OPTICAL CACHE MEMORY

Physical Optics Corporation

Dr. Chai-Pei Kuo

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

91-01815

Rome Laboratory
Air Force Systems Command
Griffiss Air Force Base, NY 13441-5700
This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-91-40 has been reviewed and is approved for publication.

APPROVED: Fred N. Haritatos
FRED N. HARITATOS
Project Engineer

APPROVED: WALTER J. SENUS
WALTER J. SENUS
Technical Director
Directorate of Intelligence & Reconnaissance

FOR THE COMMANDER:
IGOR G. PLONISCH
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL( IRAP) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.
The goal of this effort is to build an Exploratory Development Model (EDM), of an optical cache memory utilizing existing off-the-shelf components to demonstrate system feasibility, key system concepts, and validation of the overall system approach. An optical parallel addressing architecture was identified which utilizes a vector-matrix inner product as the basic framework for an optical cache memory system. Furthermore, in order to mimic the dynamic memory requisition and updating sequence that is encountered in an electronic system, a polymeric memory material was studied and its performance was tailored to provide long memory persistence, fast writing, fast erasing, and high efficiency.
# TABLE OF CONTENTS

1.0 INTRODUCTION .................................................................................................................. 1
2.0 OPTICAL CACHE SYSTEM ARCHITECTURE DEFINITION .............................................. 2
  2.1 Principle of Cache Memory ................................................................................................. 2
  2.2 Problem Areas .................................................................................................................. 4
  2.3 System Architecture Optical Implementation ................................................................. 7
    2.3.1 Highly Parallel Optical Cache Memory Management ............................................... 7
    2.3.2 The Vector-Matrix Inner Product .............................................................................. 8
3.0 RECONFIGURABLE STORAGE MATERIALS ................................................................. 21
  3.1 Basic Properties of Storage Material Using Two Photon Absorption ............................... 23
  3.2 Basic Properties of Materials Using Spatial (Spectral) Hole Burning Effect ................... 26
  3.3 ORGANIC STORAGE MATERIAL BASED ON PHOTOINDUCED ANISOTROPY ............ 29
    3.3.1 Recording Principles Using Photoinduced Anisotropy in Organic Polymer ............... 30
    3.3.1.1 Holographic Mode Recording ............................................................................ 31
    3.3.1.2 The Birefringent Mode (Operating Principle and Application) .......................... 33
    3.3.2 Material Samples .................................................................................................... 34
    3.3.3 Investigations into Photoinduced Anisotropy ........................................................... 34
    3.3.4 Physical Properties of Guest-Host Dye Polymers and Their Modifications
       for Memory Storage ......................................................................................................... 38
4.0 SUMMARY .......................................................................................................................... 45
  4.1 Construction of an Exploratory Development Model ....................................................... 45
  4.2 Conclusions and Recommendations for Future Research .............................................. 56
5.0 REFERENCES .................................................................................................................... 60
1.0 INTRODUCTION

The speed of modern supercomputing devices is primarily the result of the use of parallel multiprocessors and very fast memory. The fastest memory, which is usually located close to the processor, consists of a cache. The cache's size is limited to a fraction, about $10^{-2}$, of the main memory size. The main memory may include hundreds of megabytes in a large machine. The computer memory is organized in a hierarchy, from the fastest to the slowest memory. The basic role of the caching system is to continuously keep the needed instructions in the fastest memory in order to minimize central processing unit (CPU) access-time. A large part of the speed of supercomputing derives from this caching system which may reduce access time to the subnanosecond level. This memory, however, is orders of magnitude more expensive than the standard random-access memory whose access time is measured in the tens or hundreds of nanoseconds. The efficient combination of memory devices of various hierarchy levels will result in a super computer that operates at the speed of the fastest element, provides the bulk of its capacity at the cost of its cheapest element, and that has short access time.

Because cache memory is used to speed up the flow of instructions and data into the processor from the main memory, an improved cache memory will provide two benefits. First, the average access time of processor memory requests will be reduced; this increases processor throughput. Second, the available memory bandwidth will be used less by the processor, allowing other devices on the system bus to use the main memory without interfacing with the processor. As will be described in this report, an optical cache memory promises to increase access time, throughput rate, and storage capability with the possibility of decreasing size, power, and system cost. More importantly, the implementation of parallel memory access in an optical cache system eliminates the need to synchronize the multiprocessor and the sequentially accessed electronic memory device.

The mission in this research and development program is to demonstrate the operation of a basic optical cache system. In order to achieve this objective, two major topics have been investigated and the results are presented in this report. One investigation has been a study of architecture definition for an optical system that has nondestructive parallel memory access and can allow memory data manipulation. The second investigation is for the purpose of optical read/write/erase memory material development. In our system architecture design, the memory reading approach is based on the implementation of optical vector-matrix inner product. Because of the intrinsic chromatic characteristics of the memory storage medium that was used, the non-mechanical parallel memory accessing mechanism is achieved through a novel manipulation of light (i.e., a green light of $\lambda = 514.5$ nm and a red light of $\lambda = 632$ nm) in the system to facilitate the write/erase process.
and the read process respectively. For storage medium development, studies were made first on currently existing materials. Difficulties for their application in a practical cache memory system were identified (section 3.0). A novel material based on the effect of locally induced photoanisotropy was developed and extensively modified in this program. This material, constituted from an organic polymer, exists in a thin film format and is easily reconfigurable at a low light level. In the course of development, a great deal of effort was spent on understanding the physical mechanism that causes photo-induced anisotropy in the polymer. A theoretical electro-dynamic heat transfer model was developed (section 3.3.4) to help predict the ultimate material performance that could be achieved. Based on this model, experimental data were analyzed. The results clearly indicate that important parameters, such as time response for memory write or erase, power consumption, and memory persistence time, can all be improved through proper material synthesis and preprocessing techniques (section 3.3.4). However, a follow-on investigation is necessary to further delineate this process.

2.0 OPTICAL CACHE SYSTEM ARCHITECTURE DEFINITION

2.1 Principle of Cache Memory

The cache is composed of the fastest memory and is physically located close to the central processor in order to minimize signal transit time. Ideally, the parts of memory most likely to be accessed would be located in the cache. However, it is generally impossible to know in advance of a computation which parts of memory will be most frequently used. Therefore, in practice, sections of main memory are placed into the cache on demand.

In a general architecture, references to the large main memory are "funnelled" through the cache. As shown in Figure 2-1, when the processor issues a request for an address in the main memory, a section of memory (block) containing the requested address is copied into a part of the cache. Subsequent requests for an address within that block are then fulfilled by accessing the cache rather than the main memory, which would be slower. Statistically, there is a high likelihood that a memory request will fall within the same block as the previously requested address.
Most computer programs refer to a relatively small part of memory at a time. Once the appropriate blocks of main memory have been copied into the cache, then the majority of memory requests are met in the cache. The set of blocks frequently used by a computation is called the working set. If the cache holds the working set, comparatively few memory references must access main memory, and the average access time is only slightly greater than that of cache access.

In the event that a memory request fails to be satisfied by the cache (the request misses), the relatively slow operation of copying the appropriate block into the cache must be performed. A miss thus may cause a delay of much longer than the access time of the main memory. It is crucial that the memory request be resolved (i.e., that it be determined whether the requested address is already in the cache or not) with the greatest of speed, as this determination must take place for each memory request.
2.2 Problem Areas

Ideally, in a freely loadable cache system design (see Figure 2-2), blocks of main memory may be placed in the cache as needed and in any order, leading to the best possible match of the cache contents with the actual working set of the computation. However, current machines do not have this design. Several parameters or figures of merit are relevant in a current cache system. These are the access times of the cache memory itself and of the slower main memory, the time taken to copy a block in the event of a miss, and the time with which a memory reference can be resolved. The latter is the chief bottleneck in presently used systems.

Fully-Loadable (Associative Cache)

The resolution process may be analyzed as follows. When the processor issues a memory request, the block in which the requested address falls must be identified. This poses no special problem. Memory is organized so that the most significant bits of the memory address give the start address of the block (or simply block address). Next, the requested block address must be compared with the list of addresses of blocks which are in the cache. If a match is found, that block is determined to reside in the cache; if not, the request must be satisfied from main memory and the corresponding block copied into the cache. \textit{One of the key factors is the speed of comparison of the requested block address with the existing cache list. In a freely loadable cache}, the list may
be long and the comparison must thus be made in parallel rather than sequentially, or it will be prohibitively slow.

