An Optical Associative Parallel Processor for High-Speed Database Processing

Theoretical Concepts and Experimental Results

Ahmed Louri and James A. Hatch Jr., University of Arizona

This architecture exploits optics to perform wordparallel and bitparallel relative magnitude searches. As the present system evolves, it will substantially exceed current database processing speeds.

November 1994

ith its many communications advantages, optics continues to receive increasing attention as a way to provide the storage, speed, and massive interconnections needed in future computing systems. We have devised a novel architecture that exploits the advantages of optics for performing word-parallel and bit-parallel equivalence and relative magnitude searches of database tables in constant time. Moreover, our experimental optoelectronic implementation of the associative processing portion of this architecture has achieved encouraging preliminary results.

Associative processing is advantageous for performing symbolic computing tasks for several reasons, as explained in the Guest Editor's Introduction to this issue. However, there have been many obstacles to commercially successful associative processors. Some of these obstacles include the higher cost and poorer storage density of associative memory compared with conventional memory, the lack of efficient broadcasting and funneling, and the lack of parallel access to data.

A potential solution to many of these shortcomings is to integrate such alternate technologies as optics with conventional electronics. First, the use of free-space and fiber-based optical interconnects can alleviate the wiring complexity of associative processing systems¹ by migrating the implementation of wiring into the third dimension, that is, free space. This decreases the chip area used for routing signals between chips and boards and increases the area available for gates. Moreover, the large bandwidth of optics will provide higher interconnection densities with lower power dissipation.¹ The ease with which optical signals can be expanded (which allows for signal broadcasting) and combined (which allows for signal funneling) can also be exploited to solve the interconnect design problem and alleviate network latency problems. Furthermore, since photons do not readily interact with each other, optical signals are less prone to crosstalk, thus potentially allowing higher interconnect densities than with electronic signals.

0018-9162/94/ \$4.00 © 1994 IEEE

Acronyms used in this article

BER - Bit error rate CA — Comparand array CCD - Charge-coupled device CL - Cylindrical lens FLC --- Ferroelectric liquid crystals IA --- Intermediate array LCTV — Liquid crystal television MCU - Match/compare unit M/D --- Match/detector OA - Output array OCAPPRP --- Optical content-addressable parallel processor for relational database processing OU - Output unit POHM — Page-oriented holographic memory RA - Relation array SLM --- Spatial light modulator

comparands through the *same* database table in parallel, where *n* is the number of comparands. Regardless of the search dimensionality, relational database processing can be decomposed into two types of searches, equivalence and relative magnitude. Equivalence searches merge tables (relations) on the basis of the presence of identical entries in either of two relations. Relative magnitude searches retrieve data from a single relation, as in, say, the search for all articles published after 1990.

search increases the system throughput by a factor of n by searching multiple

In a high-speed database system, it is important to perform both of these operations in a bit-parallel manner. By bitparallel we mean that an operation's execution time is independent of the number of bits per word (the word size). To do this, all bits must factor directly into the output. As the word size increases, the wiring complexity of the operation can restrict its electronic implementation.

Optical equivalence searches are easily implemented as bit-parallel operations. Since a mismatch in *any* bit position of two words indicates their inequality, a simple funneling (beam-merging) operation determines the result. However, optical relative magnitude searches are not as simple. For two words that are not equal, the relative magnitude is not immediately



Our experimental architecture, called the optical content-addressable parallel processor for relational database processing (OCAPPRP), supports parallel relational database processing by fully exploiting the parallelism of optics. First we address the searching capacity of the architecture and the problem of execution-time differences for equivalence and relative magnitude $(<, >, \leq, \geq)$ searches.

The OCAPPRP compares a search string (comparand) with each entry of a database table in parallel. The search of a single comparand through an entire table is referred to as a one-dimensional search in this article. A two-dimensional



Figure 1. Structural organization of the OCAPPRP. The number 5 in the comparand array (lower left) is simultaneously compared with each of the four rows of the RA. The E, G, and L registers store the search results. As an example, we use the search for all RA entries that are greater in value than the number 5. As reported in the G register, rows RA₁ and RA₄ satisfy the search and are transferred to the output array by the output unit.

