Dense 3D Integrated Electronic/Photonic Computing Structures
Enabled by Diffractive Optical Elements

ARPA/AFOSR Grant No. F49620-94-1-0045

Combined Technical Report

Reporting Periods:
8/1/95 to 8/31/96 (Annual Progress Report)
11/1/93 to 8/31/97 (Final Technical Report)

Armand R. Tanguay, Jr., Principal Investigator
B. Keith Jenkins, Co-Principal Investigator
Alexander A. Sawchuk, Senior Investigator

Departments of Electrical Engineering, Materials Science,
and Biomedical Engineering,
Center for Photonic Technology,
Center for Neural Engineering,
Signal and Image Processing Institute,
and Integrated Media Systems Center
University of Southern California
Los Angeles, California 90089-0483

19980421 018
Substantial progress is reported on the development of a novel photonic multichip module technology, and on the passive optical components that provide parallel chip-to-chip interconnections within the module. This progress includes extensive characterization of a compact optical power bus to distribute an optical array of readout beams to a set of modulators; novel design-algorithm development, successful fabrication, and characterization of diffractive optical elements (DOE’s) for use in ultra-compact, short-propagation-length interconnection systems; the analysis of photonic multichip module design and performance parameters; and the preliminary investigation of applications for such multichip modules. Directly related work on flip-chip bonding between silicon (detection and signal processing) chips and GaAs-based (modulator-array and VCSEL array) chips, and on the design and test of FET-SEED spatial light modulator array chips has also been accomplished. Different operational wavelengths and module design variations potentially allow either of these smart-pixel spatial light modulator approaches to be used for the optical input/output functions within the photonic multichip module.
Dense 3D Integrated Electronic/Photonic Computing Structures
Enabled by Diffractive Optical Elements

DARPA/AFOSR Grant No. F49620-94-1-0045

Combined Technical Report

Reporting Periods:
8/1/95 to 8/31/96 (Annual Progress Report)
and
11/1/93 to 8/31/97 (Final Technical Report)

Armand R. Tanguay, Jr., Principal Investigator
B. Keith Jenkins, Co-Principal Investigator
Alexander A. Sawchuk, Senior Investigator

Departments of Electrical Engineering, Materials Science,
and Biomedical Engineering,
Center for Photonic Technology,
Center for Neural Engineering,
Signal and Image Processing Institute,
and Integrated Media Systems Center
University of Southern California
Los Angeles, California 90089-0483
(213)-740-4403; (213)-740-9823 (FAX)
atanguay@mizar.usc.edu
# Table of Contents

1. Executive Summary of the Research Effort ......................................... 3
2. Research Objectives ............................................................................ 5
3. Status of the Effort ............................................................................. 7
4. Research Accomplishments ................................................................ 8
   4.1 Introduction and system overview .................................................. 8
   4.2 Smart pixel SLM ........................................................................... 10
      4.2.1 Flip-chip bonding .................................................................. 10
      4.2.2 CMOS-SEED test and demonstration ....................................... 13
   4.3 Optical power bus ........................................................................ 16
   4.4 Diffractive optical element array ................................................... 17
      4.4.1 DOE design .......................................................................... 17
      4.4.2 DOE fabrication .................................................................... 19
      4.4.3 DOE characterization ............................................................... 21
   4.5 Photonic multichip module (MCM) analysis and applications ........ 23
      4.5.1 Interconnection system analysis .............................................. 24
      4.5.2 Multichip module system analysis .......................................... 25
      4.5.3 Applications ......................................................................... 26
      4.5.4 Integrated submodule demonstration project ............................. 27
   4.6 References .................................................................................... 28
5. Personnel Supported ............................................................................. 29
6. List of Publications and Presentations Resulting from this Grant ........ 31
7. Interactions/Transitions ..................................................................... 35
   7.1. Participation/presentations at meetings, conferences, and seminars .. 35
   7.2. Consultative and advisory functions ............................................. 39
   7.3. Transitions ............................................................................... 39
8. Inventions and Patent Disclosures ....................................................... 40
1. Executive Summary of the Research Effort

A wide variety of potential applications exist for extremely compact, ultra-high-capacity computational modules in both embedded and high performance computer environments. These applications include, among others, sensor signal processing, sensor fusion, image processing, feature identification, pattern recognition, and early vision. All of these applications are heavily computation-intensive, and in many cases must be accomplished in operational environments that are severely restricted both in available power and in the space allowed for computational elements.

In order to handle these and other "grand challenge" problems, advanced computational systems must employ distributed parallel processing elements in an architecture that can integrate multiple chips in a compact package, operate at low power, support high-bandwidth communication among processing elements and memories, and support high-bandwidth parallel input/output (I/O). Toward these ends, we have investigated densely-interconnected multilayer hybrid electronic/photonic modules that provide parallel, compact optical interconnections between electronic processor-and-memory chips in a stacked, multilayer structure. In this structure, each active layer is composed of a silicon electronic multiprocessor chip (with integrated detectors) that is flip-chip bonded to a compound semiconductor chip configured with an array of multiple quantum well modulators or vertical-cavity surface-emitting laser diodes. Interconnections between active layers are provided by both planar and volume (computer-generated) diffractive optical elements that are proximity-coupled to the active layers in order to form rugged 3-D computational blocks.

The key advantages of such an advanced packaging architecture and its associated packaging technology include the capacity for parallel transmission of intermediate computational results, and the availability of dense local and global interconnections with potential for a high degree of fan-out and fan-in. Electronic processors with both optical I/O and electronic I/O are highly suited for these tasks, with optical I/O employed for dense parallel chip-to-chip interconnections in the
vertical dimension, and electronic I/O used for lateral control signal and local cache memory access as appropriate.

As the hybrid processor/detector/modulator chips have been the subject of an intensive research and development effort under separate DARPA sponsorship, we have focused herein on the implementation of the dense vertical interconnections based on the use of diffractive optical elements in multilayer computational modules, on system integration and packaging issues, and on the design and analysis of the overall compact photonic multichip module. Computational architectures and models that stand to benefit from such a module include ultra-compact hypercube cellular-array processors, exchange/bypass interconnection networks, parallel-addressed optical/photonic cache memory, multiprocessor systems with optically-addressed shared memory, and multilayer artificial neural network systems for processing, recognition, and understanding of sensory (e.g., two-dimensional image, radar, SAR, and focal plane array) data.
2. Research Objectives

The research program described herein encompasses innovative 3-D integrated photonic computing architectures and their applications; novel passive components to make such architectures compact with high computational capacity (including complex-diffractive/refractive optical elements for optical power distribution and fan-out/fan-in interconnections, implemented as both planar and volume elements); and design and fabrication techniques that increase the capability and manufacturability of the resultant optical elements.