The comparison process is equivalent to the action of a memory in which the desired block address is used as the look-up key. The result of the look-up is the cache location of the desired block, with a null result if the block was not in the cache. If a standard memory organization is used for this, a location must be provided for each block in main memory. For instance, if there are $10^6$ blocks in main memory, the comparison requires a memory of $10^6$ storage locations. Note that this memory must be at least as fast as the cache, or the cache speed is wasted. This scheme, then, would require a fast memory as large as the cache just to manage the cache, and a memory request could not be faster than two cache access times.

In present cache architecture, this problem is avoided by restrictions on the loadability of the cache, so that the desired block address need be compared, in effect, with only a few cache block addresses. This simpler comparison is handled at high speed using special registers. The disadvantage of this is that the cache holds a much poorer approximation to the true working set, in most cases, so that the number of misses and consequent delays is large.

In order to preserve the freely loadable caching design, it is necessary to use a content-addressable memory (CAM) to perform the comparison. The addresses of the cache-resident blocks are stored in the CAM, and the desired block address is used as the look-up key. There is no actual datum stored in the CAM for each cache-resident block besides its order in the cache.

Figure 2-3 gives a schematic of a cache memory system using a CAM to resolve memory requests. The central processor (CPU) issues a memory request, an address which appears on the memory bus. The desired block address is the first $N_B$ bits of the address, where $N_B$ is the number of bits needed to specify the block address. The desired block address passes to the CAM, where cache address look-up is performed, and the output is a vector which determines that the requested address resides in the cache (case 1) or not (case 2). In case 1, the address has been determined to reside in the cache and is read from there. In case 2, the address is not in the cache and is read from main memory.

Figure 2-3 does not show the circuitry needed to copy blocks of storage to and from the cache, nor that needed to decide in which cache block a new memory block copy is to be placed. These functions are handled well using conventional means, and are beyond the scope of this research.
In summary, the functions of an ideal cache memory architecture should include the following:

1. The fast and slow memories themselves;
2. The circuitry to copy blocks between the cache and main memory;
3. The arbitration circuitry to decide where in the cache to place new block copies;
4. The CAM to determine whether and where a requested address is to be found in the cache;
5. The memory and data buses.

In the present work, only item 4, the CAM was handled optically. This is due to the limitations of our memory material, which is very fast for reading and quite slow for writing. Since the cache memory itself must be both read and written in large blocks, optical implementation of the cache itself has not yet been fully accomplished. Note that the slowness of writing of an optical CAM is not a large problem, as the CAM need be written into only when a block is copied, and then only one location need be changed and copied into the system.

Figure 2-3 A Hybrid Cache Memory System with Content Addressable Memory Implemented by Optical Means. The cache and the main memory are implemented electronically through virtual memory.
2.3 System Architecture Optical Implementation

2.3.1 Highly Parallel Optical Cache Memory Management

As identified in the previous section, the problem of rapid determination of whether a desired memory address is resident in the cache (memory request resolution) is of major importance and is the primary candidate for solution by highly parallel optical methods. As has been stated, this problem may be solved by using a content addressable memory (CAM) which may be implemented optically in a number of ways.

In previous developments [2,3], holographic techniques have effectively demonstrated optical direct matrix multiplication (parallel inner product) and the cross-correlation function. Among these optical signal processing methods, the inner product approach apparently has immediate application in the equality check of the address of the desired block and those of the blocks in the cache. However, typical holographic storage materials such as dichromated gelatin (DCG) and silver halide are storage capacity limited. Theoretically these materials can achieve up to $10^{15}$ bits/cm$^3$ user data capacity, but in order to attain such status, any storage medium utilizing volume holography effects must have large material thickness, while simultaneously preserving low optical absorption and minimum information distortion. Unfortunately, the results of the research performed so far in holographic ultra-high capacity storage are limited by exactly such fundamental material requirements. Aside from the issue of storage capacity limitation, there is an inherent obstacle that prevents holographic storage from being used in dynamic applications such as cache memory. This is the fact that holographic memories are unable to be erased or rewritten because the recording process is irreversible. Therefore, in order to demonstrate a fully loadable optical cache, the remaining challenge is two-fold. First, we have to define a modified address equality check system based on the inner product concept and especially the characteristics of the address storage medium. This system architecture may depart from the original approach that was used in standard holographic methods. Second, a dynamic memory material must be identified (or developed) and applied in the system. Although material study has been a major part of this work, for the moment we will concentrate on the issue of constructing a parallel address checking system with fully optical implementation.
2.3.2 The Vector-Matrix Inner Product

As shown in Figure 2-4, a digital vector which represents the block address in memory can usually be transformed to an optical signal by using a spatial light modulator (SLM). The spatial intensity profile emitted from the SLM has a one to one correspondence to the input digit vector (i.e., the values of 1 and 0 correspond to dark and bright). The inner product may be used to detect equality as follows. First, note that we are comparing digital vectors which represent block addresses. The vector elements are bits, so that the elements and their products take on the values 0 and 1. Let the address of the block containing the desired memory location be denoted by \( \vec{B} \), a vector whose \( N_B \) elements are the bits representing the block address. The inner product of \( \vec{B} \) with another vector \( \vec{V} \) is given by:

\[
P = \sum_i B_i \cdot V_i
\]  

(1)

The inner product is zero if the two vectors are logical complements of each other, so that \( V_i = 1 - B_i \) for all \( i \). It appears that a satisfactory equality check is obtained by taking the inner product of one block address vector with the complement of the other. However, this is not a unique result. It can be easily proven that the number of vectors \( \vec{V} \) for which \( \vec{B} \cdot \vec{V} = 0 \) is equal to \( 2^{N_Z} \) where \( N_Z \) is the number of zero bits in \( \vec{B} \).

From this it is clear that taking the inner product of the address vectors is not a foolproof equality check. However, there is a clever way to do this equality check. Let the block addresses be given as digital vectors formed by concatenating the bits of the address with the bits of the complement of the address. For instance, if a 5-bit block address is 11001, then the vector representing the block address would be (1, 1, 0, 0, 1, 0, 0, 1, 1, 0). It can be shown that the inner product of one block address vector thus defined with the conjugate complement of another is zero if and only if the block addresses are the same.

The optical implementation of this vector-matrix inner product concept is illustrated schematically in Figure 2-4(a). For simplicity, we assume the storage mechanism is still based on holographic methods. The dynamic storage case will be discussed further in the next few sections. In holographic storage, a plane wave (the reference) is used to interfere with an object wave, and this interference pattern is recorded in a storage medium. When the object wave is derived from an SLM, the spatial matrix characteristics of the SLM with each of the rows designated as a vector, is thus recorded. The position of each individual vector in the memory plane can be recalled and
identified while a proper "search" vector is illuminates on the memory. As shown in Figure 2-4(b), this vector identification process is expanded from 2 bits/vector to 4 bits/vector. When the number of bits is increased in a vector, it is necessary to use a cylindrical lens to collect those diffracted light signals from the holographic memory. The diffracted lights are added incoherently in one dimension by the cylindrical lens. It is thus clear that a designated vector in the matrix memory can be identified if and only if a null signal is detected. More important is the parallel vector-matrix inner product function performed in this scheme. In terms of the cache memory application, it is important to note that the matrix is equivalent to the desired CAM, as shown in Figure 2-5. Its product with the vector representing the desired block address is a vector with \( N_c \) elements. This vector has all non-zero elements, except for a zero in the position corresponding to the cache address matching the desired address. If there is no match, there are no non-zero elements.

The speed of existing optical matrix multipliers is limited by the speed of the spatial light modulators (SLMs). At present, a 256 x 256 SLM operates normally at about 1 kHz; this figure is expected to increase to about 1 MHz and then to 10 MHz within the next 3 to 5 years. The faster rate would imply a single raw modulation time of 4 ns. This time, slower than the fastest semiconductor memory access times, is nevertheless small enough to warrant development.
Figure 2-4(a) The Basic Principles of Using Spatial Light Modulators (SLMs) to Record and Recognize Digital Vectors. The non-erasable holographic memory is used in this case as an example for vector-matrix inner product. Three processes which correspond to memory recording, replay, and identify are shown as labeled. A 2 x 2 SLM is used in this case. Note that the lens system always produces an image with reversed parity on the mirror.
Consider the identification of the first row which is a vector of (0101). An input object must be the conjugate of this vector. Thus $= (1010)$. 