COMPUTER

known because only the first bit position to result in a mismatch, beginning with the most significant bit, is relevant.

The problem of isolating this bit position has limited relative magnitude searches to bit-serial implementations, requiring up to m iterations, where m is the word size. We exploit the interconnection capabilities of optics in developing a new word-parallel and bit-parallel technique for isolating this bit position, which allows constant-time operation, that is, O(1)operations; hence, our approach is a single-step algorithm.

OCAPPRP description. Figure 1 shows a preliminary organizational structure for the OCAPPRP. The architecture consists of a selection unit, a match/compare unit (MCU), an equality unit, a threshold unit, an output unit (OU), and a control unit (not shown). The selection unit is intended to enable word and bit slices of a search string (comparand) called the comparand array. The CA and other optical inputs can be supplied by either optical disks or page-oriented holographic memory (POHM). In a POHM, many pages of data (approximately 1,000×1,000 bits/page for an area of 1 square millimeter)² are stored as multiple subholograms on a single substrate. They offer storage densities of more than a terabyte, with transfer rates exceeding 100 Gbytes/s.

The MCU searches the comparands through a data array known as the relation array, which stores the database table (relation) being searched. The RA consists of k tuples of word size m. The search of a comparand through the RA begins with the bit-by-bit search for mismatches within input word-pairs (see Figure 2).

At this point, the search is not complete. The intermediate results, represented by the right-hand side of Equation 1 in Figure 2 and called the intermediate array, merely indicate the match/mismatch of the corresponding words on a bit-by-bit level. A zero in the IA indicates the equality of the corresponding CA and RA bits, while a one indicates their inequality. To determine the equality/inequality and relative magnitude on the word level, we need to further process the IA in the equality and threshold units, respectively.

The equality unit determines matches among input word-pairs by scanning the 1A for mismatches. This is accomplished by ORing the bits along the IA rows, since bit-by-bit mismatches are represented by ones. Thus, a single mismatch

November 1994

The search of a comparand through the RA is accomplished by bitwise XORing the CA with each RA entry, where a comparand $CA = ca_{(m-1)} \dots ca_1 ca_0$ and

$$[ca_{(m-1)} \dots ca_{1} ca_{0}] \oplus \begin{bmatrix} ra_{i(m-1)} \cdots ra_{11} ra_{10} \\ ra_{2(m-1)} \cdots ra_{21} ra_{20} \\ \vdots & \vdots & \vdots \\ ra_{k(m-1)} \cdots ra_{k1} ra_{k0} \end{bmatrix} = \begin{bmatrix} ra_{i(m-1)} \oplus ca_{i(m-1)} \oplus ca_{1} ra_{10} \oplus ca_{0} \\ ra_{2(m-1)} \oplus ca_{i(m-1)} \oplus ca_{1} ra_{20} \oplus ca_{0} \\ \vdots & \vdots & \vdots \\ ra_{k(m-1)} \oplus ca_{i(m-1)} \oplus ca_{i(m-1)} \cdots ra_{k1} \oplus ca_{1} ra_{k0} \oplus ca_{0} \end{bmatrix}$$

where the symbol \oplus represents the XOR operation. (1)

Figure 2. Searching a comparand through the relation array.





Figure 4. The equality unit determines the match/mismatch of input wordpairs from the comparand array and the relation array by detect-



ing at least one mismatched bit position in a given row of the intermediate array.

in any bit position indicates the mismatch of the two words. The result is then inverted, a step needed only if the "positive logic" representation is desired. These operations are demonstrated by the expression shown in Figure 3. This expression forms a $k \times 1$ column vector known as the equality register. The *E* register is represented as $E = \{e_1 \ e_2 \ ... \ e_k\}^T$, where a 1 in element *e*, represents the equality of the CA with the *i*th RA entry, *RA*_{*i*}. The equality unit is demonstrated schematically in Figure 4. We see that element $e_2 = 1$ indicates that row RA₂ matches the CA. This register is then vertically ORed to form the match/detector (M/D) bit. The condition M/D = 1 indicates that at least one entry of the RA matches the CA. The value of this bit gives a quick indication of whether there are any matches.