The research program objectives were to:

1. Investigate multilayer integrated photonic computing structures, including compatibility of components, physical size, and computational capability.

2. Develop, fabricate, and analyze novel passive optical components based on combinations of refractive (by means of indiffusion and ion-exchange) and diffractive (by means of patterned phase and amplitude modulation) effects, providing advanced diffractive optical elements for multiplexed vertical interconnections.

3. Study the potential application of optical interconnections based on advanced diffractive optical elements to compact 3-D integrated structures, providing beam fan-out and fan-in, routing, and optional weighting functions.

4. Develop, fabricate, and analyze compact optical power distribution elements, based on waveguide optical input, 2-D optical output into the vertical (out-of-plane) dimension, and optical pass-through functions.

5. Investigate computer-generated volume diffractive optical elements based on multiple layers in stratified volume holographic optical element (SVHOE) structures, with applications to global interconnection, beam shaping, and power distribution.

6. Develop computer-aided design, characterization, and manufacturing techniques for production of diffractive optical elements, including an assessment of the
applicability of the currently available MOSIS process to prototype DOE fabrication.

7. Investigate the computational characteristics of integrated 3-D photonic computing structures based on electronic/photonic interconnections. This investigation includes the design and analysis of different optical interconnection techniques, as well as the study of parallel computational architectures that maximally utilize the integrated photonic hardware.

8. Perform initial design work and feasibility assessments leading toward a demonstration system that vertically connects two silicon electronic chips, with compact parallel photonic interconnections that implement local weighted signal fan-out and fan-in.

Objective 8 is new (relative to the original proposal) and was added because the success of our investigations indicated that a near-term demonstration is more feasible than originally anticipated.
3. Status of the Effort

Progress on the research grant "Dense 3D Integrated Electronic/Photonic Computing Structures Enabled by Diffractive Optical Elements," ARPA/AFOSR Grant No. F49620-94-1-0045, A. R. Tanguay, Jr., P. I., B. K. Jenkins, Co-P. I., and A. A. Sawchuk, Senior Investigator, is described herein for the overlapping reporting periods 8/1/95 to 8/31/96 (Annual Progress Report) and 11/1/93 to 8/31/97 (Final Technical Report).

Substantial progress is reported on the development of a novel photonic multichip module technology, on the active CMOS-based electronic and modulator elements that enable both electronic functionality and optical I/O, and on the passive optical components that provide parallel chip-to-chip interconnections within the module. This progress includes extensive characterization of a compact optical power bus to distribute an optical array of readout beams to a set of modulators; novel design-algorithm development, successful fabrication, and characterization of diffractive optical elements (DOE's) for use in ultra-compact, short-propagation-length interconnection systems; analysis of photonic multichip module design and performance parameters; and the preliminary investigation of applications for such multichip modules. Directly related work on flip-chip bonding between silicon (detection and signal processing) chips and GaAs-based (modulator-array and vertical cavity surface-emitting cavity laser array) chips, and on the design and test of CMOS-SEED cellular-logic-processing chips containing spatial light modulator arrays has also been accomplished. Different operational wavelengths and module design variations potentially allow either of these smart-pixel spatial light modulator approaches to be used for the optical input/output functions within the photonic multichip module.
4. Research Accomplishments

4.1 Introduction and system overview

While extensive research has been performed on the capabilities of optical and photonic parallel computing systems (both at USC and other institutions), the development of an effective packaging technology (comprising parallel interconnections, interface compatibility, component mounting, and thermal management) has been largely neglected and is vital to the future realization of practical optical/photonic computing systems [1]. Our approach integrates silicon electronic chips for localized (analog or digital) processing, optical detector arrays, GaAs-based modulator arrays or vertical cavity surface emitting laser arrays (for electrical-signal-to-optical-signal conversion), waveguide-based optical power buses, two-dimensional lenslet arrays, and integrated diffractive optical elements (DOE’s). The three-dimensional integration of these electronic and photonic elements (as shown schematically in Fig. 1) allows for the implementation of parallel 3-D optical interconnection, parallel chip input/output, and optical power distribution functions, in order to achieve powerful yet compact computational modules.

A high degree of functionality can be achieved by integrating the processing, conversion, and interconnection functions into a multiple layer locally-interconnected structure, with each layer incorporating both passive (linear) and active (digital or analog nonlinear) elements. The overall structure comprises two basic submodules: (1) a silicon detector array/GaAs modulator array/optical power bus/diffractive optical element multilayer structure, as shown schematically in Fig. 2(a) (or a silicon detector array/GaAs VCSEL array/diffractive optical element multilayer structure, as shown schematically in Fig. 2(b)), and (2) an additional passive volume holographic optical element, as shown schematically in Fig. 3. Input signals are detected on the silicon chip of the input submodule by an array of parallel detectors, and are processed locally by associated electronic processing elements before being passed on to the GaAs modulator or VCSEL array. Readout of the modulator array is provided by the optical power bus; in the cases of both the modulator array and the VCSEL array, the interconnection weights and local fan-out
Fig. 1. Conceptual diagram of multilayer hybrid electronic/photonic computation/interconnection element, showing multiple silicon electronic chips, each divided into pixels or processing elements, and fanout/fanin interconnections between them.

Fig. 2. Schematic diagram (cross-sectional view of two pixels) of multilayer hybrid electronic/photonic computation/interconnection elements, showing (a) silicon VLSI chips, GaAs modulator chip, optical power bus and diffractive optical element; and (b) silicon VLSI chips, GaAs vertical cavity surface emitting laser (VCSEL) array chip, and diffractive optical element.

Fig. 3. Densely interconnected 3-D electronic/photonic computational module, showing both local (DOE) and global (VHOE) optical interconnections.
functions are provided by a planar diffractive optical element and microlens array. Global interconnections among several such submodules are provided by volume holographic optical elements as required by each processor architecture.

