Compact Holographic Read/Write Memory

Wenhai Liu and Demetri Psaltis
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125
Phone: +1-626-395-4843
Fax: +1-626-568-8437
{ wliu, psaltis }@sunoptics.caltech.edu

Abstract— We examine the primary challenges for building a practical and competitive holographic random access memory (HRAM) system, specifically size, speed, and cost. We show that a fast HRAM system can be implemented with a compact architecture by incorporating conjugate readout, a smart-pixel array, and a linear array of laser diodes. It provides faster random access time than hard disk (100 microseconds or less) and similar bandwidth as silicon storage with lower cost. Preliminary experimental results support the feasibility of this architecture. Our analysis shows that in order for the HRAM to become competitive, the principal tasks will be to reduce spatial light modulator (SLM) and detector pixel sizes to 1 μm, increase the output power of compact visible-wavelength lasers to several hundred milliwatts, and develop ways to raise the sensitivity of holographic media to the order of 1 cm/J.

1. Introduction

Holography memory is a potential technology that can provide very large storage density and high speed. The theoretical storage capacity of this technology is on the order of V/λ^3 [1] (where V is the volume of the holographic medium and λ is the wavelength of light), or equivalently, a storage density limit of about one bit per cubic wavelength. Furthermore, holography has the inherent advantage of massive parallelism. Unlike conventional storage media such as magnetic hard disks and CD-ROMs, which access only one bit at a time, each access of a holographic memory yields an entire data page -- potentially megabits at a time.

Figure 1 shows a typical angle-multiplexed holographic memory in the 90^0 geometry. Information is recorded in the holographic medium through the interference of two coherent beams of light. The information-carrying signal beam and the interfering reference beam cause an index grating (the hologram) to be written in the material through the electro-optic effect. If the hologram is subsequently illuminated with one of the original writing beams, light is diffracted from the grating in such a way that the second beam is reproduced.

Due to Bragg effects, many holograms can be multiplexed within the same volume of material by slightly changing the angle of the reference beam with each new data page. Thousands of holograms can be multiplexed this way in a small volume of crystal, offering the potential of very high storage densities.

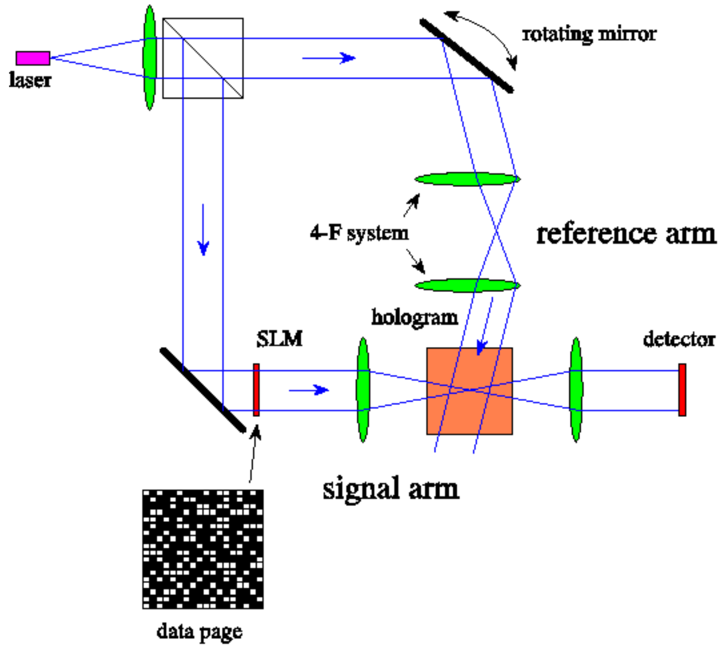


Figure 1. Typical angle-multiplexed holographic memory.

Silicon (1×1 cm^2)	$125
$LiNbO_3$ (1×1×1 cm^3)	$ 10
Liquid Crystal	$ 5
Beamsplitters and lens	$ 6
LD array (500)	$ 25 - 100
Total	$ 171 - 246

Table 1. Estimated cost of components in the holographic memory module, assuming production in large quantities

In the figure shown, the signal path consists of a spatial light modulator (SLM) and detector array with a 4-F imaging system between them, and the reference path uses another 4-F lens system in combination with a rotating mirror to provide the angular tilt to the reference beam. Recent work has shown the ability to store and retrieve many thousands of holograms [2,3]. Much of the progress that has been made can be attributed to advancements in our understanding of ways to take advantage of the Bragg selectivity of 3-D recording to multiplex holograms, as well as continued research in holographic material properties and dynamics.