The threshold unit (where the term threshold is synonymous with relative magnitude) processes a second copy of the IA for the word-parallel and bit-parallel relative magnitude search of the CA and the RA in a single step. Recall that the rows of the IA indicate the bit-by-bit



Figure 5. Optical implementation of the match/compare unit. The comparand array is searched through the relation array by expanding (broadcasting) and imaging it onto the relation array, which is stored in the spatial light modulator. The two are bitwise XORed, and the result is used for bit-parallel equivalence and relative magnitude searches.

equality/inequality of the CA and the RA. To compute the relative magnitude of a row on the word level, we must isolate the first bit position of the IA, beginning with the most significant bit, to result in an inequality. The relative magnitude of this bit position determines the relative magnitude of the entire word. Since our algorithm isolates this bit position in a single step, it eliminates the need for iterating through bit slices. We provide a complete, detailed description of this operation and its optical implementation elsewhere.^{3,4}

The threshold unit creates two sets of registers called the greater-than (G) register, where $G = \{g_1 \ g_2 \ \dots \ g_k\}^T$, and the lesser-than (L) register, where $L = \{l_1 \ l_2$... l_k ^T. The condition $g_i = 1$ indicates that RA_i is greater than the CA, and $l_i = 1$ indicates that RA_i is less than the CA. In Figure 1, elements g_1 and g_4 are set, which indicates that rows RA1 and RA4 are greater than the CA. Likewise, the value of l_3 indicates that row RA₃ is less than the CA. The OU then transfers selected tuples of the RA to the optical output array (OA). It dynamically maps nonconsecutive input tuples onto consecutive rows of the OA. In Figure 1, rows RA₁ and RA4 are mapped onto rows OA1 and OA₂, respectively.

OCAPPRP implementation. Below, we describe in detail the optical implementation of the match/compare and equality units of the OCAPPRP. Because the



Figure 6. Organization of the experimental system. The video modulators create video signals that write data patterns on the liquid crystal televisions. The output of a search is collected by linear CCD arrays, which are then read by circuits that sample, threshold, and digitize the CCD output. New inputs are created on the basis of the results, and another pass through the system begins.

threshold unit is so complex, we refer readers interested in its optical implementation to our previously published work.³

Figure 5 illustrates the detailed optical implementation of the MCU. To search the CA through each entry of the RA in parallel, the CA from the selection unit is first scaled in the vertical dimension by cylindrical lenses CL1 and CL2. It is then imaged onto the spatial light modulator (SLM), which holds the RA. An SLM is a real-time reconfigurable device capable of modifying the amplitude (or intensity), phase, or polarization of an optical wavefront as a function of position across the wavefront.⁵ The SLM for the bitwise XOR operation is a liquid crystal device that rotates the polarization of the incident light by 90 degrees in the bit positions that contain logical 1's. Thus, the polarization of the incident light is rotated by 0 or 180 degrees for the 00 and 11 cases, respectively, yielding vertically polarized light, and 90 degrees for the 01 and 10

COMPUTER



Figure 7. Laboratory arrangement of the initial experimental version of the OCAPPRP. To the right of the optical table is the microcontroller-based control unit for the system. The lens system performs the optical matching. An IBM PC displays the search results.



Figure 8. (a) The spatial encoding scheme for representing binary-valued data as optical signals. (b) The truth table for using the encoding scheme to perform matching in our experimental system.

cases, yielding horizontally polarized light in the IA. Lohmann provides more information on polarization-encoded logic.6 The polarizer "disables" the equalities in the IA by blocking the vertically polarized light. Without light, these bit positions can no longer factor into computations. The resulting data plane is then duplicated by the beam splitter, with one copy going to the equality unit and the other going to the threshold unit. In the equality unit, cylindrical lens CL3 sums all the bits along each RA row to search for mismatches. A negative logic representation of the E register is formed at the focal point of CL3. An optional SLM

November 1994

may be inserted to generate the positive logic form shown in Figure 4.

An experimental OCAPPRP. First we describe the laboratory setup, including the devices and components used, then we present some of the results we obtained experimentally.

