Optoelectronic Circuits using Holographic Elements

Thesis by

José Mumbrú

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2002

(Defended May 28, 2002)

Acknowledgments

First, I would like to thank my thesis advisor, Prof. Demetri Psaltis, for the guidance and support he has given me over the years of my work in his group. It has been a remarkable experience and a privilege to work in his Laboratory. My gratitude also goes to Lucinda Acosta for her enthusiastic assistance with all administrative matters, and to Ya-Yun Liu, the group's technician, for her work on maintaining the lasers and other lab equipment in fine condition for my experiments.

I owe a big "thank you" to Iouri Solomatine, who has been essential in fabricating polymer materials. I could not have performed the experiments on the MB-PVA film presented in Chapter 3, or the PQ-PMMA material in Chapter 5, without his valuable help. I also thank Vladimir Shelkovnikov for his collaboration in the synthesis of the modified PQ molecules, and Dr. Andrei V. Veniaminov for conversations leading to a deeper understanding of the physics of the PQ-PMMA material. I want to thank Prof. Khrishna Shenoy for enjoyable collaboration and discussion on how to apply the ideas of reconfigurable computers and OPGAs into the field of neural prosthetics, which led as a result to the publication of a chapter in the book *Replacement Parts for the Brain: Intracranial Implants of Hardware Models of Neural Circuitry*.

I would like to thank all of the Psaltis group members, past and present, for creating a professional and friendly research environment. These include Prof. George Barbastathis, Dr. Gregory Billock, Prof. Karsten Buse, Dr. Ernest Chuang, Emmanouil-Panagiotis Fitrakis, Vijay Gupta, Dr. Michael Levene, Zhenyu Li, Dr. Wenhai Liu, Dr. Zhiwen Liu, Hua Long, George Maltezos, Irena Maravic, Todd Meyrath, Dr. Christophe Moser, Dimitris Sakellariou, Benjamin Schupp, Dr. Xu Wang, Yunping Yang, and anyone else I might have unintentionally missed. Special thanks go to Prof. Ali Adibi, who helped me in getting integrated into the group when I arrived five years ago, and to Dr. Gregory Steckman for teaching me many of the skills necessary to carry out sophisticated experiments in the lab. I am indebted to Dr. Xin An, along with Dr. Gan Zhou and Dr. Fai Mok from Holoplex, for their help in making some of the OPGA achievements described in this thesis a reality. Last

but not least, many thanks to George Panotopoulos, Hung-Te Hsieh, and Martín Centurión, who have been both my officemates and friends.

I would like to express my gratitude to Prof. Peter Bossaerts for his advising in my pursue of a subject minor in Financial Economics. Special thanks to Dr. Allen Pu, not only for being a former Psaltis group member, but also for letting me see what a day in the life of a venture capitalist is like. I also thank my undergraduate thesis advisors Prof. Irina Veretennicoff and Prof. Hugo Thienpont and their research group at the Vrije Universiteit Brussel for having stimulated my interest in Optics to the extend of doing a Ph.D.

I thank my wife, Aude, for her support and understanding while working long hours in the labs of Caltech. Also, my family for their continued encouragement through all the years of my education. Finally, I would like to acknowledge *La Caixa* for their fellowship program, which provided me with financial support during my first two years at Caltech.

Abstract

This thesis presents the results of research in the use of holographic modules in optoelectronic systems, their applications, and the characterization of polymer materials on which to record volume holograms, for these modules. The first chapter makes the case that a direct interface between an optical memory and a chip integrating detectors and logic circuitry can better utilize the high parallelism inherent in holographic modules.

The second chapter introduces the idea of reconfigurable computing and Field Programmable Gate Arrays (FPGAs) as the framework in which to design a hybrid system, the Optically Programmable Gate Array (OPGA), that outperforms its electronic counterpart by reducing its reconfiguration time by three orders of magnitude.

The OPGA is the combination of three elements: an addressing device to selectively recall holographic data pages, an optical memory, and an optoelectronic chip. The third chapter investigates the issues related to each one of these elements and their integration in a compact module. Operation of the system is demonstrated by holographically programming the OPGA chip.

In the fourth chapter, experiments are performed to characterize the Aprilis ULSH500-7A photopolymer and study quantitatively its ability to store high-bandwidth holographic data pages. A method for hologram mastering and copying using Aprilis films is also demonstrated.

Chapter five investigates the recording and diffusional-amplification dynamics of the PQ-doped PMMA polymer. Different strategies to optimize the material by reducing the duration of its post-exposure are examined, and the corresponding experimental results presented. The recording of strong gratings in samples of PQ-PMMA is tested to determine the suitability of this material for components in optical fiber networks.

The final chapter deals with applications and systems that can benefit from a direct interface with a holographic module. The use of optically programmable processors in the field of neural prosthetics is explored. The design of a holographic 4-D microscope is presented and tested experimentally.

Table of Contents

CHAPTER 6 Holographic systems for processing and imaging applications

List of Figures

CHAPTER 1 Introduction

CHAPTER 2 Optically reconfigurable processors

CHAPTER 3 OPGA elements and system integration

CHAPTER 4 Holographic data storage in Aprilis photopolymer

CHAPTER 5 Optimization of PQ-PMMA polymer

CHAPTER 6 Holographic systems for processing and imaging applications

List of Tables

1 Introduction

Since its invention by Gabor in 1948 [1-1], holography has been applied in many different areas in industry as well as in science. The use of holography for optical data storage has drawn particular attention since proposed in the early 1960's [1-2],[1-3]. The high storage densities achievable with holographic memories, first estimated by van Heerden in 1963 [1-2] reaching the range of a TByte/cm^3, motivated intense research to develop holographic systems. However, it was not until the early 1990's that advancements in optoelectronic devices and materials made viable holographic memories [1-4],[1-5].

Parallelism is inherent to holographic memories in two different ways: First, due to the Bragg effect, it is possible to superimpose (i.e., multiplex) hundreds, or even thousands, of holograms in the same volume that can be retrieved either individually or simultaneously, like in optical correlators, without crosstalk. Second, the information carried by the holograms can be spatially modulated and formatted as pages of digital data that can contain information in the order of 1 Mbit. This parallelism makes it possible to achieve very high transfer rates, in the range of 1 Tbit/s, between the optical memory and the photodetector array during the readout cycle.

Although holographic memory platforms have been developed [1-6]-[1-9] to successfully demonstrate such high densities and transfer rates, typically they have not addressed the issue of how to deal with the information once it has been retrieved. There are two issues that require consideration: The first one is how can this gigantic transfer rate

be sustained after detection. In other words, how can an optoelectronic system efficiently deliver the retrieved data from the module of the detector to the processor unit of the system without loss in the transfer rate. The second issue is about the sort of processors, and by extension applications, that not only demand a large amount of input data, but also need to process it at very high rates.

This thesis addresses these questions as it looks at the optoelectronic system as a whole, in which the holographic module needs to interact efficiently with the processor. In that aspect, it will be argued that a direct interface between a holographic module and a silicon chip that integrates computing resources alongside photodetectors can meet the high throughput of the optical memory and satisfactorily solve the bandwidth bottlenecks of current systems. The thesis also deals with the issue of identifying systems in which to show the advantages derived from such an interface. Various situations are considered in which the hologram is always the pivotal element of the system, although it performs a different role in each case.

In the context of an optically reconfigurable processor, the holograms store the information needed to program the logic. The system benefits from the parallelism of the holographic data page and the distributed hardware resources of the device to increase the computing performance of the processor. In the case of reconfigurable interconnects, the hologram can store several interconnection patterns. High throughput is achieved by having simultaneously many optical channels, each one modulated at high speed, being selectively redirected by different holograms. For imaging applications, a holographic microscope utilizes the diffraction properties of volume gratings to image a specimen onto the detector. In this scenario, high transfer rate is again obtained by having the unknown

wavefront of the specimen be sensed in parallel by the multiplexed holograms to extract information of the specimen and deliver it efficiently to the detector.

Although these are just a few examples, they illustrate the basic principles through which holography can tackle the problem of high information transfer rates. Holograms can convert a simple input (e.g., a plane wave reference beam) into a complex 2-D wavefront, or even a transform a 2-D wavefront into another. Furthermore, multiple holograms can process the same wavefront simultaneously without one grating affecting the operation of the others. It is believed there are many more systems that could take advantage of these properties if they incorporated a direct interface to a holographic module.

1.1 Outline of the thesis

The work presented in this thesis can be divided into three different areas: holographic systems, optical materials, and applications. At the holographic system level, Chapter 2 presents an optoelectronic system, the so-called Optically Programmable Gate Array (OPGA), that implements a direct interface between a holographic memory and a silicon chip that combines both detectors and logic resources. The context of reconfigurable processors and, more specifically Field Programmable Gate Arrays (FPGAs), has been chosen to illustrate how a fast optical interface can overcome the limitations of current electronic devices. The OPGA consists of three major elements: an array of VCSELs to recall the holographic templates, an optical medium, and a VLSI chip. After looking at the system as a whole in Chapter 2 and evaluating two different architectures, Chapter 3 addresses the main requirements and issues for each one of the three components, and their integration in a compact module. Finally, holographic programming of the OPGA chip is demonstrated.

The second part of the thesis deals with holographic materials. The feasibility of high-performance optoelectronic systems relies heavily on the availability of holographic materials that can fulfill their demands. Photopolymers are particularly attractive for read-only applications because of their high recording sensitivity and large modulation of their index of refraction. Chapter 4 presents the experimental characterization of Aprilis polymer, testing not only its dynamic range but also performing a quantitative measurement of its ability to multiplex high-bandwidth images. Chapter 5 considers a different type of polymer, the PQ-PMMA, that exhibits hologram amplification through diffusion. Efforts to optimize the material in order to reduce the duration of its post-processing, or even make it unnecessary, are evaluated through experimentation.

The last piece of the thesis looks at the applications in which to show the advantages derived from using hybrid optoelectronic systems. The final section of Chapter 5 explores the possibility of implementing filters (or interconnects) for wavelength division multiplexing (WDM) systems using holograms stored in a thick polymer like PQ-PMMA. In the context of the OPGA, Chapter 6 applies the idea of optically reconfigurable processors into the area of neural prosthetics and analyzes a specific example. Chapter 6 also presents a holographic microscope able to extract 4-D information (3-D spatial plus color) from a specimen in real time. The architecture of such a microscope is analyzed theoretically and its operation is experimentally demonstrated.

References

[1-1] D. Gabor, "A new microscope principle," *Nature* **161**, 777 (1948).

[1-2] P. J. van Heerden, "Theory of optical information storage in solids," *Applied Optics* **2**, 393-401 (1963).

[1-3] E. N. Leith, A. Kozma, J. Upatnieks, J. Marks, and N. Massey, "Holographic data storage in three-dimensional media," *Applied Optics* **5**, 1303-1311 (1966).

[1-4] D. Psaltis, and F. Mok, "Holographic memories," *Scientific American* **273**, 70-76 (1995).

[1-5] D. Psaltis, and G. W. Burr, "Holographic data storage," *Computer* **31**, 52-60 (1998).

[1-6] A. Pu, and D. Psaltis, "High-density recording in photopolymer-based holographic three-dimensional disks," *Applied Optics* **35**, 2389-2398 (1996).

[1-7] X. An, D. Psaltis, and G. W. Burr, "Thermal fixing of 10,000 holograms in $LiNbO_3$:Fe," *Applied Optics* **38**, 386-393 (1999).

[1-8] G. Zhou, D. Psaltis, and F. Mok, "Holographic read-only memory," *Optical and Quantum Electronics* **32**, 405-417 (2000).

[1-9] G. Burr, H. Coufal, C. M. Jefferson, J. Hoffnagle, M. Jurich, R. MacFarlane, and R. Shelby, "Holographic storage delivers high data density," *Laser Focus World* **36**, 123-+ (2000).

2 Optically reconfigurable processors

2.1 Introduction

Reconfigurable processors bring a new computational paradigm where the processor modifies its structure to suit a given application, rather than having to modify the application to fit the device. The reconfigurability makes it possible for these processors to use more efficiently their resources by adjusting themselves depending on the characteristics of the input or on nonsatisfactory previous results to better implement the target task.
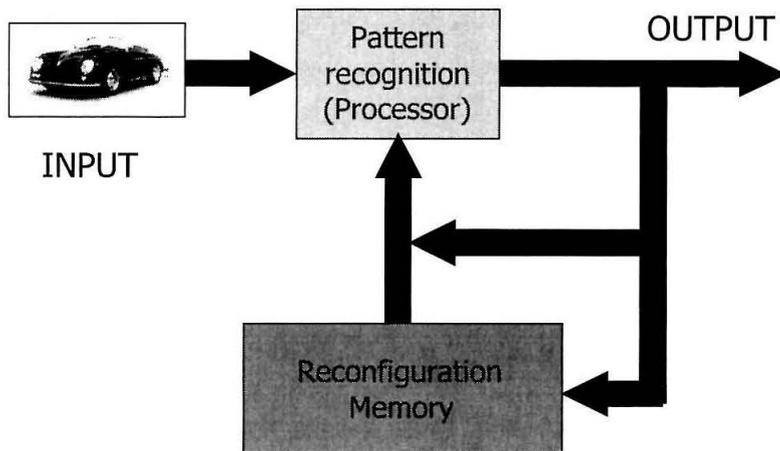
Fig. 2-1. Reconfigurable processor applied to pattern recognition. The external memory stores the configuration templates that define the functionality of the processor. Using reconfiguration the same processor can perform different tasks on the same input image.

Given an application, like pattern recognition in Figure 2-1, the reconfigurable processor can be customized to deal with a specific class of objects, but with enough flexibility that, if at a later time the salient class of objects shifts to a different one, the device can be reprogrammed to deal with the new problem without degrading its performance. Further-

more, the processor can adapt itself in order to be robust to changes of orientation or illumination of the input object. By reprogramming, the same hardware can be time-multiplexed to carry out sequentially several tasks on the same input, or perform different tasks to different parts of the same input image. Reconfiguration also makes it possible to implement learning by allowing the processor to evolve in a controlled manner in order to learn the function that needs to be computed.

In other applications, where it is necessary to implement different concurrent tasks by partitioning the hardware resources among them, a reconfigurable processor can outperform a nonreconfigurable solution by dynamically reallocating the hardware of idle tasks into those that may be temporarily overflowed. This feature, called spatial multiplexing, becomes especially attractive when partial rather than global reconfiguration is possible because it allows reprogramming part of the device without halting the execution in the rest of it.

In this chapter, we will see how optics can play a very important role in the development of reconfigurable processors. Section 2.2 provides the reader with a brief description of Field Programmable Gate Arrays (FPGAs), while Section 2.3 explains how the performance of a conventional FPGA can be enhanced by interfacing the chip to a holographic memory. Finally, the architecture of the OPGA module and the system design considerations are presented in Section 2.4.

2.2 Field Programmable Gate Arrays (FPGAs)

A Field Programmable Gate Array (FPGA) is a device where this idea of reconfigurable hardware can be implemented. FPGAs emerged as a new technology for the imple-

mentation of digital logic circuits during the mid 1980's. The basic architecture of an FPGA

consists of a large number of Configurable Logic Blocks (CLBs) and a programmable mesh

of interconnections. Both the function performed by the logic blocks and the interconnec-

tion pattern can be specified by the circuit designer. In the beginning FPGAs were mostly

viewed as large Programmable Logic Devices (PLDs) and they were usually employed for

the implementation of the "glue-logic" used to tie together complex VLSI chips like micro-

processors and memories used to build general purpose computers.

While several FPGAs were configured by static RAM (SRAM) cells, this was gen-

erally considered a limitation by users concerned about the chip's volatility. For this reason,

fuse-based FPGAs were also developed and for many applications were much more attrac-

tive, both because they were faster and smaller, due to less programming overhead, and also

because there was no volatility in their configuration since this had been burned into the

chip. Not until the late 1980's and early 1990's did it become clear that the volatility of

SRAM-based FPGAs was not a liability but could open an entirely new spectrum of appli-

cations, since the programming of such FPGAs could be changed electrically at almost any

point during operation.

These devices have gained popularity due to the fact that they are between a soft-

ware oriented solution, like a microprocessor running a program stored in memory, and a

hardware-oriented solution, like an application specific integrated circuit (ASIC)

(Figure 2-2). The FPGA-based solution is faster than a microprocessor or Digital Signal

Processors (DSPs) because the FPGA is conceived as a large array of small logic blocks

working in parallel and operating at the bit level, exactly where general purpose processors

are most inefficient. Even though microprocessors have more capabilities, in order to keep
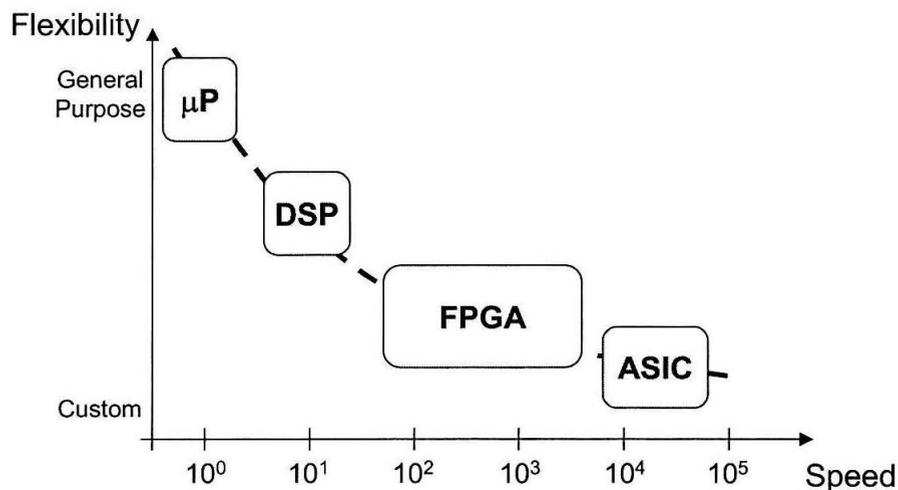
Fig. 2-2. Flexibility/speed trade-off comparison for different hardware implementations: microprocessors (µP); digital signal processors (DSP); field-programmable gate arrays (FPGA); and application specific integrated circuits (ASIC).

their generality, they are still designed to operate with fix data formats (8, 16, 32, 64 bits...). Therefore, they perform poorly when they need to deal with problems where data has "non-standard" lengths. On the other hand, the fine granularity of the computing blocks of the FPGA allows the user to better map the hardware resources of the chip to meet the demands of the problem. Using FPGA platforms, speedups of several orders of magnitude have been achieved for some applications [2-1]-[2-4]. The ASIC solution provides most of the time the optimal implementation both in terms of speed and silicon area requirement, however, it has the drawbacks of being a single-purpose processor. Compared to ASICs, FPGAs are much more flexible since they contain some hardware resources that can be programmed by the user to implement some given task and, by changing that configuration, the same hardware can be used to carry out something totally different with minimal development time and cost.

Although FPGA architecture design is a field of important ongoing research in the FPGA community, and many different implementations have already become commercially available, it is beyond the scope of this section to describe all of them. To illustrate the common features of their internal structure, one of the most widely used designs, the symmetric array [2-5], Figure 2-3, will be analyzed. In this case, the logic blocks are
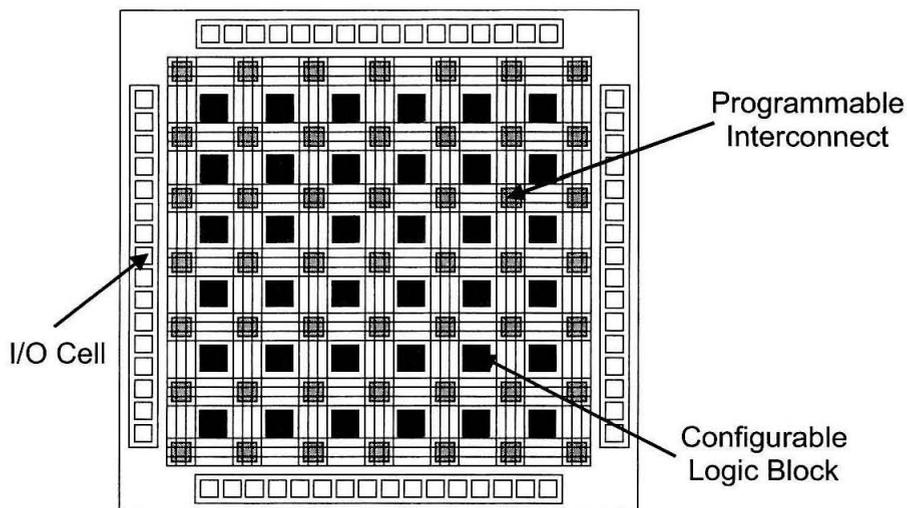
Fig. 2-3.   Architecture of a typical FPGA. A symmetric array of CLBs is surrounded by a mesh of buses and matrices of programmable interconnects that provide connectivity among the CLBs, as well as with the input/output (I/O) cells.

arranged in a two-dimensional array and interleaved with vertical and horizontal buses used to establish connectivity among them. Connections between segments in two different buses can also be performed by means of programmable interconnects in switching matrices. Finally, on the periphery of the chip, there are some input/output cells.

The basic functional unit of the FPGA is the Configurable Logic Block (CLB), which implements an elementary Boolean operation. Despite the fact that there are CLBs based on multiplexors or OR-AND arrays, the use of look-up tables (LUTs) to synthesize

logic functions enjoys much greater flexibility [2-5]. A LUT can be seen as a small bank of memory where the inputs encode the address of a position in this memory, which stores the result of a pre-programmed logic function of the inputs. By changing the bits stored in the LUT, the function computed is altered.
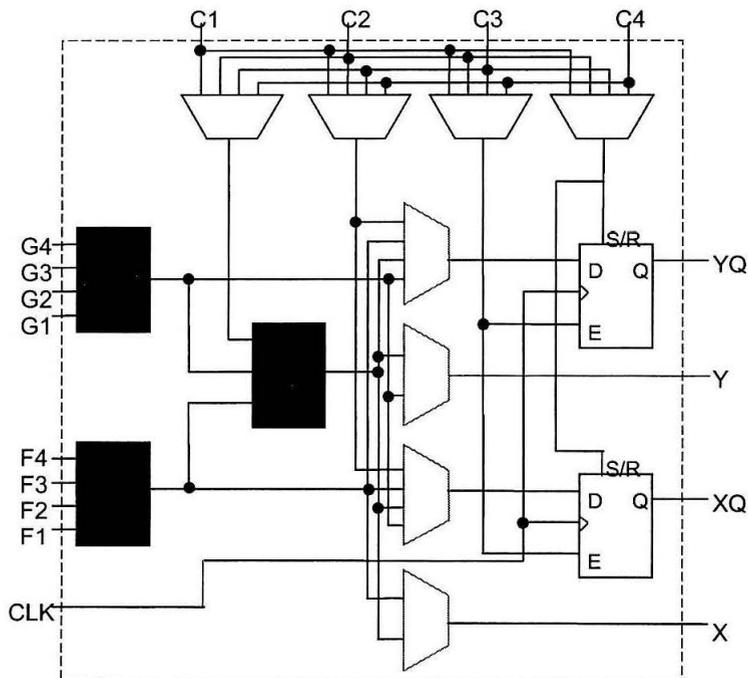
Fig. 2-4. Schematic of a LUT-based CLB. Two independent sets of inputs F[1:4] and G[1:4] feed the LUTs on the left of the figure. The outputs of the LUTs can be combined using an additional LUT for more complex Boolean functions. The control signals C[1:4] define the way the results of the LUTs are routed to the output of the CLB by means of the multiplexors. This output can be buffered, which allows implementing sequential logic.

The schematic of a LUT-based CLB is shown in Figure 2-4. In this case, two sets of inputs, on the left-hand side, feed two independent 4-input LUTs. A third LUT has the ability of combining the results of the LUTs from the previous stage, increasing the functionality of the CLB to implement more complex logic functions. The two outputs of the

CLB are on the right-hand side and they can be buffered if necessary by means of flip-flops. These registers allow implementing sequential logic in the CLB.

FPGAs have traditionally been used very successfully as accelerators in many applications like signal processing [2-1], image filtering [2-2], automated target recognition [2-3], or cryptography [2-4]. In a typical arrangement, as shown in Figure 2-5, the
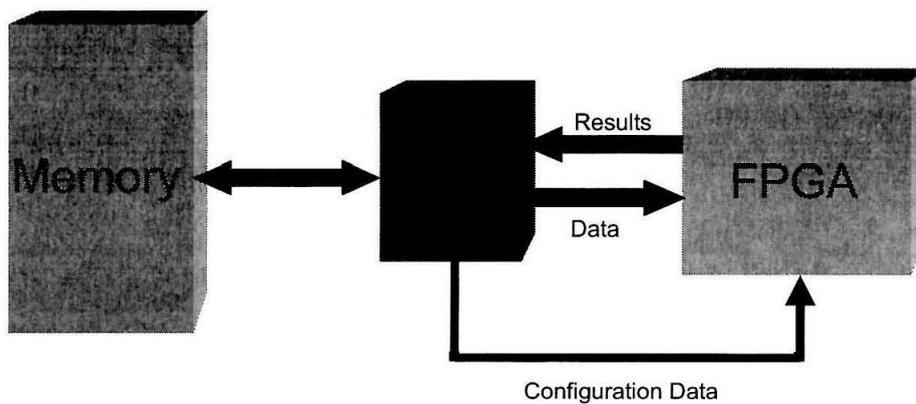
Fig. 2-5.   FPGAs are usually used as accelerators. The master processor (μP and memory) programs the slave processor (FPGA) to perform those most computationally intense tasks.

FPGA is setup as a coprocessor that is controlled by the microprocessor. For a given application, if there is some task that is computationally very expensive, the microprocessor can program the FPGA to perform that task much faster than if it was executed by the main processor. The configuration data of the FPGA, which specifies the values in the LUTs and the interconnection pattern, is stored in an external memory, in most cases an EPROM, and downloaded into the FPGA chip on demand. The microprocessor just feeds the data into the FPGA and waits for the results, and all the cumbersome computation has been hard-wired inside the FPGA.

Although the size of these devices, in terms of logic gates, can vary among different models and manufacturers, they can easily contain on the order of 10^5 gates and the trend is to keep increasing the logic density to go beyond the million-gate FPGA. This means that the configuration data page for a medium size FPGA can be as large as 1 Mbit. Despite the fact that the FPGA can be reprogrammed multiple times, the user typically does not take advantage of this feature. In most cases, the FPGA is configured only once and this configuration is downloaded into the FPGA off-line, before the execution is started. The main reason for not reconfiguring dynamically the device, i.e., changing its internal configuration once the execution has started and some data is already flowing into the device, has been the small communication bandwidth between the external configuration memory and the FPGA chip itself. The configuration bandwidth of the FPGA has not scaled well enough to keep up with the enormous data throughput. Upon programming, the configuration data is downloaded serially by shifting a long bit stream into the FPGA. The transfer data rate between memory and FPGA is only on the range of 100 Mbps, which results on configuration times of tens or even hundreds of milliseconds. These long reconfiguration times, if compared to clock cycles of just tens of nanoseconds and input/output throughputs reaching 100 Gbps, become an important overhead. Some attempts to decrease the reconfiguration times have been proposed, like providing a dedicated parallel bus to increase the bandwidth with the configuration memory, or having fast-access cache memory built in the chip [2-6],[2-7]. Both solutions only contribute to further increase the already high power dissipation of the FPGAs, which although application dependent can easily be in the range of 1–10 Watts.

2.3 Optically Programmable Gate Array

Optical memory modules possess inherently a high degree of parallelism, since the data is handled in the format of pages. Such parallelism results in a large communication bandwidth between the memory and the array of photodetectors during a readout cycle, or the Spatial Light Modulator (SLM) upon recording. The use of optical memories in information processing systems makes it necessary to consider the interface between the holographic module and the silicon circuitry that processes the data retrieved from the memory and stores computational results.

Traditionally, holographic systems have not addressed this issue, so even though the information can be delivered very fast to and from the optical memory, this parallelism is lost in the communication between the optoelectronic chips and the processor, becoming a bottleneck. Therefore, a direct interface between memory and processor would be much more effective since the parallelism would always be preserved, as suggested in Figure 2-6. The direct interface avoids the slow interchip communication by simply integrating on the same silicon die the logic circuitry and an array of photodetectors. However, the question now is to identify which computing devices have enough hardware parallelism to exchange data efficiently with the optical memory. It is here that the distributed hardware resources of the FPGA marry the parallelism of the optical memory.

Based on the FPGA architecture, the OPGA [2-8] is a device where the computation is still performed by programmable logic blocks and interconnects as in the conventional FPGA, but where the configuration data is brought into the chip optically. The design of the OPGA is conceptually different from a similar system described in [2-9], in which the information is brought into the chip via optical I/O ports while the configuration is still per-
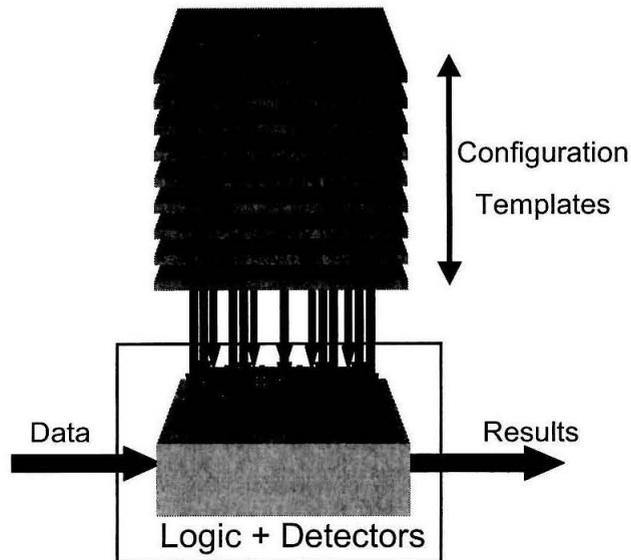
Fig. 2-6. Direct interface between the optical memory and the silicon chip carrying photodetectors and logic circuitry. The configuration templates stored as holograms (depicted as slices) can be downloaded in parallel to the chip.

formed electronically. This optical reconfiguration capability in the OPGA results from interfacing an optical memory with an optoelectronic chip, as illustrated in Figure 2-6. The holographic memory can store a large number of configuration templates that can be transferred down to the FPGA chip as a single page. By taking the reconfiguration circuitry out of the FPGA chip, the OPGA can achieve a larger logic density, i.e., more CLBs can be implemented, than in the conventional device.

In its initial implementation, the OPGA module is intended to operate as a Holographic Read-Only Memory (HROM), where a priori and for a given application, the user will decide the library of different configuration templates that needs to be stored in the memory. This frees the OPGA module from all the optics and optoelectronics required to write in the memory, like the SLM, and makes it very compact. However, it is conceivable

that future OPGA designs encompass both read and write capabilities, which would enable an increased computational flexibility.

The OPGA is basically the integration of three main components or technologies: an array of vertical cavity surface-emitting lasers (VCSELs) used to retrieve the templates stored in the memory; the optical memory that contains a large set of configuration contexts; and the VLSI chip that combines CMOS logic and photodetectors. Each one of these components presents a number of issues that will be discussed in the next chapter.

2.4 OPGA system architecture

The technique used to store and multiplex the holograms in the optical memory determines the architecture of the entire module. For this reason, it is not possible to discuss the holographic memory, the VCSEL array, or the chip separately without giving first a more general view to the system involving these three elements. The OPGA has the potential to outperform conventional FPGAs in two different aspects:

• **Number of fast-access configurations:** The OPGA module can contain one hundred holographic templates, formatted as pages of 1000×1000 pixels, so there is enough information in each hologram to program a medium-sized FPGA. Compared to cache-based FPGAs, in which between four and eight configurations are stored locally [2-6], the 100 templates stored in the optical memory represent a two-order of magnitude increase in flexibility for the OPGA.

• **Reconfiguration speed:** The parallel interface between holographic memory and chip makes it possible to decrease reconfiguration times in two or three orders of magnitude and go from the millisecond regime in electric FPGAs down to micro-

seconds in their optical counterpart. The reconfiguration time of the device depends greatly on the optical power available per VCSEL and is eventually limited by the integration time of the photodetectors, being the target configuration time between 1–100 µs.

This section introduces the two architectures that have been considered for the OPGA module, compares the main system-level requirements of the two designs, and discusses how they are interrelated.

2.4.1 Transmission-geometry OPGA module

The initial design for the OPGA consisted of a symmetric transmission-geometry module as depicted in Figure 2-7. Each one of the VCSELs in the array is used to record a different hologram in the memory, which in this case is a thin layer of red-sensitive photopolymer. For a matter of robustness, the polymer film is sandwiched between two prisms.
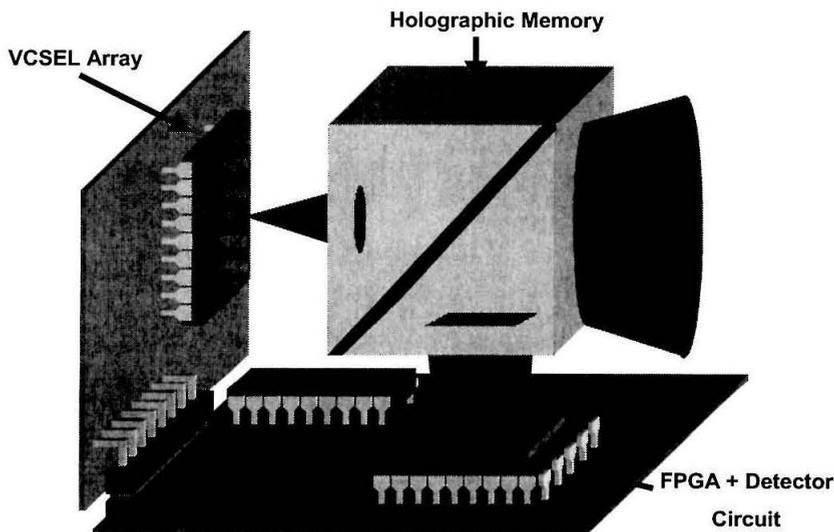
Fig. 2-7.   Transmission-geometry OPGA module. The spherical wave emitted by the VCSEL retrieves a holographic template from the optical memory and projects it onto the photodetector array in the chip.

The technique used to multiplex the holograms is shift multiplexing with spherical reference [2-10],[2-11]. Therefore, the shift selectivity of the material needs to be matched to the spacing between adjacent VCSELs in order not to have crosstalk between data pages. Phase-conjugate reference is used to read out the holographic templates, so the reconstructions self-focus on the array of photodetectors on the FPGA chip. Since no additional optics is required for the readout, the module is very compact.
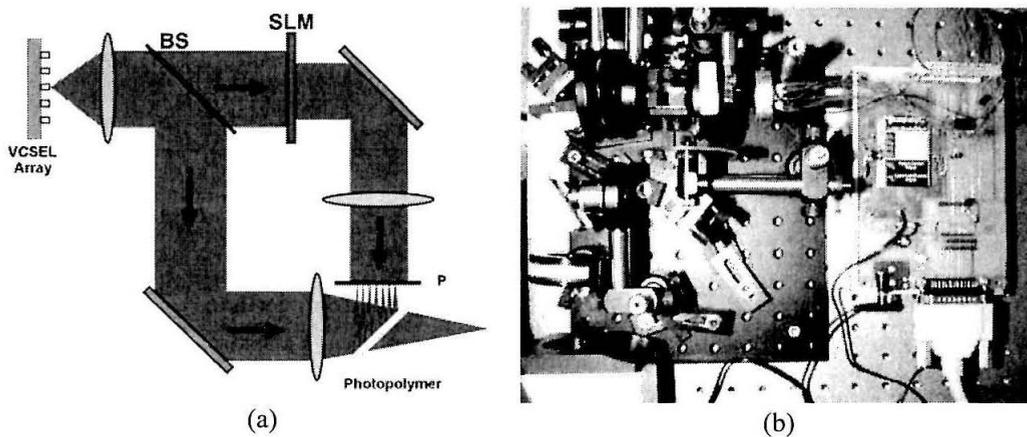
Fig. 2-8.   (a) Schematic drawing of the recording setup for the OPGA, where BS is a beam splitter and SLM is a spatial light modulator, and (b) actual setup used in the experiments.

In the recording setup, as the schematic in Figure 2-8(a) shows, the beam emitted by the VCSELs is collimated by the first lens. Then a beam splitter (BS) creates the reference and signal beams. A SLM is used to transfer the information that needs to be stored in the memory into the signal beam. The SLM image is projected to plane P, where the photodetectors will be upon readout, by a lens. The lens on the reference arm focuses the reference beam creating a converging spherical beam used to record the hologram on the photopolymer. During readout the array of VCSELs should be placed on the plane where

the converging reference beam focuses, so that the diverging beam emitted by the VCSEL creates the phase-conjugate reference that reads out the hologram.

Some experiments have been performed with this architecture to record and multiplex holograms in the memory, in this case a 100 μm thick film of Du Pont HRF700 photopolymer. In the experimental setup, Figure 2-8(b), a single mask with a chessboard pattern was used instead of the SLM and it was rotated to store different holograms. The board on the right of the picture corresponds to the circuit to drive the array of VCSELs. Figure 2-9 presents the reconstruction of two different pages. Since the VCSELs can be switched on and off in less than a nanosecond, different pages can be retrieved in very short time. This means that the OPGA module could switch rapidly among configuration contexts.

Hologram 1   Hologram 2
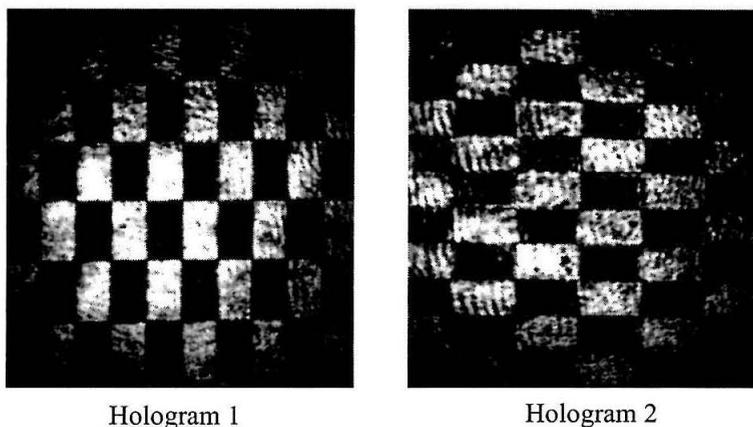
Fig. 2-9.   Detail of the reconstruction using VCSELs of two shift-multiplexed holograms in the transmission-geometry module.

A number of problems have been encountered in this architecture, not intrinsically related to the architecture itself but to the VCSELs, which were not specially designed for this application, and the medium used in the optical memory. The main problem, noticeable

in Figure 2-9, is that there is some crosstalk between holograms. The reason is that for the thickness of the medium, 100 µm, the shift selectivity of the memory is larger than the spacing between VCSELs (50 µm in particular for this array) and, consequently, Bragg-mismatch between neighboring holograms is not achieved. Placing the VCSELs much closer to the optical medium would narrow the shift selectivity of the memory. However, the small divergence angle of the VCSEL requires placing the array far enough as to overlap the entire footprint of the signal beam in order to record the hologram. Therefore, these two requirements turn out to be incompatible. Nevertheless, this limitation could be overcome by having a custom array of VCSELs with larger separation between elements and larger divergence angles. Another possibility would be to place a lenslet array on top of the VCSELs in order to increase their divergence angle.

Another issue is the quality of the reconstructed holograms. In this architecture, the spot of the signal beam on the material is relatively large, which makes the recording very sensitive to irregularities in the optical medium. However, a more serious issue is the power required per VCSEL to have short reconfiguration times in the OPGA. Since the 100 holographic pages are stored in a single location, the power requirement on each VCSEL becomes more demanding. If the target is to fully reconfigure the device in tens of µs, a straightforward calculation assuming a material with M/5 requires that each laser output a few mWatts. The characterization of the VCSEL arrays will reveal that a power requirement of 0.5 mW per element is more realistic.

2.4.2 Reflection-geometry OPGA module

Mostly due to the limitation in power, the architecture of the OPGA evolved towards a new design where the device could still have short reconfiguration times, in the range of tens of µs, but with a not so demanding requirement on the power per VCSEL. The technique used to store the holograms combines both spatial and shift multiplexing. The main difference with respect to the previous architecture is that upon recording, Figure 2-10, a lens focuses the beam before it impinges the SLM, as in a van der Lugt imag-
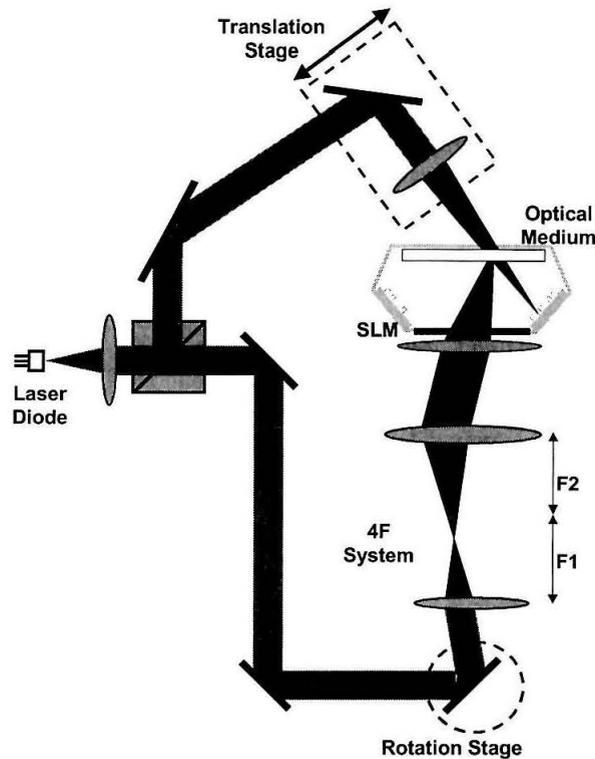
Fig. 2-10. OPGA recorder setup. The linear translation stage in the reference arm, combined with the rotation stage and the lenses in the 4-F system in the signal arm, is used for the shift multiplexing of the holograms.

ing system [2-12], down to a small spot on the recording medium. By changing the angle of incidence of the beam on the lens, the signal spot focuses on a different location in the

material, which is partially overlapping with the previous ones. The pages of data are recorded in these partially overlapping circles that span a stripe on the optical material. To achieve Bragg mismatch among holograms, a converging reference beam needs to be shifted accordingly to illuminate the corresponding signal spot. In the recording setup, a laser diode with enough coherence length can be used instead of the VCSEL array. The beam emitted by the diode is collimated and splitted into the signal and reference arm. The signal beam passes through a rotation stage and a 4-F system that changes its angle before it illuminates the SLM. The reference beam is focused by a lens mounted on a mechanical scanner used to translate the beam beyond the shift-selectivity of the optical medium.
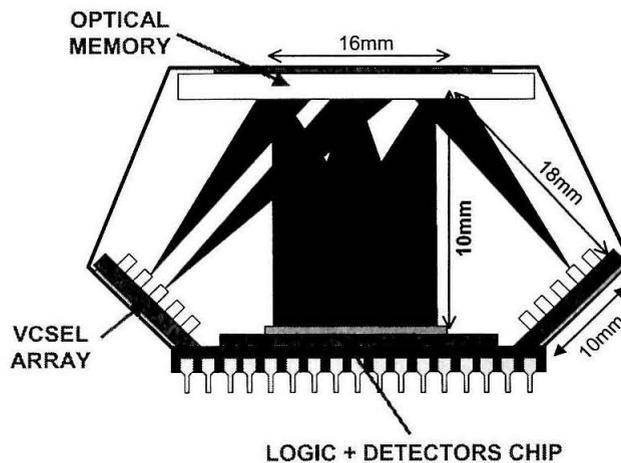
Fig. 2-11.   OPGA reader module. The light emitted by each VCSEL is used to read out a different hologram that self-focuses, due to phase-conjugation, on the array of detectors in the chip.

During readout, the system becomes very compact (Figure 2-11) for two reasons. First, the module uses reflection geometry for recording, so during readout the beam from the VCSEL and the array of photodetectors are both located on the same side of the material. Secondly, phase-conjugate readout makes unnecessary the use of any extra component. The VCSEL array is placed on the plane where the recording reference beams focus.

Thus, each VCSEL illuminates one of the spots in the memory and the reconstructed image back-propagates to the plane of the SLM where the photodetector array is located upon readout.
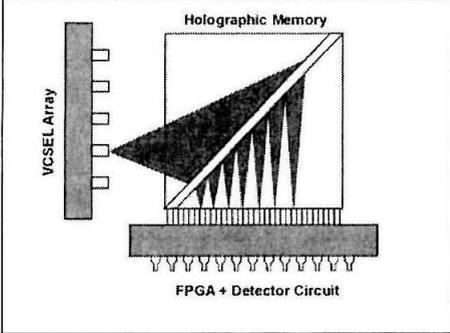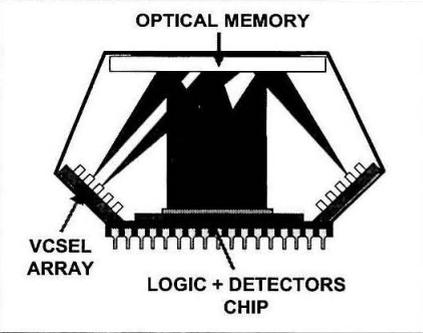
Transmission OPGA module	Reflection OPGA module
Signal spot size on material (D) $\sqrt{2}\left[Np + (\lambda/p)\dfrac{2d + Np(\lambda/p)}{1 - (\lambda/p)^2}\right]$	$\dfrac{2\lambda F}{p}$
Distance from VCSEL array (Z_0) $D \cdot \left[\dfrac{\sin(\alpha)}{\sin(\theta + \alpha)} + \dfrac{\sin(\alpha)}{\sin(\theta - \alpha)}\right]^{-1}$	$D \cdot \left[\dfrac{\sin(\alpha)}{\sin(\theta + \alpha)} + \dfrac{\sin(\alpha)}{\sin(\theta - \alpha)}\right]^{-1}$
Shift selectivity (δ) $\dfrac{\lambda Z_0}{2L\sin(\theta)}$	$\dfrac{\lambda Z_0}{L\sqrt{n^2 - \sin(\theta)^2}}$
Wavelength selectivity ($\Delta\lambda/\lambda$) $\dfrac{\lambda\sqrt{n^2 - \sin(\theta)^2}}{2L\sin(\theta)^2}$	$\dfrac{\lambda}{L[n + \sin(\theta)]}$

Table 2-1. Summary of the expressions of system design features for the transmission- and reflection-geometry OPGA modules.

A benefit of this new architecture is that the diffraction efficiency per hologram now scales not as the total number of stored holograms but the number of overlapping ones at any location. Therefore, an important increase in diffracted power is obtained. Assuming that the same 100 configuration pages are stored in the M/5 material, but only 20 holograms overlap at any given location, then the diffraction efficiency per hologram is as high as 6.25%. The increase in diffracted power makes possible to achieve µsec-scale reconfigura-

tion times with only 0.5 mW of available power. A comparison of the main features

between the transmission and the reflection modules is presented in Table 2-1, their rele-

vance towards the design of the OPGA will be discussed in the following paragraphs.

Table 2-2 provides the description of the system parameters and their values. They will

help to illustrate the differences between the two proposed designs.

Parameter name	Symbol	Transmission OPGA system	Reflection OPGA system
Pixel size	p	5μm	
Data page format	N×N	1000×1000	
VCSEL wavelength	λ	680nm	
VCSEL divergence angle (half angle)	α	4.4°	
Distance from recording plane (Fresnel hologram)	d	10mm	
Focal length of transforming lens (Fourier hologram)	F		10mm
Reference beam angle (w.r.t. material surface)	θ	45°	30°
Material thickness	L	200μm	

Table 2-2. Description of the parameters and values assumed for them in the comparison between the transmission- and reflection-geometry OPGA modules.

Besides relaxing the power requirement of the VCSELs, another benefit of the new

architecture is that recording occurs at the Fourier plane, rather than in the Fresnel zone as

in the original transmission module. In the transmission module, a 1000×1000 pixel holo-

gram with 5 μm pixel-size placed at a distance of 10 mm from the recording plane would

result in a footprint of 11 mm in diameter. On the other hand, for the reflection module, if

the SLM is placed at the same distance from the recording material, and the lens that

focuses the signal beam has a focal length of 10 mm, the signal spot size on the material

will be just 2.7 mm in diameter. Given the smaller dimension of the area where the pages

are recorded, the holograms are much less sensitive in the second case to any nonuniformity

of the medium and, consequently, the quality of the reconstructed images is better.

The size of the signal spot on the recording plane has a major impact on the shift selectivity of the memory because the spherical reference beam emitted by the VCSEL must originate far enough to be able to completely overlap the signal spot on the material. For example, the large footprint in the transmission module, compounded with the low divergence angle of the VCSELs, makes necessary to position the VCSEL array at a distance of 51 mm, which in turn results in a shift selectivity of 122 µm. In the second design, the fact that the signal spot is smaller allows to position the VCSELs much closer to the medium (the minimum Z_0 is 8.7 mm), lowering the shift selectivity to 21 µm.

Large shift selectivities are very undesirable. As δ increases, so does the spacing required between VCSELs in the array, which compromises their uniformity. One approach to narrow the shift selectivity is by increasing the thickness of the recording medium; however, this also narrows the tolerance of the holograms to differences between the nominal wavelength of emission of the array and the actual wavelength of each individual element. This can be observed in Figure 2-12, where the diffraction efficiency of the hologram is plotted as a function of the thickness of the material when the readout and recording wavelength differ by 0.17 nm, which corresponds to a dispersion of 0.05%. If we want to ensure that wavelength mismatch never accounts for more than 10% loss in diffraction efficiency, the thickness of the material should not exceed 250 µm. This shows that the thickness of the medium plays an important trade-off between shift-selectivity and dynamic range on one side (in which the larger, the better), and tolerance to wavelength mismatch on the other (in which the smaller, the better).

It is important to point out that the advantages of the second architecture over the first one originate not in that the former uses reflection geometry as opposed to transmis-
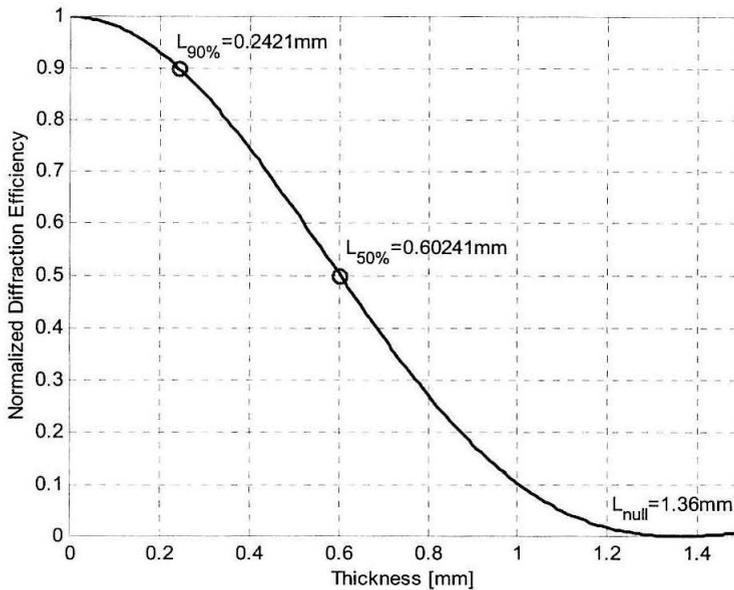
Fig. 2-12. Diffraction efficiency as a function of the thickness of the material in the reflection-geometry OPGA module when the readout wavelength differs in 0.05% with respect to the wavelength used for recording.

sion geometry, but in the difference between Fourier-plane versus Fresnel-region recording. As a matter of fact, a transmission-type architecture using a van der Lugt imaging system is also possible. This case would be particularly interesting because the system would enjoy both small shift selectivity and, at the same time, broader wavelength selectivity than in a reflection module, which means more tolerance to wavelength nonuniformity.

References

[2-1] P. Stogiannos, A. Dollas, and V. Digalakis, "A configurable logic based architecture for real-time continuous speech recognition using hidden Markov models," *Journal of VLSI Signal Processing Systems for Signal Image and Video Technology* **24**, 223-240 (2000).

[2-2] A. Benedetti, and P. Perona, "Real-time 2-D feature detection on a reconfigurable computer", *Proceedings of the 1998 IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, 1998.

[2-3] J. Jean, X. J. Liang, B. Drozd, K. Tomko, and Y. Wang, "Automatic target recognition with dynamic reconfiguration," *Journal of VLSI Signal Processing Systems for Signal Image and Video Technology* **25**, 39-53 (2000).

[2-4] J. P. Kaps, and C. Paar, "Fast DES implementations for FPGAs and its application to a universal key-search machine," *Selected Areas in Cryptography Lecture Notes in Computer Science* **1556**, 234-247 (1999).

[2-5] S. D. Brown, R. J. Francis, J. Rose, and Z. G. Vranesic, *Field-Programmable Gate Arrays*, Kluwer Academic Publishers, Norwell, 1992.

[2-6] S. Trimberger, D. Carberry, A. Johnson, and J. Wong, "A time-multiplexed FPGA," *Proceedings of the IEEE Symposium on FPGA-Based Custom Computing Machines*, 34-40, IEEE Computer Society, Los Alamitos, 1997.

[2-7] M. Motomura, Y. Aimoto, A. Shibayama, Y. Yabe, and M. Yamashina, "An embedded DRAM-FPGA chip with instantaneous logic reconfiguration," *Proceedings of the IEEE Symposium on FPGAs for Custom Computing Machines*, 264-266, IEEE Computer Society, Los Alamitos, 1998.

[2-8] J. Mumbru, G. Zhou, S. Ay, X. An, G. Panotopoulos, F. Mok, and D. Psaltis, "Optically reconfigurable processors," *SPIE Critical Review 1999 Euro-American Workshop on Optoelectronic Information Processing* **74**, 265-288 (1999).

[2-9] T. H. Szymanski, M. Saint-Laurent, V. Tyan, A. Au, and B. Supmonchai, "Field-programmable logic devices with optical input-output," *Applied Optics* **39**, 721-732 (2000).

[2-10] D. Psaltis, M. Levene, A. Pu, G. Barbastathis, and K. Curtis, "Holographic storage using shift multiplexing," *Optics Letters* **20**, 782-784 (1995).

[2-11] G. Barbastathis, M. Levene, and D. Psaltis, "Shift multiplexing with spherical reference waves," *Applied Optics* **35**, 2403-2417 (1996).

[2-12] A. van der Lugt, "Packing density in holographic systems," *Applied Optics* **14**, 1081-1087 (1975).

3 OPGA elements and system integration

3.1 Introduction

After focusing on the system as a whole in the previous chapter, this chapter will look at each one of the elements of the OPGA separately. The following sections consider specific issues of the VCSEL array, the chip and the optical material, and the requirements they have to meet for their use in the OPGA system. Section 3.5 deals with the integration of the three components into a small package and presents a prototype for the OPGA that demonstrates, in a compact module, the operation of VCSELs as addressing devices to readout holograms from the optical memory. Finally, the integration between a holographic memory and a silicon chip is demonstrated in Section 3.6, where holograms are utilized to configure the logic in the chip.

3.2 Red VCSEL arrays

Vertical Cavity Surface Emitting Lasers (VCSELs) operating in the infrared wavelengths are widely used in optical fiber data links, optical interconnects, and storage applications. In contrast to conventional laser diodes, which emit light from the edge of the chip, VCSELs emit light vertically from the wafer surface. Therefore, instead of having to cleave the wafer into single elements, they can be packaged as large arrays [3-1]. The first VCSELs emitting in the red wavelengths were reported in 1993 [3-2]. The shift towards shorter wavelengths has opened the possibility of using such devices in holography, since most optical materials are sensitive in the visible range of the spectrum.

As described in Chapter 2, a VCSEL array can be used as addressing device in the OPGA module to selectively retrieve one data page from the optical memory. Section 3.2.1 describes the requirements the VCSEL array must meet, while Section 3.2.2 presents the results of the experimental characterization of arrays with different VCSEL sizes and element count that have been fabricated by Honeywell.

3.2.1 VCSEL requirements

The two most important parameters for the VCSELs, if they are to be used in the OPGA module, are their output power and wavelength uniformity across the array.

Consider first the wavelength uniformity: As seen in Section 2.4, the holograms are recorded in the memory of the OPGA module using a laser diode. However, they are read out with the VCSELs. Any difference between the recording and readout wavelengths will have a double effect on the reconstruction of the hologram: First, the amount of diffracted light will decrease due to Bragg-mismatch, making necessary longer detector integration times and therefore degrading the reconfiguration performance of the OPGA. Second, the reconstruction of the hologram will shift on the plane of the detector, resulting in pixel misregistration and incorrect programming of the OPGA chip.