Let every row be the same.
For second row identification, the input vector again will be \((1001)^* = (0110)\)

Suppose we input a vector \((0011)\) which is not stored in the hologram. What will happen? As usual \((0011)^* = (1100)\)

Figure 2-4(b) A Row Vector in Holographic Memory with 4 x 4 Stored Data. A cylindrical lens is placed behind the memory plane to perform one-dimensional intensity summation of the emerging diffraction signal. The identification is shown by the positions of zero intensity at the focal plane. Each position determines the resident memory position. Also shown in the example is a null memory detection in which vector-matrix product does not produce any zero intensity summation.
\[ 2N_B = 2 \log_2 N_M \]

\[
\begin{bmatrix}
\{\text{block address}\}^* \\
\{\text{block address}\}^* \\
\vdots \\
\{\text{block address}\}^*
\end{bmatrix} \cdot \begin{bmatrix}
\{\text{desired block address}\} \\
\{\text{desired block address}\} \\
\vdots \\
\{\text{desired block address}\}
\end{bmatrix} = \begin{bmatrix}
\text{Result Vector} \\
\text{Contains Zero} \\
\text{In Position} \\
\text{Corresponding} \\
\text{To Cache Location} \\
\text{Of Desired Block}
\end{bmatrix}
\]

* denotes complement

Figure 2-5  Mathematical Representation of the Address-Resolving CAM Based on Matrix Multiplication.

The above method represents the inner product approach. During the course of system architecture definition, we also investigated an approach that make uses of the holographic outer product method. The mechanism of obtaining an outer product is substantially more complicated than that for obtaining an inner product. An example of a holographic outer product is illustrated in Figure 2-6. In this case, each column vector of a matrix is stored in a pixel of the memory. The whole memory plane constitutes a matrix. Originally, this matrix is recorded from the illumination of diffused SLM emission, vector by vector at each pixel. Upon reconstruction (or identification), the parallel vector-matrix out-product is performed as shown in the Figure 2-6(c).
In recording, the column vectors of a matrix are stored into individual locations of a hologram. Each hologram is recorded separately, and the holographic recording is achieved by using a diffused SLM pattern. While in reconstruction, the search light of the designated input vector is used to illuminate the memory from the opposite direction. Each component of the search vector is fallen onto a corresponding subhologram's location and the diffraction of each subhologram generates the result of an outer product.
To Identify (0011) Conjugate (b_1 b_2 b_3 b_4) = (1100)

\[
\begin{pmatrix}
0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0
\end{pmatrix}
\quad \rightarrow \quad
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
a_{31} & a_{32} & a_{33} & a_{34} \\
a_{41} & a_{42} & a_{43} & a_{44}
\end{pmatrix}
\]

\[
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
a_{31} & a_{32} & a_{33} & a_{34} \\
a_{41} & a_{42} & a_{43} & a_{44}
\end{pmatrix}
\begin{pmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4
\end{pmatrix}
=
\begin{pmatrix}
a_{11}b_1 + a_{12}b_2 + a_{13}b_3 + a_{14}b_4 \\
a_{21}b_1 + a_{22}b_2 + a_{23}b_3 + a_{24}b_4 \\
a_{31}b_1 + a_{32}b_2 + a_{33}b_3 + a_{34}b_4 \\
a_{41}b_1 + a_{42}b_2 + a_{43}b_3 + a_{44}b_4
\end{pmatrix}
\]

Figure 2-6(b) An Illustration of the 4x4 Matrix Storage.
To Perform the Matrix Multiplication

To Perform the Matrix Multiplication

The First Term is Identified as the Stored Address

Figure 2-6(c) An Illustration of the Outer-Product Matrix multiplication Using a Previously Stored Matrix.
There are a couple of reasons which prevent this approach from being adopted in our system. The first and the most obvious reason is the poor reliability of the holographic pixel to pixel recording. Because diffused light which carries column vector information is involved, it is not possible to use an acousto-optic deflector for fast, nonmechanical pixel to pixel recording. Secondly, the outer-product has advantages over the inner-product only when static nonerasable memory is considered. Due to the particular spectral and polarization characteristics of the dynamic storage material which has been developed in this program, it is physically difficult to implement the outer-product approach in a parallel processing architecture.

As a result, we have made use of the holographic inner product method in a static memory system to verify that the optical signal to noise ratio (S/N) of an address match is high enough to determine the match location. As shown in Figure 2-7, the parallelism of the system thus constructed is handled by a fan-out hologram which disperses a search vector (requested from the CPU and converted into an SLM pattern) into four different memory planes (each memory plane constitutes a page).

![Figure 2-7](image)

*Figure 2-7  The System Layout of a Non-Erasable, Parallel Accessing Cache Memory System Using the Inner Product Method.*
The vector-matrix innerproduct is performed simultaneously in these four pages. Inclusion of time delay and mechanical movement are not required in this scheme. In Figure 2-8, the system operation, which incorporates a $10 \times 10$ SLM, is shown. First we see the global set up, then a four channel fan out of the search vector onto the holographic memory plane and then the results of the simultaneous inner product.
Figure 2.8  The Demonstration of the Operation of a Non-erasable Cache Memory System. (a) The global view of the system set up. (b) the parallel 400 bit matrix multiplier. (c) Identification signals of the vector-matrix inner product. The one dimensional signal is derived from the utilization of a large aperture cylindrical lens.
Although this system can only perform parallel search on static memory, its advantage in fast data management has been verified. Furthermore, based on this preliminary system, we are able to evaluate several parameters that are essential in further development of a read/write/erasable memory addressing system. These parameters are the scaling factor of storage capacity, the access time, the throughput rate, and the bit error rate. Each of these parameters will be addressed in detail.

A. Optical Cache Storage Capacity
As shown in Figure 2-7, the memory storage arrangement in our system is based on a two dimensional bit plane approach. The storage capacity in this scheme is determined by the total number of pixels of an SLM and the ultimate number of pages of SLM pattern that would be possible to record. For instance, a 256x256 pixel SLM shall yield 60k individual addressable bits and, because of the concatenation requirement to the row vector, the binary information carrying capacity of each SLM frame is reduced to 30k bits. If the SLM frame is page multiplexed by 1000, the bit plane memory system will approach a user data capacity of $3 \times 10^7$ bits.

B. Access Time
Since the parallel optical look-up is implanted in the cache system, the time it takes to access the entire memory is essentially the same as the time needed to access a single bit. Taking into account the speed of light, we estimate that the actual access time in our scheme is less than 0.1 nsec.

C. Throughput Rate
The throughput rate is defined as

$$R_{th} = \frac{\text{No. of bits accessed}}{\text{Total access time}}$$

In the electronic cache system, a typical 1 Mbit content addressable memory requires 3 ms (assume the electronic access time for each bit is at its shortest, i.e. 3 ns) to be completely accessed if the conventional sequential accessing mechanism is considered. This will yield a throughput rate of $R_{th} = \frac{10^6\text{bits}}{3 \text{ ms}} = 3.3 \times 10^8$
bits/sec. For the parallel accessed optical cache, however, the throughput rate is significantly higher. It is

\[ R_{th} = 10^6 \text{ bits/}(0.1 \text{ nsec} + 0.1 \text{ msec} + 0.04 \text{ µsec}) \geq 10^{10} \text{ bits/sec} \]

where we have taken the total access time as a summation of the estimated parallel accessing time (0.1 ns), the SLM framing time (0.1 ms), and the detector rise time (0.04 µsec). It should be noted that the SLM framing time is the longest and thus it becomes the dominant factor in the denominator when the approximation is taken. Nevertheless, we still observe a two order of magnitude increase in \( R_{th} \) for the optical cache system.

D. Bit Error Rate (BER)

To check the Bit Error Rate, a software program (see Appendix A) which statistically counts the cache miss hit ratio was implanted into the static cache system. The simulation result shows that the BER will not converge into a narrow range but rather fluctuates between the value of \( 10^{-10} \) and \( 10^{-12} \). Since this simulation is based on a static cache system, such a broad range BER is expected. In order to further evaluate the BER rate in detail, a dynamic system which is capable of continuously updating memory storage as desired.

The basic function of a cache memory system is to be in coordinate with the central processing unit in real time environment. The reconfigurability of the cache should be addressed, therefore, as equal in importance to the fast memory look-up and addressing mechanism. In the next section, the results of study of dynamic all-optic memories is presented.

3.0 RECONFIGURABLE STORAGE MATERIALS

The search for proper materials for all-optical data storage has become a critical issue in the development of a practical system that would provide high capacity and high accessing speed operation.

In order to characterize a material to determine whether it is a candidate for optical information storage, four properties must be understood. They are:

- the material's light/matter interaction response
- the material's ability to retain information
- the driving power requirements for light/matter interaction
- and the material's response speed.

Each property will be discussed separately below.

In light/matter interaction, optical material must respond to change in intensity, polarization, or wavelength of the illuminating light. In most cases the light is coherent. An appropriate optical material need not respond to all three factors but should respond to at least two of them. Research is currently being conducted on materials sensitive to both intensity and wavelength. Good examples include materials exhibiting two-photon absorption phenomena and materials using the spatial hole burning effect. Both kinds of materials have demonstrated superb potential in performing three-dimensional, ultra-high capacity, and ultra high speed data storage. However, two-photon materials need always be driven by a very powerful laser source because of the intrinsic low absorption cross-section of the non-linear two photon process. Furthermore, memory materials using a spatial hole burning effect require cryogenic cooling.

