

Holographic data storage promises increased storage densities and improved access times that could lead to a new range of novel memory devices

Optical data storage enters a new dimension

IBM Holographic Storage Team

IN THE era of the Internet, massive amounts of information and multimedia have become easily accessible in every corner of the world. The decreasing cost of storing data, and the increasing storage capacities of ever smaller devices, have been key enablers of this revolution. Current storage needs are being met because improvements in conventional technologies – such as magnetic hard-disk drives, optical disks and semiconductor memories – have been able to keep pace with the demand for greater and faster storage.

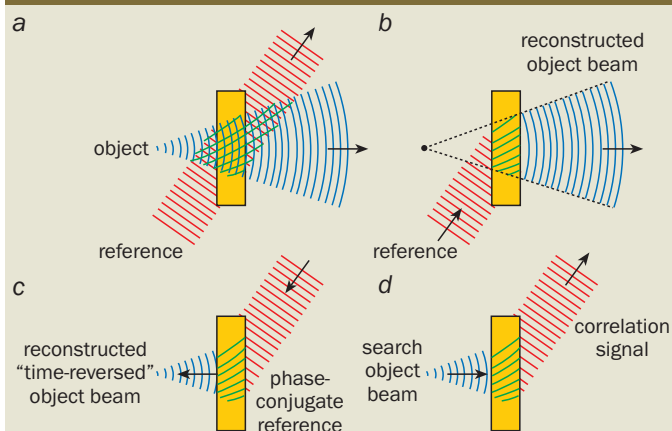
However, there is strong evidence that these surface-storage technologies are approaching fundamental limits that may be difficult to overcome as the ever-smaller regions that store bits of information become less thermally stable and harder to access. Exactly when this limit will be reached remains an open question: some experts predict these barriers will be encountered in 2–3 years, while others believe they will not be reached for at least another five years. In either case, one or more successors to current data-storage technologies will be needed in the near future.

An intriguing approach for the next generation of data-storage systems uses optical holography to store information throughout the three-dimensional volume of a material. And by superimposing many holograms within the same volume of the recording medium, it should be possible to achieve far greater storage densities than current technologies can offer.

Although holography was conceived in the late 1940s, it was not considered a potential storage technology until the development of the laser in the 1960s. The rapid development of holography for displaying 3-D images led to the realization that holograms could potentially store data at a volumetric density of one bit per cubic wavelength. Given a typical laser wavelength of around 500 nm, this density corresponds to 10^{12} bits (1 terabit) per cubic centimetre or more.

In holographic storage, data are transferred to and from the storage material as 2-D images composed of thousands of pixels, each of which represents a single bit of information. Since an entire “page” of data can be retrieved by a photodetector at the same time, rather than bit-by-bit, the holographic scheme promises fast read-out rates as well as high storage densities. If a thousand holograms, each containing a million pixels, could be retrieved every second, for example, then the output data rate would reach 1 gigabit per second. (In comparison, a DVD optical-disk player reads data 100

1 How to record and read data using holograms



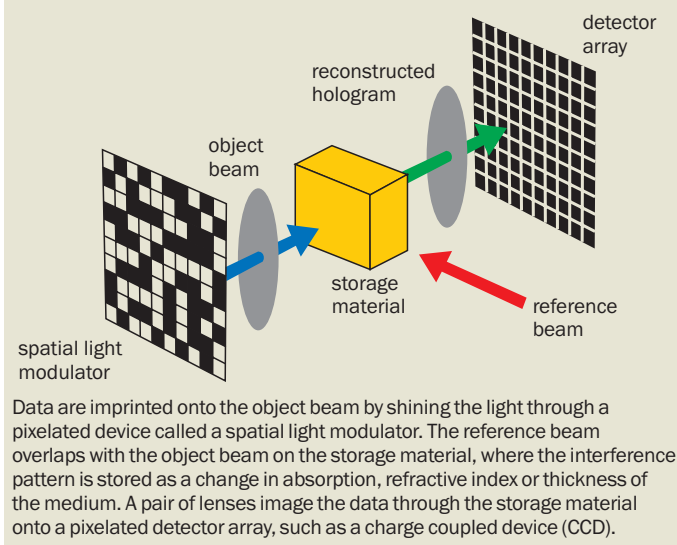
(a) Holographic storage of a single data bit. The spherical wave from a single pixel (the object shown in blue) interferes with a coherent plane wave (red) in the reference beam. The resulting interference pattern (green) changes the refractive properties of the photosensitive medium (yellow). (b) The hologram is read out using the original reference beam, which is diffracted by the interference pattern stored in the medium to reconstruct the original spherical wavefront. An image of this beam can be formed on a single detector pixel, resulting in the retrieval of a single bit. (c) The hologram can also be read out by illuminating it with a counter-propagating (or “phase-conjugate”) reference beam. The phase-conjugate beam returns to its point of origin, where the bit value can be read without a high-quality imaging system. (d) A third way to retrieve data involves illumination with a diverging beam called an object beam, which reconstructs the original plane-wave reference beam. This beam can be focused onto a detector and provides an optical measurement of the correlation between the stored data and the object beam. This technique allows us to search the stored data according to its content, rather than its address.