System setup. We are building, at the University of Arizona, Department of Electrical and Computer Engineering, an experimental OCAPPRP. Our initial version uses a bit-serial relative magnitude algorithm instead of the single-step relative magnitude algorithm discussed pre-

viously. Future versions will include this advanced single-step feature. Figure 6 illustrates the system's organization, which allows for everything to be controlled by a microcontroller board. Active-matrix liquid crystal television screens (LCTVs) are used as spatial light modulators, since they are cost-effective devices for demonstration purposes. Video generation ICs interfaced to the microcontroller generate the write patterns for the LCTVs. After a set of patterns is written to the LCTVs and the search is performed, linear CCD (charge-coupled device) arrays detect the output. Auxiliary circuits sample and threshold the CCD output to form a digital data stream, which is then read by the microcontroller. New data patterns are generated and the search continues.

The optical portion of the demonstration system in Figure 6 operates as follows: The digital patterns are impressed on a beam from an argon laser by LCTV1 and LCTV2. The two LCTVs store the CA and RA, respectively. The CA is copied electronically into each row of LCTV1, and the beam expansion optics (cylindrical lenses CL1 and CL2 in Figure 5) are eliminated to reduce complexity and cost. The search array is simultaneously written into LCTV2. The superposition of the two data planes performs the optical matching, whose result is split into two paths. One path is focused to a vertical line, forming the E register, which is then imaged onto a linear CCD array. Here, electronic circuits sample, threshold, and digitize the data for the microcontroller. This path represents the equality unit. To implement the bit-serial relative magnitude algorithm, the other path uses LCTV3 to disable bit slices during iterations. This forms the G register and the L register (not shown). Figure 7 shows the laboratory setup.

Experimental results. We performed a sample search with our experimental laboratory system. For this initial version, we are using a different encoding scheme because commercial LCTVs have a limited contrast ratio (the amount of light transmitted in the "on" state compared with that transmitted in the "off" state). Other liquid crystal devices, such as ferroelectric liquid crystals (FLC),⁵ exhibit the necessary contrast but are more expensive. In this scheme, both a binary value and its complement are used to encode a single bit, as in the expression $x\overline{y} + \overline{x}y$. The encoding scheme and truth table are illustrated in Figures 8a and 8b.



Figure 9. The theoretical and experimentally obtained system results.

A black square represents an opaque LCTV pixel, whereas a white square represents a transparent one. The patterns on the two LCTVs are superimposed to perform the logical operation. In the result, if both pixels corresponding to a single bit are dark, then the two inputs are logically equivalent. The presence of light in either pixel of a bit indicates a mismatch between the two inputs.

To map these patterns onto the devices, the LCTV display areas are partitioned into eight rows with 16 squares per row. Thus, we are demonstrating an 8×8 data array. Note that this is a proof-ofprinciple system. Our word size, *m*, is limited solely by the contrast ratio of the devices and not by the architecture. As mentioned above, other currently available devices have the necessary contrast ratio to support a minimum word size of 128 and even 256 bits.

Figure 9 illustrates the various data planes for the search. Figure 9a is a

graphical representation of the pattern written to LCTV1, while Figure 9b is the pattern written to LCTV2. The data plane in Figure 9c illustrates the theoretical results of the optical matching operation. We see that row 5 is completely dark because of the perfect match between the search string and the fifth array entry. A single pixel is illuminated in rows 4 and 6, indicating a single mismatch between these array entries and the search string. Moreover, the column vector in Figure 9d illustrates the theoretical output of the horizontal summing of Figure 9c, which is vertically inverted by the optical system. Again, row 5 is dark relative to the others.

To demonstrate the system's operation, we include photographs of the input patterns written to LCTV1 and LCTV2 (Figures 9e and 9f, respectively). In Figure 9g, we report the experimentally obtained results of the optical matching, which fulfill our expectations from Figure 9c. Furthermore, the photograph in Figure 9h illustrates the experimental generation of the output column vector. Overall, the results satisfy our expectations and successfully demonstrate the system's ability to perform optical parallel-string searches. The final step is to detect this optical result with a CCD array and report the results electronically to the microcontroller.