These two effects are depicted in the K-sphere diagram in Figure 3-1. A grating (vector $\mathbf{K_G}$) is recorded at wavelength λ (solid circumference) by the interference of the reference beam propagating at an angle θ with respect to the x-axis (vector $\mathbf{K_R}$) and the signal beam. If the hologram is now read out at a wavelength λ' (dashed circumference) by the beam $\mathbf{K'_R}$, the reconstructed signal beam will be Bragg-mismatched by an amount $\Delta\mathbf{K_Z}$ and propagate along $\mathbf{K'_S}$ at an angle α with respect the z-axis.
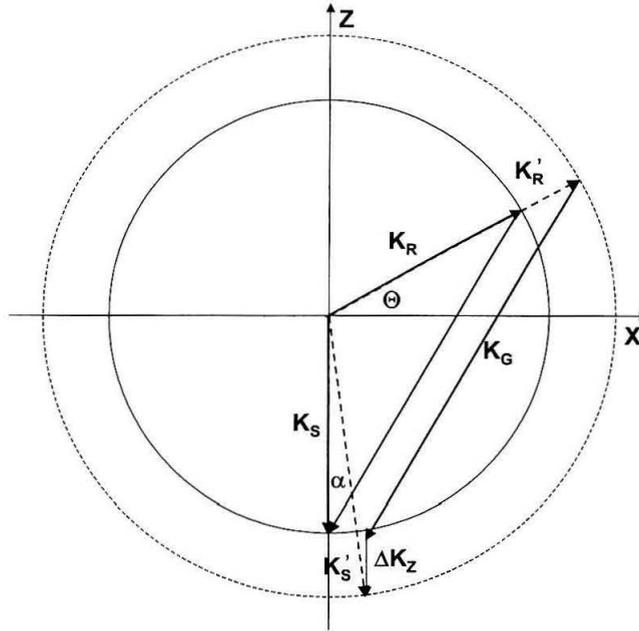
Fig. 3-1.   K-sphere diagram to illustrate the deviation angle α of the diffracted beam K'_S when the grating is read out with a reference beam at wavelength λ' instead of λ.

According to the definition of angles in Figure 3-1, and assuming that the medium has a thickness L along the z-axis and is infinite in the transverse dimensions, the wavelength selectivity of the hologram can be readily calculated and is given by

$$\left(\frac{\Delta\lambda}{\lambda}\right) = \frac{\lambda}{L(1 + \sin\theta)},\qquad(3\text{-}1)$$

where $\Delta\lambda = \lambda - \lambda'$. Furthermore, if the detector plane is perpendicular to the z-axis and located at a distance D from the material, the amount of shift along the x-axis (δx) of the reconstructed hologram on the detector plane results in

$$\delta x = D\tan\alpha \approx D\left(\frac{\Delta\lambda}{\lambda}\right)\cos\theta.\qquad(3\text{-}2)$$

The approximately equal sign is due to $\tan\alpha \approx \sin\alpha \approx \alpha$, which is true only for small $\Delta\lambda$. Equation 3-2 can be rewritten in the form of a constraint on the ratio $\left(\frac{\Delta\lambda}{\lambda}\right)$ for a maximum

acceptable shift $(\delta x)_{max}$ that the system can tolerate without suffering from pixel misregistration

$$\left(\frac{\Delta\lambda}{\lambda}\right) \le \frac{(\delta x)_{max}}{D\cos\theta}. \tag{3-3}$$

In the context of the OPGA module ($\theta=30^{\circ}$, D=10 mm, L=200 μm, and λ=680 nm) the condition in Equation 3-3 becomes more restrictive than the wavelength selectivity. To avoid a maximum shift larger than 20% of a 10 μm pixel detector, $(\delta x)_{max}=\pm2$ μm results in $\left(\frac{\Delta\lambda}{\lambda}\right) \le \pm 0.021$ %. Therefore the ratio of the difference between highest and lowest wavelengths across the VCSEL array, λ_H and λ_L, respectively, to the recording wavelength must satisfy $\left(\frac{\lambda_H - \lambda_L}{\lambda}\right) < 0.05$ %.

As far as the power requirement is concerned, the maximum optical power that the VCSEL can output has a direct impact in the reconfiguration time of the OPGA. It is important to point out that the VCSEL must operate in single mode. The presence of higher-order modes would degrade the quality of the reconstructed holograms as the wavefront of the beam emitted by the VCSEL would not match the wavefront of the laser diode used during recording.

The incident power required per VCSEL (P_{Inc}) depends on parameters of the holographic memory and the photodetectors. The former can be either related to the medium itself, like the dynamic range, or to the design of the optical memory, like the number of holograms sharing the same spatial location (M) and the number of pixels in each hologram (N $\times$ N). Among the latter, there is the integration time of the detectors (T_{Int}) and the minimum number of photons for correct detection (#photons), which in turn is a function of the detector quantum efficiency and the threshold number of signal electrons required at the

detector. The required incident power can then be parametrized as a function of the integration time (OPGA reconfiguration time) and described by

$$P_{Inc} \geq (\#photons) \cdot \left(\frac{hc}{\lambda}\right) \cdot N^2 \cdot \left(\frac{M/\#}{M}\right)^{-2} \cdot \frac{1}{T_{Int}}. \tag{3-4}$$

As described in Chapter 2, each location on the material contains only 20 holograms (M=20) of 1000×1000 pixels each ($N^2=10^6$). The number of photons required is a function of the integration time. As the integration time increases, the number of collected noise-electrons increases as well, and therefore more signal electrons (photons) are required for correct detection. Testing of the photodetectors has revealed that up to 50 μsec integration time, 2600 photons are required for correct detection [3-3].
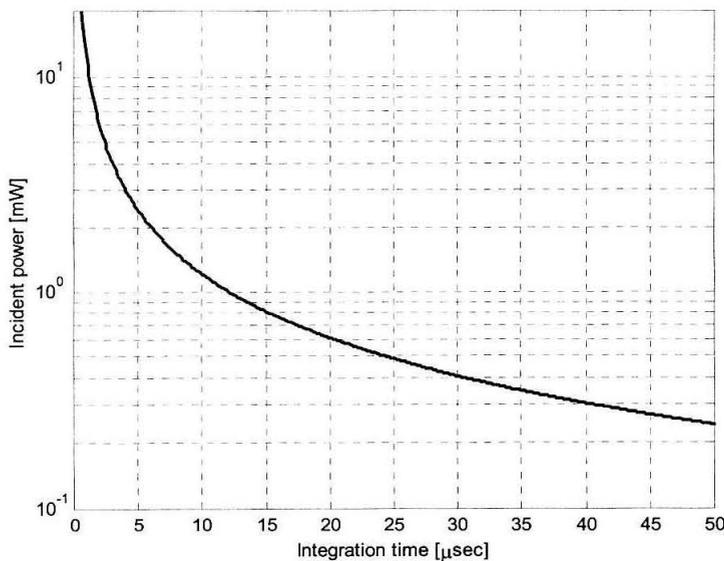
Fig. 3-2.   Minimum power emitted by the VCSEL for correct signal detection as a function of the integration time of the photodetectors.

The dependence between power and integration time described in Equation 3-4 is presented in Figure 3-2. As it can be observed, 25 μsec integration time can be attained if

each VCSEL is able to output 0.5 mW in single mode. Although shorter reconfiguration

times could be achieved if the VCSEL was able to emit a few milli-Watts of power and, for

example 5 mW in single-mode would make possible an integration time of just 2.5 μsec,

the value of 0.5 mW per VCSEL will be taken as a more realistic power requirement.

3.2.2 Array characterization

Initial experiments to characterize red VCSELs and to verify the suitability of this

type of laser diodes for applications in holographic systems were performed on a 4 × 4 sym-

metric array with a 50 μm spacing between elements.

The VCSELs operate in single mode and produce a spherical beam with a full diver-

gence angle of 8.7°. When switching on the VCSELs, the rise time is less than 100 ps,

which allows switching speeds over 1 GHz. Their output power is on the range of hundreds
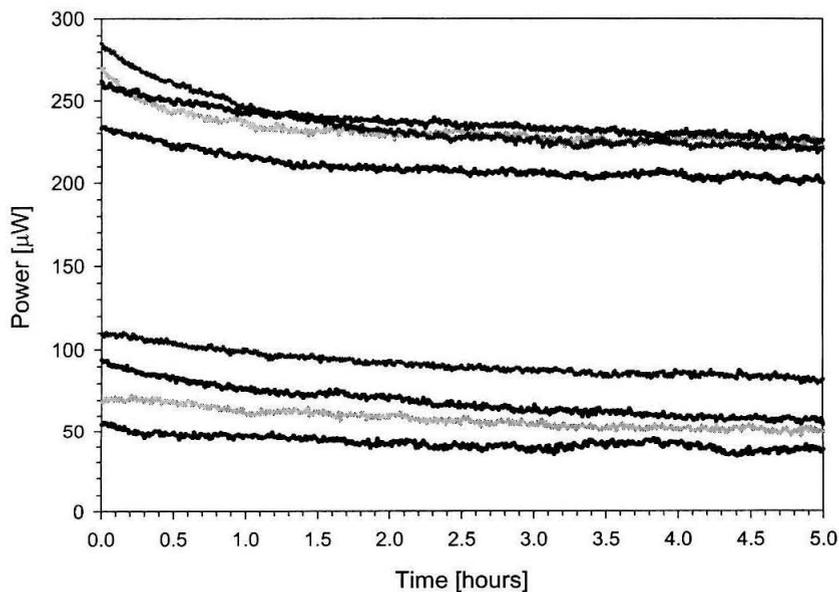
Fig. 3-3.   Temporal evolution of the power emitted in continuous mode by the VCSELs over a
period of five hours. Each curve correspond to a different element of the VCSEL array.

of μWatts, from 50 μW for the worst elements up to almost 300 μW for the best ones. The

stability over time of the output power has also been studied. The power emitted by the

VCSELs has been monitored over a period of 5 hours. Figure 3-3 shows the measurements

for eight different elements in the array. As it can be observed, there is an initial drop in

power due most probably to thermal heating of the P-N junction of the lasers. The long-

term stability measured after a one hour warm-up over the last 4 hours results to be on aver-

age 18.6%, although in the worst case it is as large as 32.9%. A more realistic parameter is

the short-term stability over a 15 minute interval after the initial warm-up, which turned out

to be on average 4.8% and never larger than 9.2%.
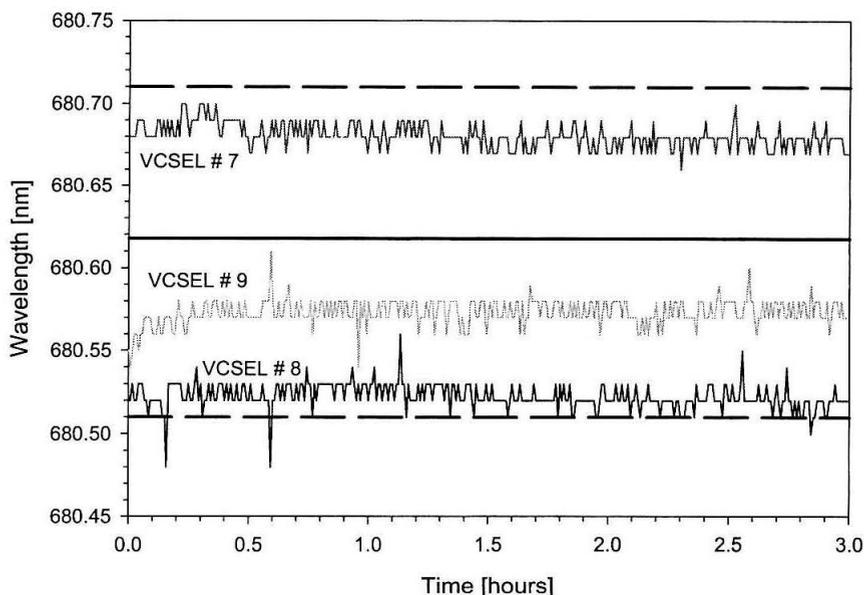
Fig. 3-4.   Drift in wavelength of emission of the VCSELs monitored for a period of three hours.
The solid line corresponds to the ensemble-averaged wavelength across the array, while the dashed
lines indicate the maximum and minimum time-average wavelength for an individual laser.

The wavelength of all the operating devices in the array has been measured (Figure 3-4). The average wavelength across the entire array is 680.62 nm, with the highest and lowest wavelengths being 680.71 nm and 680.56 nm, respectively. This dispersion of values of just 0.15 nm across the array corresponds to a uniformity of 0.022%, in compliance with the specifications.

The fluctuation of the wavelength for individual elements with time has also been investigated, and it has been found to be smaller than 0.016%. This very good stability, even without any thermal control of the VCSEL, and a coherence length better than a meter, makes the VCSELs adequate to record and readout holograms, as shown in Figure 3-5.

Fig. 3-5.   Hologram of a 34 μm pixel-size mask recorded on 100 μm thick Du Pont photopolymer using VCSELs.

The characterization of the 4 × 4 array revealed that, in order to use VCSELs in the OPGA module, the amount of optical power emitted by each device needed to be augmented and that the good wavelength uniformity in the small array might be difficult to preserve as the size of the array was increased. Both issues were taken into consideration when specifying larger arrays of 25 × 1 VCSELs. Figure 3-6 shows the die containing columns

of 25 VCSELs each, with a 100 μm pitch. The devices where fabricated in three different

diameters: 20 μm, 15 μm, and 10 μm.
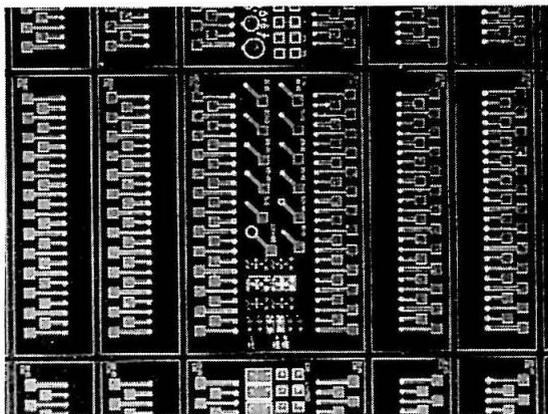
Fig. 3-6.   Picture of the die containing several 25×1 arrays of red VCSELs, arranged in columns.
The devices are the small circles at the end of the square pads.

The optical power vs. current curve has been measured for the three sizes of

devices. Figure 3-7 shows the typical results for a 10 μm, 15 μm and 20 μm cell. It was

observed that as the size of the VCSEL is augmented, so are its threshold current and opti-

cal power. However, the device becomes multimode at lower levels of current. The 10 μm

VCSEL could not produce more than 0.34 mW single-mode on average across the entire

array. On the other hand, the 20 μm VCSEL becomes multimode slightly above its thresh-

old current at 0.4mW, despite the fact that it can produce more than 1 mW. The 15 μm

VCSEL turned out to be the one that is able to put out the highest power in single-mode,

around 0.58 mW on average across the array, satisfying the power requirement for the

OPGA module.

The wavelength stability of individual VCSELs was about 0.012%. As far as the

wavelength uniformity across the 15μm-VCSEL array is concerned, the wavelength varies
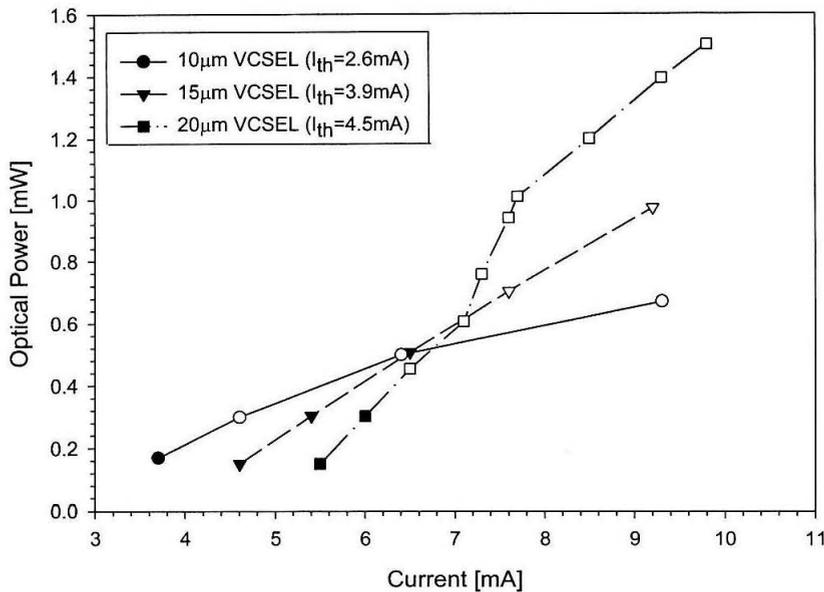
Fig. 3-7.   Typical Optical power vs. current characteristic curve for the 10 μm (circles), 15 μm (triangles) and 20 μm (squares) VCSEL. For each curve, the change from solid to clear markers indicates the transition in the VCSEL beam profile from single mode to multimode.

form 677.87 nm to 677.46 nm, being 677.65 nm the array-average wavelength. This results in a uniformity of 0.06%.

3.2.3 Conclusions

Improvements in the fabrication process of the 25×1 arrays with respect to previous generations of devices have made possible to increase their output power and meet the desired target of 0.5 mW single-mode. However, the wavelength uniformity has proved to be difficult to preserve in larger arrays as the physical distance between VCSELs increases. Despite the fact that Honeywell has been able to produce some arrays with a level of uniformity better than 0.05%, the consistency from die to die seems to still require further research in the fabrication process.

The degradation of the wavelength uniformity as the size of the VCSEL array increases is the main cause that prevents VCSELs from being used as the addressing device in the OPGA system, and forces to consider other alternatives instead, like having a MEMS array of mirrors deflecting the beam of a single laser diode.

3.3 OPGA chip

The development of pixel sensors using standard CMOS technology [3-4], the same that is used for most microprocessors and memory modules, enables the integration of photodetectors with on-chip processing circuitry and has coined the expression "camera-on-a-chip" [3-5]. The OPGA chip makes use of the system-on-a-chip idea and integrates on the same die an array of pixels to detect the reconstructed holograms, as well as the logic circuitry of a conventional FPGA.

The photodetectors can be combined with the existing logic of the FPGA using either one of the two different topologies depicted in Figure 3-8: the detectors can be sparsely distributed across the whole chip interleaving them with the logic, or conversely, they can be laid out as a single large array on a specific region of the chip.

From the electronics point of view, the first topology is more convenient, because each pixel is detected exactly where it is needed to program the logic element. This makes unnecessary to distribute the detected signals all across the chip. However, from the optics side, to have detectors spread over the entire chip means that the quality of the reconstructed hologram must be uniform over a much larger area. Therefore, the second topology makes the optics simpler because the hologram needs to be uniform in a smaller region.
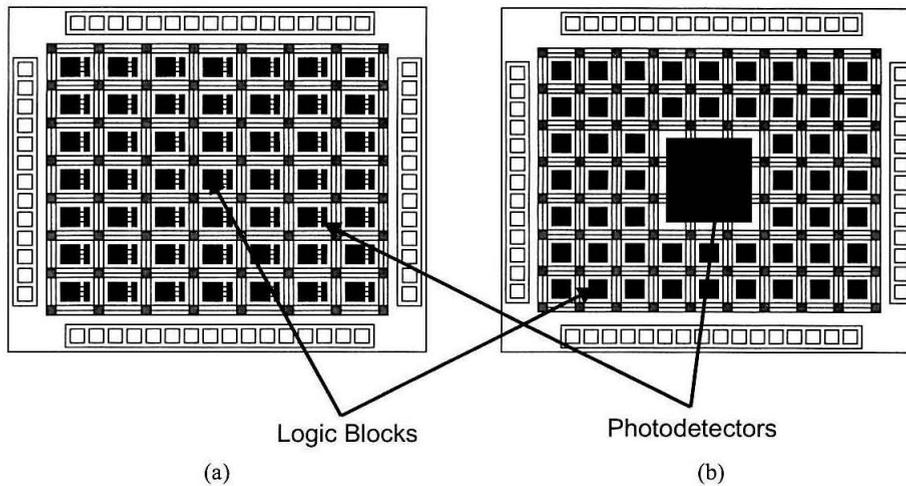
Fig. 3-8. Detector distribution on the OPGA chip: (a) Sparse, where the photodetectors are interleaved with the logic; or (b) Concentrated, where all the photodetectors are implemented as an array and detected signals need to be delivered to the programmable elements.

However, this comes at the price of having to implement a more complex mesh of buses to deliver the detected signals to the logic blocks.

Since each one of these topologies presents interesting aspects, both of them have been considered in the OPGA chips that have been designed and fabricated. In the first two generations of the chip, the photodetectors were arranged in small arrays and integrated into each logic block. However, the concentrated topology has been adopted in the final design because it simplifies the interface with a regular SLM.

3.3.1 Differential encoding

The light detected by each pixel needs to be converted into a logic value "1" or "0" by comparing its voltage to some threshold. The simplest way to perform such conversion is to set the same threshold to all the photodetectors in the chip. However, a global threshold cannot compensate for spatial variations of intensity across the entire data page. An alternative is to use different threshold levels across the area of the chip. This is not a perfect

solution either, even assuming that generating many different bias voltages for the thresh-olds is not an issue, because the spatial nonuniformity in the reconstructed hologram can change from one holographic page to another.

A very efficient way to be more robust to the intensity nonuniformity of the recon-structed data page is differential encoding [3-6], [3-7]. In this case, a pair of pixels in the hologram represents each single bit of information required to program the chip. The dif-ferential photodetector must have two photosensitive areas, referred to as the left and right pixels, which need to be matched to the pixel pair in the hologram. The logic "1" is then represented by left pixel ON and right pixel OFF and logic "0" by left pixel OFF and right pixel ON. This coding scheme makes it unnecessary to set any threshold for the photode-tectors. Since the global variation of the incident illumination is reduced, the signal-to-noise ratio is increased and therefore the bit-error rate is improved. From the optics point of view, this type of data representation is simple and does not increase the system cost.

3.3.2 Photodetector structure

CMOS pixel sensors are based on either a passive or an active structure. Passive pixel sensors (PPS), depicted in Figure 3-9(a), consist of just a photosite that converts pho-tons into electrons. The generated photocurrent discharges the reset voltage stored in the floating diffusion of the pixel. The analog pixel value is carried off the pixel to the support circuits that will amplify and digitize it. The simplicity in design of the passive pixel makes possible to dedicate most of the area of the pixel to the collection of light, obtaining high fill factors. However, this simplicity also makes the PPS more vulnerable to noise. On the other hand, active pixel sensors (APS), Figure 3-9(b), incorporate additional transistors
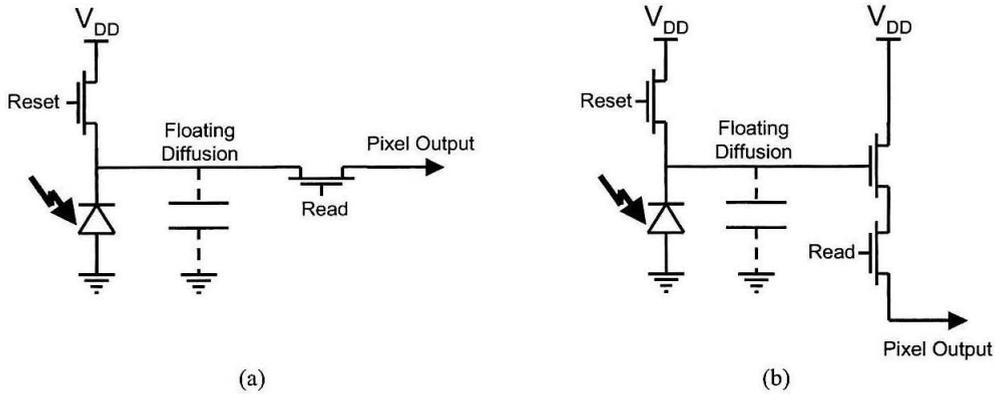
Fig. 3-9. Circuit schematic of (a) a passive pixel sensor, and (b) an active pixel sensor that includes a source-follower amplifier.

inside each pixel, like the source follower amplifier in Figure 3-9, to amplify the photo-generated signal and reduce noise. However, adding these components reduces the fill factor of the APS.

In the OPGA chip, the main two requirements on the photodetectors are their size and speed. The pixel sensors must have small pitch to result in a low overhead in silicon area. The smaller size of passive pixels makes them interesting for this application. But at the same time, the pixels must have enough sensitivity to guarantee a short integration time in the range of 1–100 μsec. This second requirement is better satisfied by active pixels, which reject better noise and crosstalk.

In order to find out which type of CMOS sensor is the most adequate for the OPGA application, and in particular to investigate the possibility of utilizing passive pixels, a first prototype chip has been designed and fabricated by Photobit using 0.35 μm CMOS process. The chip contains a 2 × 3 array of optically addressable logic blocks. The architecture of the logic block, Figure 3-10, is based on a 4-input 1-output LUT that can be programmed
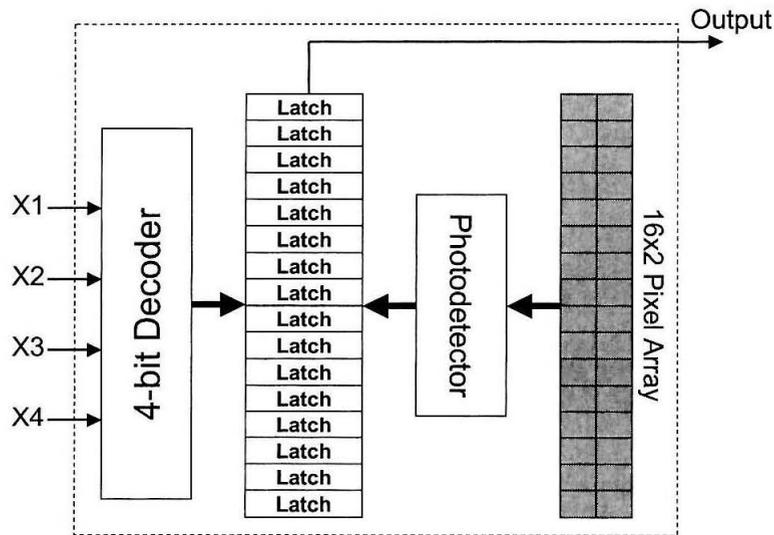
Fig. 3-10. Architecture of an optically addressable 4-input 1-output LUT. During the programming cycle, the array of 16 differential pixels time-share the photodetector. The detected bits are stored into the latches, and can be accessed during normal operation via the decoder.

by an array of 16 differential pixels with 5 μm pitch implemented inside the logic block. The areas of the OPGA chip that are not light sensitive have been shielded with metal to avoid stray radiation interfering with the logic circuit.

The optically addressable CLB in Figure 3-10 consists of a 4-bit decoder and a memory bank formed by an array of 16 latches to store the value of the computation, as in a conventional LUT; but dislike an "all-electronic" LUT, the logic block incorporates a 16×2 pixel array and a shared photodetector. The size of each CLB is about 125×85 μm, and the overhead due to optoelectronics is 24.8%. A detailed block diagram of the detection chain is provided in Figure 3-11. During the programming cycle, the photodetector will sequentially interrogate the 16 pixel pairs to read out the analog value in the left and right pixel of the pair. These voltages are taken to the first amplification stage, where a common source amplifier amplifies each channel separately before the signals are fed to the second

stage of the photodetector: a differential amplifier/comparator that produces the digitized

output that is delivered to the corresponding latch in the LUT.
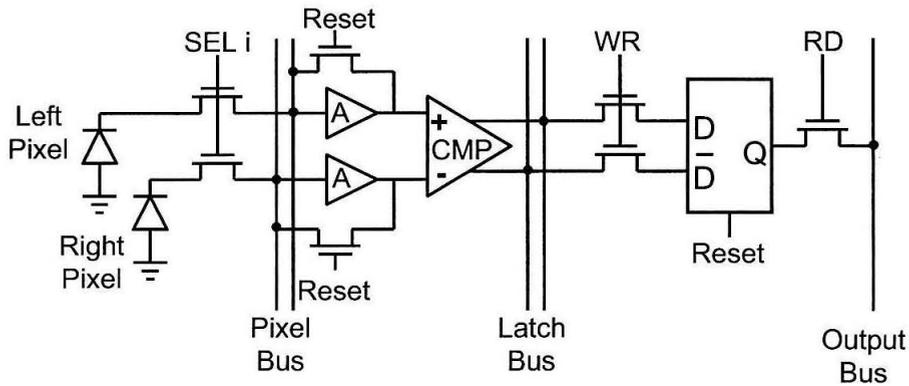
Fig. 3-11.   Block diagram of the detection chain. The first stage consists of two common source amplifiers (A) to amplify the value of each pixel in the pair separately. The second stage is a differential amplifier and comparator (CMP) that produces a 1-bit digitized output.

The preliminary testing of the passive-pixel chip showed satisfactory operation of

the logic blocks at low clock speeds around 60 KHz. The performance of the chip degrades

at higher clock speeds, as the electron injection noise in the pixel buses becomes more seri-

ous and leads to problems like blooming and crosstalk, which result in erroneous program-

ming of the LUTs. The lesson learned from the characterization of the first-generation chip

was that despite the potential savings in real estate obtained when using PPS, the chip

requires a higher level of noise suppression for high-speed operation only achievable with

the use of the APS. This is the reason why an active pixel solution was adopted for the

second and third generation chips, the latter also referred to as full OPGA chip and

described in the next section, also fabricated by Photobit.

3.3.3 Full OPGA chip design

The final version of the OPGA chip [3-3], Figure 3-12, mimics a small-scale FPGA. The chip combines a 64×32 array of differential APS sensors (the big block on the left in the picture); and the logic array (the small block on the right in the picture) containing the logic blocks and interconnecting resources.
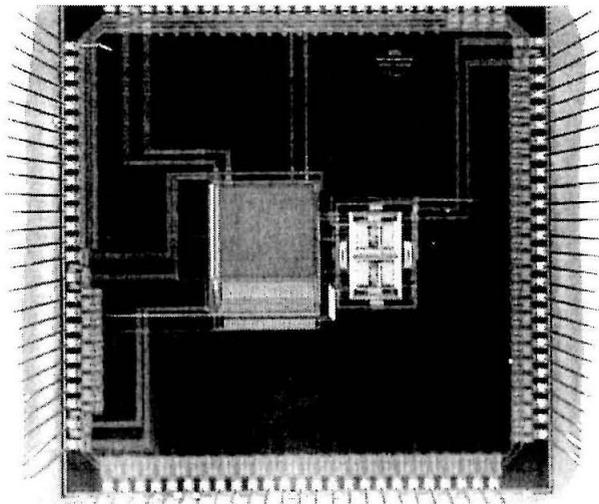
Fig. 3-12.   The full OPGA chip, designed by Photobit in 0.35 µm standard CMOS process, integrates a 64×32 array of differential photodetectors (the large block on the left in the picture) and fully connected logic array (the small block on the right).

The programmable logic array contains four logic blocks in a 2×2 arrangement, and an interconnection network based on five crossbar switching matrices and four connection matrices, and four 5-bit tri-state I/O ports. A schematic of the logic array is provided in Figure 3-13. The architecture of the CLB is based on a 5-input LUT with 1-bit buffered output. The crossbar switching matrices (S-Boxes) allow for connectivity between two segments of bus lines or between the buses and the I/O ports. On the other hand, the connection matrices (C-boxes) establish connectivity between segments of bus and the inputs of the logic blocks.
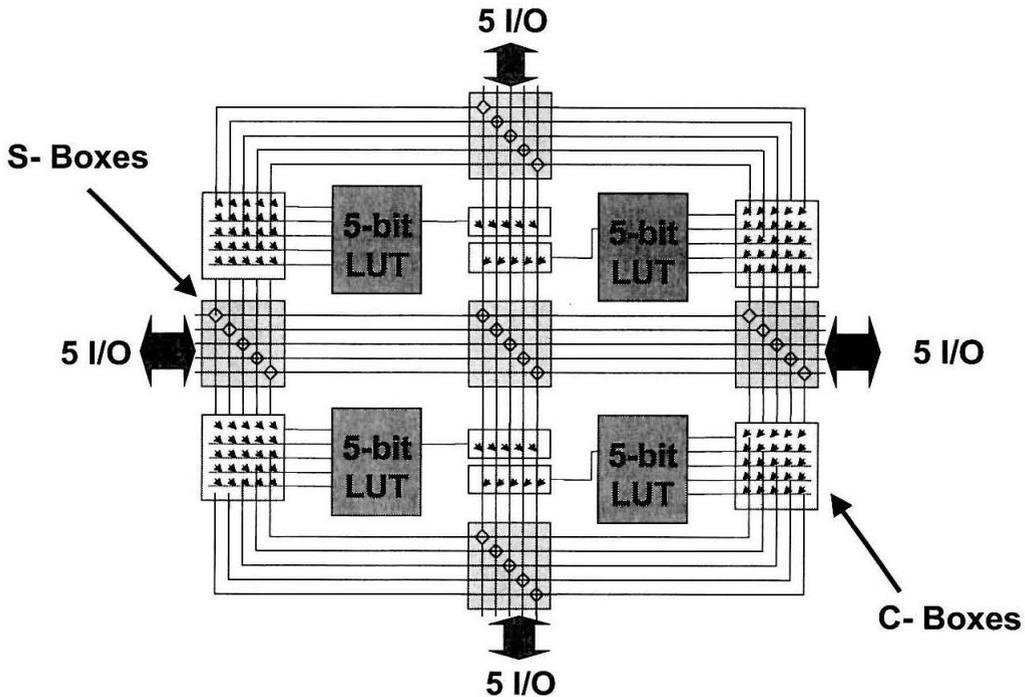
Fig. 3-13. Schematic of the programmable logic array. The logic circuitry consists of a 2×2 array of 5-input LUTs with one buffered output each. The LUTs are fully interconnected by means of 5 switching matrices located in the center, left, right, top, and bottom of the array.

The APS array has been designed with pixel size of 15 μm. Although smaller sizes can be fabricated, the 15 μm sensor size has been selected to match the pixel size of the Kopin SLM that will be used to record the holograms in the optical memory. The detection process to read out the analog signal in each APS pair and convert it into a 1-bit digitized output is performed in parallel row-wise by an array of 32 detectors, physically located on the bottom of the APS array in Figure 3-12, and time-shared within each column of the APS array. The block diagram of the photodetector is essentially the same as the one depicted in Figure 3-11, with the difference that the passive sensors have been replaced by active ones.

Once a 32-bit control word has been detected, it can be used to program the on-chip logic. A total of 13 words, or equivalently rows of the APS array, are necessary to configure

the entire logic of the OPGA, therefore 4 different configurations can be downloaded simultaneously into the chip. This feature makes it possible to switch between configuration contexts within the same hologram and implement partial reconfiguration of the OPGA chip.

3.4 Optical materials

After deciding upon the mechanism to store the configuration templates in the holographic memory of the OPGA, it is not less important to consider which optical media are appropriate for the OPGA system. One of the basic requirements on the optical material is its M/#. The medium must have enough dynamic range to support 100 holographic reconfiguration templates. As discussed in Section 2.4, a material with at least M/5 would be desirable. It is worth making the remark that this M/# must be achieved not just for plane wave holograms but for high-bandwidth data pages as well. In general, thicker media are preferred because they offer larger dynamic range and better selectivity, which allows for more dense multiplexing. However, in this particular application, a too thick material would result in a very narrow wavelength selectivity, making even more stringent the requirement in wavelength uniformity of the VCSEL array.

In order to make possible the recording of high-bandwidth data pages, the holographic material should exhibit very good optical quality and uniformity, low scattering and not undergo significant dimensional changes during recording, like shrinkage in certain photopolymers. Finally, the material should also have a broad modulation transfer function (MTF) to provide enough flexibility to work in both transmission and reflection geometries.

Since the optical memory in the OPGA is intended to operate in a read-only mode, many different polymer-based materials can be used. In these media, readout of the stored data does not result in erasure, therefore, the lifetime of the holograms is only limited by the aging of the polymer itself. Holographic polymers are interesting because they exhibit very good dynamic range and also have high sensitivity. For example, Du Pont photopolymer has been used in early experiments in the OPGA, as well as in the system demonstration described in Section 3.6. However, this polymer suffers from shrinkage and poor optical quality due to nonuniformity of the material, which distorts the reconstructed images. This problem becomes more important as the pixel size is reduced, even if phase-conjugate readout is used. Therefore, Du Pont photopolymer is not the best option for a full scale system. Another alternative that has been considered is a red-sensitive acrylamide-based PVA material (see Section 3.4.1). This medium is fairly similar to Du Pont photopolymer and in fact suffers from many of the same problems as the Du Pont films.

The possibility of using PQ-doped PMMA has also been explored. This material shows good optical quality and M/# [3-8], but it has extremely poor absorption in the red, making necessary to use green wavelength laser sources in the OPGA module. A full description of the mechanism of grating formation in PQ-PMMA, along with experiments to optimize the performance of the material, are provided in Chapter 5. Nevertheless, the most solid choice seems to be the Aprilis film [3-9]-[3-11], since it enjoys both good recording dynamics and high optical quality. Chapter 4 presents the experimental characterization of the material and its use in holographic data storage applications.

A list of optical materials would not be complete without taking into consideration the photorefractive crystals. Iron-doped lithium niobate crystals are another good option for

the OPGA (see Section 3.4.2), and an advance packaging design that uses a $LiNbO_3$ crystal is described in Section 3.5 and has been demonstrated in [3-12]. The volatility of the stored information during readout, or even in the dark, has been traditionally seen as a drawback in this class of materials, however, recent research has successfully developed nonvolatile re-writable memories using doubly doped crystals [3-13], [3-14].

	Du Pont	**Aprilis**	**$LiNbO_3$:Fe**	**MB-PVA**	**PQ-PMMA**
Thickness	20-100 µm	200-500 µm	cm	80 µm	1-10 mm
M/#*	4.5 (100 µm)	10 (200 µm)	5 (8 mm)	1 (80 µm)	4.8 (3 mm)
Sensitivity: 1% energy Saturation	$\sim$1 mJ/cm^2 300 mJ/cm^2	1 mJ/cm^2 300 mJ/cm^2	>100 mJ/cm^2 N/A	20 mJ/cm^2 40 mJ/cm^2	12 mJ/cm^2 $\sim$1 J/cm^2
Scattering	5×10^{-4}/strad	10^{-5}/strad	10^{-6}/strad	$\sim$ Du Pont	10^{-4}/strad
Optical Quality	Good	Very good	Excellent	Good	Very good
MTF	Good	0.2-10 µm	Good	--	Good
Shrinkage	$\sim$3.5%	$\sim$0.05%	N/A	$\sim$ Du Pont	N/A
Post-processing	Curing	Curing	N/A	--	Baking
Volatility	No	No	Yes w/o fixing	No	No
Lifetime: Shelf life Data retention	Months Years	Months >10 years	Years Years	-- --	Years Years

Table 3-1. Comparison of the main properties of available holographic materials that have been considered for the OPGA. (*M/# measured for the actual thickness of the material)

Before presenting the experimental characterization of each one of the materials and discussing of their particularities, Table 3-1 provides a comparative overview of the main properties of the different holographic media that have been considered for the OPGA module.

3.4.1 MB-doped acrylamide-based PVA polymer

The repertory of red-sensitive holographic media has been increased with the development of an acrylamide in polyvinyl alcohol (PVA) photopolymer film. The use of acry-

lamide material for holography was first published in [3-15]. High diffraction efficiency holograms using acrylamide in PVA solution films were reported in [3-16]. This particular material is doped with 3,7-bis(dimethylamino)phenothiazin-5-ium chloride, usually called methylene blue (MB), which is a chromophore sensitive to red wavelengths.

Films of MB-PVA polymer have been synthesized to characterize its performance. Sample preparation consists of dissolving the acrylamide, along with the MB dye and a polymerization initiator, triethanolamine, in an aqueous solution with 10% wt of PVA. The concentrations of monomer, dye and initiator were chosen to match the values for material optimization published in [3-17] and summarized in Table 3-2. The solution was poured onto glass plates and let dry in the dark to obtain films of 80 μm in thickness.

Component	Concentration
Acrylamide	3.4×10^{-1} M
Triethanolamine	2.0×10^{-1} M
Methylene blue	2.4×10^{-4} M

Table 3-2. Concentrations of the optimal composition of the photosensitive PVA solution. Concentrations are in Molar (i.e., Mol/l)

The absorption spectrum of one of the samples, Figure 3-14, shows that the material is highly absorptive in the region of the red wavelengths, resulting in an optical density of 1.17 for an 80 μm thick film at 633 nm.

The photochemistry of this material is fairly similar to the one of Du Pont polymer. Exposure of the sample to red light excites the dye molecules, which in turn activate the molecules of the initiator, triggering the free-radical polymerization of the acrylamide. Holographic recording is achieved by illuminating the sample with an interference pattern. The photoinduced polymerization of the acrylamide in the bright fringes, combined with
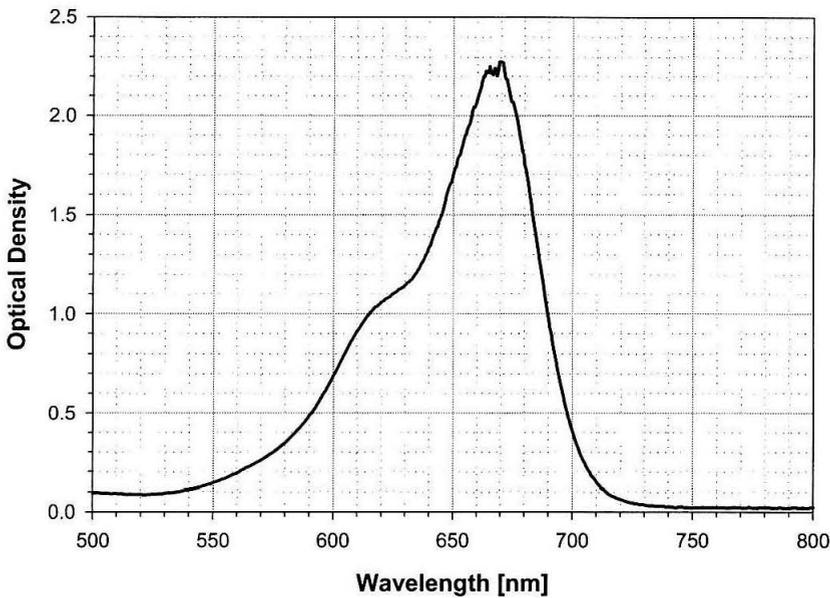
Fig. 3-14.   Absorption spectrum of an 80 μm thick MB-PVA film.

the diffusion of the monomer from the dark into the bright regions, results in significant

local changes of the index of refraction.

Recording experiments on this material have been performed using a He-Ne laser

in a symmetric transmission-geometry setup. The angle of incidence of each one of the

beams (outside the material) is 30° with respect to the surface normal of the sample. The

growth of the hologram to saturation is plotted in Figure 3-15. The recording curves satu-

rate after 250 mJ/cm^2 of exposure, achieving a level of diffraction efficiency about 35%.

The square root of these curves has been fitted to saturating mono-exponentials to deter-

mine the recording "time" constant. This constant resulted to be 99.5 mJ/cm^2 on average,

which corresponds to a sensitivity of 791 cm/J.

The material seems to require about 10 mJ/cm^2 of pre-exposure energy before there

is any growth in the diffraction efficiency of the hologram. This agrees well with the results
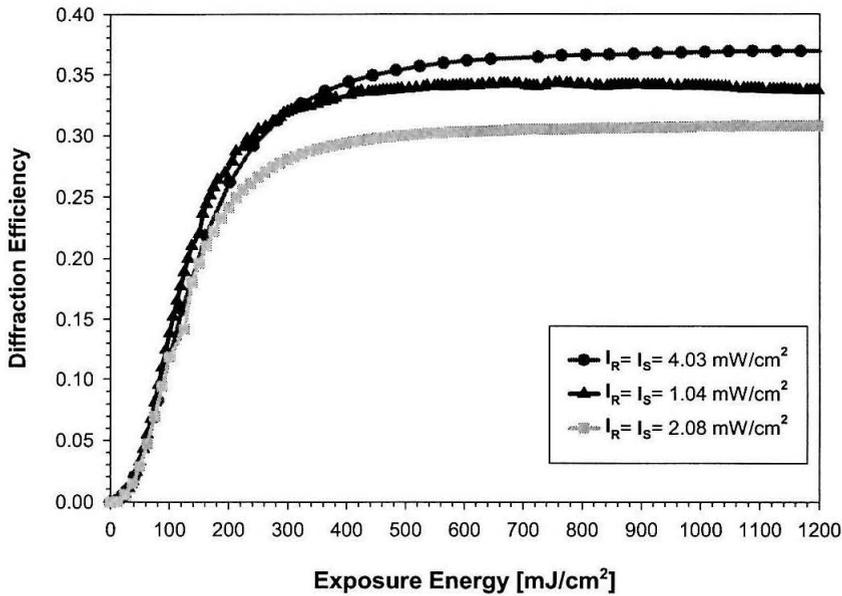
Fig. 3-15.   Hologram diffraction efficiency as a function of the recording exposure energy.

from the multiplexing experiments in which the early holograms are consistently weaker, like the first two holograms in Figure 3-18(a) and Figure 3-19.

The recorded holograms presented good quality and rendered nice selectivity curves, like the one presented in Figure 3-16. The measured angular selectivity turned out to be 0.74° (outside the material). This very wide angular selectivity is fully consistent with the 80 μm thickness of the samples.

Multiplexing experiments have been performed in order to measure the M/# of the material. The small thickness of samples, which results in large angular selectivity, greatly limits the number of holograms that can be angularly multiplexed. Alternatively, peristrophic multiplexing [3-18] can accommodate more easily a larger number of co-located holograms. The sample is mounted on a rotation stage in the experimental setup, so it can be rotated around the normal to its surface between successive exposures. As the
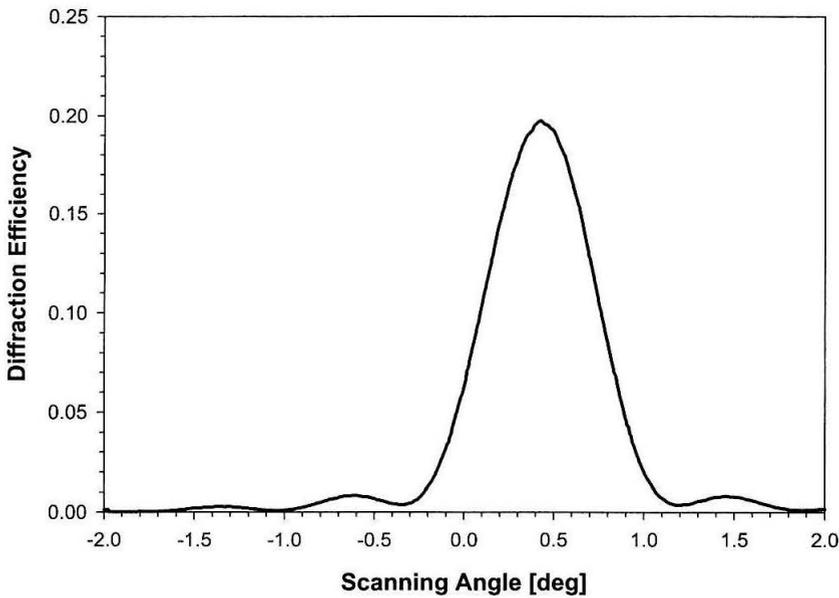
Fig. 3-16.   Angular selectivity curve for a hologram recorded in an 80 μm thick MB-PVA film.

sample rotates, the direction of propagation of the diffracted wave changes, which can then

be filtered off the detector, making it possible to record a new hologram without having

crosstalk from the previous one. In general the sample can be rotated 360°, however the

degeneracy of the symmetric recording geometry reduces that range to just 180°.

Figure 3-17(a) plots the diffraction efficiencies of 31 multiplexed holograms. Each

recording beam had 1.04 mW/cm^2 in intensity and a 7 mm diameter. The exposure energy

was 17 mJ/cm^2 for all the holograms, ensuring that the dynamic range of the sample is used

up after approximately 530 mJ/cm^2. The cumulative grating strength of the holograms is

then computed in Figure 3-17(b). The M/# can be estimated from the saturation value of

this curve, being M/0.91 for the actual thickness of the material.

The M/# can be optimized by equalizing the strengths of the holograms using a

recording schedule. A schedule was calculated by fitting a sixth order polynomial to the
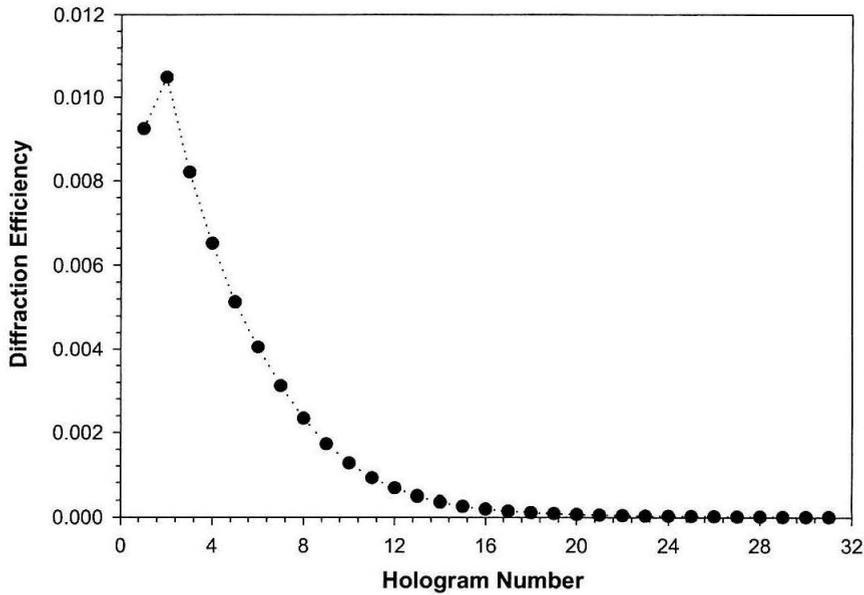
(a)
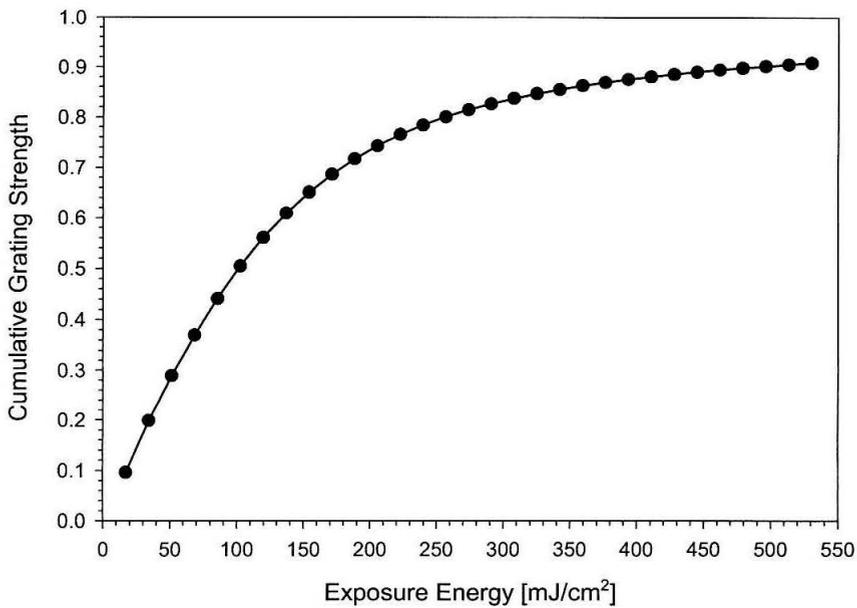
(b)

Fig. 3-17.   (a) Individual hologram strengths of 31 holograms multiplexed with equal exposures, and (b) their cumulative grating strength. The solid line in (b) is a sixth order polynomial fit.

experimental data in Figure 3-17(b) according to the method described in [3-19]. The

scheduled recording experiment produced more equal holograms, Figure 3-18(a), with an

average diffraction efficiency $\eta = 9.6 \times 10^{-4}$. The cumulative grating strength curve in Figure 3-18(b) saturated at M/0.94, which is slightly higher than in the previous case for constant exposure.

The optimization process can be iteratively refined by computing a new recording schedule for the next iteration based on the results from the previous one. As the holograms become more equalized, the cumulative grating strength should saturate at a higher value, indicating that the dynamic range of the material is used more efficiently. Although better equalization has been achieved in the experiments, Figure 3-19, it has not necessarily been coupled to higher M/# (M/0.81 in Figure 3-19, with an average $\eta = 7.5 \times 10^{-4}$). This is most probably due to differences in the thickness of the samples

The material presents a good M/# for its small thickness. Further multiplexing experiments produced a maximum M/# of 1.3. However, the large absorption of the material at 633 nm leaves little room to increase the dynamic range by making thicker samples. The consistency of the results throughout different experiments was poor due to sample-to-sample variance. Further development of the MB-PVA material is necessary if this material is to be used for holographic data storage applications, like in the OPGA system.

Another aspect that needs consideration is the image quality of the reconstructed holograms. Shrinkage during recording is expected in this material due to the process of polymerization induced by illumination with the laser beam. In that sense, the fact that the angle of incidence of the probing beam needed to be detuned to Bragg-match the hologram in Figure 3-16 is an indication of the presence of shrinkage. Other issues like material uni-
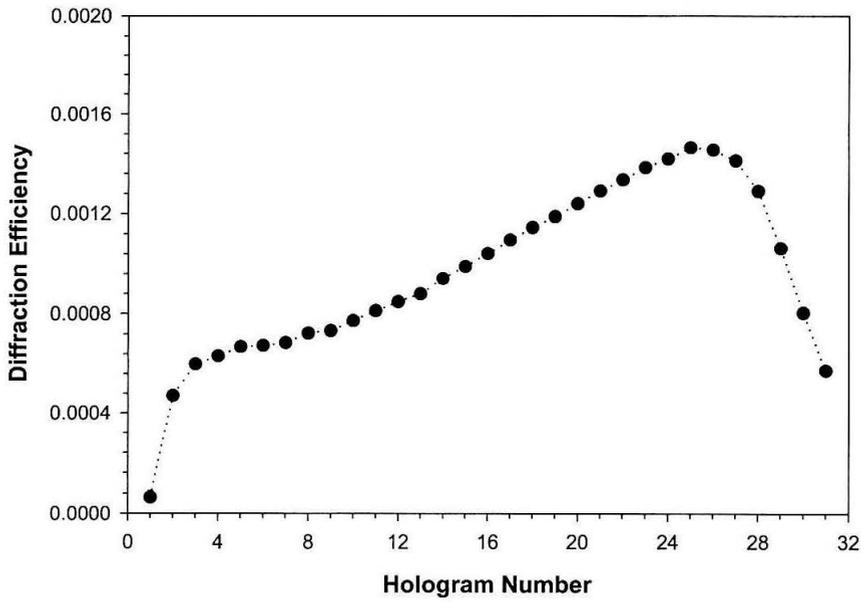
(a)
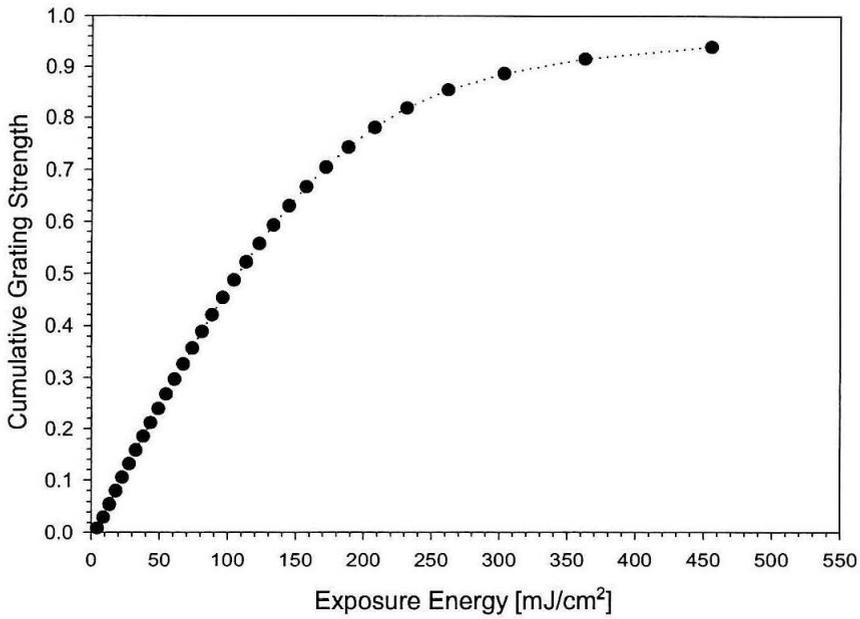
(b)

Fig. 3-18.   (a) Individual strengths of 31 multiplexed holograms recorded with an exposure schedule and (b) their cumulative grating strength.
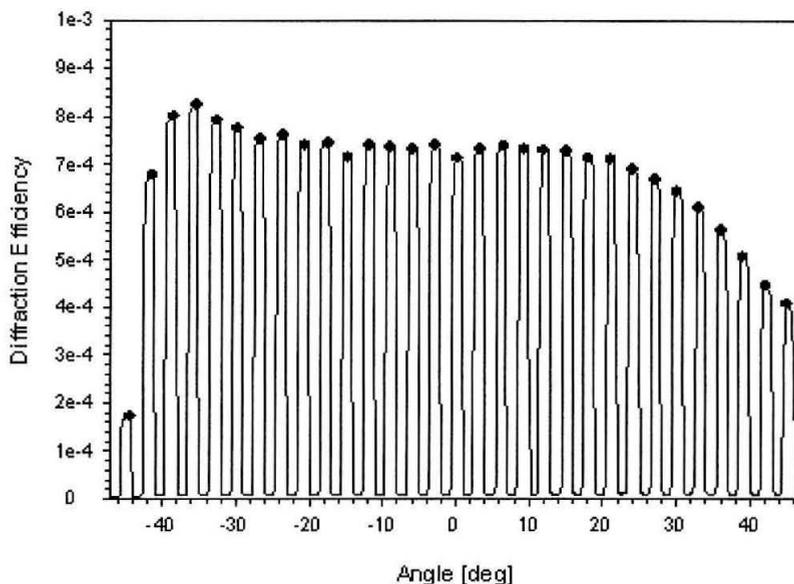
Fig. 3-19.   Comb function of 31 peristrophically multiplexed holograms recorded with exposure schedule.

formity and scattering were not investigated, although no significant improvements with respect to Du Pont photopolymer are expected.