times slower.) Despite this attractive potential, however, research into holographic data storage all but died out in the mid-1970s due to the lack of suitable devices that could transfer 2-D pixelated images.

Interest in volume-holographic data storage was rekindled in the early 1990s by the availability of charge coupled devices (CCD), semiconductor detectors, small liquid-crystal panels and other devices that can display and detect 2-D pages of data. The wide availability of these devices was made possible by the commercial success of hand-held camcorders, digital cameras and video projectors.

With these components at their disposal, researchers have

2 Basics of a holographic data-storage system



begun to demonstrate the potential of holographic storage and have shown that data can be stored at densities equivalent to 390 bits per square micron. This density exceeds the storage capabilities of DVD disks by a factor of almost 20, and of magnetic disks by a factor of five.

The potential of holographic storage has generated numerous research efforts at large multinational companies, including Lucent, Imation, Panasonic, Sony and NEC, as well as independent start-up companies such as Holoplex, Tamarak and Accuwave. Several consortia of universities and industries are funded by the Defense Advanced Research Project Agency (DARPA) in the US with industrial partners including IBM, Siros, Rockwell, Kodak, Bayer and Aprilis.

Holography basics

A hologram is a recording of the optical interference pattern that forms at the intersection of two coherent optical beams. Typically, light from a laser is split into two paths called the object and reference paths (figure 1a). The beam that propagates along the object path carries the information while the reference beam is used to record and read out the hologram. A plane wave is commonly used as the reference beam because it is simple to reproduce at a later stage.

To make the hologram, the reference and object beams are made to overlap on a photosensitive medium, such as a photopolymer or an inorganic crystal, where the resulting optical interference pattern creates chemical and/or physical changes. As a result, a replica of the interference pattern is stored as a change in the absorption, refractive index or thickness of the media.

The pattern contains information about both the amplitude and the phase of the two light beams. This means that when the recording is illuminated by the read-out beam, some of the light is diffracted to “reconstruct” a weak copy of the object beam (figure 1b). If the object beam originally came from a 3-D object, then the reconstructed hologram makes the 3-D object reappear.

If the hologram material is thin – as it is on many credit cards – the read-out beam can differ in angle or wavelength from the reference beam that was used to record the image and the scene will still appear. However, if the hologram is

Associative retrieval

With a conventional memory or data-storage device, a user must supply an address at which the desired data is located. In volume-holographic data storage, this implies that the data – which were once imprinted on an “object” beam and stored within the volume – can be read out later by illuminating the volume with the correct “addressing” reference beam (figures 1a and b).

However, this hologram can also be illuminated by the *object* beam to reconstruct all of the angle-multiplexed reference beams that were used to record data pages into the volume (figure 1d).

The amount of power diffracted into each “output” beam is proportional to the similarity between the input data page that is displayed on the spatial light modulator and the stored data page. Each set of output beams can be focused onto a detector array, so that each beam forms its own correlation “peak”. The stored pages that match the input page can be identified by setting a threshold on the detected optical signal.

