# Toward an optimal foundation architecture for optoelectronic computing. Part II.   Physical construction and application platforms

Haldun M. Ozaktas

Various issues pertaining to the physical construction of systems that are based on regularly interconnected device planes, such as heat removal and extensibility of the optical interconnections for larger systems, are discussed. Regularly interconnected device planes constitute a foundation architecture that is reasonably close to the best possible as defined by physical limitations. Three application platforms based on the foundation architecture described are offered.   © 1997 Optical Society of America

*Key words:*   Optical interconnections, optical computing.

## 1.   Introduction

In Ref. 1 (Part I of two; see pages 5682–5696, in this issue) the tree of possibilities for optoelectronic computing architectures was examined systematically and arguments that allow us to prune suboptimal branches of this tree were offered, and the conclusion was reached that electronic circuit planes interconnected optically according to regular connection patterns represent a choice that is reasonably close to the best possible, as defined by physical limitations.   Thus, it was proposed that this foundation architecture should provide a basis for future research and development in this area.   In this paper I discuss some aspects of its physical construction and also some of its applications.

Regularly interconnected device planes may be depicted schematically as shown in Fig. 1.   They consist of electronic circuit planes, with optical input and output, interconnected to each other in regular connection patterns that have efficient optical implementations.   The reader is assumed to have read Part 1, in which fundamental aspects of such systems are discussed.

## 2.   Physical Construction

In this section we discuss some of the more physical aspects of the architecture that was advocated in Ref.

1.   First, we must note that the way we have drawn multiple device-plane architectures (Fig. 1) in Ref. 1 is misleading, since this is not necessarily or preferably the way they would actually be constructed. Rather, a number of considerations indicate that it is actually best to employ device planes whose optical inputs and outputs all face in the same direction and to array the device planes side by side on the same geometrical plane.   For one thing, it is actually more convenient from the perspective of device technologies such as those described in Refs. 2 and 3 to have device planes whose sources–modulators and detectors look in the same direction and operate in reflection mode.   In the configuration depicted in the upper part of Fig. 2, where the several consecutive device planes are arrayed side by side on a single geometrical plane, interconnections are established by a reflective optical system lying on one side—say the top side—of the device planes.   The bottom side of this plane is reserved for heat-removal purposes. Overall, this configuration permits the attainment of the smallest possible system size as defined by heat removal and interconnection density considerations, while simplifying design by separating the interconnections and the heat-removal paths.   In addition, there might be additional engineering advantages that we have not considered that result from keeping all active devices in a single plane.   (In fact, now it becomes possible to integrate some of or all the several adjacent device planes on the same wafer if desired.)

Any of the other more obvious ways of configuring the consecutive planes we can think of result in suboptimal system sizes.   One example is shown in the
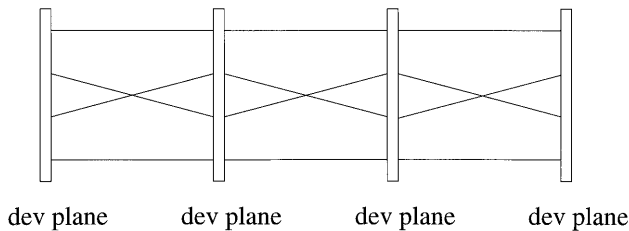
Fig. 1.   Regularly interconnected device planes.

lower part of Fig. 2, where we can see several consecutive planes situated in actual succession, with heat removal taking place from the bottom.   (Since we do not have full surface contact between the devices and the convective heat sink, the conductive resistance constitutes a bottleneck, limiting severely the amount of heat we can remove.)

The direction of flow of the coolant fluid is also critical.   Flow along the surface of the device planes would be suboptimal, whereas flow in a direction perpendicular to the device planes allows the smallest possible system size to be attained (Fig. 3).   Because interlacing optical and fluid-flow paths can be complicated, this is best achieved by use of one side of the device planes for optical interconnections and the other side for heat removal, as shown in the upper part of Fig. 2.   Although the folded coolant path shown in the lower part of Fig. 3 may be constructionally less straightforward than that shown in the upper part of the same figure, this approach results in better scaling behavior as we go to larger and larger arrays.[4]