4.2 Smart pixel SLM’s

This section describes program accomplishments on the two active layers of the multilayer computational module. These two layers essentially comprise a smart-pixel spatial light modulator (SLM) [2-5]. While work on these SLM structures has been funded primarily by other contracts and grants, it forms an integral part of the photonic multilayer computational module. Below we summarize work on flip-chip bonding of a silicon chip to a GaAs (modulator or VCSEL array) chip for a hybrid SLM (usable in multichip modules), and work on a monolithic approach (usable in submodules consisting of two electronic chips) that uses CMOS-SEED arrays fabricated by AT&T.

4.2.1 Flip-chip bonding of silicon driver chips and GaAs-based modulator and VCSEL arrays

- During the grant period, we purchased and installed a two-hearth electron beam vacuum deposition system for indium bump evaporation, which has been configured with large capacity crucibles that can easily accommodate the large layer thicknesses (10 to 20 μm) required for uniform array bump bonding. This system has also been configured with a thermal evaporation source, which yields a higher indium bump evaporation rate and thereby results in better uniformity control. A thick-photoresist liftoff process has been developed (based on earlier work at Rockwell International) that yields vertical sidewalls and excellent bump definition with bump diameters in the range of 10 to 50 μm. A Research Devices M-8 Aligner/Bonder was installed and calibrated for alignment of matched indium bump patterns over large substrate areas. Bump contact verification is accomplished by means of a Research Devices infrared microscope as well as contact resistance and continuity testing.
We have successfully demonstrated photolithographic definition and deposition of 50 × 50 μm, 25 × 25 μm, 15 × 15 μm, and 8 × 8 μm In bump contacts onto electron-beam-deposited Au electrodes on InGaAs/GaAs chips, as well as onto electron-beam-deposited Al electrodes on Si chips. We have fabricated up to 25 μm high indium bumps using a set of processing parameters that induce an intentional surface roughness to allow for greater contact strength between the two substrates. We have extended this process to a number of silicon fan-out test chips (8 × 8 arrays) with Al pads for In bump deposition, as well as larger pads for external wirebonding. The silicon fan-out chips were subsequently flip-chip aligned and bonded to pixellated InGaAs/AlGaAs chips with asymmetric cavity MQW modulators. We have demonstrated that each one of the 64 modulators in the array could be controlled by an external voltage applied through the Si chip and the In bump contacts. Furthermore, we examined the reflectivity characteristics of individual modulators before and after flip-chip bonding, and found that the bonding process does not alter the modulator performance.

Under the auspices of this program, an optically-addressed artificial neural network (ANN) chip has been designed and fabricated using USC's MOSIS foundry service that is specifically designed for flip-chip bonding to GaAs-based modulator or VCSEL arrays. The first generation device comprises a 16 × 16 array of dual-input, dual-output pixels on 100 μm × 100 μm centers, fabricated within the MOSIS 1.2 μm HP scalable n-well process. The chip has demonstrated a 2 to 9 volt sigmoidal large-signal response in excess of 100 KHz, and a small-signal response in excess of 1 MHz (limited by the highest bandwidth tested in both cases). Spice simulations indicate a small-signal response in excess of 4 MHz. The estimated chip power dissipation is 2 mW per pixel, or about 0.5 W per chip using these design rules. Currently, a second generation design has emerged from fabrication and packaging, with delivery in February, 1998. This redesigned chip contains a 32 × 32 array, again using the MOSIS 1.2 μm HP scalable n-well process, and has been specifically designed to address issues regarding scalability, especially with regard to power dissipation and signal routing. This advanced design should help determine the ultimate scaling boundaries available for use within this VLSI linewidth technology.
In parallel with the effort to fabricate and evaluate flip-chip bonded hybrid silicon-GaAs SLMs based on modulator arrays, we have continued to investigate the potential for hybrid integration of low-power vertical-cavity surface-emitting-laser arrays. Such VCSEL arrays offer the potential advantage of simplicity of integration, as no optical power bus is required for optical readout. However, the state-of-the-art to date has produced VCSEL arrays that are too power consumptive to allow for scalable integration in this type of compact multichip module architecture. In order to further understand the potential advantages and disadvantages of VCSEL arrays in the context of photonic multichip module integration, we proposed to the Joint (USA-Japan) Optoelectronics Program for several sets of VCSEL arrays that were amenable to flip-chip bonding with matching silicon driver chips. To date, we have received VCSEL arrays from both NEC (Japan) and MODE (USA) for integration and comparison testing. The NEC devices, for example, comprise $8 \times 8$ arrays (64 individually-addressable lasers) with a center wavelength of 980 nm (the design wavelength of our photonic multichip module structures), a threshold current of 2 mA per laser, an operational current of 5 mA per laser at the operational voltage of 4 V (compatible with the complementary silicon driver chips), thereby producing an output power of 2 mW per laser. Although this level of power dissipation is unacceptably high, several key VCSEL array and array-integration issues can be tested with these devices, including the array output power uniformity as a function of drive current, the individual-device wavelength uniformity across the array, the polarization properties of each laser within the array, the switching speed for small-signal and large-signal modulation following flip-chip bonding to the silicon driver chip, and the degree of mutual coherence between individual lasers within the array (which is in turn important for understanding the fan-in properties of the photonic multichip module following fan-out by diffractive optical element arrays).

**Significance to field and relationship to original goals.** We have established an extensive capability for the electronic vertical interconnection of pairs of chips within hybrid multichip modules by means of flip-chip bonding. Our work is based on the development of a manufacturable process technology for the dense vertical
interconnection of two-dimensional arrays of InGaAs/GaAs and InGaAs/AlGaAs compound semiconductor Fabry-Perot modulators or VCSELs with Si photodetector/drive electronics arrays to form high-bandwidth spatial light modulators. This research makes possible the hybridization of a wide spectrum of possible electronic/photonic devices, including, for example, Si-addressed vertical cavity surface-emitting laser (VCSEL) arrays. This accomplishment represents an essential step forward toward the development of high computational capacity multichip modules, which is the principal goal of the research program.