If the patterns that make up these pages correspond to the various data fields of a database, and if each stored page represents a data record, then this optical-correlation process can be used to simultaneously compare the entire database against the search argument. This parallelism gives content-addressable holographic data storage an inherent speed advantage over a conventional serial search through large databases, in particular. For example, it would take a conventional software-based search 40 seconds to go through one million records each containing 1 kilobyte of data. In comparison, an appropriately designed holographic system built using off-the-shelf components could search the same records in about 30 ms – over 1200 times faster. Custom-built hardware could reduce this search time to 1 ms or less.

In addition to searching for exact matches to a query, we have introduced a novel encoding scheme that allows us to perform a “fuzzy search” for similar matches. To demonstrate parallel searching of a holographic content-addressable memory, we stored and searched a small multimedia database of 100 images in our modified DEMON I system. We achieved high fidelity with searches for similarly coloured images, searches for specific colour percentages, and searches on keyword. Each search could be performed with a single optical exposure.

With the ability to perform fuzzy searches and to search many thousands of holograms in parallel, volume-holographic content-addressable data storage is an attractive method for rapidly searching vast databases with complex queries.

recorded in a thick material, the reconstructed object beam will only appear when the read-out beam is almost identical to the original reference beam.

Since the diffracted wavefront accumulates energy from throughout the thickness of the storage material, a small change in either the wavelength or angle of the read-out beam generates enough destructive interference to make the hologram effectively disappear.

The sensitivity of the holographic reconstruction to changes in wavelength and angle increases with the material thickness, which means that the laser and read-out optics need to be stable and give repeatable results. However, destructive interference also opens up a tremendous opportunity: a small storage volume can now store multiple superimposed holograms, each one distributed throughout the entire volume. The destructive interference allows each of these stored holo-

grams to be independently accessed with its original reference beam.

Several different techniques have been developed to define a set of suitable reference beams by, for example, slightly changing the angle, wavelength or phase of the original light beam. Using so-called angle multiplexing, as many as 10 000 holograms have been stored in a 1 cm³ volume.

Storing and retrieving digital data

To use volume holography as a storage technology, digital data must be imprinted onto the object beam for recording and then retrieved from the reconstructed object beam during read out (figure 2).

The device for putting data into the system is called a spatial light modulator (SLM) – a planar array consisting of thousands of pixels. Each pixel is an independent microscopic shutter that can either block or pass light using liquid-crystal or micro-mirror technology. Liquid-crystal panels with 1000 × 1000 pixels and micro-mirror arrays with 1000 × 800 elements are commercially available due to the success of computer-driven projection displays. The pixels in both types of device can be refreshed over 1000 times per second, allowing holographic data-storage systems to reach input data rates of 1 gigabit per second – assuming that the laser power and material sensitivities permit.

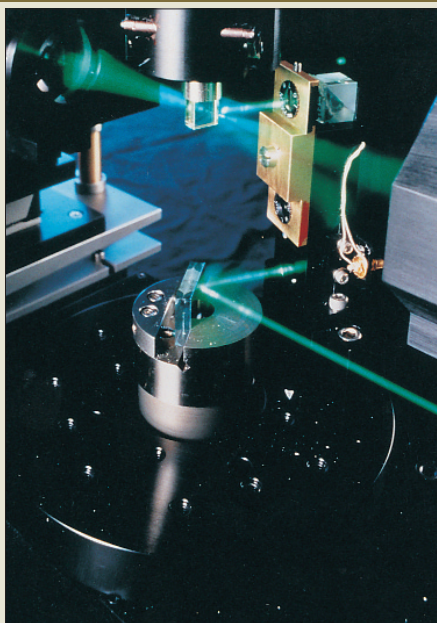
The data are read using an array of detector pixels, such as a CCD camera or a semiconductor sensor. The object beam often passes through a set of lenses that image the SLM pixel pattern onto the output pixel array (figure 2). To maximize the storage density, the hologram is usually recorded where the object beam is tightly focused.

When the hologram is reconstructed by the reference beam, a weak copy of the original object beam continues along the imaging path to the camera, where the optical output can be detected and converted to digital data.

The speed of a storage device is described by the read-out rate (in bits per second) and the latency, or time delay, between asking for and receiving a particular bit of data. To access holographically stored data, the correct reference beam must be directed to the appropriate spot within the storage media. The hologram is then reconstructed and the optical signals processed and decoded to extract the desired digital data. The latency tends to be dominated by mechanical movement, especially if the storage media has to be moved.