3.4.2 Lithium niobate crystals

Photorefractive crystals have been widely used as support for holographic data storage applications [3-20], [3-21]. Unlike the other materials compared in Table 3-1, hologram formation in $LiNbO_3$ is achieved by means of the photorefractive effect. During recording, the interference pattern of two light beams excites electrons in the bright fringes into the conduction band of the crystal. At the end of their lifetime, the excited electrons return to either the valence band or to centers in the bandgap created by dopants like iron, where they can be promoted again into the conduction band. However, the electrons that recombined into centers in the dark fringes of the interference pattern can no longer be excited and remain trapped. The redistribution of charges in the crystal, and their eventual

trapping, builds a spatially variant electric field that induces in turn the modulation of the refractive index of the crystal through the electro-optic effect, known as Pockel's effect. Of course, the process of index modulation is reversible, so uniform illumination will erase any local distribution of electric charges.

The recording sensitivity in $LiNbO_3$ crystals is between one and two orders of magnitude lower than in photopolymers, making the recording process more slowly. However, this is relatively unimportant in the context of the OPGA, because the OPGA module is intended to operate as a read-only memory and therefore the recording of the holographic configuration templates occurs off-line. Furthermore, unlike photopolymers, $LiNbO_3$ crystals offer the possibility of implementing a fully read-write module. A holographic RAM type of memory [3-22] makes it possible to update the set of configuration templates that the OPGA chip can access at any particular time, giving a higher level of flexibility to the OPGA system.

Although a comprehensive characterization of $LiNbO_3$:Fe is beyond the scope of this section, experiments have been performed using Fe-doped crystals in the reflection geometry for red wavelengths to verify that their performance meets the requirements for this particular application. The experimental setup is a symmetric reflection-geometry system in which each recording beam impinges the crystal at an angle of $10°$ (outside the medium) and has an intensity of 4.03 mW/cm^2.

Among the z-cut $LiNbO_3$ crystals available in the laboratory, the best results have been obtained in an 8 mm thick crystal with 0.02% wt Fe that was reduced in order to increase its absorption at 633 nm. The evolution of a grating in that crystal is monitored in

Figure 3-20, in which the hologram is recorded for 30 minutes, growing up to 4.8% in diffraction efficiency. The square root of the rising edge of the curve has been fitted to a line and the sensitivity of the crystal has been estimated to be 0.031 cm/J. The hologram is then exposed to non-Bragg-matched uniform illumination for an interval of 6 hours to study the erasure dynamics. This portion of the curve has been fitted to a monoexponential, resulting in an erasure time constant (τ_e) of 29304 sec.
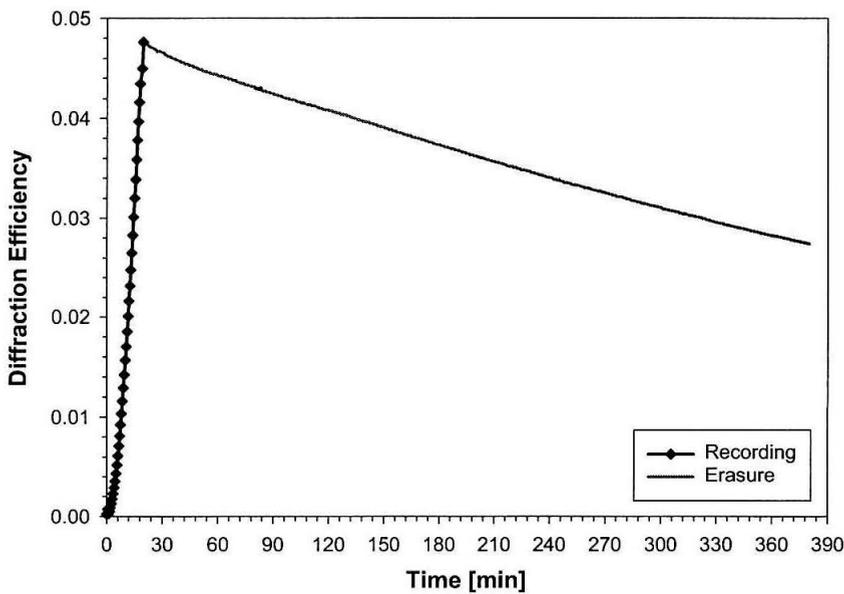
Fig. 3-20.   Recording and Bragg-mismatched erasure curves at 633 nm for an 8 mm thick z-cut LiNbO$_3$:Fe 0.02% wt crystal.

In photorefractive materials, the dynamic range is directly related to the asymmetry between the recording and erasure rates. The recording process follows an exponential growth to saturation (A$_o$) with recording time constant τ_w. The M/# is obtained as the product of the recording slope at the origin and the erasure time constant [3-23]:

$$M/\# = \frac{A_0}{\tau_w} \cdot \tau_e,$$

$$(3-5)$$

resulting in M/5.8 for the thickness of this particular LiNbO$_3$ crystal.
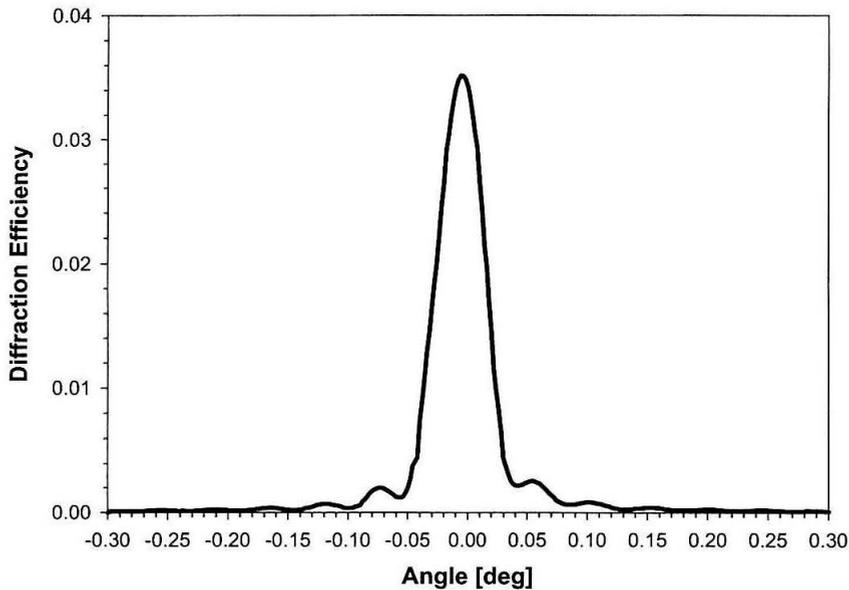
Fig. 3-21.   Angle selectivity curve of a hologram in an 8 mm thick crystal.

An alternative way to measure the dynamic range, also less subject to errors due to the estimation of the recording and erasure time constants, is by multiplexing several holograms at the same location. To minimize crosstalk between neighbor holograms, the angular separation between them was chosen to be 0.15°, which corresponds to the third null of the selectivity curve shown in Figure 3-21. The results of two different experiments, in which 41 and 101 holograms were angularly multiplexed, are presented in Figure 3-22.

The holograms have been equalized using a recording schedule computed as explained in [3-24], and according to which the recording time for the mth hologram (t_m) is given by the following expression:

(a)

(b)
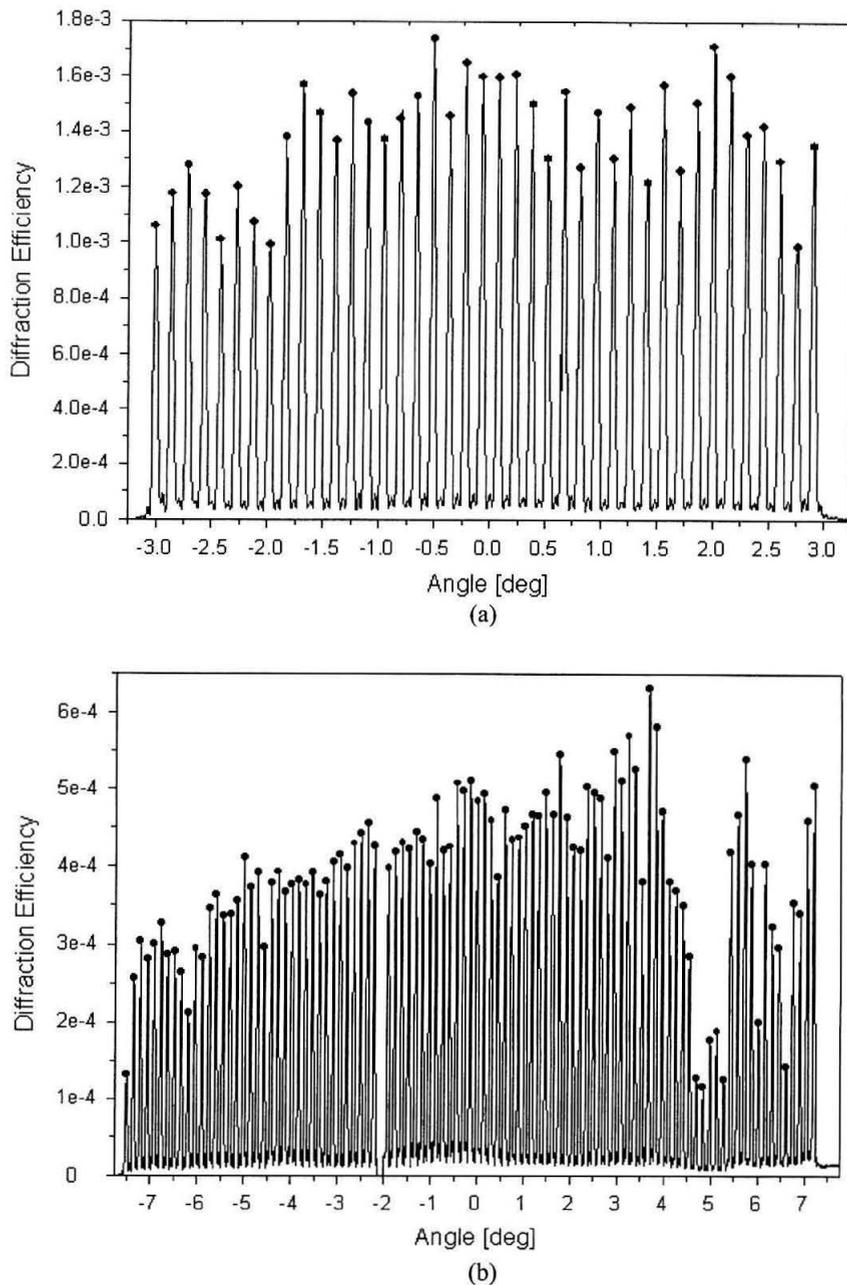
Fig. 3-22.   Comb function of (a) 41 and (b) 101 angularly multiplexed holograms equalized using recording schedule.

$$t_m = \frac{\tau_e}{m - 1 + \dfrac{\tau_e}{t_1}}, \tag{3-6}$$

where τ_e is the erasure time constant of the crystal (in this case τ_e=29304 sec) and t_1 is the recording time of the first hologram. If t_1 is chosen to be τ_e, then the expression of t_m reduces to $\dfrac{\tau_e}{m}$. However, t_1=τ_e may lead to undesirably long recording times, as it occurs in this case in which the erasure time constant is more than 8 hours.

If t_1 is smaller than τ_e, it is still possible to equalize the multiplexed holograms, although not the entire dynamic range of the material will be used. To see how the used portion of the dynamic range, $(M/\#)_{Used}$, relates to the real $(M/\#)$ of the material, consider the Mth hologram. Since it is the last hologram, it sees no erasure and therefore its grating strength is given by

$$\sqrt{\eta_M} = \frac{\partial}{\partial t}\sqrt{\eta} \cdot t_M. \tag{3-7}$$

Plugging in the expressions for the recording slope and t_M, Equation 3-7 becomes

$$\sqrt{\eta_M} = \frac{A_o}{\tau_w} \cdot \frac{\tau_e}{M - 1 + \dfrac{\tau_e}{t_1}}, \tag{3-8}$$

which can be re-written as

$$\sqrt{\eta_M} = (M/\#)_{Real} \cdot \frac{1}{M - 1 + \dfrac{\tau_e}{t_1}}. \tag{3-9}$$

On the other hand, given the fact that the holograms have been equalized, the used portion of the $(M/\#)$ satisfies the relation $(M/\#)_{Used} = \sqrt{\eta_m} \cdot M$, in which $\sqrt{\eta_m}$ is the grating strength of the mth hologram. Since all the holograms are equally strong, it suffices

to take the case m=M and replace into Equation 3-9 to obtain finally the relation between

$(M/\#)_{Used}$ and $(M/\#)_{Real}$:

$$(M/\#)_{Used} = (M/\#)_{Real} \cdot \frac{M}{M - 1 + \dfrac{\tau_e}{t_1}} \qquad (3\text{-}10)$$

In Figure 3-22(a), the holograms have a diffraction efficiency of 1.39×10^{-3} on average, resulting in $(M/\#)_{Used}$=1.53. Since t_1 is just 1.02% of τ_e, only the 29.77% of the dynamic range was used, which means in fact M/5.13 for the thickness of the crystal. A similar result is obtained in Figure 3-22(b), in which the measured $(M/\#)_{Used}$=2.26 corresponds to just the 45.47% of the dynamic range, being M/4.97 the true value for the crystal. In both cases, the M/# is somewhat lower than the value computed using Equation 3-5.

The optical quality of $LiNbO_3$ is much higher than in any other of the materials that have been considered. This property, combined with a very low level of scattering of the material, makes possible to store data pages with very small pixel size. Figure 3-23 shows
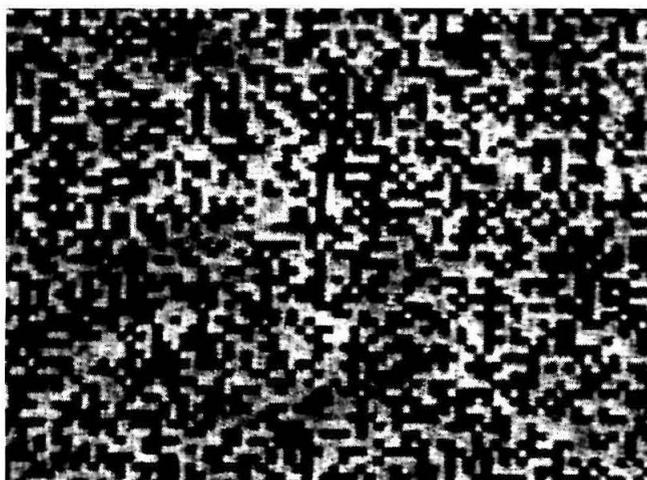
Fig. 3-23.   Phase-conjugate reconstruction of a 4 μm pixel hologram recorded in $LiNbO_3$.

a phase-conjugate reconstruction of a 4 µm pixel hologram recorded in the same 8 mm thick crystal. In the experiment, the hologram has been magnified using a ×40 microscope objective to image it onto a CCD camera with 9 µm pixel pitch. However, the system could have been made totally lensless by matching the pixels in the hologram to a 4µm-pitch CMOS detector. Two different metrics are commonly used to measure the quality of a detected hologram: the signal-to-noise ratio (SNR) and the bit-error rate (BER).
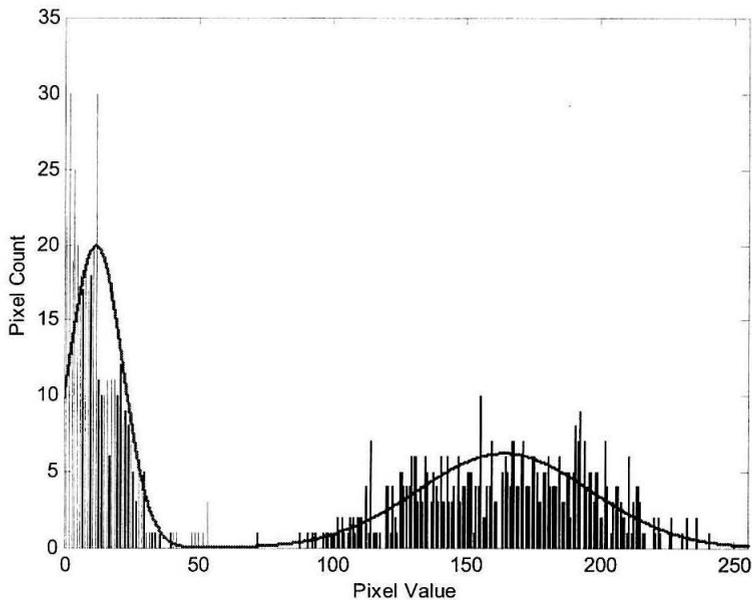
Fig. 3-24. Histogram of pixel intensity of hologram shown in Figure 3-23 fitted to two normal distributions to compute the BER.

The SNR measures the contrast between the ON and OFF pixels of the hologram, which can be easily observed from the histogram of the intensities (digitized into 256 levels in an 8-bit resolution CCD) of the hologram pixels, as shown in Figure 3-24. Imperfections and local defects in the material, as well as in the optics of the setup, lead to intensity non-

uniformity of the reconstructed hologram, which results in a distribution of pixel intensities that cluster into two groups in the histogram. The SNR is then defined as

$$\text{SNR} = \frac{\mu_1 - \mu_0}{\sqrt{\sigma_1^2 - \sigma_0^2}}, \tag{3-11}$$

where μ_1 and σ_1 are the mean and standard deviation of the cluster corresponding to the ON pixels (logic "1"); and similarly μ_0 and σ_0 for the cluster of OFF pixels (logic "0").

The BER provides information about the probability of erroneous detection, and can also be obtained from the histogram by fitting each cluster to some type of probability distribution: $f_1(\cdot)$ for "1", and $f_0(\cdot)$ for "0". For example, the data in Figure 3-24 has been fitted to two normal distributions. The BER is then defined as

$$\text{BER} = P_1 \int_{-\infty}^{\gamma} f_1(\xi) \cdot d\xi + P_0 \int_{\gamma}^{\infty} f_0(\xi) \cdot d\xi, \tag{3-12}$$

where P_1 and P_0 are the symbol priors, usually taken $\frac{1}{2}$ because the holograms are random-pixel images, and γ is the decision threshold, which according to the maximum likelihood criterion is optimal when determined as the solution to $P_1 \cdot f_1(\gamma) = P_0 \cdot f_0(\gamma)$. If $f_1(\cdot)$ and $f_0(\cdot)$ are normally distributed, Equation 3-12 reduces to

$$\text{BER} = P_1 \cdot \text{erfc}\left(\frac{\mu_1 - \gamma}{\sqrt{2} \cdot \sigma_1}\right) + P_0 \cdot \text{erfc}\left(\frac{\gamma - \mu_0}{\sqrt{2} \cdot \sigma_0}\right). \tag{3-13}$$

The hologram in Figure 3-23 presents a global SNR across the entire image of 4.98, with regions where the SNR reaches a maximum of 5.18. The resulting raw BER (i.e., before any error-correction techniques are implemented) is 5.25×10^{-4}.

The above presented experimental results show that $LiNbO_3$ is able to meet both the optical quality requirements for small-sized pixels and the dynamic range considerations (i.e., M/5) for rapid configuration of the OPGA system. However, the larger thickness of the crystals (a few millimeters rather than a few hundred microns in polymers) puts more strain in the level of wavelength uniformity of the VCSELs. It is also important to point out that the volatility of the holograms has to be addressed for a long-term operation of the OPGA module. Solutions to this problem exists, like thermal fixing [3-21], although nonvolatility is obtained at the expense of M/#.

3.5 Module packaging

After presenting the properties and main issues of each one of the three major components of the OPGA module, it is necessary to discuss next the integration of the addressing device, optical memory, and CMOS chip, in a single package. The main goal is that the OPGA module needs to be small enough to be mounted on a computer board, thus the module should have a low profile. As already mentioned in Chapter 2, the height of the module depends only on the focal length of the lens used before the SLM, and this distance can be made as little as 1 cm. The module is very compact due to the lensless readout and to the small size of the area of the recording medium used to store the holograms.

The package shown in Figure 3-25 houses the optical memory on the top rectangular window. The VCSEL arrays, integrated on both sides, retrieve the holograms detected on the chip located on the bottom of the package. The package also needs to be robust to ensure the proper alignment between all of its components. It is important to preserve the one-to-one correspondence between the pixels in the hologram and the photodetectors on
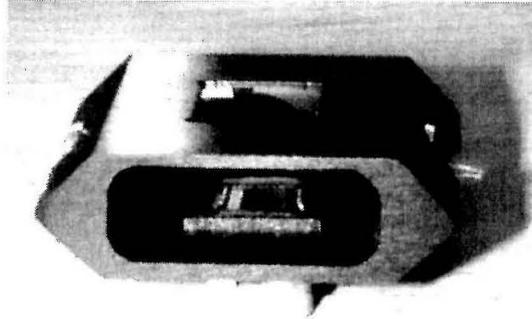
Fig. 3-25.   Mechanical design of the OPGA module integrating in a compact package the optical memory (window on the top), VCSEL arrays (one on each side) and the chip (bottom).

the chip and also to avoid any change on the areas illuminated by the VCSELs on the optical material.

A first prototype has been developed to demonstrate that it is possible to integrate all three elements in a compact module. The module uses a 5×1 array of VCSELs to read out the holograms that have been stored in a 100 µm thick layer of DuPont photopolymer used as the optical memory. For this demonstrator, instead of the OPGA chip, a simple CCD camera chip was interfaced to the optical memory to detect the reconstructed holograms.

During recording, a laser diode stores two shift-multiplexed holograms in the memory at the two locations that match the position of the two VCSELs at each end of the array. The laser diode has been thermally controlled to match its wavelength to the one of the VCSELs. The experimental setup, shown in Figure 3-26, combines a reflection geometry architecture and lensless readout making the OPGA module very simple and compact. Once the recording operation is finished, the VCSEL array is assembled into the OPGA reader module, and the module removed from the setup of the recorder.
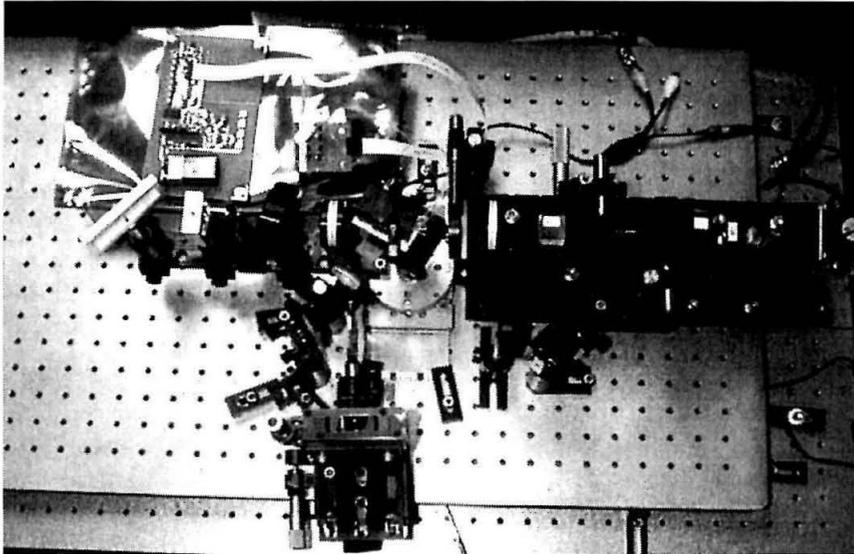
Fig. 3-26. Picture of the experimental setup used to record the holograms into the optical memory of the OPGA module.
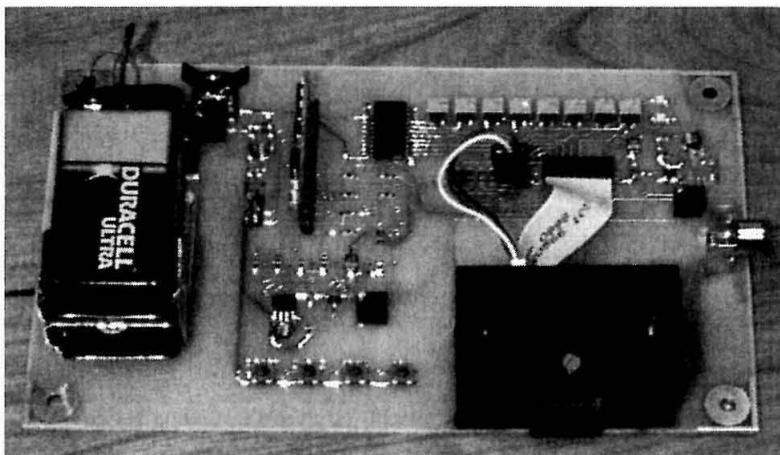
Fig. 3-27. First generation OPGA prototype mounted on the board that carries the circuitry to drive single VCSELs in the array and to power up the module.

The OPGA module is then mounted on a demo board, shown in Figure 3-27, that

contains the additional circuitry to drive the VCSELs and select which element in the array

is active. The board also carries the interface to the monitor where the holograms are dis-

played as they are read out by the VCSELs (Figure 3-28).
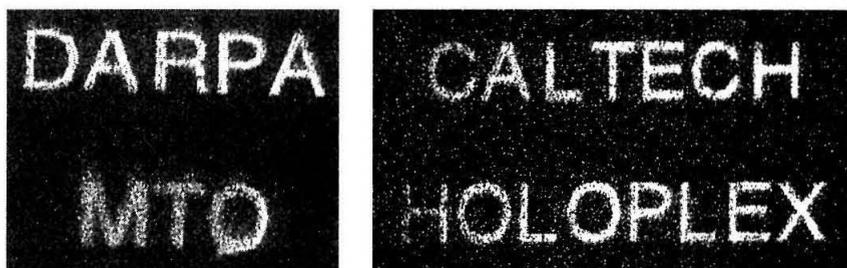
Fig. 3-28.   Reconstruction of the two holograms stored in the optical memory of the OPGA prototype. Two different VCSELs are used to retrieve one hologram at a time.

A more advanced packaging for the OPGA module has been proposed and demonstrated in [3-12]. The module achieves an extremely low profile by directly mounting the optical memory, a thin slab of $LiNbO_3:Fe$, on top of the OPGA chip as depicted in Figure 3-29. The system makes use of the folded shift multiplexing technique[3-12] to store holograms into the memory. The light from the VCSELs is coupled into the crystal from the edge and guided along the crystal as it bounces off the top and bottom surfaces due to total internal reflection to read out the holograms, which are projected onto the OPGA chip.
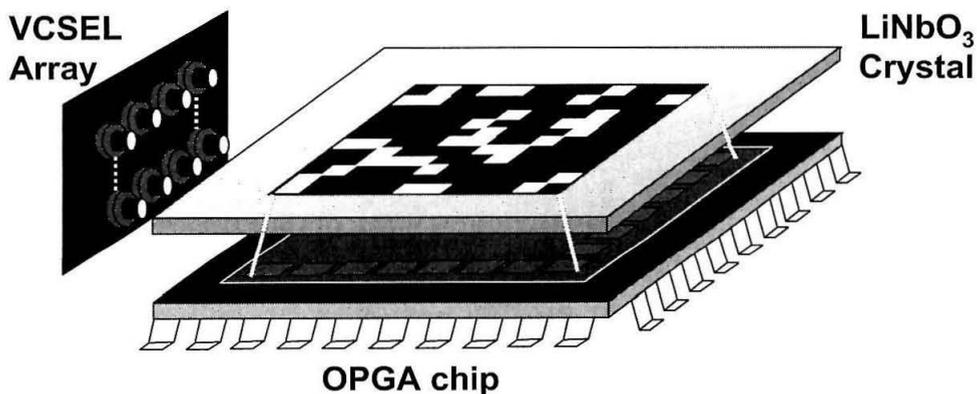
Fig. 3-29.   Advanced packaging for the OPGA in which a thin slab of $LiNbO_3$ is directly mounted on top of the OPGA chip. The beam from the VCSEL is guided inside the crystal.

3.6 System demonstration

The previous section discussed the integration of the three components of the OPGA into a compact module. A demo OPGA module was developed in which VCSELs were used to read out holographic data pages from the optical memory. However, the system used a CCD camera instead of the OPGA chip. In this section, the holographic programming of the OPGA chip will be demonstrated. The system described here interfaces the full OPGA chip with an optical memory containing the configuration templates that are downloaded onto the chip upon readout.

3.6.1 Demonstrator setup

The 514 nm line of an Argon ion laser is used in the experimental setup sketched in Figure 3-30. In order to record the holograms in the OPGA system, the laser light beam is split into two arms. The signal beam illuminates a 15 μm pixel 320×240 Kopin SLM. The pattern displayed on the SLM is projected onto the chip using two 5 cm focal length Nikon lenses arranged in a 4-f system and pixel-matched to the APS array of the chip.

The holographic memory, in this particular demonstration a 38 μm thick sample of Du Pont photopolymer, is placed between the SLM and the first Nikon lens. The optical medium is mounted on a rotation stage to provide peristrophic multiplexing. The iris placed at the Fourier plane of the first lens in the 4-f system, initially intended to filter out the higher diffraction orders of the signal beam produced by the pixelation of the SLM, will also serve to block all the peristrophically multiplexed holograms that are not selected to program the chip at any given time. Simple additional optics is necessary to bring a plane wave reference beam onto the holographic material at an angle of incidence of 45° with
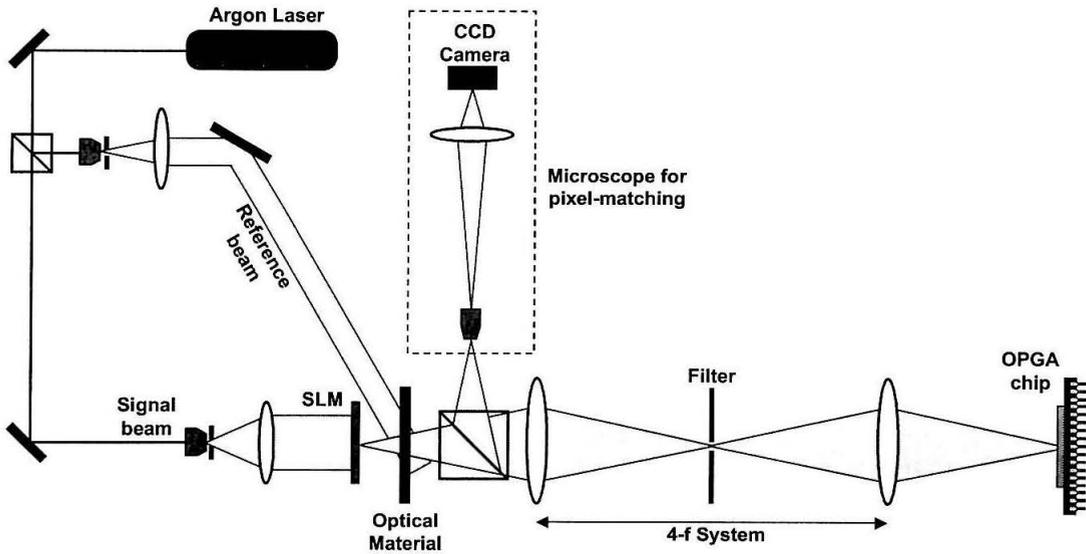
Fig. 3-30. Experimental setup for the holographic programming of the OPGA chip. The microscope (inside the dashed-line box) is used to monitor the pixel-matching process.

respect to its surface normal. The total intensity of the recording beams has been set to 1.5 mW/cm^2.

3.6.2 Pixel alignment process

The differential-encoding scheme used by the OPGA chip makes the process of pixel-matching the SLM to the APS array more difficult. As described in Section 3.3.1, each pair of pixels in the SLM is mapped into one bit of information. The alignment is complicated by the fact that the output of the chip does not correspond to the image on the SLM as seen by the APS array on the chip (i.e, a 64 × 64 pixel array), but to the differentially detected configuration bits (i.e., a 64 × 32 data page). Therefore, the digitized output of the chip alone does not provide enough information to guide the alignment process.

Instead, pixel-matching between the SLM and the chip has been performed with the aid of a microscope as shown in Figure 3-30. Due to the imaging property of the 4-f system,

when the chip is illuminated with an incoherent light source (e.g., a white LED illuminator), the light reflected off the surface of the chip forms an image of the array of photodetectors at a distance equal to one focal length in front of the first Nikon lens. The microscope magnifies this image, Figure 3-31(a), and projects it onto a CCD camera.

<div align="center">(a)   (b)   (c)</div>
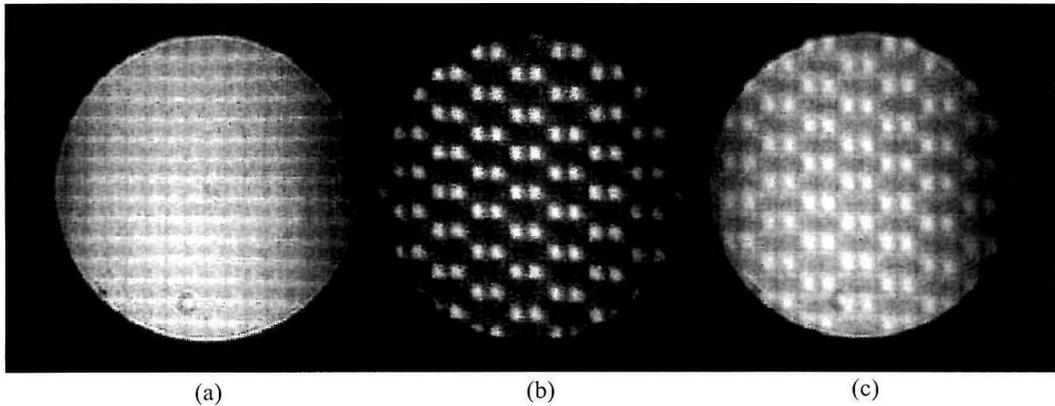
Fig. 3-31.   Pixel alignment process showing the image of (a) the APS array on the chip, (b) the pixels on the SLM, and (c) the SLM pixels superimposed on the APS array.

Similarly, the pattern displayed on the SLM can also be observed on the microscope. The pixels of the SLM are imaged first on the OPGA chip. However, the chip can be viewed as a partial mirror that reflects some of the light that comes from the SLM. This light travels back through the 4-f system and forms an image again at one focal length away from the first lens, right in front of the microscope, Figure 3-31(b). The pixel-matching condition is achieved when the pattern displayed on the SLM coincides with the layout of the APS array, Figure 3-31(c).

3.6.3 Holographic programming

Once the optical interface is properly aligned, holograms carrying configuration data can be recorded in the optical memory, and their reconstructions used to drive the logic

of the OPGA. Figure 3-32 shows the reconstruction of a hologram as seen by the APS array

on the chip. The differential nature of the encoded data is evident from the fact that columns

of mostly ON pixels are immediately followed by a column in which the only ON pixels

occur at the vacancies of the previous column.
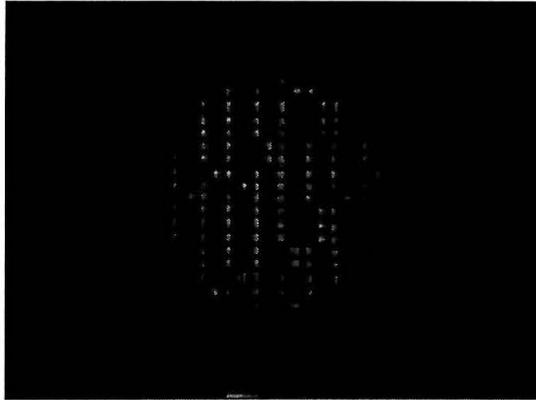
Fig. 3-32.   Reconstruction of a hologram containing configuration data. The extension of the region of the hologram that can be visualized on the CCD is limited by the aperture of the objective lens in the microscope.

The digitized output of the APS array is accessible from the outside of the chip and

can be visualized on a computer screen via a parallel port that connects the board that car-

ries the OPGA chip to a computer. The 64×32 output pattern in Figure 3-33, correspond-

ing to the hologram showed in Figure 3-32, is retrieved free from error. The detected data

is then used to program the logic array of the OPGA to perform some type of computation.

Given that only 13 rows out of the 64 available in the APS array are needed to fully con-

figure the logic, several optical programs can be stored in the hologram and downloaded

into the chip in less than 8 μs (APS integration time).

Two different operations can be programmed using this hologram: Configuration

#1 routes a 5-bit input signal from the I/O port on the left of the chip throughout the chip
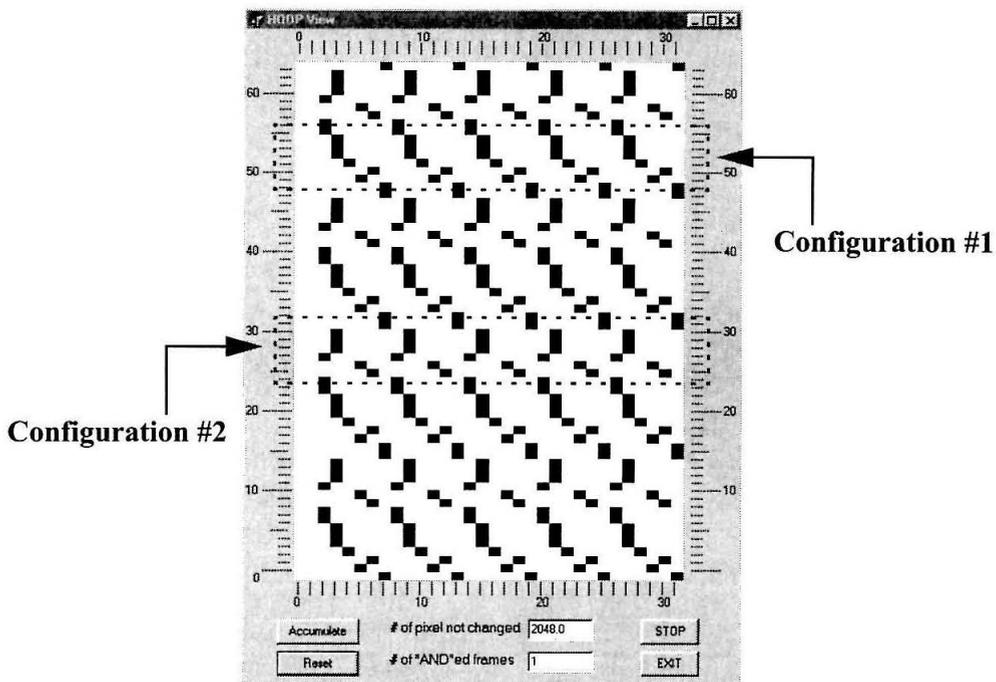
Fig. 3-33.   Digitized output of the APS array showing the programming information contained in the hologram after being differentially detected. The hologram contains two configuration templates in the rows encircled by dotted lines.

and is finally delivered to the I/O port on the top of the logic array (labeled as Output Port A) in Figure 3-34(a). On the other hand, configuration #2 modifies the internal state of the crossbar switching matrices, so the input signal is re-routed towards the I/O port on the right of the chip (labeled as Output Port B) in Figure 3-34(b). The operation of the chip is exper-imentally verified with a test signal generated by a 5-bit counter that cycles through all its 32 logic states. The results of the holographic programming can be observed in the scope trace in Figure 3-35. As the programming of the chip toggles continuously between config-uration #1 and #2, the input signal is alternatively delivered to either port A or port B.

The ability to reconfigure the OPGA using different rows of the APS array is equiv-alent to the concept of partial reconfiguration used in the FPGA community, when only part

(a)

(b)
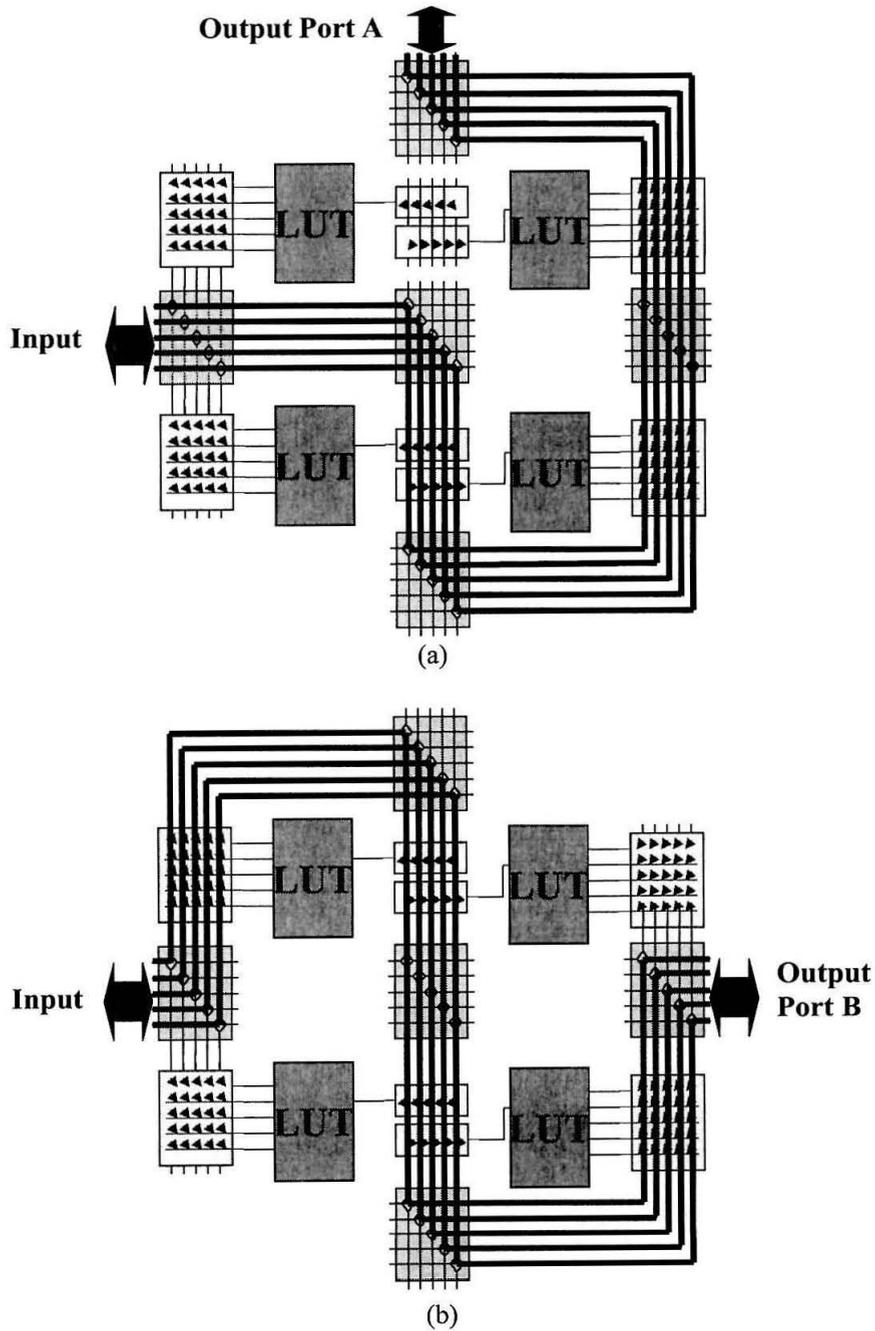
Fig. 3-34. Routing scheme programmed by each one of the two configurations contained in the hologram. (a) Configuration #1 delivers the 5-bit input to Output Port A (on the top), while (b) Configuration #2 sends it to Output Port B (on the right).
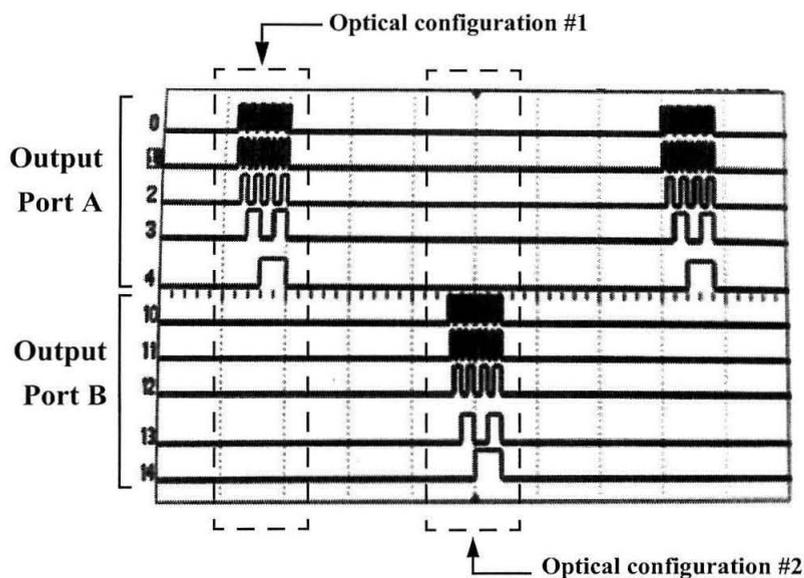
Fig. 3-35. Scope trace of the output of the logic block array as programmed optically. A 5-bit input signal is delivered cyclically to either Output Port A during Configuration #1, or Output Port B during Configuration #2.

of the logic of the FPGA needs to be reprogrammed. Having different configurations coexisting on the same hologram adds flexibility to the OPGA.
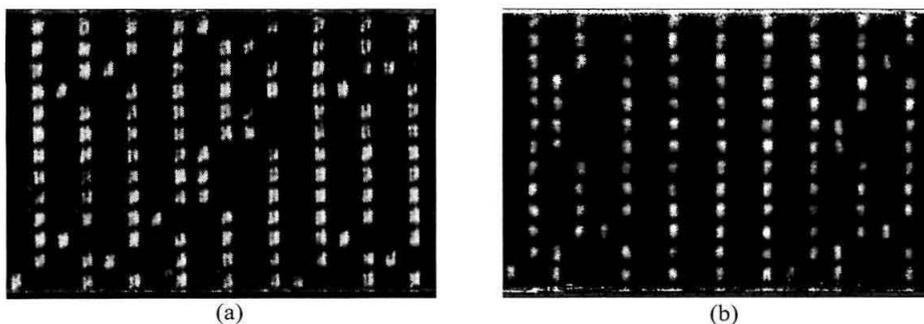
Fig. 3-36. Reconstruction of the two multiplexed holographic templates as seen by the APS array. Each hologram carries two configuration programs: (a) *Hologram I* contains Configuration #1 and #2, while (b) *Hologram II* contains #3 and #4.

Finally, full reconfiguration using different holograms has been demonstrated by

peristrophically multiplexing two holographic patterns in the optical memory with an angu-

lar separation of 20° between them. The reconstruction of the two holographic templates is presented in Figure 3-36. By rotating the optical medium, it is possible to select between *Hologram I* and *Hologram II* to implement global reconfiguration of the chip. Furthermore, by choosing which rows to use within each hologram, two different functions can be programmed. Therefore, 4 different configurations can be downloaded into the chip. While the two operations performed by *Hologram I* have already been described (see Figure 3-34), the ones performed by *Hologram II* (configurations #3 and #4) are presented in Figure 3-37 and explained below.

Configuration #3 broadcasts the signals from the input port to both output ports, while configuration #4 splits the input lines into two groups: those corresponding to the even bits are sent to Output Port A, and the lines carrying the odd bits are sent to Output Port B. The experimental results are presented in Figure 3-38. The 5-bit counter used as test input signal appears on both output ports during configuration #3, as shown in the scope trace. When the OPGA is programmed with configuration #4, the bits with weights 2^0, 2^2 and 2^4 are detected on Output Port A, while the ones with weights 2^1 and 2^3 appear on Output Port B.

The reconfiguration time (T_{config}) can be defined as the interval between the end of the execution of the program under the current configuration (i.e., once the 5-bit counter has cycled through its 32 logic states) and the beginning of the execution of a different program with the new configuration (i.e., the first logic state of the counter). The configuration time includes the readout cycle of the selected rows of the APS array, plus the delivery time of the configuration bits to the logic blocks. In the experiment shown in Figure 3-38, the measured T_{config} was 127 µs for full chip configuration. This time has to be interpreted as
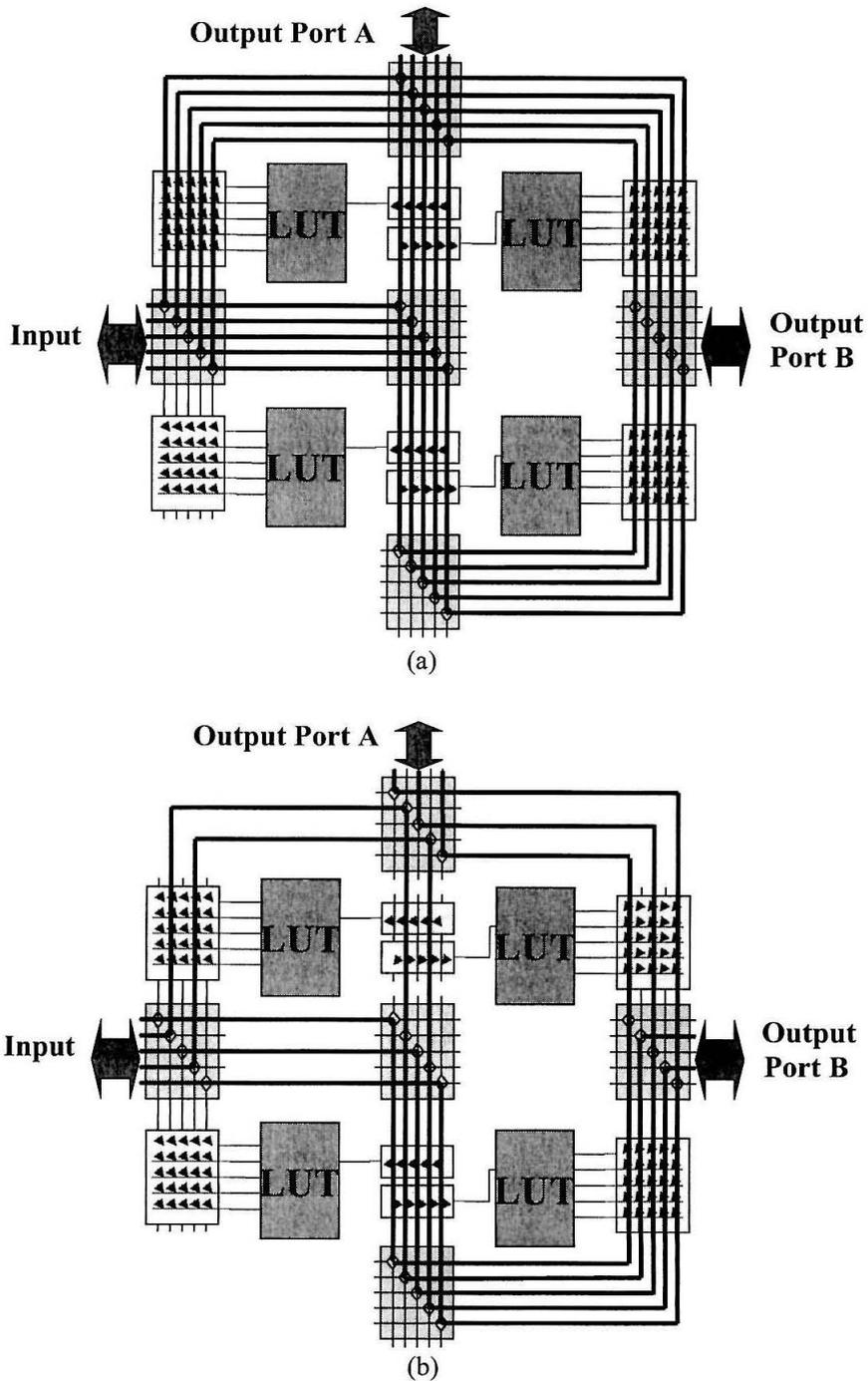
Fig. 3-37.   Routing scheme programmed by the two configurations contained in *Hologram II*. (a) Configuration #3 broadcasts the 5-bit input to both Output Port A and B, while (b) Configuration #4 sends the even lines of the input to Output Port A and the odd lines to Output Port B.
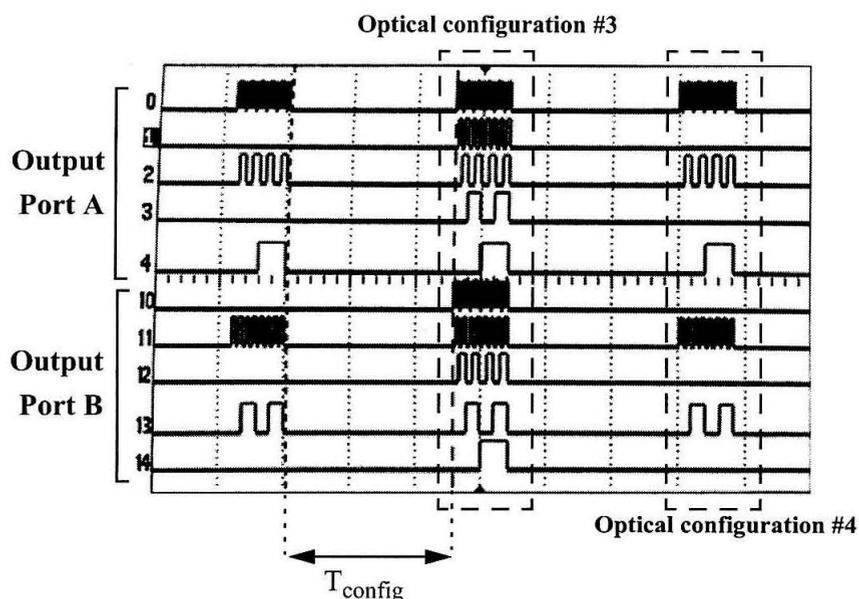
Fig. 3-38. Scope trace of the output of the logic block array as programmed optically by *Hologram II*. The 5-bit input signal is broadcast to both output ports during Configuration #3, and split into even and odd lines during Configuration #4. The reconfiguration time T_{config} is 127 µs.

an upper bound, because it is limited by the master clock signal of 500 KHz generated by

the OPGA board.

This last experiment demonstrates the feasibility of transferring information in par-

allel from a holographic database to a silicon chip, and the high degree of flexibility derived

from the reconfiguration capabilities. The system can be expanded to include more holo-

graphic templates using materials with larger dynamic range and better optical quality, like

Aprilis photopolymer.

Automation of the rotation stages to implement either peristrophic or angular mul-

tiplexing should also be carefully explored. The reconfiguration time measured in the

experiments did not take into consideration the time it takes to retrieve a particular holo-

graphic data page from the optical memory. The use of a VCSEL array in the design of the

OPGA made the readout time be negligible because no mechanical components were involved. However, if a single laser is to be used to overcome the limitations of the red VCSELs, the use of a fast steering device is required to keep the reconfiguration time low. Fortunately, the increase in available output power that a laser diode provides with respect to the VCSELs makes possible to trade off longer retrieval times with shorter photodetector integration times.

References

[3-1] A. V. Krishnamoorthy, K. W. Goossen, L. M. F. Chirovsky, R. G. Rozier, P. Chandramani, W. S. Hobson, S. P. Hui, J. Lopata, J. A. Walker, and L. A. D'Asaro, "16×16 VCSEL array flip-chip bonded to CMOS VLSI circuit," *IEEE Photonics Technology Letters* **12**, 1073-1075 (2000).

[3-2] R. P. Schneider Jr., and J. A. Lott, "Cavity design for improved electrical injection in InAlGaP/AlGaAs visible (639–661 nm) vertical-cavity surface-emitting laser diodes," *Applied Physics Letters* **63**, 917-919 (1993).

[3-3] J. Mumbru, G. Panotopoulos, D. Psaltis, X. An, F. Mok, S. Ay, S. Barna, and E. Fossum, "Optically programmable gate array," *Proc. SPIE of Optics in Computing 2000* **4089**, 763-771 (2000).

[3-4] S. Mendis, S. E. Kemeny, and E. R. Fossum, "CMOS active pixel image sensor," *IEEE Transactions on Electronic Devices* **41**, 452-453 (1994).

[3-5] E. R. Fossum, "CMOS image sensors: Electronic camera-on-a-chip", *IEEE Transactions on Electronic Devices* **44**, 1689-1698 (1997).

[3-6] J. F. Heanue, M. C. Bashaw, and L. Hesselink, "Volume holographic storage and retrieval of digital data," *Science* **265**, 749-752 (1994).

[3-7] D. Psaltis, and G. W. Burr, "Holographic data storage," *Computer* **31**, 52-60 (1998).

[3-8] G. J. Steckman, I. Solomatine, G. Zhou, and D. Psaltis, "Characterization of phenanthrenequinone-doped poly(methyl methacrylate) for holographic memory," *Optics Letters* **23**, 1310-1312 (1998).

[3-9] D. A. Waldman, H.-Y. S. Li, M. G. Horner, "Volume shrinkage in slant fringe gratings of a cationic ring-opening holographic recording material," *Journal of Imaging Science and Technology* **41**, 497-514 (1997).

[3-10] D. A. Waldman, H.-Y. S. Li, and E. A. Cetin, "Holographic recording properties in thick films of ULSH-500 photopolymer," *Proc. SPIE of Diffractive and Holographic Device Technologies and Applications* **3291**, 89-103 (1998).

[3-11] L. Paraschis, Y. Sugiyama, L. Hesselink, "Physical properties of volume holographic recording utilizing photo-initiated polymerization for nonvolatile digital data storage," *Proc. SPIE of Conference on Advanced Optical Data Storage: Materials, Systems, and Interfaces to Computers* **3802**, 72-83 (1999).

[3-12] C. Moser, *Optical Information Processing*, Ph.D. thesis, California Institute of Technology, 2001.

[3-13] K. Buse, A. Adibi, and D. Psaltis, "Non-volatile holographic storage in doubly doped lithium niobate crystals," *Nature* **393**, 665-668 (1998).

[3-14] K. Buse, A. Adibi, and D. Psaltis, "Efficient non-volatile holographic recording in doubly doped lithium niobate," *Journal of Optics A-Pure and Applied Optics* **1**, 237-238 (1999).

[3-15] J. A. Jenney, "Holographic recording with photopolymers," *Journal of the Optical Society of America* **60**, 1155-1161 (1970).