In this paper, we describe a holographic random access memory (HRAM) with phase conjugate reconstruction and present experimental results from this architecture. It has the potential of faster random access time than hard disk (100 microseconds or less) and similar bandwidth as silicon storage with lower cost. The phase conjugation leads to high-resolution signal image recovery with a compact and inexpensive optical system. And we believe that HRAM can be a competitive memory technology if optoelectronics technology can achieve the following three milestones in the next few years:

1). small SLM and detector pixel sizes on the order of 1 μm;

2). high recording sensitivity of the holographic material with no more than 1 J/cm^2 to reach saturation;

3). inexpensive high spatial density laser diodes with at least 500 mW of output power in the near-infrared or visible wavelength.

2. Conjugate Readout Method

Despite the high theoretical limit on the storage density of volume holographic storage (one bit per cubic wavelength of material), the practical implementation of holographic systems is often bulky due to the large space occupied by the various components that are necessary to provide the recording and readout mechanisms for the crystal. The system of Figure 1 is fairly simple with a relatively small number of components, however the spacing requirements of the imaging lenses imposes constraints on how closely these components can be placed. For example, assuming SLM and detector array dimensions of 1cm and high quality lenses with F/#=1, the focal distance between the arrays, lenses, and crystal must also be at least 1cm. The system of Figure 1 would then occupy a volume of approximately 6cmx5cmx1cm, which is 30 times larger than the volume of the recording material.

The reason we normally need to place lenses within the signal path is to undo the effects of diffraction. When we record a hologram of the signal beam diverging from the input SLM and reconstruct it with the original reference beam, we produce a virtual image of the input data page and thus require a lens to refocus it onto the detector array. We can eliminate the lens system between the SLM and detector array if we reconstruct a real image instead of a virtual one. One way to do this is to use phase conjugate readout [4-6] as illustrated in Figure 2. Using this method, a hologram is recorded in the usual manner between the signal and reference beams, but the hologram is read out with the phase conjugate of the reference beam, propagating in the opposite direction as the one used for recording. This causes the signal reconstruction from the hologram to propagate back along the direction from which it originally came, reversing the original signal diffraction, and refocusing exactly at the plane of the SLM array. To generate the conjugate reference we may use a phase-conjugate mirror [5], or in the case of a planar reference beam, we may simply use a counter-propagating plane wave at the each angle.

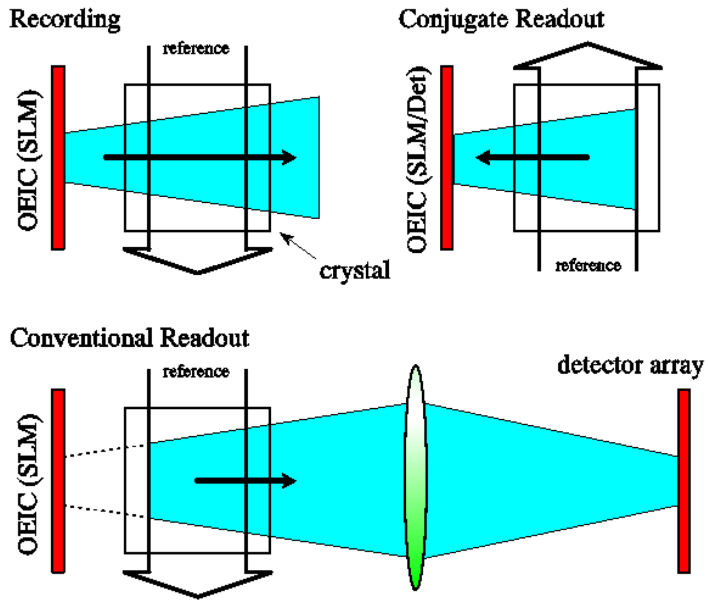


Figure 2. Comparison of phase conjugate readout method with conventional readout using imaging lenses.

Experimentally, we compared the reconstructed image fidelity that can be obtained with conventional reconstruction using high-quality, custom-designed lenses to the image fidelity we get with the conjugate readout method of planar reference beams. An SLM and detector array each with pixel spacing of 24μm were used for these tests, allowing one-to-one matching of the SLM and detector pixels. Both methods yielded SNR (signal-to-noise ratio) values ranging from about 3.8 to 4.5, verifying that the conjugate readout method produces results that can only be achieved with quality lenses, while using a much more compact and inexpensive optical system.