The read-out rate is often dictated by the camera integration time: the reference beam reconstructs a hologram until a sufficient number of photons accumulate to differentiate bright and dark pixels. A frequently mentioned goal is an integration time of about 1 millisecond, which implies that 1000 pages of data can be retrieved per second. If there are 1 million pixels per data page and each pixel stores one bit, then the read-out rate is 1 gigabit per second. This goal requires high laser power (at least 1 W), a high-quality storage

3 Testing times for materials



A close-up of the PRISM apparatus used to test holographic materials. The spatial light modulator is a chrome-on-glass mask (not shown), and the detector array is a portion of a large CCD camera comprising 1024 × 1024 pixels (not shown). A pair of precision rotation stages (centre) allow the reference beam to enter the storage material (upper centre) at any horizontal incidence angle.

material and a detector that has a million pixels and can be read out at high “frame rates”.

Frame rates of 1 kHz have been demonstrated in such “mega-pixel” CCDs, but these are not yet commercially available. However, low-noise mega-pixel semiconductor-detector arrays that can support 500 frames per second are expected to become available soon. Even with these requirements, faster read-out rates and lower latency could be reached by steering the reference-beam angle non-mechanically, by using a pulsed laser and by reading only the desired portion of the detector array.

Both the capacity and the read-out rate are maximized when each detector pixel is matched to a single pixel on the SLM, but for large pixel arrays this requires careful optical design and alignment. At IBM’s Almaden Research Center in California, we have built several test platforms on which this pixel-to-pixel matching has been achieved. The first such platform was designed to test holographic recording materials for the DARPA-sponsored Photorefractive Information Storage

Materials (PRISM) program.

We have designed the PRISM tester to have customized imaging optics and a high degree of alignment (figure 3). This means that any errors that occur while the data pages are stored and retrieved can be unambiguously assigned to the storage material under study.

These high-performance features have allowed our partners in materials development – including Aprilis and Bayer – to evaluate and markedly improve their materials and fabrication processes. Since the reference beam can be delivered to the front, back or sides of samples, the PRISM system has been able to test and study the recording physics of a wide variety of storage materials, including both write-once and erasable materials.

Recording materials

Materials for writing permanent volume holograms generally involve irreversible photochemical reactions that are triggered by the bright regions of the optical interference pattern. A photopolymer material, for example, polymerizes in response to optical illumination: material diffuses from darker to brighter regions so that short monomer chains can bind together to form long molecular chains. And in a so-called direct-write or photochromic material, the illuminated molecules undergo a local change in their absorption or index of refraction, which is driven by photochemistry or photo-induced molecular reconfiguration.

Both types of materials are inexpensive to make in bulk, but both can have problems reproducing the object beam faithfully. With a photopolymer, problems arise because the material shrinks during recording, distorting the reconstructed pixelated image. The direct-write material responds both to

the rapid variations of the interference pattern encoded with data and to long-range brightness variations across the illuminated spot. Such effects distort the reconstructed data pages. Fortunately, all these problems can be minimized by careful system design. For example, we have developed a novel optical-illumination system that delivers beams with extremely uniform brightness.

One advantage of a photopolymer is that leftover monomers can be polymerized after recording without affecting the holograms. In contrast, a direct-write material requires a separate chemical or optical step after the hologram-recording process to deplete the remaining absorbers. This is necessary otherwise the light from the read-out beam would induce further photochemical reactions, which would reduce the contrast between the bright and dark fringes in the recorded interference pattern and thereby erase the stored holograms. As with photographic film, both these types of write-once media must be protected from ambient light before use, and tend to lose their effectiveness as they age.

