Holographic data storage

by J. Ashley
    M.-P. Bernal
    G. W. Burr
    H. Coufal
    H. Guenther
    J. A. Hoffnagle

    C. M. Jefferson
    B. Marcus
    R. M. Macfarlane
    R. M. Shelby
    G. T. Sincerbox

We present an overview of our research effort on volume holographic digital data storage. Innovations, developments, and new insights gained in the design and operation of working storage platforms, novel optical components and techniques, data coding and signal processing algorithms, systems tradeoffs, materials testing and tradeoffs, and photon-gated storage materials are summarized.

Introduction

With its omnipresent computers, all connected via the Internet, the Information Age has led to an explosion of information available to users. The decreasing cost of storing data, and the increasing storage capacities of the same small device footprint, have been key enablers of this revolution. While current storage needs are being met, storage technologies must continue to improve in order to keep pace with the rapidly increasing demand.

However, both magnetic and conventional optical data storage technologies, where individual bits are stored as distinct magnetic or optical changes on the *surface* of a recording medium, are approaching physical limits beyond which individual bits may be too small or too difficult to store. Storing information throughout the *volume* of a medium—not just on its surface—offers an intriguing high-capacity alternative. Holographic data storage is a volumetric approach which, although conceived decades ago, has made recent progress toward practicality with the appearance of lower-cost enabling technologies, significant results from longstanding research efforts, and progress in holographic recording materials.

In holographic data storage, an entire page of information is stored at once as an optical interference pattern within a thick, photosensitive optical material (**Figure 1**). This is done by intersecting two coherent laser beams within the storage material. The first, called the object beam, contains the information to be stored; the second, called the reference beam, is designed to be simple to reproduce—for example, a simple collimated beam with a planar wavefront. The resulting optical interference pattern causes chemical and/or physical changes in the photosensitive medium: A replica of the interference pattern is stored as a change in the absorption, refractive index, or thickness of the photosensitive medium. When the stored interference grating is illuminated with one of the two waves that was used during recording [**Figure 2(a)**], some of this incident light is diffracted by the stored grating in such a fashion that the other wave is reconstructed. Illuminating the stored grating with the reference wave reconstructs the object wave, and vice versa [**Figure 2(b)**]. Interestingly, a backward-propagating or phase-conjugate reference wave, illuminating the stored grating from the "back" side, reconstructs an object wave that also propagates backward toward its original source [**Figure 2(c)**].

A large number of these interference gratings or patterns can be superimposed in the same thick piece of media and can be accessed independently, as long as they are distinguishable by the direction or the spacing of the gratings. Such separation can be accomplished by changing the angle between the object and reference wave or by changing the laser wavelength. Any particular data page can then be read out independently by illuminating the stored gratings with the reference wave that was used to
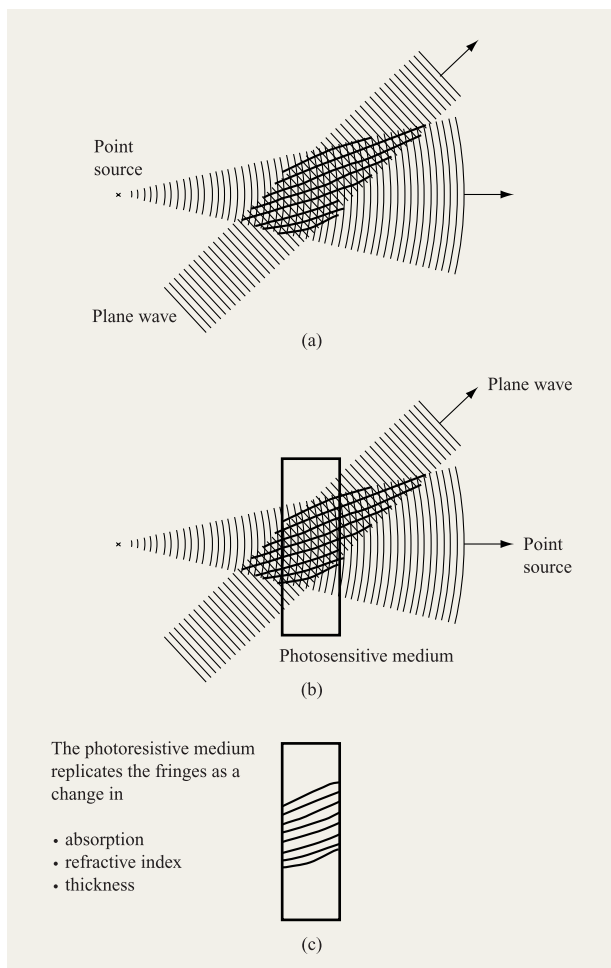
341

Figure 1

Storage of one bit of information as a hologram: (a) Superposition of the spherical wave from one bit with a coherent plane wave reference beam forming an interference pattern. (b) Exposure of a photosensitive medium to the interference pattern. (c) Record of the interference grating, stored as changes in the refractive properties of the medium.

number of relatively slow, and therefore low-cost, parallel channels.

The data to be stored are imprinted onto the object beam with a pixelated input device called a spatial light modulator (SLM); typically, this is a liquid crystal panel similar to those on laptop computers or in modern camcorder viewfinders. To retrieve data without error, the object
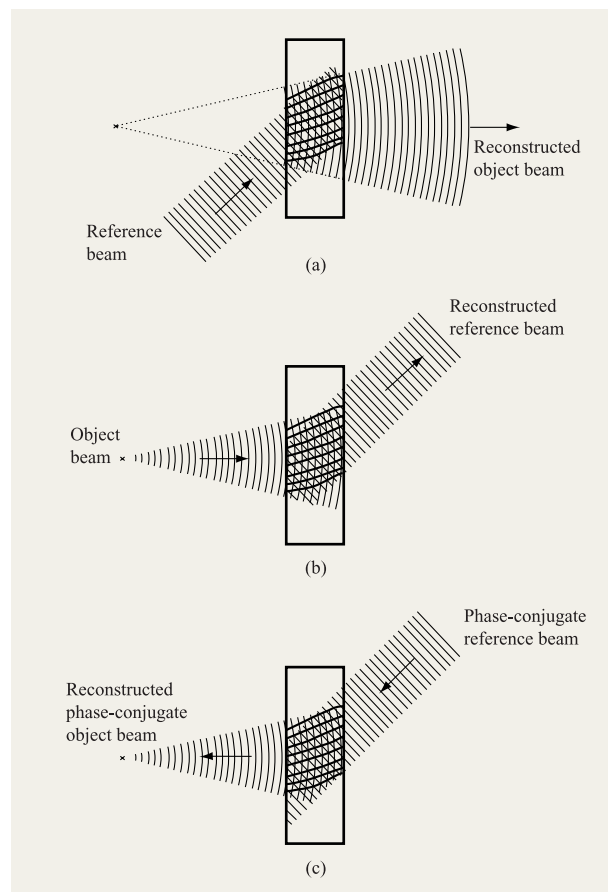

Figure 2

Reading of holographic information by (a) Illumination with the reference beam, which is diffracted by the stored interference pattern to reconstruct the original spherical wavefront of the object beam. This beam can be imaged to a single small detector, resulting in the retrieval of a single bit. (b) Illumination with the diverging object beam, which is diffracted by the stored interference pattern to reconstruct the original plane wave reference beam. This beam can be focused to a detector, representing an optical measurement of the correlation between the stored data and the illuminating object beam, allowing content-addressable searching. (c) Illumination with a counter-propagating (or "phase-conjugate") reference beam, which is diffracted by the stored interference pattern to reconstruct a phase-conjugate copy of the original object beam. This phase-conjugate object beam returns to its original point of origin, where the stored bit value can be read without requiring a high-quality imaging system.

store that page. Because of the thickness of the hologram, this reference wave is diffracted by the interference patterns in such a fashion that only the desired object beam is significantly reconstructed and imaged on an electronic camera. The theoretical limits for the storage density of this technique are around tens of terabits per cubic centimeter.

In addition to high storage density, holographic data storage promises fast access times, because the laser beams can be moved rapidly without inertia, unlike the actuators in disk drives. With the inherent parallelism of its pagewise storage and retrieval, a very large compound data rate can be reached by having a large

beam must contain a high-quality imaging system—
one capable of directing this complex optical wavefront
through the recording medium, where the wavefront
is stored and then later retrieved, and then onto
a pixelated camera chip (**Figure 3**). The image of the
data page at the camera must be as close as possible to
perfect. Any optical aberrations in the imaging system or
misfocus of the detector array would spread energy from
one pixel to its neighbors. Optical distortions (where
pixels on a square grid at the SLM are not imaged to a
square grid) or errors in magnification will move a pixel
of the image off its intended receiver, and either of
these problems (blur or shift) will introduce errors in
the retrieved data. To avoid having the imaging system
dominate the overall system performance, near-perfect
optics would appear to be unavoidable, which of course
would be expensive. However, the above-mentioned
readout of phase-conjugated holograms provides a partial
solution to this problem. Here the reconstructed data page
propagates backward through the same optics that were
used during the recording, which compensates for most
shortcomings of the imaging system. However, the
detector and the spatial light modulator must still be
properly aligned.

A rather unique feature of holographic data storage is
associative retrieval: Imprinting a *partial* or search data
pattern on the object beam and illuminating the stored
holograms reconstructs all of the reference beams that
were used to store data. The intensity that is diffracted
by each of the stored interference gratings into the
corresponding reconstructed reference beam is
proportional to the similarity between the search
pattern and the content of that particular data page. By
determining, for example, which reference beam has the
highest intensity and then reading the corresponding data
page with this reference beam, the closest match to the
search pattern can be found without initially knowing its
address.

Because of all of these advantages and capabilities,
holographic storage has provided an intriguing alternative
to conventional data storage techniques for three decades.
However, it is the recent availability of relatively low-cost
components, such as liquid crystal displays for SLMs
and solid-state camera chips from video camcorders for
detector arrays, which has led to the current interest in
creating practical holographic storage devices. Recent
reviews of holographic storage can be found in [1–4].
A team of scientists from the IBM Research Division
have been involved in exploring holographic data storage,
partially as a partner in the DARPA-initiated consortia
on holographic data storage systems (HDSS) and on
photorefractive information storage materials (PRISM).
In this paper, we describe the current status of our
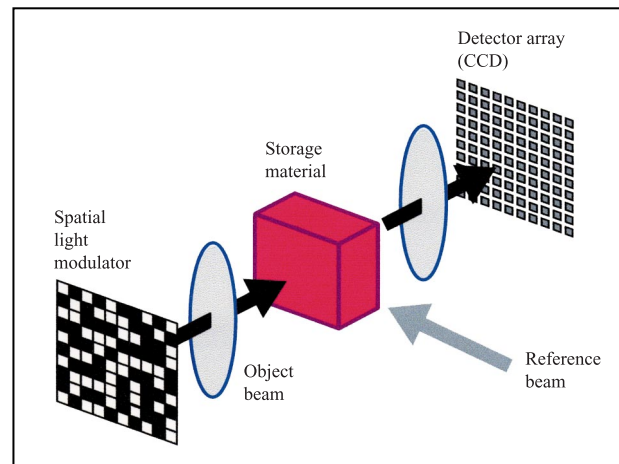effort.

Figure 3

Basic holographic data storage system. Data are imprinted onto the
object beam with a pixelated input device called a spatial light
modulator (SLM). A pair of lenses image the data through the
storage material onto a pixelated detector array such as a charge-
coupled device (CCD). A reference beam intersects the object
beam in the storage material, allowing the storage and later
retrieval of holograms.

The overall theme of our research is the evaluation
of the engineering tradeoffs between the performance
specifications of a practical system, as affected by the
fundamental material, device, and optical physics.
Desirable performance specifications include data fidelity
as quantified by bit-error rate (BER), total system
capacity, storage density, readout rate, and the lifetime of
stored data. This paper begins by describing the hardware
aspects of holographic storage, including the test platforms
we have built to evaluate materials and systems tradeoffs
experimentally, and the hardware innovations developed
during this process. Phase-conjugate readout, which eases
the demands on both hardware design and material
quality, is experimentally demonstrated. The second
section of the paper describes our work in coding and
signal processing, including modulation codes, novel
preprocessing techniques, the storage of more than one bit
per pixel, and techniques for quantifying coding tradeoffs.
Then we discuss associative retrieval, which introduces
parallel search capabilities offered by no other storage
technology. The fourth section describes our work in
testing and evaluating materials, including permanent
or write-once read-many-times (WORM) materials,
read–write materials, and photon-gated storage materials
offering reversible storage without sacrificing the lifetime
of stored data. The paper concludes with a discussion
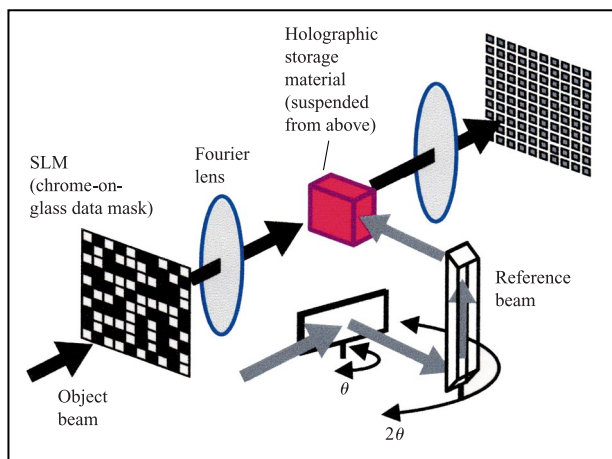of applications for holographic data storage.

343

Hardware for holographic data storage

Figure 3 shows the most important hardware components in a holographic storage system: the SLM used to imprint data on the object beam, two lenses for imaging the data onto a matched detector array, a storage material for recording volume holograms, and a reference beam intersecting the object beam in the material. What is not shown in Figure 3 is the laser source, beam-forming optics for collimating the laser beam, beamsplitters for dividing the laser beam into two parts, stages for aligning the SLM and detector array, shutters for blocking the two beams when needed, and waveplates for controlling polarization. Assuming that holograms will be angle-multiplexed (superimposed yet accessed independently within the same volume by changing the incidence angle of the reference beam), a beam-steering system directs the reference beam to the storage material. Wavelength multiplexing has some advantages over angle-multiplexing, but the fast tunable laser sources at visible wavelengths that would be needed do not yet exist.

The optical system shown in Figure 3, with two lenses separated by the sum of their focal lengths, is called the "4-f" configuration, since the SLM and detector array turn out to be four focal lengths apart. Other imaging systems such as the Fresnel configuration (where a single lens satisfies the imaging condition between SLM and detector array) can also be used, but the 4-f system allows the high numerical apertures (large ray angles) needed for high density. In addition, since each lens takes a spatial Fourier transform in two dimensions [5], the hologram stores the Fourier transform of the SLM data, which is then Fourier-transformed again upon readout by the second lens. This has several advantages: Point defects on the storage material do not lead to lost bits, but result in a slight loss in signal-to-noise ratio at all pixels; and the storage material can be removed and replaced in an offset position, yet the data can still be reconstructed correctly. In addition, the Fourier transform properties of the 4-f system lead to the parallel optical search capabilities offered by holographic associative retrieval. The disadvantages of the Fourier transform geometry come from the uneven distribution of intensity in the shared focal plane of the two lenses, which we discuss in the axicon section below.