Phase conjugation read-out not only eliminates the lenses and associated path lengths that are normally required in the signal path, it also provides a possibility to record and reconstruction signal beams with high spatial frequencies. The holographic recording and reconstruction possesses a basic spatial frequency bandwidth for the holograms, which limits the smallest feature size to be record and reconstructed. The theoretical calculation and experimental measurements indicates a width bandwidth for holographic recording and reconstruction in the photorefractive $LiNbO_3$. This makes it possible to record and reconstruction holograms with very small pixel sizes, which has important effects on the system storage density and cost efficiency as discussed in section 3. Figure 3 shows the theoretical simulation of the holographic recording and reconstruction bandwidth inside a $LiNbO_3$ with 90^0 geometry, with the consideration of the interface losses. The hologram strength is a function of the spatial frequency, or the incident angle of the signal beams due to the different grating period, orientation and interference modulation depth [7,8]. The experimental measurement of the bandwidth confirms the theoretical prediction as shown in figure 3. Holograms with sub-micron pixels were recorded and conjugate

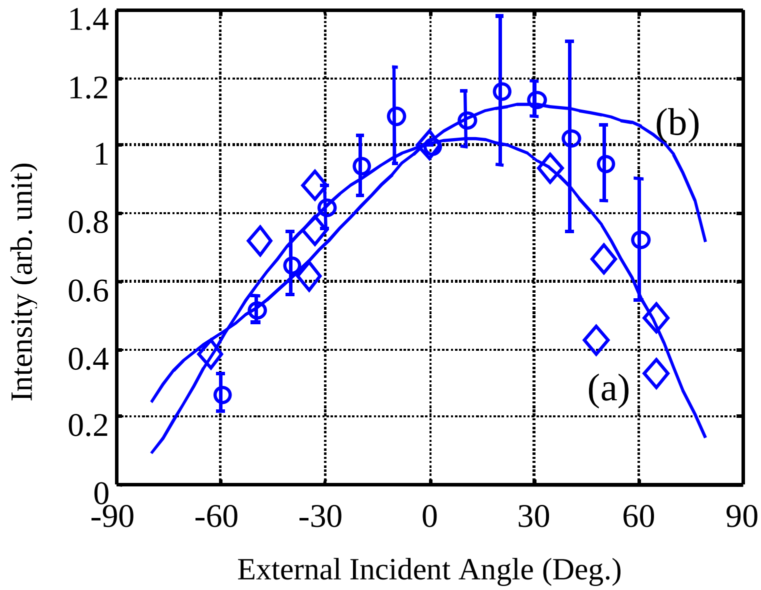
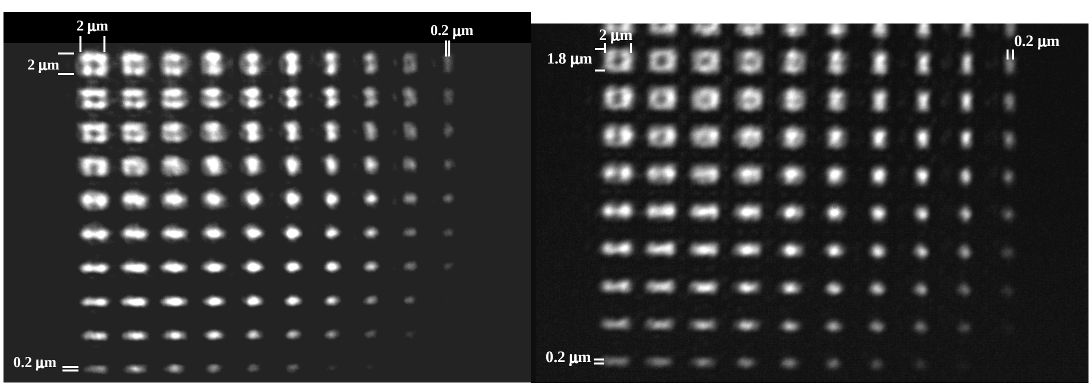


Figure 3. (a) The experimental data (diamond) and the theoretical calculation of holographic efficiency in the signal reference plane; (b) the experimental data (circle) and the theoretical calculation of the holographic efficiency out of signal reference plane.

reconstructed, which further demonstrated the resolving power of the phase conjugate reconstruction. Figure 4 shows the mask image and the phase conjugation. There are no image degradation detected for the hologram reconstruction from the direct image of the mask.