4.2.2 CMOS-SEED test and demonstration

- We have designed, fabricated, and tested a SIMD-type 2-D parallel pipeline processor. This processor is based on a CMOS-SEED chip that contains a processing element (PE) array in which each PE has an independent free-space optical I/O data port operating at on-chip clock rates. These optical ports allow the 2-D parallel transport of data between processor arrays, memory, and input/output devices, such as video sensors and displays. The 2-D parallel data channels are capable of eliminating an I/O bottleneck that produces severe latencies in current SIMD machines.

- The CMOS-SEED chip was designed and fabricated through the AT&T (now Lucent Technologies) CMOS/MQW foundry run, arranged through the DARPA sponsored CO-OP program at George Mason University. This foundry run provided each participant with five copies of a $2 \times 2 \text{ mm}^2$ CMOS chip containing an array of 200 flip-chip-bonded MQW diodes formed in a $20 \times 10$ array. The chip designer can use all or some subset of these diodes for free-space optical I/O. With appropriate circuitry and contacts, the diodes can function as a detector, a light-emitting diode, or an optical modulator at 850 nm. The CMOS circuitry was fabricated through the MOSIS foundry service by Hewlett-Packard using their 0.8 µm CMOS process. The HP process provides three metal layers. The third metal layer is used to make 20 micron-square flip-chip bond pads for contacting the MQW diodes, as well as for circuitry interconnections. The diodes have optical windows of 18 microns square and are placed on a $62.5 \mu \text{m} \times 125 \mu \text{m}$ pitch, covering an area of $1.25 \times 1.25 \text{ mm}^2$. 
• We used the core $1.25 \times 1.25 \text{ mm}^2$ of the CMOS-SEED chip to implement a $5 \times 10$ mesh-connected array of processing elements, each of size $250 \times 125 \text{ \mu m}^2$. Accordingly, each CMOS-SEED chip has the capability to process 50 data elements (channels) in parallel. On the perimeter of the $5 \times 10$ smart pixel array are memory elements and buffers for driving global instruction and clock lines. The remaining outer ring of circuitry is dedicated to the 40 wire-bond pads, and uses 1.75 mm$^2$ or 43.5% of the chip's total area.

• We designed the CMOS-SEED processing core to consist of a $5 \times 10$ array of identically replicated PE circuits. Each PE contains 187 transistors implementing a cellular logic processing element within an area of $250 \times 125 \text{ \mu m}^2$. Each PE contains (1) circuitry for storing three bits of data, (2) logic for performing the complement (logical NOT) and union (logic OR) operations on stored data, (3) logic for performing morphological dilation operations with neighboring pixel data, (4) logic for selecting memory for storing the computed result, (5) an optical input port, and (6) an optical output port.

• We designed the optical ports to use a dual-rail representation of the optical signals. Thus, each single-bit optical port receives or transmits data on a pair of optical beams. The transmitter sets the state of two reflecting MQW diode modulators so that one is more absorbing than the other. A pair of equal intensity beams reflects from the modulator pair and enters a pair of detectors. The receivers make a decision based on the difference between two optical intensities in these detectors. Using this dual-rail representation eliminates the need for a global threshold optical intensity value. This simplifies the optical system design, because optical channels can be designed with differing amounts of optical loss (for example, each stage in a multi-stage system may have a different amount of loss) as long as pairs of beams have identical losses. This also reduces the complexity of the receiver design by eliminating the need for threshold adjustment circuitry.

• We designed CMOS-SEED transmitter-to-receiver parallel optical links that can each operate at 500 Mb/s with 600 fJ of supplied optical power. We do not plan to operate over 100 Mb/s due to the limited bandwidth of our instruction/data buffer interface. Because of the analog nature of the receiver circuit design, each
of our receivers draws over 5 mW of static electrical power and dominates the overall electrical power drawn on each CMOS-SEED chip. This power dissipation can be greatly reduced by using a clock-sense-amplifier-based smart pixel receiver (CSABSPR), since all CMOS-SEED optical links are synchronous and the clock signal is available at each receiver. Each Smart Pixel Array Cellular Logic (SPARCL) optical receiver circuit occupies an area of $32 \times 36 \mu m^2$ and each transmitter occupies $12 \times 46 \mu m^2$. These I/O circuits are each connected to flip-chip bonding pads of size $20 \times 20 \mu m^2$.

- We constructed an optomechanical package that houses up to five interconnected CMOS-SEED chips in a $2'' \times 10'' \times 14''$ volume to form the 2-D parallel pipeline processor. The package uses a slotted baseplate and mounting rings to create a rugged and stable optical system, compatible with printed circuit board packaging constraints. This optical system was constructed entirely with commercial off-the-shelf optical devices, except for a diffractive optical element and a patterned mirror, which we custom designed and fabricated.

- We have successfully cascaded CMOS-SEED chips with parallel optical links and demonstrated high performance parallel-pipeline image processing. We have performed high-speed edge detection, motion estimation, noise removal, parallel addition, parallel subtraction, and parallel multiplication on our CMOS-SEED processing system. For example, we have demonstrated image edge detection in three 30 ns clock cycles of a CMOS-SEED chip.

- We are currently using the constructed CMOS-SEED system to demonstrate real-time digital video MPEG encoding. In our demonstration, the CMOS-SEED system performs one of the most complex operations required in MPEG encoding, namely, motion vector calculation. This requires that a block-wise correlation between two adjacent frames be performed at the full frame rate. The correlation is performed using a block search method, with the minimum difference block chosen as the best-matched block. The offset of this block is encoded as a motion vector. A collection of difference blocks and motion vectors are then transmitted as a digital video stream.
Significance to field and relationship to original goals. The intent of this chip was to demonstrate that simple digital logic optoelectronic processing elements and circuits can be cascaded with optical channels to perform complex image or data processing. The intent was also to demonstrate high-speed 2-D parallel free space optical data links. This is an important first step in demonstrating the feasibility of vertically-interconnected photonic multichip modules for digital information processing applications.

4.3 Optical power bus

- We have fabricated slab waveguides in lithium niobate by titanium metal deposition followed by indiffusion in a controlled atmosphere furnace. The uppermost layer (1 μm thickness) of the titanium-indiffused lithium niobate region (Ti:LiNbO₃) has a slightly higher refractive index than that of the bulk lithium niobate crystal substrate. This leads to light confinement within the Ti:LiNbO₃ region, forming the slab waveguide. A series of first generation optical power bus modules with 660 individual rib waveguides and vertical outcoupling gratings were then fabricated by using planar microelectronic processing techniques involving photolithography and ion beam etching [6].