[3-16] M. J. Jeudy, and J. J. Robillard, "Spectral sensitization of a variable index material for recording phase holograms with high efficiency," *Optical Communications* **13**, 25 (1975).

[3-17] S. Blaya, L. Carretero, R. Mallavia, A. Fimia, R. F. Madrigal, M. Ulibarrena, and D. Levy, "Optimization of an acrylamide-based dry film used for holographic recording," *Applied Optics* **37**, 7604-7610 (1998).

[3-18] K. Curtis, A. Pu, and D. Psaltis, "Method for holographic storage using peristrophic multiplexing," *Optics Letters* **19**, 993-994 (1994).

[3-19] A. Pu, K. Curtis, and D. Psaltis, "Exposure schedule for multiplexing holograms in photopolymer films," *Optical Engineering* **35**, 2824-2829 (1996)

[3-20] F. H. Mok, "Angle-multiplexed storage of 5000 holograms in lithium niobate," *Optics Letters* **18**, 915-917 (1993).

[3-21] X. An, D. Psaltis, G. W. Burr, "Thermal fixing of 10,000 holograms in $LiNbO_3$:Fe," *Applied Optics* **38**, 386-393 (1999).

[3-22] E. Chuang, W. H. Liu, J. J. P. Drolet, D. Psaltis, "Holographic random access memory (HRAM)," *Proceedings of the IEEE* **87**, 1931-1940 (1999).

[3-23] F. H. Mok, G. W. Burr, D. Psaltis, "System metric for holographic memory systems," *Optics Letters* **21**, 896-898 (1996).

[3-24] G. W. Burr, *Volume holographic storage using the 90° geometry*, Ph.D. thesis, California Institute of Technology, 1996.

4 Holographic data storage in Aprilis photopolymer

4.1 Introduction

The use of photopolymers in optical data storage and other holographic applications has fuelled an extensive research work to develop new materials [4-1]-[4-9]. However, it has proven difficult to find the one that, while enjoying a large dynamic range and high sensitivity, can also remedy the limitations that most of the conventional polymers suffer from, like shrinkage and scattering. In that regard, the material developed by Aprilis Inc. [4-10], [4-11] seems to have successfully addressed many of these issues. The excellent optical quality of the Aprilis material, with $\lambda/10$ wavefront distortion over a 2×2 mm area, combined with an extremely low level of scattering and very wide modulation transfer function (MTF), which extends from 0.2 μm to 10 μm, makes this polymer meet satisfactorily the demands of high performance holographic systems.[4-12]

Shrinkage in Aprilis film is negligible thanks to the use of the cationic ring-opening polymerization (CROP) technology [4-12]. If compared to DuPont photopolymer, for example, Aprilis experiences a 0.05% dimensional change during recording (mostly in the transversal direction), which represents a two order of magnitude improvement with respect to the 3.5% in DuPont. Figure 4-1 shows the different effects of shrinkage on the reconstruction of a digital image recorded in a sample of DuPont and in Aprilis. For the hologram recorded in DuPont, shrinkage makes impossible to reconstruct the entire image at once. The leftmost part of the image in Figure 4-1(a) is lost due to Bragg mismatch. On

the other hand, as shown in Figure 4-1(b), there is no problem to retrieve the entire holo-gram recorded in the Aprilis.

(a)   (b)
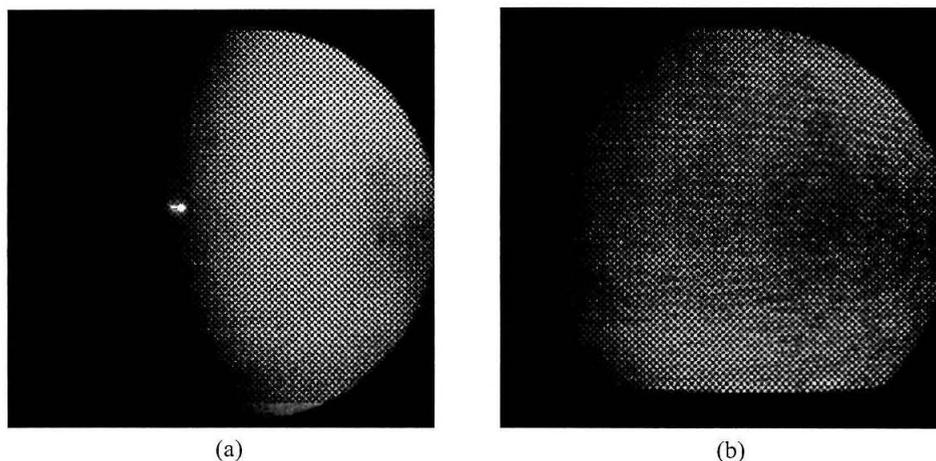
Fig. 4-1.   Qualitative comparison of the effect of shrinkage on the reconstruction of the hologram of a digital image recorded in (a) Du Pont and (b) Aprilis polymer.

Although the holographic medium was initially thought to be the most critical ele-ment in the OPGA module, the optical memory is no longer the limiting factor thanks to the good performance of the Aprilis material. Section 4.2 presents further experimental characterization of Aprilis ULSH500-7A, a 200 μm thick film designed to operate at both 514 nm and 532 nm. Section 4.3 deals with the ability of the medium to store images. Finally, Section 4.4 introduces a technique for hologram duplication using Aprilis films.

4.2 Material characterization

The effect of the pre-exposure on the quality of the holograms recorded in Aprilis material has been investigated. Experiments were performed with an Argon laser in a trans-mission-geometry setup. The 514 nm recording beams had an intensity of 15 mW/cm^2 and were incident at 20° with respect to the surface normal of the material. A white light lamp

with a green filter was used for pre-exposure. Its emission spectrum had a peak at 532 nm,

and its FWHM bandwidth was approximately 90 nm. The intensity of the lamp was set to

1 mW/cm^2.

While in other photopolymers, like Du Pont, holographic recording does not begin

until a given amount of energy has been absorbed by the material, interestingly hologram

growth can be observed in Aprilis films without any pre-exposure applied to the samples.

The real-time growth of a hologram in a sample without pre-exposure has been monitored

with a He-Ne laser and is shown in Figure 4-2. The recording curve does not exhibit a dead
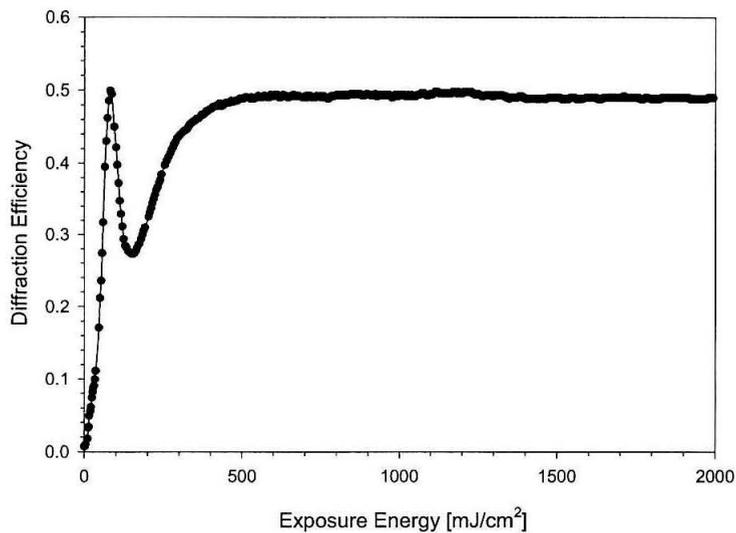
Fig. 4-2.   Real-time growth of a hologram recorded without energetic pre-exposure.

zone at the beginning before hologram formation takes off, on the contrary the diffraction

efficiency grows very rapidly to its maximum before it goes through a hump and finally

settles at around 50%.

The angular selectivity curve of this hologram, measured in Figure 4-3(a), is highly distorted and presents substantial enhancement of its side lobes. The actual selectivity curve (solid dots) differs greatly from the theoretical plot (solid line) for the thickness of the material. Another experiment was performed in which the sample of Aprilis was pre-exposed for 150 mJ/cm^2, and then a grating was recorded for nearly 1 J/cm^2. A 60% strong hologram was obtained at the end of the experiment, whose angular selectivity curve, Figure 4-3(b), rendered very good agreement with the theory. The reason for pre-exposure in Aprilis material seems to be the need to structurally stabilize the polymer before holographic recording takes place. Insufficient pre-exposure will doubtlessly result in distorted holograms.

Experiments to multiplex plane-wave holograms were conducted to determine the dynamic range of the material. In the setup, the sample was centered on a rotation stage to allow for peristrophic multiplexing. Moreover, angular multiplexing was also available thanks to the combination of a mirror mounted on a second rotation stage and a 4-f system, both placed in the path of the reference beam.

The issue of hologram distortion reappears in the multiplexing experiments. If the sample receives inadequate pre-exposure, the addition of more holograms causes the degradation of the previously recorded ones. Eventually the material becomes stable, so the recording of more holograms does not damage the previous ones. To visualize this effect, Figure 4-4 presents the angular selectivity curves of some of the 320 holograms multiplexed in a deficiently pre-exposed sample. The early holograms, in particular the first, second and third holograms in Figure 4-4(a), are severely distorted. By the time the 58th, 59th and 60th holograms were recorded, Figure 4-4(b), the sample was more stable, so their

(a)

(b)
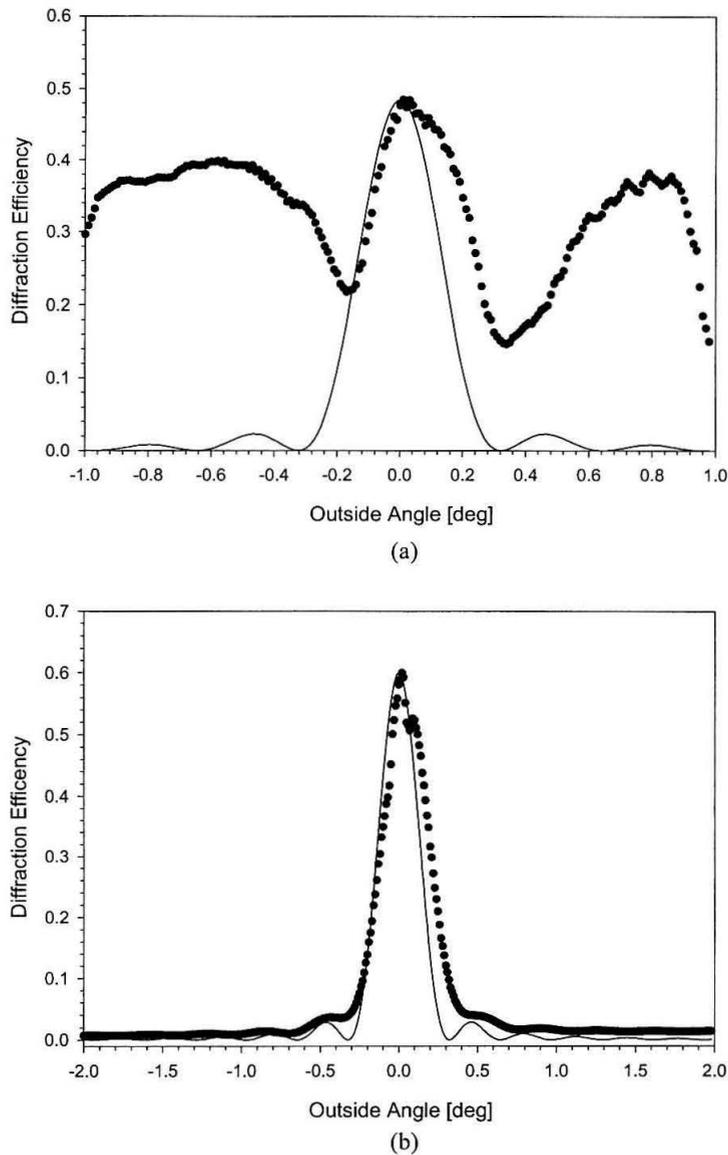
Fig. 4-3. Comparison of the angular selectivity curves for a hologram recorded (a) without pre-exposure and (b) with 150 mJ/cm^2 of pre-exposure. In each case, the actual measurement (solid dots) is compared to the theoretical selectivity curve (solid line).

selectivity curves started to agree with what is expected from theory. Finally, the 178th, 179th and 180th holograms, Figure 4-4(c), rendered very nice selectivity even after storing 140 more holograms.
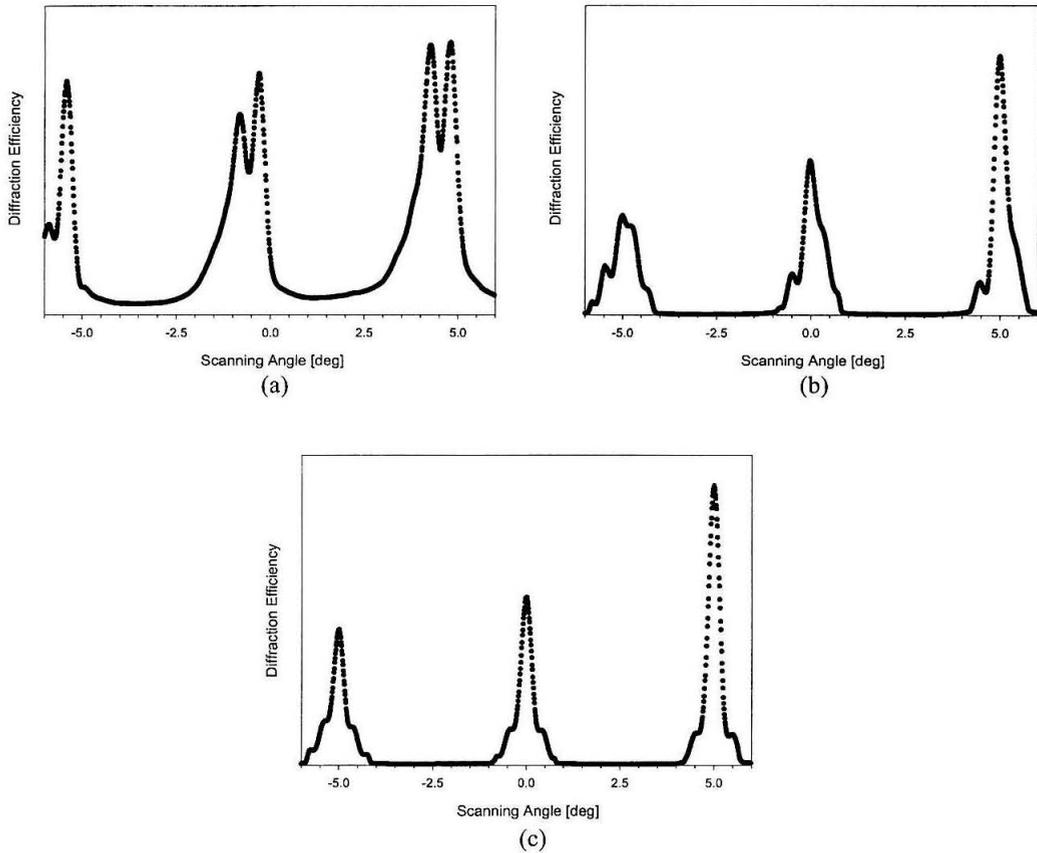
Fig. 4-4.   Angular scan of the multiplexed holograms at three peristrophic location showing the selectivity curves of the (a) first, second and third; (b) 58th, 59th and 60th; and (c) 178th, 179th and 180th holograms.

Despite the fact that large values of dynamic range have been measured (reaching M/10), the concept of M/# is not meaningful if part of the multiplexed holograms are distorted, because they have no use in any application. It is therefore necessary to increase the amount of energy delivered to the sample during pre-exposure to avoid distortion since the very first hologram. Unfortunately, this will be at the expense of sacrificing part of the available dynamic range of the material.

(a)

(b)
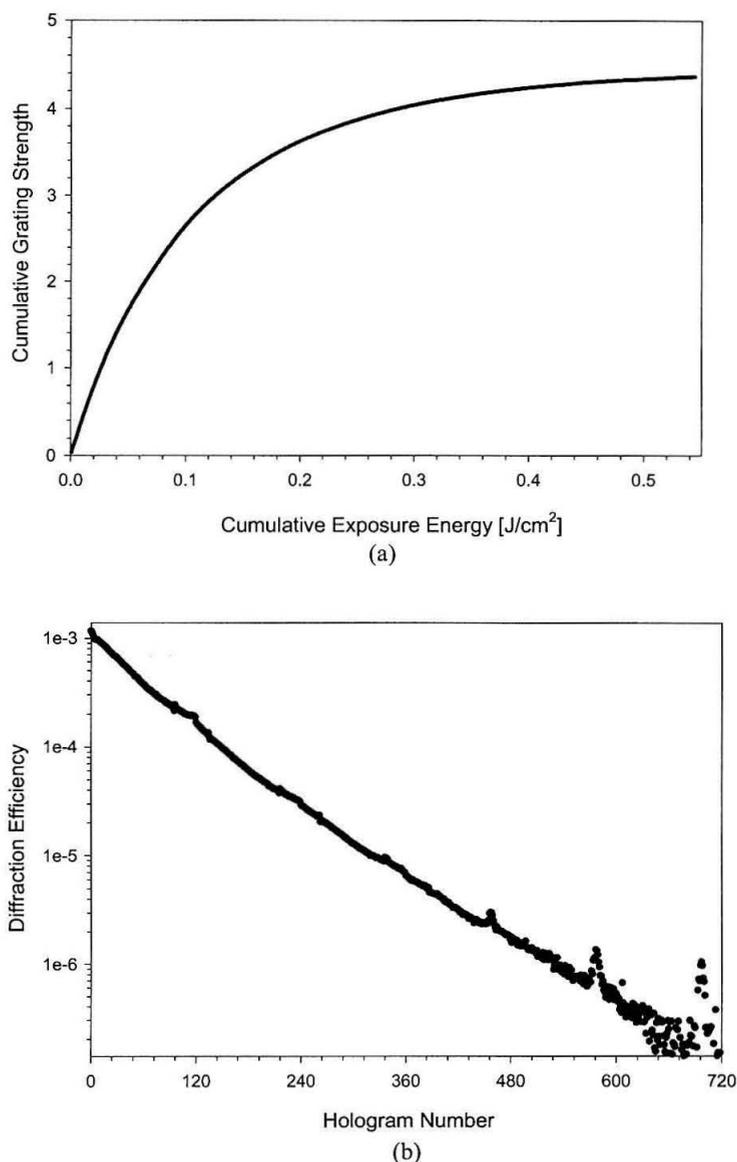
Fig. 4-5. (a) Cumulative grating strength of 720 plane-wave holograms and (b) their individual hologram strengths with equal exposures for each hologram.

Figure 4-5 shows the results of a multiplexing experiment in which 720 holograms were recorded in a single location of the material. The holograms were disposed in six angular groups separated by 1°. Each group was integrated by 120 holograms spaced by 1.5° in the peristrophic direction. A pre-exposure of 120 mJ/cm² was delivered to the

sample before the holograms were recorded with constant exposure, 0.75 mJ/cm^2 each. The cumulative grating strength curve in Figure 4-5(a) saturates around 4.4. Figure 4-5(b) presents the diffraction efficiency of the holograms ordered according to the sequence in which they were recorded. As a result of the constant exposure, the strength of the holograms decays exponentially.

In order to equalize the holograms, a recording schedule has been derived using the results presented in Figure 4-5(a) as described in [4-13]. Another experiment was carried out to store 720 holograms in a new sample. In this case, scheduled recording was adopted, but the disposition of the holograms and pre-exposure parameters were kept the same. The results are in Figure 4-6. The saturation value of the cumulative grating strength, Figure 4-6(a), increased with respect to the unscheduled case, revealing an M/# of 6.8. The holograms exhibited good equalization (within 20%), as it can be observed in Figure 4-6(b), and the average diffraction efficiency was 0.9×10^{-4}.

4.3 Holographic data storage

The holograms stored in the optical memory of the OPGA module will be actual pages of encoded data, rather than simple plane waves, thus the ability of the Aprilis material to store multiple high bandwidth images needs to be tested. The signal-to-noise ratio (SNR) and the bit-error rate (BER) introduced in Chapter 3 will be used again as metrics of the image quality of the holograms.

4.3.1 Image-plane holograms

The signal arm of the experimental setup has been modified as shown in Figure 4-7 to relay the image of a 40 μm random-pixel mask onto the CCD camera. The lenses have a
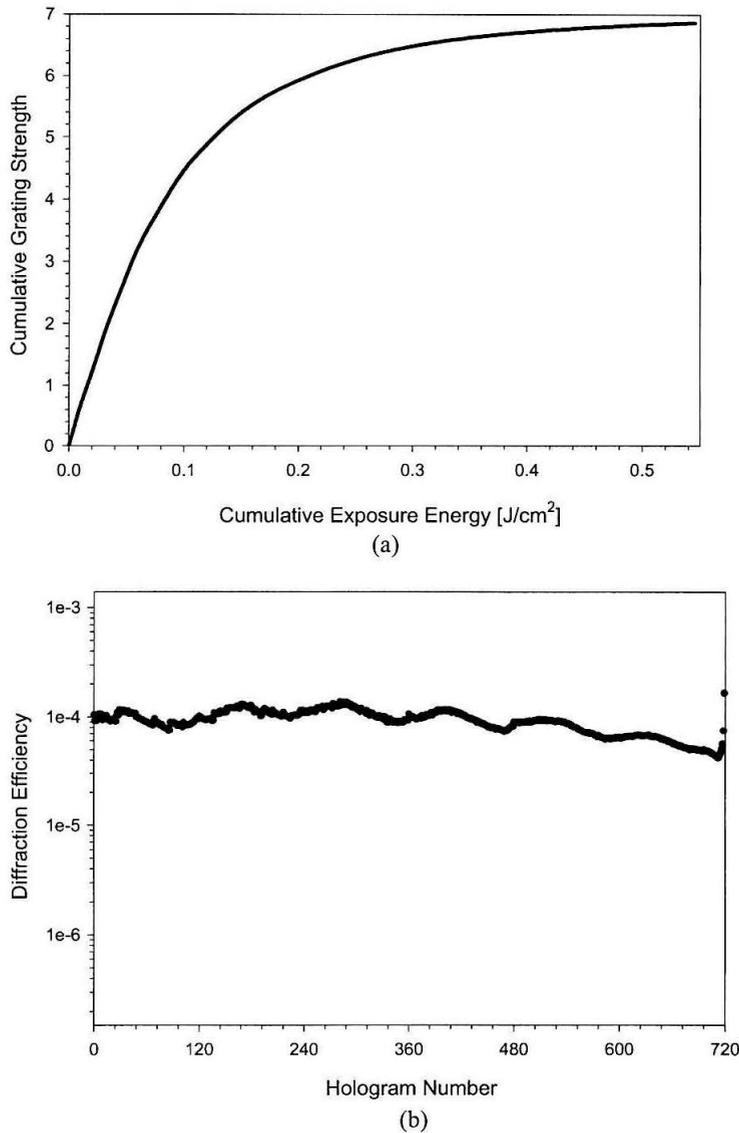
Fig. 4-6.   (a) Cumulative grating strength of 720 plane-wave holograms and (b) their individual hologram strengths with an exposure schedule.

focal length of 8cm and the material is placed at the image-plane of the 4-f system formed by the first two lenses. The sample has been mounted on a double rotation stage to provide both angle and peristrophic multiplexing. The angle between the reference and signal beam is now 32.5° (in air), with the material placed symmetrically between the two when at the
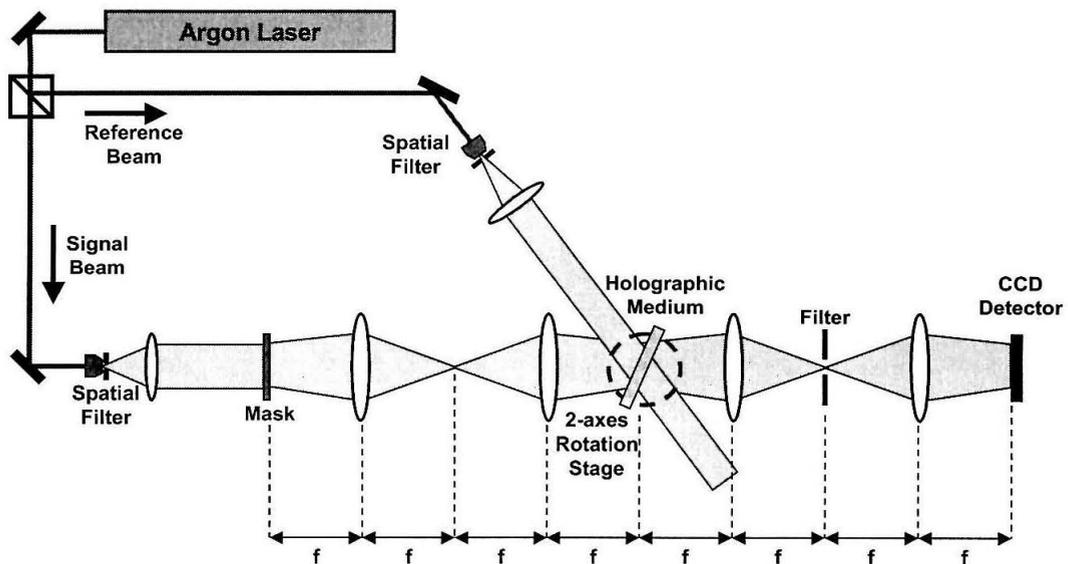
Fig. 4-7. Schematic drawing of the setup used to record image holograms. The material is mounted on a two-axis rotation stage and is located at the image plane of the first 4-f system. All lenses have a focal length (f) of 8 cm.

zero position. The purpose of the second 4-f system is twofold: First, it images the recording plane onto the CCD detector. Second, it implements a spatial filter, when an iris is placed at the Fourier plane of the third lens, necessary to select which one of the peristrophically multiplexed holograms is presented to the detector [4-14].

The SNR of the mask imaged through the optical system is measured in first place, resulting in a value of 8.20. Since the Aprilis polymer is sandwiched between two 1.5mm-thick glass slides, the effect of inserting the glass slides in the signal path was considered next, leading to an SNR of 8.12. Finally, the quality of the imaging through a sample of Aprilis material bleached with uniform illumination produced an SNR of 8.65. The fact that there was no degradation of the SNR evidences the good optical quality of the material.

A single hologram was recorded with beam intensities of 2.5 mW/cm^2 for a total exposure of 1.5 mJ/cm^2. The image of the reconstructed holographic data page is in
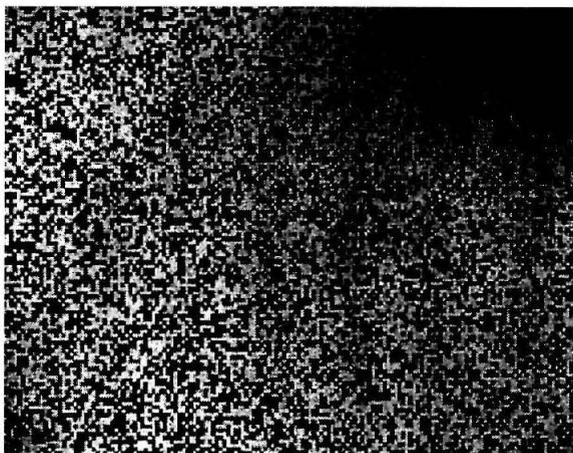
Fig. 4-8. Reconstruction of a single image-plane hologram of a 40 μm random-pixel mask.

Figure 4-8. The SNR of the reconstructed hologram was 6.45, and the BER measured from its pixel histogram (Figure 4-9) resulted in 2.38×10^{-8}. The next step was to multiplex 270 holograms in a single location. The holograms were arranged in 6 sets separated 2° in the angular direction. In turn, each set consisted of 45 peristrophic holograms separated 4°
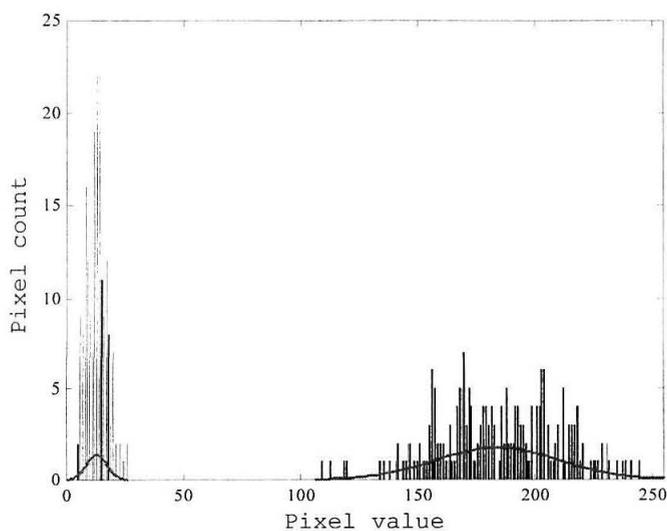
Fig. 4-9. Histogram of pixel intensities of the holographic data page in Figure 4-8. The histogram has been fitted to two normal distributions to compute the BER.

from each other. The holograms were recorded with equal exposure for 0.75 mJ/cm^2, while

the beam intensities were kept the same as in the single hologram case. Figure 4-10 shows
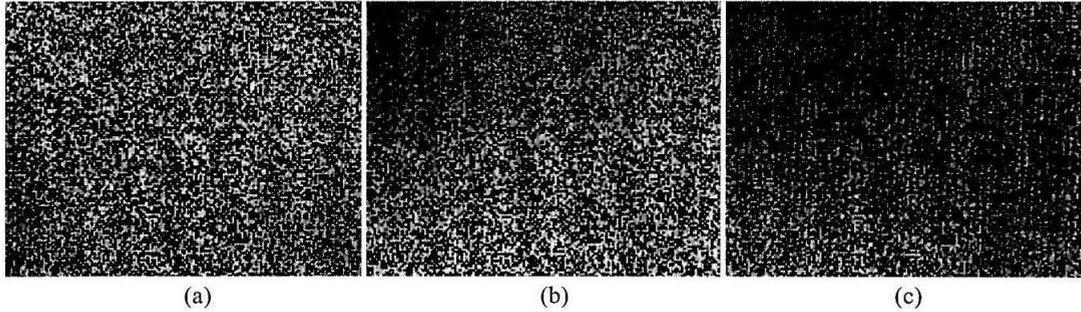
Fig. 4-10.   Reconstruction of the (a) first, (b) 135th, and (c) 270th hologram from a set of 270 image holograms recorded with equal exposures.

the reconstruction of the first, 135th and 270th holograms, ranked according to the order in

which they were recorded during the experiment. These holograms presented an averaged

SNR of 5.25. A recording schedule was derived, using again the method described in

[4-13], by fitting a sixth-order polynomial to the cumulative strengths of the measured

holograms. Using the schedule, and keeping all other parameters the same as in the

unscheduled experiment, a new set of 270 holograms was recorded. The reconstruction of

three of the holograms is presented in Figure 4-11. The measured SNR was 5.11 on aver-
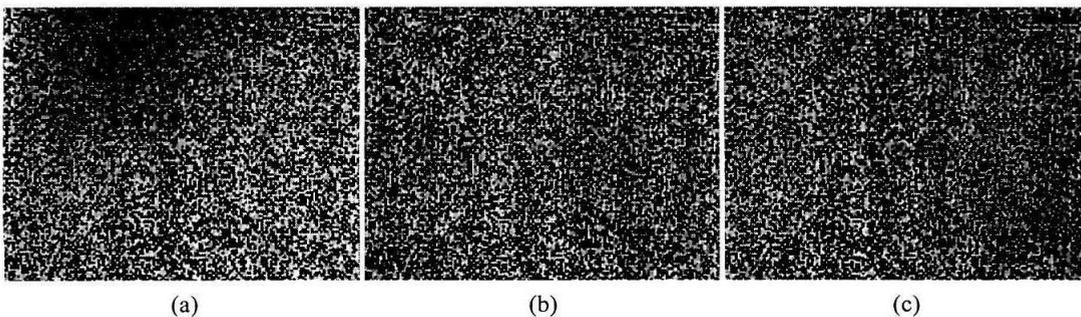
Fig. 4-11.   Reconstruction of the (a) first, (b) 135th, and (c) 270th hologram from a set of 270 image holograms recorded with exposure schedule.

age, fairly similar to the constant-exposure case. For the sake of comparison with the results

presented in Figure 4-6, the cumulative grating strength curve of this set of holograms is

shown in Figure 4-12. At the end of the experiment, after a total exposure of 0.2 J/cm^2, the

(a)

(b)
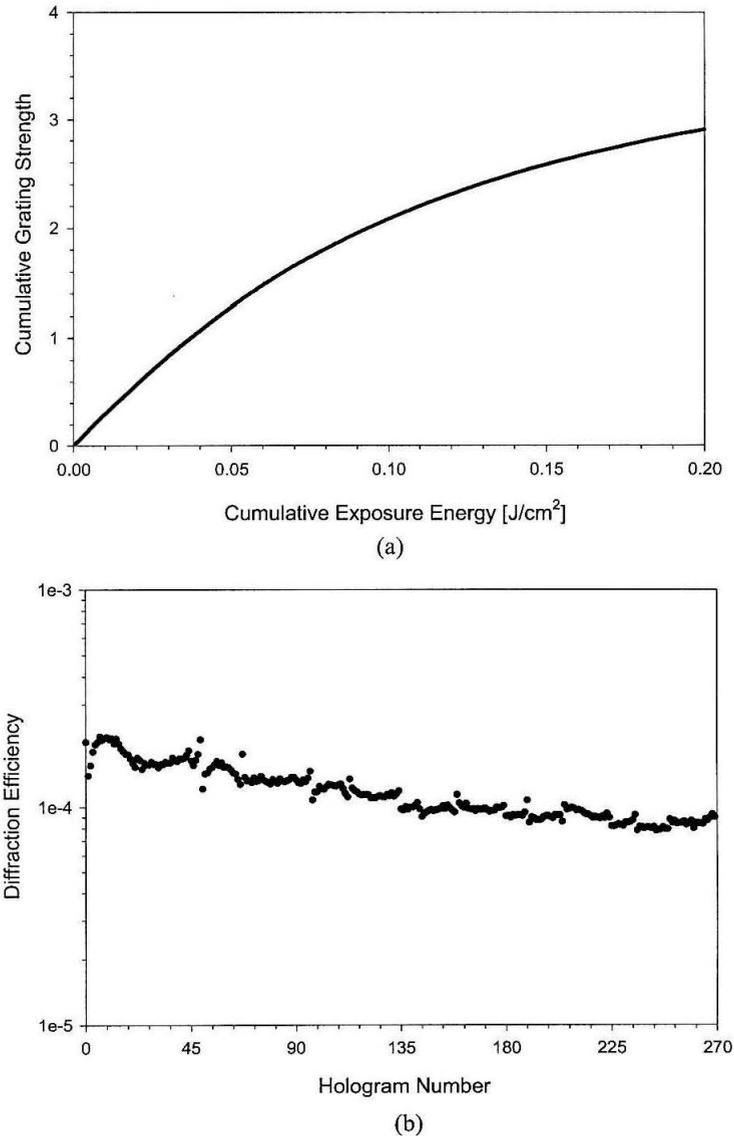
Fig. 4-12.   (a) Cumulative grating strength of 270 image-plane holograms and (b) their individual hologram strengths with an exposure schedule.

curve in Figure 4-12(a) reaches 2.91, although it has not saturated yet. Compared to the curve in Figure 4-6(a), a 50% reduction in M/# can be expected when multiplexing random-pixel images with respect to the case of plane-wave holograms due to the fact that only half of the pixels in the mask are bright.

The imaging performance of the material for reflection holograms has also been tested. In the reflection-geometry setup, the reference beam impinges the sample from behind, so the angle between the two recording beams is 147.5°. Since peristrophic multiplexing is not possible in reflection geometry[1], only angle-multiplexing was employed to store 11 holograms with constant exposure of 0.75 mJ/cm^2 and angular separation of 2° Three of the holograms are shown in Figure 4-13. Their averaged SNR was 5.20, virtually the same as the value obtained for transmission-geometry holograms.

| (a) | (b) | (c) |
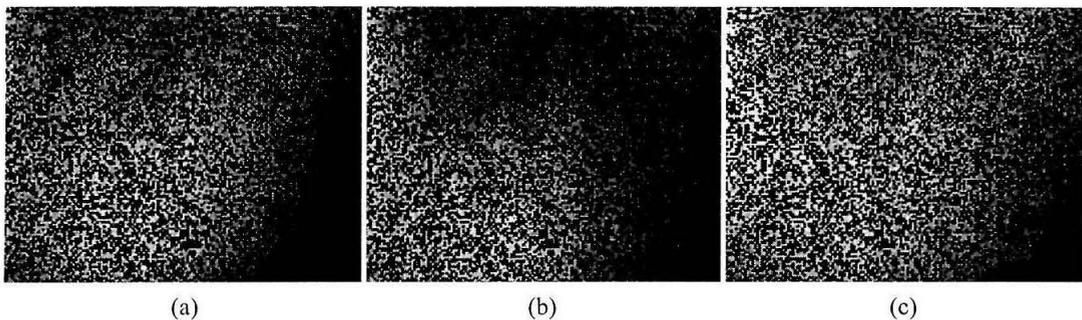
Fig. 4-13. Reconstruction of the (a) first, (b) 2nd, and (c) 11th hologram from a set of 11 reflection-geometry holograms recorded with equal exposures.

A summary of the results for image-plane holograms is provided in Figure 4-14, from which it can be seen that Aprilis material is able to store holograms with high fidelity.

---

1. The grating vector of a symmetric reflection-geometry hologram is parallel to the axis of rotation of the material, therefore the peristrophic direction becomes degenerate. In a nonsymmetric case, the amount of peristrophic rotation necessary to have the reconstruction blocked out by the spatial filter is still large enough to make multiplexing impractical.
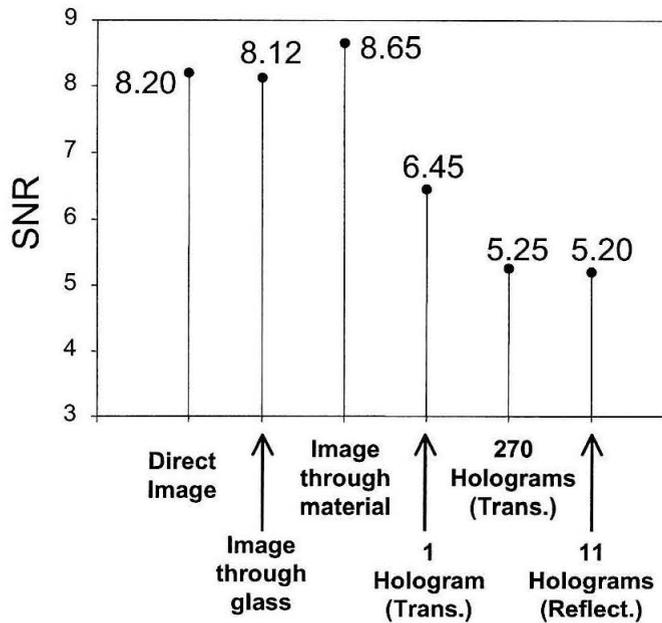
Fig. 4-14. Summary of the values of SNR obtained for image-plane recording. The horizontal axis specifies the case for which the SNR was measured.

The decay in SNR with respect to the original level above 8 is the result of distortion produced by the successive recording of holograms in the sample. As more holograms are multiplexed, local changes of the index of refraction of the medium leads to a degradation of the reconstruction of the previously recorded holograms and thus a drop in the SNR. However, due to the good optical quality of this material, high values of SNR can still be obtained.

4.3.2 Fourier-plane holograms

In many applications, like in the reflection-geometry OPGA module, it is the spectrum of the data page rather than its image that is recorded into the optical medium. While in an image-plane hologram, defects or nonuniformity in the recording medium reveal in the form of changes of intensity across the extension of the reconstructed hologram, as it

can be appreciated in the upper left corner of Figure 4-8; in the case of a Fourier-plane holo-gram, these defects will affect different parts of the signal spectrum, having a more global effect on the reconstructed image. Therefore, it is also necessary to characterize the Aprilis material when holographic recording takes place at the Fourier plane. The setup shown in Figure 4-7 has been slightly modified so the overlapping between the signal and reference beam occurs at the Fourier-plane of the first lens after the mask, while keeping the angle between the two beams and the lenses in the signal path the same as before.
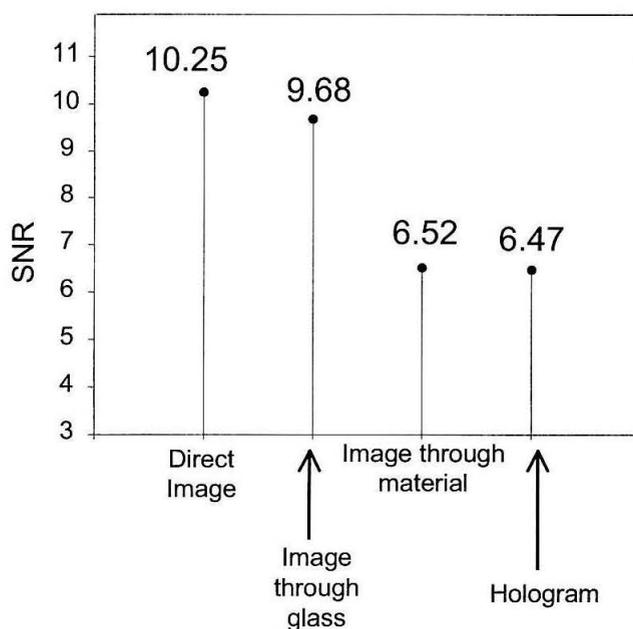
Fig. 4-15.   Summary of the values of SNR obtained for Fourier-plane recording. The horizontal axis specifies the case for which the SNR was measured.

Direct imaging, and imaging through glass, of the mask on the CCD camera resulted in very high SNR, about 10 as indicated in Figure 4-15. However the SNR dropped to 6.52 when the mask was imaged through a sample of Aprilis that had been bleached under uniform illumination. The big drop in SNR reveals how sensitive imaging is to local

defects in the recording medium. Better imaging through the sample can be obtained if this is moved slightly off the Fourier plane of the first lens, so that the focused signal spot is not inside the material. The inset in Figure 4-16 shows the reconstruction of a hologram exposed for 4.1 mJ/cm^2. Holographic recording did not significantly reduce the SNR of the system, since the hologram has an SNR of 6.47, approximately the same as for the image-plane hologram. The BER calculated from the pixel histogram in Figure 4-16 is 5.62×10^{-9}.
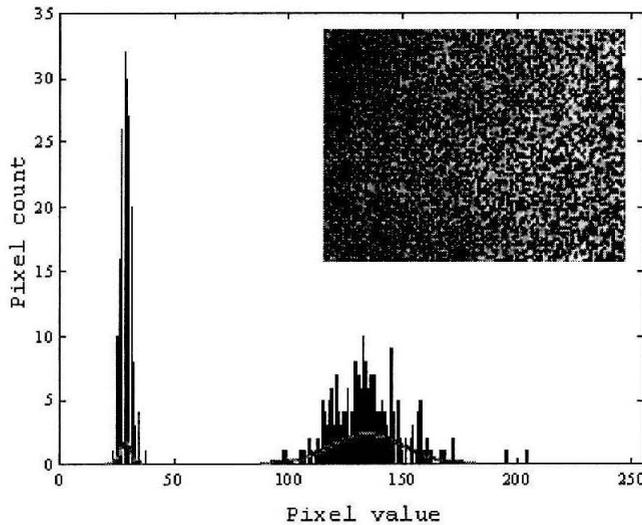
Fig. 4-16. Histogram of pixel intensities of the reconstruction of a Fourier-plane hologram (shown in the inset), and fitted normal distributions to calculate the BER.

The multiplexing of Fourier-plane holograms has also been investigated. For this particular experiment, the 40-μm pixel mask was replaced by a 640×480 Kopin SLM with 24-μm pixel pitch. The lenses of the second 4-f system in the recording setup were changed in order to demagnify the size of the pixels of the SLM, so that they match the size of the pixels in the CCD (approximately 9 μm). The intensity of the recording beams was balanced at the recording plane, being the reference 5 mW/cm^2.

A set of 99 random-pixel data pages with a minimum feature size of 4×4 SLM-pixels was generated, and stored in a single location of the material combining angular and peristrophic multiplexing. The holograms were arranged in 11 angular subsets, spaced 2.5^o from each other, each subset consisting of 9 holograms separated by 20° in the peristrophic orientation. Scheduled recording was used to make equal the strength of the holograms. Six of the multiplexed holograms are shown in Figure 4-17. The reconstructions rendered high-
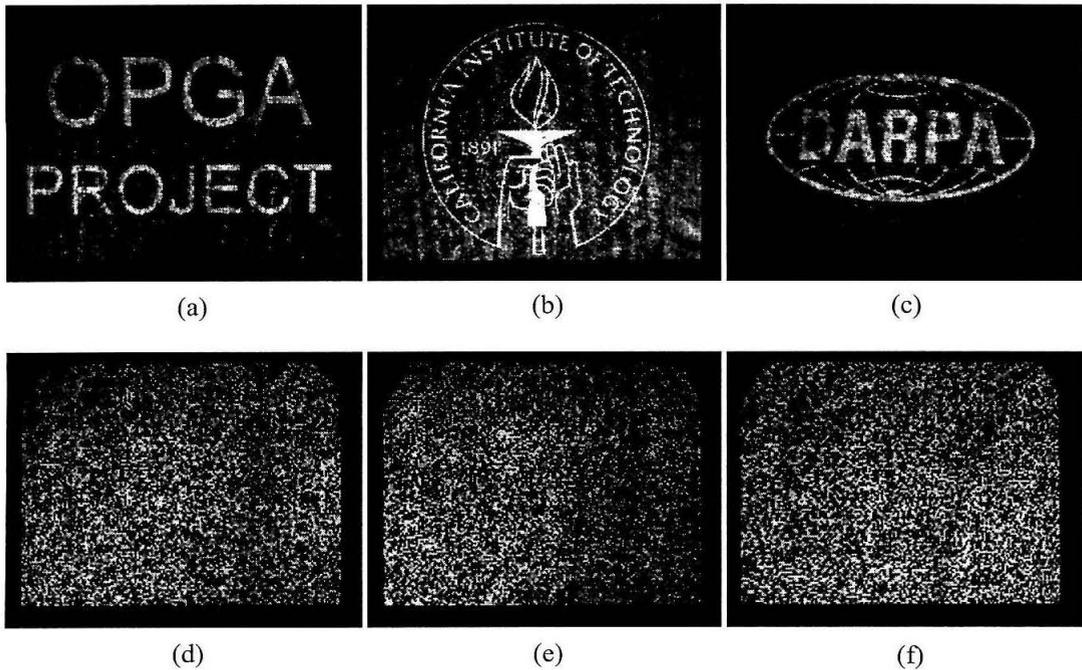
Fig. 4-17. Holographic reconstruction of the (a) first, (b) second, (c) third, (d) 10th, (e) 45th, and (f) 99th data page created with an SLM and recorded in the Fourier-plane at a single location of the material.

quality images and exhibited good equalization, as it can be observed from their comb function in Figure 4-18. It is not surprising that the first three holograms in Figure 4-18 have a lower diffraction efficiency. Unlike in random-pixel patterns, the percentage of bright

pixels in these three holograms is lower than 50%. The average diffraction efficiency is 2.35×10^{-3}, which corresponds to M/4.83.
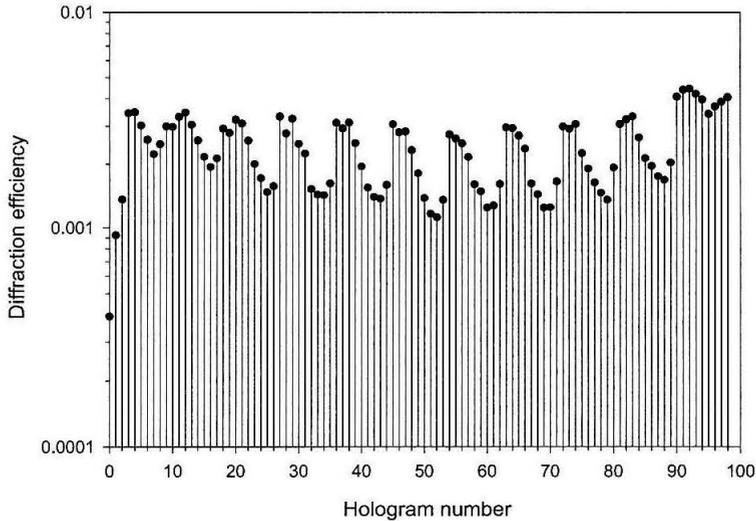
Fig. 4-18.   Individual diffraction efficiency of the 99 angular and peristrophically multiplexed holographic data pages recorded with exposure schedule.

This last result demonstrates that the Aprilis material provides a large dynamic range for the storage of high bandwidth images, and can meet the $(M/\#)_{Image}$ required to be used effectively in the OPGA. Furthermore, the performance of the material for both image-plane and Fourier-plane recording makes indeed this material an excellent candidate for such application.

4.4 Hologram mastering and duplication

As discussed in Chapter 3, one of the requirements for the OPGA module is that it needs to be compact. In order to achieve a very simple design, the technique to read out the holograms must not require any lenses. The ability of the Aprilis material to store high-quality holograms makes possible to devise a process of hologram mastering and duplica-

tion that efficiently simplifies the retrieval of the holographic templates in the OPGA module without increasing the complexity of the recorder, nor compromising the quality of the reconstructions.

Other techniques that allow for lensless readout have been traditionally employed. One of them is to use a single-lens imaging system (e.g., a unit magnification imaging system) and record at the back focal plane of the lens, before the detector. The main disadvantage of this approach is that it is extremely difficult for a single lens, even for a custommade one, to obtain good imaging quality over a large area. As a variation of this method, two lenses arranged in a 4-f system can be utilized to relay the contents of the SLM onto the detector. In this case, holographic recording takes place on the Fresnel region between the second lens and the detector. A pair of lenses can more effectively correct for aberrations in the system and achieve better image quality. However, customized lenses would still be necessary for large data pages. Furthermore, sometimes this solution may not be viable if the lenses have very short working distance.

An alternative approach is phase-conjugate readout [4-15]–[4-17], which unlike the other two solutions does not require any special optics in the system. The quality of the reconstructed holograms does not depend on the imaging optics, but on how well the wavefront of the counter-propagating readout beam matches the one of the recording beam. This turned out to be problematic in the context of the OPGA, because the reference beam during recording should be a converging spherical wave that exactly matches the divergent beam emitted by each VCSEL, thus making the requirements on the VCSELs more stringent.

The technique to master and duplicate holograms investigated in this section

obtains the same imaging quality as in the phase-conjugate readout case, even with off-the-

shelf lenses in the recorder, while avoiding the complexity of phase-conjugating a spherical

reference beam in the OPGA. This technique splits the holographic recording process into

two steps. During the first one, a master hologram (i.e., a high-quality hologram with high

diffraction efficiency) is recorded in a 200 μm thick Aprilis film by interfering a focused

signal beam with a plane-wave reference, as sketched in Figure 4-19. The master hologram
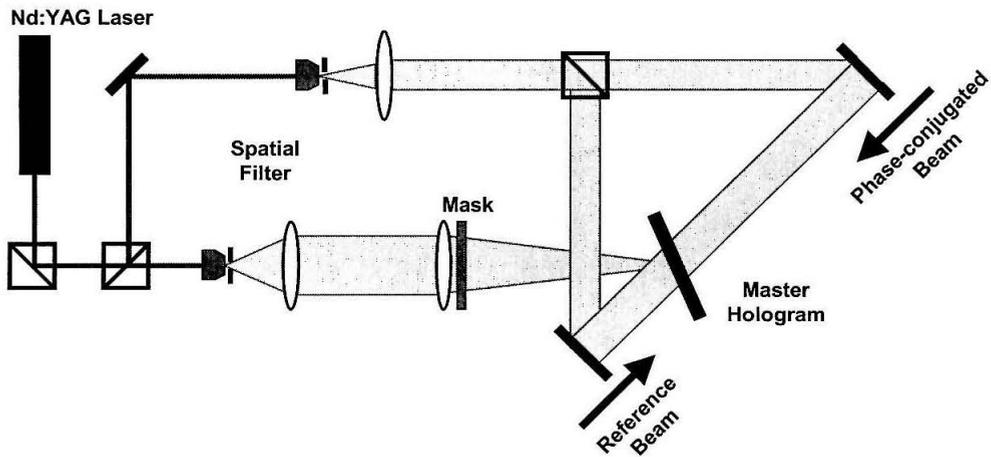
Fig. 4-19.   Schematic of the setup for the mastering process. Recording occurs between the signal beam and the forward-propagating reference beam. Upon read-out, the phase-conjugated reference is used to pixel match the hologram to the CMOS imager (not shown in the figure).

is read out using a phase-conjugate of the reference beam. Since the reference beam is a

plane-wave, it is easier to generate a high-fidelity phase-conjugated wavefront. At this

time, the mask and the focusing lens are removed from the setup and replaced by a CMOS

imager, so the reconstruction of the hologram can be pixel-matched.

In the experimental setup, the 532 nm line of a Nd:YAG laser was used to record the master hologram. The signal travels through a 65 mm focal-length lens before it impinges a 24-μm pixel mask. The phase-conjugated reconstruction of the hologram, displayed in Figure 4-20, was pixel-matched to a CMOS imager with 12 μm pitch, thus each pixel in the hologram is double sampled by the detector.
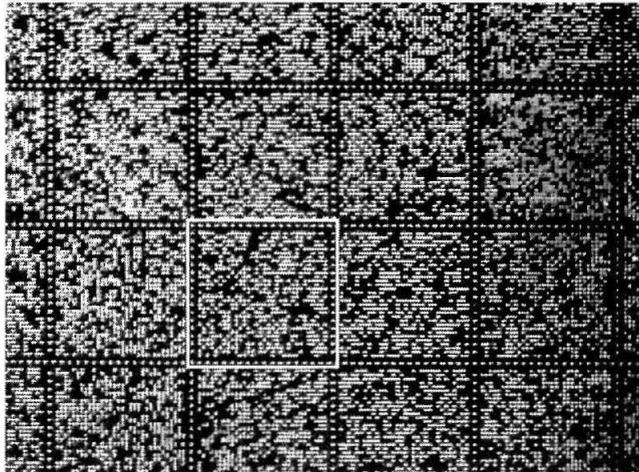
Fig. 4-20.   Picture of the pixel-matched and phase-conjugated master hologram as detected by the CMOS imager. The hologram pixel size is 24 μm.

Once the master hologram has been pixel-matched, it can be copied to other samples of Aprilis. During the second step, the duplication process, a fresh film is placed between the master hologram and the CMOS camera, as suggested in Figure 4-21. The duplicated hologram is recorded by the interference between the phase-conjugated reconstruction of the master and a spherical reference beam. The light emitted from the tip of an optical fiber provided a spherical wavefront to simulate the beam emitted by the VCSEL. Moreover, the sample for duplication was mounted on a translation stage so several copies of the master hologram can be multiplexed by translating the material.
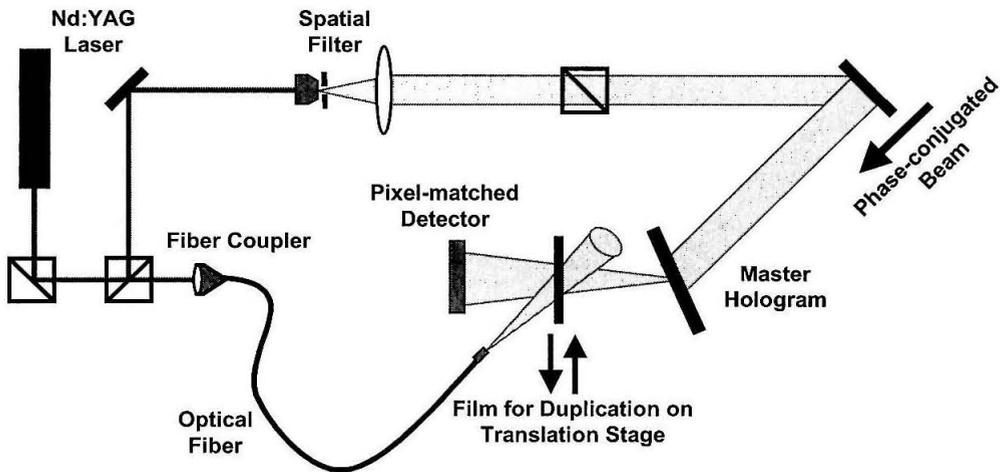
Fig. 4-21.   Setup for the hologram copying process.The duplicated hologram is recorded by the interference between the reconstruction of the master hologram and a spherical wave reference.

In the actual experiment, the raw beam of the laser was expanded and used directly as reference to record a copy of the master hologram. The region of the hologram encompassed by the white square in Figure 4-20 has been magnified in Figure 4-22(a) to facilitate the comparison with the same portion of the duplicated hologram, which is presented in Figure 4-22(b). Despite the poor quality of the reference beam, the results are noticeable. From the two pictures, no degradation in the duplicated hologram can be appreciated, proving the high degree of fidelity of the copying process.

(a)   (b)
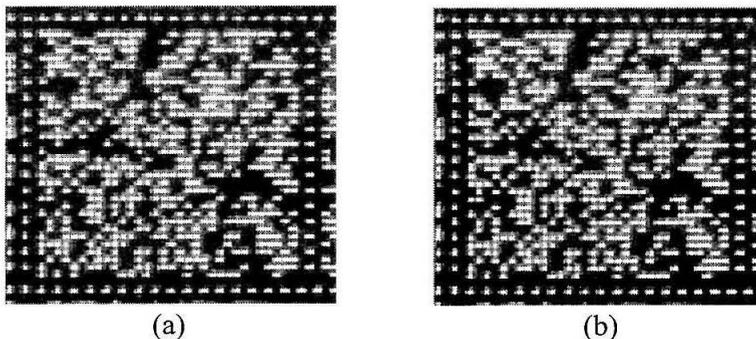
Fig. 4-22.   (a) Detailed view of the master hologram (square block encircled by the white line in Figure 4-20), and (b) the same part of the data page retrieved from the copied hologram.