(a) (b)

Figure 4. (a) The direct image of a resolution photo mask with pixels from 2x2 μm^2 down to 0.2x0.2 μm^2. (b) The holographic phase conjugate reconstruction of the photo mask. Both images were magnified by a Nikon objective lens with NA=0.65.

3. Compact Fast Access Architecture

While conjugate readout eliminates the lenses in the signal path of the memory system, we still require a compact design to rapidly deflect the reference beam for multiplexing purposes. The 4-F system shown in Figure 1, while reliable, is bulky and slow due to the limited mechanical speed of the rotating mirror.

With the recent development of compact laser emitters, such as laser diodes and Vertical-Cavity Surface-Emitting Laser (VCSEL) devices [9,10], it has become feasible to consider the possibility of incorporating arrays of hundreds of microscopic laser sources in a holographic memory. We can then design a system in which each angle multiplexed hologram is addressed by a dedicated laser source. This architecture is shown in Figure 5. A Fourier transforming lens is used to convert the spatial shifts between the laser elements into angularly offset plane waves incident on the crystal. In this implementation, the time it takes to produce the proper read-out reference beam is determined by the switching time of the laser sources, which is in the nanosecond regime. Using a 1cm-thick crystal and a wavelength of 630nm, the first null of the angular selectivity function occurs at an angular spacing of 0.0036^{o}. Using a lens with a focal length of 2cm would require the laser elements to be placed only 1.3 μm apart to produce this angular separation. In practice, we would separate them by 10 μm or more in order to reduce interpage crosstalk while also making the array easier to fabricate.

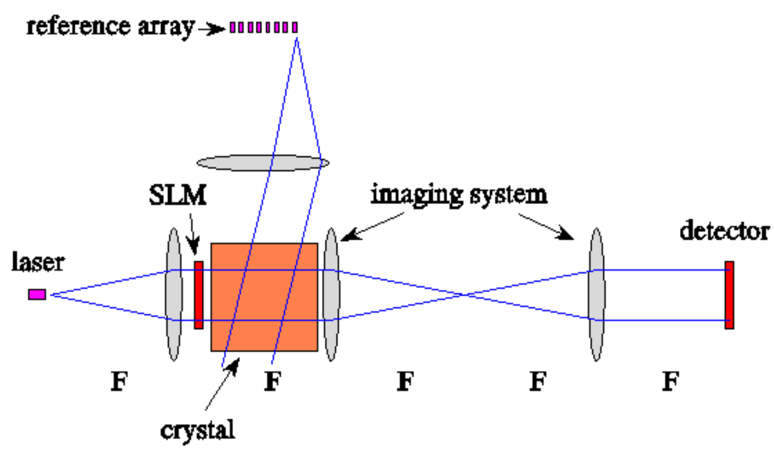


Figure 5. Use of a laser array in the reference arm of an angle multiplexed memory for fast page access.

This approach is also compatible with the conjugate readout method as shown in Figure 6. With a properly aligned laser array and a mirror placed on the opposite face of the crystal such that it lies at the focus of the Fourier transforming lens, the proper conjugate beam can be generated with the symmetrically opposite laser source. A beamsplitter must also be introduced to accommodate both the SLM and detector devices. The combination of conjugate read-out in the signal beam path and laser diode arrays in the reference beam path results in a very compact holographic memory module with fast access. It is not completely lensless, since one lens still remains in the system, but such a lens would be required to collimate the laser source in any optical system that uses plane waves.

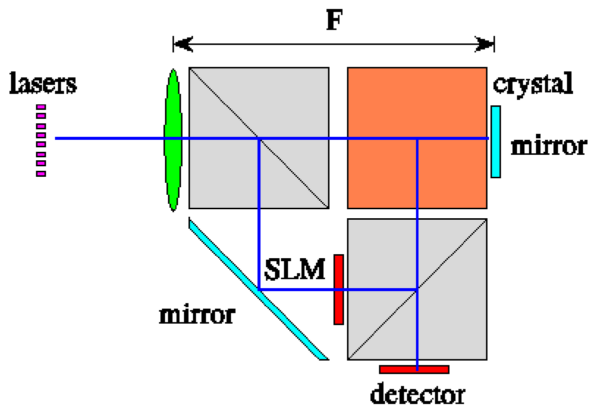


Figure 6. Compact memory module with phase conjugation incorporating separate SLM and detector devices.