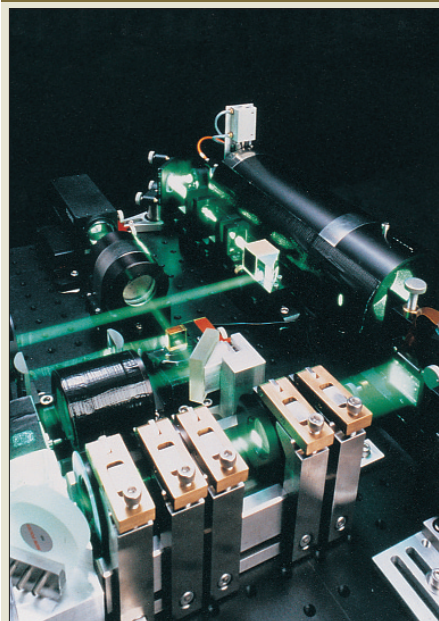
Although problems with shrinkage, scattering and dynamic range remain, recent developments in these write-once materials have overcome previous difficulties with poor optical quality and excessive absorption. And fairly thick samples (0.5–1 mm) have been demonstrated. Together with recently developed multiplexing techniques that use spherical or speckled reference beams to increase the number of holograms that can be superimposed in thin media, these developments bring “write-once/read-many” holographic storage systems close to the prototype stage.

Most erasable holographic materials are inorganic photorefractive crystals doped with transition metals or rare-earth ions. These crystals are often available in centimetre-thick samples and include lithium niobate, strontium barium niobate and barium titanate doped with iron, cerium, praseodymium or manganese.

These materials react to the light and dark regions of an interference pattern by transporting and trapping electrons, which subsequently leads to a local change in the index of refraction. The trapped charge can be rearranged by later illumination, so it is possible to erase recorded holograms and replace them with new ones.

This would seem to enable a read–write storage device in which small blocks of data can be written, read and erased equally well. However, the recording rate of photorefractive materials is typically 5–50 times slower than the achievable read-out rate. In addition, it is quite difficult to erase individual holograms from a small storage volume without affecting the other superimposed holograms. As a result, a holographic storage system built from photorefractive crystals is not a conventional read–write system. Instead such a storage device would record data slowly and in large 100–1000 megabyte

4 Holographic storage in action



The DEMON I holographic data-storage system. The information is encoded in light that passes through a spatial light modulator (right foreground), which is then focused onto the photosensitive inorganic crystal (centre) and finally imaged onto a pixelated detector (centre background). The coherent reference beam is directed by a mirror (left edge) onto the crystal.

blocks but could then provide rapid access to any small chunk of data. The large data blocks could be erased and replaced as desired.

The process of charge re-excitation means that the stored holograms can be erased during normal read out – a feature that is clearly undesirable. Moreover, the holograms can gradually fade in the dark via thermal excitation.

Recorded holograms can be “fixed” – that is made semi-permanent and resistant to fading during read out – by using separate thermal or electronic processes. Since this fixing process affects all the stored holograms within a volume simultaneously, it tends to be slow and cumbersome.

An alternative method for achieving non-volatile storage in photorefractive materials is to record at a wavelength that is only absorbed by the crystal in the presence of a third “gating” beam of a different wavelength. This third beam is present only during recording and is switched off while the information is read out, allowing the data to be retrieved without being erased.

Conventional photorefractive materials can be optimized for this gated, two-colour recording process by changing the way in which they are fabricated or by adding multiple dopants. For example, one of us (RMM) and Harald Guenther, now at the Gemfire Corporation in California, have enhanced the two-colour response of lithium niobate by using material in which the ratio of lithium to niobium was increased. Meanwhile, Demetri Psaltis and co-workers at the California Institute of Technology have improved the response of the crystals by doping them with manganese and iron atoms.

In the last two years, research into gated, two-colour photorefractive materials has led to improvements in the sensitivity and dynamic range of the materials. This, in turn, has increased both the speed with which information can be written and the data-storage capacity. However, further improvements are still needed before prototypes can be built.

Coding and signal processing

In most data-storage systems, designers maximize important figures of merit – such as the storage density and data rate – by pushing the physical components of the system well beyond the point where the system is error-free. Coding and signal-processing algorithms are then introduced to reduce the proportion of erroneous bits to acceptable levels.

Whereas the PRISM tester was built to test the properties of holographic storage materials, we built a second platform called DEMON I to evaluate the coding and signal-processing techniques, including the algorithms used to combat noise (figure 4).