Holographic digital data storage testers

In order to study the recording physics, materials, and systems issues of holographic digital data storage in depth, we have built three precision holographic recording testers. Each of these platforms is built around the basic design shown in Figure 3, implementing mapping of single SLM pixels to single detector pixels using precision optics in the object beam, and angle-multiplexing in the reference beam. In addition, care has been taken in the design and assembly of the components listed above but not shown in Figure 3, in order to allow experimental access to a wide range of holographic data storage parameters with minimal instrumental contributions to the raw error rate. The three testers, described in the following sections, are called the PRISM tester, the DEMON I platform, and the DEMON II platform.

• *PRISM tester*
The PRISM tester [6, 7], built as part of the DARPA PhotoRefractive Information Storage Materials consortium, was designed to allow the rigorous evaluation of a wide variety of holographic storage materials. This tester was designed for extremely low-baseline BER performance, flexibility with regard to sample geometry, and high stability for both long recording exposures and experimental repeatability. The salient features of the PRISM tester are shown in **Figure 4**. The SLM is a chrome-on-glass mask, while the detector array is a low-frame-rate, 16-bit-per-pixel CCD camera. Custom optics of long focal length (89 mm) provide pixel matching over data pages as large as one million pixels, or one megapel. A pair of precision rotation stages direct the reference beam, which is originally below the incoming object beam, to the same horizontal plane as the object beam. By rotating the outer stage twice as far as the inner, the reference-beam angle can be chosen from the entire 360-degree angle range, with a repeatability and accuracy

of approximately one microradian. (Note, however, that over two 30-degree-wide segments within this range, the reference-beam optics occlude some part of the object-beam path.) The storage material is suspended from a three-legged tower designed for interferometric stability (better than 0.1 μm) over time periods of many seconds. The secondary optics occupy approximately 2 feet by 4 feet of optical table space, and the tower and stages approximately 4 feet by 4 feet.

The system is equipped with an argon (514.5-nm) and a krypton (676-nm) laser, and all optics are optimized to work at both wavelengths. Beam-forming optics and shutters control the power and polarization of the object and reference beams, and relay optics overexpand the object beam to ensure a uniform illumination of the data mask. Precision linear stages control the position of the data mask in two axes (allowing selection from a set of multiple patterns), the Fourier lenses in one axis each (to control magnification), and the crystal position in three axes. In addition, the crystal can be rotated about two axes, and the camera position controlled in three linear axes and one rotational axis. All stages and shutters are under computer control, allowing direct operator control of the system as well as unsupervised execution of long experiments. While the camera uses 1024×1024 detector pixels on 9-μm centers, data masks are available with pixel pitch of 36 μm (resulting in 65536 data pixels), 18 μm (262144 pixels), and 9 μm (1048576 data pixels, also known as a "megapel"). The baseline BER performance of the system without a storage material (limited only by the imaging system) was estimated to be 1×10^{-18} with the low-resolution mask, 1×10^{-12} with the medium-resolution mask, and 1×10^{-7} with the megapel data mask.

Figure 5 shows the experimental demonstration of holographic storage and retrieval of a 1Mb data page, with object and reference beams entering orthogonal faces (90-degree geometry) of an Fe-doped lithium niobate (LiNbO$_3$) crystal [7]. This histogram shows the occurrence of intensity levels in the data page detected by the camera. Since the data mask pattern of bright ("1") and dark ("0") pixels is known, the intensity levels of each of these classes can be plotted separately. In the absence of random noise and deterministic variations, all bright pixels would have the same detected intensity, which would be well separated from the intensity of all dark pixels, resulting in two spikes. Instead, the distribution of intensities makes it more difficult to apply a single threshold and separate the bright and dark pixels in the real data-retrieval scenario (for which the data mask pattern is, by definition, unknown). While this particular page has no detected errors, the distributions can be fitted with Gaussian approximations to provide a BER estimate of 2.4×10^{-6}. Since this hologram was retrieved using a readout pulse of 1 ms, this experiment implements the
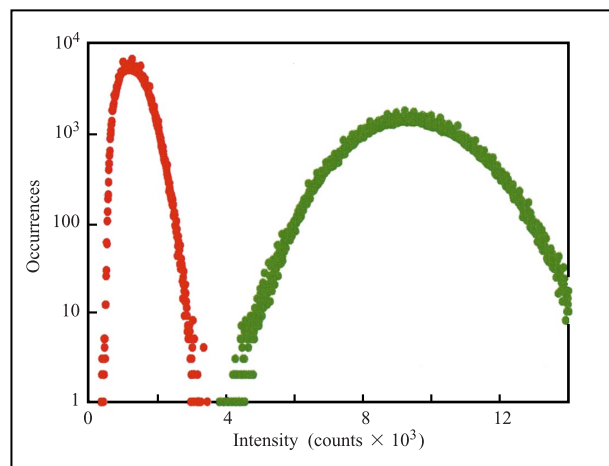

Figure 5

Histogram of received data values for an error-free one-million-pixel data page reconstructed with a one-millisecond pulse of read-out light. There are two measured distributions of received intensities: dark (binary "0") pixels on the left, and bright ("1") pixels on the right. The intensity scale covers the 12-bit range of the CCD camera. In the absence of noise, each distribution would be a single spike (all detector pixels which were supposed to receive a bright pixel would measure exactly the same intensity). If the two distributions were to blur into each other, a simple intensity threshold would not be able to assign binary values without error. From [7], reproduced with permission.

optical signal (but not the subsequent fast electronic readout) of a system with a readout rate of 1 Gb/s.

• DEMON I

While PRISM was designed to handle any conceivable material testing requirement, the DEMON I platform [8], shown in **Figure 6**, was built to be a platform for evaluating coding and signal processing techniques. The reference/object-beam geometry was restricted to the 90-degree geometry, and the reference beam deflected with a galvanometrically actuated mirror through a simple 4-f system, limiting the variation of the angle to ±10 degrees. A transmissive liquid crystal SLM, capable of displaying arbitrary data patterns, was pixel-matched onto a small, 60-Hz CCD camera in two stages. First, a precision five-element zoom lens demagnified the SLM (640×480 pixels with 42-μm pitch) to an intermediate image plane (same pixel count on 18-μm pitch). Then a set of Fourier lenses identical to those in the PRISM imaged this plane 1:1 onto the detector array (640×480 pixels, but 9-μm pitch). Because of the finer pitch on the CCD, only the central 320×240 field of the SLM was detected. To implement true pixel matching, the detector was aligned so that light from each SLM pixel fell squarely on a single
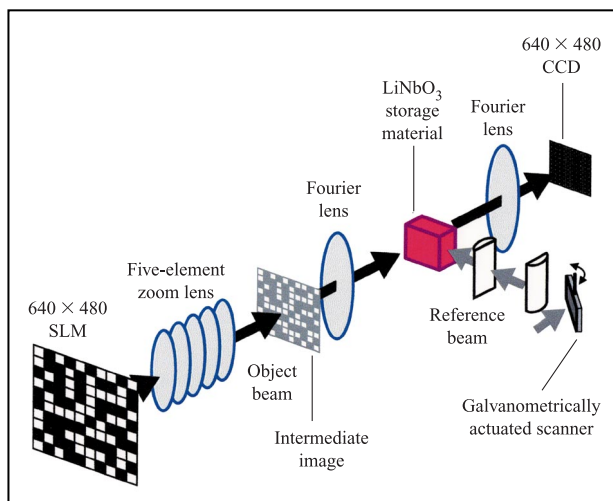
345

Salient features of the DEMON I holographic digital data storage engine. A five-element zoom lens demagnifies the SLM to an intermediate image plane, which is then imaged to the CCD detector with a pair of lenses. The reference beam and object beams enter orthogonal faces of a $LiNbO_3$ crystal; a galvanometrically actuated scanner changes the reference beam angle over ± 10 degrees about the normal.

Primary features of the DEMON II holographic digital data storage engine. Utilizing 30-mm-focal-length Fourier transform lenses in the 90-degree geometry with a one-million-pixel SLM, this system has demonstrated areal storage densities in excess of 100 bits/μm^2.

detector pixel (thus ignoring three of every four pixels on the CCD). Laser light from the green 514.5-nm line of an argon-ion laser was delivered to the platform with a single-mode polarization-preserving optical fiber, which produces a clean Gaussian intensity profile. Optical power delivered to the apparatus prior to the object/reference beamsplitter was as much as 400 mW. Simple linear stages move the SLM in two axes and the CCD in three axes for alignment. The entire system, not including the laser, occupies 18×24 inches of optical table space.

The first experiment performed on the DEMON I tester was the demonstration of multiple hologram storage at low raw BER (BER without error correction) using modulation codes, which allow decoding over smaller pixel blocks than the global thresholding described above. Using an 8-mm-thick $LiNbO_3$:Fe crystal storage medium and a strong modulation code (8:12), 1200 holograms were superimposed and read back in rapid succession with extremely low raw BER ($<2 \times 10^{-8}$) [8]. Subsequently, many of the codes and signal-processing algorithms and techniques described later in this paper were either refined or invented on the DEMON I system. In addition, the DEMON I platform has been used to implement both associative retrieval and phase-conjugate readout, as described below.
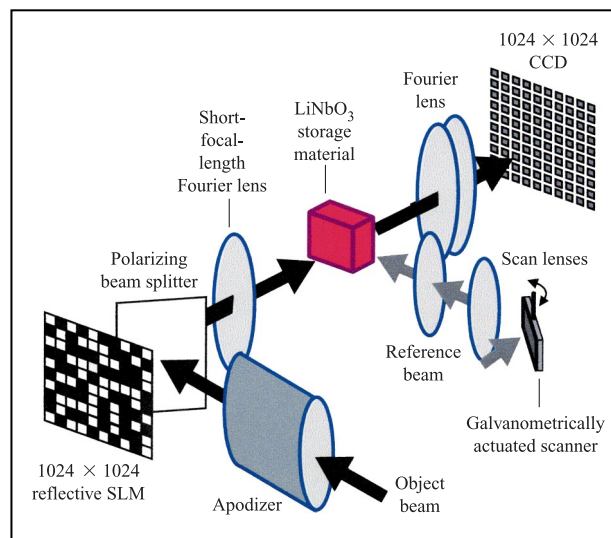
• DEMON II

The DEMON II holographic storage platform, shown in **Figure 7**, was designed to achieve high-density holographic data storage using short-focal-length optics, while including aspects of the previous two test platforms. DEMON II combines the large data pages of the PRISM tester with the dynamic SLM and the 90-degree geometry configuration of the DEMON I platform. Here, the SLM is a reflective device fabricated by IBM Yorktown [9], containing 1024×1024 pixels and illuminated via a polarizing beamsplitter cube. A novel apodizer, described in the next section, provides uniform illumination over the entire data page without sacrificing input power. The magnification from the 12.8-μm pitch of the SLM pixels to the 12-μm pitch of the 41-Hz CCD camera (1024×1024 pixels, 41 frames per second) is built into the Fourier optics (effective focal length 30 mm). A pair of scan lenses provide an improved relay of the reference beam from the galvanometrically actuated mirror to the $LiNbO_3$ crystal, providing diffraction-limited performance over an angular scan range of ± 15 degrees.

The laser light is provided by a diode-pumped solid-state laser (532 nm, doubled Nd-YAG); waveplates and polarizing beamsplitters provide control over the power in the reference beam and object beam. The use of two separate elements in the back Fourier lens (between the storage material and the detector array) allows the magnification of the optical system to be varied over a

range of ±0.5%. Linear stages provide two axes of motion for the storage material and three axes of motion for the detector array. The entire system, including the laser, occupies 2 feet × 2 feet. As with the PRISM and DEMON I systems, all stages and shutters are under computer control, allowing both direct operator control of the system and unsupervised execution of complex scripted experiments.

The short focal length of the DEMON II optics allows the system to demonstrate high areal storage densities (the storage capacity of each stack of holograms, divided by the area of the limiting aperture in the object beam). Since the lenses in the object beam implement a two-dimensional spatial Fourier transform, an aperture placed in the central focal plane of the 4-f system (just in front of the storage material) can be described as a spatial low-pass filter. The smaller the volume allocated to each stack of holograms, the larger the capacity of a given large block of storage material. However, if the aperture is decreased too far, some of the information from the SLM fails to pass through the aperture. The size of the smallest tolerable aperture corresponds to the spatial equivalent of the Nyquist sampling condition, in which the spatial frequency sampling on the SLM (one over its pixel pitch) is twice the maximum spatial frequency allowed to pass the limiting aperture. Only for apertures equal to or larger than this so-called "Nyquist" aperture is the information from all pixels of the SLM guaranteed to pass to the detector array. Since both "positive" and "negative" spatial frequencies are represented in a centered aperture, the Nyquist aperture turns out to be equal to the inverse of the pixel pitch of the SLM, scaled by the wavelength and the focal length of the lenses. The design of the imaging optics is then complicated by this need for short focal length, since the maximum ray angle (and thus the potential for optical aberrations) is greatly increased. The optical distortion (displacement of pixel centers from a rectangular grid) in the DEMON II platform is consequently much larger than in the other two testers, reaching approximately 0.03% (0.3 pixels) in the corners of the received data page. The development of signal-processing algorithms to compensate for this misregistration between SLM and CCD pixels is a research topic that we are currently pursuing, with some initial success.

- *Innovative optics*

In the course of development of PRISM, DEMON I, and DEMON II, a number of challenging optical design problems arose. Here we describe two innovative hardware solutions that have been developed.

Axicon

As previously noted, the Fourier transform process used to focus the object beam into the storage media has the
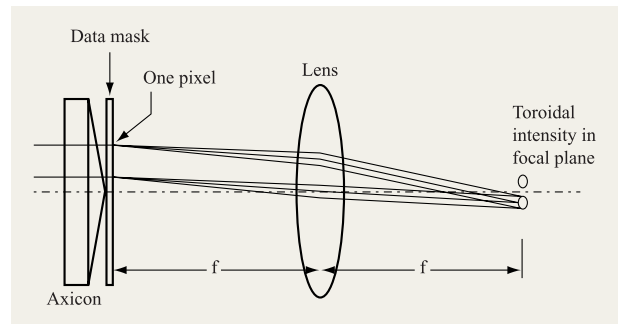
A convex axicon deflects the Fourier components of the pixelated data mask into a toroid. This spreads out the rays, which would otherwise focus to a single large spike at the optical axis of the focal plane and distort holographic data pages.

side effect of producing an undesired high-intensity peak on the optical axis. This intensity spike can easily saturate the photosensitive response of the storage media, resulting in severe degradation of both transmitted images and stored holograms. It has been known for many years that a potential solution to this problem can be implemented by superimposing a random phase distribution on the pixels of the SLM. In work performed by M.-P. Bernal et al. at IBM Almaden [10], it was shown that although such a "random phase mask" does redistribute the intensity in this spike, the alignment of such a phase mask is critical, and new optical artifacts (dark lines and interference fringe effects) are introduced in the transmitted image. These artifacts, along with the difficulty of maintaining the alignment of yet another pixelated component, have made it improbable that random phase masks will be the solution to the coherent saturation problem.