1. Cost

The cost is perhaps the most important metric for accessing the commercialization prospects of HRAM. We will compare the costs of HRAM and DRAM with reference to Figure 7. We can think of HRAM as a holographic module that sits on top of a page of DRAM. The ability of the HRAM to multiplex holograms essentially allows us to store M DRAM data page, hence saving us the cost of fabricating M-1 additional DRAM pages in silicon. However, it is not quite that simple. First, the silicon device in the HRAM is not really a DRAM page, but rather the DHR chip described earlier or and SLM/detector pair. Because of the necessity of fabricating SLM and detector pixels (either in the same optoelectronic device or in two separate devices), the page density of the DHR will be less than that of a true DRAM. We call this ration of the page densities R>1. Moreover, the cost of the holographic module also includes the optical elements C_{Opt}, and laser diode array C_{LD}, in addition to the cost of the silicon C_{Si}. The projected costs of the optical elements (assuming production in large quantities) are summarized in Table 1. We assume the silicon cost to be purely based on area, and therefore will be identical to that for an equal-sized DRAM. The cost of the laser array is not well known at this time, since large arrays have not yet been produced for visible wavelengths; however, we estimate the cost to be in the range of $25-$100 per array.

The cost ratio per megabyte CR of holographic memory to the silicon storage will be:

$$\mathbf{CR} = \frac{\mathbf{C_{Si}} + \mathbf{C_{Opt}} + \mathbf{C_{VCSEL}}}{\mathbf{C_{Si}}} \cdot \frac{\mathbf{R}}{\mathbf{M}} \qquad \textbf{(1)}$$

where the R is the pixel area ratio of the SLM and detector to the silicon area of each bit on DRAM, M is the number of holograms multiplexed in the crystal on top of the silicon. With the fixed cost of silicon area C_{Si}, optical elements C_{Opt}, and LD array C_{LD}, the key to have a small cost ratio CR is to have small R and large M, which means a high storage density in holographic memory comparing with the DRAM.

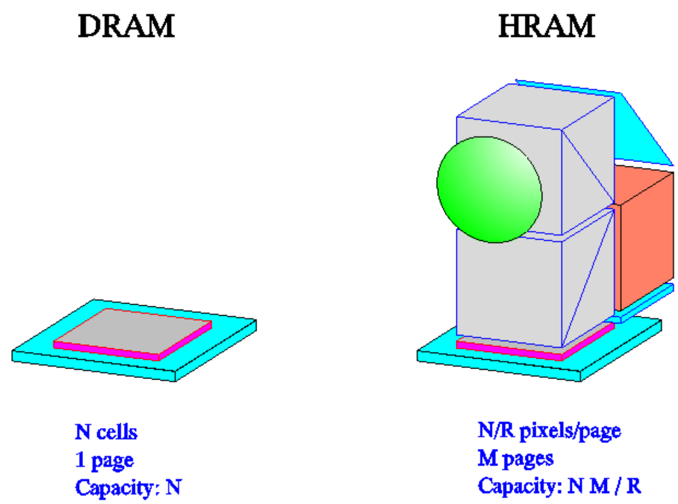


Figure 7. Model for cost comparison between HRAM and DRAM.

The number of holograms to be recorded and readout with reasonable bit error rate, is limited by the dynamic range and sensitivity, or the M/# of the material. Recording and reading 10,000 holograms at one location of a $LiNbO_3$ crystal was demonstrated with a similar system. However limited by the material M/# [11], the LD array number and power, and reasonable recording/readout rates, it is practical to keep M below 1000.

For current commercial SLM and detector array, the pixel area is typically 4x4μm^2. And the current commercial DRAM is 1 μm^2/bit, which leads R=16. With typical M=1000, we have R/M=1.6%, which leads to a small and promising CR. However if the DRAM keeps the history trend as the NTRS97 [12] projected, the DRAM cell will be 0.04μm^2/bit in 2006. To keep the R around 25, the pixel size of the holographic data pages has to be 1x1μm^2 or even smaller, which is previous proved achievable for the holographic memory system.

Figure 8 shows the experimental demonstration of conjugate hologram reconstruction of a 1x1 μm^2 random pixel mask as SLM, which gives Bit Error Rate (BER) at 7x10^{-5}. This finite BER indicates the requirement for error correction coding for the holographic memory.