- We have also fabricated slab waveguides in GaAs in which the waveguide structure consists of Al₀.₃Ga₀.₇As and GaAs. An epitaxial AlGaAs layer was grown on a <100>-cut semi-insulating GaAs substrate to act as a barrier layer for the 1-micron-thick GaAs guiding layer. Successful fabrication of first generation optical power bus modules in this GaAs-based technology was also achieved, integrating arrays of rib waveguides with vertical outcoupling gratings. The steps for fabrication of the rib waveguides and the outcoupling gratings parallel the processing steps taken for the case of Ti:LiNbO₃.

Significance to field and relationship to original goals. In designing and manufacturing a feasible ultra-dense compact photonic multichip module that uses modulators for optical readout from each optoelectronic chip, a compact microoptical component is essential to be able to supply the optical power for
reflective readout of the modulator array. The optical power bus was conceptualized to fill this need for a compact "microoptical beam splitter".

4.4 Diffractive optical element array

In order to obtain the desired reconstructed intensity pattern from a diffractive optical element, the phase distribution in a diffractive optical element (DOE) array must be carefully designed. For our purposes a DOE is considered to be a pure phase-only element, although it is well known that in general, no phase-only distribution exists that can generate a given exact desired reconstructed intensity pattern. Thus, the goal of the design process is to optimize the phase distribution of the DOE so that it can provide a reasonable approximation to the desired reconstruction.

4.4.1 DOE design

- In this program we have concentrated on the development of new design methods for diffractive optical elements (DOE's). In current implementations, DOE's are pure phase devices, and therefore we derive algorithms to design DOE phase elements that generate a specified power spectrum in the Fourier plane with very high precision. We have developed a nonlinear least-squares (NLS) algorithm to design DOE's that reconstruct diffraction patterns with higher uniformity, efficiency, and signal-to-noise ratio as compared with previous design methods. The technique also uses a phase-shifting quantization procedure that greatly reduces the quantization error for DOE's. We have compared the simulated reconstruction results of DOE's designed by combining these methods with results obtained by the commonly used two-stage iterative Fourier transform (IFT) design algorithm of Wyrowski [7].

- In the specific cases examined above, the DOE's designed by use of the NLS method have lower nonuniformity (~0.5% less) and a better SNR (~2.3 dB larger) than those obtained using the two-stage IFT method, while sacrificing only a little in terms of diffraction efficiency (~0.5%). Similar results occur for many other analog or digital DOE patterns. One disadvantage of the NLS method is its
computation cost. The cost of the NLS algorithm for each iteration is \(O(M^2N^2 \log_2 MN)\) for an \(M \times N\) pattern, whereas it is \(O(MN \log_2 MN)\) per iteration for the two-stage IFT method. However, the NLS method requires fewer iterations and fewer initial guesses than the two-stage IFT method. In addition, designing DOE's is an off-line operation; thus, the computation cost is generally a minor consideration.

- We have continued to develop design programs for two-phase-level diffractive optical elements based on simulated annealing, and multiple-phase-level diffractive optical elements based on the Gerchberg-Saxton algorithm. We have further designed DOE's for cellular hypercube and uniform array fan-out interconnections intended for digital optical interconnection applications, and weighted interconnections intended for artificial neural network applications [8]. We have also developed a modified version of the Gerchberg-Saxton algorithm to achieve crosstalk reduction for a limited-fan-out system. The crosstalk analysis and reductions achieved are described below in Sect. 4.5.1, "Interconnection System Analysis".

- We have developed new design programs for multiple-phase-level diffractive optical elements based on simulated annealing. Comparisons were made with Gerchberg-Saxton design of two-phase-level and multiple-phase-level DOE's, and with simulated annealing design of two-phase-level DOE's. We have demonstrated that DOE's designed using our new approach exhibit a higher-performance combination of diffraction efficiency, accuracy, and dynamic range of neural-network interconnection weights, and signal-to-noise ratio than those designed by these more traditional algorithms. The tradeoff is an increased computation time for the design, which translates into more time and effort spent on the optimization of the design algorithm program.

- We have developed an algorithm to design and analyze diffractive microlens arrays. This algorithm maximizes the lens numerical aperture, under the constraint of a minimum feature size available, by adjusting the number of phase quantization levels as a function of the lens radius.
• We have developed programs for the analysis of fabrication errors, such as feature shrinkage, mask misalignment, and etch-depth errors. This program allows us to trace irregularities found in the DOE performance back to the fabrication process. This feedback is essential for optimizing the fabrication parameters. This program also creates DOE solutions that are more tolerant to fabrication errors.

• We have analyzed multiple reflection effects on DOE performance, and have discovered several situations in which antireflection (AR) coating of DOE devices is highly desirable. We have characterized DOE’s before and after AR coating to verify our analysis, as described further below.

4.4.2 DOE fabrication

• Designs of individual DOE elements were transferred onto chrome masks using an optical laser-write technique and an electron-beam-rastering system. The electron-beam-write technology with an e-beam spot size of 0.1 microns provided the highest resolution and best feature-definition masks.

• Binary-phase-level DOE's have been fabricated using chrome-coated photomasks to transfer the pattern onto a fused silica substrate through ion beam etching. Ion beam etching of fused silica was chosen over other alternative etching techniques such as wet chemical etching or reactive ion etching methods in the DOE fabrication due to the very accurate resulting pattern transfer of the overlaying mask. The smallest feature sizes of the DOE elements fabricated varied from 2 microns to 16 microns, and various DOE elements that perform different weighted fan-out interconnections in the direction normal to the DOE plane have been fabricated and characterized.

• To enhance the diffraction efficiencies of the DOE elements, we have investigated multilevel DOE elements. Such DOE’s provide greater flexibility in the design and implementation of optical interconnections as compared to binary-phase DOE's. Apart from the increased diffraction efficiency possible with the multilevel DOE’s, additional focusing power can be built into the multilevel
DOE array. We anticipate that this will reduce the requirements on the refractive (lens array) element to be used in conjunction with the DOE.