As an alternative, we have developed several optical structures which also spread the energy in the undesired intensity spike across the Fourier transform plane, without requiring precision alignment. One particular structure of interest is the axicon, a simple cylindrically symmetric cone of glass, typically with an oblique vertex angle. Introducing the axicon in the illumination beam of the SLM distributes the undesired intensity spike along a ring in the Fourier plane. The diameter of the ring depends on the vertex angle of the conical optic, the index of refraction, and the focal length of the Fourier lens. The axicon can either be placed directly behind the data mask or SLM, as shown in **Figure 8**, or, preferably, imaged onto the SLM using some relay optics. In the latter case, there is some slight longitudinal alignment sensitivity (but little sensitivity to transverse position). These relay optics can double as the beam expander used to fill the SLM aperture,
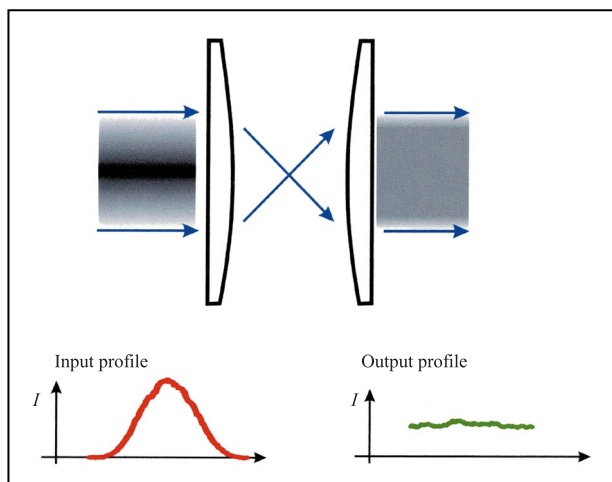
347

with the axicon placed at its input focal plane. The axicon has been shown to slow down the degradation of the object-beam imaging path with optical exposure to the same degree as the random phase mask [11], without requiring precision alignment or increasing interpixel crosstalk [10].

Aspherical apodizer
Typical laser beams have a spatial profile dictated by the oscillation mode of the laser resonator, with the simplest mode having a Gaussian or bell-shaped profile. The simplest method for generating a beam with a uniform (or flat) spatial profile is to simply expand a Gaussian beam and use only the center portion. The power efficiency then trades off directly with the desired flatness of illumination: If an illumination flatness of 5% is required over a certain area, only 5% of the incident beam power can actually be used. It has long been desirable in laser physics to be able to efficiently generate a laser beam with a uniform cross section. Although many ingenious solutions have been proposed, the few that have been implemented generally work only over the first $1/e$ field points of the original Gaussian beam, and commonly suffer from poor flatness, severe diffraction effects, and distortion of the wavefront quality of the apodized beam. In addition, many solutions, including diffractive optics, create a beam which attains uniform intensity in one plane in space, but then diverges and distorts away from that plane.

As part of the design of DEMON II, the creation of "flat-top" beams was studied. This was germane not only to DEMON II, but also to ongoing work in deep-UV

lithography. A new insight was obtained after a review of historical efforts in this field. A two-element telescope with transmissive optical elements was designed that produces a highly efficient flat-top laser beam with the capability of propagating for several meters with little distortion and diffraction-limited wavefront quality. The Gaussian-beam-to-flat-top converter utilizes a convex aspheric lens to introduce aberrations into the beam, redistributing the laser power from a particular incident Gaussian profile to the desired flat-top profile with a rapid-intensity roll-off at the edge. A second aspheric optic recollimates the aberrated beam, restoring the wavefront quality and allowing it to propagate for long distances without spreading. As a result, the central 60% of the output power will be uniform in intensity to 2%, and 99.7% of the incident laser beam power is used in the output apodized beam. The roll-off of the intensity profile was carefully crafted to minimize diffraction effects from the edge of the beam during propagation. Although the input and output beam dimensions are fixed for a given apodizer, it was discovered that a single apodizer could be used from the deep UV into the far IR with only a simple focus adjustment.

Fabrication of such aspheric elements has long been very difficult and costly. Recently, new computer-controlled polishing technology has become available which can make the fabrication of such aspheric surfaces routine. Working closely with the vendor who developed these fabrication capabilities has allowed the DEMON II design team to build such an apodizer and to demonstrate that it works. **Figure 9** shows an example of input and output intensity profiles (not showing the roll-off) measured using the apodizer. A second design will achieve tighter specifications through the use of more sophisticated optical testing devices (computer-generated holograms) during fabrication. This apodizer represents a real step forward in the area of laser illumination control, and many potential applications in a variety of areas have already surfaced.

• *Phase-conjugate readout*
As described in the previous sections on tester platforms, the need for both high density and excellent imaging requires an expensive short-focal-length lens system corrected for all aberrations (especially distortion) over a large field, as well as a storage material of high optical quality. Several authors have proposed bypassing these requirements by using phase-conjugate readout of the volume holograms [12–15]. After the object beam is recorded from the SLM with a reference beam, the hologram is reconstructed with a phase-conjugate (time-reversed copy) of the original reference beam. The diffracted wavefront then retraces the path of the incoming object beam in reverse, canceling out any

accumulated phase errors. This should allow data pages to be retrieved with high fidelity with a low-performance lens, from storage materials fabricated as multimode fibers [12, 13], or even without imaging lenses [14, 15] for an extremely compact system.

Most researchers have relied on the visual quality of retrieved images or detection of isolated fine structure in resolution targets as proof that phase-conjugate retrieval provides high image fidelity. This, however, is no guarantee that the retrieved data pages will be correctly received by the detector array. In fact, the BER of pixel-matched holograms can be used as an extremely sensitive measure of the conjugation fidelity of volume holograms. Any errors in rotation, focus, x-y registration, magnification, or residual aberrations will rapidly increase the measured bit-error rate (BER) for the data page. Using the pixel-matched optics in both the DEMON I platform and the PRISM tester, we have implemented low-BER phase-conjugate readout of large data pages. On the PRISM tester, phase conjugation allowed the readout of megapel pages through much smaller apertures than in the original megapel experiment mentioned above, which was performed without phase conjugation [7]. This demonstrates a thirtyfold increase in areal density per hologram.

Figure 10 shows a simplified diagram of the PRISM tester, modified for this phase-conjugate experiment. The Fourier lenses were removed, and the object beam was focused by a lens through the megapel mask onto a mirror placed halfway between the mask and CCD. After deflection by this mirror, the object beam was collected by a second lens, forming an image of the mask. Here an Fe-doped $LiNbO_3$ crystal was placed to store a hologram in the 90-degree geometry [16]. After passing through the crystal, the polarization of the reference beam was rotated and the beam was focused into a self-pumped phase-conjugate mirror [17] using a properly oriented, nominally undoped $BaTiO_3$ crystal. In such a configuration, the input beam is directed through the $BaTiO_3$ crystal and into the far corner, creating random backscattering throughout the crystal. It turns out that counter-propagating beams (one scattered upon input to the crystal, one reflected from the back face) are preferentially amplified by the recording of real-time holograms, creating the two "pump" waves for a four-wave-mixing process. Since momentum (or wave-vector) must be conserved among four beams (energy is already conserved because all four wavelengths are identical), and since the two "pump" beams are already counter-propagating, the output beam generated by this process must be the phase-conjugate to the input beam [17].

The crystal axes of the $LiNbO_3$ were oriented such that the return beam from the phase-conjugate mirror wrote the hologram, and the strong incoming reference beam was used for subsequent readout [16]. (Although both
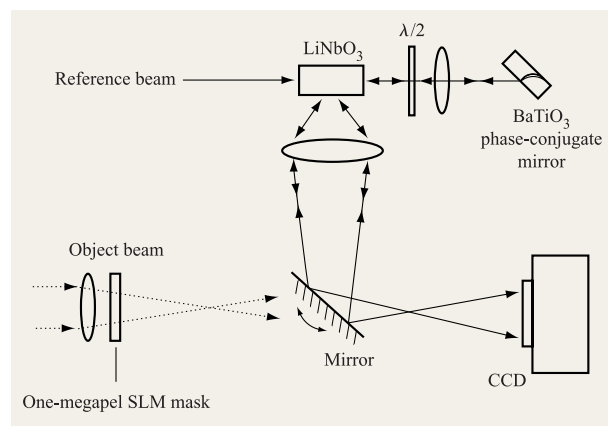
mutually phase-conjugate reference beams were present in the $LiNbO_3$ during recording, only the beam returning from the phase-conjugate mirror wrote a hologram because of the orientation of the $LiNbO_3$ crystal axes. For readout, the phase-conjugate mirror was blocked, and the incoming reference beam read this hologram, reconstructing a phase-conjugate object beam.) By turning the mirror by 90 degrees, this phase-conjugate object beam was deflected to strike the pixel-matched CCD camera. We were able to store and retrieve a megapel hologram with only 477 errors (BER $\approx 5 \times 10^{-4}$) after applying a single global threshold. The experiment was repeated with a square aperture of 2.4 mm on a side placed in the object beam at the $LiNbO_3$ crystal, resulting in 670 errors. Even with the large spacing between SLM and CCD, this is already an areal density of 0.18 bits per μm^2 per hologram. In contrast, without phase-conjugate readout, an aperture of 14 mm × 14 mm was needed to produce low BERs with the custom optics [7]. The use of phase-conjugate readout allowed mapping of SLM pixels to detector pixels over data pages of 1024 pixels × 1024 pixels without the custom imaging optics, and provided an improvement in areal density (as measured at the entrance aperture of the storage material) of more than 30.

In a second experiment, we modified the DEMON I platform in an analogous manner, using a $BaTiO_3$ crystal for phase conjugation and $LiNbO_3$ for recording data-bearing holograms of 320 pixels × 240 pixels. To demonstrate the phase-conjugation properties, the two retrieved pages of **Figure 11** illustrate the results of passing the object beam through a phase aberrator
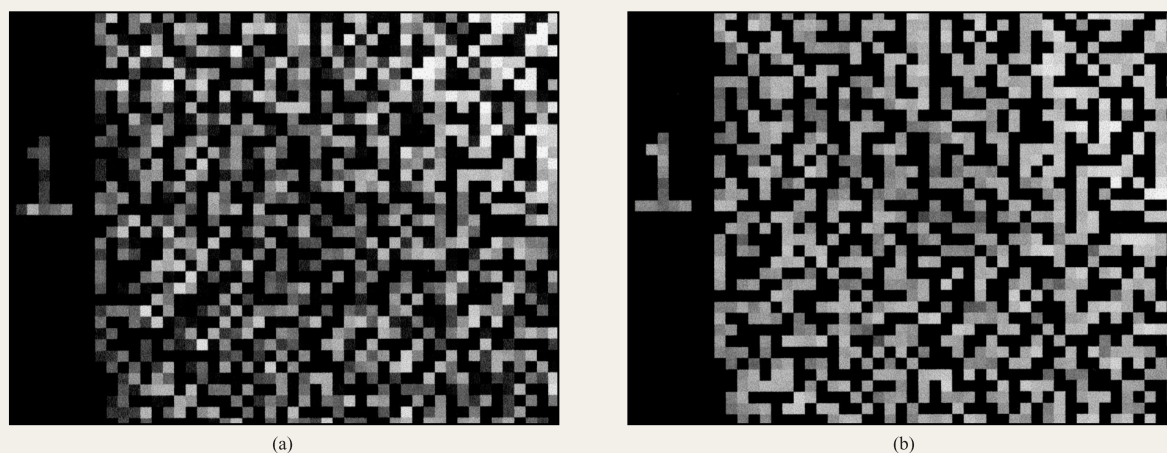
349

(a)   (b)

Figure 11

Portions of data pages holographically reconstructed through a phase aberration (a) without phase-conjugate readout (BER: 5×10^{-2}); (b) with phase-conjugate readout (BER $<10^{-5}$), thus canceling out accumulated phase errors.

(a 1-mm-thick plastic plate). Figure 11(a) shows the data page with only one pass through the plastic plate, demonstrating conventional, non-phase-conjugate readout, while Figure 11(b) demonstrates phase-conjugate readout, where the object beam passes through the plate once during hologram storage and then again upon readout with the phase-conjugate reference beam, correcting the phase aberrations.

One of the practical issues affecting the use of phase-conjugate readout is the need to multiplex the reference beam in order to attain meaningful capacities. Instead of the single pair of reference beams shown in Figure 10, a practical system would require as many as a thousand pairs of reference-beam angles. If the two reference beams are not true phase-conjugate pairs, the differences between them will distort the resulting reconstructed data page. It is not yet clear how a practical system would be able to guarantee this phase-conjugation relationship among many reference beams.

Having discussed the optical components that imprint and detect information, we move to a discussion of coding and signal processing, and the best possible use of these components to record and retrieve digital data from a holographic data storage system.

Coding and signal processing

In a data-storage system, the goal of coding and signal processing is to reduce the BER to a sufficiently low level while achieving such important figures of merit as high density and high data rate. This is accomplished by

stressing the physical components of the system well beyond the point at which the channel is error-free, and then introducing coding and signal processing schemes to reduce the BER to levels acceptable to users. Although the system retrieves raw data from the storage device with many errors (a high raw BER), the coding and signal processing ensures that the user data are delivered with an acceptably low level of error (a low user BER).

Coding and signal processing can involve several qualitatively distinct elements. The cycle of user data from input to output can include interleaving, error-correction-code (ECC) and modulation encoding, signal preprocessing, data storage in the holographic system, hologram retrieval, signal postprocessing, binary detection, and decoding of the interleaved ECC.

The ECC encoder adds redundancy to the data in order to provide protection from various noise sources. The ECC-encoded data are then passed on to a modulation encoder which adapts the data to the channel: It manipulates the data into a form less likely to be corrupted by channel errors and more easily detected at the channel output. The modulated data are then input to the SLM and stored in the recording medium. On the retrieving side, the CCD returns pseudo-analog data values (typically camera count values of eight bits) which must be transformed back into digital data (typically one bit per pixel). The first step in this process is a postprocessing step, called equalization, which attempts to undo distortions created in the recording process, still in the pseudo-analog domain. Then the array of pseudo-

350

analog values is converted to an array of binary digital data via a detection scheme. The array of digital data is then passed first to the modulation decoder, which performs the inverse operation to modulation encoding, and then to the ECC decoder. In the next subsections, we discuss several sources of noise and distortion and indicate how the various coding and signal-processing elements can help in dealing with these problems.

• *Binary detection*
The simplest detection scheme is threshold detection, in which a threshold T is chosen: Any CCD pixel with intensity above T is declared a 1, while those below T are assigned to class 0. However, it is not at all obvious how to choose a threshold, especially in the presence of spatial variations in intensity, and so threshold detection may perform poorly. The following is an alternative.

