

A layer of oxide quickly grows at the silicon air interface. A thin layer of oxide will even grow in room temperature conditions. After the initial thin layer is grown, however, the oxide layer can only grow when oxygen or silicon atoms pass through the existing oxide layer. The temperature is increased to promote diffusion. Oxygen diffuses much faster than silicon so the oxide grows at the oxide silicon interface.

Once the oxide layer is grown, its thickness is measured with an ellipsometer. It is important to remember that oxide also grows on the backside of the wafer because this layer must be removed before the substrate can be electrically connected to ground.

4.2.2 Polysilicon Deposition

After the oxide is grown the poly process calls for a one micron layer of polysilicon to be deposited. Polysilicon is silicon composed of many small crystals rather one large crystal. The polysilicon is deposited in a low pressure chemical vapor deposition (LPCVD) furnace. The furnace is pumped down to around 150 mTorr after being flushed with nitrogen and the temperature is raised to around 580 degrees C. The overall reaction is

\[
\text{SiH}_4(g) \rightarrow \text{Si}(s) + \text{H}_2(g)
\]

Deposited polysilicon is not as smooth as single crystal silicon. Subsequently any layers deposited on top of it will also be rough. The roughness is readily seen in the scanning electron micrographs in figure
4.1, but can even be detected under an optical microscope. Future processes should include a polish after the polysilicon deposition to smooth out the surface so that the beams will be flat reflectors.

**Figure 4.1** The left image is a scanning electron micrograph of the undercut beams in the oxide process without gold electrodes. The right image shows the beams in the poly process with gold electrodes and traces deposited. Note that the surface of the nitride beams is much smoother for the oxide process than for the poly process.

### 4.2.3 Nitride Deposition

The beam structural material is silicon nitride. 158 nm of low-stress silicon nitride is deposited in both processes. Nitride application happens in an LPCVD furnace also. Some possible nitride reactions\(^{16}\) are

\[
\text{3SiH}_4(g) + 4\text{NH}_3(g) \rightarrow \text{Si}_3\text{N}_4(s) + 12\text{H}_2(g) \quad 4.3
\]

\[
\text{3SiCl}_2\text{H}_2(g) + 4\text{NH}_3(g) \rightarrow \text{Si}_3\text{N}_4(s) + 6\text{HCl}(g) + 6\text{H}_2(g) \quad 4.4
\]

Once the nitride is in place the wafers are ready for lithography and etching. The thickness of the nitride can be checked with an ellipsometer. A wafer should be kept out of the oxide growth process so that the nitride thickness can be measured as a single film thickness. The more thin film layers that are present in a film stack the more difficult it is to get an accurate reading of film thickness with an ellipsometer.
4.3 Lithography

4.3.1 Photoresist
Before the wafers can be etched a layer of photoresist must be patterned on the wafers to define the etch features. To improve adhesion of the photoresist to the wafers, they are first baked in a hexamethyldisilazane (HMDS) environment. HMDS is meant to promote adhesion to oxides, but experience has shown that it helps photoresist stick better to nitride also. If the backside of the wafer is being etched and the resist is required only to protect the frontside of the wafer, the HMDS step may be omitted. After HMDS application the coater deposits liquid photoresist onto a wafer while it is spinning. This spreads the resist to a thickness of about one micron. After the resist is spun onto the wafer it is baked for a minute at 115 °C to preharden it. That is the final step before exposure.

4.3.2 Stepper
The stepper exposes the image of a mask into the photoresist coated on a wafer. Masks for the stepper are 10 cm square pieces of fused silica with chrome plated on one side. Chrome is selectively ablated with an e-beam to create the desired pattern. The stepper actually demagnifies the mask image to 1 cm square. This small image covers only a portion of the wafer so the stepper tiles the image across the wafer repeating it in a two-dimensional, 8x8 array.

When fabricating features about the same size as the resolution limit of the stepper (~1 μm), the focus and exposure time of the tool must be carefully calibrated. A test wafer receives an array of images with varying exposure times and focus levels. After development the die with the clearest pattern dictates the proper exposure time and focus. Because the lens produces a slightly curved image surface on the wafer, the edges of the dies tend to have a different focus than the center of the image. Thus the best die on the test wafer should be chosen by looking at the critical features on the die. During development of the MEMS SLM process, choosing the exposure time and focus by looking at features on the periphery caused the beams in the center of the die to be out of focus and too narrow as shown in figure 4.2. After being exposed, the photoresist is developed and then postbaked. The postbake step further hardens the photoresist so that it becomes a durable etch mask.
4.4 Plasma Etching

4.4.1 Nitride Etch
The layers on the wafer get etched in the opposite order than they were applied. The first etch steps are performed in a plasma etcher. The plasma etcher breaks down gases into chemically active components with a high power RF plasma. For this process CF₄ enters the chamber to etch trenches in the nitride at 350 watts. If the right balance of chemistry and RF power is used, nearly vertical sidewalls can be etched. The speed of the etch can be characterized by measuring the nitride thickness on the ellipsometer before and after the etch.