Noise can arise in a holographic storage system from a number of sources, including poor imaging of the data pattern onto the detector, optical scatter, cross-talk between multiplexed holograms, spatial brightness variations and

electronic detector noise. The fundamental trade-off between the levels of signal and noise is a consequence of the finite dynamic range of the storage material. As the number of holograms or the read-out rate increases, the amount of power diffracted towards the detector array decreases, reducing the signal-to-noise ratio and increasing the number of incorrect bits.

Using the DEMON I platform, we have developed several powerful data-modulation codes and novel signal-processing techniques for the removal of deterministic variations from holographic data pages. We have also demonstrated ways to record more than one bit per pixel by controlling the amount of light that each pixel receives during recording.

In addition, we have developed a strong understanding of the trade-offs involved in recording holograms at high areal density, where focusing a large data page through a small hole introduces cross-talk between pixels.

This has allowed us to build a third platform, called DEMON II, specifically to demonstrate high areal density. (High volumetric density then follows from a combination of high areal density and a thin material.)

DEMON II combines large, mega-pixel data pages together with short-focal-length optics that have a high resolving power, and novel post-processing routines to correct for optical aberrations. Recently, we used DEMON II to store and retrieve volume-holographic data at an areal density of 390 bits per square micron. (In comparison, DVD disks have an areal density of 20 bits μm^{-2} and magnetic disks that can store up to 80 bits μm^{-2} have been produced in the lab.)

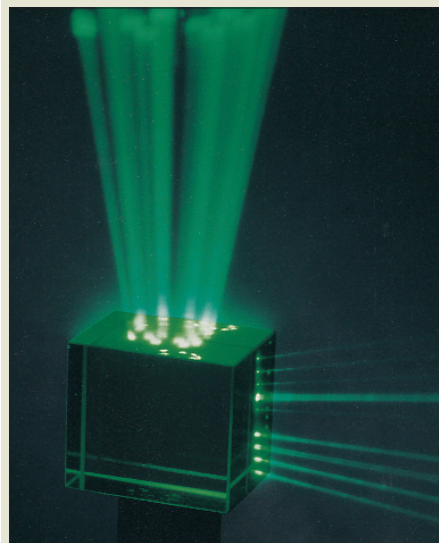
Phase-conjugate read out

Experiments show that an expensive short-focal-length lens system and a storage material with high optical quality are needed to combine high density storage with excellent imaging capabilities. Moreover, the optical system must be corrected for all aberrations, particularly distortion, over a large field of view.

In the past, several researchers – including Lambertus Hesselink and co-workers at Stanford University, Feng Zhao at the Accuwave Corporation and Psaltis at Caltech – have proposed bypassing these requirements by reading the volume holograms using a “phase-conjugate” approach (figure 1c).

Once the light from the spatial light modulator has been recorded with a reference beam, the resulting hologram can be reconstructed with a phase-conjugate or “time-reversed” copy of the original reference beam. The wavefront diffracted by the phase-conjugate read-out beam then retraces the path of the incoming object beam in reverse, cancelling out any accumulated phase errors from lens aberrations or material imperfections. This allows data pages to be retrieved with high fidelity using an inexpensive lens, or even without imaging lenses for an extremely compact system.

5 Searching data in an instant



Holographic data storage has the unique ability to locate similar features stored within a crystal instantly. A data pattern projected into a crystal from the top searches thousands of stored holograms in parallel. The holograms diffract the incoming light out of the side of the crystal, with the brightest outgoing beams identifying the address of the data that most closely resemble the input pattern. This parallel search capability is an inherent property of holographic data storage and allows a database to be searched by content.

However, two uncertainties prevented earlier work from proceeding. First, researchers were worried that imperfections in producing the phase-conjugated reference beam would introduce errors in each retrieved data page. In the simplest case this phase-conjugate beam is a separate beam that is carefully aligned to propagate in exactly the opposite direction to the original reference beam. However, even minor differences between the two beams will distort the reconstructed data pages.

Alternatively an extremely accurate phase-conjugate beam can be produced by a so-called phase-conjugate mirror. In this device, an incoming beam enters a barium-titanate crystal and interferes with its own backscattered reflection. The resulting hologram eventually diffracts the phase-conjugate of the input beam. Thus there are two holograms involved: one to record the data-bearing object beam, and one to produce the phase-conjugated reference beam.