- The DOE CAD designs have been executed, and the mask set for an 8-level DOE has been fabricated on a chrome-coated photomask using a high resolution (0.1 micron) electron-beam-mask-writing system.

- The fabrication of the multilevel DOE's is accomplished by a combination of photolithographic patterning techniques and ion beam etching. Four-level DOE's have been fabricated at USC by using a two-step process, involving two mask levels and the etching of each separate level by means of ion beam etching. These four-level USC-fabricated DOE's are currently undergoing characterization.

- Future work involves optical characterization of the multilevel DOE's, and simulations for fabrication errors due to mask misalignment between stages, etch-depth variations, and feature shrinkage.

- In parallel, we have participated in a diffractive optics workshop sponsored by DARPA through George Mason University's CO-OP program. In this workshop we designed (and Honeywell, Inc. fabricated) DOE's for integration in our digital and neural-network CMOS-SEED processing systems. All DOE designs were contained within two separate 1 cm$^2$ substrates. The minimum feature size of this run was 1.5 microns. The digital-system DOE substrate contained a 5 x 20 uniform intensity spot array generator; a 3 x 10 uniform intensity spot array generator; and nine microlens arrays and DOE's for 2-D fan-out interconnections among the CMOS-SEED PEs. The nine microlens arrays included: a 6 x 6 array of 200 micron aperture F/2 lenses, a 2 x 2 array of 625 micron aperture F/2 lenses, a 10 x 20 array of 62.5 micron aperture F/10 lenses, a 10 x 20 array of 125 micron aperture F/7 lenses, a 10 x 20 array of 62.5 micron aperture F/7 lenses, and four 10 x 20 arrays of 62.5 micron aperture F/3 lenses using various minimum feature sizes. The fabrication process included a chrome masking layer to create apertures for DOE's that allowed easy alignment in our CMOS-SEED processing system.
• We have successfully integrated several of the Honeywell-fabricated DOE's into our digital CMOS-SEED processing system package. We have used the $5 \times 20$ uniform intensity spot array generator to transfer, in parallel, 50 bits of data from an array of modulators on one CMOS-SEED chip to an array of detectors on another CMOS-SEED chip. We are currently designing a more compact system that uses diffractive microlenses to create a micro-channel for each optical signal.

• In order to improve the diffraction efficiency of the DOE elements and to reduce reflection losses, crosstalk, and feedback effects for multilayer cascaded systems such as the proposed photonic MCM, we have investigated the use of antirefection coatings for individual DOE arrays. Preliminary completed work showed that single layer (quarter wave) magnesium fluoride coatings that were electron-beam evaporated onto a fused silica substrate in a Balzers BAK-640 thin-film deposition system reduced the reflectance from 4% for an uncoated surface to 1.6% per surface.

• Further reduction in the reflectivity of the coated surface can be achieved by using a two-layer magnesium fluoride/magnesium oxide coating. These two-layer coatings have been designed, modeled, and fabricated, with the desired result that the reflectance was reduced to 0.08%.

• In addition, a novel four layer dielectric antireflection coating has been designed and modeled, which should decrease the reflectance even further.

4.4.3 DOE characterization

• We have developed a program that uses a moving window to scan the DOE diffraction pattern, and calculates the sum of pixel values within the window. The resulting value at the appropriate location is proportional to the intensity of the corresponding diffraction order.

• Using the moving-window integration method, we have characterized several of the DOE's described above. For example, one such DOE has 16 phase elements per period, $4 \times 4$ periods, a $4 \mu \text{m} \times 4 \mu \text{m}$ minimum feature size, and was designed
to operate at a wavelength of 633 nm. The design algorithm predicted a
diffraction efficiency of 89% and a uniformity variation of 4%. This DOE was
designed to reconstruct nine signal orders in a $3 \times 3$ array, with relative intensity
values given by 4:2:1 for the central (zero) order, the four nearest-neighbor
orders, and the four next-nearest-neighbor orders, respectively. The measured
diffraction efficiency was 74%, and the measured uniformity variation was 33%.

Another Honeywell-fabricated DOE described above, the $5 \times 20$ spot array
generator, was designed to have less than a 7% spatial nonuniformity and a
minimum of 15 dB cancellation of unwanted diffraction orders (including the
zero order). We measured these values with a DOE characterization system and
found a 14% nonuniformity and a 14 dB cancellation of unwanted diffraction
orders (including the zero order), showing excellent performance for a first
fabrication run.

We measured the Honeywell DOE performance before and after antireflection
coating and found the AR coating to be a significant benefit. We tested 10 spot
array generators and found that on average the diffraction efficiency was boosted
11%, the cancellation of unwanted diffraction orders increased by 2 dB, and the
uniformity was improved by 23%.

We have designed and are currently constructing a system for characterizing the
diffractive microlens arrays described above. The most important parameter for
microlenses is the degree of wavefront aberration. Our microlens
characterization system uses interferometric methods for determining the optical
wavefront uniformity across an array of microlenses.

We have theoretically analyzed and simulated the effect of differences between
the DOE design and illumination wavelengths. Such shifts can occur in the final
system, for example, because of optical source instabilities or variances in the
central emission wavelength across an array of source elements. Knowledge of
the effect of differences in wavelength also allows characterization of the DOE at
experimentally convenient wavelengths, and conversion of the resultant data to
the wavelength to be used in the final system. In the case of a binary-phase-level
DOE, the DOE generates a symmetric intensity pattern at any wavelength, and all
nonzero diffracted orders change by a single factor that depends on the change in illumination wavelength. In the case of multiple-phase-level DOE's, the relation is more complicated and depends on the particular DOE pattern.

- We have theoretically analyzed and simulated the effect of etch-depth error on reconstructed pattern of DOE's. For binary-phase-level DOE's, the effect of etch-depth error on the DOE reconstruction is identical to the effect of an incorrect illuminating wavelength. For multiple-phase-level DOE's, the change in the intensities of the reconstructed patterns depends on the etch-depth error in each etch step, as well as on the phase distribution of the DOE.