4.4.2 Oxide Etch
In the oxide process a chemistry of CF₄ and CHF₃ etches the sacrificial oxide at 350 watts.

4.4.3 Polysilicon Etch
Achieving straight sidewalls with a polysilicon etch is challenging. An etch using HBr and Cl₂ at 300 watts seemed to attack the photoresist and etch the nitride such that it produced slanted sidewalls. Figure 4.3 shows the slanted sidewalls formed by this etch. The sidewalls shown in the right figure were produced by etching first with HBr and Cl₂ at 300 watts and then with just HBr at 100 watts. Using only HBr for the entire etch should create much more vertical sidewalls.

The sidewalls of the sacrificial layer must be perfectly vertical or even slightly undercut for the metal deposition. If the walls are slanted at all then the metal will cover them and connect the bottoms of the trenches to the tops of the beams. This will prevent the beams from being freed when the final release etch is performed. Even perfectly vertical sidewalls may not prevent gold from covering the sidewalls.
Figure 4.3 A plasma etch with HBr and Cl$_2$ of the polysilicon layer on the left. The etch also attacks the photoresist and the nitride which creates slanted sidewalls. The etch on the right used Cl$_2$ only during the first part of the etch, but completed using just HBr.

4.5 Metal Liftoff

4.5.1 KOH Etch of Polysilicon

To prevent the sidewalls from getting covered in gold, the polysilicon must be isotropically etched to slightly undercut the beams. The etch should be timed precisely so that it will not entirely undercut the beams and release them. If they were released early the beams would probably become stuck to the bottom of the trench due to stiction during subsequent wet processing. Figure 4.4 shows the results of slanted and vertical sidewalls in the left picture and the desired results of a slight undercut in the right image.

Figure 4.4 The beams in the SEM image on the left have gold on the sidewalls even though they are nearly vertical. Also the gold seems to bridge across the trench in the trench second from the left. In the right image the gold shouldn’t cover the sidewalls if they slightly undercut the beams.
The KOH etch needs to be properly characterized before it can be used with confidence in this process. After the KOH etch and before the lithography for the metal liftoff the wafers must go through an extensive cleaning process. Three separate "piranha etches" clean the wafers so that they can be placed back into the photolithography equipment. A piranha etch is the first step of the RCA clean except with a 7:3 mixture of sulfuric acid to hydrogen peroxide.

4.5.2 Stepper Lithography for Liftoff
The lithography for metal liftoff is the same as for etching except that the post-bake is omitted from the liftoff procedure. Hard baking photoresist can cause it to reflow slightly, which rounds the photoresist features. By not post-baking the photoresist it keeps sharp features and dissolves more easily in acetone later. For thicker metal layers, image reversal photoresist may give better liftoff results. Image reversal liftoff washes away in areas not exposed to light.

4.5.3 Metal Deposition for Electrodes and Traces
The MEMS SLM needs a layer of gold 20 to 50 nm thick to form the electrodes and traces. The second layer of metal which thickens the bond pads is 300 nm thick. 10 nm of titanium is first deposited to enhance the adhesion of the gold. An e-beam vacuum chamber deposits the titanium and gold. The wafers are mounted in a liftoff plate which holds them nearly perpendicular to the metal source. The evaporated metal travels in a nearly line-of-sight path because of the vacuum conditions (5e-7 torr) of the evaporation chamber. A high-energy electron beam sweeps back and forth across the metal charge in the crucible. This heats the metal and causes it to evaporate. The metal layers cover the wafer and photoresist liftoff layer thereon.

4.5.4 Metal Liftoff for Traces
The metal electrodes and traces form when the wafer is bathed in acetone. The acetone dissolves any photoresist left on the wafer taking with it the metal that was on the resist. Small features or thick metal layers sometimes have trouble lifting away as in figure 4.5. Immersing the acetone beaker in an ultrasound bath dislodges excess metal. The thicker bond pad layer of gold partially delaminated from the bond pads in the ultrasonic bath. Putting down a second layer of titanium before the
second layer of gold may help the gold stick better. After the metal has lifted off, the wafer is rinsed in DI water and dried from the center out with compressed nitrogen. For difficult liftoff patterning it was found that applying an adhesive tape (in this case Scotch Tape™) and peeling it off removed the liftoff metal very well. Perhaps a clean room compatible version of this liftoff could be developed.