Within a sufficiently small region of the detector array, there is not much variation in pixel intensity. If the page is divided into several such small regions, and within each region the data patterns are balanced (i.e., have an equal number of 0s and 1s), detection can be accomplished without using a threshold. For instance, in sorting detection, letting N denote the number of pixels in a region, one declares the $N/2$ pixels with highest intensity to be 1s and those remaining to be 0s. This balanced condition can be guaranteed by a modulation code which encodes arbitrary data patterns into codewords represented as balanced arrays. Several such codes are reported in [8, 18]. Thus, sorting detection combined with balanced modulation coding provides a means to obviate the inaccuracies inherent in threshold detection. The price that is paid here is that in order to satisfy the coding constraint (forcing the number of 0s and 1s to be equal), each block of N pixels now represents only M bits of data. Since M is typically less than N, the capacity improvement provided by the code must exceed the code rate, $r = M/N$. For example, for $N = 8$, there are 70 ways to combine eight pixels such that exactly four are 1 and four are 0. Consequently, we can store six bits of data (64 different bit sequences) for a code rate of 75%. The code must then produce a $>33\%$ increase in the number of holographic pages stored, in order to increase the total capacity of the system in bits.

One problem with this scheme is that the array detected by sorting may not be a valid codeword for the modulation code; in this case, one must have a procedure which transforms balanced arrays into valid codewords. This is not much of a problem when most balanced arrays of size N are codewords, but for other codes this process can introduce serious errors. A more complex but more accurate scheme than sorting is correlation detection, as proposed in [8]. In this scheme, the detector chooses the codeword that achieves maximum correlation with the array of received pixel intensities. In the context of the 6:8 code described above, 64 correlations are computed for each code block, avoiding the six combinations of four 1 and four 0 pixels that are not used by the code but which might be chosen by a sorting algorithm.

• *Interpixel interference*
Interpixel interference is the phenomenon in which intensity at one particular pixel contaminates data at nearby pixels. Physically, this arises from optical diffraction or aberrations in the imaging system. The extent of interpixel interference can be quantified by the point-spread function, sometimes called a PSF filter. If the channel is linear and the PSF filter is known, the interpixel interference can be represented as a convolution with the original (encoded) data pattern and then "undone" in the equalization step via a filter inverse to the PSF filter (appropriately called deconvolution). Results on deconvolution with data collected on DEMON I at IBM are described in [19, 20].

Deconvolution has the advantage that it incurs no capacity overhead (code rate of 100%). However, it suffers from mismatch in the channel model (the physics of the intensity detection makes the channel nonlinear), inaccuracies in estimation of the PSF, and enhancement of random noise. An alternative approach to combating interpixel interference is to forbid certain patterns of high spatial frequency via a modulation code. According to the model in [21], for certain realistic and relatively optimal choices of system parameters (in particular at the Nyquist aperture described above [21]), if one forbids a 1 surrounded by four 0s (in its four neighbors on the cardinal points of the compass), areal density can be improved provided that the modulation code has a rate >0.83. Such a code at rate $8:9 = 0.888\ldots$ is described in [22]; in fact, [22] describes such codes of much higher rate, but at the expense of increased complexity.

A code that forbids a pattern of high spatial frequency (or, more generally, a collection of such patterns of rapidly varying 0 and 1 pixels) is called a low-pass code. Such codes constrain the allowed pages to have limited high spatial frequency content. A general scheme for designing such codes is given in [23], via a strip encoding method in which each data page is encoded, from top to bottom, in narrow horizontal pixel strips [24]. The constraint is satisfied both along the strip and between neighboring strips. Codes that simultaneously satisfy both a constant-weight constraint and a low-pass constraint are given in [22].

• *Error correction*
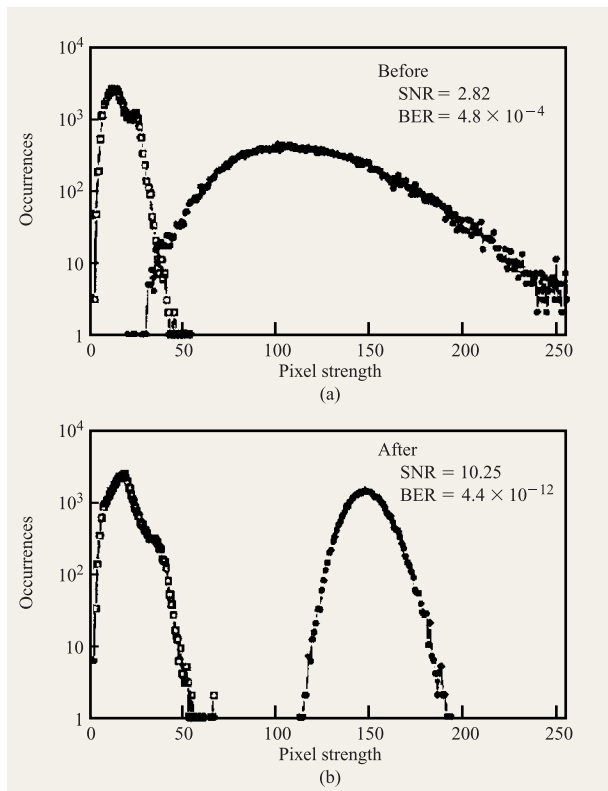In contrast to modulation codes, which introduce a distributed redundancy in order to improve binary

351

Figure 12

Intensity histogram for high-areal-density holograms (a) before and (b) after applying the predistortion technique. Before, inter-pixel crosstalk broadens the brightness distribution; after, these deterministic variations are reduced, improving the BER of the system.

detection of pseudo-analog intensities, error correction incorporates explicit redundancy in order to identify decoded bit errors. An ECC code receives a sequence of decoded data (containing both user and redundant bits) with an unacceptably high raw BER, and uses the redundant bits to correct errors in the user bits and reduce the output user BER to a tolerable level (typically, less than 10^{-12}). The simplest and best-known error-correction scheme is parity checking, in which bit errors are identified because they change the number of 1s in a given block from odd to even, for instance. Most of the work on ECC for holographic storage has focused on more powerful Reed–Solomon (RS) codes [25]. These codes have been used successfully in a wide variety of applications for two reasons: 1) They have very strong error-correction power relative to the required redundancy, and 2) their algebraic structure facilitates the design and implementation of fast, low-complexity decoding algorithms. As a result, there are many commercially available RS chips.

In a straightforward implementation of an ECC, such as an RS code, each byte would be written into a small array (say 2 times 4 for 8-bit bytes), and the bytes in a codeword would simply be rastered across the page. There might be approximately 250 bytes per codeword. If the errors were independent from pixel to pixel and identically distributed across the page, this would work well. However, experimental evidence shows that the errors are neither independent nor identically distributed. For example, interpixel interference can cause an error event to affect a localized cluster of pixels, perhaps larger than a single byte. And imperfections in the physical components can cause the raw BER to vary dramatically across the page (typically, the raw BER is significantly higher near the edges of the page).

Assume for simplicity that our choice of ECC can correct at most two byte errors per codeword. If the codewords are interleaved so that any cluster error can contaminate at most two bytes in each codeword, the cluster error will not defeat the error-correcting power of the code. Interleaving schemes such as this have been studied extensively for one-dimensional applications (for which cluster errors are known as burst errors). However, relatively little work has been done on interleaving schemes for multidimensional applications such as holographic recording. One recent exception is a class of sophisticated interleaving schemes for correcting multidimensional cluster errors developed in [26].

For certain sources of error, it is reasonable to assume that the raw-BER distribution is fixed from hologram to hologram. Thus, the raw-BER distribution across the page can be accurately estimated from test patterns. Using this information, codewords can then be interleaved in such a way that not too many pixels with high raw BER can lie in the same codeword (thereby lowering the probability of decoder failure or miscorrection). This technique, known as matched interleaving, introduced in [27], can yield a significant improvement in user BER.

• *Predistortion*
The techniques we have described above are variations on existing coding and signal-processing methods from conventional data-storage technologies. In addition, a novel preprocessing technique unique to holographic data storage has been developed at IBM Almaden. This technique, called "predistortion" [28], works by individually manipulating the recording exposure of each pixel on the SLM, either through control of exposure time or by relative pixel transmission (analog brightness level on the SLM). Deterministic variations among the ON pixels, such as those created by fixed-pattern noise, nonuniformity in the illuminated object beam, and even interpixel crosstalk, can be suppressed (thus decreasing BER). Many of the spatial variations to be removed are

present in an image transmitted with low power from the SLM directly to the detector array. Once the particular pattern of nonuniform brightness levels is obtained, the recording exposure for each pixel is simply calculated from the ratio between its current brightness value and the desired pixel brightness [28].

At low density, raw-BER improvements of more than 15 orders of magnitude are possible [28]. More significantly, at high density, interpixel crosstalk (which is deterministic once each data page is encoded) can be suppressed and raw BER improved from 10^{-4} to 10^{-12} [28]. **Figure 12** shows this experimental result, implemented on the DEMON I platform with a square aperture of 2.8 mm $\times$ 2.8 mm placed at the Fourier transform plane of the imaging optics. Another use of the predistortion technique is to increase the contrast between the 1 and 0 pixel states provided by the SLM. By using interferometric subtraction while recording the hologram, the amount of light received at the 0 detector pixels can be reduced [28].

- *Gray scale*

The previous sections have shown that the coding introduced to maintain acceptable BER comes with an unavoidable overhead cost, resulting in somewhat less than one bit per pixel. The predistortion technique described in the previous section makes it possible to record data pages containing gray scale. Since we record and detect more than two brightness levels per pixel, it is possible to have more than one bit of data per pixel [29]. The histogram of a hologram with six gray-scale levels made possible by the predistortion technique is shown in **Figure 13**. To encode and decode these gray-scale data pages, we also developed several local-thresholding methods and balanced modulation codes [29].

If pixels take one of g brightness levels, each pixel can convey $\log_2 g$ bits of data. The total amount of stored information per page has increased, so gray-scale encoding appears to produce a straightforward improvement in both capacity and readout rate. However, gray scale also divides the system's signal-to-noise ratio (SNR) into $g - 1$ parts, one for each transition between brightness levels. Because total SNR depends on the number of holograms, dividing the SNR for gray scale (while requiring the same error rate) leads to a reduction in the number of holograms that can be stored. The gain in bits per pixel must then outweigh this reduction in stored holograms to increase the total capacity in bits [29].

- *Capacity estimation*

To quantify the overall storage capacity of different gray-scale encoding options, we developed an experimental capacity-estimation technique [12]. In this technique, the dependence of raw BER on readout power is first measured experimentally. A typical curve is shown in
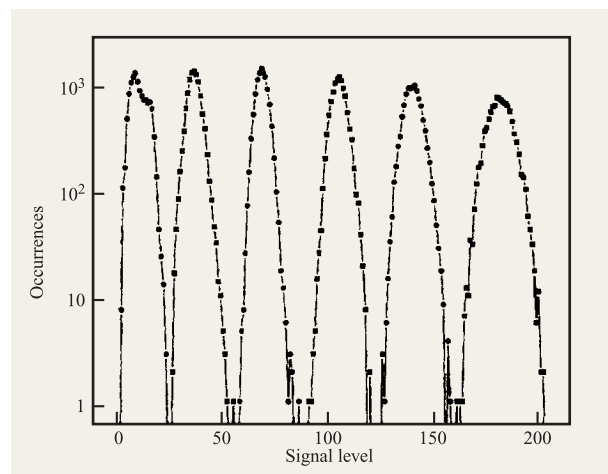
Histogram of a hologram with six gray-scale levels recorded on the DEMON I platform using the predistortion technique. From [29], reproduced with permission.

Figure 14(a). The capacity-estimation technique then produces the relationship between M, the number of holograms that can be stored, and raw BER [**Figure 14(b)**]. Without the capacity-estimation technique, producing Figure 14(b) would require an exhaustive series of multiple hologram experiments.

In general, as the raw BER of the system increases, the number of holograms, M, increases slowly. In order to maintain a low user BER (say, 10^{-12}) as this raw-BER operating point increases, the redundancy of the ECC code must increase. Thus, while the number of holograms increases, the ECC code rate decreases. These two opposing trends create an "optimal" raw BER, at which the user capacity is maximized [30]. For the Reed–Solomon ECC codes we commonly use [29], this optimal raw BER is approximately 10^{-3}. By computing these maximum capacities for binary data pages and gray-scale data pages from $g = 2$ to $g = 6$, we were able to show that gray-scale holographic data pages provide an advantage over binary encoding in both capacity and readout rate. The use of three gray levels offered a 30% increase in *both* capacity and readout rate over conventional binary data pages.

Associative retrieval

As mentioned in the Introduction, volume holographic data storage conventionally implies that data imprinted on an object beam will be stored volumetrically [**Figure 15(a)**], to be read out at some later time by illumination with an addressing reference beam [**Figure 15(b)**].
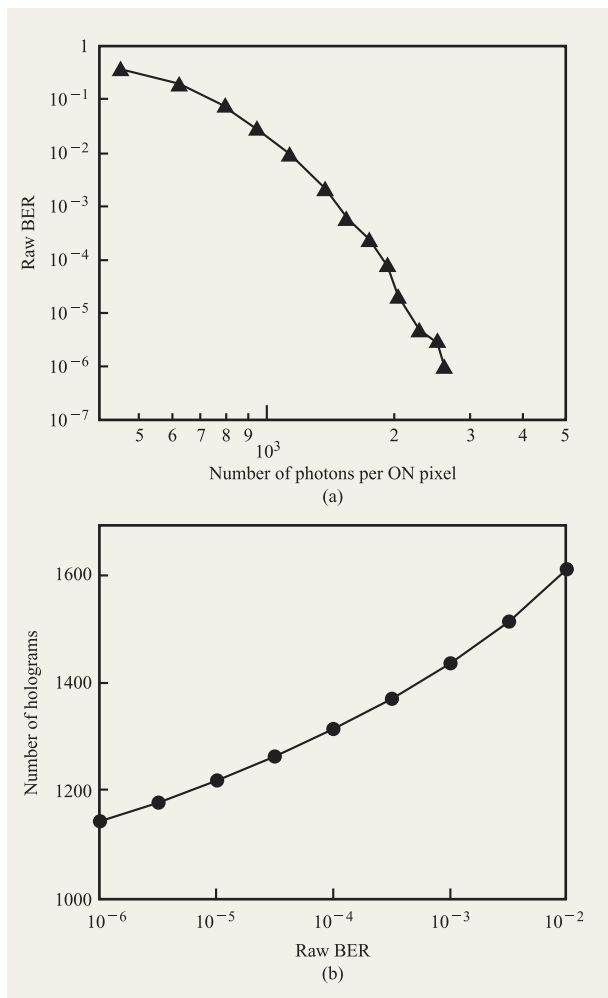
353

Figure 14

Capacity-estimation technique begins with (a) a simple experimental measurement of raw BER of a few holograms as a function of the reconstructed intensity, and produces (b) an estimation of the number of holograms that could be superimposed as a function of the raw BER that the system is asked to maintain. Without this technique, one would need to perform repeated multiple-hologram experiments to obtain these data.