Theoretical performance analysis

The system's execution time (T_{ex}) can be expressed by

$$T_{\rm ex} = T_{\rm set} + T_{\rm proc} + T_{\rm trf} \tag{3}$$

where T_{set} is the setup time, T_{proc} is the processing time of the optical system, and T_{trf} is the time needed to transfer the result to the host computer. Since T_{set} and T_{trf} can be overlapped with the processing time (assuming heavy pipelining), they need not be considered in this preliminary analysis. The term T_{proc} can be expanded as

$$T_{\text{proc}} \approx T_{\text{page}} + T_{\text{prop}} + 2T_{\text{SLM}} + T_{\text{detect}}$$
 (4)

 T_{proc} represents the equivalence search path, T_{page} is the time needed to read a page of data from holographic memory,

COMPUTER

 $T_{\rm prop}$ is the light propagation time between components, T_{SLM} is the SLM switching time, and T_{detect} represents the speed of the detector. The value of T_{prop} is negligible (on the order of picoseconds) and $T_{\text{page}} \approx 20$ nanoseconds, while $T_{\text{detect}} \approx 30 \text{ ns and } T_{\text{SLM}} \text{ is on the order of}$ microseconds. Therefore, the dominant factor in Equation 4 becomes $2T_{SLM}$. Johnson et al.7 note that currently available optically addressable ferroelectric liquid crystal arrays have been demonstrated in sizes of 128×128 . Electrically addressable versions of FLC arrays have been demonstrated in sizes of $1,280 \times$ 1,120 pixels. Although the switching time of the liquid crystal itself is approximately 2 microseconds, the time needed to charge the photoconductive layer currently limits the optical addressing time to about 3 milliseconds. However, with improved integration techniques, we believe this value can eventually be reduced to tens of microseconds, judging by the incredible (gigahertz) operating speeds of current discrete phototransistors

The number of tuples that the system can match per second is found by dividing the SLM size by the execution time. We estimate that for $T_{ex} = 6 \text{ ms} (2T_{SLM})$, an OCAPPRP built with this technology would be capable of matching 2×10^4 tuples/second in a 1D system and 3×10^{6} tuples/s in a 2D system. Using a power analysis that relates the necessary optical power to the bit rate, we estimate that the optical power requirements for the 1D system operating at a BER (bit error rate) = 10^{17} is approximately 50 milliwatts for equivalence searches. Expectations are that $1,028 \times 1,120$ optically addressed arrays will be operating at $T_{\rm SLM} = 30 \ \mu {\rm s.}^7$ For future 1D systems with $T_{ex} = 60 \,\mu s$, we can expect to match 1×10^7 tuples/s with an optical power requirement of approximately 1W. Using another technology, such as self-electrooptic-effect devices,5 which have write times in the nanosecond range, we should be able to match approximately 1 $\times 10^{10}$ tuples/s. Since data transfer rates greater than 100 Gbytes/s are currently possible with page-oriented holographic memory,² an OCAPPRP will have no problem supporting these predicted processing rates. Such performance is possible because data is retrieved from memory in parallel as pages and is then processed as pages.

Note that the above execution-time analysis excludes the speedup realized

November 1994



Figure 10. The 2D match/compare unit implements Equation 1 (Figure 2) for each entry of the comparand array. The results of these parallel operations appear in the intermediate array, which will later be processed by the equality and threshold units.

when selection is based on relative magnitude searches. The use of our new single-step algorithm, even for current word sizes of 64 bits, will provide even greater performance increases. Others are performing similar research in this area,² and our predicted performance is very close to their most recent estimates of 1×10^9 tuples/s.

Future work

Work on the OCAPPRP is ongoing. One of the most substantial improvements is the extension of the MCU to 2D. Although we have already completed this extension, we omitted it from this article so that the architecture described would match that of the experimental system. Nevertheless, we can briefly explain the main concepts.