Comparing the cost per megabyte for the DRAM projection of 42 cents/Mbyte in 2006, we have the cost estimation for the holographic module in table 1, where we assume the same cost per area for silicon usage. With the R=25 for 1x1μm^2 pixel size and M=500, the cost for holographic memory is around 4 cents/Mbyte, one order of magnitude lower than the DRAM in 2006.

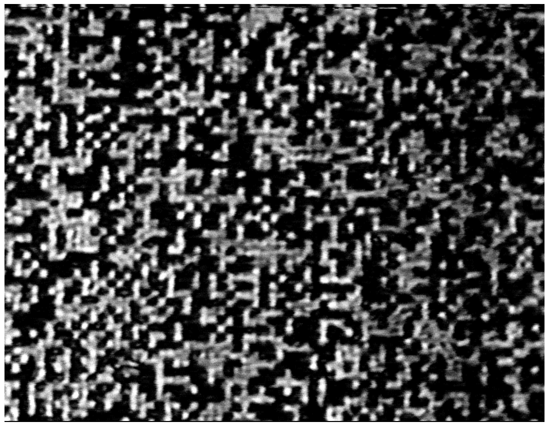


Figure 8. The phase conjugate reconstruction of 1x1 μm^2 random data mask holograms.

2. System volume density

An analysis of the system storage density of the holographic memory module (including the recording medium and all the optical components) in Figure 6 shows that the module storage density peaks at about 40Mb/cm^3 for an optimum pixel size of 5μm. There is an optimum pixel size because as the pixel size decreases the light in the signal path spreads more due to diffraction, causing us to use larger apertures for the crystal and beamsplitters.

A more aggressive concept for minimizing the volume is shown in Figure 9. This design relies on total internal reflection to contain the beam diffraction within the boundaries of the module, so that the optical elements can be made the same size as the SLM array. Preliminary experiments indicate that accurate recordings are obtained using the internally reflected light. In this case, the system density can be raised to the order of 2Gb/cm^3, if SLM pixel sizes fall to 1μm. At this density, a gigabyte of data could be stored in a single module with a volume of 1x2x2 cm^3. The challenges in achieving such high densities are several: Development of SLM and detectors with 1 micron pixels, designing the optical system so that we have uniform illumination throughout, and further characterization of the performance of the module when the light is allowed to undergo total internal reflection.

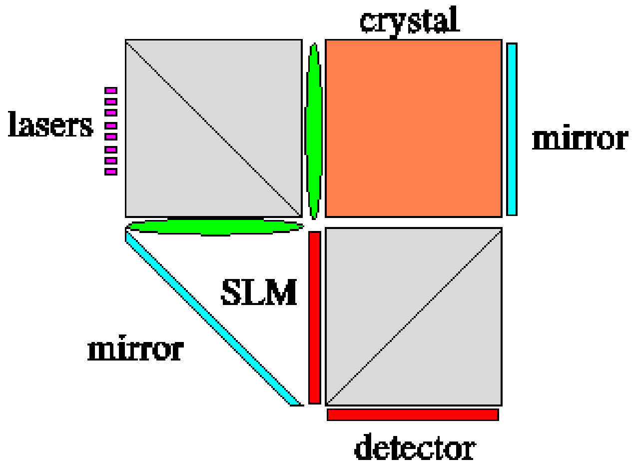


Figure 9. Variation of compact memory module for minimum volume.

3. Readout and recording rate

Since the laser diode array discussed in the previous section allows us to switch between multiplexed data pages with negligible delay (on the order of nanoseconds), the random access time and the readout rate become limited by the required integration time of the detector. We can write the integration time as

$$\textbf{Detector integration time} = \frac{\mathbf{Ne}\,h\nu N^2}{\left(\frac{M/\#}{M}\right)^2 P_i} \qquad (2)$$

where Ne is the number of electrons per pixel that we need to integrate for the given detector sensitivity and level of background noise, h is Planck's constant (6.63×10^{-34} J•s), ν is the light frequency, N^2 is the total number of pixels in the detector array, M/# is the system metric [11] of the holographic medium, M is the number of multiplexed holograms, and P_i is the incident readout power. For example, if we use a crystal of M/#=10 to record 500 holograms of a 1000x1000 pixel array, and we read out with 100mW of laser power, requiring 300 electrons per pixel, the integration time, and hence the random access time, would be 2.4 μs. This corresponds to a sustained readout transfer rate, from the hologram to the silicon detectors, of 53GB/s.