For optical materials to respond to the change of light intensity, a coherent light source with high peak intensity is generally needed to drive the material. This is because the intensity dependent material characteristic change is attributed to the material's small third order nonlinear optical coefficient. Although there is a great deal of academic interest in the intensity dependent effect, so far, practical application of materials which respond to changes in intensity of illuminating light has been limited due to the high power requirement. This leads to the conclusion that materials which are sensitive to polarization and wavelength are better choices for optical memory storage.

In analogy to magnetic storage, optical memory should have the capability to maintain information. The capability to allow information to be quickly erased is also an important requirement for a dynamic optical storage medium. In the cache application, optical memory is required to persist indefinitely, until an external perturbation is applied to change the medium's storage mechanism. In responding to the perturbation, a relaxation of the memory state to the original unwritten state is desired. Ideally, in order not to consume CPU execution time, the memory erasing shall be completed within a few CPU cycle times, on order of a few nanoseconds. Up to the present, there have been no well publicized reports of such a material. In this report, we describe a candidate material that could fulfil the above requirements.

As mentioned earlier, the material's response to light necessitates that writing/recording be done at a low power level. At the same time, that material's sensitivity must be high enough to facilitate reading. The sensitivity of a material is indicated by the maximum achievable refractive index
modulation that is induced by the external light source. For a static material, such as dicromated gelatin (DCG), $\Delta n$ (index modulation) can reach 0.1. Information stored in this material can not be erased. Dynamic photo-refractive materials, on the other hand, can respond to low light intensity but also have low $\Delta n$. This seriously impairs their storage capacity.

The previously described materials have response speeds varying from nanoseconds to milliseconds. A judicious choice will naturally be materials that respond in the nanosecond range. However, materials that fall into this region are those that use the third order nonlinearity. In this report, we present a material that potentially has a potential response speed of several milliseconds.

The promising optical storage material alluded to above is made of a composite polymer. It can be formed into a thin film relatively easily and it is extremely sensitive to a broadband wavelength, from 500 to 530 nanometers. Its sensitivity depends on the state of the polarization change of blue-green light. In this material, optical information can be recorded and addressed by two methods: holographic and direct birefringent induction. The holographic method requires that two orthogonally polarized single frequency green beams be involved in the recording process and a red probe beam for memory accessing. In the birefringent induction method, only one linear polarized green beam is sufficient for recording and one red beam for reading. Detailed principles of these two methods are given in Section 3.3. We have used both methods to record optical information in the polymer material. In the birefringent induction method, the index modulation has been measured to be as high as 0.01 at a recording power of 1mW. The persistence of information thus recorded was tested two weeks after recording and it was determined that the storage efficiency had less than 20% decay. The polymer sample under testing was kept at room temperature in an open environment and green light perturbation to the sample was completely avoided.

3.1 Basic Properties of Storage Material Using Two Photon Absorption

It has long been proposed that a 3-D parallel optical memory device\cite{4} may be used to construct a 3-D parallel optical memory system. Such an optical memory device is based on volume and erasable optical storage in an amplitude-recording medium Figure 3-1. Originally, the memory mechanism in this device was achieved by two-photon excitation\cite{5} of the two forms of a Spirobenezopyran\cite{6} embedded in a spiropyran form (I) via heterolytic cleavage (Figure 3-2). The merocyanine form absorbs in the green-red region of the visible spectrum and emits red-shifted fluorescence when excited with green light.
Figure 3-1  Schematic Diagram of a 3-D Optical Memory Based on a Two-Photon Recording Process. The overlapping spot where two beams intersect will exhibit a different fluorescence spectrum than the material normally does.

Figure 3-2  The Two-Photon Absorption Material Which Has Been Reported is Based on the Photochromic Change From the Spiropyran Form (I) to the Merocyanine Form (II).

The entire excitation scheme in this method is based on a "virtual" two-photon process in which neither of the two photons can be absorbed individually: both must be absorbed simultaneously, which necessitates that the two beams overlap in time and space within the storage medium. This is in contrast to a sequential biphotonic process, or photon-gated hole burning, in which each
photon is absorbed independently in real quantum levels. Since the two photon absorption process occurs only at the place where the two beams overlap, within the volume of the memory, this process can be truly implanted into 3-D information storage.

The advantage of such a 3-D memory are (i) immense information storage capacity, (ii) random access, (iii) parallel addressing, (iv) very fast optical writing and reading speed, nanoseconds and faster, (v) small size and low cost, (vi) absence of mechanical or moving parts, (vii) minimal cross-talk between adjacent bits, and (viii) high reading sensitivity. However, there is a major drawback which prevents such a material from being used in a real application. This drawback is due to the nature of the two-photon process. As mentioned previously, a two-photon process relies on the simultaneous absorption of two photons (since both the writing and reading processes have to be excited at different spectra) to excite a material system (Figure 3-2). Being a higher-order process, two photon absorption cross section often is many orders of magnitude smaller than that of a one-photon absorption. Traditionally, the transition probability per unit time per unit volume per unit energy interval of a two photon absorption is measured by the two photon absorption coefficient $\gamma$ which is [7]

$$\gamma = \frac{8\pi^2}{C^2 \epsilon^{1/2}(\omega_1) \epsilon^{1/2}(\omega_2)} \text{Im} \chi^{(3)}$$

$\epsilon(\omega)$ is the dielectric function, and $\chi^{(3)}$ is the third order nonlinear susceptibility of the material system. Because $\chi^{(3)}$ is the third coefficient term in a polynomial expansion for the materials susceptibility, its value is usually smaller than $\chi^{(1)}$ (i.e. the first order susceptibility that is associated with a single photon absorption coefficient) by at least seven orders of magnitude. It is now clear that in order for such material to be used in digital data storage purposes, a powerful picosecond laser is absolutely required to serve as the pump source in initiating the two-photo absorption for writing and reading.
3.2 Basic Properties of Materials Using Spatial (Spectral) Hole Burning Effect

A new class of optical memory has recently been proposed [8-10]. It holds the promise of making new atomic-level data storage a reality. This approach to storage, generally referred to as frequency-selective optical data storage, works by addressing atoms spectrally as well as spatially (Figure 3-3). The constituent atoms/molecules within many materials display spectrally narrow resonances (with a width referred to as \( \Delta v_h \)), and the resonance of individual atoms/molecules is spread throughout a rather broad frequency range (referred to as \( \Delta v_l \)). Atoms located within a minimally sized spatial volume can be subdivided and hence addressed on the basis of their frequencies. In some materials, up to \( \Delta v_l/\Delta v_h = 10^7 \) frequency subdivisions can be made, making it possible to subdivide the billions of atoms/molecules within a minimally sized spatial storage volume into separately addressable groups containing only a few hundred atoms each. The atoms/molecules belonging to each spectral group are generally positioned randomly throughout the spatially addressed storage volume with which they are associated.
A simple configuration for temporally addressed memory is shown in Figure 3-4. Two pulses are involved in the storage of a bit stream in a single spatial storage cell. The first pulse, referred to here as the preparation pulse, excites the active atoms/molecules into a superposition state involving the ground state and excited state 1. The intensity of this pulse should be adjusted so that it transfers roughly half of the ground state atoms within its bandwidth to the excited state. The preparation pulse may be temporarily short compared to the bandwidth of the bit stream to be stored, or it may be longer and sweep in frequency across the bandwidth of the bit stream. The
second pulse (i.e. the data pulse) consists of the actual bit stream to be stored. Its intensity, like that of the preparation pulse, should be adjusted so that it would, by itself, transfer approximately half of the population in the ground state and within its bandwidth to excited state 1. The central frequency of both storage pulses should lie within the inhomogenous absorption profile of the storage material. The total temporal interval between the state of the preparation pulse and the end of the data pulse should be less than approximately $\Delta \nu^{-1}$. The combined action of the preparation and data pulses acts to record the Fourier transform of the data pulse as a feature of the absorption profile of the storage material. If the preparation pulse is not short, the transform is recorded in a complex form.

![Diagram of pulse sequence and material](image)

Figure 3-4  An Optical Arrangement for Temporally Addressed Memory Using a Spectral Hole Burning Effect.