Significance to field and relationship to original goals. We view diffractive optical element (DOE) arrays as key components for chip-to-chip interconnections. The DOE provides fan-out, interconnection routing, and optional interconnection weighting (to implement synaptic weights in the case of artificial neural network interconnections). We have demonstrated: (1) designs that provide reasonably high optical throughput efficiency and reconstruction accuracy while providing reduced channel-to-channel interconnection crosstalk; (2) fabrication techniques that are reliable and potentially amenable to volume manufacture of DOE's; (3) testing methodologies capable of characterizing and evaluating the design and fabrication techniques; and (4) the application of DOE's to perform dense, point-to-point, 2-D parallel optical interconnects. Our overall goal is to develop DOE's tailored to highly dense, parallel systems for use in compact photonic multilayer computational structures.

4.5 Photonic multichip module (MCM) analysis and applications

We have investigated two interconnection architectures for photonic implementations of neural networks: one for global space-variant interconnections with full connectivity, and one for space-variant interconnections with limited fan-out from each node. Both are based on planar DOE's. The purpose of this part of the investigation was to study the feasibility of both architectures for compact multilayer structures, to use these results to develop performance projections for full-scale
multilayer systems based on our understanding of their technological limitations, and to assess the potential use of this system in a variety of application areas.

4.5.1 Interconnection system analysis

- We have analyzed the scaling properties of both fully-connected and limited-fan-out space-variant interconnection systems as applied to physically compact structures. The limited-fan-out system allows a very high density of chip-to-chip interconnections in a short-propagation-length system, can interconnect a large number of nodes, and tends to scale favorably with the number of interconnection nodes. The fully-connected system is restricted to a lower overall density of interconnections, requires a longer-propagation-length system, interconnects a smaller number of nodes, and tends not to scale as favorably with the number of nodes being interconnected.

- We have analyzed the theoretical degree of crosstalk due to the DOE in the fully-connected space-variant interconnection system [9]. The theoretical crosstalk is quite low for typical values of parameters characteristic of the DOE design and fabrication process. In practice, the overall crosstalk in such a system is likely to be determined by effects such as lens aberrations, component misalignment, and other sources of stray light. Many of these effects can be controlled through optimized optical system design and component manufacturing processes.

- We have also analyzed the theoretical degree of crosstalk due to the DOE array in the limited-fan-out space-variant interconnection system. For systems based on conventional DOE designs, the crosstalk is much higher than in the fully-connected system. The DOE is likely to be the main contributor to crosstalk in these systems.

- Because of these findings, we have subsequently developed a crosstalk reduction technique based on a modified DOE design. The resulting DOE designs can lower crosstalk in the limited-fan-out systems significantly, making the dense, compact interconnections of limited-fan-out space-variant system viable.
• To evaluate these crosstalk effects numerically, we have simulated a weighted optical interconnection system with 16,384 (≈ 128²) nodes in both the input and output planes. Each of the required 16,384 sub-DOE's (each with the number of phase levels = 16 and the size of the periods = 8 × 8 phase elements) connects an input node to the nearest 5 × 5 neighborhood in the output plane with connection weights between zero and one, and with an average diffraction efficiency of 79%. The results of this simulation verify the above qualitative statements on the relative degrees of crosstalk in the different systems. For example, in the limited-fan-out system, our crosstalk reduction technique reduced the two most significant contributions to crosstalk by more than an order of magnitude for a set of typical values of parameters.

Significance to field and relationship to original goals. We have developed a design methodology for the first, to our knowledge, optical 3-D space-variant interconnection system that provides dense interconnections with very short system propagation lengths, on-axis illumination and reconstruction, and (optionally weighted) fan-out provided by DOE's. Crosstalk reduction was necessary for the system to be viable in an analog (e.g., neural-network) system. These results remove a significant bottleneck to the implementation of complex interconnections in a compact multilayer computational structure or vertically interconnected photonic multichip module.

4.5.2 Multichip module system analysis

• We have analyzed a multichip module in which each layer has limited-fan-out connections to the chip in the next layer, has 10⁴ optical input/output ports per cm², and operates at 50 MHz bandwidth per input/output port. The fan-out is 16 (i.e., fan-out to a 4 × 4 neighborhood), which provides an aggregate bandwidth between vertically-adjacent chips of 8 Tb/s. The overall layer thickness is less than or equal to 2 mm. Power dissipation in the GaAs modulator-array chip, in the worst case (with all ports operating at maximum bandwidth at full duty cycle), is 1 W/cm², and in the silicon chip is determined primarily by the signal processing electronics. The total external optical-laser-diode power required is in the range of 0.2 to 4 W for each cm² of chip area. The pixel size is 100 μm × 100
μm, and each pixel includes: a 10 μm × 10 μm modulator (on the GaAs chip); a 10 μm × 10 μm detector and 10 μm × 10 μm bonding pad (with an 8 μm high bump bond) on the silicon chip; and the remainder of the silicon chip area within each pixel is available for signal processing and driving electronics. Each sub-DOE is 100 μm × 100 μm in size.

Significance to field and relationship to original goals. We have analyzed the photonic multilayer structure from the point of view of component density, power dissipation, bandwidth, and crosstalk. While there are many interrelated parameters and many possible sets of specifications, we have found a set of specifications that is mutually self-consistent, and within our current understanding and knowledge is achievable with the technology we are developing.

4.5.3 Applications

We have been assessing the applicability of this photonic multilayer computational structure to a variety of areas, as summarized below.

- The application areas of sensory data processing, image processing, and vision provide a natural mapping onto the photonic multilayer computational module architecture. In the early (low-level) stages of processing, an image size of 512 × 512 may be input and processed in a highly-parallel feedforward manner. The processing may be modeled after artificial neural networks (which in turn may be modeled after biological vision systems), or may be modeled after digital parallel image/vision processing structures and algorithms. In some applications it may be desirable, or even crucial, for the processing module to be small, rugged, and low power, e.g., for use in smart cameras or other rich sensory environments.

- Other applications involve various aspects of parallel processing. Multistage reconfigurable interconnection networks are highly parallel and hardware intensive, and stand to benefit substantially from photonic technology based on fixed optical interconnections and smart pixel SLM's. The compact multilayer computational structure we are developing provides a potential implementation
of such networks in a physically small, rugged package that uses manufacturable components.

- Parallel processors based on distributed or shared memory consist of a set of processing elements, each with a cache memory. Allowing parallel access among the cache memories and a possible main memory can decrease computational bottlenecks and significantly reduce the time spent on insuring cache coherence; although not specifically investigated in this effort, this application area may also benefit from the multilayer computational structure we have been developing. In a distributed-memory parallel processor, the communication among processing elements is crucial, and the use of photonics to increase the aggregate communication bandwidth, for example using cellular hypercube interconnections, is likely to be beneficial as well.