However, the same hologram (the interference pattern between a reference beam and a data-bearing object beam) can also be illuminated by the *object* beam [**Figure 15(c)**]. This reconstructs all of the angle-multiplexed reference beams that were used to record data pages into the volume. The amount of power diffracted into each "output" beam is proportional to the 2D cross-correlation between the input data page (being displayed on the SLM) and the stored data page (previously recorded with that particular reference beam). Each set of output beams can be focused onto a detector array, so that

each beam forms its own correlation "peak." Because both the input and output lenses perform a two-dimensional Fourier transform in spatial coordinates

Figure 15

Holographic data storage system: (a) Two coherent beams, one carrying a page of information, interfere within a photosensitive material to record a hologram. (b) Illuminating the hologram with the reference beam reconstructs a weak copy of the original information-bearing beam for capture with a detector array. (c) Illuminating multiple stored holograms with a new page of search information reconstructs all of the reference beams, computing in parallel the correlation between the search data and each of the stored pages. From [33], reproduced with permission.

[5], the optical system is essentially multiplying the Fourier transforms of the search page and each data page and then taking the Fourier transform of this product (thus implementing the convolution theorem optically). Because of the volume nature of the hologram, only a single slice through the 2D correlation function is produced (the other dimension has been "used" already, providing the ability to correlate against multiple templates simultaneously).

The center of each correlation peak represents the 2D inner product (the simple overlap) between the input page being presented to the system and the associated stored page. If the patterns which compose these pages correspond to the various data fields of a database, and each stored page represents a data record, the optical correlation process has just simultaneously compared the entire database against the search argument [31–33]. This parallelism gives content-addressable holographic data storage an inherent speed advantage over a conventional serial search, especially for large databases. For instance, if an unindexed conventional "retrieve-from-disk-and-compare" software-based database is limited only by sustained hard-disk readout rate (25 MB/s), a search over one million 1 KB records would take ~40 s. In comparison, with off-the-shelf, video-rate SLM and CCD technology, an appropriately designed holographic system could search the same records in ~30 ms — a 1200× improvement. Custom components could enable 1000 or more parallel searches per second.

For this optical correlation process to represent a database search, the spatial patterns of bright (ON) pixels on the holographic data pages must somehow represent the digital data from fixed-length database fields. The SLM is divided into separate regions, each dedicated to a particular fixed-length field of the database. For example, a two-bit data field might be encoded by four blocks of pixels at a particular point within the SLM page. Such an encoding implements an exact search through the database. By thresholding the detected optical signal (essentially an analog quantity), any matching records are identified. Thresholding becomes commensurately more difficult, however, when many fields are being searched simultaneously. And when the threshold does not work correctly, completely unrelated records are identified as matches because near matches between pixel block patterns do not represent near matches in encoded data value.

We have developed a novel data-encoding method which allows similarity or fuzzy [34] searching, by encoding similar data values into similar pixel block patterns. As shown in **Figure 16(a)**, data values are encoded by the position of a block of ON pixels within a vertical track, creating a "slider" (like the control found on a stereo's graphic equalizer, for instance). As an example, the data value 128 might be encoded as a pixel
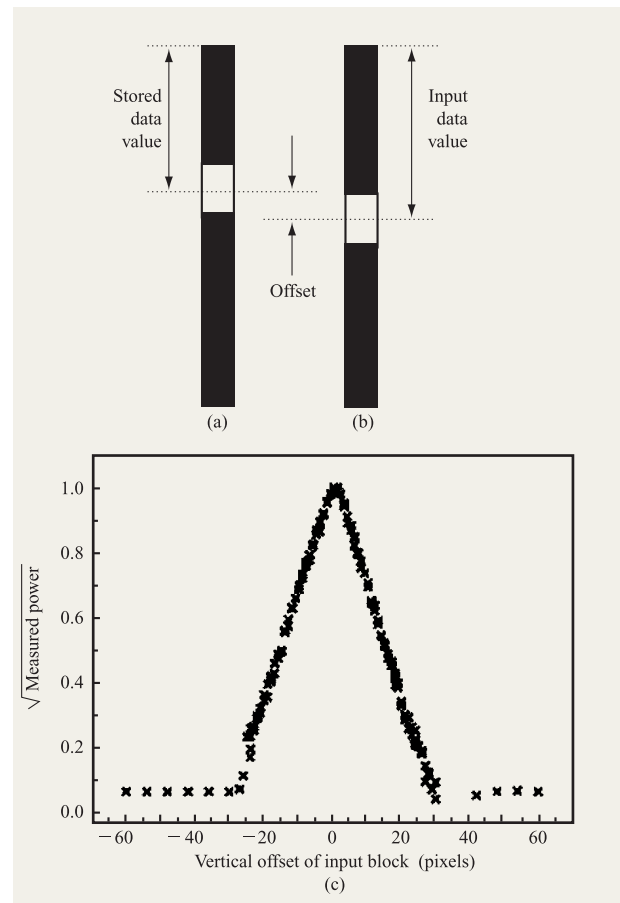

Figure 16

Data encoding for fuzzy searching: (a) When storing a hologram, a small block of SLM pixels are turned ON at some location within a predefined rectangular portion ("slider" track) of the data page. (b) For correlation readout, an input query is encoded as a similar block within the *same* track. (c) Any offset between the two blocks causes the brightness of the correlation peak to decrease. By encoding data values with the center position of the pixel block, the holographic system can now measure the similarity between data records and the input query, implementing fuzzy searching. From [33], reproduced with permission.

block of height h_s, centered within a column of 256 pixels. During the search for data values near 128, the partial overlap between the input slider block [**Figure 16(b)**] and the stored slider block causes the resulting correlation peak to indicate the similarity between the input query and the stored data. The holographic content-addressable system is optically measuring the inner product between an input data page (containing a pixel block at some position along this slider column), and each stored page (possibly containing a pixel block at the same position in the same slider column). This is the same result that

355

would be produced by cutting holes at nearly the same spot on two sheets of black cardboard, aligning their edges, and then holding them up to a light. The holographic system is merely condensing this partial overlap into a single intensity result, and is performing the same test on a large number of holograms simultaneously.

More compact data representations can be realized by combining both fuzzy and exact search encodings. The higher-order bits would be encoded compactly with binary-type encoding, while the low-order bits remained available for fuzzy searching. This trades search flexibility for more capacity (in terms of fields per database record). By adding a correlation camera to the DEMON I platform, we experimentally demonstrated this fuzzy search encoding. **Figure 16(c)** shows results from a search of a single fuzzy-encoded data field as the input data value approached and then exceeded the stored value. The amplitude response (the square root of measured power as a function of the relative position of the input slider block) formed a triangularly shaped function [35]. The correlation of identical rectangles creates the triangle; the signals add in field amplitude yet are detected in intensity; thus, this triangle shows up after taking the square root of the measured signals. With this fuzzy encoding technique, the analog nature of the optical output becomes an enabling feature instead of a drawback.

To demonstrate high-fidelity parallel searching of a holographic content-addressable memory, we stored a small multimedia database in our modified DEMON I system. Each hologram represented one record from an IBM query-by-image-content (QBIC) database [36]. In the QBIC system, searches are performed across feature vectors previously extracted from the images, rather than on the images themselves. Each record included several alphanumeric fields (such as image description and image number) encoded for exact searches, and 64 fuzzy sliders containing the color histogram information (percentage of each given color within the associated image). A separate portion of the SLM page, pixel-matched onto a CCD detector for conventional address-based holographic readout, was encoded with the binary data for the small binary image [33]. One hundred holograms were recorded in a 90-degree-geometry [8] $LiNbO_3$ crystal, with the reference angles chosen so that each reference beam was focused to a unique portion of the correlation camera [33].

Each search, initiated by a user query, ran under computer control, including display of the appropriate patterns, detection of the correlation peaks (averaging eight successive measurements to reduce detector noise), calibration by hologram strength, identification of the eight highest correlation scores, mapping of correlation bins to reference-beam angle, address-based recall of these eight holograms, decoding of the pixel-matched data pages, and, finally, display of the binary images on the computer monitor. The optical readout portion occupied only 0.25 s of the total ~5-s cycle time. To find images based on color similarity, the 64 sliders were used to input the color histogram information for the upper left image in **Figure 17(a)**. The slider patterns for this color histogram were input to the system on the SLM, resulting in 100 reconstructed reference beams. After detection, calibration, and ranking of these 100 correlation peaks, the reference beams for the brightest eight were input to the system again, resulting in eight detected data pages and thus eight decoded binary images. Figure 17(a) shows the first four of these images, indicating that the holographic search process found these images to be those which most closely matched the color histogram query. **Figure 17(b)** quantifies the search fidelity by plotting the detected correlation peak intensity as a function of the overlap between the object-beam search patterns. Perfect system performance would result in a smooth monotonic curve; however, noise in the real system introduces deviations away from this curve. As expected, the feature vector for the left-hand image correlated strongly with itself, but the system was also able to correctly identify the images with the highest cross-correlation.

These sliders could also be used to select images by color distribution [33]. **Figures 17(c)** and **17(d)** correspond to a search for images containing 20% white and 20% light gray. Although several images were ranked slightly higher than they deserved (red circle), the system performance was impressive, considering that the background "dark" signal was twice as large as the signal. In **Figures 17(e)** and **17(f)**, the alphanumeric description field was used to search for the keyword *shore*. Note that because many characters are involved, both the expected and measured scores are large. However, we obtained similar results for exact search arguments as small as a single character.

With the fuzzy coding techniques we have introduced, volume holographic content-addressable data storage is an attractive method for rapidly searching vast databases with complex queries. Areas of current investigation include implementing system architectures which support many thousands of simultaneously searched records, and quantifying the capacity–reliability tradeoffs.

Recording materials

● *Materials and media requirements for holographic data storage*
Thus far, we have discussed the effects of the hardware, and of coding and signal processing, on the performance

Figure 17

Three experimental search results from an all-holographic search-and-retrieve engine, operating on a database of 100 feature vectors from the IBM Query-by-Image-Content (QBIC) image database [36]: (a) The best four images found when the search query was the color feature vector for the leftmost image. (b) Measured correlation score (ratio of the detected signal to the "dark" calibration value), for each of the 100 database records, as a function of the expected response (number of SLM pixels in common between the input and each stored page). (c) The best four images found when the color sliders for 20% white and 20% light gray were input. (d) Measured vs. expected correlation score. (e) The best four images found when searching for the key word *shore*, encoded into five characters with three nonbinary subfields per character. (f) Measured vs. expected correlation score. From [33], reproduced with permission.

of holographic data storage systems. Desirable parameters described so far include storage capacity, data input and output rates, stability of stored data, and device compactness, all of which must be delivered at a specified (very low) user BER. To a large extent, the possibility of delivering such a system is limited by the properties of the materials available as storage media [37]. The connections between materials properties and system performance are complex, and many tradeoffs are possible in adapting a given material to yield the best results. Here we attempt to outline in a general way the desirable properties for a holographic storage medium and give examples of some promising materials.

Properties of foremost importance for holographic storage media can be broadly characterized as "optical quality," "recording properties," and "stability." These directly affect the data density and capacity that can be achieved, the data rates for input and output, and the BER.

As mentioned above, for highest density at low BER, the imaging of the input data from the SLM to the detector must be nearly perfect, so that each data pixel is read cleanly by the detector. The recording medium itself is part of the imaging system and must exhibit the same high degree of perfection. Furthermore, if the medium is moved to access different areas with the readout beam, this motion must not compromise the imaging performance. Thus, very high standards of optical homogeneity and fabrication must be maintained over the full area of the storage medium. With sufficient materials development effort and care in fabrication, the necessary optical quality has been achieved for both inorganic photorefractive crystals and organic photopolymer media. As discussed above, phase-conjugate readout could ultimately relax these requirements.

A more microscopic aspect of optical quality is intrinsic light scattering of the material. The detector noise floor produced by scattering of the readout beam imposes a fundamental minimum on the efficiency of a stored data hologram, and thus on the storage density and rate of data readout [38]. Measurements on the PRISM tester have shown that, in general, the best organic media have a higher scattering level than inorganic crystals, by about a factor of 100 or more.

Because holography is a volume storage method, the capacity of a holographic storage system tends to increase as the thickness of the medium increases, since greater thickness implies the ability to store more independent diffraction gratings with higher selectivity in reading out individual data pages without crosstalk from other pages stored in the same volume. For the storage densities necessary to make holography a competitive storage technology, a media thickness of at least a few millimeters is highly desirable. In some cases, particularly for organic

materials, it has proven difficult to maintain the necessary optical quality while scaling up the thickness, while in other cases thickness is limited by the physics and chemistry of the recording process.

Holographic recording properties are characterized in terms of sensitivity and dynamic range. Sensitivity refers to the extent of refractive index modulation produced per unit exposure (energy per unit area). Diffraction efficiency (and thus the readout signal) is proportional to the square of the index modulation times the thickness. Thus, recording sensitivity is commonly expressed in terms of the square root of diffraction efficiency, η:

$$S_{\eta_2} = (\eta^{1/2})/(I\ell t), \tag{1}$$

where I is the total intensity, ℓ is the medium thickness, and t is the exposure time; this form of sensitivity is usually given in units of cm/J. Since not all materials used are the same thickness, it is a more useful comparison to define a modified sensitivity given by the usual sensitivity times the thickness:

$$S'_{\eta_2} = S_{\eta_2} \times \ell. \tag{2}$$

This quantity has units of cm^2/J and can be thought of as the inverse of the writing fluence required to produce a standard signal level. The unprimed variable, S_{η_2}, might be used to convey the potential properties of a storage material, given that the particular sample under test is extremely thin; in contrast, S'_{η_2} quantifies the ability of a specific sample to respond to a recording exposure.

For high output data rate, one must read holograms with many pixels per page in a reasonably short time. To read a megapixel hologram in about 1 ms with reasonable laser power and to have enough signal at the detector for low error rate, a diffraction efficiency around $\eta = 3 \times 10^{-5}$ is required. To write such a hologram in 1 ms, to achieve input and output data rates of 1 Gb/s, the sensitivity for this example must be at least $S'_{\eta_2} = 20$ cm^2/J.

The term *dynamic range* refers to the total response of the medium when it is divided up among many holograms multiplexed in a common volume of material; it is often parameterized as a quantity known as *M#* (pronounced "M-number" [39]), where

$$M\# = \sum \eta^{1/2}, \tag{3}$$

and the sum is over the M holograms in one location. The M# also describes the scaling of diffraction efficiency as M is increased, i.e.,

$$\eta = (M\#/M)^2. \tag{4}$$

Dynamic range has a strong impact on the data storage density that can be achieved. For example, to reach a density of 100 bits/μm^2 (64 Gb/in.^{2}) with megapixel data pages, a target diffraction efficiency of 3×10^{-5}, and area

Table 1 Comparison of properties of prospective materials for holographic data storage media.