My comments on the physical implementation of the optical interconnections are brief, since the various issues and options are well known.   (For instance, see Refs. 5–15 and the references therein.)   A well-studied option that is also consistent with our general assumptions is the use of microfacets directly above the light sources–modulators and detectors
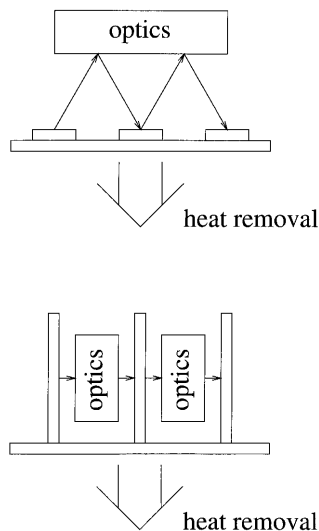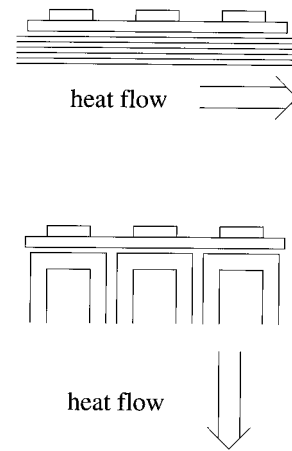


Fig. 3.   Alternative directions of heat flow.

and a mirror or reflective optical element at the top (for instance, as shown in Fig. 4 and as described in Refs. 6, 12, 13, and 16).   Because only a regular pattern of interconnections is to be implemented, the optical apertures interconnecting each stage to the next must be divided into only a small number of subapertures so that the size of a single stage, as imposed by interconnection density, will be nearly equal to the smallest size that is physically possible: $\propto N^{1/2}$ (this is the size of the lateral extent of a single stage as well as the distance of the reflective optics from the device planes).   Planar free-space packaging approaches that fix the distance of the reflective optics from the device planes to a much smaller value would not permit the desired connections to be implemented in a system of size $\propto N^{1/2}$ since the interconnection capabilities of such systems are essentially similar to two-dimensional systems.[13]

The resulting system is depicted in Fig. 5.   The devices are partitioned into several stages, which are interconnected to each other with regular connection patterns.   Alternatively, we might choose to use the more simple single-stage system shown in Fig. 6 by incurring a moderate increase in the total time of computation, as discussed in Ref. 1.

One of several general issues and problems that might be treated with respect to a particular application is the problem of adjusting the optical pixel density with respect to the electronic circuit density.



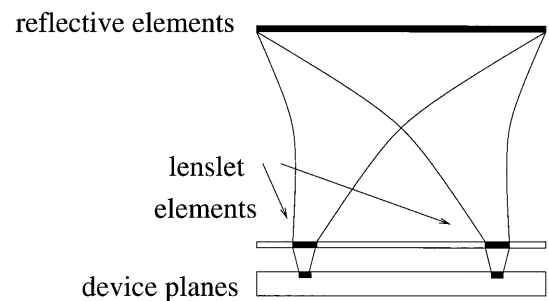Fig. 2.   Arraying device planes for heat removal.



Fig. 4.   Schematic depiction of a single-facet space-invariant architecture.
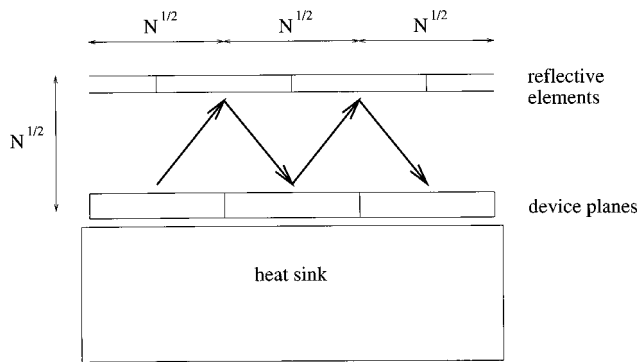
Fig. 5. Foundation architecture (multistage version).

This problem is essentially the same as that treated in Refs. 17–19, so that one can arrive at specific conclusions for a particular situation by modifying the treatment in these references. One result emerging from the simulations reported in Ref. 18 that we might note is that it is often more efficient to cluster the electronic circuits into islands rather than spreading them out uniformly over the plane of devices. This clustering decreases the electronic power dissipation, reducing the overall system size.