We can write the recording rate of the memory module as

$$\text{Recording rate} = \frac{N^2 ISLp}{(M/\#)/M} \qquad (3)$$

where N^2 is the total number of pixels per data page, I is the incident recording intensity, S is the sensitivity per unit length of the recording medium, L is the crystal thickness, and p is the light efficiency of the SLM. Again assuming a crystal of M/#=10 to record 500 holograms of a 1000x 1000 pixel array, with I=100mW/cm^2, S=0.1cm/J, L=1cm, and p=50%, we obtain a recording rate of 31kB/s. This is typical for experiments currently performed. Increasing the recording rate to make it comparable to the read-out rate is highly desirable for a practical system. We will discuss possible methods for achieving this goal later on.

4. Roadmap for A Competitive HRAM Technology

From the preceding discussion, we can summarize that it would be commercially competitive for a holographic memory system with parameters: M/#=10, S=1 cm/J, laser diode array with output 500 mW/cm^2 for each element and 1000 holograms storage of 10,000x10,000 pixels each page. This module expect to deliver a recording rate >100 Mbyte/sec, access time <100 μsec, and cost <$0.04 /Mbyte. For comparison, the DRAM is projected to be $0.40/MB in 2006[12].

Presently, the greatest challenge for the HRAM is to raise its recording rate by several orders of magnitude. To achieve this, we must rely in part on improvements in SLM technology to bring the pixel sizes down to 1 μm. This will allow us to increase the size of each data page to 10,000x10,000 pixels while still holding the array size to about 1cm^2. By increasing the page size in this way, we immediately gain two orders of magnitude in the sustained recording rate due to the increased parallelism. Experimentally, we have used a mask fabricated with e-beam, lithography to record and reconstruct data pages with 1 μm pixels holographically with good image fidelity. Figure 10 shows an experimental measurement of the SNR for various pixel size holograms. The reconstruction for 1 μm pixels gives SNR =4.

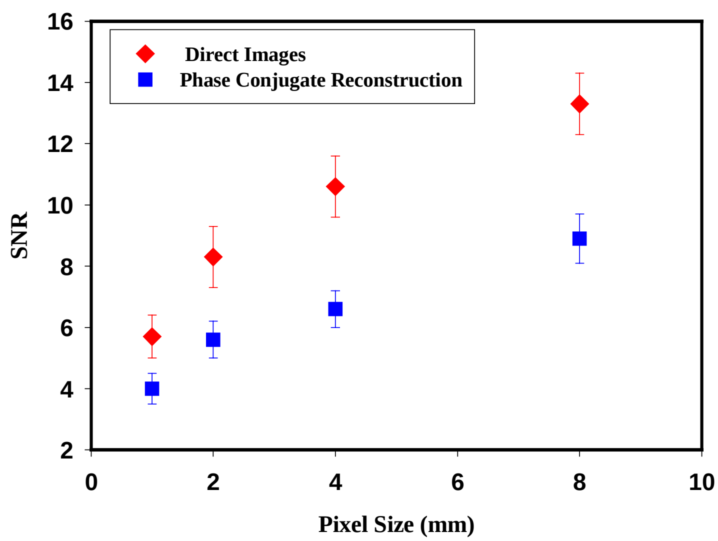


Figure 10. The SNR for the direct images and the holographic phase conjugate reconstruction of random binary data of pixel size from 8x8 down to 1x1 μm^2.

Reducing the pixel sizes to 1 μm is not only necessary for raising the recording rate, but also for maintaining the cost advantage of HRAM over DRAM. By 2006, the DRAM cell pitch is expected to fall to 0.2 μm [12]. By bringing the SLM pixel pitch down to 1 μm, we can hold the factor R in Equation (1) at 25, and beat the cost of DRAM by an order of magnitude.

Because the HRAM readout rate is limited by the electronic transfer rate out of the detector chip, we can afford to give up some readout speed in favor of increasing the recording speed. We do this by intentionally reducing the strength of the holograms so that we can record with shorter exposures, at the cost of increasing the detector integration time. In Equations (2) and (3), this is equivalent to recording in a medium with lower M/#, but without sacrificing sensitivity. Unfortunately, as we increase the required integration time we increase at the same time the random access time of the memory. In order to maintain an advantage of at least an order of magnitude over magnetic disks in random access time, we can only afford to increase the integration time to several hundreds of microseconds.