Immediately following the writing sequence, the bit stream is stored in complementary frequency-dependent population distributions in the ground state and in excited state 1. If excited state 1 decays exclusively to the ground state, the stored information will disappear over the lifetime of the excited state 1. For the information to be retained, the excited state population must be prevented from recombining with the population in the ground state. In some materials, the ground state contains sublevels and the excited state decays back to them in a way which differs from the way in which the writing pulses depopulate them. In these cases, information can be retained as a frequency dependent non-thermal distribution of population among the ground state sublevels.
In the level scheme shown in Figure 3-3(a) long term data storage is achieved by applying a gating pulse, resonant with the transition between excited states 1 and 2, following the writing pulses and before excited state 1 decays. If excited state 2 decays primarily to the alternate ground state, the complementary population created during the storage sequence will be prevented from recombining over the ground state-alternate ground state thermalization time.

The recall sequence consists of an input read pulse followed by a single pulse. The read pulse must be essentially identical to the preparation pulse; however, it may in principle be resonant with a transition coupling the ground state with another excited state. In cases studied to date, the read pulse and the preparation pulse have been resonant with the same transition. If a second transition is to be employed, its inhomogeneous broadening must be correlated with the inhomogeneous broadening of the transition used in the writing sequence. The read pulse stimulates the storage material to emit a signal pulse having the same temporal envelope as the data pulse. The signal pulse is a coherently emitted signal, and as such it is highly directional. If all input pulses are collinear (Figure 3-4), the signal pulse will have directional properties essentially identical to the data pulse.

Temporally addressed frequency-selective optical data storage has been demonstrated in the laboratory [11,12], and the basic optical physics related to the storage and retrieval process is now well understood. The question of whether this form of memory will ever find significant application will depend primarily on the success or failure of attempts to find suitable storage materials. All long-term storage materials demonstrated to date are useful only when cooled to near liquid helium temperatures (4K) in order to make AV reasonably small, and this situation is unlikely to change. Materials exhibiting the ideal level system shown in Figure 3-3 have not yet been found. However, if an ideal material is found, near atomic scale memories with storage densities up to $10^{14} \text{ - } 10^{18}$ bits/cm$^3$ and input/output rates in the multigigahertz range may become achievable.

3.3 Organic Storage Material Based on Photoinduced Anisotropy

As identified in the beginning of this report, in addition to materials which make use of two photon absorption phenomenon or the hole burning effect, a new class of organic material that is unique in its response to the illuminating light has been applied in this program as a dynamic memory storage medium. All the disadvantages of the two photon absorption and spatial hole burning mechanisms
are invisible to the synthetic polymer because its storage mechanism does not rely on the quantum mechanic state excitation but on fast photoinduced electric dipole rotation.

The basic structure of the polymer in use can be best described as a guest-host system. Organic dye molecules which have high polarity were chemically cast into a suitable polymer matrix in which each molecular dipole is oriented with maximum randomness. The original development of this photosensitive polymer was for holographic recording purposes. Thus, no attention was paid previously to its dynamic characteristics or to its performance in non-volatile memory storage. However, we were able to investigate this material for the cache memory application. In the course of studying the formation of the polymer film, several modifications were made to enhance its performance in time response and memory persistence. Meanwhile, based on classical electrodynamic theory, we developed a local field correction model to assist in the analysis of data measured according to the modification process.

In order to present this peculiar material in more detail, this section is organized around several topics. Each topic deals with one particular aspect of the material. The operation modes or storage mechanisms will be presented first. The following subsection will show the microscopic nature of this material system and then an electromagnetic theory derivation which describes a closed dipole system will be presented. In accordance with the modeling theory prediction, we will show quality as well as quantity improvements in rise time, persistence, and modulation bandwidth of a modified material system when compared with the original material.

3.3.1 Recording Principles Using Photoinduced Anisotropy in Organic Polymer

There are two operating modes which exist for using a synthetic dye polymer in memory storage. They are the holographic mode and the birefringent mode. In principle, as discussed later, these two modes are based on the same physical mechanism. However, due to different application requirements, the birefringent mode is more convenient for use in reconfigurable cache memory because of the much simpler optical arrangement (only one green polarized beam and one red beam are sufficient for memory writing and erasing) that can be configured into a parallel accessing system. Furthermore, the birefringent mode poses the advantage of insensitivity to environmental instability and therefore higher memory recording efficiency in a short time period.
3.3.1.1 Holographic Mode Recording

As in a typical volume hologram recording geometry, the optical memory is stored in the polymer via two beam interference (Figure 3-5). However, in order to achieve the maximum storage efficiency with these two beams (one actually carries the information and the other is a reference plane wave) they have to be orthogonally polarized to each other. The reconstruction (or diffraction) efficiency reaches a maximum (~70%) when the two orthogonal beams are circularly polarized. In contrast to normal holographic recording, this orthogonal polarization requirement indicates that the memory storage is achieved through periodic birefringent modulation instead of direct refractive index modulation.

According to the absorption spectrum of the polymer, the recording beam's wavelength is usually chosen to be close to the absorption peak. After the recording process, the memory is non-volatile as long as it is retrieved or probed by a coherent light whose wavelength is outside the polymer's absorption band. In erasing the hologram, a uniform illuminating light with arbitrary polarization but the same wavelength as the writing beams can restore the polymer to its original unwritten state. Typical wavelengths that have been tested for recording and erasing are the 488nm and 514nm lines of the Ar+ ion laser. The reading wavelength is 630nm.

There is a stumbling block that prevents the dynamic holographic recording in the polymer from becoming more practical. This arises from the stability requirement of the two beam interference. As encountered in all conventional hologram recording, in order to facilitate efficient memory storage, high temporal coherence of the interfering light must be maintained. Any environmental fluctuations drastically influence the memory storage process.
Figure 3-5 Recording Geometry for a Polarization Hologram: (a) R and S are the two orthogonal recording waves lying in the plane normal to the XY plane and subtending an angle $\theta$. The horizontal axis $\xi$ lies in the XY plane at angles of $\pm 45^\circ$ to the coordinate axes $x$ and $y$. (b) Arrangement for parallel image recording using polarization holography. EPD-60 is the internal code number of the birefringent polymer developed at Physical Optics Corporation.

The consequence of any disturbance in recording is reduction in speed. In fact, we have experienced an average time response of approximately 2 minutes in every one of our polymer samples used in the holographic mode.

Although we have shown that reversible and orthogonal holographic recording can be performed via utilization of the synthetic dye polymer, no explanation has yet been made to show in detail the time evolution of the material's microscopic structure change.

Nevertheless, in another operating mode (which is substantially simpler than the holographic mode), we are able to perform basic analysis of the material's response to the writing beam, and
from it gain knowledge needed to control the material's properties by modifying its microscopic structure.

3.3.1.2 The Birefringent Mode (Operating Principle and Application)

As shown in Figure 3-6, a piece of dye sensitized polymer film is inserted into a pair of cross polarizers, probed by a HeNe laser light ($\lambda = 632$nm), and illuminated from the side with an arbitrary linear polarized argon ion laser light ($\lambda = 514$nm). Initially, when the green pump light is blocked, the red probe light has no difficulty passing through the first polarizer and the polymer film. It is then stopped by the analyzer. However, immediately after the illumination with the green light, a red transmitting light is observed after the analyzer. With time, integrated intensity builds up. Because the green light and the red light fall on the same spot on the polymer film, it is apparent that the green light induced local anisotropy has turned the material from its isotropic state into a birefringent state. When the red polarized light is transmitting through this anisotropic area, a rotation of the polarization will be experienced by the red light just as it is passing through the crystal wave plate.

![Figure 3-6](image)

**Figure 3-6** The Optical Arrangement for Birefringent Mode Recording.
It is evident that local birefringence can only be induced at a location where the polymer film interacts with a polarized green light. This birefringent mode operation can be further generalized for two dimensional digital memory storage if the polarized green beam is embedded with a binary spatial pattern (see Figure 3-6). Prior to direct utilization in memory storage, we shall further address the material physics which explains this polarization sensitive mechanism. By understanding the microscopic nature of the material, we hope to make improvements in its synthetic and processing techniques which can optimize the material's kinetic performance. Basic features that need to be addressed with respect to the material's performance are the sensitivity, power consumption, and persistency.

3.3.2 Material Samples

In accordance with previous experiments\textsuperscript{[13]}, we start our sample preparation by introducing methyl red dye into a polymer matrix. Methyl red is an azo dye in which, under the action of linearly polarized light, some degree of ordering of the molecules can be achieved, leading to dichroism. Our experiments showed that the molecules of methyl red can be oriented reversibly many times without fatigue. The reverse process (i.e., destruction of the stereoregularity) depends on the polymer matrix. In darkness it may last from a few seconds to several days.

In regard to the matrix, polyvinyl-alcohol (PVA) turned out to be most suitable for our purposes. It is a well known material\textsuperscript{[14]} in which an optical axis along the direction of the force can be induced through mechanical deformation. Thus, the material becomes optically anisotropic -- it acquires birefringence ($\Delta n=n_e-n_o$) over the whole visible region.