References

[4-1] W. S. Colburn, and K. A. Haines, "Volume hologram formation in photopolymer materials," *Applied Optics* **10**, 1636–1641 (1971).

[4-2] S. Calixto, "Dry polymer for holographic recording," *Applied Optics* **26**, 3904–3909 (1987).

[4-3] C. Carre, D. J. Lougnot, Y. Renotte, P. Leclere, and Y. Lion, "Photopolymerizable material for holographic recording in the 450–550 nm domain: characterization and applications II," *Journal of Optics* **23**, 73–79 (1992).

[4-4] D. J. Lougnot, and C. Turck, "Photopolymers for holographic recording: II. Self-developing materials for real-time interferometry," *Journal of Optics* **1**, 251–268 (1992).

[4-5] A. Fimia, N. Lopez, F. Mateos, R. Sastre, J. Pineda, and F. Amat-Guerri, "New photopolymer used as a holographic recording material," *Applied Optics* **32**, 3706–3707 (1993).

[4-6] S. Martin, C. A. Feely, and V. Toal, "Holographic recording characteristics of an acrylamide-based photopolymer," *Applied Optics* **36**, 5757–5768 (1997).

[4-7] S. Blaya, L. Carretero, R. Mallavia, A. Fimia, R. F. Madrigal, M. Ulibarrena, and D. Levy, "Optimization of an acrylamide-based dry film used for holographic recording," *Applied Optics* **37**, 7604–7610 (1998).

[4-8] F. Zhao, E. E. E. Frietman, and X. Li, "Novel type of red-sensitive photopolymer system for optical storage," *Proceedings SPIE on Advanced Optical Memories and Interfaces to Computer Storage* **3468**, 317–321 (1998).

[4-9] P. Cheben, and M. L. Calvo, "A photopolymerizable glass with diffraction efficiency near 100% for holographic storage," *Applied Physics Letters* **78**, 1490–1492 (2001).

[4-10] D. A. Waldman, H.-Y. S. Li, M. G. Horner, "Volume shrinkage in slant fringe gratings of a cationic ring-opening holographic recording material," *Journal of Imaging Science and Technology* **41**, 497-514 (1997).

[4-11] D. A. Waldman, H.-Y. S. Li, and E. A. Cetin, "Holographic Recording Properties in Thick Films of ULSH-500 Photopolymer," *Proc. SPIE of Diffractive and Holographic Device Technologies and Applications* **3291**, 89-103 (1998).

[4-12] D. Waldman, R. Ingwall, E. Cetin, M. Horner, R. Minns, and P. Panchu, "Advantages of the CROP Technology in Aprilis Photopolymer Recording Media for High Performance Holographic Storage Systems," *Eighth NASA Goddard Conference on Mass Storage Systems and Technologies*, 2000.

[4-13] A. Pu, K. Curtis, and D. Psaltis, "Exposure schedule for multiplexing holograms in photopolymer films," *Optical Engineering* **35**, 2824-2829 (1996).

[4-14] K. Curtis, A. Pu, and D. Psaltis, "Method for holographic storage using peristrophic multiplexing," *Optics Letters* **19**, 993-994 (1994).

[4-15] J. P. Huignard, J. P. Herriau, G. Rivet, P. Gunter, "Phase-conjugation and spatial-frequency dependence of wave-front reflectivity in $Bi_{12}SiO_{20}$ crystals," *Optics Letters* **5**, 102-104 (1980).

[4-16] J. J. P. Drolet, E. Chuang, G. Barbastathis, D. Psaltis, "Compact, integrated dynamic holographic memory with refreshed holograms," *Optics Letters* **22**, 552-554 (1997).

[4-17] W. H. Liu, and D. Psaltis, "Pixel size limit in holographic memories," *Optics Letters* **24**, 1340-1342 (1999).

5 Optimization of PQ-PMMA polymer

5.1 Introduction

The optical properties of thick polymers, like PQ-doped PMMA, make them well-suited in many holographic applications. PQ-PMMA is a material that exhibits hologram amplification through diffusion. The long post-processing of the PQ-PMMA samples after recording, necessary to reveal the holograms, is seen as one of the disadvantages of this material. Therefore, this chapter will focus on the optimization of the material in order to reduce the duration of the post-processing, or even make it unnecessary.

After a brief description of the material and an explanation of the diffusional mechanism of grating formation in PQ-PMMA in Section 5.2[1], different options to increase the speed of hologram formation are investigated. Section 5.3 explores the impact of the residual concentration of monomer in the polymeric matrix on the diffusion speed by means of a comparative analysis between the material used at the California Institute of Technology and the one fabricated at the National Chiao Tung University. Section 5.4 tackles the effect of adding plasticizer to the material. Although the plasticizer is very effective in increasing the speed of hologram formation, it will compromise the stability of the gratings. Finally, in Section 5.5, the performance of the material is tested when the basic PQ molecule is replaced by other PQ-based compounds.

---

1. A comprehensive characterization of the optical properties of this material, and its use as support for optical data storage, can be found in [5-1].

The last section in this chapter explores the area of applications for which the PQ-PMMA material could be utilized. Section 5.6 deals particularly with the possibility to implement narrow band filters for WDM applications using this material.

5.2 Phenanthrenequinone-doped PMMA

Phenanthrenequinone- (PQ-) doped poly(methyl methacrylate) (PMMA) [5-2], [5-3] has been used as a recording material in optical memories and other holographic systems [5-4]-[5-7]. This material consists of a polymeric basis doped with chromophores, the PQ molecules. This material is lightweight and durable, and does not suffer from shrinkage. High-optical quality samples of different shapes and thicknesses can be obtained. These properties make it an excellent candidate for holographic memory modules.

5.2.1 Recording and hologram formation process

The process of hologram formation in PQ-PMMA is depicted in Figure 5-1. Persistent holographic recording is achieved by interfering two beams of coherent light inside the material. An interference pattern with a spatial period Λ of the form

$$I = I_o\left[1 + \cos\left(\frac{2\pi x}{\Lambda}\right)\right] \tag{5-1}$$

induces the modulation of the index of refraction of the material by optically activating the PQ molecules, in the bright fringes, and inducing their attachment to the PMMA matrix. Recording occurs at room temperature, at which the diffusion of PQ molecules is negligible [5-2],[5-8]. Therefore, the attachment of PQ molecules to PMMA chains results in local changes of the concentration of index of refraction of the material with respect to those regions, the dark fringes, where the PQ molecules remain unattached. If C is the concentra-
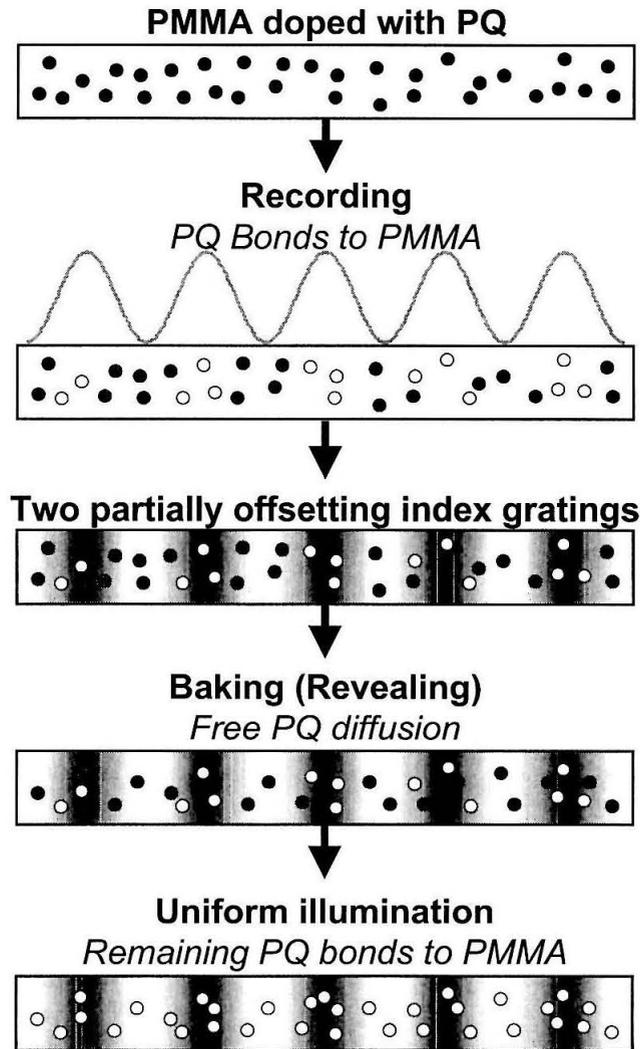
Fig. 5-1. Process of hologram formation in PQ-PMMA by means of photoinduced attachment of PQ molecules to the PMMA matrix, and subsequent hologram amplification due to diffusion of nonattached PQ molecules.

tion of PQ molecules that are uniformly distributed in a fresh PQ-PMMA sample, then once recording is completed (referred to as time t=0) the concentration of bonded PQ (C_2) follows a spatial distribution like

$$C_2(x, t = 0) = C_o\left[1 + \cos\left(\frac{2\pi x}{\Lambda}\right)\right],$$ (5-2)

in which C_o is a constant related to the exposure energy of the hologram. The distribution

of free PQ molecules (C_1) has also been modulated by the light pattern, leading to

$$C_1(x, t = 0) = C - C_2(x, 0) = C - C_o\left[1 + \cos\left(\frac{2\pi x}{\Lambda}\right)\right]. \qquad (5-3)$$

The diffraction efficiency of the hologram is proportional to the square of the mod-

ulation of the index of refraction (under the assumption of weak grating). Before any

baking is applied to the sample, the expression of the Δn is the result from the two out-of-

phase gratings and is given by

$$\Delta n(x, t = 0) = \alpha_2 \tilde{C}_2(x, 0) - \alpha_1 \tilde{C}_1(x, 0) = (\alpha_2 - \alpha_1) \cdot C_o \cos\left(\frac{2\pi x}{\Lambda}\right), \qquad (5-4)$$

where α_1 and α_2 are the polarizabilities, and $\tilde{C}_1(x, 0)$ and $\tilde{C}_2(x, 0)$ are the first Fourier

component of the concentrations of free PQ and PQ attached to PMMA respectively.

Since one grating is partially compensating the other one, the diffraction efficiency

of the hologram is weak; however, it can be enhanced by raising the temperature, for exam-

ple to 55°C in our experiments. At higher temperatures free PQ can diffuse, thus the grating

formed by unattached molecules $C_1(x, t)$ is erased, while the one formed by the attached

PQ $C_2(x, t)$ remains unaffected. The one-dimensional diffusion equation for each one of

the two gratings can be written as

$$\frac{\partial}{\partial t}C_1(x, t) = D_1\frac{\partial^2}{\partial x^2}C_1(x, t)$$

$$\frac{\partial}{\partial t}C_2(x, t) = 0 \qquad (5-5)$$

with D_1 being the diffusion constant of free PQ molecules, and assuming the diffusion con-

stant of bonded molecules $D_2 = 0$. The solution to Equation 5-5 is of the form

$$C_1(x, t) = (C - C_o) - C_o\left[\cos\left(\frac{2\pi x}{\Lambda}\right) \times \exp\left(-\frac{4\pi^2}{\Lambda^2}D_1 t\right)\right], \qquad (5\text{-}6)$$

in which the exponential decay models the erasure of the grating formed by free PQ molecules with baking. As $t \to \infty$, the erasure process is completed and the expression for the modulation of the index of refraction becomes:

$$\Delta n(x, t \to \infty) = \alpha_2 \tilde{C}_2(x, t \to \infty) = \alpha_2 \cdot C_o \cos\left(\frac{2\pi x}{\Lambda}\right). \qquad (5\text{-}7)$$

The erasure of one of the gratings results in the enhancement of the diffraction efficiency of the hologram. Photopolymers in which this type of diffusional enhancement occurs are usually called diffusion-amplified materials. The amplification of the revealing process can be defined as the ratio of Δn between the after and before baking values

$$A = \frac{\Delta n(x, t \to \infty)}{\Delta n(x, t = 0)} = \frac{\alpha_2}{\alpha_2 - \alpha_1}. \qquad (5\text{-}8)$$

and purely depends on the contrast of the polarizability of the photoproduct (the PQ molecules bonded to the PMMA chains) with respect to the original PQ molecules.

5.2.2 Experimental setup

The 514 nm line of an Argon ion laser is used in a symmetric transmission-geometry setup, Figure 5-2, to record the holograms in the PQ-PMMA samples. The recording laser beams are spatial-filtered and collimated (not shown in Figure 5-2), and the angle of incidence outside the material of each one of them is approximately 30° with respect to the surface normal of the sample, resulting in a grating period Λ of 0.5 μm. A 633 nm He-Ne laser beam, outside the absorption band of PQ-PMMA, is Bragg-matched to the hologram to monitor the real-time growth of the hologram during recording, the post-exposure

dynamics of the sample once the 514 nm beams have been shut off, and the evolution of

the grating strength with baking.
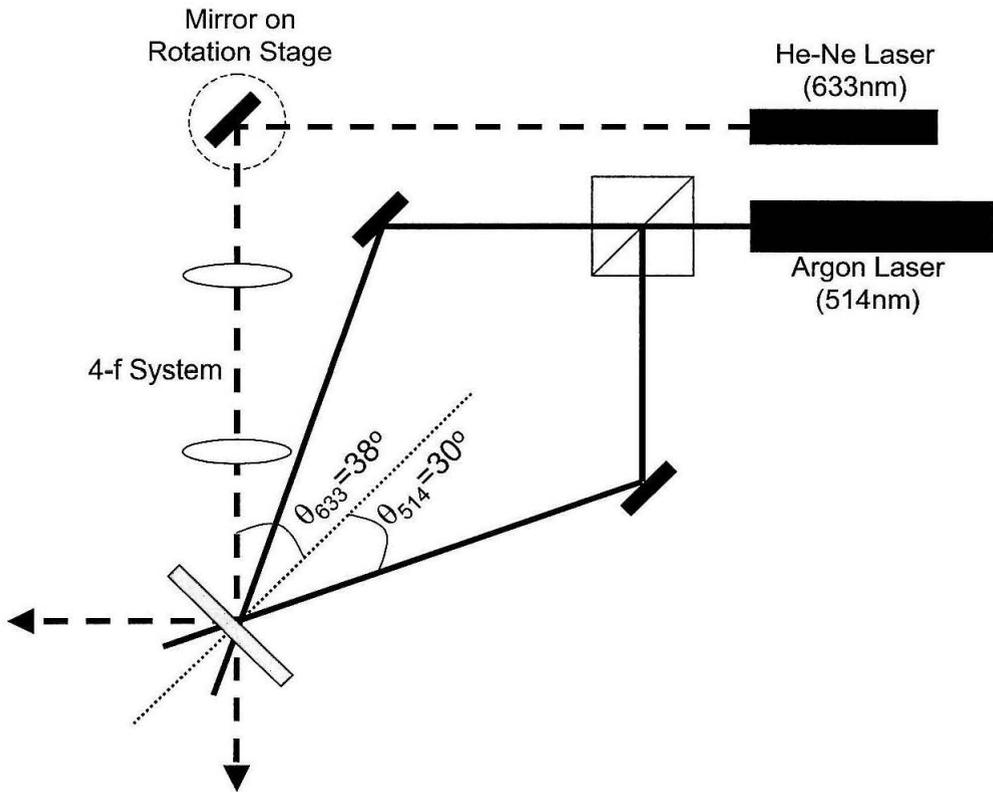
Fig. 5-2.   Experimental setup used to record gratings in PQ-PMMA samples. The He-Ne beam can be Bragg-matched to the sample with a mirror on a rotation stage and a 4-f system to monitor the post-recording dynamics of the sample.

The red beam is reflected off a mirror mounted on a motorized rotation stage and

passes through a 4-f system before impinging the optical medium. This allows to correct

the angle of incidence of the red beam for changes in the bulk index of refraction of the

material, or even take into account the possibility of residual shrinkage, while monitoring

the post-exposure dynamics of the material. Furthermore, the fact that the experimental

setup makes possible to take selectivity curve measurements of the holograms with the He-

Ne beam is essential to study the effect of baking on the holograms because the baking does not take place *in situ*, and therefore the hologram needs to be perfectly Bragg-matched every time that the sample is repositioned into the setup after having been baked in the oven for a period of time.

5.3 Effect of the residual concentration of MMA

A different type of PQ-PMMA material has been developed at the National Chiao Tung University (NCTU) and used for optical data storage applications. The material fabricated at NCTU seems to contradict the mechanism of hologram formation and amplification described in Section 5.2.1, since high diffraction efficiency holograms have been reported in these PQ-PMMA samples with recording at room temperature without the need for any post-recording thermal treatment of the samples [5-6],[5-9]. It is believed that in this case, the higher concentration of monomer in the sample plays a key role in the recording process, because the PQ molecules seem to prefer to attach to MMA rather than to the PMMA matrix [5-9]. An alternative explanation for hologram formation is then regarded, in which at room temperature MMA and PQ molecules can diffuse in the material. However, the PQ-MMA groups, formed due to photoexcitation of the PQ during recording, experience only minimal diffusion. Thus no baking is required to amplify the hologram, since continuous diffusion of PQ is building up the PQ-MMA grating while at the same time erasing any out-of-phase grating of free PQ.

In order to find a common theory that can satisfactorily explain the distinct hologram dynamics exhibited by the PQ-PMMA material fabricated at NCTU and the one developed at Caltech, comparative experiments to monitor the recording and baking pro-

cesses have been performed on samples from both universities. The samples used in the comparison are a 1 mm thick disk of 40 mm in diameter made at Caltech and a bar 20×36 mm in cross section and 2.4 mm thick made at NCTU. The concentration of PQ is 0.5% wt, the same in both samples.

5.3.1 Sample preparation process

Sample preparation consists of dissolving PQ molecules ($\leq 0.7\%$) in liquid methyl methacrylate (MMA) together with azo-bis-isobutirylnitrile (AIBN), a polymerization thermal initiator. This solution is poured into molds and allowed to polymerize in a pressure chamber. The preparation process followed at Caltech differs from the one followed at NCTU in the temperature at which the pressure chamber is set during polymerization. For the Caltech material, the temperature of the chamber is set to 80°C. On the other hand, at NCTU the polymerization process is split into two steps [5-5],[5-9]. First, the solution is let to rest at room temperature for approximately 120 hours until the solution turns homogeneously viscid. At this point, the temperature of the chamber is increased to 45°C for 24 hours to complete the process. After the polymerization is completed, the residual concentration of monomer in the Caltech samples (~1%) is much smaller than in those from NCTU (~10%). This excess of monomer will be, as it will be argued, responsible for the very distinct properties of both PQ-PMMA materials.

5.3.2 Comparison of the recording dynamics

A weak hologram is recorded in the NCTU and Caltech samples using the setup described in Section 5.2.2. The exposure energies are 0.3 and 0.5 J/cm2 respectively. This energy is delivered to the sample in an interval of less than 10 sec, minimizing thus the

effect of instability in the recording setup. Figure 5-3 shows the dark enhancement of the

diffraction efficiency of the hologram in the NCTU (solid line) and Caltech (dashed line)

samples during a period of 5 hours after recording. The square root of each curve, the grat-

ing strength, has been fitted to a mono-exponential function with time constant 23.05 min-

utes for the NCTU sample and 16.96 minutes for the Caltech sample. From these results,

there are two issues that need to be interpreted: First of all, the reason why this time con-

stant seems to be the same in both cases. And secondly, why the saturation level in the

NCTU sample occurs at a much higher value of diffraction efficiency than in the Caltech

sample.
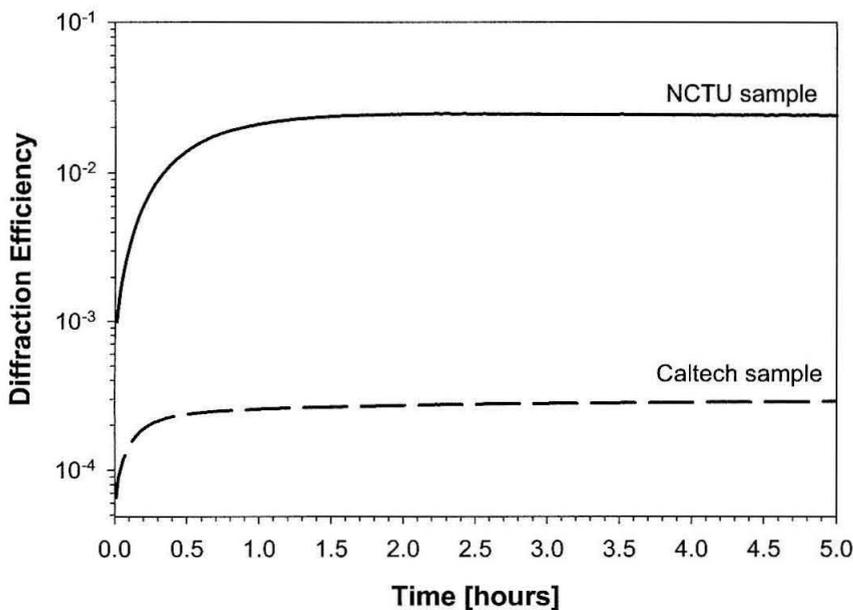
Fig. 5-3. Comparison between the evolution of the diffraction efficiency of a hologram recorded in the NCTU sample (solid line) and in the Caltech sample (dashed line) for five hours after exposure at room temperature.

As far as the time constant is concerned, the experimental results on both PQ-

PMMA media suggest that the same mechanism is taking place in both samples and dom-

inates the post-recording dynamics. Then the question is to identify what this process is. It

is central to the understanding of the dynamics of the material to verify the possibility that

free PQ molecules may diffuse in the polymeric matrix at room temperature, at which the

experiments were performed. This hypothesis can be checked by solving the one-dimen-

sional case of the diffusion equation at room temperature, equivalently to what was done in

Section 5.2.1. The diffusion constant of the free PQ molecules (D_1) relates to the measured

time constant τ as:

$$\frac{4\pi^2}{\Lambda^2} D_1 = \frac{1}{\tau} \tag{5-9}$$

Taking $\Lambda=0.5$ μm and $\tau\approx20$ minutes, the diffusion constant would be 10^{-11} m^2/s at 25°C.

This value is far from the value of 10^{-21} m^2/s reported in [5-2], even if the latter is adjusted

to account for the difference of plasticity of the two samples. Therefore, it can be concluded

that the observed dynamics does not correspond to the diffusion of free PQ molecules. In

fact, it is believed that the enhancement of the hologram after recording is due to the attach-

ment of photoactivated PQ radicals to either MMA or PMMA. Consequently, the observed

dynamics is the rate at which this attachment is taking place. This process is much faster

than the diffusion of PQ molecules and, in the experiment, it is completed within the first

two hours after exposure.

Regarding the question why the diffraction efficiency at its saturation value in the

NCTU sample is much higher than it is in the Caltech sample, in the experiment shown in

Figure 5-3 it is about 2 orders of magnitude, and having already ruled out the possibility of

PQ molecules diffusing at 25°C from dark into bright regions of the material, the only spe-

cies that could still diffuse would be the MMA. The fact that MMA is about two and a half

times lighter than PQ can explain an increase of its diffusion constant in one or two orders

of magnitude, however this would still be insufficient to account for the value obtained

from the experiment.

Conversely, it is believed that the origin of such a difference in diffraction effi-

ciency is partly found in the ratio of the polarizabilities of the PQ molecule in its different

states.[5-2],[5-8] If α_1, α_2 and α_3, are the polarizabilities and $C_1(x,t)$, $C_2(x,t)$ and $C_3(x,t)$

are the concentrations of free PQ, PQ attached to PMMA, and PQ attached to MMA,

respectively, then upon recording the two offsetting gratings are in the case of the Caltech

PQ-PMMA:

$$\Delta n(x, t) = \alpha_2 C_2(x, t) - \alpha_1 C_1(x, t), \qquad (5\text{-}10)$$

where the minus sign indicates the π phase shift between the gratings, and for the NCTU

PQ-PMMA:

$$\Delta n(x, t) = \alpha_3 C_3(x, t) - \alpha_1 C_1(x, t). \qquad (5\text{-}11)$$

It is then conjectured that the ratio between α_3 and α_1 must be larger than the ratio

between α_2 and α_1 resulting in a larger amplification and thus a stronger hologram. It is

unlikely though that this higher contrast obtained when PQ attaches to MMA can totally

account for the different behavior of the two samples. Furthermore, if we take into account

the fact that the number of PQ molecules (0.5%wt) is in either case much lower than the

number of molecules of monomer (1%wt in the Caltech and 10%wt in the NCTU samples),

it is not unreasonable then to think that even in the Caltech PQ-PMMA part of the grating

is recorded in PQ molecules attaching to the monomer.

What we consider to be the main reason for the discrepancy in behavior of the samples can be found in [5-7]. As described in the reference, PQ can act as photoinitiator of polymerization in those situations where there is an important excess of residual monomer, as it happens in the NCTU material. In these situations, the photopolymerization is combined, and can even dominate, the diffusion mechanism of hologram formation. In other words, in the NCTU PQ-PMMA, the photoactivated PQ molecules act as initiators of the polymerization before they eventually attach to MMA, resulting in substantial modulation of the index of refraction, and thus strong holograms.

5.3.3 Comparison of the baking dynamics

The effect of baking the samples on the diffraction efficiency of the holograms has also been investigated. Six holograms were recorded on different locations on the Caltech sample with exposure energies ranging from 2 J/cm^2 up to 20 J/cm^2. The recording sensitivity of the material can be obtained by fitting the grating strength of the holograms to a line, Figure 5-4(a). In the figure, only five of the six holograms are presented. The hologram with exposure energy of 20 J/cm^2 is omitted because of saturation of the material. From the measurements, a prebaking sensitivity of 0.103 cm/J is obtained. The sample was then taken out of the recording setup and placed in an oven at 55°C. The diffraction efficiency of the holograms has been monitored and it is plotted in Figure 5-5(a) as a function of the baking time. As expected, at 55°C the diffusion of free PQ molecules is greatly enhanced so after just 18 hours in the oven, the free PQ grating has been practically totally erased. On average, the holograms amplified their diffraction efficiency by a factor of 22.78, achieving values as high as 60% for the longest exposures. At this point, the "after

baking" sensitivity is measured, resulting in 0.396 cm/J, Figure 5-4(a). This increase is consistent with the amplification of the holograms during baking.

A similar experiment is performed on the NCTU sample. Five holograms were recorded with energetic exposures from 0.25 J/cm^2 to 1.5 J/cm^2. The holograms have been allowed to develop for 2.5 hours in the dark before placing the sample in the oven. At this point the holograms were measured, and their grating strengths fitted to a line, Figure 5-4(b), resulting in an equivalent sensitivity of 1.34 cm/J, which is approximately 3.4 times higher than the one obtained in the Caltech sample. The diffraction efficiency of the holograms was monitored after incremental baking was being applied to the sample. The result that was obtained, as shown in Figure 5-5(b), is clearly different from the one experienced by the Caltech sample. The holograms are not enhanced but degraded by the baking. The evolution of the strength of the gratings can be fitted to exponential functions, and the time constant is estimated to be 9.65 hours.

The reason for the degradation is that at 55°C both free PQ molecules and PQ-MMA groups can diffuse in the polymer matrix, so it is not that one of the two out of phase gratings is being erased but both of them, resulting in the destruction of the hologram. Once again, the estimated time constant can be converted to a diffusion constant using Equation 5-9. For Λ=0.5 μm and τ=9.65 hours, the equivalent diffusion constant for free PQ and PQ-MMA would be 1.9×10^{-19} m^2/s at 55°C, which is in good agreement with the value of 1.6×10^{-19} m^2/s extrapolated from data reported in [5-2].

(a)

(b)
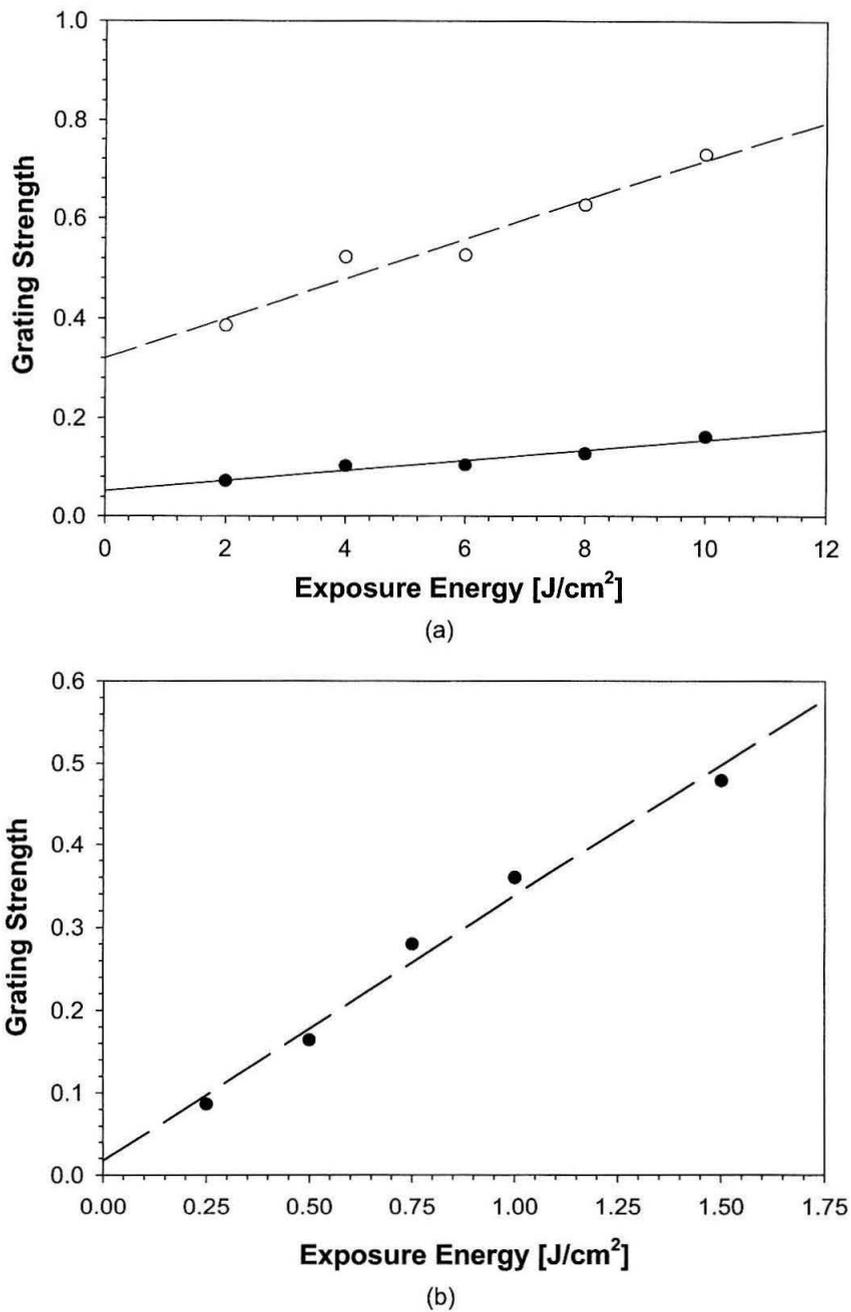
Fig. 5-4.   Comparison of the recording sensitivity: (a) Caltech sample: Steady-state grating strength of the holograms before baking (black dots and solid line) and after 18 hour baking (white dots and dashed line) as a function of the exposure energy; (b) NCTU sample: Steady-state grating strength of the holograms (before baking only).

(a)

(b)
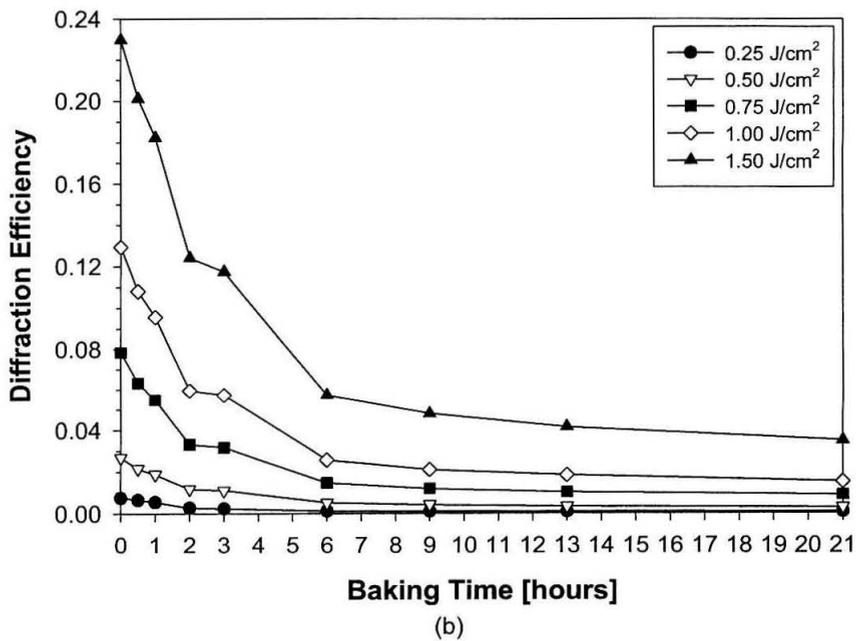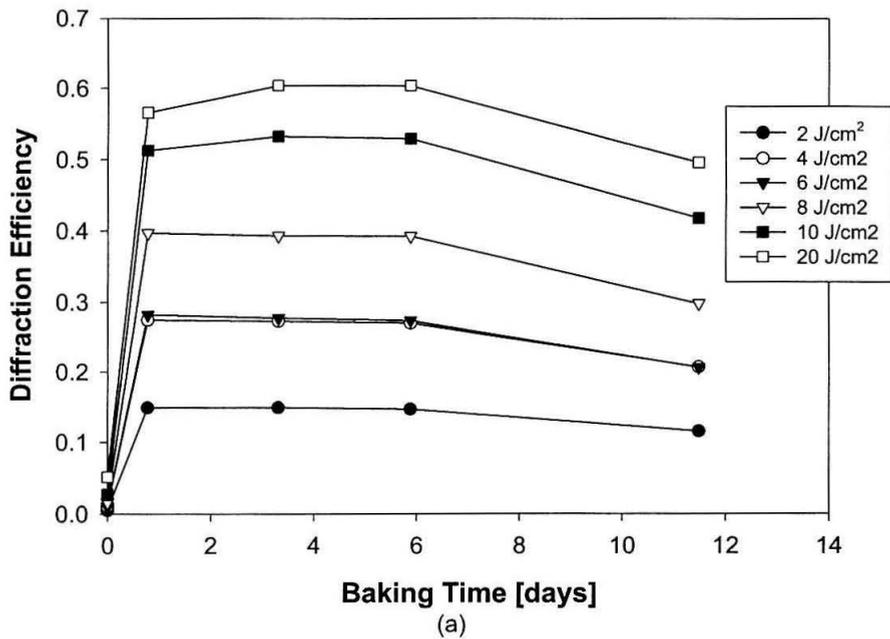
Fig. 5-5. Effect of baking at 55°C on the diffraction efficiency as a function of time: (a) Amplification of the holograms recorded in the Caltech sample; (b) Erasure of the holograms recorded in the NCTU sample.

5.3.4 Prolonged baking

The effect of prolonged baking of the sample has also been analyzed. In the case of the Caltech sample, the strength of two holograms, each with exposure energy of 20 J/cm^2, has been monitored for up to 44 days of continuous baking, Figure 5-6(a). After the initial amplification, further baking did not significantly degrade the grating formed by the PQ attached to the PMMA. On the other hand, in the NCTU material the destruction of the hologram is not complete. As it can be seen in Figure 5-6(b), the grating strength curves initially decay due to the mentioned diffusion process, but eventually flatten off to some final value. The remaining grating is due to the modulation of index refraction caused by the polymerization of monomer in the illuminated areas, as well as by the fraction of PQ that attached to the PMMA. This grating is stable and will not degrade with additional baking. The presence of a residual grating supports the hypothesis that the polymerization mechanism is responsible for the hologram recording in the NCTU material.

5.3.5 Conclusion

From the comparison between the NCTU and Caltech PQ-PMMA materials, it is concluded that the difference in their behavior lies in the different concentration of residual MMA in the samples. Experimental evidence shows that during recording PQ molecules attach to MMA but no diffusion takes place at room temperature. However, the excess of monomer during recording enables photoinduced polymerization as a mechanism for hologram formation leading to high diffraction efficiencies without the need of baking. The grating formed by the PQ-MMA groups is unstable and it can be erased within a few hours of baking.
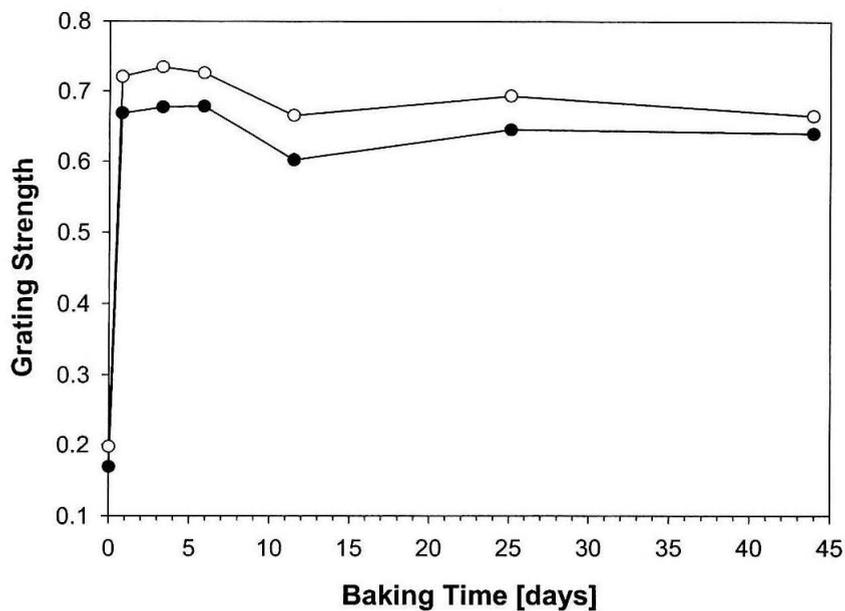
(a)
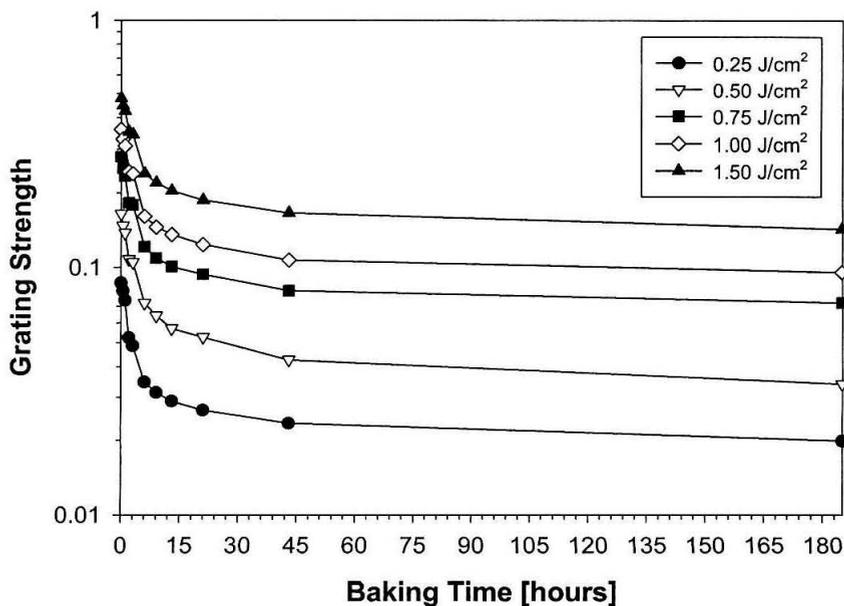
(b)

Fig. 5-6.   Effect of prolonged baking: (a) Evolution of the grating strength of 2 holograms (exposure energy 20 J/cm^2) recorded in the Caltech sample as a function of baking time; (b) Decay of the grating strength of the holograms recorded in the NCTU sample as a function of baking time.

5.4 Role of the plasticizer

The slow diffusion of free PQ molecules in PMMA is regarded as being one of the major problems of the PQ-PMMA material. The samples have to be baked in an oven to facilitate the diffusion of PQ molecules, responsible for the revealing of the hologram. This process typically takes about one or two days at 55°C. It would be then desirable to be able to modify the composition of the material to increase the diffusion speed of PQ molecules and, consequently, reduce the revealing time of the hologram. In this section, the possibility of achieving this objective by adding plasticizer to PQ-PMMA has been explored.

Adding plasticizer to the material has the effect of lowering the glass transition temperature of the polymer.[5-10] The glass transition is a property of amorphous polymers, i.e., polymers whose chains are not arranged in ordered crystals, but just disseminated around. At high temperatures, the polymer chains have no difficulty in moving around and the material becomes rubber-like. This motion usually is not translational, in the sense that the chain as a whole is not moving any distance in one direction. Instead, the motion is mainly due to the fact that segments of the chain can wiggle around, folding or opening the chain. When the temperature is low, the chains tend to stay immobile in the material. The temperature at which this change of mobility occurs is called the glass-transition temperature (T_g). In the case of PMMA, T_g is between 100 and 120°C [5-10].

The glass-transition temperature of a polymer depends on the flexibility of the backbone of the polymer chain (e.g., the "methacrylate" part of PMMA), as well as on the groups that hang from the backbone (e.g., the "methyl" part of PMMA). These pendant groups limit how closely polymer chains can pack together. Therefore the larger the pen-

dant groups, the farther apart the chains are and the more easily they can move, thus the lower T_g.

The plasticizer plays a role equivalent to big pendant groups. The molecules of plasticizer fill in the space between the polymer chains, separating them from each other and lowering the glass-transition temperature.
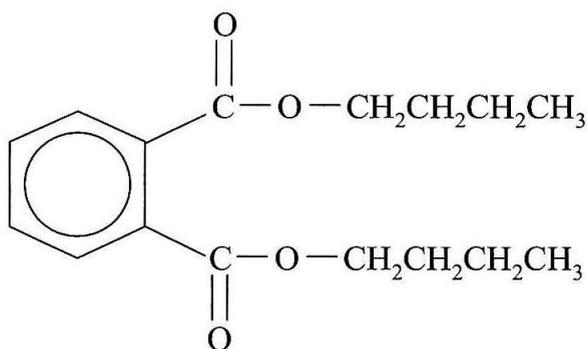
Fig. 5-7.   Dibutyl phthalate (DBPH)

The plasticizer used in the PQ-PMMA was dibutyl phthalate (DBPH). Plasticized samples were prepared by pouring DBPH into the solution with the monomer and AIBN and PQ before the polymerization process is initiated. The behavior of these samples was studied experimentally and the results are presented here.

5.4.1 Experimental method and results

The effect of the plasticizer on the hologram dynamics in PQ-PMMA can be studied by monitoring the evolution of its post-recording strength as incremental baking is applied to PQ-PMMA samples with different concentration levels of DBPH.

Three different 2mm-thick samples with a DBPH concentration of 1%, 2% and 3%, respectively, and a 3 mm thick sample containing no DBPH were used for this experiment.

Six holograms were recorded using the setup described in Section 5.2.2 on different locations of each one of the samples with exposure energies from 0.50 J/cm^2 up to 6 J/cm^2. After recording, the samples were kept in the dark at room temperature for 2 hours to allow the process of attachment of photoactivated PQ radicals to PMMA to be completed before baking the samples at $55^{o}C$.

The evolution of the strength of the holograms is shown in Figure 5-8 and Figure 5-9 as incremental baking is applied to the samples for up to 112 hours. It can be observed in the figures that as the concentration of plasticizer increases, there is a faster growth of the holograms during the post-exposure, since the diffusion of free PQ molecules is favored. However, at the same time the lifetime of the hologram is reduced, since the PMMA chains can diffuse in the medium more easily too.

The amplification factor of the hologram, defined as the ratio between the maximum strength achieved by the hologram during baking and its strength measured right after recording, is plotted in Figure 5-10. The amplification of the gratings is fairly consistent across the cases 2%, 1%, and no DBPH, resulting in 4.70 on average (solid line). The case 3% DBPH is distinct as the average amplification factor is 2.97 (dashed line). The lower amplification is due to higher levels of grating strength obtained right after exposure. This suggests that, for this concentration of plasticizer, the glass-transition temperature has been lowered to the extend that the diffusion of free PQ is greatly enhanced even at room temperature.

(a)

(b)
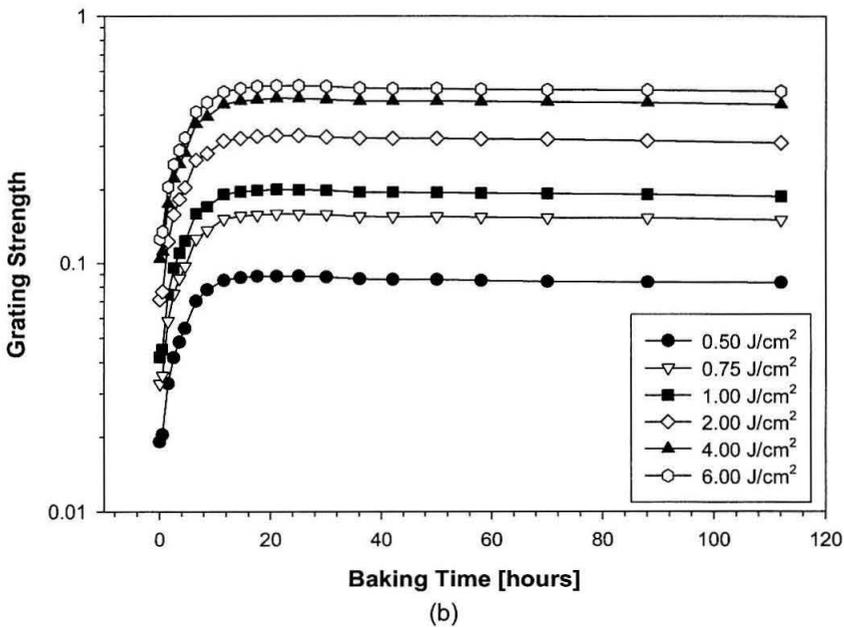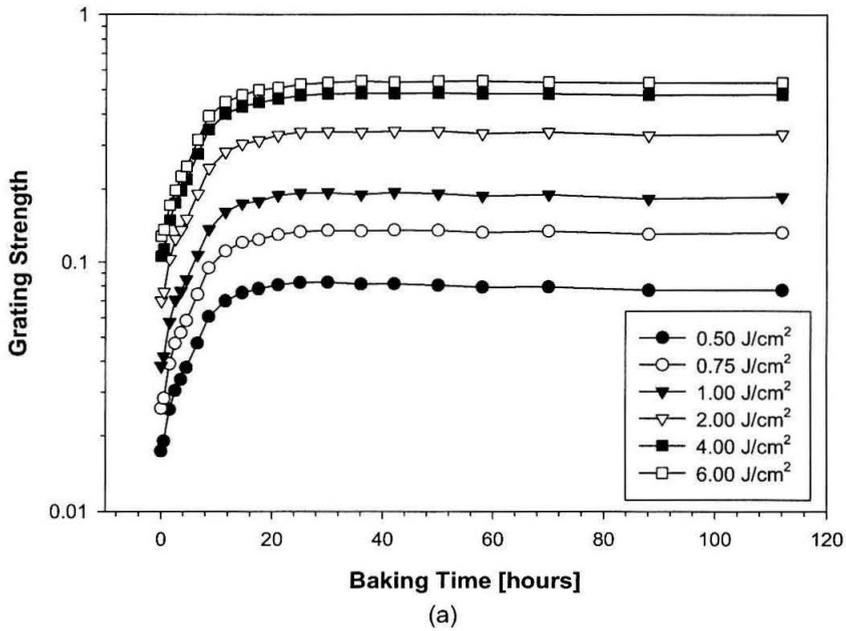
Fig. 5-8. Effect of baking at 55°C on the grating strength of the holograms as a function of time for the samples with (a) no DBPH, and (b) 1% DBPH.

(a)

(b)
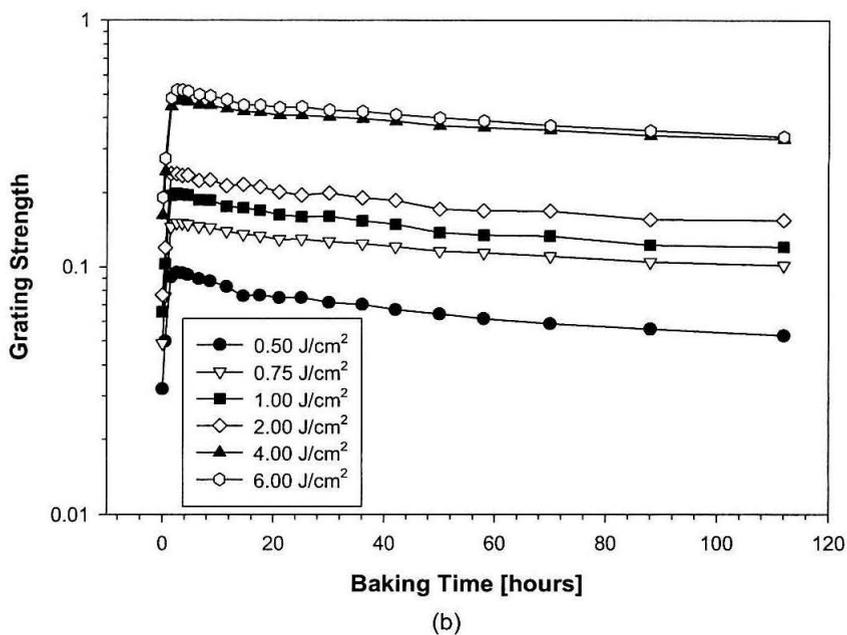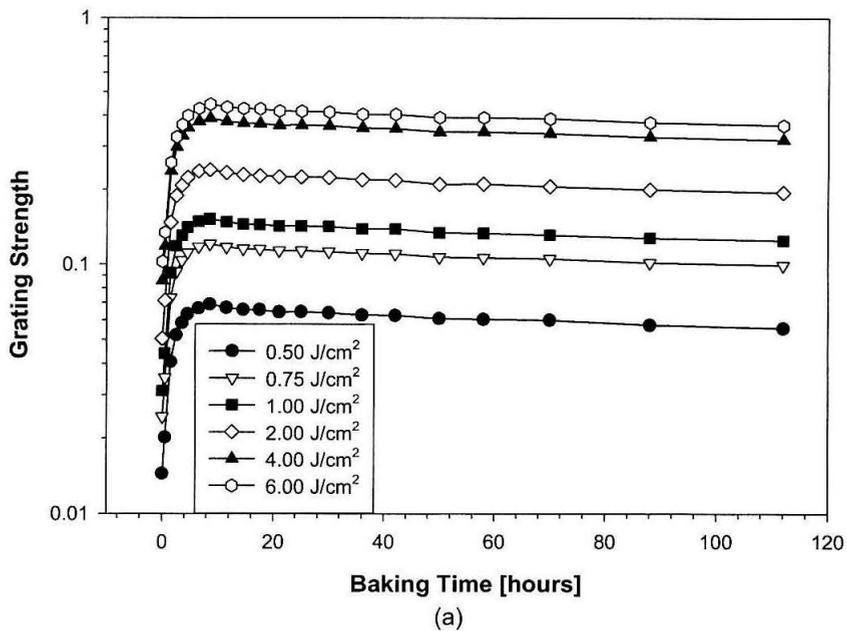
Fig. 5-9.   Effect of baking at 55°C on the grating strength of the holograms as a function of time for the sample with DBPH concentration of (a) 2%, and (b) 3%.
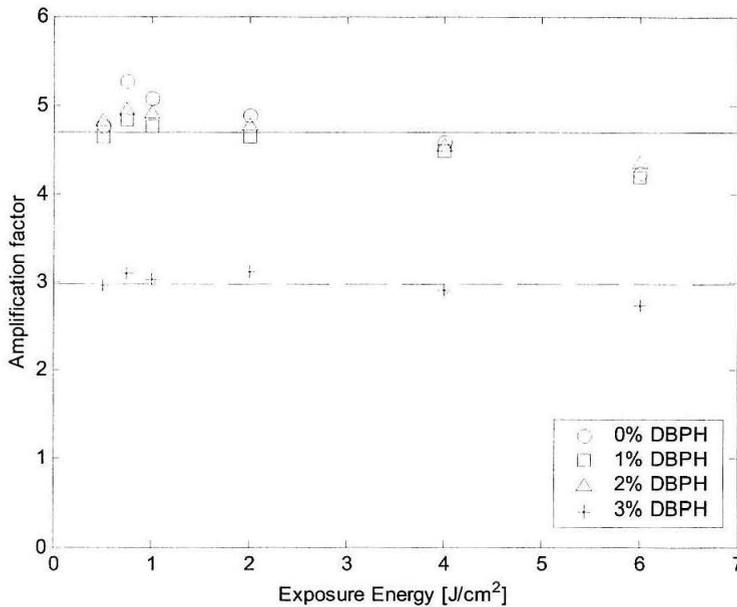
Fig. 5-10. Grating strength amplification factor as a function of the hologram exposure energy for samples with different concentration of plasticizer. The solid line is the average amplification for holograms in samples with 0%, 1%, and 2% DBPH, while the dashed line is the average for the holograms in the sample with 3% DBPH.

5.4.2 Theoretical model

The temporal behavior of the hologram is the result of the competition of two diffusional processes: One of them is the decay of the grating formed by the unattached PQ molecules, which reveals the grating formed by the PQ molecules bonded to the PMMA chains, resulting in the enhancement of the hologram. The other process is due to the diffusion of PQ-PMMA chains themselves, which leads to the decay of the grating formed by the bonded PQ molecules and therefore to the erasure of the hologram. As suggested in Figure 5-11, the evolution of the grating strength can be modeled as the combination of an exponential growth to saturation with a fast time constant τ_1 (associated to the diffusion of free PQ) and an exponential decay with a much slower time constant τ_2 (related to the diffusion of the macromolecules) in the following way

$$\sqrt{\eta} \;=\; A_o \left[1 - \exp\left(-\frac{t}{\tau_1}\right) \right] \exp\left(-\frac{t}{\tau_2}\right), \tag{5-12}$$

where A_o would be the saturation grating strength if there was no erasure.
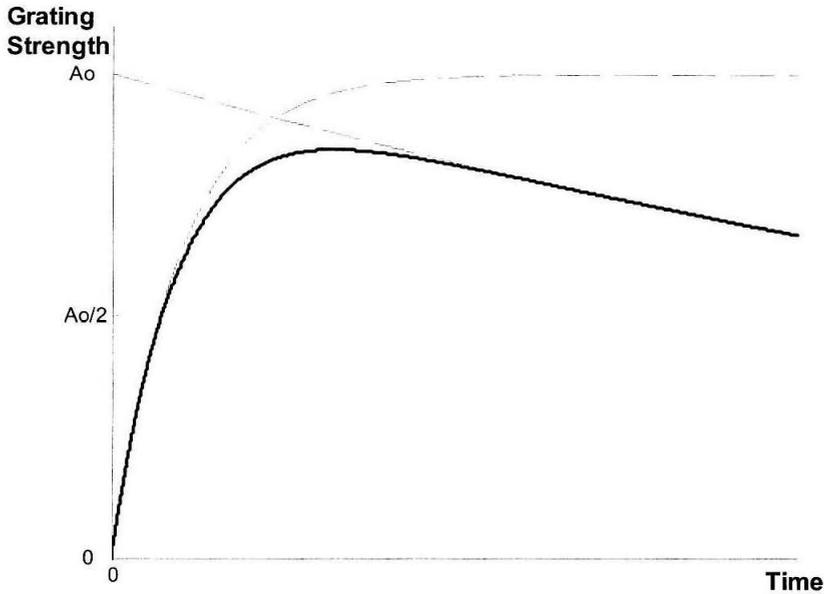
Fig. 5-11.   Theoretical model of the diffusional processes in the PQ-PMMA sample. The solid line obtains from the combination of an exponential growth (PQ diffusion) and an exponential decay (macromolecule diffusion), both depicted with dashed lines.