Other opportunities for increasing the recording rate can arise from improvements in laser output powers or from improving the sensitivity of the recording materials. Compact laser arrays with outputs of 500mW per emitter may be possible by 2006, or if not, we may consider sharing a larger, more powerful tunable laser among multiple HRAM modules. Increasing material sensitivity presents more of a challenge. The sensitivity of $LiNbO_3$:Fe, by far the most commonly used recording material today, Is typically around 0.02cm/J in the 90-degree geometry. In order to get recording rates on the order of 100 MB/s, we must find ways to boost the material sensitivity to about 1cm/J by improving lithium niobate's properties. For instance, switching to transmission geometry and increasing the doping level result in large increases in M/# which can be traded for better sensitivity as we discussed previously. Alternatively, we can switch to alternative materials such as doubly doped $LiNbO_3$, in which sensitivity $S > 1$ cm/J was measured in the transmission geometry. However, this is a relatively new material and much more expensive at present.

5. Conclusion

In order to develop a competing HRAM technology, three main challenges must be met: reducing pixel size to 1 μm, producing arrays of high-power laser diodes, and increasing the sensitivity of holographic recording media. Each of these tasks is difficult, but if they can be achieved by 2006, then the projected HRAM performance levels shown in previous section become feasible. Attaining these goals will position the HRAM as a viable alternative memory technology to magnetic storage, offering performance that is at least one order of magnitude better in terms of random access and transfer rate than magnetic hard disks, and at least one tenth the cost compared to fabricating an equivalent memory in DRAM.

Acknowledgment:
The authors would like to thank Drs. E. Chuang, J-J. Drolet, G. Barbastathis, and A. Adibi for helpful discussions.

References:

[1] P.J. van Heerden, “Theory of optical information storage in solids,” *Applied Optics*, vol.2, no.4, pp.393--400, 1963.

[2] F. H. Mok, “Angle-multiplexed storage of 5000 holograms in Lithium Niobate,” *Optics Letters*, vol. 18, no. 11, pp. 915-917, June 1993.

[3] X. An, D. Psaltis, G. W. Burr, “Thermal fixing of 10,000 holograms in LiNbO3 : Fe,” *APPL OPTICS* 38: (2) 386-393 JAN 10 1999.

[4] J. Drolet, G. Barbastathis, J. Patel, and D. Psaltis, “Liquid crystal devices for volume holographic memories,” in *Proc. OSA Annu. Meeting*, Portland, OR, Sept. 1995.

[5] J. Drolet, E. Chuang, G. Barbastathis, D. Psaltis, “Compact, integrated dynamic holographic memory with refreshed holograms,” *OPTICS LETTERS*, 22: (8) 552-554 APR 15 1997

[6] Z. O. Feng and K. Sayano, “Compact read-only memory with lensless phase-conjugate holograms,” *Optics Letters*, vol. 21, pp.1295--1297, August 1996

[7] H. Zhou, F. Zhao, and F. Yu, “Angle-dependent diffraction efficiency in a thick photorefractive hologram”, *Applied Optics*, Vol. 34, No. 8, pp. 1303-1309, Mar. 1995.

[8] W. Liu, and D. Psaltis, “Pixel size limit in holographic memories”, *optics letters*, vol. 24: pp. 1340-1342 Oct. 1999.

[9] J. L. Jewell, K. F. Huang, K. Tai, Y.H. Lee, R. J. Fischer, S. L. Mccall, and A. Y. Cho, “Vertical cavity single quantum well laser,” *Applied Physics Letters*, vol.55, no.5, pp.424--426, 1989.

[10] W. W. Chow, K. D. Choquette, M. H. Crawford, K. L. Lear, and G. R. Hadley, “Design, fabrication, and performance of infrared and visible vertical-cavity surface-emitting lasers,” *IEEE Journal of Quantum Electronics*, vol.33, pp.1810-1824, October 1997.

[11] F.H. Mok, G.W. Burr, and D. Psaltis, “System metric for holographic memory systems,” *Optics Letters*, 21: (12) 896-898 JUN 15 1996.

[12] “The national technology roadmap for semiconductors,” Semiconductor Industry Association, Tech. Rep., San Jose, Ca 95110, 1997.