The second concern was that many pairs of phase-conjugate reference beams would be needed to read the many different holograms recorded within the same volume, and that it

would be impossible from a practical point of view to maintain these beams over long periods of time. Furthermore, researchers were prevented from using phase-conjugate mirrors because barium-titanate crystals take some time to respond to changes in the input beam.

To solve this problem we proposed – and are currently testing – a novel architecture that allows phase-conjugation and multiplexed holographic storage to co-exist. The technique involves separating the phase-conjugation and hologram-storage processes into two successive steps using a “buffer” hologram. This approach allows many holograms to be stored in separate storage locations using a single spatial light modulator and just one detector. Since such a system contains a single pair of phase-conjugate beams, it never needs to wait for the barium-titanate crystal to respond.

We anticipate that the successful use of phase-conjugation in holographic storage will enable compact and affordable high-capacity systems, with only a moderate increase in the overall system complexity.

Applications of holography

Holographic storage is an emerging candidate for the next generation of data-storage systems. It could potentially lead to new devices, including fast, erasable holographic memory, write-once 3-D disks, pre-recorded 3-D disks and holographic “content-addressable” database machines (figure 5).

A fast, erasable holographic memory could provide sub-millisecond access to information and support terabytes of data with read-out rates in excess of 1 gigabit per second. Suitable applications might include “video on demand” and large Web servers.

Write-once 3-D disks could support more than 100 giga-

bytes per disk, each 120 mm in diameter. Blocks of data 100 megabytes in size could be accessed in 10–100 milliseconds and could perhaps be read out at rates of 500 megabits per second. Suitable applications might include the archiving of data that require permanent storage yet rapid access, such as medical information, and high-resolution maps and satellite images. There are initial indications that some write-once materials could possibly be made reversible, leading to erasable 3-D disks.

On the other hand, holographic data storage might be the successor to the DVD standard, offering pre-recorded 3-D disks for distributing computer programmes, movies and multimedia. One advantage of holographic data storage is that the 3-D disks cannot be replicated via physical stamping. This would make it difficult to pirate disks on a large scale.

Finally, holography could provide a unique hardware device for searching databases rapidly (see box on page 38). This device might offer 1000 searches through more than 1 million database records per second, supporting applications such as data-mining, genome databases, fingerprint or iris-print databases and bio-informatics.

Bright future?

The research efforts of the last few years have demonstrated that holographic storage systems with desirable properties can be engineered and built in the laboratory. However, existing and competing storage technologies are also continuing to evolve and improve at a tremendous pace, making the next few years crucial for holographic storage.

We are collaborating with (and competing against) a large number of scientists around the world to study the technical feasibility of holographic storage and memory devices that are relevant for real-world applications. The next steps are to optimize the storage media, to demonstrate these systems outside the laboratory environment, and to design and build systems at a cost that is competitive with existing technologies.

If suitable recording materials become available from the research efforts currently under way, holographic memory, 3-D disks, and database-searching machines may very well become a reality.

Further reading

- J Ashley *et al.* 2000 Holographic data storage *IBM Journal of Research & Development* **44** at press
- J F Heanue, M C Bashaw and L Hesselink 1994 Volume holographic storage and retrieval of digital data *Science* **265** 749
- J H Hong *et al.* 1995 Volume holographic memory systems: techniques and architectures *Optical Engineering* **34** 2193
- D Psaltis and G W Burr 1998 Holographic data storage *Computer* **31** 52–60
- D Psaltis and F Mok 1995 Holographic memories *Scientific American* **273** 70
- G T Sincerbox (ed) 1994 *Selected papers on Holographic Storage SPIE Milestone Series MS 95* (SPIE, Bellingham)

The members of the IBM Holographic Storage Team are **Geoffrey W Burr, Hans Coufal, John A Hoffnagle, C Michael Jefferson, Mark Jurich, Brian Marcus, Roger M Macfarlane and Robert M Shelby**. The team is at the IBM Almaden Research Center, D2-K18, 650 Harry Road, San Jose, CA 95120, USA, e-mail burr@almaden.ibm.com