4.6 Release Etch

4.6.1 Release Etch

Difficulties in the release etch motivated switching from the oxide process to the poly process. The original process called for the oxide sacrificial layer to be isotropically etched in HF acid to release the beams. Two complications arose. Because HF attacks titanium the gold electrodes and traces quickly delaminated from the die. Also the surface tension from the wet etch pulled the beams down to the substrate causing them to stick and not release. Releasing the beams in HF vapor was also tried but again the metal surface was attacked quickly leaving the surface mottled.

The polysilicon process utilizes a release etch that is becoming more familiar in MEMS – a xenon difluoride dry release. For this etch a wafer is placed in a chamber filled with XeF₂ gas which etches the polysilicon isotropically. This etch worked well to release the nitride beams that weren't coated in gold. As stated above the gold covering the sidewalls of the coated beams prevents them from moving. Figuring out a way to prevent gold from sticking to the sidewalls of the trenches will be key to getting this process to work.

4.7 Packaging

4.7.1 Packaging

After the dies are successfully patterned and released they must be packaged. The first step of packaging is die sawing the wafer into individual dies. This mask set produces an eight-by-eight array of dies on each wafer. Three dies in each corner of the array are clipped by the edges of the wafer. Thus each wafer can potentially produce 52 functional MEMS SLMs. After the die saw cuts up the wafer each die is placed in a 256 pin package. Each bond pad is then individually gold
wire bonded to a pin. Electrical contact with the bottom electrode is made through the back of the wafer. The back of the wafer is set at zero volts.

4.7 Conclusion

Each step in the oxide process and the poly process are crucial to creating a working device. Many steps were well characterized in both the oxide process and the poly process. The lithography, liftoff, and release steps of the poly process were worked out well. The sidewall etching of the polycrystalline silicon is a major issue that has not been properly resolved, but suggestions for working out solutions to this problem have been presented.
5. Conclusions and Suggestions for Future Work

5.1 Review

5.1.1 Thesis Review
In the pursuit of truly convincing visual communications, many solutions have been imagined. A few have enjoyed relative success, but far more have fallen short of their intended goals. This thesis has reviewed the technology of a few holographic displays. A variation on the Scophony-style holographic display has been presented. The capabilities of several spatial light modulators were compared. A MEMS diffractive grating was built in conjunction with this work, however, more fabrication challenges remain before we may arrive at a usable solution. The process techniques have been presented to encourage completion of the device.

5.2 Suggestions for Future Work
Following is a suggestion for a scalable MEMS SLM based holographic display. If an SLM can be built on the same substrate as multiplexing circuitry, then the SLMs could be tiled very closely together to form a very long one-dimensional array of diffracting beams. If a sufficiently long line can be tiled together then horizontal scanners would no longer be required to build MEMS-based holographic display. A single long vertical scanning mirror could scan the entire line. By coupling this setup with a concave mirror to refocus the image, a complete display can be made. The suggestions for implementing this display fall under four sections: SLM fabrication, electronics configuration, and optical system design.

5.2.1 SLM Fabrication
One shortcoming of the SLM presented in this thesis is that only a small length in the center of the beams is effectively available to diffract light. This happens because the beams curve when they are bent so that the phase delay away from the center of the beams is less than that in the center where the deflection is greatest. This means that input light must be focused down to a very small line centered on the middle of the
beams. Longer flat beams like those used in the Polychromator would allow the focus of the line to be relaxed. However, the added complexity of the Polychromator process is not attractive. Here I present a potentially simple fabrication process that would create longer flat areas in the beams and could alleviate problems due to wet release etching of the beams and sidewall metallization. The process is summarized in figure 5.2.

A silicon wafer receives an oxide layer and a nitride layer. A photoresist layer defines the beams' shape. Ideally the beams would be wider than the gaps between them, for example 1 micron wide beams with .5 micron wide gaps. Trenches are then etched into the nitride layer using the photoresist as an etch mask. The etch would likely be a plasma etch such as CF$_4$. The photoresist is then removed and a second resist layer is patterned to expose only the middle of the beams (starting from about 10 microns from the ends of the beams) to an HF etch. This etch exposes the underlying single crystal silicon substrate. The wafer is then placed in a KOH etch to create a groove under the middle of the beams. KOH etches single crystal silicon along two directions much faster than the third so that the etch is shaped like a trench. The remaining oxide under the beam ends is then etched to fully release the beams. Because the distance from the middle of the beams to the underlying substrate is more than 28 microns the beams should remain released during subsequence wet etches. When the gold is deposited, there should be no problems with the gold connecting the beam sidewalls to the bottom of the trench.