Both methyl red and PVA are water soluble. Hence, the mixing of these two components easily produces a uniformly colored polymer in liquid phase. The samples are then made by spin coating the liquid onto microscope glass slides and dehydrating them in a 69°C oven for 2-10 hours.

3.3.3 Investigations into Photoinduced Anisotropy

Figure 3-7 shows the absorption spectrum of a 60-μm thick layer of PVA containing 0.02-wt% methyl red. Since PVA is transparent in the visible region, absorption is due entirely to the dye whose stable form is its transisomer. In later experiments, samples with various dye
concentrations were used. It is shown that there is a direct correspondence between high dye concentration and high photoinduced anisotropy.

After absorbing the light of an Ar+ laser beam with a wavelength of 514 nm, a large part of the dye molecules undergoes spatial (trans-cis) isomerization (see Figure 3-8). The eisform of azo dye is thermally unstable and a reverse thermal process of restoration begins. Also, if the light is linearly polarized, the molecules tend to order themselves in such a way that the direction of their actual transition is perpendicular to the polarization direction of the light [13]. Thus, the optical transmission for light, polarized along the polarization direction of the acting light, increases; for light polarized perpendicular to the polarization direction of the action light, it decreases. Dichroism is induced.

Figure 3-7  Absorption Spectrum of a Methyl Red-PVA Polymer, the Peak Wavelength of Absorption is at 430 nm.
The kinetics of the anisotropic change in optical transmission is shown in Figure 3-9. Measurement was accomplished with a low-intensity light beam of the same wavelength. The intensity of the acting Ar laser beam was 100 mW/cm². After the green acting light is removed, the sample begins to restore to its original state and the ordering of the dye molecules is lost. The degree of restoration after the recording light is switched off and the rate of this process depend on the polymer matrix and in PVA it can be altered by thermal treatment. In the case corresponding to Figure 3-9(a), this treatment was not carried out and the original optical transmission was restored almost completely in ~30 seconds.

Along with dichroism, light birefringence is induced in the films under the action of Ar⁺ light. Photoinduced birefringence, particularly for wavelengths outside the optical transmission, is of greater interest from a memory storage point of view. We investigated birefringence qualitatively at λ=633nm. For this purpose, the samples were placed in the birefringent mode set up shown in

Figure 3-8 Trans-Cis Reversible Photoisomerization Mechanism.

Figure 3-9 (a) Optical Transmission Change at λ=632nm Under the Action of an Ar⁺ laser beam. The moment of switching on and off the acting light are marked with arrows. (b) Kinetics of the light signal-change after passing through a pair of crossed polarizers, due to the thermal treatment the birefringent state has long "memory" after the acting light is removed.
Figure 3-6. They were irradiated with an Ar⁺ laser beam, linearly polarized at ± 45° with respect to the directions of transmission of P and A (P=Polarizer; A=Analyzer). The optical anisotropy induced in the sample under the action of the blue light causes a light signal to appear after the analyzer A. At λ=633 nm, this transmission is due solely to the photoinduced birefringence which is given by

\[ I = I_0 \sin^2 \left( \frac{\delta \phi}{2} \right) \]

where I is the intensity of the transmitted light after A, \( I_0 \) is the intensity of the light passing through the pair of parallel polarizers,

\[ \delta \phi = \frac{2\pi}{\lambda} \delta n \cdot d \]

d is the thickness of the sample (100 μm), and \( \delta n \) is the birefringence.

The values of \( \delta n \) depend on the intensity of the Ar⁺ beam, dye concentration, and preliminary treatment of the polymer. In optimal conditions, \( \delta \phi \) of the sample exceeded \( \pi/2 \) (\( \delta n = 0.01 \)). Such large values for the birefringence make us think that the polymer matrix does not remain unchanged in the recording. The exact mechanism of this process is not clear, but probably PVA molecules are deformed in accordance with the photoinduced ordering of the dye molecules, and that is the reason for the large values of the induced birefringence.

As shown later, the thermal erasure of the birefringence reached depends on the sample and may be slowed down substantially by preliminary thermal treatment. The sample, whose photoinduced change is shown in Figure 3-9(b), was preheated to 80°C for 30 minutes. After this treatment, a large (> 50%) part of the photoinduced birefringence is retained for a long time after the acting light is removed. These samples have the ability to store information for several days.
3.3.4 Physical Properties of Guest-Host Dye Polymers and Their Modifications for Memory Storage

The organic material that has been investigated belongs to a guest-host material system. Dye molecules reside in the polymer matrix in the absence of any bonding force. In general, neither cross-link bonding nor copolymerization exist between the dye molecule and the monomer.

Organic dye has long been recognized by its chain structure. This structure is usually characterized by a high degree of polarizability. Therefore, each molecule can be treated as an electric dipole moment whose direction is randomly oriented in the polymer matrix. Ideally, the randomness of such dye orientation shall stay at a maximum to avoid any local interaction (i.e., it is a macroscopically isotropic system). In order to see how such an isotropic system interacts with incident light, and to reduce the complexity of our analysis, the approach is to start with an ideal isotropic guest-host system. As shown in Figure 3-10, when such a system is illuminated by a polarized light whose field vector is denoted by $\vec{E}$, the coupling of matter to the electromagnetic field is described by the interaction potential [115]:

$$V_{\text{int}} = \sum_a \vec{P}_a \cdot \vec{E}(\vec{r}_a)$$
Figure 3-10 Modeling of an Isotropic Guest-Host Polymer System Under the Illumination of a Polarized Light Whose Wavelength Lies at the Peak of the Dye Molecules's Absorption Band.

Here, \( \vec{P}_a \) is the dipole and \( \vec{r}_a \) is the coordinate of the \( a \)th molecule, respectively. Letting \( q_a \) be the electronic coordinate of the dipole and \( \hat{z}_a \) the effective charge (which is a measure of the strength of the dipole), then,

\[
\vec{P}_a = q_a \hat{z}_a
\]

There is a nonzero force (or torque) acting on the molecule because \( \partial V_{\text{int}} / \partial q_a \) is nonzero. This torque tends to align the dipole along the field direction.

Although the applied electromagnetic field tends to cause each dipole moment to rotate, the rotation is prohibited because the dye molecules are physically "frozen" inside the homogeneous polymer matrix at room temperature. However, if the wavelength of the electromagnetic field lies in the single photon absorption band of the dye molecule, electronic transition of the molecules from ground state to the excited state may occur, followed by the release of the absorbed energy through fluorescence and heat. Figure 3-11 shows this transition mechanism. Note: optical phonon generation is the consequence of heat dissipation.
Figure 3-11 The Transition of an Excited Dye Molecule.

While the heat (phonon) is dissipated, "local" temperature rises at the site and causes the surrounding polymer matrix to undergo a phase transition, from a glassy state to a rubber state. Subsequently, the dipole moments gain much more freedom to rotate within the trapping matrix. According to the polarization of the incident field, locally induced anisotropy is thus introduced by reordering the dipoles from a complete random orientation to a distinct direction.

It must be emphasized that the isotropic system just described is an ideal case. In practical situations, complicated factors such as defects and inhomogeneity of dipole distribution may drastically affect the global performance of such a system when interacting with polarized light. Because photoinduced anisotropy is based on three successive processes in our model (i.e., single photon absorption, heat transfer, and dipole rotation) and in which the time constant for dipole absorption and rotation is on the picosecond and microsecond scale, respectively. It is thus apparent that the heat transfer property of a polymer host will be a dominant factor in determining the rise time of induced birefringence. However, based on our experience, there exist two other key factors in determining the forming time and the strength of induced birefringence. They are:

1. Dye molecule concentration and polymer thickness, and
2. Homogenity of dye distribution and domain structure.
The first factor is purely based upon our empirical observation and no analytical studies have yet been made. The second factor, which originates from the material's microscopic nature can, however, be understood in detail. Because the dye molecule, especially azo dye, possesses large polarizability, the typical formation of dye immersed in polymer matrix is in microcrystalline powder structure if the bonding (between each dye molecule) has not been completely broken down. Figure 3-12 shows schematically such a microcrystalline domain structure. In this structure, the material is no longer macroscopically isotropic.

\[ \bar{E}_{\text{loc}} = \sum_j \bar{E}_{\text{P}_j} \]

Figure 3-12  The Microcrystalline Domain Structure and Local Field Interactions in a Quasi-Isotropic System.
Instead, at each light-matter interaction region, there will always be local fields. These local fields originate from a non-interacting region by the group behavior of the surrounding microcrystal present against the real external pulling field. Due to the presence of the local fields, the polarization reorientation process is perturbed. The consequences of this perturbation are