Material	Image quality	Scatter	Holographic fidelity	S'_{η_2} (cm^2/J)	M#	Stability	Thickness (mm)
LiNbO$_3$:Fe	+++	+++	+	0.02	1	0	10
LiNbO$_3$ (Two-color)	++	++	+	0.02*	1*	++	10
Polaroid photopolymer	+++	−	0	20	1.5	+	0.5
PQ/PMMA	+	−	+	0.2–0.5	1	++	2
Bayer photo-addressable polymer	+++	0	++	0.002–0.02			0.1

*Values depend on writing intensity (see Table 2).

at the medium of 0.1 cm^2 would require $M\# = 5$, a value that is barely achievable with known recording materials under exposure conditions appropriate for recording high-fidelity data holograms.

Stability is a desirable property for any data storage system. In the case of holographic storage, the response of the recording medium, which converts the optical interference pattern to a refractive index pattern (the hologram), is generally linear in light intensity and lacks the response threshold found in bistable storage media such as magnetic films. In the case of write-once-read-many (WORM) media such as photopolymers, the material response is irreversible; once the material has been fully exposed, further optical irradiation produces no further response, and the data can be interrogated by the readout beam without erasing it or distorting it. Much basic research in holographic storage has been performed using photorefractive crystals as storage media (e.g., [2, 8, 40], the experiments described above). Of these crystals, Fe-doped lithium niobate has been the workhorse. Its sensitivity is sufficient for demonstration purposes, but lacks a factor of 100 for practical application. Since photorefractives are reversible materials, they suggest the possibility of a rewritable holographic storage medium. However, because they are linear and reversible, they are subject to erasure during readout. Several schemes have been investigated for stabilizing or "fixing" the recording so that the data can be read without erasure. One scheme that does this without compromising the ability to erase the data, known as two-color recording, has received a good deal of attention recently [41–43]. Recording is enabled by simultaneous irradiation of the crystal by a gating beam of different wavelength than the usual object and reference beams. In the absence of the gating wavelength, the data can be read without causing erasure. More details are given in the next section.

Stability in the dark over long periods is also an issue; organic photopolymer materials are often subject to aging processes caused by residual reactive species left in the material after recording or by stresses built up in the material during recording. Erasure may occur because of residual thermal diffusion of the molecules which record the hologram. Index modulation in photorefractives results from a space charge that is built up by the optical excitation and migration of mobile charge carriers. Stability in the dark depends on the trapping of these carriers with trap energies that are not thermally accessible at room temperature.

Many kinds of materials have been investigated as holographic storage media. **Table 1** is a comparison of the properties of several that are among the best available as data storage media. Five materials are compared on the basis of optical imaging quality, scattered light level, hologram fidelity, sensitivity, M#, stability, and available thickness. These include the much-studied Fe-doped lithium niobate, two-color recording in reduced stoichiometric lithium niobate [41], and three organic materials that were chosen to typify the range of properties available from various organic materials systems. Photopolymers [44–47] are very promising because of their high sensitivity and dynamic range; they are discussed in more detail below. Phenanthrenequinone-doped polymethylmethacrylate (PQ/PMMA) [48] has excellent optical quality and is based on a photoreaction between the dopant and polymer followed by diffusion of unreacted chromophore; this requires a long thermal treatment, which is a disadvantage from a system perspective. Finally, photo-addressable polymers [49] are also promising but are still at an early stage of development.

• *Summary of polymer work*
Polymer materials are important candidates for holographic storage media. They promise to be inexpensive to manufacture while offering a wide variety of possible recording mechanisms and materials systems. The opportunity for fruitful development of polymer holographic media is thus very broad, and a variety of approaches to using organic materials for holography have

359

been pursued, including organic photorefractive materials [50, 51], triplet-sensitized photochromic systems [52], photo-addressable polymers [49], and materials which produce index modulation via material diffusion. Of the latter class, PQ/PMMA is a polymer glass in which a photoreaction binds the phenanthrenequinone chromophore to the PMMA [48]. During a thermal treatment, typically for about 24 hours, unbound PQ diffuses, and the resulting concentration gradients are frozen in place by a final uniform illumination that binds the remaining unreacted chromophore to the PMMA backbone, leading to a fixed hologram. This material has the excellent optical quality of the PMMA matrix, it is available in reasonable thickness, and its sensitivity, while somewhat low, is reasonably good. However, the current need for lengthy thermal treatment makes it unacceptable for most storage applications.

The diffusion-driven photopolymer systems [44–47] offer very high sensitivity and need no such postexposure processing. The basic mechanism is a photosensitized polymerization, coupled with diffusion of monomer and other components of the material formulation under influence of the resulting concentration gradients. The medium is usually partially prepolymerized to produce a gel-like matrix, allowing rapid diffusion at room temperature. Refractive index modulation and recording of holograms result from both the density change and the difference in polarizability of the polymerized material. The magnitude of this refractive index modulation can be very high, resulting in a high dynamic range. For simple plane-wave holograms, an M# as high as 42 has been observed [47]. For digital data holograms, the contrast of the interference pattern between object and reference beams is lower than in the plane-wave case, and the recording conditions do not produce as large an index modulation. Even so, the M# observed for digital holograms on the PRISM materials tester is around 1.5, one of the highest yet observed; this value can undoubtedly be improved by optimization of the recording conditions.

The recording mechanism for photopolymers also leads to some disadvantages, including the shrinkage of the material with polymerization and the possibility of nonlinear response. Both of these distort the reconstructed holograms and thus cause errors in decoding the digital data. For some photopolymers, significant advances have been made toward eliminating these undesired properties; for example, shrinkage has been reduced to less than 0.1% while sufficient useful dynamic range for recording of data has been retained [44]. There are additional problems in increasing the thickness of these materials to the millimeter scale that is desirable for holography, and even then the Bragg angle selectivity is not sufficient to allow enough holograms to be written in a common volume to achieve high data density. However, through the use of nonselective multiplexing methods, it is possible to increase the density to a competitive level. One of these methods, known as peristrophic multiplexing [53], involves the rotation of the medium about an axis normal to its plane such that the reconstructed hologram image rotates away from the detector, allowing another hologram to be written and read. We have recently demonstrated the recording and readout with very low error rate of 70 holograms of 256 Kb each on the PRISM tester, using a combination of Bragg angle and peristrophic multiplexing.

Photopolymer materials have undergone rapid development and show great potential as write-once holographic media. Because of this rapid development, there is relatively little research addressing the issue of long-term data integrity and stability after recording. Work in this area is ongoing.

Another class of organic materials undergoing rapid development is the photo-addressable polymer systems [49]. These systems incorporate azo-dye chromophores that are highly optically anisotropic and that undergo optically induced reorientation. Thus, optical irradiation produces a large refractive index change through the birefringence induced by this reorientation process. The index change can be stabilized by incorporating the chromophores into a polymer matrix containing liquid crystal components. At this point, these materials lack a convenient means of desensitization once the data have been written, so that they do not saturate and overwrite the holograms during readout. However, the index change available via this mechanism is very large; a recording medium of this type could have very high dynamic range, and thus the potential for high data storage density, and perhaps be reversible, thus enabling rewritable storage.

The best of the photopolymers are promising as storage media for WORM data storage. The photorefractive crystals have traditionally been the favorite candidates for reversible, rewritable storage; recent work on two-color recording has shown the way to a possible solution of the volatility of reversible media during readout. The following section describes this concept.

• *Two-color or photon-gated holography*
Two main schemes for providing nondestructive readout have been proposed, both in lithium niobate, although the concepts are applicable to a broader range of materials. The first was thermal fixing [54, 55], in which a copy of the stored index gratings is made by thermally activating proton diffusion, creating an optically stable complementary proton grating. Because of the long times required for thermal fixing and the need to fix large blocks of data at a time, thermally fixed media somewhat resemble reusable WORM materials. Another class of fixing process uses two wavelengths of light.
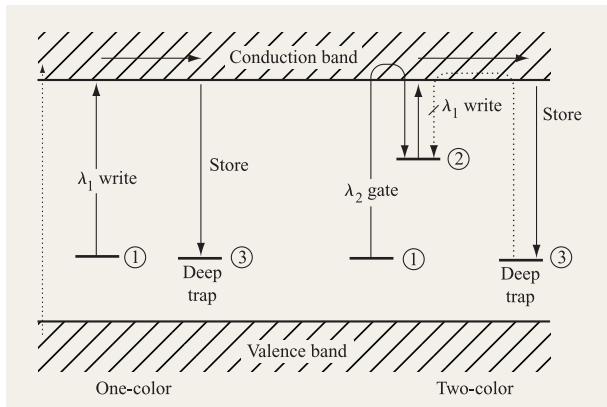
One approach uses two different wavelengths of light for recording and reading [56], but for storage applications this suffers from increased crosstalk and restrictions on the spatial frequencies that can be recorded. The most promising two-color scheme is "photon-gated" recording in photorefractive materials, in which charge generation occurs via a two-step process [57]. Coherent object and reference beams at a wavelength λ_1 record information in the presence of gating light at a wavelength λ_2. The gating light can be incoherent or broadband, such as a white-light source or LED [58]. Reading is done at λ_1 in the absence of gating light. Depending on the specific implementation, either the gating light acts to sensitize the material, in which case it is desirable for the sensitivity to decay after the writing cycle, or the gating light ionizes centers in which a temporary grating can be written at the wavelength λ_1. **Figure 18** shows a schematic of energy levels comparing the two-color and one-color schemes for a photorefractive material with localized centers in the bandgap. A very important and unique figure of merit for photon-gated holography is the gating ratio, the ratio between the sensitivity of the material in the presence and absence of gating light.

Reduced stoichiometric lithium niobate shows both one-color sensitivity in the blue-green spectral region and two-color sensitivity for writing in the near IR and gating with blue-green light [41, 59–61]. From this it can be seen that the gating light also produces erasure. This is a consequence of the broad spectral features of
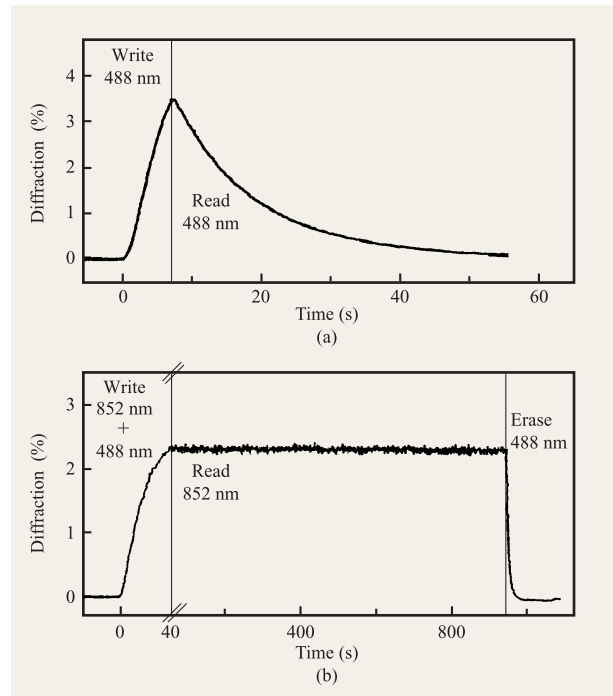
reduced or Fe-doped lithium niobate. Considerable progress is envisaged if a better separation of gating and erasing functions can be achieved by storing information in deeper traps and/or using wider-bandgap materials. **Figure 19** compares one-color and two-color writing in a sample of reduced, near-stoichiometric lithium niobate to illustrate the nondestructive readout that can be achieved. The gating ratio in this case was in excess of 5000.

Conventionally, lithium niobate is grown in the congruent melting composition, expressed by the quantity $c_{Li} = [Li]/([Li] + [Nb]) = 48.5\%$, because the identical compositions of the melt and the crystal promote high optical quality and large boules. Crystals of nominally undoped lithium niobate, grown with a stoichiometry (SLN) of 49.7% by a special double-crucible technique [62], were compared with those of the congruent composition (CLN). Strong differences were observed,

361

Table 2 Summary of data and comparison of two-color and one-color results, for stoichiometric (SLN) and congruent (CLN) lithium niobate.

Material	Recording scheme	Fe concentration (ppm)	$10^3\,S_{\eta_2}$ (incident) (cm^2/J)	$10^3\,S_{\eta_1}$ (absorbed) (cm^2/J)	M#/cm**	Gating ratio @ 852 nm
Reduced SLN	Two-color* 852 + 488	1.0	8	160	0.8	1600
Reduced SLN + Fe	Two-color* 852 + 488	100	9	150	0.5	10,000
CLN	Two-color* 852 + 488	residual	0.02	>20	0.05	—
Reduced CLN + Fe	One-color 488 nm	200	100	170	24	N/A

*I_w = 4 W/cm^2, 852 nm; I_g = 1 W/cm^2, 488 nm; Λ = 6 μm; E parallel to c-axis.
**For plane-wave, small-angle geometry.

as shown in **Table 2**. Materials were evaluated in a plane-wave geometry in which two collimated 852-nm beams from a single-frequency diode laser were incident on the sample at an external crossing angle of 20 degrees. Gating light was provided either by an Ar$^+$ laser at 488 nm or by several GaN LEDs. Further details of the experimental setup were recently published [41].

Reduction of lithium niobate (heat treatment in an oxygen-poor atmosphere) induces a broad visible absorption band. This band is attributed primarily to absorption by a bipolaron consisting of an electron trapped on a regular Nb site and another trapped at a Nb$_{Li}$ antisite, together with a strong lattice distortion [63]. In addition, there is some contribution to the band from residual impurities such as Fe^{2+}. Irradiating with blue-green light is the gating or sensitizing step, which produces a transient absorption around 1.6 eV [64]. This absorption is assigned to a small polaron, or electron trapped at Nb$_{Li}$, produced by dissociation of the bipolaron [65], and is responsible for the sensitivity at 852 nm.

As we have seen, the most important photorefractive properties for two-color holographic data storage are the gating ratio (measuring the degree of nonvolatility), sensitivity, M# or dynamic range, dark decay, and optical quality. Table 2 shows most of these properties for stoichiometric and congruent compositions compared to the behavior of conventional one-color Fe-doped lithium niobate. Photorefractive sensitivity for two-color recording in lithium niobate is linear in the gating light intensity, I_g, only at low values of I_g because of competition between gating and erasing. Hence, the sensitivity in terms of incident intensities S_{η_2} is defined similarly to that for one-color processes [see Equation (2)], but for a fixed and reasonably low value of I_g = 1 W/cm^2. The sensitivity in terms of absorbed power is $S_{\eta_1} = S_{\eta_2}/\alpha$, where α is the absorption coefficient at the writing wavelength. In terms of this sensitivity, all samples studied, including the single photon Fe-doped material written at 488 nm, are almost

equally sensitive. This suggests that the sensitivity is determined by the amount of light that can be absorbed at the writing wavelength. So far, the maximum absorption of writing light that we have found in reduced SLN is 6% for I_g = 1 W/cm^2.

Summarizing the results of Table 2, the sensitivity gains for two-color recording in reduced, nearly stoichiometric lithium niobate with respect to the congruent material are 15× for increased stoichiometry and 20× for degree of reduction. In addition, lowering the gating wavelength from 520 nm to 400 nm gains a further factor of 10, and cooling from 20°C to 0°C a factor of 5.