- We have demonstrated that the use of dense parallel optical data links between VLSI planes offers potentially large performance gains in SIMD-type computing applications. We have distributed processing elements across several two-dimensional planes linked with dense (62.5 micron pitch), high-speed (we have demonstrated 30 Mb/s per channel, and estimate that > 500 Mb/s is possible) optical interconnections, and have then performed parallel image processing routines without the array loading/unloading latency common to current SIMD architectures. We can link our 2-D parallel pipeline processor to memory and video or display devices to perform real-time image processing functions.

4.5.4 Integrated submodule demonstration project

- We have performed initial design and fabrication work towards the demonstration of a two-chip photonic multichip module (MCM), based on the technology described above. The system is designed to be capable of processing real-time image data received from a CCD chip. The MCM is designed to be mountable on a standard printed-circuit board, and to be used in a system that can recognize and track moving objects in real time.
4.6 References


5. Personnel Supported

Prof. Armand R. Tanguay, Jr., Principal Investigator
Prof. B. Keith Jenkins, Co-Principal Investigator
Prof. Alexander A. Sawchuk, Senior Investigator
Dr. Allan Weber, Research Scientist
Dr. Chris Kyriakakis, Postdoctoral Fellow
Dr. Charles B. Kuznia, Postdoctoral Fellow
Dr. Clare Waterson, Postdoctoral Fellow
Kartik Ananthnarayanan, Graduate Research Assistant
Scott DeMars, Graduate Research Assistant
Patrick Nasiatka, Graduate Research Assistant
Edward J. Herbulock, Graduate Research Assistant
Jen-Ming Wu, Graduate Research Assistant
Jeng-Feng Lin, Graduate Research Assistant
Wei-Feng Hsu, Graduate Research Assistant
Dan Olster, Graduate Research Assistant
Kuang-Yu Li, Graduate Research Assistant
Sabino Piazzolla, Graduate Research Assistant
Nainjeet Ramlagan, Graduate Research Assistant
Chingchu Huang, Graduate Research Assistant
Adam A. Goldstein, Graduate Research Assistant
Karen Tierney, Budget and Business Analyst
Gloria Halfacre, Administrative Services Coordinator I
Erlinda Varilla, Administrative Services Coordinator II
Mariela Ortiz, Student Worker
Xinh Berreman, Student Worker
Kevin Oliver, Student Worker
6. List of Publications and Presentations Resulting from this Grant

Listed below are peer-reviewed publications that were enabled by this grant.


10. W. -F. Hsu and A. A. Sawchuk, "Improved Usable Capacity for Optical Page-Oriented Memories with Smart Pixel Interfaces," International Symposium on Optical Memory and Optical Data Storage (ISOM/ODS 96), Optical Society of America, Maui, Hawaii (July 8-12, 1996).


on Spatial Light Modulators, Trends in Optics and Photonics Series (Optical Society of America, Washington, DC, 1997).


Listed below are Ph. D. theses that were in part enabled by this grant.


* Available from Signal and Image Processing Institute, 3740 McClintock Ave., EEB 404, University of Southern California, Los Angeles, CA 90089-2564; email gloria@sipi.usc.edu, or world-wide-web at http://sipi.usc.edu.

** Available from Optical Materials and Devices Laboratory, University of Southern California, University Park MC-0483, Los Angeles, CA 90089-0483); email ktierney@mizar.usc.
7. Interactions/Transitions

7.1. Participation/presentations at meetings, conferences, and seminars


15. A. R. Tanguay, Jr., B. K. Jenkins, and A. A. Sawchuk, "Dense 3-D Integrated Electronic/Photonic Computing Structures Enabled by Diffractive Optical

Dense 3D Integrated Electronic/Photonic Computing Structures 36


18. A. A. Sawchuk and W.-F. Hsu, "Smart Pixel Interfaces for Page-Oriented Optical Memories," Workshop on Data Encoding for Page-Oriented Memory (DEPOM'96), Phoenix, AZ, March 27-28, 1996 (associated with written paper 8, above).


21. W.-F. Hsu and A. A. Sawchuk, "Improved Usable Capacity for Optical Page-Oriented Memories with Smart Pixel Interfaces," International Symposium on Optical Memory and Optical Data Storage (ISOM/ODS 96), Optical Society of America, Maui, Hawaii, July 8-12, 1996 (associated with written paper 10, above).


7.2. Consultative and advisory functions

Below are listed consultative and advisory functions to other laboratories and agencies, especially Air Force and other DoD laboratories.

Workshop on Superconductive Electronics: Devices, Circuits, and Systems, Farmington, Pennsylvania, October 8-12, 1995, for Dr. Fernand Bedard of the National Security Agency.

Focused on critical aspects of packaging superconducting electronic chips with high input/output port requirements for applications in telecommunications (e.g., a high-bandwidth non-blocking crossbar switch), special purpose computational submodules (e.g., a $1024 \times 1024$ point FFT accelerator), and hybrid computational modules with dynamically reconfigurable interconnections (e.g., a rendering engine for computer animation and graphics).

Free-Space Optics Workshop, AFOSR, Washington, DC, (November 3-4, 1995)

7.3. Transitions

In an ongoing collaboration with TRW (with funding from NSA), we are continuing to investigate several technological issues that relate to indium bump flip-chip bonding of superconducting electronics for multichip module integration. Under this collaboration we also plan to explore hybrid electronic/photonic solutions.

Recently, a key aspect of the USC collaboration with TRW (assisting with the indium bump process technology needed for a fully functional superconductive crossbar switch) was very successful. The USC-developed "velcro" indium bumps exhibited superior mechanical yield strengths (much greater than reported in the literature thus far). A multiple chip attach to the "mother" 4" substrate was performed successfully, and the superconductive multichip module is currently undergoing electrical characterization. Preliminary results indicate successful
electrical continuity of all indium bonds even at temperatures below 4 degrees Kelvin.

8. Inventions and Patent Disclosures

No patents have been applied for, and no inventions have been disclosed, based on the work supported by this grant during this reporting period.