The actuation electrodes no longer act across the entire length of the beams, but only towards the ends. This type of actuation is known as leveraged bending. Combined with the strain stiffening effect of the thin beams the travel range of the beams should be increased past one-third the original gap. This means that the silicon dioxide layer and hence the gap can be made thinner at the outset. These steps should fit into a modified cmos process.

Leveraged bending requires higher voltages to actuate a beam than full length actuation. When the voltage becomes too great the drive circuitry becomes unreasonable to fabricate on the chip. A rough estimate of the drive voltage can be arrived at by assuming that the contribution of the unactuated middle section of the beam doesn't contribute to setting the drive voltage. Then if the beam length is set at the actuation electrode length the pull-in voltage can be estimated as before. Figure 5.3 shows the electrode length vs. pull-in voltage.

![Figure 5.1 Circuitry required to drive one beam.](image)
After plasma (CF4) etch of nitride layer to form beams

After HF etch of middle of oxide sacrificial layer with resist in place

After KOH etch of silicon substrate to form a 28 micron deep v-trench

After HF etch to get rid of sacrificial oxide and after aluminum deposition

Figure 5.2 New MEMS process profiles.
Figure 5.3 The effect of actuation electrode length on the pull-in voltage can be seen here for a beam that is 100 microns long. The actuation electrode length is the sum of the electrode lengths extending toward the middle of the beam from the beam ends. For this simulation $g_0 = 500$ nm, $\sigma_0 = 110$ MPa, $W = 1.5e-6$, and $H = .15e-6$.

Because most of the space on the MEMS diffractive SLM presented in this thesis is dedicated to the pin out of the electrode traces, the number of beams that can be independently controlled is limited. If cmos transistors could be built in the available space instead, then the number of beams could be increased. Because the beams are deflected by a simple voltage, the necessary drive electronics for each beam could be made similar to those in an active matrix LCD. A transistor and a capacitor as shown in figure 5.1 should be sufficient to drive one beam. This circuit is addressable by rows and columns.

Addressing the beams by rows and columns would greatly reduce the number of bond pads required to control all of the beams. If $N$ is the total number of beams on the chip then the number of bond pads required to column/row address the beams is $2*N^3$. If half of the beams are addressed from one side of the die and half from the other then the number of bond pads becomes

$$bondpads = 4 \left( \frac{N}{2} \right)^5$$

5.1
If the beams can be fabricated 1.5 microns wide with .5 micron gaps between them then 5000 beams would fit side to side on a 1-cm wide die. Thus 2500 transistors and capacitors would have to be addressed on either side of the beam array. Each side would need 100 bond pads. This is 56 fewer bond pads than are needed to address the current design. However if these are to be tiled together very tightly then the limiting dimension is not the size of the bond pad but how closely traces on a printed circuit board can be packed. A 2-cm wide die would fit 10,000 beams across but only require 142 pads per side or 284 pads total. This leaves 140 microns for each bond pad.

The area available for control circuitry for each beam is the beam width plus the gap width times the distance from the end of the beam to the edge of the bond pad. For a 1 cm wide die this is 2 \mu m * (5000 \mu m - (50 \mu m + 140 \mu m)) = 9620 \mu m^2. That equates to a square of silicon 98 microns on a side in which to build a transistor, a capacitor, interconnects, and an isolation barrier between transistors. A TI DMD only uses a square 17 microns on a side to house its drive transistors and address interconnects.

To equal the horizontal spatial bandwidth product of the Mark II display 26 chips would need to be tiled together and then demagnified by a factor of 3.5. Figure 5.4 shows the layout of the die and how they would be tiled together.

Figure 5.5 shows a possible optical setup using the tiled MEMS diffractive SLMs and a single vertical scanning mirror. This setup would likely be capable of thousands of horizontal hololines because it would scan vertically in a way similar to the horizontal scanning of the GLV two-dimensional projection display. Of course, this display would be subject to the same computational bandwidth problems of every other holographic display that has been built.
Figure 5.4 26 two-centimeter wide chips with 1.5 micron beams and .5 micron gaps tiled side-by-side with onboard drive circuitry and demagnified 3.5 times would have the same spatial bandwidth product as the Mark II. Holographic Video Display.

Figure 5.5 Optical layout for a holographic display based on a one-dimensional array of MEMS diffractive SLMs.
5.3 Conclusions

5.3.1 Conclusions
Holographic display technology is a challenging and exciting new visual medium. The art and science of building holographic displays will continue as light modulating methods are explored. A MEMS diffractive SLM may be a significant addition to the tools that electro-holographers have to work with. As the right complement of tools becomes available these researchers should be able to give to the world a new means of communication. When holographic displays become commonplace our communications abilities will improve dramatically (and husbands still won’t understand what their wives are telling them).

The End of a Beginning
References


3 Okolicsanyi, F., Wireless Engineer, 14, 527, 1937.