1. Slower response time,
2. Lower induced birefringence, and
3. Short birefringence persistence.

In order to prevent this level of complication and to improve the PVA-dye performance in time response as well as to retain its induced birefringence, we have developed a proprietary pre-processing technique that applies thermal treatment in a synthetic process to minimize the microcrystalline dye-formation and a special cooling technique to attain a homogeneous polymer medium. Figure 3-13 shows a comparison of time response, time persistence, and modulation bandwidth for the modified sample and the original published sample. In Figure 3-13(a), a comparison is made to show the advancement of the material's time response in birefringent mode operation. The rise time of Methyl Red (MR) doped PVA has nearly a four fold improvement over the Ethyl Orange doped PVA. Both material samples in this test are raw materials, no preprocessing was applied. In Figure 3-13(b), the thermally pretreated MR-PVA samples were tested for memory persistence. Both samples were operated in the birefringent recording mode and one sample was treated by slow cooling while the other sample was treated by fast cooling after the heat treatment. It is observed that the memory recording response of the slow cooled sample differs from the fast cooled sample in three aspects. In comparison with the fast cooled PVA, the slow cooling technique improves the MR-PVA material's sensitivity (note the doubled recording efficiency), the response time (less than 5 second rise time), and the memory persistence after the writing beam is removed. Specifically, the slow cooled sample reveals an improvement that it is capable of retaining 80% of the memory state 60 seconds after the writing beam is removed. In contrast, the fast cooled MR-PVA sample shows a much faster and steeper decay in retaining memory efficiency. In Figure 3-13(c), the modulation bandwidth of the MR-PVA material was evaluated. The basic principle of this measurement is based on the AC birefringent mode in which the polarization state of the green writing beam (see Figure 3-6) is alternated at a constant frequency. The material's response to a periodic polarization change is thus recorded by measuring the transmission of a red probe light. Because the polymer will respond to the polarization change up to the kHz range, the ultimate memory write/erase time is expected to be in the millisecond range.
Figure 3-13
(a) Comparison of the Rise Time Modified Methyl Red-PVA and Ethyl Orange-PVA in Birefringent Mode Operation. (b) The improvement in memory persistence of thermally pretreated methyl red-PVA polymer is evidenced by the slow degradation of He-Ne transmission after the acting Ar+ laser light is removed. (c) The modulation bandwidth is a figure of merit that determines at what frequency the temporal response of material will fall off by half. As shown in the figure, we have tested our material at various frequency levels and, in general, the material's AC birefringent efficiency will reduce to half at the kHz range.
4.0 SUMMARY

4.1 Construction of an Exploratory Development Model

As specified in the contract, the goal of this research is to build an Exploratory Development Model (EDM), of an optical cache memory utilizing existing off-the-shelf components to demonstrate system feasibility, key system concepts, and validation of the overall system approach. In the first half of this report, we have identified an optical parallel addressing architecture which utilizes a vector-matrix inner product as the basic framework for an all optical cache memory system. Furthermore, in order to mimic the dynamic memory requisition and updating sequence that is encountered in an electronic system, a polymeric memory material was studied and its performance was tailored to provide long memory persistence, fast writing, fast erasing, and high efficiency.

The consideration for constructing an EDM system is thus how to perfectly match the memory material to the addressing system and interface the whole memory system with the CPU in real time operation. Because the material configuration can be altered by polarized light whose wavelength lies in the blue-green region, and the memory information can be accessed through red light, two different light sources, one for addressing purposes, and one for memory reconfiguration purposes must be used in the cache system.

In addition to light source consideration, parallel memory access was also implemented through the use of chromatically corrected holographic grating pairs in the read and write/erase channels. These two special purpose grating pairs are designed and fabricated at POC's laboratory. The specific function of the grating pairs is that the first grating can disperse incoming light uniformly into three distinct channels and then collimate it parallelly through the second grating. Because the light dispersion and collimation are accomplished through the diffraction grating pairs, wavelength compensation are thus necessary to ensure that the grating pairs in the memory reconfiguration channels (in which the green light is manipulated) have exactly the same dispersion and collimation angles as the grating pairs that are used in the read channel (in which the red light is manipulated).

The layout of our demonstration cache memory system (as shown in Figure 4-1) is a relatively complicated one. Two light sources of different wavelengths are used. The red HeNe light is dedicated to memory searching. At this wavelength the memory stored in the polymer will not be perturbed. The green Argon laser light, however, is used for the purpose of reconfiguring, rewriting, or erasing the memory. In addition to the light sources, several unique optical components were also put into the system. They are the liquid crystal (LC) light valve and
polarizing beam splitter assembly, the 514 nm transmission/633 nm reflection dichroic beamsplitters, the 10x10 ferroelectric liquid crystal SLM, the nematic liquid crystal polarization rotator, and two sets of image relay optics.

The active components, such as the LC light valve and the polarization rotator, are used to condition the green beam. The light valve is used to switch on and off the green light that enters the system and the polarization rotator is used to change the polarization status of the green beam in either writing or erasing the memory. The dichroic beam splitters which allow total transmission of green light and maximum reflection of red light are used in several places in the system. Their function is to manipulate the propagation routes of the green and red lights which are initially combined before passing through the SLM and reach the polymer memory plane from opposite directions.

Figure 4-1 The Optical Layout of a Cache Memory EDM System.
In the demonstration, the system operation is divided into two sequential modes, i.e., the cache memory searching mode and the memory reconfiguration mode. In the memory searching mode, the green light is removed by using a liquid crystal light valve in its beam path. When the memory search operation is completed, the green light will then be activated and propagate through the SLM (to encode proper information), the image collimating lenses, the dichroic beam splitter (which is used to separate the "read" light and the "write" light), the electro-optic polarization rotator (for proper polarization control of the green beam in memory write and erase processes), the diffraction collimating grating pair (for parallel memory location accessing) and finally fall onto the memory plane. Because of the special properties of the storage material, old memory must be erased first by a uniform green light that has a random polarization state. The new memory can be recorded by an encoded green light of a fixed (usually 45°) polarization. The detailed operation of the read and write beam, including the function of each optical component corresponding to the system layout, is shown in the flow chart in Figure 4-2.
Figure 4-2  Flow Chart of the EDM Demonstrated Processes in which Cache Memory Read and Write/Erase Process is Partitioned into Two Sequential Processes.
In addition to the memory accessing system and the plane memory, we have employed a 10x10 spatial light modulator as the interface between the CPU and the memory system. During the storage sequence, the CPU (mimicked by a microcomputer) issues a string of data which is then displayed on the SLM in a digital vector format. When the green writing light (a plane wave) passes through the SLM, this designated information will be incoded into the light stream. The exact same mechanism is applied to the red reading light during the memory search sequence, except that the read (red) light and write (green) light are manipulated into different routes by the utilization of dichroic beam splitters.

Although the encoded green and red lights reach the memory plane through opposite directions, they fall on the same memory location. In other words, both the control (green) light and search (red) light are acting on the memory in real time. This arrangement allows one to physically monitor multichannel memory writing and erasing processes. Because the red light physically transmits through the memory plane, a vector-matrix inner product (analyzed in previous sections) takes place simultaneously in the three parallel channels and the results can be resolved through an apparatus which is composed of an analyzer, a cylindrical lens, and a CCD line scanner. The CCD array, which runs at a scanning speed of 20 MHz in this demonstration, is the output interface device where the result of the optical memory search is translated back to the CPU for the next processing routing. Figure 4-3 shows the set-up of such a system during operation.
During the final EDM demonstration, the dynamic features of this system were demonstrated by configuring the system in several modes. They are

1. Blank Memory Search
2. Memory Writing (Write through Cache)
3. Memory Persistence
4. Parallel High Throughput Memory Search
5. Designated Memory Erasure
6. Memory Write Back (Write Back Cache)