Given the fact that $\tau_2 \gg \tau_1$, Equation 5-12 can be simplified for the cases $t \gg \tau_1$ and $\tau_1 \gg t$. In the first case, it can be assumed that the growing exponential has reached saturation, so Equation 5-12 reduces to

$$\ln(\sqrt{\eta}) \cong \ln(A_o) - \frac{t}{\tau_2}. \tag{5-13}$$

In the case $\tau_1 \gg t$, Equation 5-12 can be approximated by the first-order Taylor's series expansion around the origin as

$$\sqrt{\eta} \cong \sqrt{\eta}\big|_{t=0} + \frac{d}{dt}\sqrt{\eta}\bigg|_{t=0} t \;=\; \frac{A_o}{\tau_1}\, t. \tag{5-14}$$

The experimental curves presented in Section 5.4.1 have been fitted in their decaying portions to a line in order to determine the parameters τ_2 and A_o as described by Equation 5-13. The saturation grating strength as a function of the exposure energy is presented in Figure 5-12 for different concentrations of DBPH. For the sake of comparison, the value of A_o corresponding to the sample without plasticizer has been normalized by a factor $\times \frac{2}{3}$ to take into account the fact that this sample is 1mm thicker than the others. The fact that there is little spread in the values of A_o across different levels of plasticizer for each exposure energy indicates that the plasticizer has no effect on the energetic sensitivity or on dynamic range of the material.
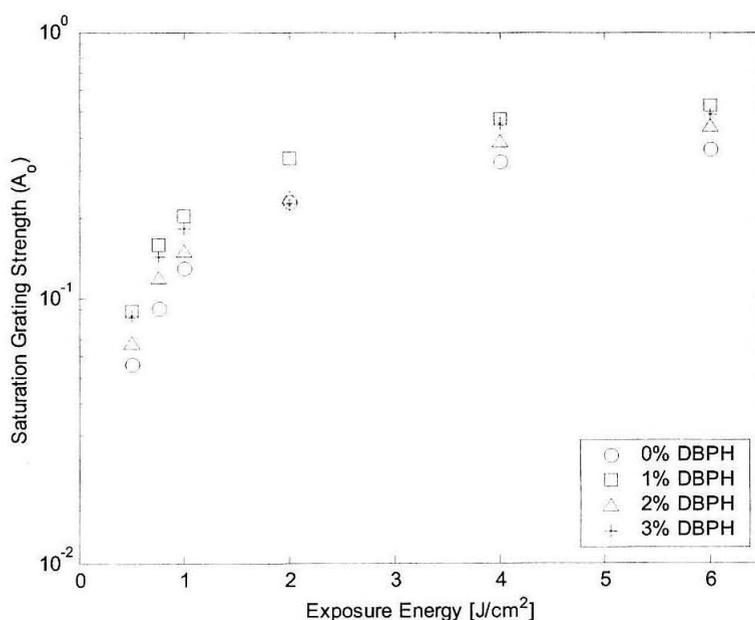
Fig. 5-12.   Saturation value of the grating strength as a function of the exposure energy for samples with concentration of DBPH between 0% and 3%.

The slope at the origin of the curves has also been fitted and τ_1 has been estimated using Equation 5-14 and the previously calculated value for A_o. Table 5-1 summarizes the

values obtained for the growth and decay time constants for different concentrations of

DBPH. The entries in the table correspond to the mean and standard deviation of each one

of the time constants after averaging over the six holograms that were recorded in each one

of the samples. The last column in the table contains the ratio τ_2/τ_1, which serves to verify

that τ_2 is more than two orders of magnitude larger than τ_1 and therefore the assumption

$\tau_2 \gg \tau_1$ is satisfied for all concentrations of plasticizer.

DBPH Concentration [%]	Growth Time Constant τ_1 [hours]		Decay Time Constant τ_2 [hours]		Ratio τ_2/τ_1
	Average	Stand.Dev.	Average	Stand.Dev.	
0	17.40	0.15	4536.26	4059.87	260.67
1	11.35	0.35	1639.08	179.87	144.54
2	3.21	0.18	590.08	22.72	183.85
3	0.75	0.54	261.12	40.78	349.29

Table 5-1. Summary of the estimated growth and decay time constants for different concentrations of DBPH.

As far as the decay time constant is concerned, τ_2 decreases with increasing pres-

ence of plasticizer in the sample. Furthermore as Figure 5-13 suggests, the decrease seems

to follow an exponential trend at least for concentrations up to 3% of DBPH. The linear fit

to the data considers only the values of τ_2 for 1%, 2%, and 3% DBPH because the data point

for 0% DBPH is not significant.

Similarly, the growth time constant decreases with increasing concentration of plas-

ticizer. The values for τ_1 have been plotted in Figure 5-14 and their average values have

been fitted to a line. The data point corresponding to the sample with 3% concentration of

DBPH has been excluded from the fit because of its large estimation error. The quality of

the fit in Figure 5-14 suggests also an exponential dependence with the concentration of

plasticizer in the material. However, the fact that the fit is not as well behaved as the one
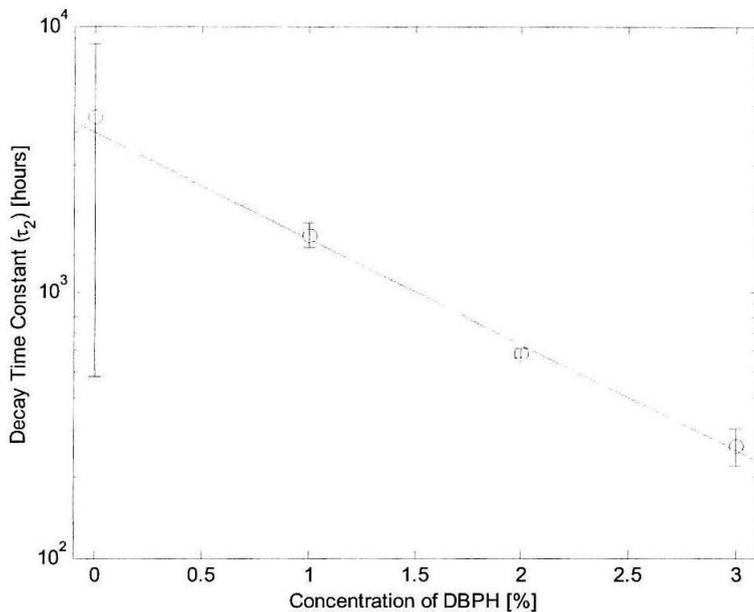
Fig. 5-13.   Hologram decay time constant as a function of the concentration of DBPH in the sample (dots) and monoexponential fit to the experimental data (solid line).

obtained for the decay time constant (Figure 5-13), implies that the dependence of τ_1 on the concentration of DBPH is more complex than a monoexponential relation. This is supported by the fact that the line in Figure 5-14 should be below the data point τ_1 for 3% DBPH since this time constant has been underestimated. Notice in this case that the monitoring of the evolution of the hologram after exposure, Figure 5-9(b), was impaired by the rapid growth of the hologram with baking.

5.4.3 Conclusions

The use of plasticizer in PQ-PMMA samples helps to reduce the revealing time of the holograms, but at the expense of compromising the stability of the material due to the diffusion of the polymer chains and, with it, the lifetime of the holograms. Furthermore, there does not seem to be much room for optimization of the material because the rate at
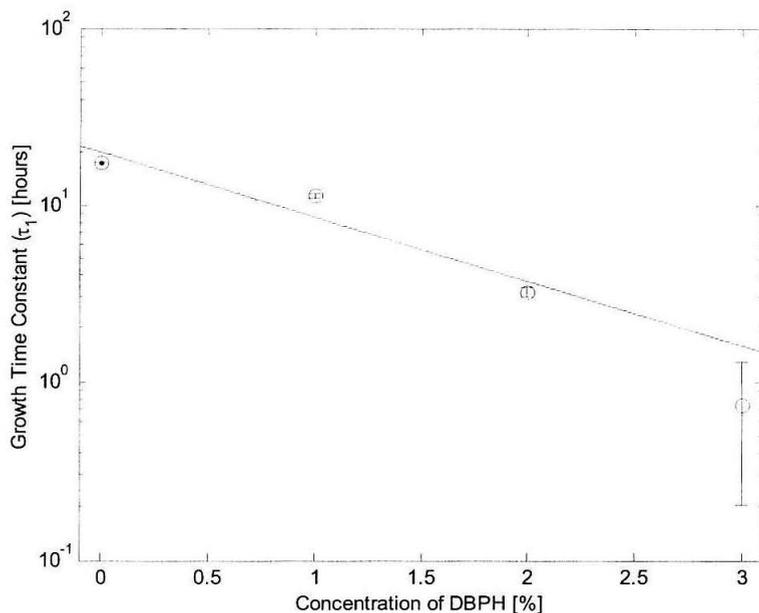
Fig. 5-14. Hologram growth time constant as a function of the concentration of DBPH in the sample (dots) and monoexponential fit to the experimental data (solid line).

which the growth and decay time constants decrease with increasing concentration of DBPH (linear fits in Figure 5-13 and Figure 5-14) is approximately the same.

As a fringe benefit of these experiments, the diffusion constant at 55°C of PMMA can be estimated using the linear fit in Figure 5-13. The value $\tau_2 \approx 4000$ hours at 0% DBPH can be extrapolated, which according to Equation 5-9 on page 10 and taking $\Lambda = 0.5$ μm, translates into a diffusion constant of 4.40×10^{-22} m^2/sec at 55°C for the PMMA macro-molecules.

Finally, for highly plasticized samples, the diffusion of free PQ molecules is greatly enhanced even at room temperature. This phenomenon, noticeable in the case of the sample with 3% DBPH in which the amplification factor is smaller due to higher prebaking diffraction efficiencies, becomes dramatic for samples with 5% DBPH. Figure 5-15 shows that the

prebaking grating strengths of holograms measured after recording in a 2 mm thick sample with 5% DBPH are equivalent to the after-baking strengths in a nonplasticized sample. Unfortunately, the lifetime of these holograms is just a few days at room temperature.
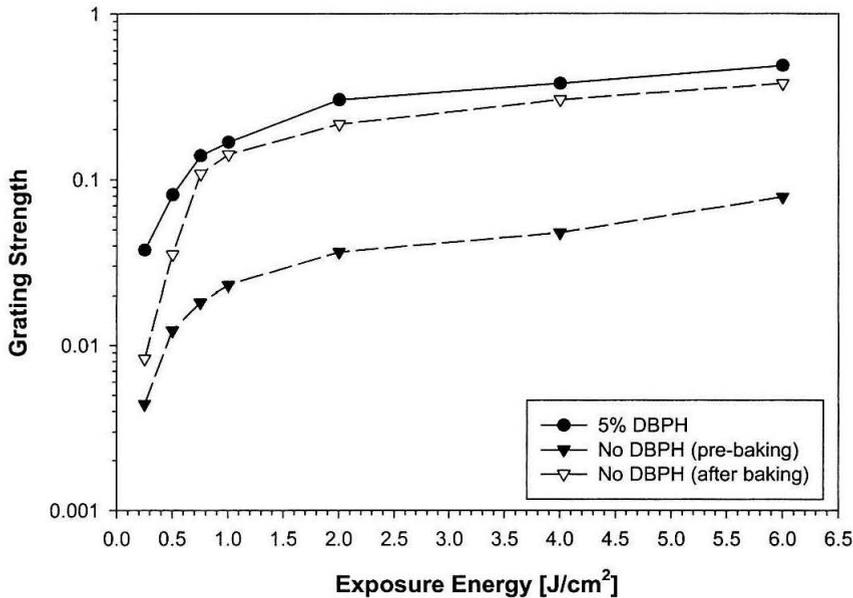
Fig. 5-15. Grating strength as a function of the exposure energy for a sample with no DBPH, measured before (solid triangles) and after (clear triangles) baking, compared with a sample with 5% DBPH (solid circles) without any baking.

5.5 New PQ molecules

After the experimental study of the effect of the plasticizer in the PQ-PMMA material done in the previous section, the main result is that the plasticizer has the positive effect of facilitating the diffusion of PQ molecules, but at the same time the diffusion of PMMA macromolecules. In other words, by modifying the global composition of the medium, it is not possible to uncouple one diffusional process from the other.

A different approach to the optimization of the material focuses on the optimization of the PQ molecules themselves. If the structure of the PQ molecules could be modified, or substituted with some other molecules from the same family, in a way in which their diffusion coefficient in a given environment is increased, this would impact only in the revealing time of the hologram without compromising its lifetime. In this section, the possibility of replacing PQ molecules with other PQ-based compounds is investigated.

5.5.1 Modified PQ molecules

The following three variations on the basic PQ molecule (9,10-phenanthrenequinone) have been taken into consideration: 1-isopropyl-7-methyl-9,10-phenanthrenequinone (referred to as IPMPQ), 2,4-dimethyl-6-brom-9,10-phenanthrenequinone (or DMBPQ) and 2,4,5,7-tetramethyl-9,10-phenanthrenequinone (or TMPQ).
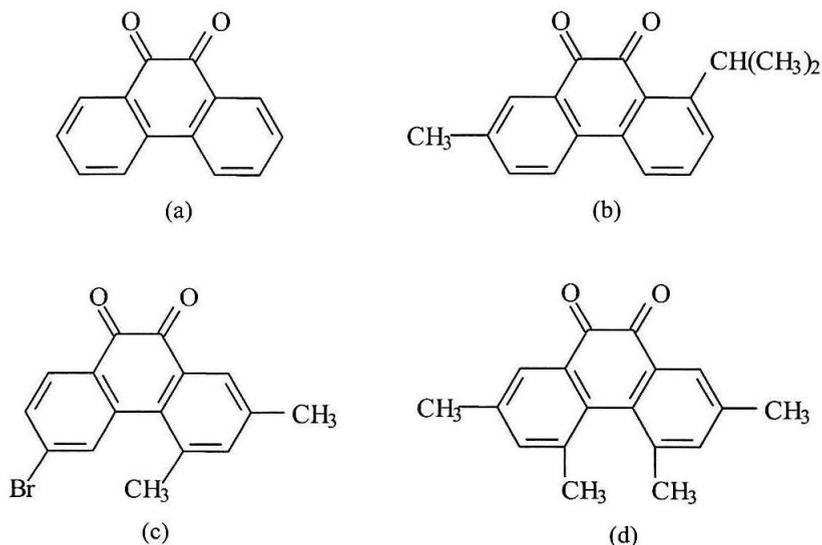
Fig. 5-16. (a) 9,10-phenanthrenequinone, (b) 1-isopropyl-7-methyl-9,10-phenanthrenequinone, (c) 2,4-dimethyl-6-brom-9,10-phenanthrenequinone, and (d) 2,4,5,7-tetramethyl-9,10-phenanthrenequinone.

Different 1 mm thick PMMA samples doped with 0.8% IPMPQ, 0.5% DMBPQ, and 0.5% TMPQ were fabricated in order to characterize the holographic properties of each one of the new molecules. For reference, their performance is compared to a 0.5% PQ sample with the same thickness.

The absorption spectrum of the IPMPQ-PMMA sample, showed in Figure 5-17, follows closely the spectrum of the PQ-PMMA sample, which is in fact the reason why the concentration of IPMPQ was chosen to be 0.8% instead of 0.5% as for the other samples.
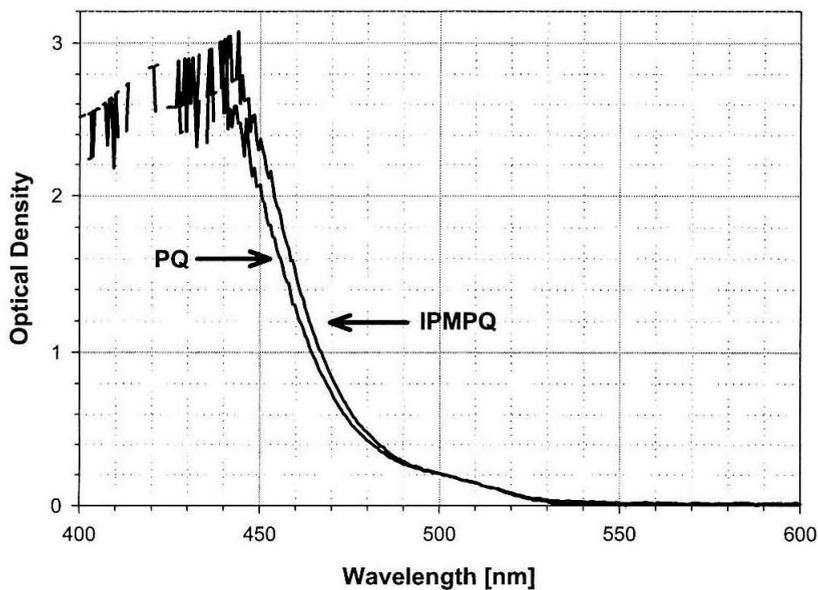
Fig. 5-17.   Absorption spectra of 1 mm thick unexposed PMMA samples doped with either 0.5% PQ or 0.8% IPMPQ.

The absorption spectra of the DMBPQ-PMMA and the TMPQ-PMMA samples are presented in Figure 5-18. Compared to regular PQ-PMMA, the optical density of DMBPQ-PMMA is slightly higher at both 488 nm and 514 nm wavelengths. The TMPQ-PMMA has an optical density 2.6 times bigger than PQ-PMMA at 514 nm and almost 4 times larger at

488 nm. Like in PQ-PMMA, the absorption at 532 nm is also very low for the three new materials.
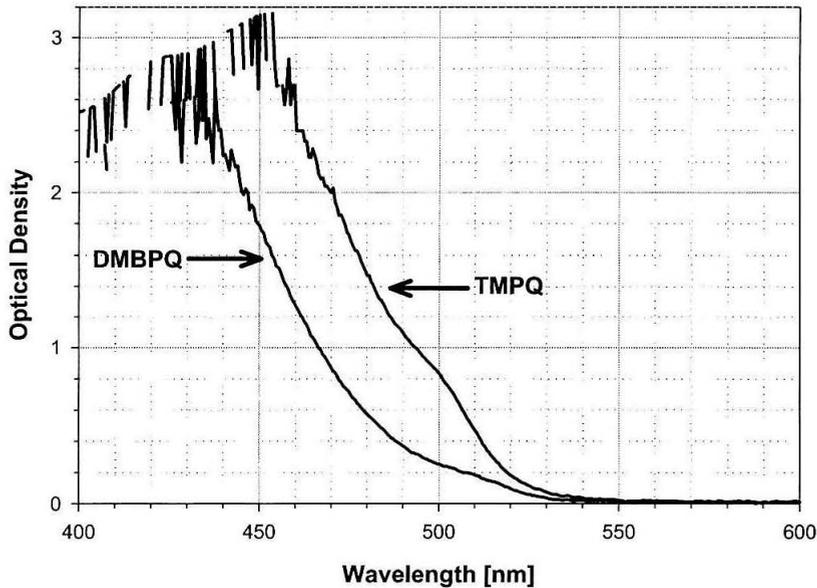
Fig. 5-18.   Absorption spectra of 1 mm thick unexposed PMMA samples doped with 0.5% of DMBPQ or TMPQ.

5.5.2 Experimental comparison

Experiments to characterized the performance of these different PQ-based PMMA materials at 514 nm have been performed using the setup described in Section 5.2.2. To study the recording dynamics, the growth of a hologram is monitored in real time as the exposure energy of 20 J/cm^2 is delivered to the sample at room temperature. The recording curves for the four materials are compared in Figure 5-19. Although the IPMPQ-PMMA sample grows faster and saturates at a higher level of diffraction efficiency than the PQ-PMMA sample, they both exhibit a very similar behavior. The saturation value of diffraction efficiency is around 1.6% and it requires about 8 J/cm^2 to reach the 90% of the saturation value.
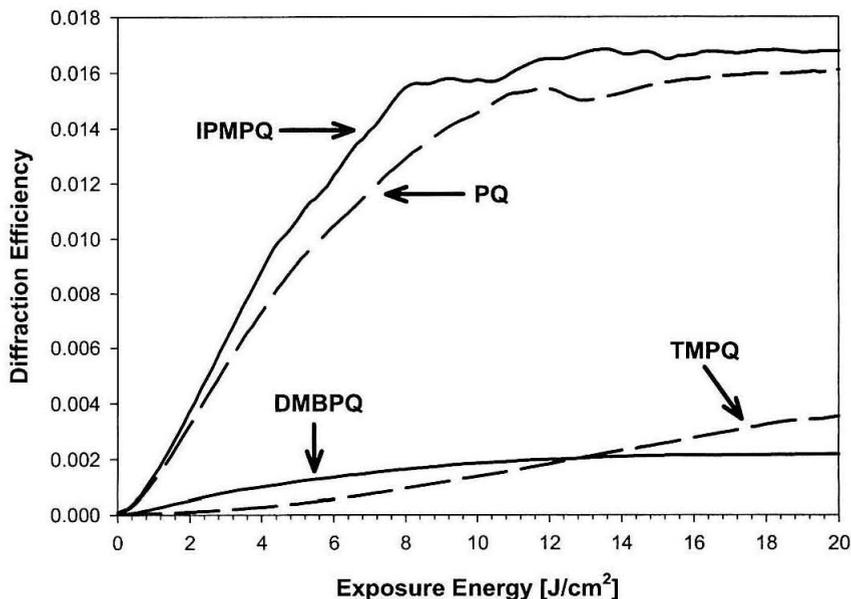
Fig. 5-19.   Comparison of the hologram recording dynamics for the PQ-, IPMPQ-, DMBPQ-, and TMPQ-PMMA samples.

The dynamic range of the DMBPQ-PMMA material seems to be much lower than in the other two previous samples, since the diffraction efficiency saturates below 0.2%. However, the recording speed is still comparable to IPMPQ-PMMA and PQ-PMMA because it takes 9.2 J/cm^2 to reach the 90% of the saturation diffraction efficiency. Finally, the TMPQ-PMMA sample exhibits very inefficient recording, as the hologram did not reach saturation after an exposure of 20 J/cm^2. The slow recording can be mostly attributed to the high absorption of the sample at 514 nm. By fitting the experimental data to a saturating mono-exponential curve, the dynamic range of the TMPQ-PMMA sample is estimated to be 0.8% and that it requires 43.4 J/cm^2 to reach the 90% level.

The post-exposure behavior of the samples has also been investigated. A hologram has been recorded with a 10 J/cm^2 exposure in each one of the samples. Right after expo-

sure, the evolution of the diffraction efficiency of the holograms has been monitored with

the He-Ne laser for 10 minutes as they rest in the dark at room temperature.
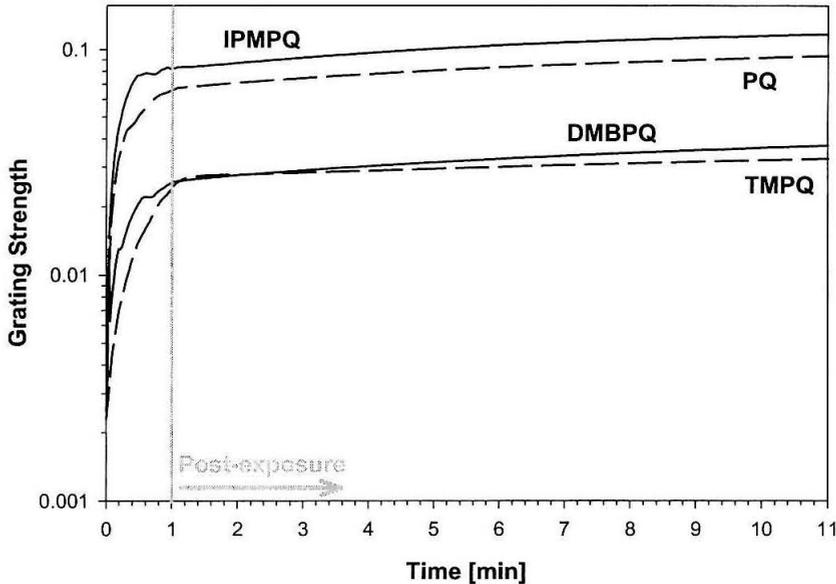
Fig. 5-20. Comparison of the post-exposure dark dynamics at room temperature for the PQ-, IPMPQ-, DMBPQ-, and TMPQ-PMMA. The vertical solid line indicates the end of the recording.

The experimental results are presented in Figure 5-20, in which the vertical line

separates the exposure (left-hand side) and the post-exposure regimes (right-hand side). It

comes as no surprise, after the results presented in Figure 5-19, that the strength of the holo-

grams at the end of the recording time is higher in IPMPQ-PMMA and PQ-PMMA than it

is in DMBPQ-PMMA and TMPQ-PMMA. During the post-exposure, the holograms con-

tinue growing in the dark: the grating strength of the holograms amplified by a factor of

1.44 in the first 10 minutes after exposure for the IPMPQ, PQ and DMBPQ materials, and

1.35 for TMPQ-PMMA. The fact that the enhancement of the hologram after recording is

fairly similar in all the samples is fully consistent with the explanation provided in

Section 5.3.2, according to which the enhancement during post-exposure is due to the

attachment of photoactivated PQ-based radicals to PMMA. The oxygen atoms in the PQ molecule are responsible for the attachment to the PMMA chains. Therefore, the same dynamics is expected in all the PQ-based materials.

Finally, the evolution of the hologram with baking of the different PQ-based materials is studied. Six holograms with energetic exposures ranging between 2 and 20 J/cm^2 have been recorded in the PQ-PMMA, IPMPQ-PMMA, and DMBPQ-PMMA samples. The range chosen for the TMPQ-PMMA was between 4 and 45 J/cm^2 due to the lower sensitivity of the material. The diffraction efficiency of the holograms has been monitored as a function of the baking time. The evolution of the gratings in IPMPQ-PMMA is shown in Figure 5-21. The results are once again very similar to the PQ-PMMA case that was already
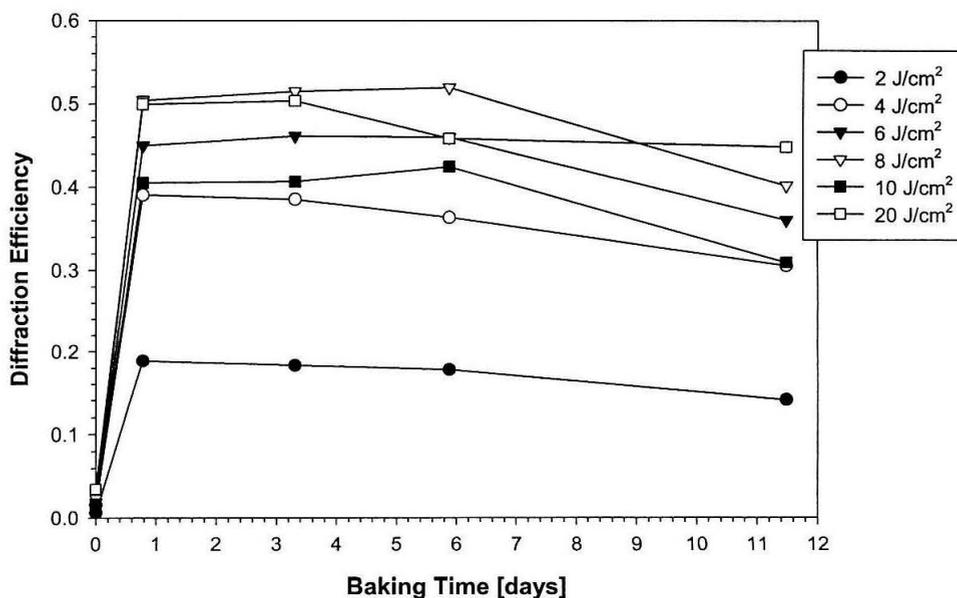
Fig. 5-21.   Effect of baking at 55°C on the grating strength of the holograms as a function of time in an IPMPQ-PMMA sample.

presented in Figure 5-5(a). The holograms have been fully revealed after 18 hours in the oven and their diffraction efficiency amplified by a factor of 22.8 on average.

The baking dynamics in the cases of DMBPQ-PMMA, Figure 5-22(a), and TMPQ-PMMA, Figure 5-22(b), is distinct from the two previous materials in the sense that the diffusion of the PQ-based molecules seems to be much slower than in the original PQ-PMMA material. The revealing of the hologram in the DMBPQ-PMMA material, due to the erasure of the grating formed by free DMBPQ molecules, was not completed even after baking the sample for 12 days. The TMPQ-PMMA material presented even a poorer performance, because the holograms hardly amplified with baking.

5.5.3 Conclusions

The main results of the experimental comparison of the PQ-based PMMA materials are summarized in Table 5-2. Although the use of new PQ-based molecules has led to some improvements in the material as far as the recording process is concerned, none of the new molecules seems to have a big impact in reducing the baking time required for complete revealing of the holograms. In that sense, the DMBPQ and TMPQ molecules have exactly the opposite effect as they seem to diffuse with much higher difficulty in the polymeric matrix, slowing down the growth of the hologram during baking The IPMPQ-PMMA material has higher dynamic range and recording sensitivity than the regular PQ-PMMA; however, these improvements do not seem to be more relevant than to make IPMPQ-PMMA material a perfect substitute of PQ-PMMA.
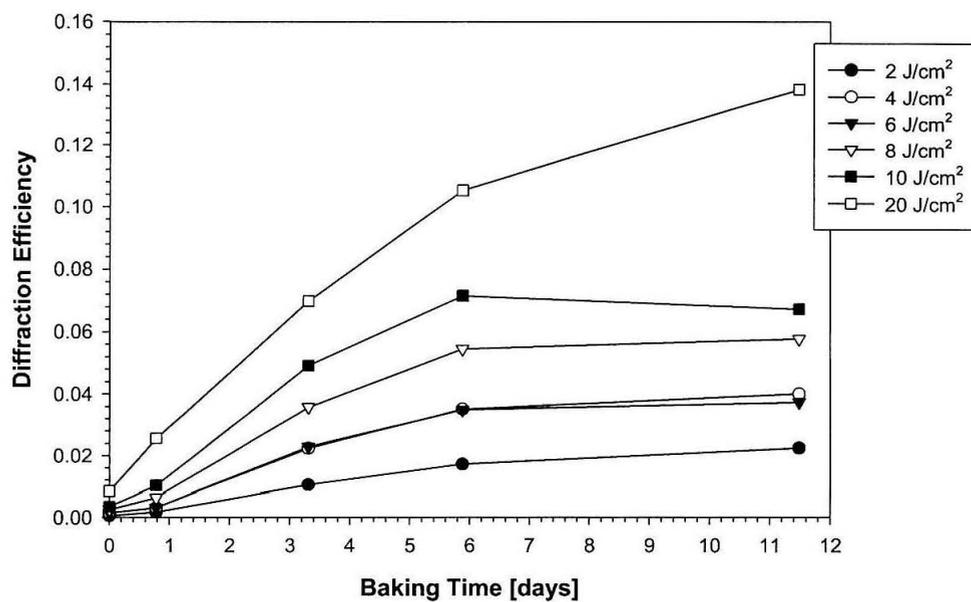
(a)
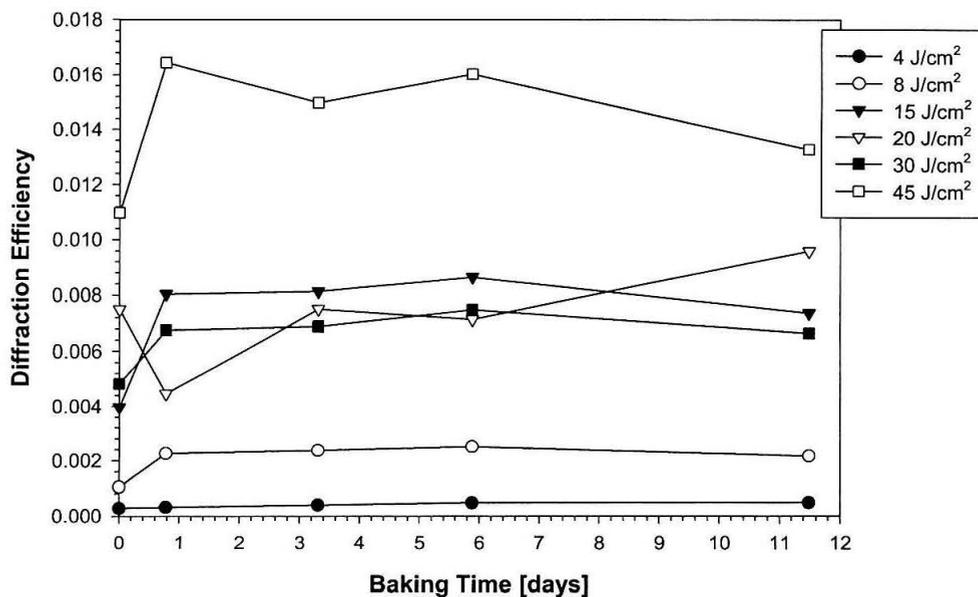
(b)

Fig. 5-22.   Effect of baking at 55°C on the grating strength of the holograms as a function of time in (a) a DMBPQ-PMMA sample, and (b) a TMPQ-PMMA sample.

Molecule	RECORDING		POST-EXPOSURE	BAKING
	Dynamic Range (prebaking)	Saturation Energy (90% level) [J/cm^2]	Amplification (first 10 min)	Amplification[†]
PQ	1.49%	8.49	2.02	22.31
IPMPQ	1.63%	7.35	2.05	22.61
DMBPQ	0.17%	9.19	2.13	27.83[*]
TMPQ	0.80%	43.44	1.82	2.16[*]

Table 5-2. Comparison of the recording, post-exposure and baking parameters for the four PQ-based molecules. [†]Amplification for diffraction efficiency after 12 days bake. [*]Measured after 44 days bake instead of 12 days.

A different approach to optimize the material by replacing the polymeric matrix has been proposed and tested in [5-11]. PQ-doped poly(bisphenol-A-carbonate) (PC) is used instead of PQ-PMMA. It seems to be that PQ molecules can diffuse in PC more easily than they can in PMMA, resulting therefore in shorter revealing time of the holograms. One might think that PC could behave in a similar manner to plasticized PMMA, lowering the glass-transition temperature of the material. However, PC has in fact a higher glass-transition temperature than PMMA (T_g~150°C versus T_g~100°C), which means that at any given temperature the diffusion of PC macromolecules is smaller than it is for the PMMA chains, increasing the lifetime of the holograms. Therefore, PQ-PC overcomes the limitations of PQ-PMMA as it enjoys both faster hologram growth and longer lifetime.

5.6 Strong gratings for WDM applications

Bragg gratings in fibers are widely utilized to implement drop-filters and demultiplexors for WDM applications. One of the techniques used to create the grating in the fiber is holography, in which the interference pattern of two UV-light beams records a sinusoidal modulation of the index of refraction of the silica fiber. The possibility of implementing

free-space optical components for optical fiber networks is particularly interesting. Since more degrees of freedom exist in the design of free-space components, higher flexibility can be added to these devices, like tunability of the filters for example. This explains why holographic media, like photorefractive materials and LiNbO$_3$ in particular, have been used as support for Bragg gratings in WDM applications [5-12],[5-13]. The low sensitivity of LiNbO$_3$ is seen as a drawback because it makes a long recording process in a stabilized setup necessary. Furthermore, the volatility of the gratings stored in the crystal requires the use of hologram fixing techniques.

Thick polymers are a good alternative to photorefractives because they possess high recording sensitivity and do not suffer from erasure. In addition to this, they are inexpensive and more durable than LiNbO$_3$. This section will investigate the suitability of the PQ-PMMA material to implement narrow filters for WDM applications [5-7].
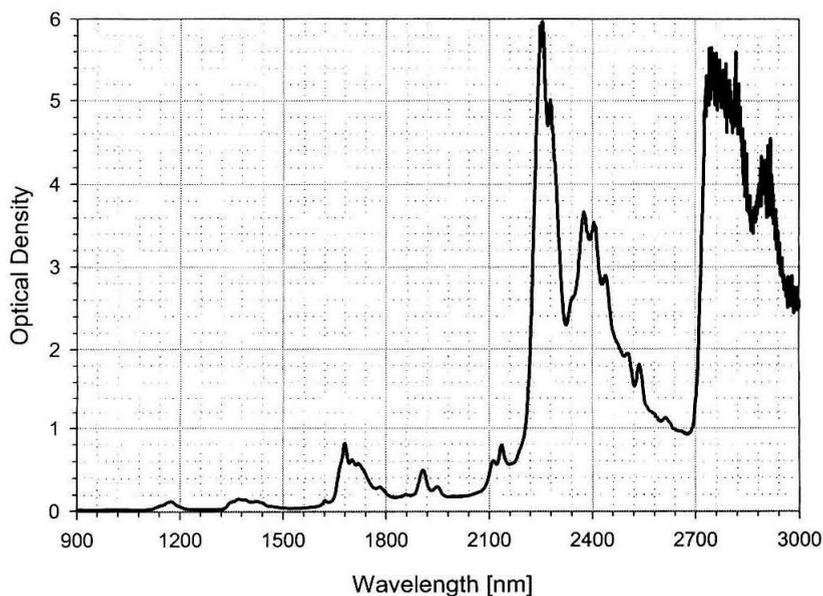
Fig. 5-23.   Absorption spectrum in the infrared band for a 1 mm thick PQ-PMMA sample.

The absorption spectrum of a PQ-PMMA sample in the infrared has been measured

using a Varian Cary 500 spectrometer. From Figure 5-23, the optical density of the material

at 1550 nm is 0.03 for a 1 mm thick disk. The fact that the material is fairly transparent in

the band around 1550 nm is beneficial for this application, because it leads to little insertion

losses, however it also means that holographic recording at these wavelengths is poor. The

formalism of the K-sphere (Figure 5-24) is useful to see that the grating can be recorded in

transmission geometry in the visible, where the absorption spectrum of the material has

been optimized, and then readout in the infrared in reflection geometry.
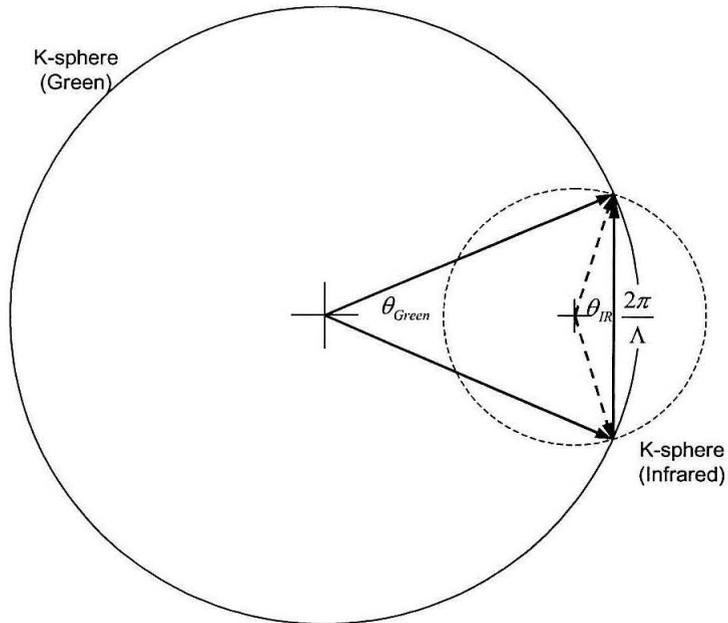
Fig. 5-24.   K-sphere diagram to illustrate the Bragg-matching in the infrared wavelengths of a grating recorded in the visible.

According to Bragg theory, a grating of spatial period Λ can be matched at different

wavelengths by appropriately choosing the angle of incidence of the light beam. A holo-

gram recorded in the green (λ_{Green}) with an angle of incidence for each one of the recording

beams θ_{Green}, can be read out in the infrared (λ_{IR}) if the incidence angle θ_{IR} satisfies the condition[5-14]

$$\frac{\sin(\theta_{Green})}{\lambda_{Green}} = \frac{2\pi}{\Lambda} = \frac{\sin(\theta_{IR})}{\lambda_{IR}}. \tag{5-15}$$

In the particular context of the filters for WDM, we are interested in a perfect reflection grating (i.e., $\theta_{IR}=90^{\circ}$).

5.6.1 Experimental setup and results

Three millimeter thick PQ-PMMA disks have been used in the experiments. Squares of about 2.4×2.4 cm in cross section have been obtained by cutting off the edges of the disks. The flat facets have then been polished to allow the propagation of a laser beam
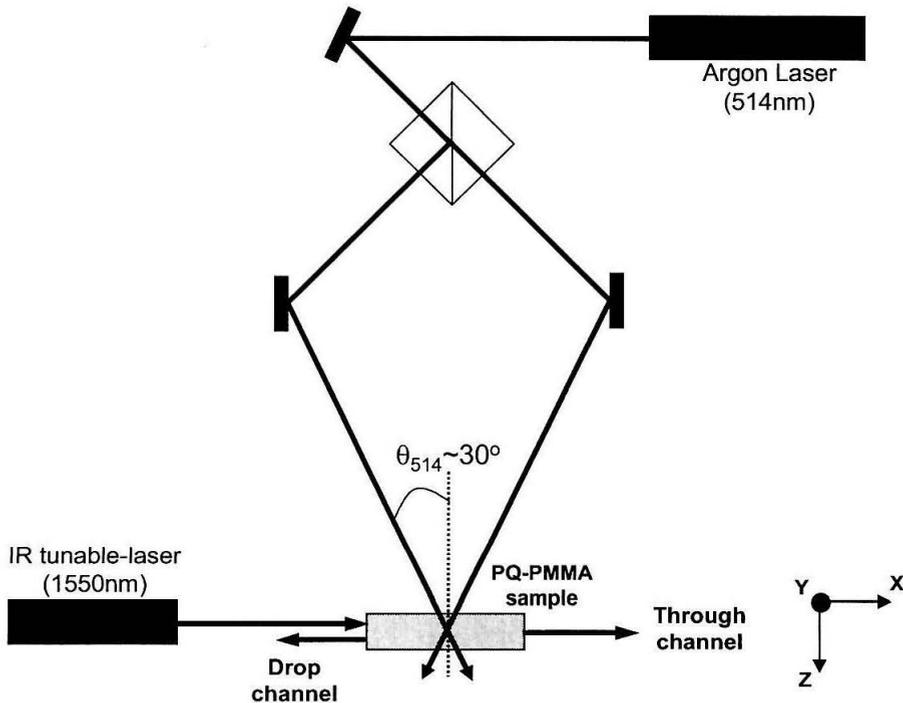
Fig. 5-25. Experimental setup for recording in transmission-geometry using the 514 nm line of the Argon laser, and readout in reflection-geometry with a tunable infrared laser.

between opposite edges without experiencing significant scattering. As suggested in the sketch in Figure 5-25, the grating is recorded from the "top" of the sample (i.e., the 2.4 ×2.4 cm facet). The green light beams from an Argon laser are incident on the sample at approximately 30° (outside the medium) in a symmetric transmission geometry. The recording beams have been sufficiently expanded to cover the entire cross section of the sample, thus maximizing the extension of the grating to the entire width of the sample during readout. The hologram is read out using a Photonetics OSICS external cavity laser (ECL), with a tuning range from 1520 nm to 1600 nm. The hologram will be Bragg-matched in reflection, therefore the infrared beam is normally incident to one of two edges of the sample that is perpendicular to the grating vector.
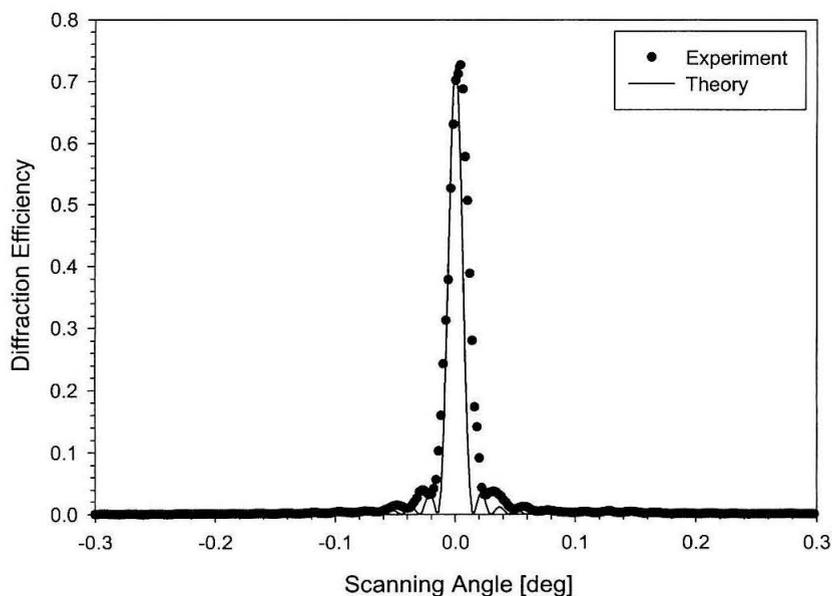
Fig. 5-26. Experimental angular selectivity curve of the hologram measured at 514 nm (dots) and theoretical prediction (solid line).

Figure 5-26 shows the experimental result of a hologram recorded with an exposure energy of 5 J/cm^2. After baking the sample at 55°C in the oven for 42 hours, the grating has

a diffraction efficiency of 72.7% when measured in transmission with green light. The measured selectivity curve (dots) shows a good agreement with the theoretical prediction (solid line).

The wavelength selectivity of the hologram in the 1550 nm band is analyzed in Figure 5-27. The amount of transmitted power, usually referred to as the through channel by the terminology for optical fiber components, is monitored as the wavelength of the infrared laser is swept across its tuning range. The hologram suppresses the wavelength for which the Bragg condition is satisfied, producing a notch in Figure 5-27. The minimum in
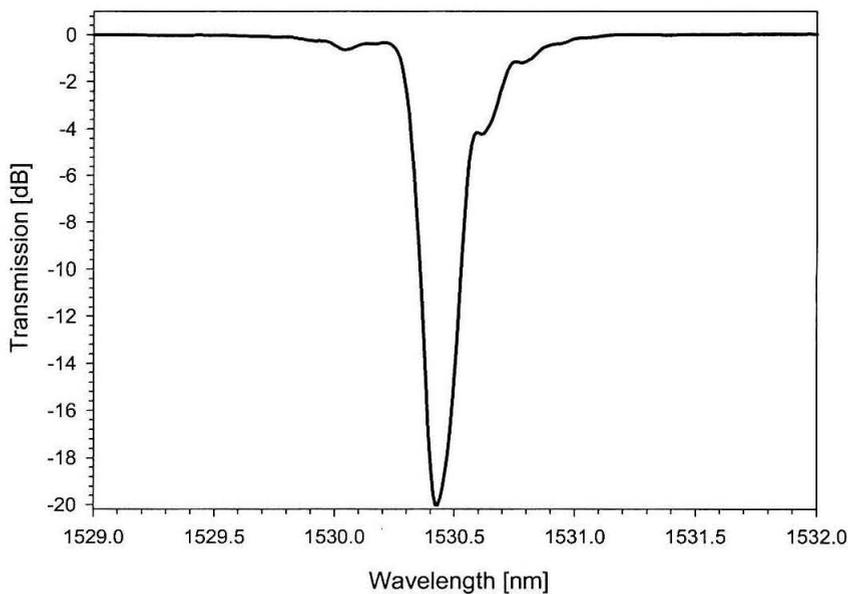
Fig. 5-27. Wavelength selectivity of the hologram measured in the infrared. The vertical axis represents the transmitted power in the through channel.

transmission efficiency occurs at 1530.43 nm, for which 99.01% of the input channel is rejected, or equivalently the rejection is 20.03 dB. Using Kogelnik formula,[5-15] the measured diffraction efficiency of the reflection hologram corresponds to a modulation of the

index of refraction $\Delta n = 0.64 \times 10^{-4}$, which is consistent with the value $\Delta n = 0.53 \times 10^{-4}$ derived from the measurement of the grating in transmission geometry at 514 nm.

In order to better characterize the shape of the wavelength selectivity of the hologram, the bandwidth of the filter is measured for different levels of rejection of the through channel. A commonly used metrics for the width of the stopband of the filter is the distance between the points for which the suppression of the transmitted signal drops 3 dB with respect to the maximum, which in this particular case is 0.09 nm (or equivalently 11.53 GHz). The information about the flat-top behavior of the filter is captured by the 0.5 dB bandwidth of the hologram, which turned out to be 0.04 nm (or 5.12 GHz). Finally, the bandwidth of the filter at the baseline level is 0.42 nm (or 53.80 GHz). This is also an interesting parameter because it determines how closely different holograms can be multiplexed without crosstalk to form a filter bank.

The filter shape in Figure 5-27 presents an asymmetry between the left and right edge of the notch. Although the asymmetry is not severe in the particular experiment shown in the figure compared to measurements in other samples, one can still observe that the edge corresponding to larger wavelengths is less steep and shows some sort of secondary lobes. One reason to justify the observed behavior has to do with the quality of the grating, either because the wavefronts that impinge on the PQ-PMMA sample are not perfectly collimated, or because they become distorted as they propagate inside the material. In either case, the recorded hologram differs from a pure sinusoidal grating, giving rise to asymmetry. Another reason for deviations from the theoretical filter shape, although for shorter wavelengths, is the finite aperture of the infrared beam, about 0.5 mm in diameter. Diffraction effects as the beam travels across the sample result in some of its components diverg-

ing from the original direction of propagation (i.e., the x-axis). These components can be Bragg-matched at shorter wavelengths, as the radius of K-sphere becomes larger.

5.6.2 Recording considerations: Secondary gratings

The recording of strong gratings has revealed the presence of distortion in the hologram when is read out in the green, once the baking process has been completed. This distortion manifests in the form of fringes over the aperture of the diffracted beam, leading to a degradation of the angular selectivity curve of the hologram. Although the distortion was thought initially to be related to stress applied to the sample by the metal mount that holds it during recording, or maybe during the post-processing of the sample in the oven due to the difference between the thermal expansion coefficient of the metal and PMMA, it turned out in fact to be related to the build-up of a secondary grating during exposure. The presence of this secondary grating leads to interference with the original hologram in a symmetric recording setup, but can be avoided by using a nonsymmetric geometry and recording a slanted grating instead, as can be observed in Figure 5-28.

The sketches in Figure 5-29 help to visualize this effect and understand the difference between the symmetric and slanted cases. During exposure, the interference between two plane waves propagating at angles θ_R and θ_S with respect to the z-axis (i.e., the surface normal of the material), and with wave vectors $\mathbf{K_R}$ and $\mathbf{K_S}$ defined as

$$\mathbf{K_R} = K \cdot [\hat{z} \cdot \cos\theta_R - \hat{x} \cdot \sin\theta_R]$$
$$\mathbf{K_S} = K \cdot [\hat{z} \cdot \cos\theta_S + \hat{x} \cdot \sin\theta_S]$$

(5-16)

results in a hologram with the following grating vector:

$$\mathbf{K_G} = \mathbf{K_S} - \mathbf{K_R} = K \cdot [\hat{z} \cdot (\cos\theta_S - \cos\theta_R) + \hat{x} \cdot (\sin\theta_S + \sin\theta_R)].$$

(5-17)
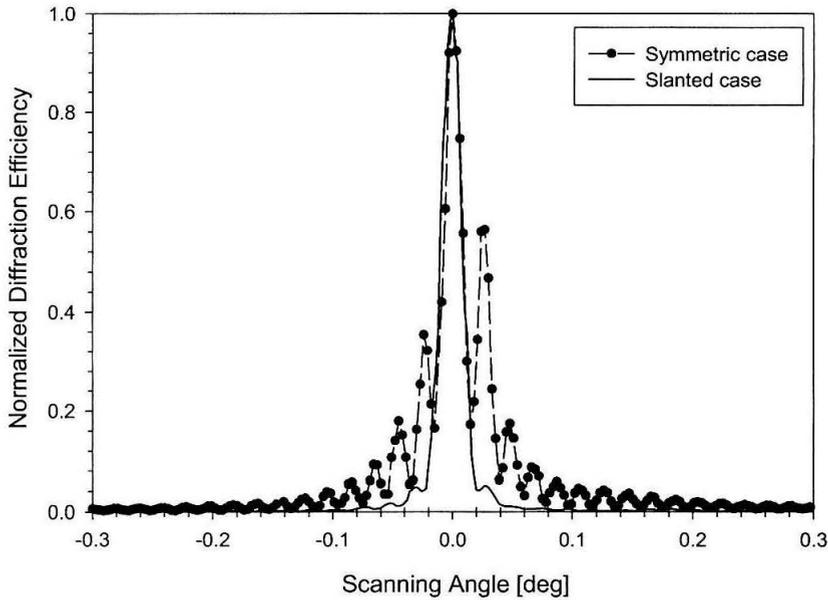
Fig. 5-28.   Angular selectivity curve for a transmission hologram recorded in a perfect symmetric geometry (dashed line and dots) compared to the curve for a 10°-slanted grating (solid line).

As each one of the recording beams impinges the rear surface of the sample, part of the light

is reflected back into the material, producing the components $\mathbf{K'_S}$ and $\mathbf{K'_R}$

$$\mathbf{K'_R} = -K \cdot [\hat{z} \cdot \cos\theta_R + \hat{x} \cdot \sin\theta_R]$$
$$\mathbf{K'_S} = -K \cdot [\hat{z} \cdot \cos\theta_S - \hat{x} \cdot \sin\theta_S]$$

(5-18)

which will record a secondary grating $(\mathbf{K'_G})$ given by the expression

$$\mathbf{K'_G} = K \cdot [\hat{z} \cdot (\cos\theta_R - \cos\theta_S) + \hat{x} \cdot (\sin\theta_S + \sin\theta_R)].$$

(5-19)

In the symmetric case ($\theta_R = \theta_S = \theta$) shown in Figure 5-29(a), the $\hat{z}$-component of the

two gratings vanishes and they become colinear. In practice, the fact that the front and rear

surfaces of the sample are not perfectly parallel to each other introduces a slight tilt between

$\mathbf{K_G}$ and $\mathbf{K'_G}$. Therefore, when probing the hologram with a green light beam, both gratings

are Bragg-matched and two diffracted spots that propagate almost colinearly but slightly
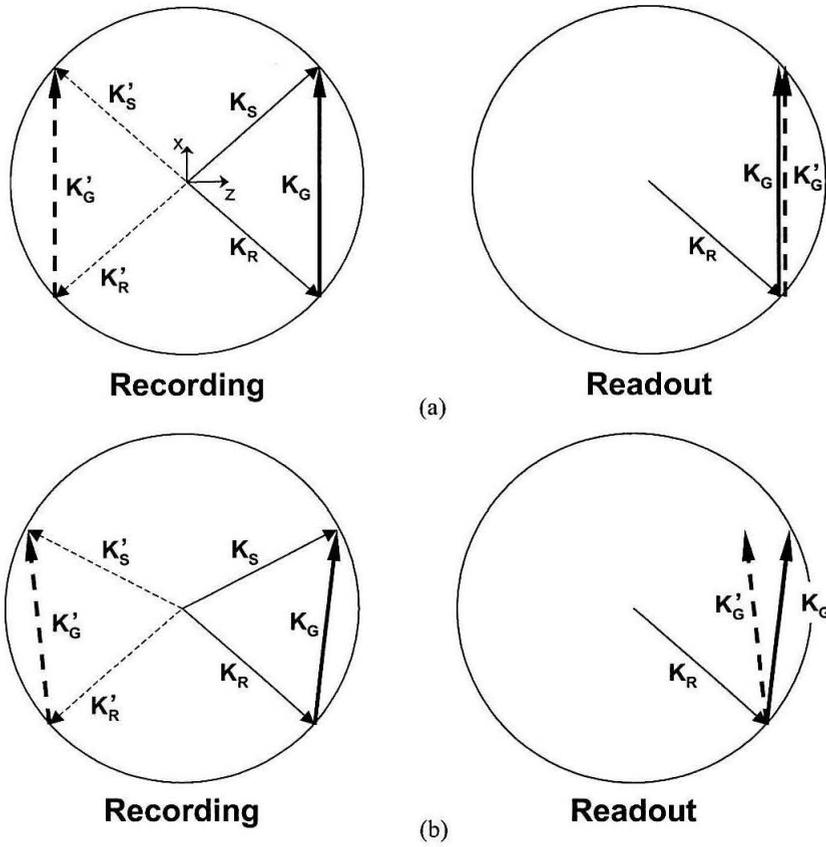
Fig. 5-29. K-sphere diagrams to illustrate the origin of the secondary grating. (a) In the symmetric case, the back reflection of the recording beams records a noise grating that will be Bragg-matched during readout; (b) for slanted recording, the noise grating due to back reflections has a different orientation and will not be Bragg-matched during readout.

shifted with respect to each other can be observed. The interference pattern between the two beams is presented in Figure 5-30. As the angle of incidence of the probing beam is changed, the interference pattern changes proving that the two spots exhibit Bragg-selectivity. For the image in Figure 5-30, the interference is destructive over most of the region of overlap and only the far edges of the two beams are bright.

This problem can be eliminated by using a nonsymmetric geometry. As shown in Figure 5-29(b), if the angles of the recording beams are chosen to be $\theta_R = \theta + \frac{\Delta\varphi}{2}$ and
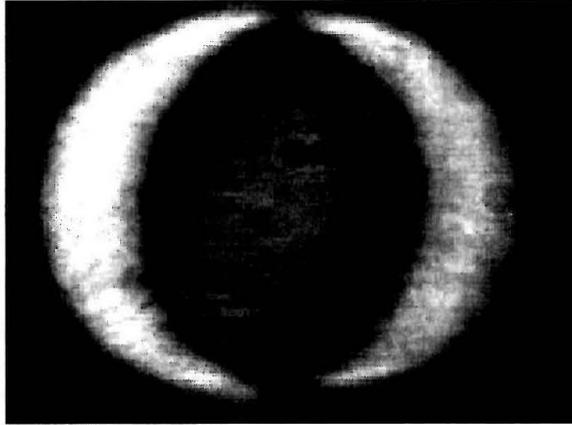
Fig. 5-30.   Profile of the diffracted spot showing an O-shaped pattern as a result of the destructive interference between the beams diffracted by each one of the two holograms.

$\theta_S = \theta - \dfrac{\Delta\varphi}{2}$, then the tilt between the two gratings is $\Delta\varphi$, which can be made larger than the angular selectivity of the material so only one grating ($\mathbf{K_G}$) is Bragg-matched. Since the interference with the spurious grating is avoided, the holograms exhibit good quality angular selectivity curves, as it can be observed in Figure 5-28.