There is an interesting difference in the behavior of one- and two-color materials with regard to dynamic range. In a one-color material, the M# is proportional to the modulation index or fringe visibility of the optical interference pattern, $m = 2(I_1 I_2)^{1/2}/(I_1 + I_2)$. However, in a two-color material, the writing light ($I_1 + I_2$) does not erase the hologram, and the M# is proportional to $(I_1 I_2)^{1/2}$. As a result, for object and reference beams of equal intensity, the M# is proportional to the writing intensity. While this provides a general way of increasing the dynamic range in a two-color material, the writing power requirements in the present material system become rather high in order to achieve a substantial increase in M#.

Instead of amplifying the role of the intrinsic shallow levels with stoichiometry, an alternative scheme for implementing two-color holography in lithium niobate is the introduction of two impurity dopants [66, 67]. One trap, such as Mn, serves as the deep trap from which gating occurs, while a more shallow trap, such as Fe, provides the more shallow intermediate level for gated recording. While this scheme provides more opportunities for tuning through choice of dopants, in general it is difficult in LiNbO$_3$ to separate the two absorption bands enough to provide high gating ratios and thus truly nonvolatile storage. In addition, while M# improves

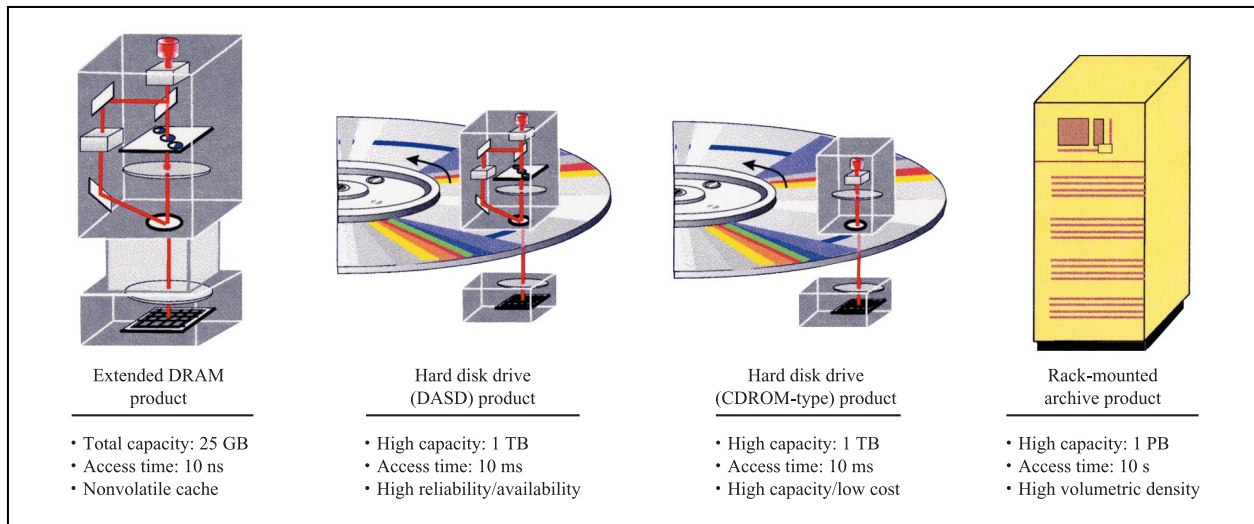| Extended DRAM product	Hard disk drive (DASD) product	Hard disk drive (CDROM-type) product	Rack-mounted archive product
• Total capacity: 25 GB • Access time: 10 ns • Nonvolatile cache	• High capacity: 1 TB • Access time: 10 ms • High reliability/availability	• High capacity: 1 TB • Access time: 10 ms • High capacity/low cost	• High capacity: 1 PB • Access time: 10 s • High volumetric density

Figure 20

Four scenarios highlighting the properties of holographic data storage: an all-solid-state memory module, which takes advantage of the potential for short access times; two rotating-disk geometries, with either erasable or WORM-type media; and, finally, a data warehouse with removable media. With its high volumetric density, holographic data storage has the potential to affect all types of data storage.

monotonically with writing intensity for stoichiometric lithium niobate, with the two-trap method M# is maximized at a particular writing intensity, thus creating an undesirable tradeoff between recording rate and dynamic range.

Two-color, photon-gated holography provides a promising solution to the long-standing problem of destructive readout in read/write digital holographic storage. In lithium niobate, optimization of the sensitivity requires control over stoichiometry (or doping), degree of reduction, temperature, gating wavelength, and gating intensity. Two-color materials differ fundamentally from one-color materials in that the dynamic range or M# can be increased by using higher writing intensity, and the sensitivity can be increased with higher gating intensity. Another route to increasing the M# would be to find a material which exhibits a two-color erase process. Substantial progress has been made in recent years in the field of two-color holography, and further progress can be expected on this complex and challenging problem.

Outlook

Holographic data storage has several characteristics that are unlike those of any other existing storage technologies. Most exciting, of course, is the potential for data densities and data transfer rates exceeding those of magnetic data storage. In addition, as in all other optical data storage

methods, the density increases rapidly with decreasing laser wavelength. In contrast to surface storage techniques such as CD-ROM, where the density is inversely proportional to the square of the wavelength, holography is a volumetric technique, making its density proportional to one over the third power of the wavelength. In principle, laser beams can be moved with no mechanical components, allowing access times of the order of 10 μs, faster than any conventional disk drive will ever be able to randomly access data. As in other optical recording schemes, and in contrast to magnetic recording, the distances between the "head" and the media are very large, and media can be easily removable. In addition, holographic data storage has shown the capability of rapid parallel search through the stored data via associative retrieval.

On the other hand, holographic data storage currently suffers from the relatively high component and integration costs faced by any emerging technology. In contrast, magnetic hard drives, also known as direct access storage devices (DASD), are well established, with a broad knowledge base, infrastructure, and market acceptance. Are there any scenarios conceivable for holographic data storage, where its unique combination of technical characteristics could come to bear and overcome the thresholds faced by any new storage technology?

Four conceivable product scenarios are shown in **Figure 20**. The first two scenarios use read/write media,

363

while the latter two are designed for WORM materials, which are much easier to develop but must support data retention times as long as tens of years. The first scenario [Figure 20(a)] takes advantage of rapid optical access to a stationary block of media, resulting in a random-access time of the order of 10 μs. The capacity is limited to about 25 GB by the size of the block of media that can be addressed by simple, inexpensive optics. Such a device could bridge the gap between conventional semiconductor memory and DASD, providing a nonvolatile holographic cache with an access time that is between DASD and dynamic random-access memory (DRAM).

Using the same optical components but replacing the stationary block of media with a rotating disk results in performance characteristics similar to those of a disk drive, albeit with terabytes (10^{12} bytes) of capacity per platter [Figure 20(b)]. In the CD-ROM type of embodiment [Figure 20(c)], holographic data storage takes advantage of the fact that single-exposure full-disk replication has been demonstrated [68]. The player for the holographic ROM is conceptually very simple: The photodiode from a conventional ROM player is replaced by a CMOS camera chip, and the reconstructed data page is then imaged with suitable optics onto that camera.

Combining one of the DASD-type R/W heads and possibly a number of CD-ROM-type readers, a robotic picker, and sufficient tiles of media, a data warehouse with petabyte (10^{15} bytes) capacity in a standard 19-inch rack is conceivable [Figure 20(d)]. While the access time to any of the stored files is determined by the robotic picker and will be of the order of tens of seconds, the aggregate sustained data rate could be enormous. In this scenario, the relatively high component cost of a read/write holographic engine is amortized over a large volume of cheap media to obtain competitive cost per gigabyte.

Will one of these scenarios with data stored in holograms materialize and become reality in the foreseeable future? In collaboration and competition with a large number of scientists from around the globe, we continue to study the technical feasibility of holographic storage and memory devices with parameters that are relevant for real-world applications. Whether this research will one day lead to products depends on the insights that we gain into these technical issues and how well holography can compete with established techniques in the marketplace.

Acknowledgments

References

1. D. Psaltis and F. Mok, "Holographic Memories," *Sci. Amer.* **273,** No. 5, 70 (1995).
2. J. F. Heanue, M. C. Bashaw, and L. Hesselink, "Volume Holographic Storage and Retrieval of Digital Data," *Science* **265,** 749 (1994).
3. J. H. Hong, I. McMichael, T. Y. Chang, W. Christian, and E. G. Paek, "Volume Holographic Memory Systems: Techniques and Architectures," *Opt. Eng.* **34,** 2193–2203 (1995).
4. D. Psaltis and G. W. Burr, "Holographic Data Storage," *Computer* **31,** No. 2, 52–60 (1998).
5. J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill Book Co., Inc., New York, 1968.
6. M.-P. Bernal, H. Coufal, R. K. Grygier, J. A. Hoffnagle, C. M. Jefferson, R. M. Macfarlane, R. M. Shelby, G. T. Sincerbox, P. Wimmer, and G. Wittman, "A Precision Tester for Studies of Holographic Optical Storage Materials and Recording Physics," *Appl. Opt.* **35,** 2360–2374 (1996).
7. R. M. Shelby, J. A. Hoffnagle, G. W. Burr, C. M. Jefferson, M.-P. Bernal, H. Coufal, R. K. Grygier, H. Guenther, R. M. Macfarlane, and G. T. Sincerbox, "Pixel-Matched Holographic Data Storage with Megabit Pages," *Opt. Lett.* **22,** 1509 (1997).
8. G. W. Burr, J. Ashley, H. Coufal, R. Grygier, J. Hoffnagle, C. M. Jefferson, and B. Marcus, "Modulation Coding for Pixel-Matched Holographic Data Storage," *Opt. Lett.* **22,** 639–641 (1997).
9. J. L. Sanford, P. F. Greier, K. H. Yang, M. Lu, R. S. Olyha, Jr., C. Narayan, J. A. Hoffnagle, P. M. Alt, and R. L. Melcher, "A One-Megapixel Reflective Spatial Light Modulator System for Holographic Storage," *IBM J. Res. Develop.* **42,** No. 3/4, 411–426 (1998).
10. M.-P. Bernal, G. W. Burr, H. Coufal, R. K. Grygier, J. A. Hoffnagle, C. M. Jefferson, E. Oesterschulze, R. M. Shelby, G. T. Sincerbox, and M. Quintanilla, "Effects of Multilevel Phase Masks on Interpixel Crosstalk in Holographic Data Storage," *Appl. Opt.* **36,** No. 14, 3107–3115 (1997).
11. M.-P. Bernal, G. W. Burr, H. Coufal, J. A. Hoffnagle, C. M. Jefferson, R. M. Shelby, and M. Quintanilla, "Experimental Study of the Effects of a Six-Level Phase Mask on a Digital Holographic Storage System," *Appl. Opt.* **37,** No. 11, 2094–2101 (1998).
12. F. Ito, K.-I. Kitayama, and H. Oguri, "Compensation of Fiber Holographic Image Distortion Caused by Intrasignal Photorefractive Coupling by Using a Phase-Conjugate Mirror," *Opt. Lett.* **17,** 215 (1992).
13. M. C. Bashaw, A. Aharoni, and L. Hesselink, "Phase-Conjugate Replay for a-Axis Strontium Barium Niobate Single-Crystal Fibers," *Opt. Lett.* **18,** 2059 (1993).
14. F. Zhao and K. Sayano, "Compact Read-Only Memory with Lensless Phase-Conjugate Holograms," *Opt. Lett.* **21,** 1295 (1996).
15. J. J. P. Drolet, E. Chuang, G. Barbastathis, and D. Psaltis, "Compact, Integrated Dynamic Holographic Memory with Refreshed Holograms," *Opt. Lett.* **22,** 552 (1997).
16. G. W. Burr and R. M. Shelby, "Pixel-Matched Phase-Conjugate Holographic Data Storage," *SPIE Holography Newsletter*, p. 8 (January 1999).
17. J. Feinberg, "Self-Pumped, Continuous-Wave Phase Conjugator Using Internal Reflection," *Opt. Lett.* **7,** 486 (1982).
18. G. W. Burr, W.-C. Chou, M. Neifeld, H. Coufal, J. Hoffnagle, and C. M. Jefferson, "Experimental Evaluation of User Capacity in Holographic Data Storage System," *Appl. Opt.* **37,** 5431–5443 (1998).
19. V. Vadde and V. Kumar, "Channel Estimation and Intra-Page Equalization for Digital Volume Holographic Data Storage," *Proc. SPIE* **3109,** 250–255 (1997).

20. V. Vadde, V. Kumar, G. W. Burr, H. Coufal, J. Hoffnagle, and C. M. Jefferson, "Equalization for High Density Volume Holographic Storage," *Proc. SPIE* **3409,** 194–200 (1998).

21. M.-P. Bernal, G. W. Burr, H. Coufal, and M. Quintanilla, "Balancing Interpixel Crosstalk and Detector Noise to Optimize Areal Density in Holographic Storage Systems," *Appl. Opt.* **37,** 5377–5385 (1998).

22. J. Ashley and B. Marcus, "Constant-Weight/Lowpass Modulation Codes for Holographic Recording," *Research Report RJ-10089 (91905)*, IBM Research Laboratory, San Jose, CA, October 1997.

23. J. Ashley and B. Marcus, "Two-Dimensional Lowpass Filtering Codes for Holographic Storage," *IEEE Trans. Commun.* **46,** 724–727 (1998).

24. J. Ashley, M. Blaum, and B. Marcus, "Coding and Signal Processing for Holographic Data Recording" (Coding and Signal Processing Report for PRISM), *Research Report RJ-10013 (89104)*, IBM Research Laboratory, San Jose, CA, 1998.

25. I. S. Reed and G. Solomon, "Polynomial Codes Over Certain Finite Fields," *J. Soc. Ind. Appl. Math.* **8,** 300–304 (1960); R. J. McEliece, *The Theory of Information and Coding*, Addison-Wesley Publishing Co., Reading, MA, 1977.

26. M. Blaum, J. Bruck, and A. Vardy, "Interleaving Schemes for Multidimensional Cluster Errors," *IEEE Trans. Info. Theory* **44,** No. 2, 730–743 (1998).

27. W.-C. Chou and M. Neifeld, "Interleaving and Error Correction in Volume Holographic Memory Systems," *Appl. Opt.* **37,** No. 29, 6951–6968 (1999).

28. G. W. Burr, H. Coufal, R. K. Grygier, J. A. Hoffnagle, and C. M. Jefferson, "Noise Reduction of Page-Oriented Data Storage by Inverse Filtering During Recording," *Opt. Lett.* **23,** 289–291 (1998).

29. G. W. Burr, G. Barking, H. Coufal, J. A. Hoffnagle, C. M. Jefferson, and M. A. Neifeld, "Gray-Scale Data Pages for Digital Holographic Data Storage," *Opt. Lett.* **23,** 1218–1220 (1998).

30. M. A. Neifeld and J. D. Hayes, "Error-Correction Schemes for Volume Optical Memories," *App. Opt.* **34,** 8183–8191 (1995).