In the first mode, we simulate system pre-checking. This is performed prior to loading the data requested by the CPU into cache from main memory. In this condition, the memory plane, which is a thin polymer film, has no induced birefringence and the CCD array should not be able to detect any signal read by red light. The result of the blank memory search is shown in Figure 4-4(a). Except for a small amount of noise, the CCD reading is null. In the second mode, the polarized
and encoded green light, split into three channels, illuminates the memory plane and one can observe in real time the parallel growth of the stored memory (see Figure 4-4(b)). The memory storage is shown in digital form. Each of the signal peaks shown on the screen constitutes 1 bit and the S/N ratio achieved in writing is 20. In the search mode, the memory shall not be perturbed by polarized green light. The parallel high throughput memory search is performed prior to checking the persistence in order to show the of long memory persistence in this unique dye polymer. The parallel search is activated by sending the matrix data from the CPU to the SLM (in other words, the SLM is showing different patterns at a high frame rate). Each matrix is thus carried by the HeNe laser light into the read channel, then split into three parallel pages and imaged onto the three memory slots. On each memory slot, there will be an individual bit to be identified (in this simulation, the information stored in the memory slots are identical) if the vector-matrix product is non-zero. Because the memory search is performed simultaneously in these three slots, the throughput rate is incredibly high. Figure 4-4(c) shows the bit to bit search by observing the position shift (from right to left) of each bit identified in the memory slots. After this high throughput demonstration, the SLM was reconfigured so that prestored memory could be read out completely. Figure 4-4(d) shows this read out signal. In comparison with the signal beam shown in Figure 4-4(b), it is clear that only minor decay (<5%) in memory storage has taken place during writing. In Figure 4-4(e) and Figure 4-4(f), selective erasure and writeback are shown. In selective erasure, bits at locations of 341, 1023, and 1643 on the monitor screen are wiped out simultaneously by sending a green laser light with orthogonal polarization (in contrast to the writing light) and desired bit pattern. The last demonstration, which shows the write back cache, is done by reversing the previous erasure process. Because every bit on the memory plane is constituted by localized birefringence, no interconnects are established in writing the bits. One can thus manipulate these birefringent states on an individual basis without perturbing the others. Both the erase and the write back processes have demonstrated this unique property.
Figure 4-4

(a) Rank Memory Search
(b) Memory Writing - Real Time Response (Write Through Cache)
(c) Parallel High Throughput Memory Search (bit position shifting from right to left)
(d) Memory Persistence
(e) Designated Memory Erasure ((1) Designated Erasing Bits (2) After Erase)
(f) Memory Write Back (Write Back Cache)
The above presented optical cache memory system has a current data storage capacity of 300 bits of user data distributed into three parallel memory locations (i.e., each memory location contains 100 bits of user data). The limitation on data capacity is due to the limited number of pixels (10x10) that are available on the spatial light modulator. In fact, the full-scale capacity of this memory module can be beyond $10^7$ bits if a 256x256 SLM is utilized together with a 1000 channel grating divider. Neither the 256x256 SLM or the 1000 channel grating divider is commercially available yet. However, the fabrication of the grating divider should be less technically difficult than that of the large array SLM because holographic technology is more mature. Current developments in SLM technology have shown that a 128x128 array is achievable. Undoubtedly, in the near future, a 256x256 array will be fabricated.

The parallel accessing time of the 300 bits of user data is defined by the frame rate of the SLM, which is typically in the μsec region for the fastest ferroelectric liquid crystal device. However, due to the powerful parallel accessing process, we attain a throughput rate of 300 Mbits/sec in the final demonstration.

Appendix B contains the software which was developed to configure and control the SLM, the liquid crystal light value, the electro-optic polarization rotator, and the CCD line scanner. The main purpose of this software is to link the CPU of a microcomputer with all the dynamic hardware that exists in the memory system.

4.2 Conclusions and Recommendations for Future Research

In conclusion, in this program we have shown the feasibility of an optical cache memory system that is capable of operating in conjunction with a central processing unit (CPU). The major features that are built into this system are a static multichannel parallel accessing scheme (through two pairs of chromatically compensated holographic gratings) and the novel usage of a dynamic memory storage material.

The whole system architecture design is based on the well-known vector-matrix inner product method and modified to accommodate the mode operation of the memory material. The dynamic memory material utilized in the cache system is a new class of Azo dye doped PVA polymer film. The bit memory storage in the polymer film is based on the localized photoinduced birefringent effect. Two operation modes of this polymer were evaluated to characterize their performance in
cache memory storage. It was determined that the birefringent mode will yield much higher storage efficiency and tolerate less stable recording conditions. The unique properties of this material for optical memory storage are that it allows information to be stored at extremely low optical power, no significant memory degradation occurs after the recording process, and the information can be erased or reconfigured locally whenever it is desired. All of these properties exist at room temperature and in an open environment. Table 4-1 presents the results of POC's experimental evaluation of this material. In addition to the Azo-dye-PVA material, we have studied the possibility of employing memory materials that make use of two photon absorption or the spatial hole burning effect. Because both of these materials have a distinct quantum mechanical nature, high optical power or low temperature is required to manipulate them. In a practical system, the ideal optical memory to be developed will operate in an environment that is at least compatible with current electronic memory. For this reason, it is not feasible in the near term to see any direct application of those materials, although they have been recognized for their great operating speed and high data storage capacity.

Table 4-1  Quantified Memory Storage Properties of Azo-PVA Polymer in Birefringent Mode Operation.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit Memory Writing Speed (to &gt; 5% efficiency)</td>
<td>~ 0.1 sec. at Recording Power of &gt;0 mW and Bit Spot 1 mm²</td>
</tr>
<tr>
<td>Local Bit Memory Reversibility</td>
<td>&gt; 0.5 Million Cycles Without Fatigue</td>
</tr>
<tr>
<td>Modulation Bandwidth</td>
<td>0--3.1 KHz (frequency response roll off at 2.8 KHz)</td>
</tr>
<tr>
<td>Maximum Reduced Birefringent (Δn)</td>
<td>Δn_max ~ 0.01</td>
</tr>
<tr>
<td>Memory Reconfiguration Time (to 9% of original efficiency)</td>
<td>T_erase + T_rewrite ~ 2 sec.</td>
</tr>
<tr>
<td>Spatial Resolution</td>
<td>3000 lines/mm</td>
</tr>
<tr>
<td>Operating Temperature</td>
<td>0°C -- 40°C</td>
</tr>
<tr>
<td>Memory Persistence Time</td>
<td>&gt; 2 Weeks with Degredation less than 20%</td>
</tr>
</tbody>
</table>

In regard to the Azo-dye doped polymer, we made a fundamental study of its microscopic origin. The results revealed by this study show that several improvements, in both the guest and the host material, can be made to further enhance the performance of this material and push it to its theoretical limit. According to the rotating dipole model, the ultimate time response in reorienting a
molecular dipole moment and the strength of induced birefringence are determined by the speed of
dipole rotation and the strength of local field interaction. Therefore, future material research (and
improvement) shall be aimed at several goals:

1. Choose host polymer systems with matching low electric constant; the lower the
   polymer's index, the less restriction of polymer matrix to dipole rotation.
2. Search for low molecular weight Azo dye compound; the lower the dye molecular
   weight, the faster the dipole rotation.
3. Perform a heat transfer study of the guest-host compound and develop a standard
   thermal preprocessing technique.
4. Modify the current system into a copolymer system.

Although a breadboard model of the optical cache memory was successfully demonstrated in this
program, the optical cache memory concept merits further intense study for transformation from
optical bench into engineering prototype. We wish to recommend that the future direction of this
study concentrate on three areas. The first is storage capacity enhancement. This requires
advancement in 256x256 SLM development. The second is producing polymeric material that
ultimately can be driven at very low optical intensity (<5mW/cm²) with high sensitivity (at 80%
birefringent efficiency), fast rise time (< 0.1 sec), low fatigue, and with long memory persistence.
More importantly, in order to reduce system dimensions, the study should focus on tailoring the
material's spectral response toward the near IR regime where compact diode lasers can be utilized
as the driving source instead of the large frame Argon laser.

The third area that requires further investigation is detailed system architecture design. The general
weakness of current architecture originates from the insufficient understanding of the memory
material's special limitations and characteristics. When enough knowledge is gained by studying
the polymeric material, a more appropriate architecture will result. This architecture will more
efficiently route the data input/output and control reconfiguration. For instance, Figure 4-5 shows
the modified architecture design of a compact optical cache memory system that could fulfill this
need. In summary, we are confident that major advancements can be made in all three areas, thus
making optical cache memory a practical reality.
Figure 4-5 Illustration of a Modified Reconfigurable Real-Time Cache Memory System in Compact Packaging
5.0 REFERENCES


ROME LABORATORY

MISSION
OF

ROME LABORATORY

Rome Laboratory plans and executes an interdisciplinary program in research, development, test, and technology transition in support of Air Force Command, Control, Communications and Intelligence (C³I) activities for all Air Force platforms. It also executes selected acquisition programs in several areas of expertise. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C³I systems. In addition, Rome Laboratory's technology supports other AFSC Product Divisions, the Air Force user community, and other DOD and non-DOD agencies. Rome Laboratory maintains technical competence and research programs in areas including, but not limited to, communications, command and control, battle management, intelligence information processing, computational sciences and software producibility, wide area surveillance/sensors, signal processing, solid state sciences, photonics, electromagnetic technology, superconductivity, and electronic reliability/maintainability and testability.