5.6.3 Readout considerations: Depth scan

During readout in the infrared, the laser beam is coupled from the edge of the sample. Since the diameter of the infrared beam, about half millimeter, is smaller than the thickness of the sample, typically 3 mm, the wavelength selectivity of the hologram can be analyzed as the point of incidence of the laser beam is gradually scanned from the top (i.e., the front surface of the PQ-PMMA sample during recording) to the bottom of the sample along the direction of the z-axis in Figure 5-25. Figure 5-31 shows in a 3-D plot how the wavelength selectivity of a hologram varies from the top to the bottom surface in steps of 200 μm. The first depth slice, labeled as 0 μm, corresponds to the minimum depth at which the entire diameter of the laser beam is guided inside the sample. The next depth slices are

measured relatively to this point, being 2600 µm the deepest before the beam is clipped by

the optical mount on which the sample rests.
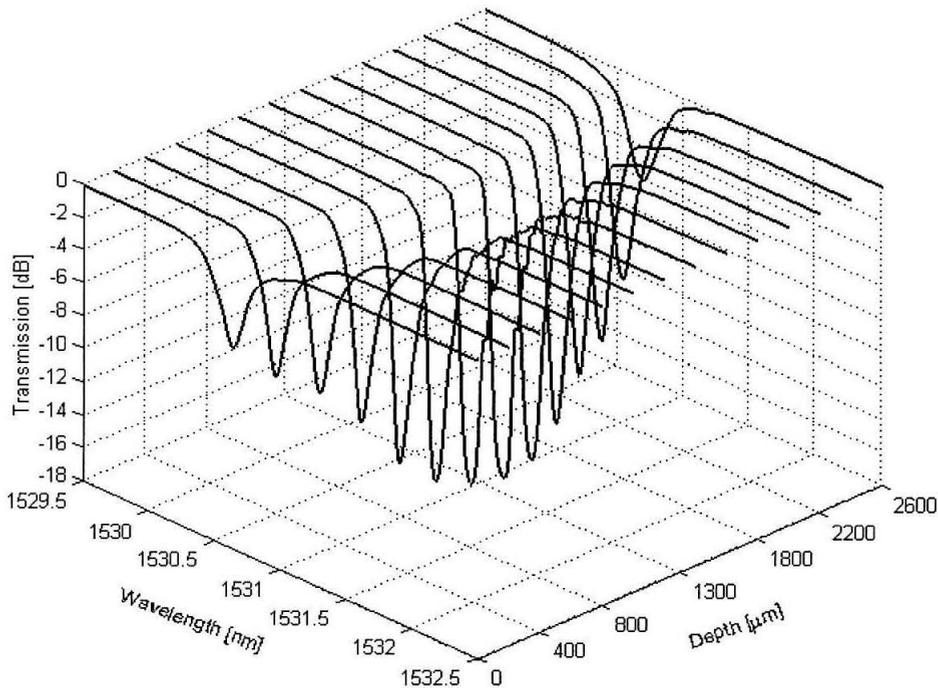
Fig. 5-31.   Suppression of the transmitted channel as a function of the wavelength of the laser measured at different depths in 200 µm steps across the 3 mm thickness of the sample.

From the experimental data, the following two trends can be noticed: First, the dif-

fraction efficiency of the filter seems to be stronger, and also the hologram exhibits a better

flat-top behavior, when the beam propagates through the sample at a depth that corresponds

to the middle, around 1300 µm in the experiment in Figure 5-31. As the wavelength selec-

tivity measurement is performed closer to either the top or bottom surfaces, the hologram

becomes weaker.

Second, the wavelength at which the transmitted channel is maximally suppressed

shifts from shorter wavelengths near the top of the sample to longer wavelengths around

the center, and then from longer to shorter again as the wavelength selectivity measurement is taken near the bottom of the sample. The swing of the Bragg wavelength can be better observed in Figure 5-32, in which this wavelength is plotted as a function of the depth.
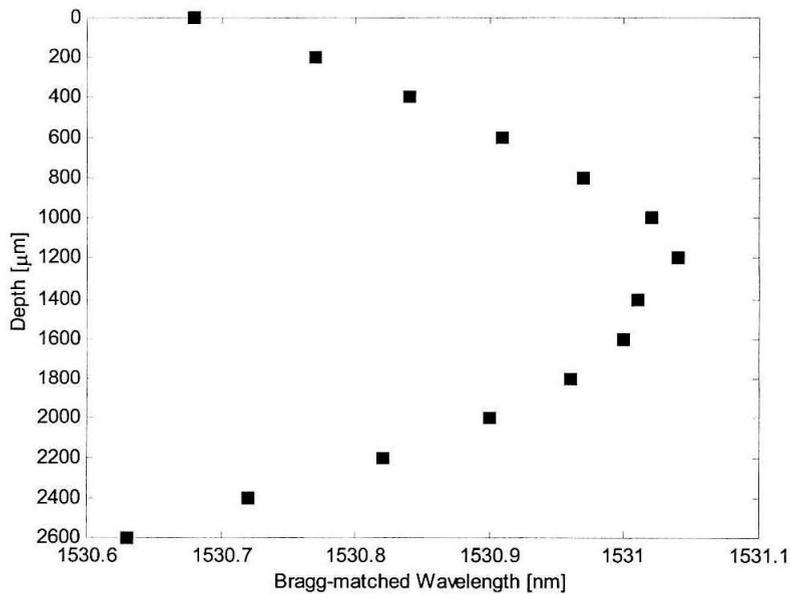
Fig. 5-32. Dependence of the Bragg wavelength of the filter with depth across the thickness of the PQ-PMMA sample.

Both effects have their origin in the process of cutting the PQ-PMMA disks into squares, and then polishing the edges of the sample. Despite careful processing, the edge of the sample, rather than being a flat surface, still presents some curvature. The curvature of the edge is more pronounced near the top and bottom surfaces of the sample, and less important in its central portion. Near the surfaces, the curvature of the interface changes the direction of propagation of the normally incident infrared beam in the x-z plane, according to the coordinate system of Figure 5-25. Since the beam travels at an angle with respect to the x-axis, a shorter wavelength (i.e., longer K-vector) is required to Bragg-match the holo-

gram. Therefore, the Bragg wavelength is longer in the central part of the edge, where there is almost no deflection, and shorter on the top and bottom of the sample, where the curvature is more prominent.

The variation of the diffraction efficiency of the hologram with depth is the result of the finite aperture of the readout beam combined with the curvature of the edge of the sample. When the point of incidence of the beam is near the center portion of the edge, almost all the rays across the aperture of the beam travel parallel, thus all they are Bragg-matched at the same wavelength and the hologram appears to be stronger. As the incident beam is shifted to either the top or the bottom, and the curvature becomes more severe, different rays across the aperture of the beam undergo different amounts of deflection. Since each angle is Bragg-matched at a different wavelength, as indicated in Equation 5-15, the resulting filter shape looks weaker and slightly wider, mainly in its tails.

Better polishing techniques can efficiently reduce both effects, and obtain a more consistent measurement of the selectivity of the filter across the depth of the sample. Figure 5-33 shows the behavior of the Bragg wavelength as a function of depth for another PQ-PMMA sample with better polished edges. For the purpose of comparison with the data in Figure 5-32, the range of the horizontal axis has been fixed to 0.5 nm in both cases to emphasize the reduction in the variation of the Bragg wavelength in the second case.

5.6.4 Discussion

This section has explored the possibility of using strong holograms recorded in PQ-PMMA samples as narrow wavelength selectors for the WDM band. Judging from the experimental results, the material does not seem to provide enough modulation of its index
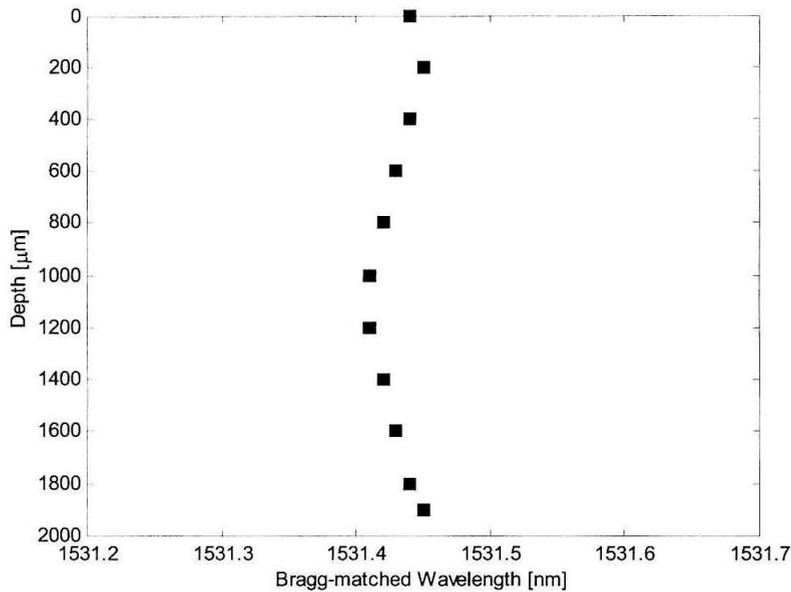
Fig. 5-33. Dependence of the Bragg wavelength of the filter with depth across the thickness of the PQ-PMMA sample.

of refraction, Δn is barely in the range of 10^{-4}, to obtain flat-top bandpass filters with enough bandwidth and suppression of neighbor channels to meet the standards for telecom applications.

High optical quality samples are essential for the successful use of this material as support for holographic filters. As already seen in Section 5.6.3, the grating is very sensitive to the flatness of the facets through which the readout laser beam comes in and out of the sample. In addition to this, the quality of the polishing has also an impact on scattering and fanning of the beam, which translates into higher insertion losses when the light of the through or the drop channels has to be coupled back into an optical fiber.

The experiments have also revealed that spurious gratings can develop during recording. Although this problem is not unique to this particular application, prolonged

exposures, necessary to saturate the dynamic range of the material and obtain strong holograms, make more likely the build-up of secondary gratings.

Finally, further material development would be necessary to understand how sensitive the diffusional dynamics of grating enhancement and erasure are to changes in the process of sample preparation and polymerization, as well as to the aging of the material on the shelf.

References

[5-1] G. J. Steckman, *Holographic Recording in Polymeric Materials with Applications*, Ph.D. thesis, California Institute of Technology, 2001.

[5-2] A. V. Veniaminov, and Y. N. Sedunov, "Diffusion of phenanthrenequinone in poly(methylmethacrylate): Holographic measurements," *Polymer Science A* **38**, 59 (1996).

[5-3] V. I. Sukhanov, "Three-dimensional deep holograms and materials for recording them," *Journal of Optical Technology* **61**, 49-56 (1994).

[5-4] J. E. Ludman, J. R. Riccobono, N. O. Reinhand, I. V. Semenova, Y. L. Korzinin, S. M. Shahriar, H. J. Caulfield, J. M. Fournier, and P. Hemmer, "Very thick holographic nonspatial filtering of laser beams," *Optical Engineering* **36**, 1700-1705 (1997).

[5-5] G. J. Steckman, I. Solomatine, G. Zhou, and D. Psaltis, "Characterization of phenanthrenequinone-doped poly(methyl methacrylate) for holographic memory," *Optics Letters* **23**, 1310-1312 (1998).

[5-6] K. Y. Hsu, S.-H. Lin, W. T. Whang, and W. Z. Chen, "Holographic data storage using photopolymers," *Proceedings SPIE on Photorefractive Fiber and Crystal Devices: Materials, Optical Properties, and Applications V* **3801**, 66-74 (1999).

[5-7] A. Popov, I. Novikov, K. Lapushka, I. Zyuzin, Y. Ponosov, Y. Ashcheulov, and A. Veniaminov, "Spectrally selective holographic optical elements based on thick polymer medium with diffusional amplification," *Journal of Optics A: Pure and Applied Optics* **2**, 494-499 (2000).

[5-8] A. V. Veniaminov, V. F. Goncharov, and A. P. Popov, "Hologram amplification by diffusion destruction of out-of-phase periodic structures," *Optika i Spektroskopiya* **70**, 864-869 (1991).

[5-9] S. H. Lin, K. Y. Hsu, W. Z. Chen, and W. T. Whang, "Phenanthrenequinone-doped poly(methyl methacrylate) photopolymer bulk for volume holographic data storage," *Optics Letters* **25**, 451-453 (2000).

[5-10] F. Rodriguez, *Principles of polymer systems 4th Edition*, Taylor & Francis, Washington D.C., 1996.

[5-11] A. Veniaminov, and E. Bartsch, "Diffusional enhancement of holograms: phenanthrenequinone in polycarbonate," submitted to *Journal of Optics A: Pure and Applied Optics*.

[5-12] V. Leyva, G. A. Rakuljic, and B. O'Conner, "Narrow bandwidth volume holographic optical filter operating at the Kr transition at 1547.82 nm," *Applied Physics Letters* **65**, 1079-1081 (1994).

[5-13] S. Breer, H. Vogt, I. Nee, and K. Buse, "Low-crosstalk WDM by Bragg diffraction from thermally fixed reflection holograms in lithium niobate," *Electronics Letters* **34**, 2419-2421 (1998).

[5-14] E. Chuang, and D. Psaltis, "Storage of 1000 holograms with use of a dual-wavelength method," *Applied Optics* **36**, 8445-8454 (1997).

[5-15] A. Yariv, *Optical Electronics in Modern Communications Fift Edition*, Oxford University Press, New York, 1997, Chapter 13.

6 Holographic systems for processing and imaging applications

6.1 Introduction

Chapter 2 introduced the concept of reconfigurable computing. It was argued there that reconfigurable processors have the advantage that can tailor its hardware resources to deal more efficiently with any given problem. In this context, the OPGA added the feature of high-speed optical interface allowing for run-time reconfiguration. This chapter investigates applications that could fully exploit the unique capabilities of such a device.

Although there are many applications that require real-time processing, it is not evident which ones could benefit from reconfiguration. Let alone, those in which reconfiguration would take place so often that would justify the need for an interface with a holographic memory. In that sense, the OPGA defines a new paradigm of applications. Section 6.2 explores its boundary by examining when optical reconfiguration starts to be advantageous over all-electronic solutions.

The prospect of helping paralyzed patients, by translating neural activity from the brain into control signals for prosthetic devices, has flourished in recent years thanks to discoveries in systems neuroscience, and to the rapid advance of MEMS [6-1]-[6-3]. However, potential barriers to continued progress in neural-prosthetic systems exist, like the relatively limited computational power available for mobile real-time processing of neural signals. Section 6.3 addresses this issue by proposing the use of reconfigurable processors,

and OPGAs in particular, to meet the computational challenges of neural-prosthetic systems.

Finally, Section 6.4 presents a different type of application: a holographic microscope capable of sensing 4-D information (i.e., 3-D spatial plus color) of a specimen. This microscope is morphologically similar to an OPGA, since the system can be still regarded as the combination of a silicon chip, a holographic memory and addressing device (e.g., a fluorescent specimen). Unlike in the OPGA, information is not stored in the holograms. The information is inherent to the addressing device (i.e., the specimen). Reconfigurability is still embedded in the system inasmuch as various imaging tasks (e.g., extracting spectral-band information, focusing at a particular depth) are performed by different holograms.

6.2 OPGA vs. cache-based FPGA

A fast interface with memory makes possible to bring the reconfiguration time down to the range of microseconds. The OPGA uses this fact to reprogram the chip in real-time. However, the OPGA is not the only approach to achieve dynamic reconfiguration. Some FPGA manufacturers, like Xilinx [6-4], have tried a different solution. They have developed FPGAs where a cache memory is built in the chip to store locally, inside the FPGA, a limited number of configurations, Figure 6-1. The cached configurations can be accessed very fast, in just a few nanoseconds, and transferred in parallel to the logic. The cache memory can store those configuration templates that are more frequently used, reducing the number of accesses to the slow external memory to fetch a new configuration. The asymmetry between many accesses to a fast cache memory and only a few accesses to the slow external memory helps to amortize the cost of the reconfiguration during run time.
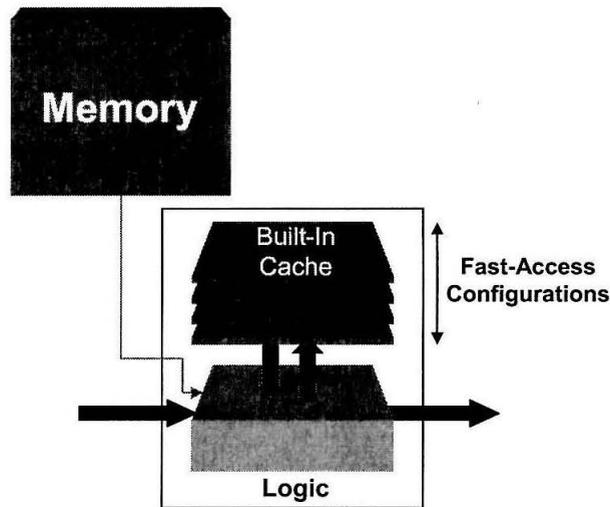
Fig. 6-1. Structure of an on-chip cache memory FPGA.

The on-chip cache memory FPGA decreases the reconfiguration time at the expense of decreasing the amount of logic implemented on the chip, since the same die area needs to be shared to implement banks of SRAM.

(a)   (b)
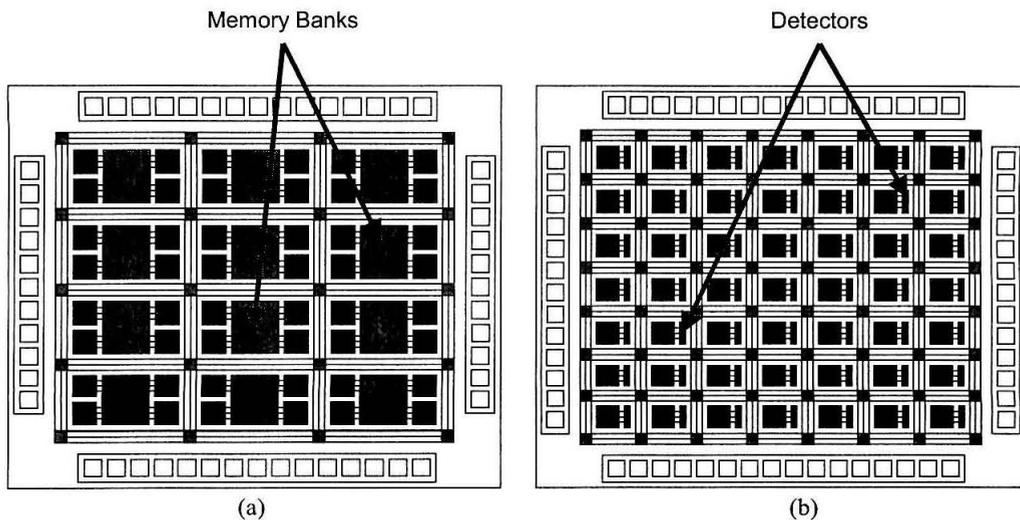
Fig. 6-2. Area trade-off between logic blocks and (a) banks of memory in the built-in cache memory FPGA, or (b) photodetectors in the OPGA.

This section presents a comparison between the two architectures (Figure 6-2), OPGA versus cache-memory-based FPGA, in terms of logic density. In both cases, it is necessary to trade off some area of the chip to implement either an array of photodetectors, in the OPGA, or banks of RAM memory, in the cache-based FPGA. The goal is to investigate which one of these two approaches allows for a higher computational capacity defined in terms of CLBs.

The model analyzes the number of CLBs that can be implemented in a fixed die area as a function of the number of configuration templates that the system is dealing with. This model does not make a distinction between logic blocks and programmable interconnects and buses. It only counts how many CLBs can be implemented in the available area after having subtracted the area dedicated to photodetectors or cache memory. This number can be understood as CLB-equivalents, so that a switching matrix or bus segments can have a cost in terms of CLB-equivalents. Thus, it is left as a chip design decision how to balance that number of CLB-equivalents among real CLBs, interconnect matrices and buses to reach some optimum. The number of CLBs, N_{CLB}, can be expressed as the ratio of the total die area and the area per logic block, A_{CLB}, and its area overhead due to either photodetectors or cache memory. If b_{CLB} stands for the number of bits required to configure a CLB, then the optical overhead is b_{CLB} times the area of a pixel detector, $A_{Detector}$. Therefore, the number of CLBs for the OPGA is given by

$$N_{CLB} = \frac{\text{Die Area}}{A_{CLB} + b_{CLB} \times A_{Detector}}. \qquad (6\text{-}1)$$

On the other hand, the cache-memory overhead is given by the size in pages of the RAM, N_{Config}, and the area of a 1-bit memory cell, A_{RAM}. The number of CLBs that can be implemented in a cache memory based FPGA results in

$$N_{CLB} = \frac{\text{Die Area}}{A_{CLB} + N_{Config} \times b_{CLB} \times A_{RAM}}. \qquad (6\text{-}2)$$

If some numerical values are taken into consideration, the number of CLBs for both situations can be plotted as a function of the number of templates. For the simulation in Figure 6-3, the CLB considered is similar to the one in the XC3000 series from Xilinx [6-5]. The logic block requires 64 bits of configuration and its area is 291×156 µm. For the banks of memory, both dynamic RAM (DRAM) and static RAM (SRAM) have been considered, although due to the nonvolatility of the optical memory it is more fair a comparison with SRAM rather than DRAM. The size for 1-bit memory cell is 8 µm^2 for SRAM and 1.5 µm^2 for DRAM. The analysis has studied two sizes of detectors: 5×5 µm and 2×2 µm pixel size.

As it can be observed in Figure 6-3 and described by Equation 6-1, the overhead due to the photodetectors remains constant independently of the number of templates in the application. However, the performance of the cache-based FPGA decreases as more and more templates need to be stored in the chip using up the area dedicated to logic. Note that in Figure 6-3 the horizontal axis is in logarithmic scale. For a small number of templates, the conventional FPGA outperforms the OPGA since almost no reconfiguration is involved in the application. Nevertheless, the break-even point occurs around 3 or 4 configuration pages. As expected from the difference in area, the decrease in performance for a SRAM cache starts sooner than for the DRAM cache. It is important to mention that this model
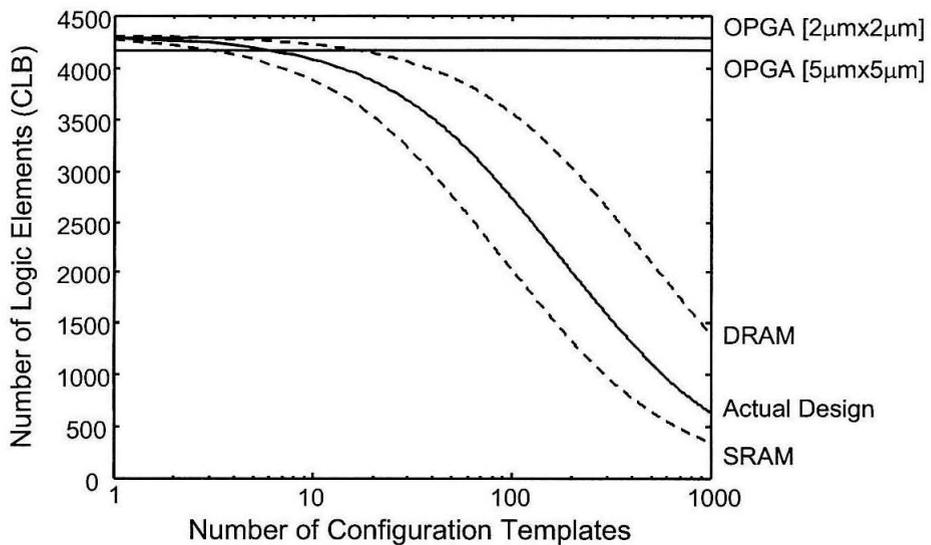
Fig. 6-3. Performance comparison between OPGA and cache-memory FPGA. The number of logic elements as a function of the number of configurations is plotted for two different pixel-sizes in the case of the OPGA and compared to SRAM and DRAM (dashed lines), and to an actual design from [6-6] (solid line)

does not consider the overhead in the memory due to additional circuitry like sense amplifiers or row/column decoders. It is difficult to quantify such overhead because it can significantly vary from one chip to another. However, the figure contains the curve that corresponds to an actual design of a DRAM-cache-based FPGA [6-6]. The effect of this overhead is to shift the curve downwards. The OPGA outperforms clearly the cache-based FPGA for applications involving the order of hundreds of templates. This defines a domain of applications that can be carried out much more efficiently using an OPGA.

6.3 Neural prosthetics

In broad terms, neural-prosthetic systems or brain-computer interfaces aim to provide disabled patients with new options for interacting with the world [6-1]. Sensory prosthetics, such as cochlear implants or artificial vision, encode information from the

environment and deliver it to the nervous system by appropriate electrical stimulation. Motor prosthetics work in reverse, by translating neural activity into control signals for prosthetic devices, to assist patients with upper spinal cord injuries, neuro-degenerative diseases or amputations [6-7]-[6-10]. Finally, intra-central-nervous-system prosthetic systems intervene with ongoing neural processing by stimulating brain regions or by recording from one region and stimulating another region [6-11].

Despite the great variety of sensors and actuators needed to address the dysfunctions mentioned above, sensory, motor, and intra-central-nervous-system prosthetic systems have many computational requirements in common. From a computational perspective, these systems will likely grow to look even more alike as, for example, motor-prosthetic systems evolve to also include supplementary sensory feedback to the nervous system and all prosthetic systems incorporate learning and adaptation to contend with, and take advantage of, neural plasticity. The anticipated similarities among these systems, as well as the demanding requirements imposed by mobile real-time processing of neural and/or sensory data, motivate the development of a processor optimized for neural-prosthetic systems.

Reconfigurable processors are well suited for neural prosthetics for three principle reasons. First, neural-prosthetic systems require many diverse computations and a single processor capable of being "rewired" rapidly can efficiently perform a wide range of calculations. Second, neural-prosthetic systems run in real time and reconfigurable processors can meet these real-time demands by being "wired" nearly optimally for any given task, which often includes a parallel processing topology. Finally, neural-prosthetic systems are likely to require greater computational resources as, for example, the number and variety

of sensors (e.g., electrodes) used to collect information expands. Reconfigurable processors, as with other high-speed electronic systems, are likely to scale well as the number of recorded neural signals increases due to the relatively slower time scale of the biological system, which allows time-multiplexing schemes to absorb the increasing computational demands.

How can reconfigurable processors in general, and OPGAs in particular, potentially be applied to neural-prosthetic systems? In order to see more easily the relevance and potential merits of reconfigurable processing, a specific example of a motor-prosthetic system, in particular a prosthetic arm, will be introduced. The goal of the discussion is simply to point out a few of the principles when considering reconfigurable processors in the context of neural-prosthetic systems. It is beyond the scope of this section to present a detailed design.

6.3.1 Example of a neural-prosthetic system

Figure 6-4 is a block diagram of a system that translates cortical activity into control signals for stimulating the musculature of a paralyzed arm. A person typically sees (senses) an object that he/she wishes to reach toward, forms a mental plan for where and how to move the arm, and finally sequences through the movement commands. Different attributes of this movement are manifest in different regions of cortex, with any one attribute (e.g. reach location) encoded across numerous neurons. This neural activity can be sensed in many ways, typically with permanently implanted electrodes. It is thought that tens to hundreds of electrodes, implanted in several cortical areas, will eventually be needed to gather enough detail of the motor plan to accurately reconstruct the desired movement in real time.

Therefore numerous neural-signal channels need to be amplified, filtered, and digitized for subsequent processing, and much of this circuitry may eventually be integrated with or near the recording electrodes. After passing through this front end, a digitized signal stream from each electrode must be processed to associate action potentials (spikes) with particular neurons, to estimate spectral power density, and to estimate other spatio-temporal signal features that researchers are continuing to relate to movement parameters. These signals can then be compared with previously characterized responses of each neuron and electrode to arrive at a moment-by-moment estimate of the desired movement parameter, such as the arm movement direction or end-point location. Common estimation methods include maximum likelihood and neural networks.
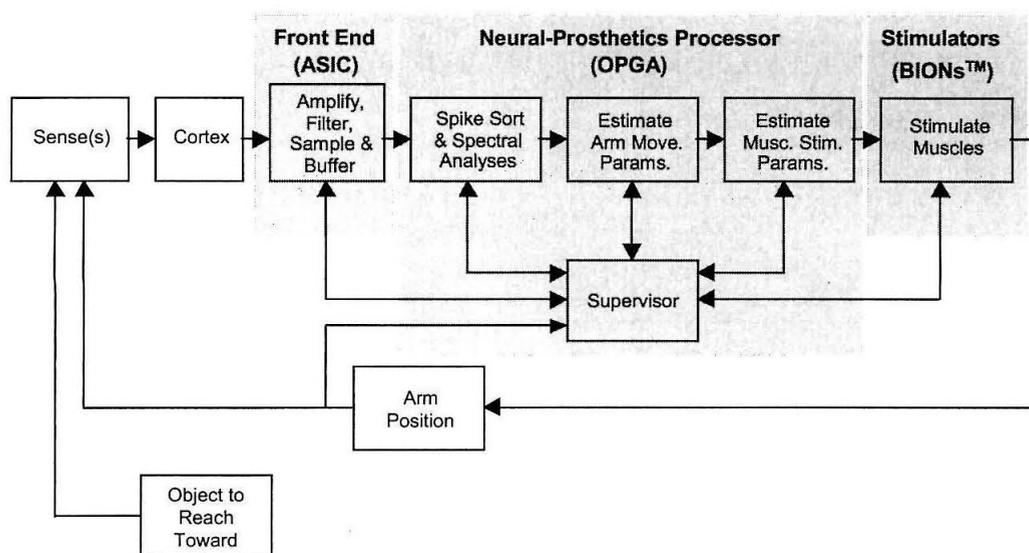
Fig. 6-4. Block diagram of a motor-prosthetic system (prosthetic-arm system). Information can be transmitted down subcutaneous wires or with telemetry, in which case additional transmit and receive circuitry is required. The block elements are grouped, according to their possible implementation, as three subsystems (boxes shaded in gray). The front end requires specialized circuitry that may best be implemented as an ASIC. The Reconfigurable Neural-Prosthetics Processor (RNPP), which could be implemented with a reconfigurable processor like the OPGA, performs many of the block element operations. Muscle stimulation could be performed with BION[TM]-like stimulators, injected into muscles.

Once the movement parameters of interest have been decoded from the neural measurements, the neural-prosthetic system must generate estimates for muscle-stimulation parameters (i.e., inverse kinematics). In this particular example, the goal is to electrically stimulate the paralyzed arm's musculature to achieve arm movements. Control of this prosthetic-arm system is achieved through negative feedback by comparing visually the arm's new position to the desired location and iterating as necessary. Importantly, even with careful calibration of the entire system such that a person's desired arm movements are executed accurately, the system will change with time and with experience. As time passes, recording quality changes due to electrode drift and neurons dying, and the efficacy of muscle stimulation can also change. As users gain more experience, neurons will almost certainly adapt (plasticity) in order to improve the performance of the system, as the brain does whenever presented with a demanding new task. Without also adapting the neural-prosthetic system to contend with and take advantage of these changes, system performance will eventually deteriorate to the point of being useless. A supervisor could monitor these time-dependent and experience-dependent changes, and adjust the system accordingly. For example, if one of the neurons controlling the system drifts out of recording range or even dies, the supervisor can remove this neuron from the database thereby making performance robust against such events. Then, if a new neuron becomes detectable the supervisor can monitor its response to ongoing prosthetic-arm movements in order to learn the encoding characteristics of this new neuron.

Several blocks in Figure 6-4 can be grouped together according to how these elements could be implemented. There are three of these groups, delimited by the gray boxes in the figure, which will be referred to as subsystems. The first subsystem is the front end,

which consists of amplifiers, filters, and A/D converters needed to transform the continuous-time and continuous-voltage neural waveform from each electrode into discrete-time and discrete-voltage signals. As these functions require highly specialized and optimized circuits (e.g., low-noise amplifiers, fast high-precision A/D converters) they are likely to be best implemented in an ASIC. As it will be discussed below, it is also possible that additional front-end functions such as signal buffering, which is straightforward to include on an ASIC, could simplify overall system design.

The second subsystem is the neural-prosthetics processor, which could be implemented in an OPGA. The reconfigurable neural-prosthetics processor (RNPP) as drawn in Figure 6-4 would handle four major functions described previously: spike sorting and spectral analyses; estimating (decoding) arm-movement parameters from the neural data; estimating (inverse kinematics) appropriate muscle stimulation parameters given the estimated arm-movement parameters; and supervising performance and adjusting the parameters of the system accordingly. In addition, the RNPP could perform two additional functions that are needed in the system architecture envisioned for this example. First, when a block of neural data is needed, the RNPP requests this data from the front-end subsystem that temporarily buffers data for each electrode channel (see Section 6.3.2). Second, when stimulation parameters have been estimated, the RNPP sends these parameters onto the stimulation subsystem (see Section 6.3.2).

The final subsystem illustrated in Figure 6-4 is the muscle stimulators. Current state-of-the-art muscle stimulators, such as the BION^{TM} [6-12], are capable of being injected into muscle with a hypodermic needle, powered wirelessly, and receiving a wireless digital transmission including a stimulator-specific identifier and stimulation parame-

ters. Thus to move the arm, the RNPP need only encode the muscle stimulation parameters appropriately and issue the data packets.

6.3.2 Reconfigurable Neural-Prosthetics Processor (RNPP) Subsystem

Figure 6-5 illustrates how a reconfigurable processor might operate in a neural-prosthetic system. This example applies an OPGA to the prosthetic-arm system described above and is intended to suggest a general architecture, not a fully functional design. Neural data flows in from the left, through the RNPP (depicted in gray), and muscle-stimulator commands flow out on the right. The RNPP subsystem figure consists of an OPGA chip area versus time plot, which suggests how OPGA electronics could be reconfigured through time, as well as the OPGA holographic memory (at top of figure).
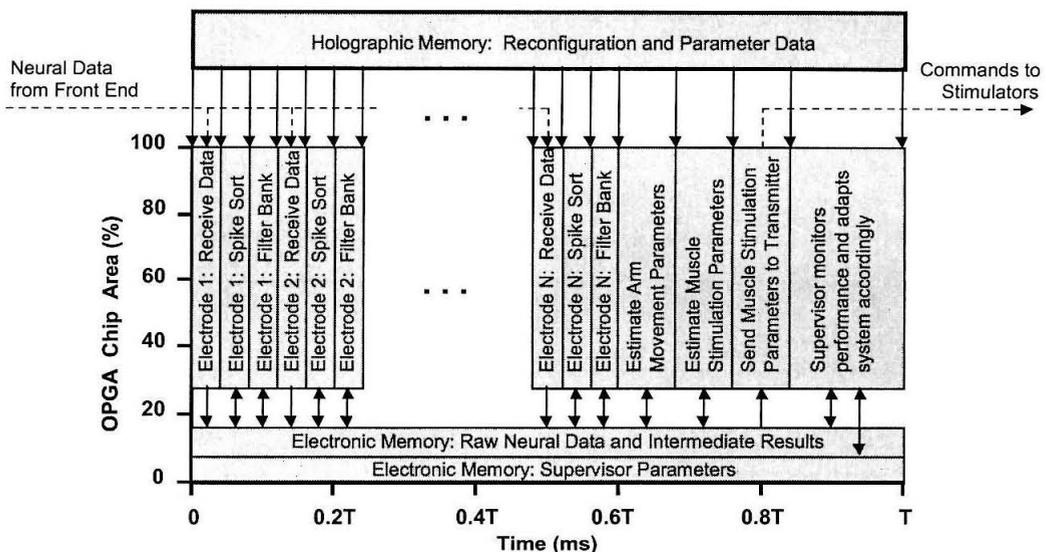
Fig. 6-5. Possible Reconfigurable Neural-Prosthetics Processor (RNPP) architecture using an OPGA. Neural signals from the front-end ASIC enter from the left (dashed lines) and muscle-stimulation parameters exit on the right (dashed lines). The OPGA is represented in two main parts, both shaded in gray. The electronic portion is represented as an OPGA Chip Area vs. Time plot, which illustrates how the OPGA electronics are allocated and reconfigured through time. The RNPP operates in cycles, with a period of T (ms). The holographic memory is depicted at the top of the figure. Lines indicate information flow, with arrows indicating the direction of this flow.

When operation begins, the front-end ASIC starts storing neural data samples in an on-chip buffer. This buffer can be thought of as a memory page with rows corresponding to sample number (time) and columns corresponding to electrode number. The number of rows equals the sampling rate termed R (e.g., 40 Ksamples/s or 40 samples/ms) multiplied by the period of time required to service all data in the buffer, represented by T (ms). The number of columns (electrodes) is termed as N, which could be on the order of tens to hundreds. Each memory element is B bytes, which is typically two (e.g., 12–16bit samples). From time 0 to T, the first buffer is filled and from time T to 2T a second buffer is filled. This allows the RNPP T ms (the interval T–2T in real time) to service the neural data stored in the first buffer. The first buffer is overwritten from 2T to 3T, while the RNPP services the second buffer, and so on. This architecture allows the ASIC to implement most of the data memory (e.g., 2 buffers × RT samples × N electrodes × B bytes/sample = 2RTNB bytes) freeing the RNPP to store only a small amount of neural data at any time (e.g., RT samples × 1 electrode × B bytes/sample = RTB bytes). For example, if a 10×10 electrode array were implanted, and the electrical signals sampled at 40 kHz for 20 ms time frame and afterwards digitized with 12 bit precision, a 120 Kbyte RAM memory for each page would be enough to store all the neural data in the ASIC. The OPGA would require 1/N as much on-chip memory, 1.2 Kbytes.

As shown in Figure 6-5, the RNPP must complete all of its operations within one period (T ms). This period is bounded by the maximum allowable time between muscle-stimulator updates, which could range from hundreds of ms for coarse motor control to just a few ms for fine motor control. While T may be made as small as current technology will allow, the ultimate limit rests with the time scale of neural representations and the informa-

tion transmission rate achieved by the front-end sensors. An important point is that while T ms pass between muscle-stimulator updates, the absolute latency of the system (i.e., time from neural event until muscle stimulation) is 2T ms, as T ms is consumed in the front end and T ms is consumed in the RNPP. Therefore, the maximum allowable period T must also take into consideration the maximum allowable latency in the closed-feedback prosthetic system.

The first RNPP operation in this example architecture is to configure the OPGA to request, receive and store neural data. The OPGA is configured by reading a page of configuration data from the holographic memory (downward-directed arrow from holographic memory at far left of Figure 6-5) and is fast enough (1–10μs) that no appreciable time is allocated to this operation in the figure. The second OPGA operation is to request T ms of Electrode 1 data from the front-end ASIC, receive these data, and place these data into on-chip OPGA memory. This operation consumes some finite amount of time, illustrated as the width of the *"Electrode 1: Receive Data"* bar, and occupies some fraction of the total OPGA electronics area, illustrated as the height of the bar. Again, this figure is meant simply to be suggestive of architectural principles. A certain fraction of the OPGA electronics area is configured as memory, with *"raw neural data and intermediate results"* and *"supervisor parameters"* each having their own reserved regions.

The next RNPP operations are to configure the OPGA for spike sorting, spike sort T ms of Electrode 1 data, and place the results in the OPGA memory. Spike sorting requires the OPGA electronics to be configured as an efficient DSP-like processor in order to correlate the neural data with neural-waveform templates [6-13]. Electrodes typically sense action potentials originating from many neurons with 1–4 of these neurons being identifi-

able based on their voltage-time waveform shapes. Statistical neural templates appropriate for each electrode can be learned off line, stored in the holographic memory, and then retrieved just before spike sorting data from a given electrode. Cross-correlating these neural-waveform templates (1–4 templates, roughly 2–3 ms in duration each) with the neural data from Electrode 1 (T ms in duration) requires shift (delay), multiplication and addition operations that are straightforward to implement in FPGAs/OPGAs. Let's consider a hypothetical, although not unrealistic, OPGA device containing 3500 CLBs and a 30 Kbyte bank of on-chip RAM memory. In such a device, four template-matching filters for a 2msec-long sequence each could be simultaneously implemented using slightly less than 1000 CLBs, assuming that we take advantage of efficient distributed computing algorithms, like the ones described in [6-14]. This low CLB count would even allow for the processing of multiple electrodes in parallel. The identity and time of each action potential in the T ms of Electrode 1 data is stored for later analysis. This does not consume large amounts of memory since spike rates are relatively low (e.g., <100 spikes/s on average) and identity and time information can be compact. Just 80 bytes per electrode, for example, if 10 matches are found on average in each one of the filters, and each match is encoded as 2 bytes.

The final RNPP operation that must be performed on data from each electrode is spectral analysis or digital filtering. Digital filters appropriate for estimating the power in a given frequency band, for example, can be designed off line given off-line data from each electrode. These filter coefficients are likely specific to each electrode. Since the power in multiple frequency bands may be of interest, the OPGA could be configured as a DSP filter bank, which again requires delays, multiplication and addition operations. A 256 tap filter,

using 12-bit precision complex coefficients, can be implemented in the OPGA using just 715 CLBs [6-14]. Therefore, there is enough hardware available to implement a bandpass-filter bank to obtain the spectral information of the neuronal signals in three different regions of the spectrum simultaneously. This suggests that even FFT analyses are possible. Filter coefficients for the digital filters of each electrode and the filter-bank electronics configuration data is stored in the holographic memory and is downloaded just before analyzing data from a given electrode. After analysis, the relatively compact spectral estimates, maybe just 16 bytes per electrode, are stored for later use.

This sequence of operations — receive T ms of buffered neural data, spike sort, and filter — repeats until all N channels of electrode data have been processed. Together all such operations must consume less than T ms (sketched as 0.6T ms in Figure 6-5) to allow sufficient time to complete the remaining RNPP operations (sketched as 0.4T ms). Continuing with the example, if the OPGA could be clocked at 166 MHz, the time required to compute the spike sorting and the spectral analysis could be as low as 116 μs, which with the reconfiguration overhead of 2 μs for the two reconfigurations becomes 118 μs per electrode. Therefore the array of 100 electrodes could be processed in 12 ms (within 60% of 20 ms). Importantly, the key reconfigurable-processing architectural principles can already be seen at work. First, the diverse neural-prosthetic computations are accommodated by the OPGA rewiring rapidly, in order to efficiently perform a range of calculations. Second, the OPGA meets the real-time demands by being "wired" nearly optimally for any given task, which includes parallel processing topologies. Finally, the OPGA is able to scale well with the number of electrodes (sensors) delivering neural data by time-multiplexing its operations. The number of electrodes this OPGA can handle is set by the speed of the processor,

not by the number of parallel circuits that will fit within the area of the chip. This ability to time-multiplex the processing of data from increasing number of electrodes, as opposed to adding additional physically parallel circuits which consume more chip area, is afforded by the relatively slower biological time scale and by the fast and parallel circuits possible in FPGAs/OPGAs.

The next RNPP operation is to estimate the arm-movement parameters. The goal is to estimate how the arm should move (e.g., new x, y, z location in space) given the new neural observations extracted from the preceding T ms in time (e.g., spike times and spectral power density). Though the best way to perform this estimation is a matter of current research, all methods require a database for how each neuron/electrode responds for real or intended reaches in numerous directions. This database can be constructed off line, stored in the holographic memory, and retrieved when the RNPP needs to estimate arm movements. The OPGA should also be configured to perform any of a number of estimation algorithms (e.g., maximum likelihood, Bayesian analysis, neural network) and, again, this configuration data is stored in the holographic memory. As before these algorithms reduce to multiplications and additions where the FPGA can perform millions of those per millisecond [6-14]. The results of this estimation are quite compact, potentially as small as the new x, y, z arm location, for example, just 6 bytes. Importantly, arm-movement estimation scales well as the number of neurons/electrodes increase.

After estimating and storing the new arm location, for example, the RNPP must estimate how each of several muscle stimulators should be activated in order to direct the arm to this desired location. Estimates of this sort require a model for where each muscle-stimulator is implanted, how muscle stimulation leads to muscle contraction, and how this

contraction moves the arm. These models are then run in reverse to arrive at stimulation parameters given the desired arm location. These reverse models, and the OPGA electronics needed to run them efficiently, are stored in the holographic memory. Upon completion, the identity of each stimulator (1 byte) and its stimulation current level (1 byte) and duration (1 byte) are stored in the on-chip memory, for delivery to BION[TM]-like stimulators [6-12].

The last RNPP operation before the arm starts moving is to send these muscle-stimulation parameters to the transmitter. As with the front-end ASIC to RNPP connection, the RNPP to muscle-stimulator connection is envisioned to contain a wireless link. Therefore, the RNPP must simply send the wireless transmitter the identity of each muscle-stimulator and its stimulation parameters, perhaps in an appropriately encoded packet format, and the transmitter will broadcast the instructions. Each muscle stimulator will activate accordingly.

The final RNPP operation is to perform the supervisory duties described in the previous section. Many signals and conditions are expected to change throughout the lifetime of such a neural-prosthetic system, therefore ongoing adjustments are almost certainly needed for adequate performance over months or years. At the front end, electrodes may drift, thereby changing the recording characteristics. The RNPP should adjust the spike-sorting and spectral-analysis algorithms accordingly. At the arm-movement estimation stage, neural plasticity can change the response characteristics of neurons. Thus, the response database and/or estimation algorithms must learn/adapt accordingly. Finally, at the back end, it is likely that over time muscle stimulation will eventually lead to slightly different arm movements. Again, the supervisor should adapt the model parameters appro-

priately. While the basic supervisory logic circuitry can be stored in, and retrieved from, the holographic memory, much of the information needed by the supervisor must be stored in OPGA electronic memory (see *"Electronic Memory: Supervisor Parameters"* in Figure 6-5). The supervisor must analyze past neural signals and system performance, store intermediate assessments, and store numerous adjusted parameter values to be accessed by the other RNPP operations (e.g., new spike-sort parameters).

6.3.3 Discussion

Section 6.3 attempted to illustrate how a reconfigurable processor, the OPGA, might be applied to an example neural-prosthetics system, the prosthetic-arm system. This so called reconfigurable neural-prosthetics processor (RNPP) subsystem performs most, if not all, of the signal processing, estimation, and control essential for the prosthetic-arm system. The OPGA-based RNPP is able to achieve this level of performance by virtue of its inherent reconfiguration speed, parallel and optimized circuitry, and the use of a time-multiplexing scheme. According to the stated assumptions and approximations, a single OPGA built in current FPGA technology would be able to perform all RNPP tasks for approximately 100 electrodes. Importantly, as semiconductor and optical technologies continue to advance a single OPGA could be capable of processing neural signals from more electrodes, perform more complex computations and control, or both. To again emphasize the broad range of neural-prosthetic systems that can be envisioned, it is also conceivable that OPGAs could contribute meaningfully to sensory-prosthetic systems (e.g., transforming images into electrical-stimulation patterns) since FPGAs have already made an important impact on image processing.

6.4 Holographic 4-D imaging

Optical data storage is one of the areas in which holography has been extensively used. In a holographic memory, the information is encoded into the hologram and retrieved with a laser beam, usually a simple plane wave. This section deals with the use of holography for imaging applications. In this case, the information is no longer stored in the hologram itself, but in the spatially encoded wavefront of the probing beam. The Bragg-selectivity property of a volume hologram that is extremely useful for high-density data storage, can in fact be also efficiently utilized in imaging applications, like confocal microscopy [6-15], [6-16] or tomography [6-17].

The problem here considered is the use of holograms to extract spatial and color information (hence 4-D imaging) of a specimen and project it onto a 2-D space on the detector. Apart from holography, other techniques are used to perform such transformation. For example, the confocal microscope places a pinhole at the finite conjugate distance of the imaging lens before the detector in order to discriminate one point at a particular depth in the specimen. Light originating from neighboring points will miss the pinhole, or will be defocused, so the pinhole filters out background noise improving the resolution of the system. A variation of this technique is the two-photon excitation microscopy (TPEM) [6-18], [6-19], in which 3-D localization is achieved by focusing high-energy femtosecond pulses inside the specimen, instead of using a pinhole.

Both confocal and two-photon microscopy suffer from the same limitation, namely only the information of one point of the specimen is gathered at a time. Consequently, there needs to be a scanning mechanism to reconstruct the entire volume of the specimen. Moreover, if color information is required, several measurements need to be taken using differ-

ent color filters. The inherent parallelism that holographic systems enjoy can satisfactorily overcome these limitations. Unlike the aforementioned techniques, holographic imaging renders simultaneously a 2-D slice of the specimen on the detector. Furthermore, by multiplexing several holograms, many of such slices (at different depths and in different color bands) can be sensed in parallel, making unnecessary the use of sophisticated and time-consuming scanning schemes. Therefore, the holographic microscope has an enormous potential for applications that demand real-time microscopy.
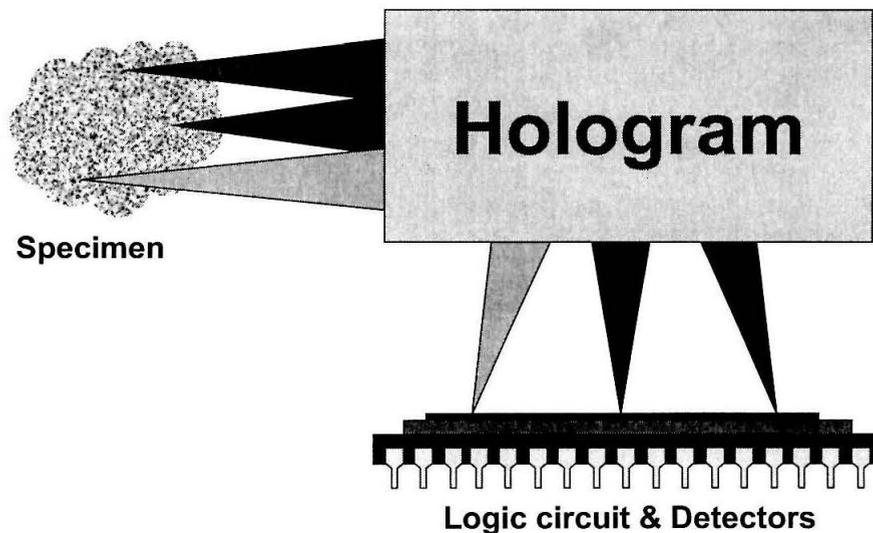
Fig. 6-6. 4-D microscope interfacing a holographic module with a chip that integrates photodetectors and processing resources.

Although essentially different in operation to the OPGA, the 4-D microscope still shares important commonalities with the former. Like an OPGA, the holographic microscope (sketched in Figure 6-6) consists of the integration of three major components: a holographic element, a silicon chip (e.g., CCD camera, CMOS imager...) and an addressing device (e.g., a fluorescent specimen). However in the microscope, instead of having

VCSELs reading out complex datapages stored in the optical memory, the spatially encoded light from the specimen is diffracted by simple gratings.

The 4-D microscope also enjoys the concept of reconfigurability, since the holograms stored in the device perform different tasks, like focusing at a particular depth or filtering a specific spectral band. Depending on the light wavefronts emanating from an unknown fluorescent specimen, various holographic templates will be selected, and the information delivered to different areas of the detector array in the chip. As in the OPGA, the adaptability and parallelism of the holographic microscope are vital for the real-time processing of the information contained in the specimen.

Finally, the 4-D microscope could also benefit from a direct interface between its holographic module and a silicon chip that integrates both logic circuitry and a detector array, since it would allow to implement a variety of image-processing tasks, like template matching for example, or even perform on-chip analysis to extract information from the specimen.

6.4.1 Principle of operation

To illustrate the principle of operation of the holographic microscope, consider the case of a transmission-geometry volume hologram, as depicted in Figure 6-7. Assume that a pure sinusoidal grating has been recorded by the interference between a plane wave propagating along the $-\hat{x}$ direction and a collimated beam propagating along the $\hat{z}$ axis generated by a point source located on-axis at the focal plane of lens L1.

Because of the angular Bragg selectivity of the hologram, strong diffraction is only achieved when the position of a probing monochromatic point source on the (x, y)-plane
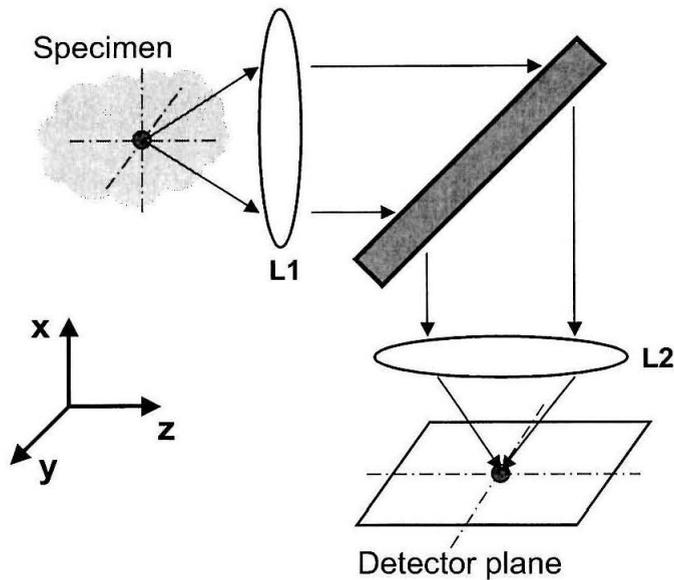
Fig. 6-7.   Schematic of a microscope that utilizes a transmission-geometry volume hologram. L1 and L2 are lenses.

matches the location of the point source used during recording. The diffracted light is focused into a point on the detector plane. If the probing point source is shifted along the x-axis, the hologram becomes quickly Bragg-mismatched and no light is diffracted towards the detector. One exception to this is what occurs along the y-axis. The recording geometry makes this direction be degenerate, which means that as the probing point source is shifted along this axis, the Bragg condition is always satisfied. Therefore the points in the specimen along a line parallel to the y-axis map into a line on the detector plane.

Similarly, the grating also exhibits Bragg selectivity in wavelength, which means that if the wavelength of emission of the probing point source differs from the wavelength at which the hologram was recorded, then the amount of diffracted power on the detector becomes negligible very rapidly.

In order to produce a full 2-D image on the detector plane, consider the Bragg condition for a symmetric grating given by

$$\Lambda = \frac{\lambda}{2 \cdot \sin\theta},$$

(6-3)

where Λ is the spatial period of the grating, λ the wavelength of the probing beam, and θ its angle of incidence (measured with respect to the surface normal of the material). Equation 6-3 establishes a coupling between wavelength and angle (or equivalently position along the x-axis) that makes it possible for points along the x-axis to Bragg-match the hologram, provided that the appropriate wavelength is selected. Therefore, when a specimen is illuminated with an extended broadband source, the hologram performs a one-to-one mapping between points on the (x, y)-plane and the detector plane.

As in a confocal microscope, the hologram is also able to discriminate light originating from different depths. The selectivity in depth is related to the angular Bragg selectivity of the grating. As the probing point source is shifted along the z-axis away from the focal plane of lens L1, the wavefront after the lens becomes spherical. A spherical wave can be regarded as the superposition of plane-wave components propagating at different angles. As the wave acquires more sphericity, fewer components remain Bragg-matched, therefore the amount of diffracted power decreases to zero.

6.4.2 Experimental imaging

A holographic microscope has been built to test its performance and use it in imaging applications. The sketch in Figure 6-8 depicts a basic transmission-geometry setup, in which the angle of incidence of each arm is 45° (in air) with respect to the surface normal of the optical medium. In fact, the holographic microscope consists just of a ×40/0.65NA
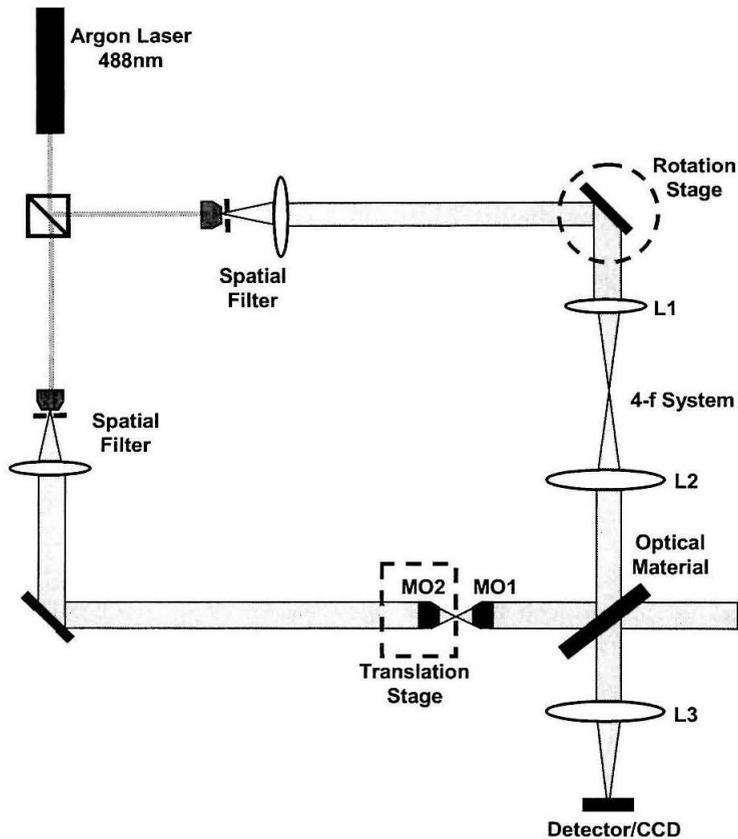
Fig. 6-8.   Experimental setup of the recorder for the 4-D microscope. MO1 and MO2 are two microscope objectives, while L1, L2, and L3 are lenses. During normal operation, the microscope consists only of MO1 and L3, in addition to the optical material.

microscope objective (MO1 in Figure 6-8), the holographic element and an imaging lens

(an 8 cm focal-length lens labeled as L3) to relay the specimen on the CCD.

The recording process occurs *in situ*, and requires the additional optical elements

shown in the figure. In order to simulate a point source, a second ×40/0.65NA microscope

objective (MO2) is placed in the signal path before MO1 to focus the beam. Microscope

objective MO1 collimates the signal wavefront before it impinges the optical medium. To

tune the hologram to get information from a particular depth, the position of the simulated

point source can be varied by means of a motorized translation stage on which microscope

objective MO2 has been mounted. After microscope objective MO1, a slightly converging (diverging) signal beam interferes with the reference beam, a simple plane wave, to record a hologram sensitive to a depth longer (shorter) than the focal length of the microscope objective.

The information of multiple depth-slices can be obtained by multiplexing several holograms. In this case, each depth-slice can be separated on the detector plane (the Fourier-plane of L3) by slightly changing the angle of the reference beam during recording. In the experimental setup, this is achieved with a mirror mounted in a rotation stage and a pair of lenses (L1 and L2) forming a 4-f system. Finally, information from different spectral bands can be gathered by tuning the holograms accordingly, which would require to adjust the angle of both recording beams. Although possible, this feature has not been added to the experimental setup.