We generalize the search of Equation 1 to 2D by representing the comparand array as CA_h , where h = 1, 2, ..., n and $CA_{h} = ca_{h(m-1)} \dots ca_{h1} ca_{h0}$. Fundamentally, 2D searches can be viewed as n 1D searches in parallel. To perform 2D searches, the MCU must be scaled from a 1D to a 2D implementation. The operation of a 2D MCU is described schematically in Figure 10, where the n intermediate results form the IA. Corresponding to each of the intermediate results, we will also have n E, G, and L registers (not shown), one for each comparand. Similar to the 2D notation for the CA, these column vectors are denoted as $E_h = \{e_{1h}, e_{2h}, \dots \}$ e_{kh} $T, G_h = \{g_{1h}, g_{2h}, \dots, g_{kh}\}^T$, and $L_h = \{l_{1h}, l_{2h}, \dots, q_{kh}\}^T$ l_{kk} ^T. More details regarding the operation and optical implementation of a 2D MCU are available in the literature.⁴

We have targeted some additional

work for the experimental system. The first task is to replace the electronically connected paths with optics to obtain purely optical feedback. Other tasks will include the experimental demonstration of the 2D MCU and the single-step relative magnitude search algorithm. In short, future generations of the system should more closely meet our goal of producing a hybrid optoelectronic machine that exploits the domains of both electronics and optics.

s we have shown, the proposed architecture has the potential to process approximately 1×10^{10} tuples/s. Nevertheless, with available optical switching devices (SLMs) the proposed system clearly cannot compete with its electronic counterparts because the switching time of current optical device technology is slow. However, many improvements are being made in optical and optoelectronic device technologies. With advances in SLM technology, we can expect the proposed optical system to become a viable and cost-effective alternative for parallel and high-speed database processing.

Acknowledgment

This research was supported by National Science Foundation Grant No. MIP 9113688.

References

 F.B. McCormick, *Photonics in Switching*, Vol. 2, Ch. 4, Academic Press, Boston, 1993.

- P.B. Berra et al., "Optical Database/ Knowledge Base Machines," *Applied Optics*, Vol. 29, No. 2, Jan. 1990, pp. 195-205.
- A. Louri and J.A. Hatch Jr., "Optical Implementation of a Single-Iteration Thresholding Algorithm with Applications to Parallel Database/Knowledge Base Processing," *Optics Letters*, Vol. 18, 1993, pp. 992-994.
- A. Louri and J.A. Hatch Jr., "An Optical Content-Addressable Parallel Processor for High-Speed Database Processing," to be published in *Applied Optics*, 1994.
- J.A. Neff, R.A. Athale, and S.H. Lee, "Two-Dimensional Spatial Light Modulators: A Tutorial," *Proc. IEEE*, Vol. 78, May 1990, pp. 826-854.
- A.W. Lohmann, "Polarization and Optical Logic," *Applied Optics*, Vol. 25, Mar. 1986, pp. 1,594-1,597.
- K.M. Johnson, D.J. McKnight, and I. Underwood, "Smart Spatial Light Modulator Using Liquid Crystals on Silicon," *IEEE J. Quantum Electronics*, Vol. 29, No. 2, 1993, pp. 699-714.



Ahmed Louri is an associate professor in the Department of Electrical and Computer Engineering at the University of Arizona, and he is director of the Optical Computing and Parallel Processing Laboratory there. His research interests include computer architecture, parallel processing, optical computing, and optical interconnects. In 1991, he received the "Best Article of 1991 Award" from *IEEE Micro*. His research has been supported by the National Science Foundation (NSF) as well as industry. In 1988 he was the recipient of the NSF Research Initiation Award.

Louri earned the Diplome D'engenieur (engineer degree) from the University of Science and Technology, Oran, Algeria, in electrical engineering, and a master's degree and a PhD degree in computer engineering from the University of Southern California. He is a member of IEEE, ACM, the Optical Society of America, and the Society of Photooptical Instrumentation Engineers.



James A. Hatch Jr. is with Trimble Navigation in Sunnyvale, California, as a research engineer. From 1992 to 1994, he was a research assistant in the Optical Computing and Parallel Processing Laboratory at the University of Arizona, Tucson. His research interests include optical computing, optical interconnects, and applications of spatial light modulators.

Hatch received a BS (cum laude) in electrical engineering from Rensselaer Polytechnic Institute, Troy, New York, in 1991 and an MS in electrical engineering with a minor in optics from the University of Arizona in 1994. He is a member of the IEEE Computer Society, the Lasers and Electro-Optics Society, and the Optical Society of America.

Readers can contact Louri at the University of Arizona, Dept. of Electrical and Computer Engineering, Tucson, AZ 85721. His e-mail address is louri@ece.arizona.edu.