The device-plane concept may be generalized to allow for multilayer structures, be they multilayer wiring substrates or multiple-active-plane technologies.[20] Optical input–output devices would be mounted on the outermost layers. The resulting compound device planes would be able to provide more silicon circuitry per optical pixel, should that prove desirable.

Many things considered, both physical and technological, it is hard to think of any other architecture in which almost everything fits in place so harmoniously and in which the various constraints are balanced and physical optimums are closely achieved. In fact, there was no *a priori* reason to expect the existence of an option that is as satisfactory as the one discussed. In Ref. 21, Miller sets ou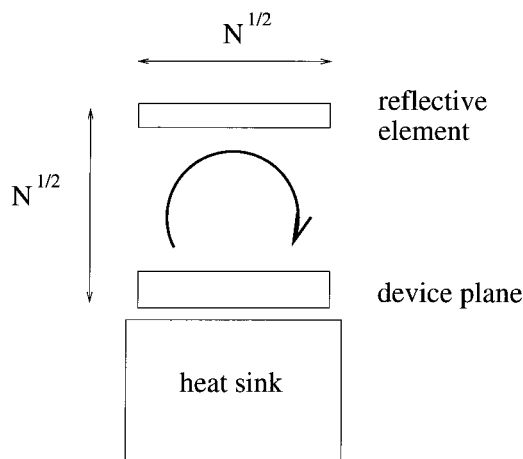t the requirements that an optical switch to be used in a large-scale digital system would need to satisfy. He then argues that, whereas most proposed devices fail in this respect, there do exist devices (self-electro-optic-effect devices) that indeed meet these requirements. I believe that the use of these devices, together with the foundation architecture discussed in this study, can provide a path for success in opto-electronic computing systems.

## 3. Future Research Directions

### A. Limits to Scaling

Figure 5 depicts the general physical architecture that is suggested as a foundation for optoelectronic system development. We arrived at this architecture through arguments involving how things scale with increasing numbers of elements, claiming that it is optimal in many respects.

The heat sink should pose no problems. The analysis of Ref. 4 is valid for systems at least as large as 1 m, and in any event it is quite evident that there is no physical barrier against removing constant power per area. Although the construction of the heat sink would probably not be altogether trivial, it does not seem that it would require any technological breakthrough.

Scaling of the device planes (the smart pixels) is limited by the decrease in yield and the difficulty of maintaining uniformity for larger planes. However, it seems that we can surmount both of these limitations by tiling (patching) together individually tested chips or wafers.

What is less evident is whether the optical system can also be scaled to large dimensions. Most optical systems are limited by the space–bandwidth product of the optical elements used. Although it has been shown that there is no fundamental upper limit for the space–bandwidth product of spherical lenses,[22] it may still not be possible to manufacture lenses of large size while maintaining the necessary uniformity and quality. In that case, it will be necessary to devise ways to patch together optical elements of a given size to create systems of ever-larger sizes. Let us note that it is necessary for the devised system to provide the global pattern of connections over the full effective aperture interconnecting each stage to the next and not merely over the constituent element apertures (as the latter would correspond to mere replication of the single-element system).

The objective may be summarized as follows:

Array optical elements of constant size to design a scalable optomechanical platform that can provide a regular connection pattern among an array of $N^{1/2} \times N^{1/2}$ points lying on a plane, such that the overall system can be housed in a box whose dimensions are $\propto N^{1/2}$ and the constant of proportionality is not much greater than the order of unity. Choose the number of surfaces each light beam must pass through to grow as slowly as possible to limit excess attenuation; $\leq \log N$ would be a reasonable rate of growth.

Given that there seem to be no fundamental barriers involved, we believe that among the many ways



Fig. 6. Foundation architecture (single-stage version).

such a system can be constructed, at least one design that is also acceptable from a manufacturing and packaging viewpoint will be found. (A study of the capacity of single- and multiple-aperture lens systems that might be relevant for the above discussion is given in Refs. 8 and 23.)

### B. Characterization of System-Level Trade-Offs

Let us also mention briefly a number of different trade-offs that arise in the types of systems proposed. If greater numbers of optical channels and pixels cannot be realized because the limitations on scaling cannot be overcome efficiently, one might consider the option of trading speed for numbers of pixels. The speed of the devices involved may be much larger than that needed for certain purposes, such as parallel pipelined video processing, for which the number of pixels available with current technology is not sufficient. Multiplexing several image pixels through a single optical pixel may permit a beneficial trade-off to be realized.