The resolution of the system has been checked in first place. A single hologram was recorded using the 488 nm line of the Argon laser in a 5 mm thick x-cut $LiNbO_3$:Fe crystal with 0.05% wt doping. The intensity of each beam was 9 mW/cm^2 and the crystal was exposed for 3 hours. At the end of the experiment the diffraction efficiency of the hologram reached 13%. The selectivity in depth of the hologram as the simulated point source is shifted along the z-axis (as defined in Figure 6-7) from its position during recording is presented in Figure 6-9. The experimental data (dots) is compared to the theoretical prediction (solid line) calculated using the method described in Section 6.4.3.3. The measured depth resolution (Δz) is 4 μm FWHM, which is larger than the predicted value of 2 μm. However, an experimental value larger than the prediction is expected due to aberrations, mainly in the imaging lens but also in the microscope objective.
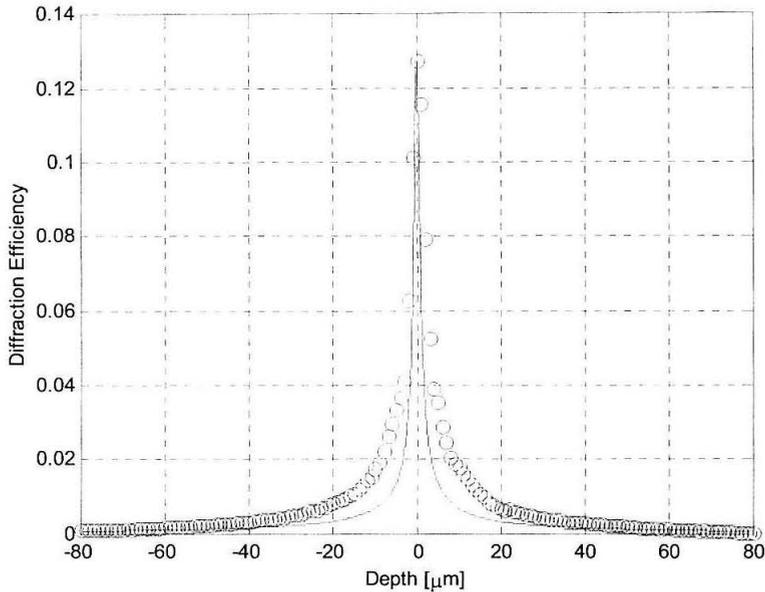
Fig. 6-9.   Depth selectivity measurement (dots) of a hologram recorded on 5 mm thick LiNbO$_3$ crystal compared to the theoretical prediction (solid line).

The angular selectivity curve of the hologram has also been measured and compared with the theory (Figure 6-10). The Bragg selectivity is approximately 1.25×10^{-2} degrees (outside the crystal), which means that, using a $\times 40/0.65$ microscope objective, if a monochromatic point source is shifted by more than 1µm along the x-axis (as defined in Figure 6-7) on the input plane of the microscope, the hologram will not be Bragg-matched. In other words, the hologram decomposes the input plane into 2 µm thick monochromatic stripes.

The ability of the hologram to image has been tested next. During normal operation (i.e. once the recording process has finished) objective MO2 is retracted so that a specimen can be placed in front of objective MO1 and observed under the microscope using white light illumination. Figure 6-11 shows the image of an 11 µm pixel mask as rendered by the
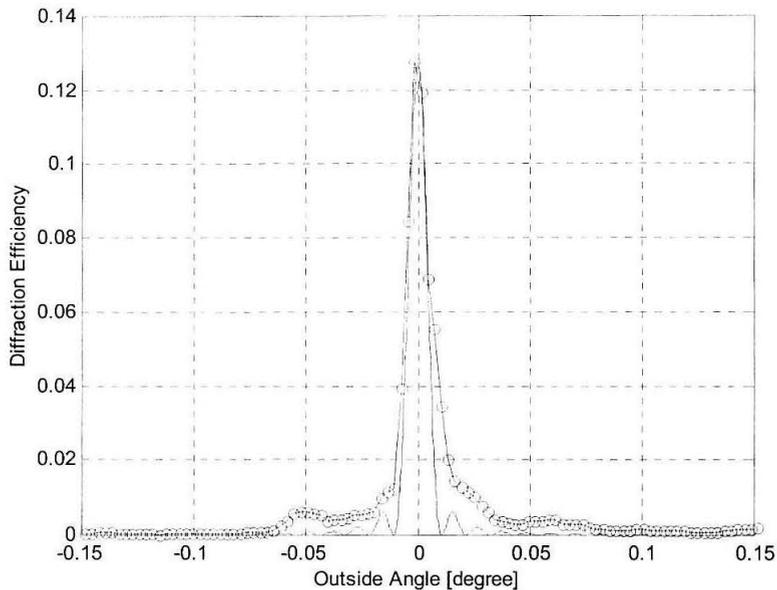
Fig. 6-10.   Experimental angular selectivity curve (dots) of a hologram recorded on 5 mm thick LiNbO$_3$ crystal compared to the theory (solid line).

hologram on the CCD camera. For this particular experiment, a ×10/0.25NA objective was used, so a larger field of view can be observed under the holographic microscope.

Fig. 6-11.   Image rendered by the hologram on the CCD of an 11 μm pixel mask when illuminated with a white light lamp.

One of the advantages of the holographic microscope over the more traditional confocal techniques is that a 2-D image of a depth/color slice of the specimen is obtained

simultaneously rather than just one point of it at a time, thus making unnecessary any mechanical scanning. This property becomes particularly interesting in applications in which information of the specimen needs to be acquired in real time, like, for example, the movement of micron-sized particles in turbulence. In the experiments, an aqueous solution containing 15 μm diameter polystyrene beads was prepared and poured into a container made of micro cover-glass walls. The evolution of the microspheres as they float in the solution was observed under the microscope. Figure 6-12 shows a sequence of four frames in which a single microsphere is tracked as it travels across the field of view of the micro-scope, indicated by the bright vertical stripe in the picture. From frame (a) through (d), the bead goes from the upper-left portion of the bright band to its lower-right corner.

(a)   (b)

(c)   (d)
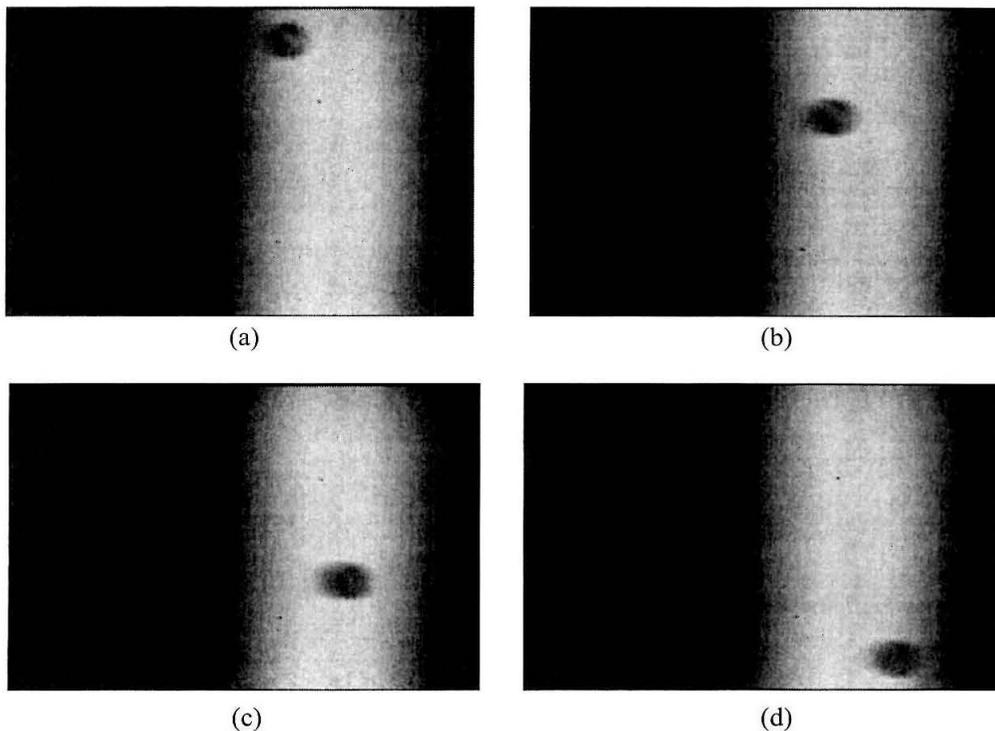
Fig. 6-12. Sequence of four frames, in which (a) is first and (d) is last, obtained with the holographic microscope using a ×40/0.65NA objective of a 15 μm microsphere as it travels across the field of view of the microscope.

Finally, imaging with multiple holograms has been considered. Five holograms were multiplexed using scheduled recording in the 5 mm thick crystal with a total recording intensity of 12 mW/cm^2. The narrow depth selectivity of the system makes it possible to multiplex depth-slices separated by just 15 μm from each other. Figure 6-13 presents the
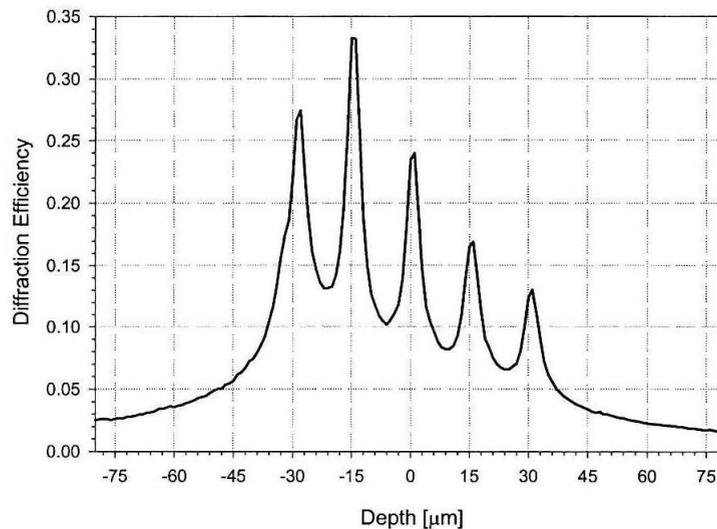
Fig. 6-13.   Scan along the z-axis showing five holograms multiplexed in a 5 mm thick crystal. The holograms are tuned at different depths, with 15 μm spacing, spanning a 60 μm range.

comb function of the five holograms as the probing point source is scanned in depth (z-axis). Each hologram is projected on the CCD camera with an angular separation of 0.5°. Again, a solution of suspended microspheres was used as specimen. Figure 6-14 shows the reconstruction of the five holograms when the specimen is illuminated with the laser beam. The holograms appear as narrow vertical stripes on the detector plane, due to the small bandwidth of the laser.

More interesting results can be obtained if broadband light is used to illuminate the specimen, since each one of the stripes in Figure 6-14 broadens and makes it possible to
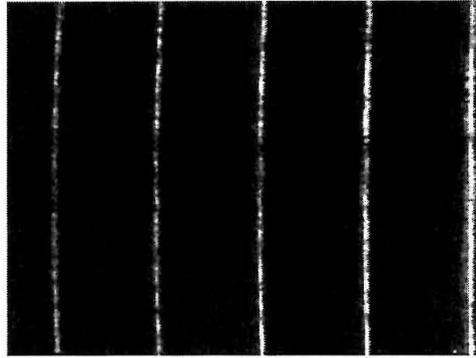
Fig. 6-14.   Simultaneous reconstruction of five depth-multiplexed holograms when a specimen is illuminated with the laser beam.

observe the microspheres at different depths without having to physically move the specimen. Figure 6-15 contains a sequence of three pictures in which a microsphere can be seen switching from one stripe to another one as it moves in the solution changing its depth with respect to the microscope objective. As the bead crosses from the left to the right hand side

(a)   (b)   (c)
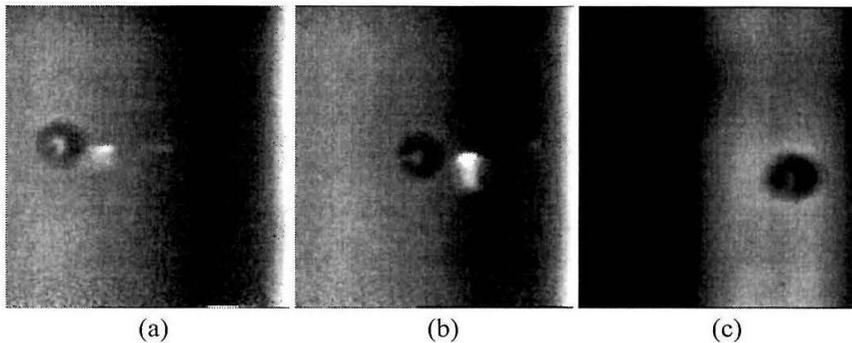
Fig. 6-15.   Sequence of three frames showing a microsphere being imaged by two different holograms (on the left and on the right sides of each frame) as the microsphere changes its depth inside the liquid solution from (a) to (c).

of the image, it also changes its depth. Frame (a) in Figure 6-15, shows a bead on the left-hand side of the image. As the bead moves towards the right and gets to the center of the image, frame (b), a second circular spot appears. This is due to the fact that the microsphere

begins to be Bragg-matched by a different hologram (i.e., a different depth-slice). By the time the microsphere reaches the right side, frame (c), the original spot has completely vanished and only the second one is visible.

6.4.3 Imaging properties of a reflection-geometry volume hologram

A symmetric reflection-geometry grating has the particularity of being degenerate both in the in-plane and out-of-plane directions. This property can be of special interest when applied to imaging, because an entire 2-D slice of the object would be Bragg-matched to the hologram and thus imaged onto the CCD. This section considers the suitability of a reflection hologram in a system like the 4-D microscope by analyzing the transformation performed by the grating to map the object space into the image plane, and studying how its angular and spectral selectivity influence the transversal and longitudinal (i.e., depth) resolution of the system.

Based on the design of the transmission-geometry microscope presented in Figure 6-7, the reflection-geometry module depicted in Figure 6-16 consists basically of the same elements. The light emitted by the specimen is collimated by lens L1 before it illuminates the hologram. The grating senses the collimated wavefront and, depending on its spatial and spectral properties, some of its components are reflected. The beam splitter is used to separate the detector plane, referred to as the (x', y')-plane, from the input plane (i.e., the x, y space); and redirects the diffracted light towards the imaging lens (L2).

6.4.3.1 Selectivity and image transformation

Consider a monochromatic point source placed at the focal distance of lens L1 in Figure 6-16 (specified by f_{mo}) and located on the optical axis of the lens (i.e., x=0, y=0).
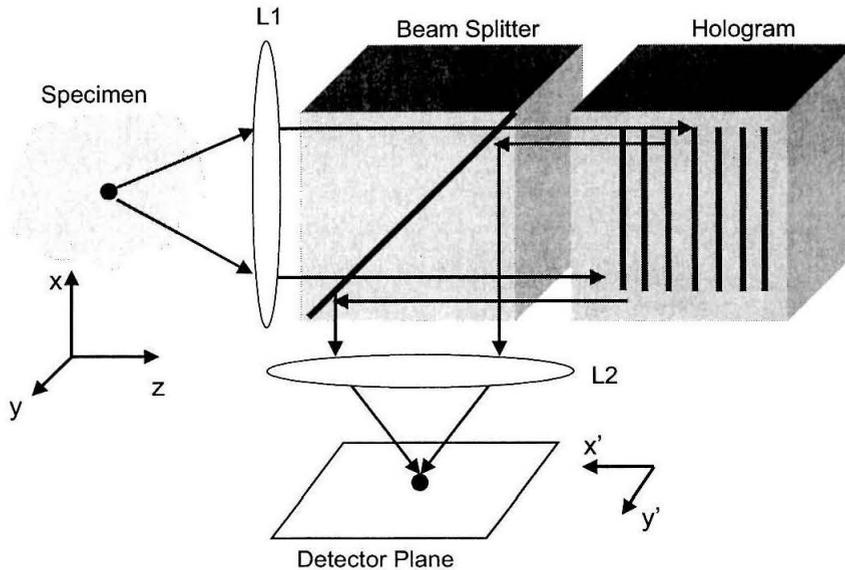
Fig. 6-16.   Schematic diagram of a holographic microscope using a reflection-geometry volume grating. The hologram performs a transformation between the object (i.e., specimen) space (x, y, z) and the image space (x', y'). L1 and L2 are the collimating and imaging lens, respectively.

The interference of its collimated wavefront with a plane wave propagating along $-\hat{z}$ records a perfect reflection-geometry grating inside the optical medium defined by the vector $\mathbf{K_G} = -2\mathrm{k} \cdot \hat{z}$.

If the point source is shifted from its original position along the x-axis by an amount Δx_p, the wavefront after the lens will still be collimated but will propagate with a certain angle $\Delta\theta_R$ with respect to the z-axis. In the paraxial approximation (i.e., $\Delta x_p \ll f_{mo}$), the mapping between position and angle performed by the collimating lens is given by

$$-\frac{\Delta x_p}{f_{mo}} = \Delta\theta_R . \qquad (6\text{-}4)$$

As sketched in Figure 6-17, the grating is read out with a slightly tilted reference beam, as a result the diffraction propagates along $-\hat{z}$ at an angle $\Delta\theta_S$. Therefore, the wave vectors of the reference ($\mathbf{K_R}$) and signal ($\mathbf{K_S}$) beams take the form:
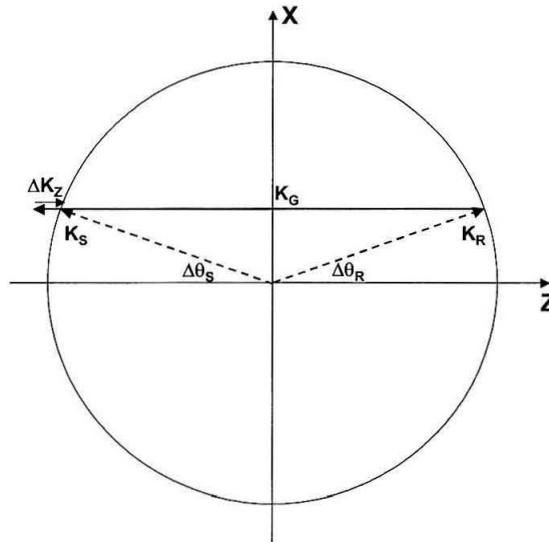
Fig. 6-17.   K-sphere diagram to illustrate the angular Bragg selectivity of a reflection grating.

$$\mathbf{K_R} = k \cdot (\cos\Delta\theta_R \cdot \hat{z} + \sin\Delta\theta_R \cdot \hat{x})$$
$$\mathbf{K_S} = k \cdot (-\cos\Delta\theta_S \cdot \hat{z} + \sin\Delta\theta_S \cdot \hat{x})$$

(6-5)

If the optical medium is assumed to have infinite transversal dimensions and finite

thickness along the z-axis (L_z), only the z-component of the diffracted beam can be Bragg-

mismatched. In other words, the following vector relation must be satisfied:

$$\mathbf{K_S} = \mathbf{K_R} + \mathbf{K_G} + \Delta K_z \cdot \hat{z}$$

(6-6)

The relation $\Delta\theta_S = \Delta\theta_R$ is obtained from the x-component of Equation 6-6, after making

the approximation $\sin(\alpha) \approx \alpha$ and $\cos(\alpha) \approx 1 - \dfrac{\alpha^2}{2}$ given that both $\Delta\theta_R$ and $\Delta\theta_S$ are

small. This result can be applied in the equation for the z-component to solve for the detun-

ing term (ΔK_z)

$$\Delta K_z = k \cdot (\Delta\theta_R)^2.$$

(6-7)

Under Born approximation, the diffraction efficiency of the hologram is related to the amount of detuning by

$$\sqrt{\eta} \sim \text{sinc}\left(\frac{L_z}{2\pi} \cdot \Delta K_z\right). \tag{6-8}$$

The angular Bragg selectivity of the hologram is defined by the first null of the sinc function, which occurs for

$$\Delta\theta_R = \sqrt{\frac{\lambda}{L_z}}. \tag{6-9}$$

Similarly to the collimating lens (L1), the imaging lens (L2 in Figure 6-16) translates the propagation angle of the diffracted beam into a position on the detector plane (i.e., the focal plane of lens L2) as defined by

$$-\frac{\Delta x'_d}{f_i} = \Delta\theta_S, \tag{6-10}$$

where f_i is the focal length of the imaging lens. Combining Equations 6-4 and 6-10, and using the result $\Delta\theta_S = \Delta\theta_R$ from Equation 6-6,

$$\Delta x'_d = -\left(\frac{f_i}{f_{mo}}\right) \cdot \Delta x_p, \tag{6-11}$$

in which the sign reversal comes from the reflection of the incident beam on the hologram. Thus, the hologram performs a linear transformation between points along the x-axis on the input space and points along the x'-axis on the detector plane.

It is worth noticing that the system presents rotational symmetry around the z-axis. Despite the fact that the expression of the angular selectivity has been derived assuming incident and diffracted beams contained in the (x, z)-plane, the orientation of the x- and

y-axes is completely irrelevant as the grating is always degenerate. Therefore, the grating

is able to perform by itself a 2-D mapping of the (x, y)-plane onto the (x' , y')-plane.
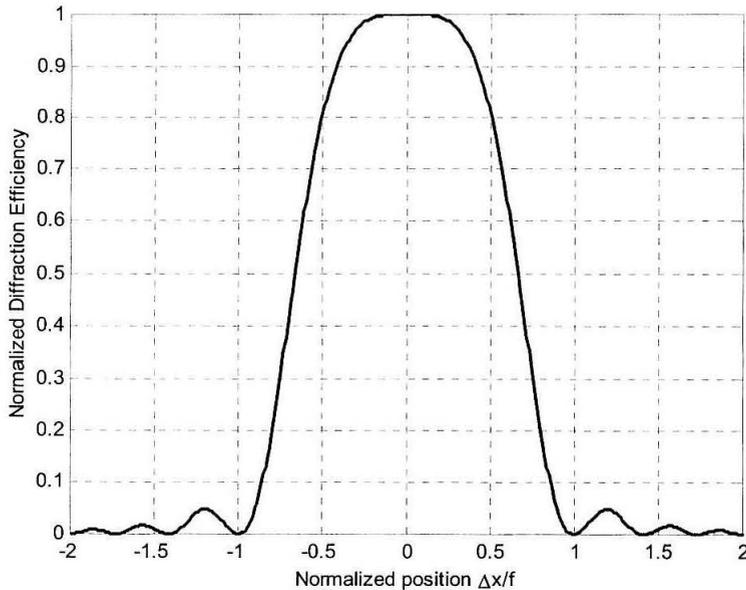
Fig. 6-18. Normalized diffraction efficiency on detector as a function of the position of a monochromatic point source as it moves along the x-axis.

The amplitude of the imaged points is affected by the sinc function of the hologram

(Equation 6-8) as plotted in Figure 6-18. Points along a line that extends beyond the angular

selectivity of the grating, will not be relayed to the detector plane. Fortunately, the square-

root dependence of $\Delta\theta_R$, due to the degeneracy of the pure reflection grating, results in

fairly wide selectivity curve. In the context of the 4-D microscope, the angular selectivity

does not determine the transversal resolution of the system (which is solely limited by the

numerical aperture of the lenses for a monochromatic point source), but the size of the cir-

cular window over which the microscope can observe the specimen. The radius of this disk

(ΔX_p) is obtained from combining Equations 6-4 and 6-9, and results in

$$\Delta X_p = f_{mo} \cdot \sqrt{\frac{\lambda}{L_z}} . \tag{6-12}$$

To illustrate the mapping performed by the hologram, a numerical simulation produces the 2-D intensity distribution of the diffracted spot on the detector plane as a monochromatic point source is shifted along the x-axis. Figure 6-19 shows four instances of a
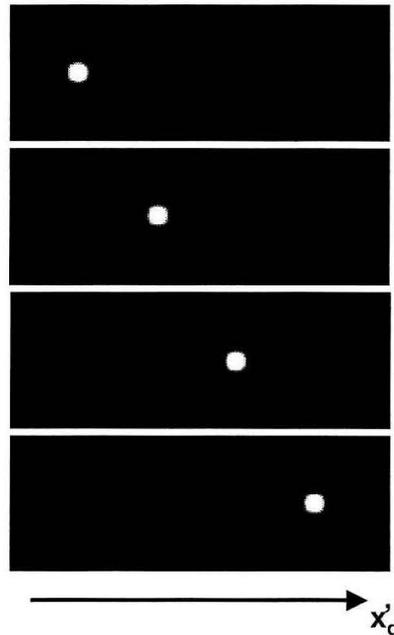
Fig. 6-19. Simulation of the intensity profile and position of the diffracted spot on the detector plane as a monochromatic point source is shifted along the x-axis in $4.5 \times \lambda$ steps. Simulation parameters: λ=514 nm, L_z=200 μm, f_{mo}=4.5 mm, f_i=4.5 mm and NA=0.65.

diffraction-limited spot on the detector plane, whose position x'_d changes as the probing point source is shifted in steps of $4.5 \times \lambda$ along the x-axis, but still well within the Bragg selectivity of the hologram.

Consider now the situation in which the readout wavelength (λ') differs from the wavelength used during recording (λ), while the point source remains in the same location
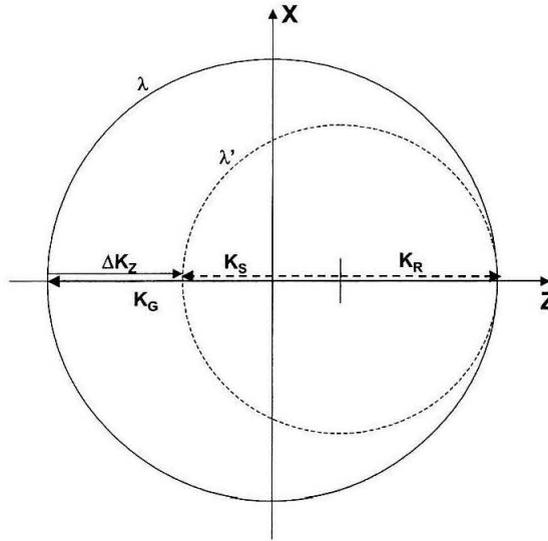
Fig. 6-20.   K-sphere diagram to illustrate the wavelength Bragg selectivity of a reflection grating.

as the one during recording. In this case, depicted in Figure 6-20, the wave vectors of the

incident and diffracted beams are given by

$$
\begin{aligned}
\mathbf{K_R} &= k' \cdot \hat{z} \\
\mathbf{K_S} &= -k' \cdot \hat{z}
\end{aligned}
\tag{6-13}
$$

so in order to satisfy Equation 6-6, the detuning term must be $\Delta K_z = -2 \cdot (k' - k)$. Assum-

ing $k' = k - \Delta k = k \cdot \left(1 + \dfrac{\Delta\lambda}{\lambda}\right)$, the expression in Equation 6-8 hits its first null when

$$
\frac{\Delta\lambda}{\lambda} = \frac{\lambda}{2L_z},
\tag{6-14}
$$

which is known as the wavelength selectivity of the hologram. Unlike with its angular

behavior, the hologram is very sensitive to changes in wavelength. Strong diffraction will

occur only when the wavelength used for readout is well within the Bragg selectivity of the

hologram.

The numerical simulation in Figure 6-21 shows how the intensity of the diffracted beam profile decreases very rapidly as the wavelength of the probing monochromatic point source is shifted from the Bragg condition in steps of 20% of the value of the wavelength selectivity (Equation 6-14).
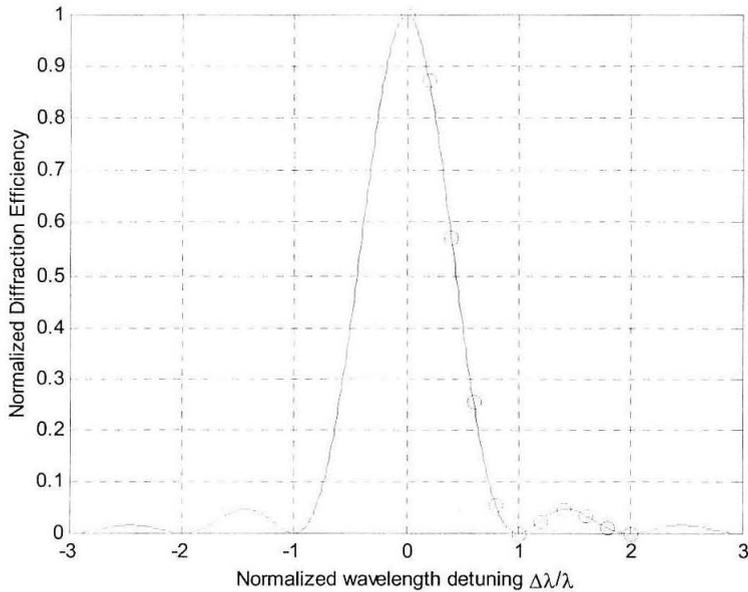
(a)
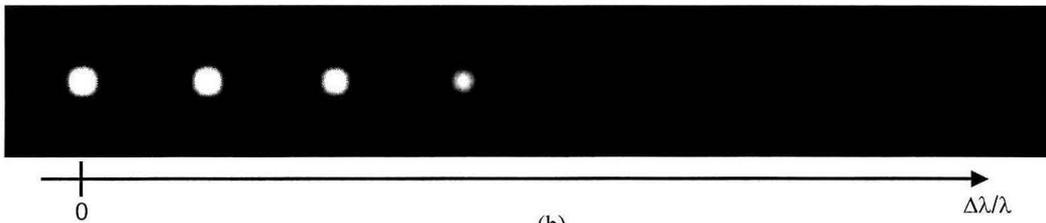
(b)

Fig. 6-21.   (a) Normalized diffraction efficiency on detector as a function of the wavelength detuning for a monochromatic point source. (b) Simulation of the intensity profile of the diffracted spot on the detector plane as a monochromatic point source changes its emission wavelength in $0.2 \times \Delta\lambda$ steps. The dots in (a) correspond to the points at which wavelength detuning has been simulated. Simulation parameters: λ=514 nm, L_z=200 µm, f_{mo}=4.5 mm, f_i=4.5 mm and NA=0.65.

To understand how the wavelength selectivity affects the transversal resolution of the microscope, it is necessary to consider the more general case of a slanted (i.e., asym-

metric) reflection grating. Suppose two counter-propagating beams with arbitrary angles θ_R and θ_S with respect to the z-axis, their interference results in a grating vector

$$\mathbf{K_G} = k \cdot [(\sin\theta_S - \sin\theta_R) \cdot \hat{x} - (\cos\theta_S + \cos\theta_R) \cdot \hat{z}]. \qquad (6\text{-}15)$$

Upon readout, the hologram is probed with a beam ($\mathbf{K_R}$) incident at the same angle θ_R but with different wavelength (λ'). Due to the change in wavelength, and in order to verify Equation 6-6, the diffracted beam ($\mathbf{K_S}$) may have to deviate from its initial angle θ_S to a new $\theta_S + \Delta\theta_S$. Given the expressions of the wave vectors of the incident and diffracted beams

$$\begin{aligned} \mathbf{K_R} &= k' \cdot [\cos\theta_R \cdot \hat{z} + \sin\theta_R \cdot \hat{x}] \\ \mathbf{K_S} &= k' \cdot [-\cos(\theta_S + \Delta\theta_S) \cdot \hat{z} + \sin(\theta_S + \Delta\theta_S) \cdot \hat{x}] \end{aligned}, \qquad (6\text{-}16)$$

and following a similar derivation as before, it can be easily verified that the wavelength selectivity of the hologram is given by

$$\frac{\Delta\lambda}{\lambda} = \frac{\lambda}{L_z \cdot [1 + \cos(\theta_S + \theta_R)]}, \qquad (6\text{-}17)$$

which reduces to Equation 6-14 when $\theta_S = \theta_R = 0$. Furthermore, the deflection angle of the diffracted beam is related to the shift in wavelength as

$$\Delta\theta_S = -\frac{\sin\theta_R - \sin\theta_S}{\cos\theta_S} \cdot \left(\frac{\Delta\lambda}{\lambda}\right). \qquad (6\text{-}18)$$

Finally, the imaging lens converts the change in angle $\Delta\theta_S$ into a shift in position on the detector plane. The point source that used to be imaged at x'_d for wavelength λ, will be imaged at $x'_d + \Delta x'_d$ for λ', where $\Delta x'_d$ takes the form

$$\Delta x'_d = f_i \cdot [\sin\theta_R - \sin\theta_S] \cdot \left(\frac{\Delta\lambda}{\lambda}\right). \qquad (6\text{-}19)$$

This effect has been modeled into the numerical simulation. A slanted grating with angles θ_R=0 and θ_S=5° is read out by the collimated beam of a monochromatic point source. The sequence of pictures in Figure 6-22 shows the shift along the x'-axis (vertical axis in Figure 6-22) as the wavelength of the probing point source changes in steps of 20% of the spectral selectivity around the Bragg wavelength.
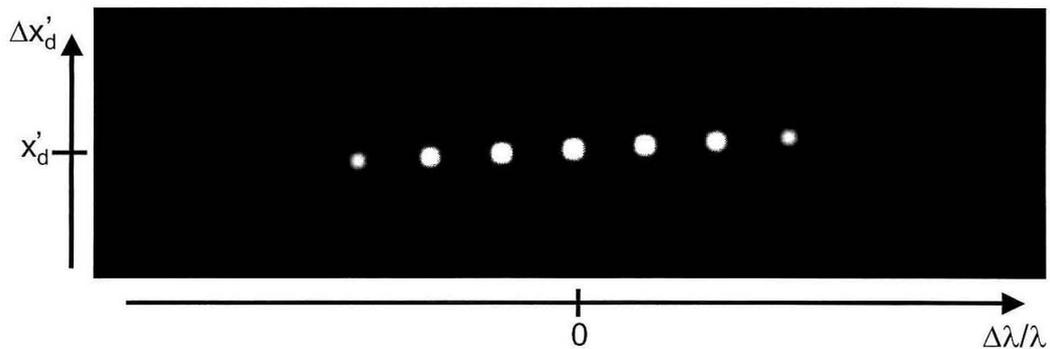
Fig. 6-22.   Simulation of the intensity profile and position of the diffracted spot on the detector plane (along the x'-axis) for a slanted grating as a monochromatic point source changes its emission wavelength in $0.2 \times \Delta\lambda$ steps. Simulation parameters: θ_R=0°, θ_S=5°, λ=514 nm, L_z=200 μm, f_{mo}=4.5 mm, f_i=4.5 mm and NA=0.65.

Equation 6-19 is especially important when the point source has a bandwidth, as each spectral component within the selectivity of the hologram will be mapped into a different position on the detector plane, resulting in some sort of rainbow and thus blurring the image of the specimen. Figure 6-23 compares the simulated diffracted spot obtained for a monochromatic source of wavelength λ and the one for a point source with bandwidth from λ−$1.2 \times \Delta\lambda$ to λ+$1.2 \times \Delta\lambda$, shown in Figure 6-23(a) and (b), respectively. The dispersive behavior of the grating degrades the resolution of the microscope, as it can be appreciated in the widening of the diffracted spot in the cross-sectional plot in Figure 6-23(c).

(a)   (b)

(c)
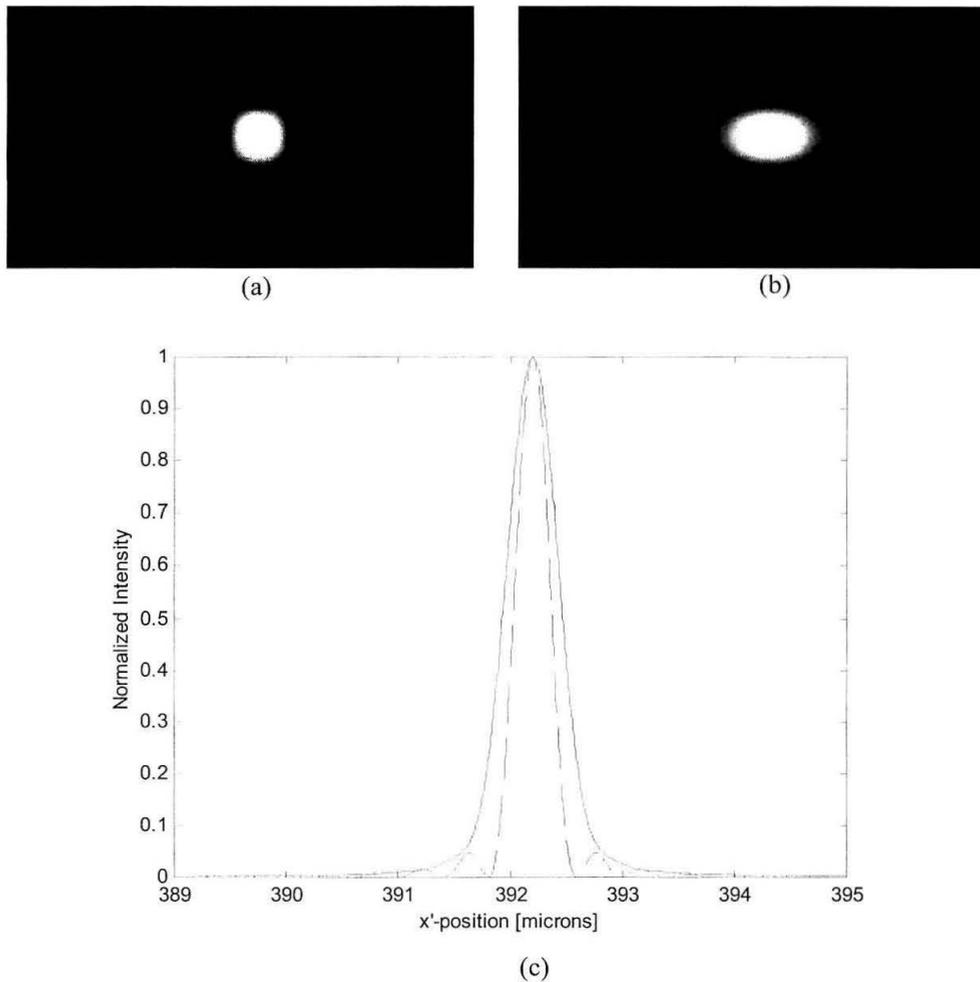
Fig. 6-23. Simulation of the intensity profile of the diffracted spot on the detector plane produced by a slanted grating for (a) a monochromatic point source with emission wavelength λ and (b) a chromatic point source with bandwidth $\lambda \pm 1.2 \times \Delta\lambda$. (c) Comparison of the intensity of the diffracted spot for a cross-sectional cut along the x'-axis for the monochromatic (dashed line) and the chromatic (solid line) point sources. Simulation parameters: $\theta_R=0°$, $\theta_S=5°$, $\lambda=514$ nm, $L_z=200$ μm, $f_{mo}=4.5$ mm, $f_i=4.5$ mm and NA=0.65.

Therefore, the wavelength selectivity plays a key role in defining the transversal resolution of the holographic microscope. The best resolution can be achieved using a symmetric grating, because $\Delta x'_d = 0$ when $\theta_S=\theta_R$. In this particular case the transversal resolution is only limited by the numerical aperture of the lenses in the system. However, the

need to multiplex several holograms to image various color bands and depths on different

regions of the detector plane, makes necessary to introduce some slant in the gratings.

6.4.3.2 Spatial and wavelength coupling

The previous section showed that the hologram is selective in wavelength and capa-

ble of imaging points in a neighborhood of the z-axis (i.e., point x=0, y=0). This section

will consider the coupling between position and color that, like in the transmission-geom-

etry microscope, permits to extend the field of view of the microscope beyond the spatial

selectivity of the hologram; and will also study the dependence of the spatial and wave-

length selectivity for an arbitrary point on the (x, y)-plane.

To analyze the coupling between position and wavelength, consider again a sym-

metric reflection grating. The detuning term is given by

$$\Delta K_z = -2k' \cdot \cos\theta + 2k, \tag{6-20}$$

where θ is measured with respect to the z-axis. Coupling is achieved when the hologram is

Bragg-matched (i.e., $\Delta K_z = 0$). Let $r = \sqrt{x^2 + y^2}$ be the distance of point (x, y) to the z-axis,

then $\tan\theta = -\dfrac{r}{f_{mo}}$. In the Bragg-matched case, Equation 6-20 can be re-written as

$$r^2 = f_{mo}^2 \cdot \left[\left(\frac{k'}{k}\right)^2 - 1\right], \tag{6-21}$$

which means that the points on a circle of radius r around the z-axis can be imaged onto the

detector plane if they emit light at λ'. Figure 6-24 indicates the relative change in wave-

length required to satisfy Equation 6-21. Therefore, a full 2-D image can be obtained on the

detector by sectioning the input plane into monochromatic concentric circles when the

specimen is illuminated with an extended broadband source. It is worth noticing in Equa-
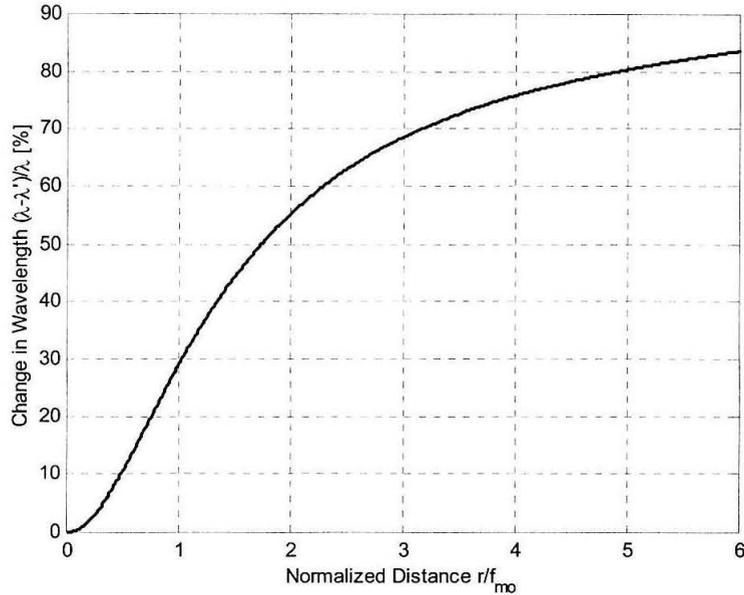
Fig. 6-24.   Relative change with respect to λ of the emission wavelength of a monochromatic point source required to satisfy the Bragg condition as a function of the distance of the point source from the z-axis.

tion 6-21 that coupling is only possible for $\lambda' < \lambda$, which is a consequence of $\mathbf{K_G}$ being a 180° grating for wavelength λ.

Consider first the case of angular detuning. Assume $\theta = \theta_o + \Delta\theta$, with θ_o being the Bragg angle. After cancelling out the Bragg-matched terms, Equation 6-20 becomes

$$\Delta K_z = k' \cdot \cos\theta_o \cdot \Delta\theta^2 + 2k' \cdot \sin\theta_o \cdot \Delta\theta \; . \tag{6-22}$$

Using the relation $\Delta\theta = -\dfrac{\Delta r}{f_{mo}}$, valid under the paraxial approximation, the radius of spatial selectivity (Δr) for an arbitrary point located at a distance r from the z-axis is given by

$$\left(\frac{\Delta r}{f_{mo}}\right)^2 + 2\left(\frac{\Delta r}{f_{mo}}\right) \cdot \frac{r}{f_{mo}} = \frac{\lambda}{L_z} \; . \tag{6-23}$$

The solution to Equation 6-23 is plotted in Figure 6-25, assuming a medium of thickness L_z=390 × λ. The more off-axis a point is located on the (x, y)-plane, the smaller its spatial
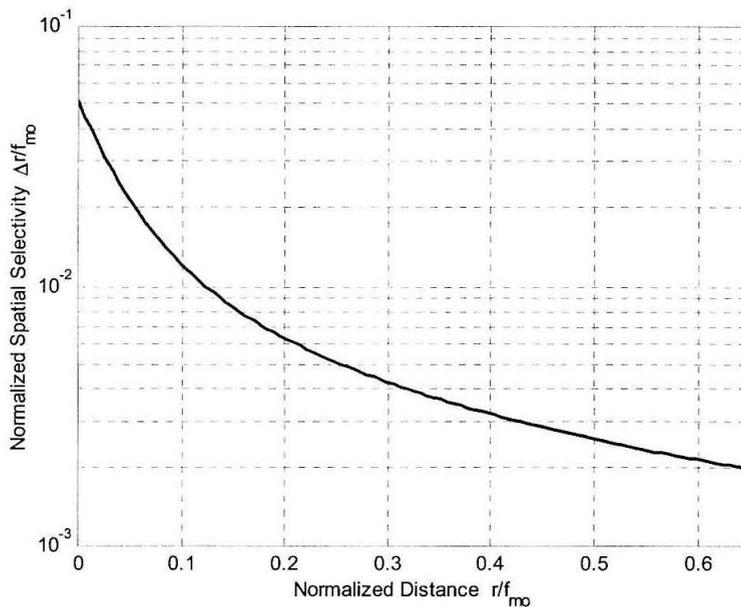
Fig. 6-25. Radius of the disk of spatial selectivity (Δr) for a monochromatic point source as a function of the distance of the source from the z-axis, assuming a material of thickness $L_z = 390 \times \lambda$.

selectivity, and consequently the smaller the size of the window that can be imaged by the 4-D microscope.

Similarly, the wavelength detuning can be analyzed assuming $k' = k'_o + \Delta k'$, with Bragg match occurring for $\Delta k' = 0$, in Equation 6-20. After simplifying the Bragg-matched terms, the spectral selectivity for a point whose Bragg wavelength λ'_o results in

$$\frac{\Delta \lambda'}{\lambda'_o} = \frac{\lambda}{2 L_z}. \tag{6-24}$$

Equation 6-24 says that the relative wavelength selectivity of the hologram is constant over the entire (x, y)-plane. Plugging Equation 6-21 into 6-24, the spectral selectivity can be explicitly related to position as

$$\Delta\lambda' = \frac{\lambda^2}{2L_z} \cdot \frac{1}{1 + \left(\dfrac{r}{f_{mo}}\right)^2}. \tag{6-25}$$

Thus, the farther the point source is from the z-axis, the narrower the portion of its spectrum that can be Bragg-matched by the hologram.

6.4.3.3 Depth selectivity

As the point source shifts along the z-axis, the wavefront after the collimating lens (L1) is no longer a plane wave. As the probing beam changes its sphericity, the hologram gradually becomes Bragg-mismatched. A simple analytical approximation can be obtained by decomposing the spherical wave into its spatial frequency components. Each component is a plane wave propagating with angle θ, for which the response of the hologram can be easily calculated. The diffraction efficiency for a monochromatic point source shifted by Δz_p from the focal plane of lens L1 is approximately

$$\eta(\Delta z_p) \cong \frac{1}{\theta_{max}} \int_0^{\theta_{max}} sinc\left(\frac{L_z}{2\pi} \cdot [2k - 2k' \cdot \cos\theta]\right)^2 \cdot d\theta, \tag{6-26}$$

where the range of transverse spatial components is determined by the amount of defocusing of the point source and the numerical aperture of the lens L1 and results in

$$\theta_{max} = NA \cdot \frac{\Delta z_p}{f_{mo}}. \tag{6-27}$$

Equation 6-26 is solved numerically and plotted in Figure 6-26 for the case L_z=200 μm, λ=514 nm and f_{mo}=4.5 mm. The wide angular selectivity of the reflection hologram results, as seen in Figure 6-26, in a poor performance to discriminate point sources located at different depths. Therefore, the hologram does not improve the depth of focus of the microscope, which is limited by the numerical aperture of the lenses. On the contrary, it increases
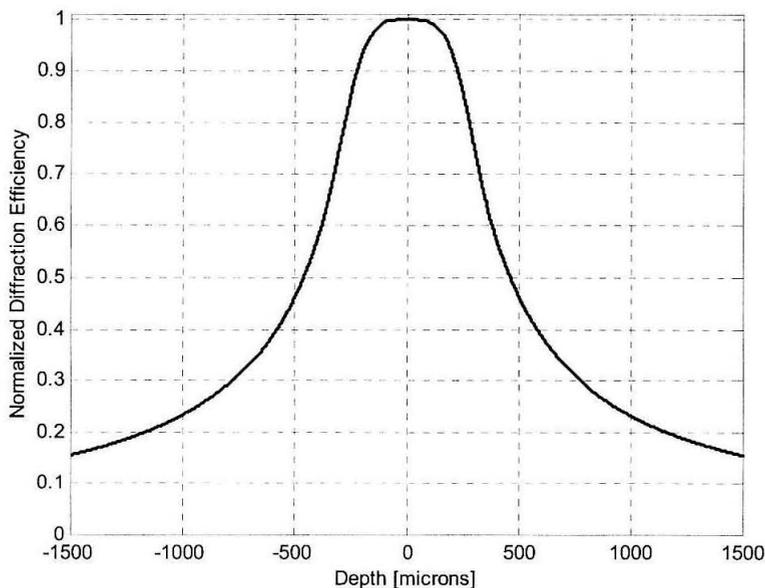
Fig. 6-26.   Normalized diffraction efficiency as a function of the position along the z-axis of a monochromatic point source. Simulation parameters: $\lambda=514$ nm, $L_z=200$ μm, $f_{mo}=4.5$ mm, $f_i=4.5$ mm and NA=0.65.

the level of background noise as the hologram cannot filter out light originating at different depths.

Finally, it is interesting to point out the coupling between depth and wavelength for a chromatic point source. Figure 6-27 presents the simulation results of the depth scan along the z-axis for a point source as its wavelength of emission is changed. As the figure indicates, the hologram can be partially Bragg-matched at shorter wavelengths if the point source is shifted away from the focal plane of the collimating lens. This happens because some of the components of the defocused probing beam can better satisfy the Bragg condition. The figure exhibits an asymmetric behavior in wavelength. This is due to the fact that coupling between depth and wavelength is not possible for wavelengths longer than the

Bragg wavelength, which is consistent with the result from Equation 6-21 and, as mentioned earlier, is just the consequence of $\mathbf{K_G}$ being a 180° grating for wavelength λ.
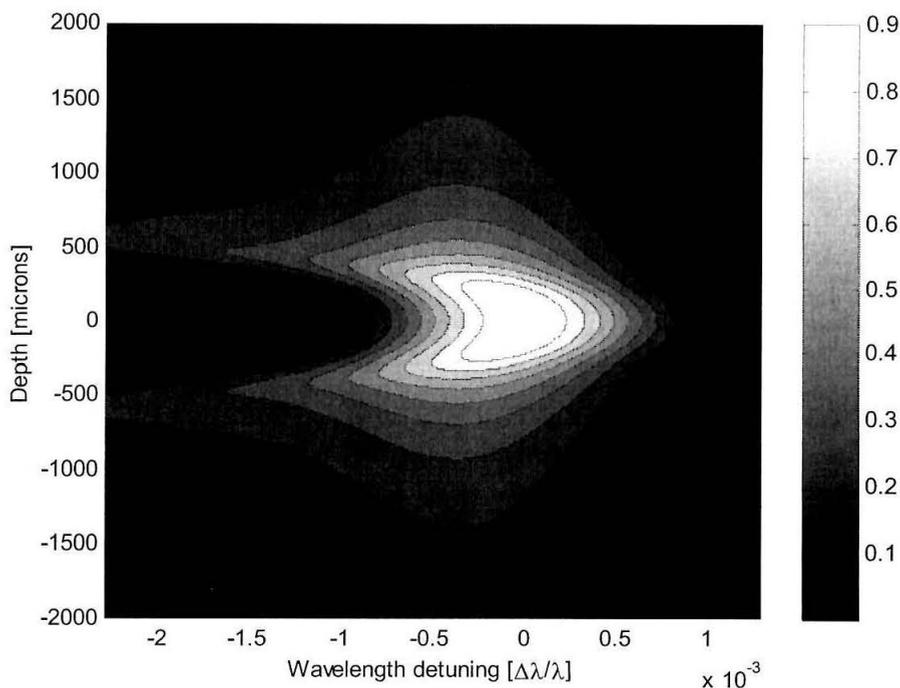
Fig. 6-27. Normalized diffraction efficiency of the hologram (color-coded in grayscale) as a function of the position along the z-axis of a monochromatic point source (vertical axis in the figure) and detuning of its wavelength of emission (horizontal axis).

6.4.4 Discussion

Both the results from experiments done with the transmission-geometry microscope and the theoretical analysis of the reflection-geometry module reveal that the resolution of the system is in large measure determined by the quality of the collimating and imaging lenses. In this sense, the holographic microscope does not outperform other more conventional solutions used in microscopy. Nonetheless, the 4-D microscope can show its superior performance over competing technologies in the arena of applications that require real-time visualization of the specimen. The 4-D microscope is better suited for those situations in

which it is paramount to gather multispectral or multidepth information of the specimen. While its competitors need to trade in resolution in order to gain speed, the holographic microscope can keep the same level of resolution while offering high speed.

The parallel processing ability of the 4-D microscope, in fact, inherent to most holographic systems, derives from the capability of multiplexing various holograms in the optical medium. However, this parallelism comes at the expense of dividing the photon count on the microscope. Therefore, to increase the light efficiency of the system, it is necessary to act either on the holographic material to improve its performance (e.g., larger M/#), or on the hologram to widen its spectral response (e.g., chirping the hologram), or even on the specimen itself by using more efficient and narrow-band fluorescent markers (e.g., quantum dots) [6-20], [6-21].

Finally, the wavelength selectivity of the hologram poses a trade-off between light collection efficiency and resolution. A broad wavelength selectivity benefits the photon count in the microscope, but at the same time degrades its transversal resolution. Fortunately, the reflection geometry can be of assistance in breaking this trade-off as for this geometry the resolution of the system is less sensitive to the bandwidth of the hologram.

References

[6-1] J. R. Wolpaw, N. Nirbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Transactions of Rehabilitation Engineering* **8**, 164-173 (2000).

[6-2] M. Barinaga, "Turning thoughts into actions," *Science* **286**, 888-890 (1999).

[6-3]   E. B. Fetz, "Real-time control of a robotic arm by neuronal ensembles," *Nature Neuroscience* **2**, 583-584 (1999).

[6-4]   S. Trimberger, D. Carberry, A. Johnson, and J. Wong, "A time-multiplexed FPGA," *Proceedings of the IEEE Symposium on FPGA-Based Custom Computing Machines*, 34-40, IEEE Computer Society, Los Alamitos, 1997.

[6-5]   Product information can be obtained from http://www.xilinx.com/.

[6-6]   M. Motomura, Y. Aimoto, A. Shibayama, Y. Yabe, and M. Yamashina, "An embedded DRAM-FPGA chip with instantaneous logic reconfiguration," *Proceedings of the IEEE Symposium on FPGAs for Custom Computing Machines*, 264-266, IEEE Computer Society, Los Alamitos, 1998.

[6-7]   R. T. Lauer, P. H. Peckham, K. L. Kilgore, and W. J. Heetderks, "Applications of cortical signals to neuroprosthetic control: A critical review," *IEEE Transactions on Rehabilitation Engineering* **8**, 205-208 (2000).

[6-8]   J. R. Chapin, K. A. Moxon, R. S. Markowitz, and M. A. L. Nicolelis, "Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex," *Nature Neuroscience* **2**, 664-670 (1999).

[6-9]   R. E. Isaacs, D. J. Weber, and A. B. Schwartz, "Work toward real-time control of a cortical neural prosthesis," *IEEE Transactions of Rehabilitation Engineering* **8**, 196-198 (2000).

[6-10] K. V. Shenoy, S. A. Kureshi, D. Meeker, B. L. Gillikin, D. J. Dubowitz, A. P. Batista, C. A. Buneo, S. Cao, J. W. Burdick, and R. A. Andersen, "Toward prosthetic systems controlled by parietal cortex," *Society for Neuroscience* **25** (1999).

[6-11] A. L. Benabid, P. Pollak, D. Gao, D. Hoffmann, P. Limousin, E. Gay, I. Payen, and A. Benazzouz, "Chronic electrical stimulation of the ventralis intermedius nucleus of the thalamus as a treatment of movement disorders," *Journal of Neurosurgery* **84**, 203-214 (1996).

[6-12] G. E. Loeb, and F. J. R. Richmond, "BION^{TM} implants for therapeutic and functional electrical stimulation," *Neural Prostheses for Restoration of Sensor and Motor Function*, CRC Press, Boca Raton, 2000.

[6-13] B. C. Wheeler, "Automatic discrimination of single units," *Methods for Neural Ensemble Recordings*, CRC Press, Boca Raton, 61-77, 1999.

[6-14] R. Andraka, and A. Berkun, "FPGAs make a radar signal processor on a chip a reality," *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, Monterrey, 1999.

[6-15] G. Barbastathis, M. Balberg, and D. J. Brady, "Confocal microscopy with a volume holographic filter," *Optics Letters* **24**, 811-813 (1999).

[6-16] G. G. Yang, H. S. Chen, and E. N. Leith, "Volume reflection holographic confocal imaging," *Applied Optics* **39**, 4076-4079 (2000).

[6-17] G. Barbastathis, and D. J. Brady, "Multidimensional tomographic imaging using volume holography," *Proceedings of the IEEE* **87**, 2098-2120 (1999).

[6-18] D. W. Piston, "Imaging living cells and tissues by two-photon excitation microscopy," *Trends in Cell Biology* **9**, 66-69 (1999).

[6-19] A. Diaspro, and M. Robello, "Two-photon excitation of fluorescence for three-dimensional optical imaging of biological structures," *Journal of Photochemistry and Photobiology B-Biology* **55**, 1-8 (2000).

[6-20] M. Bruchez, M. Moronne, P. Gin, S. Weiss, and A. P. Alivisatos, "Semiconductor nanocrystals as fluorescent biological labels," *Science* **281**, 2013-2016 (1998).

[6-21] M. Han, X. Gao, J. Z. Su, and S. Nie, "Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules," *Nature Biotechnology* **19**, 631 - 635 (2001).