31. P. D. Henshaw and S. A. Lis, "Content Addressable Optical Data Storage System," U.S. Patent 5,319,629, 1994.

32. B. J. Goertzen and P. A. Mitkas, "Volume Holographic Storage for Large Relational Databases," *Opt. Eng.* **35,** 1847–1853 (1995).

33. G. W. Burr, S. Kobras, H. Hanssen, and H. Coufal, "Content-Addressable Data Storage Using Volume Holograms," *Appl. Opt.* **38,** 6779–6784 (1999).

34. L. A. Zadeh, "Fuzzy Sets," *Inf. Cont.* **8,** 338–353 (1965).

35. S. Kobras, "Associative Recall of Digital Data in Volume Holographic Storage Systems," Diplomarbeit, Technische Universitat Munchen, Germany, 1998.

36. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, B. Qian Huang Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System," *Computer* **28,** 23–32 (1995).

37. M.-P. Bernal, G. W. Burr, H. Coufal, R. K. Grygier, J. A. Hoffnagle, C. M. Jefferson, R. M. Macfarlane, R. M. Shelby, G. T. Sincerbox, and G. Wittmann, "Holographic Data Storage Materials," *Mater. Res. Bull.* **21,** 51–60 (1996).

38. J. F. Heanue, M. C. Bashaw, and L. Hesselink, "Channel Codes for Digital Holographic Data Storage," *J. Opt. Soc. Amer. A* **12,** 2432–2439 (1995).

39. F. H. Mok, G. W. Burr, and D. Psaltis, "A System Metric for Holographic Memory Systems," *Opt. Lett.* **21,** 896–899 (1996).

40. L. Hesselink and M. Bashaw, "Optical Memories Implemented with Photorefractive Media," *Opt. & Quant. Electr.* **25,** 611–651 (1993).

41. H. Günther, R. Macfarlane, Y. Furukawa, K. Kitamura, and R. Neurgaonkar, "Two-Color Holography in Reduced Near-Stoichiometric Lithium Niobate," *Appl. Opt.* **37,** 7611–7623 (1998).

42. Y. S. Bai, R. R. Neurgaonkar, and R. Kachru, "High-Efficiency Nonvolatile Holographic Storage with Two-Step Recording in Praseodymium-Doped Lithium Niobate by Use of Continuous-Wave Lasers," *Opt. Lett.* **22,** 334–336 (1997).

43. L. Hesselink, S. S. Orlov, A. Liu, A. Akella, D. Lande, and R. R. Neurgaonkar, "Photorefractive Materials for Nonvolatile Volume Holographic Data Storage," *Science* **282,** 1089–1094 (1998).

44. D. A. Waldman, H.-Y. S. Li, and E. A. Cetin, "Holographic Recording Properties in Thick Films of ULSH-500 Photopolymer," *Roc. SIP* **3291,** 89–103 (1998).

45. D. A. Waldman, H.-Y. S. Li, and M. G. Horner, "Volume Shrinkage in Slant Fringe Gratings of a Cationic Ring-Opening Holographic Recording Material," *J. Imag. Sci. Tech.* **41,** 497–514 (1997).

46. V. Colvin, R. Larson, A. L. Harris, and M. Schilling, "Quantitative Model of Volume Hologram Formation in Photopolymers," *J. Appl. Phys.* **81,** 5913–5923 (1997).

47. L. Dhar, A. Hale, H. E. Katz, M. L. Schilling, M. G. Schnoes, and F. C. Schilling, "Recording Media that Exhibit High Dynamic Range for Digital Holographic Data Storage," *Opt. Lett.* **24,** No. 7, 487–489 (1999).

48. G. J. Stechman, I. Solomantine, G. Zhou, and D. Psaltis, "Characterization of Phenanthrenequinone-Doped Poly(methylmethacrylate) for Holographic Memory," *Opt. Lett.* **23,** 1310–1312 (1998).

49. Johannes Eickmans and Thomas Bieringer, "Photoaddressable Polymers for Optical Data Storage," Paper T1.9, presented at the Materials Research Society Fall Meeting, November 30–December 4, 1998; J. Eickmans, T. Bieringer, S. Kostromine, H. Berneth, and R. Thoma, "Photoaddressable Polymers: A New Class of Materials for Optical Data Storage and Holographic Memories," *Jpn. J. Appl. Phys.* **38,** 1835–1836 (1999).

50. P. M. Lundquist, C. Poga, R. G. DeVoe, Y. Jia, W. E. Moerner, M.-P. Bernal, H. Coufal, R. K. Grygier, J. A. Hoffnagle, C. M. Jefferson, R. M. Macfarlane, and G. T. Sincerbox, "Holographic Digital Data Storage in a Photorefractive Polymer," *Opt. Lett.* **21,** 890–892 (1996).

51. M. D. Rahn, D. P. West, K. Khand, J. D. Shakos, and R. M. Shelby, "High Optical Quality and Fast Response Speed Holographic Data Storage in a Photorefractive Polymer," *J. Appl. Phys.*, submitted for publication.

52. R. Wortmann, P. M. Lundquist, R. Twieg, C. Geletneky, C. R. Moylan, Y. Jia, R. G. DeVoe, D. M. Burland, M.-P. Bernal, H. Coufal, R. K. Grygier, J. A. Hoffnagle, C. M. Jefferson, R. M. Macfarlane, R. M. Shelby, and G. T. Sincerbox, "A Novel Sensitized Photochromic Organic Glass for Holographic Optical Storage," *Appl. Phys. Lett.* **69,** 1657–1659 (1996).

53. K. Curtis, A. Pu, and D. Psaltis, "Method for Holographic Storage Using Peristrophic Multiplexing," *Opt. Lett.* **19,** 993–995 (1994).

54. J. J. Amodei and D. L. Staebler, "Holographic Pattern Fixing in Electro-Optic Crystals," *Appl. Phys. Lett.* **18,** 540–542 (1971).

55. A. Yariv, S. S. Orlov, and G. A. Rakuljic, "Holographic Storage Dynamics in Lithium Niobate: Theory and Experiment," *J. Opt. Soc. Amer. B* **13,** 2513–2523 (1996).

56. H. C. Kuelich, "A New Approach to Read Volume Holograms at Different Wavelengths," *Opt. Commun.* **64,** 407–411 (1987).

365

57. D. von der Linde, A. M. Glass, and K. F. Rodgers, "Multiphoton Photorefractive Processes for Optical Storage in LiNbO$_3$," *Appl. Phys. Lett.* **25,** 155–157 (1974).

58. H. Guenther, G. Wittmann, R. M. Macfarlane, and R. R. Neurgaonkar, "Intensity Dependence and White Light Gating of Two Color Photorefractive Gratings in LiNbO$_3$," *Opt. Lett.* **22,** 1305–1307 (1997).

59. F. Jermann, M. Simon, and E. Kratzig, "Photorefractive Properties of Congruent and Stoichiometric Lithium Niobate at High Light Intensities," *J. Opt. Soc. Amer. B* **12,** 2066–2070 (1995).

60. Y. S. Bai and R. Kachru, "Nonvolatile Holographic Storage with Two Step Recording in Lithium Niobate Using CW Lasers," *Phys. Rev. Lett.* **78,** 2944–2947 (1997).

61. D. Lande, S. S. Orlov, A. Akella, L. Hesselink, and R. R. Neurgaonkar, "Digital Holographic Storage System Incorporating Optical Fixing," *Opt. Lett.* **22,** 1722–1724 (1997).

62. Y. Furukawa, M. Sato, K. Kitamura, and F. Nitanda, "Growth and Characterization of Off-Congruent LiNbO$_3$ Single Crystals Grown by the Double Crucible Method," *J. Cryst. Growth* **128,** 909–914 (1993).

63. O. F. Schirmer, S. Juppe, and J. Koppitz, "ESR-, Optical and Photovoltaic Studies of Reduced Undoped LiNbO$_3$," *Cryst. Latt. Def. & Amorph. Mater.* **16,** 353–358 (1987).

64. J. L. Ketchum, K. L. Sweeney, L. E. Halliburton, and A. F. Armington, "Vacuum Annealing Effects in Lithium Niobate," *Phys. Lett. A* **94,** 450–453 (1983).

65. O. F. Schirmer, H.-J. Reyher, and M. Woehlecke, in *Insulating Materials for Optoelectronics: New Developments*, F. Agullo-Lopez, Ed., World Scientific Publishers, New York, 1995, Ch. 4 and references therein.

66. K. Buse, A. Adibi, and D. Psaltis, "Non-Volatile Holographic Storage in Doubly Doped Lithium Niobate Crystals," *Nature* **393,** 665–668 (1998).

67. L. Hesselink, S. S. Orlov, A. Liu, A. Akella, D. Lande, and R. R. Neurgaonkar, "Photorefractive Materials for Nonvolatile Volume Holographic Data Storage," *Science* **282,** 1089–1094 (1998).

68. F. Mok, G. Zhou, and A. Chugh, "Read-Only Data Storage System," Paper MAAA8, presented at the 1996 OSA Annual Meeting.

Jonathan Ashley *Infineon Technologies, 375 Encinal Street, Santa Cruz, California 95060 (Jonathan.Ashley@infineon.com).* Dr. Ashley received his Ph.D. degree in 1987 from the University of California at Santa Cruz. He was an IBM Thomas J. Watson Postdoctoral Fellow in the Mathematical Sciences and a National Science Foundation Postdoctoral Fellow until 1990, when he joined the IBM Almaden Research Center as a Research Staff Member. In 1998, he joined Infineon Technologies.

Maria-Pilar Bernal *Institut d'Optique Appliquée, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Ecublens, Lausanne, Switzerland (Maria-Pilar.Bernal@epfl.ch).* Dr. Bernal received her B.S. and Ph.D. degrees in physics from the University of Zaragoza, Spain, in 1993 and 1998, respectively. From 1994 to 1998, she participated in a joint study between the University of Zaragoza and the IBM Almaden Research Center, where she worked in the field of holographic data storage. She is currently a research assistant at the Swiss Federal Institute of Technology, Lausanne, where she works on scanning near-field optical microscopy.

Geoffrey W. Burr *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (burr@almaden.ibm.com).* Dr. Burr received his B.S. degree in electrical engineering from the State University of New York at Buffalo in 1991. That year Eta Kappa Nu selected him as the Alton B. Zerby Outstanding Electrical Engineering Senior in the United States. He received his M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology in 1993 and 1996, respectively. He has been at the IBM Almaden Research Center since 1996, first as a Visiting Scientist and, since July 1999, as a Research Staff Member. Dr. Burr's research interests include volume holography, coding and signal processing, and optical information processing and display.

Hans Coufal *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (coufal@almaden.ibm.com).* Dr. Coufal received his Ph.D. degree in physics from the Technical University of Munich. After several years on the faculty at this university and at the Free University in Berlin, he joined the IBM San Jose Research Laboratory as a member of the research staff in 1981. He is a Fellow of the Optical Society of America and of the International Union of Pure and Applied Chemistry. Dr. Coufal is currently manager of the New Directions in Science and Technology Department at the IBM Almaden Research Center; he has managed IBM's effort in holographic data storage since 1993. In 1996, Dr. Coufal became one of the two principal investigators of the DARPA-sponsored HDSS and PRISM consortia.

Harald Guenther *Gemfire, 2471 East Bayshore Road No. 600, Palo Alto, California 94303 (h.guenther@gemfirecorp.com).* Dr. Guenther received his M.S. and Ph.D. degrees in physics from the University of Heidelberg, Germany, in 1992 and 1995, respectively. For the following year he continued his work at the University of Heidelberg. From 1996 to 1998, he joined the IBM Almaden Research Center in San Jose, California, as a Visiting Scientist, investigating rare-earth-doped lithium niobate for two-photon holographic data storage. In 1998, Dr. Guenther joined Gemfire Corporation.

John A. Hoffnagle *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (hoffnagl@almaden.ibm.com).* Dr. Hoffnagle received his Ph.D. degree in 1982 from the Swiss Federal Institute of Technology, Zurich. He joined the IBM Almaden Research Center in 1985, working on spectroscopy and nonlinear dynamics of trapped ions, and is currently a Research Staff Member. His current interests are applications of coherent optical technology for holographic data storage and interferometric lithography. Dr. Hoffnagle is a member of the American Physical Society and the Optical Society of America.

C. Michael Jefferson *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (michael@almaden.ibm.com).* Dr. Jefferson received his M.S. degree in scientific instrumentation from the University of California at Santa Barbara in 1975. After spending a year as Station Science Leader at the Amundsen–Scott South Pole Station, where he conducted geophysical research, he entered the physics program at UCSB. He graduated with a Ph.D. degree in physics in 1984, joining the IBM Almaden Research Center as a Research Staff Member the following year. His work at IBM has included precision instrumentation and mechanics for magnetic storage, applications of persistent spectral hole-burning for digital data storage, and the development of precision test stands for high-density holographic digital data storage.

Brian Marcus *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (marcus@almaden.ibm.com).* Dr. Marcus received his B.A. degree from Pomona College in 1971 and his Ph.D. degree in mathematics from the University of California at Berkeley in 1975. He held the IBM Thomas J. Watson Postdoctoral Fellowship in mathematical sciences from 1976 to 1977. From 1975 to 1985 he was Assistant Professor and then Associate Professor of Mathematics (with tenure) at the University of North Carolina at Chapel Hill. Since 1984 he has been a Research Staff Member at the IBM Almaden Research Center. Dr. Marcus was a co-recipient of the Leonard G. Abraham Prize Paper Award of the IEEE Communications Society in 1993. His current research interests include symbolic dynamics and coding for storage devices.

Roger M. Macfarlane *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (macfarla@almaden.ibm.com).* Dr. Macfarlane received his Ph.D. degree in physics from the University of Canterbury, New Zealand. After a postdoctoral position in physics at Stanford University, he joined the IBM Research Division as a Research Staff Member. He has also held temporary visiting positions at Oxford University and the University of Grenoble. Dr. Macfarlane is a Fellow of the American Physical Society and of the Optical Society of America; he has published extensively on the optical properties of materials, spectroscopy, and optical-coherence phenomena.

Robert M. Shelby *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (shelby@almaden.ibm.com).* Dr. Shelby received his B.S. degree from the California Institute of Technology and his Ph.D. degree from the University of California at Berkeley,

367

both in chemistry. In 1978 he joined the IBM Research Division, where he is a Research Staff Member, working on quantum optics, fundamental noise processes in optical fibers, coherent transient phenomena, and high-resolution laser spectroscopy. His current interests are recording physics for holographic data storage and spectroscopic applications of near-field optical microscopy. Dr. Shelby is a Fellow of the Optical Society of America.

Glenn T. Sincerbox *Optical Sciences Center, Meinel Building, University of Arizona, Tucson, Arizona 85721 (sinbox@u.arizona.edu).* Mr. Sincerbox is a Professor of Optical Sciences and Director of the Optical Data Storage Center in the Optical Sciences Center at the University of Arizona in Tucson, Arizona. Previous to this, he was with the IBM Almaden Research Center in San Jose, California, where he held several management positions. He was one of the principal investigators of the DARPA-sponsored PRISM and HDSS consortia until 1996, when he retired from IBM after 34 years of service to assume his current position in Arizona.