The number of optical pixels that can be fitted onto a chip, as well as the number of optical beams that can be handled with aberrations small enough to meet the requirements of digital systems, both currently seem to be limited to approximately $100 \times 100$. On the other hand, typical images that one would like to work with might be $1000 \times 1000$. One option is to multiplex local groups of image pixels through a single optical pixel. If a video image at 100 frames/s has 10 bits/pixel, we would have to deal with $10^5$ bits/s per optical pixel, which is quite manageable and still leaves room for the possibility of iterative processing.

Of course, the silicon beneath the optical pixel will have to latch, store, and process 100 image pixels' worth of information in conjunction with neighboring pixels. Channeling the information associated with several image pixels through a single optical pixel will also introduce constraints on the types of global connection patterns possible, an issue that has to be considered and analyzed explicitly for a given situation.

A final trade-off that we barely mention is that between globality and speed. Most global operations can be broken down (or factored) into a succession of local operations. The cumulative effect of consecutive local operations can allow the necessary global transfer of information to take place. An example is the use of small generating kernels for performing convolutions with large kernels (Ref. 24, p. 253). In this technique, a convolution kernel is factored into a large number of kernels, each of which is of small extent. This in effect allows us to eliminate the need for global connections but requires a greater number of time steps for completion of the operation. This is similar to the use of a locally connected computer network to simulate a globally connected one by routing of the information in several hops. Thus an operation requiring global information flow may either be implemented with physical interconnections matching the global connection pattern demanded by

the problem or it may be reformulated and implemented with physical interconnections exhibiting a lesser degree of globality. The latter case will require a greater number of steps or iterations and also perhaps a significantly greater capacity per physical channel (since many global communication paths may have to share the same local interconnection channels). The trade-offs here are similar to those analyzed in Ref. 25. In conclusion, if our devices are sufficiently fast, we may be able to give up some of the benefit of their speed to relax the demand for full globality of connections, which in turn will permit an increase in parallelism through hardware replication.

### 4. Application Platforms Based on Regularly Interconnected Device Planes

Recently there has been significant progress in making so-called smart pixels—two-dimensional arrays of electronic processing units, each with optical inputs and outputs.[2,3,26–34] Such smart pixels—or device planes, as we have preferred to refer to them—have the potential to capitalize on the advanced state of certain electronics technologies, together with the advantages of optics as an interconnection medium. One significant challenge with such electronic arrays is to devise concepts and architectures that provide a platform for realizing this potential.

The regularly interconnected device-plane architecture advocated in this work is a foundation architecture on which to build platforms on which, in turn, more specific applications can be implemented. The general features of the regularly interconnected device-plane architecture were determined by purely physical considerations. The more specific features of an evolved platform must be determined by the class of applications one wishes to implement. An analogy with programming languages may be useful. The regularly interconnected device-plane architecture is in some ways analogous to a low-level programming language, such as machine language. It is not a very efficient process for one to design a system for a particular application by starting directly from the bare, regularly interconnected device-plane concept, much as it is inefficient to write an application directly in machine language. The platforms described below may be likened to high-level programming languages, with which it is much more efficient to implement a desired application.

Although all programming languages are equivalent in a general sense, they are not equally efficient for a given purpose. The choice of a suitable platform, much like the choice of a suitable programming language, depends on the type of application that is to be implemented. New classes of applications might require the development of new platforms. Basing such platforms on regularly interconnected device planes will assure us that they exhibit desirable or optimal characteristics from a physical point of view.

Perhaps one of the most successful platforms in information processing is that based on the microprocessor. The importance of the microprocessor con-

cept stems from its having provided a platform on which computer technology could develop in the way it has. The ingredients of the microprocessor paradigm are the stored-program concept, VLSI technology, and some secondary ideas, such as the data-path concept, etc. Similar (although not necessarily directly analogous) ingredients must be put together if we wish to come forward with a new platform with comparable impact.

Some general paradigms on which alternative platforms may be based already exist, such as cellular automata, connectionist systems, etc. Unfortunately, these are not sufficiently well developed. For instance, the techniques for only the lowest levels of abstraction are developed for cellular automata; no one has a high-level programming language that they can compile into some kind of assembly language that will run on some kind of cellular-automata hardware (which consists of several levels of abstraction down to the level of a single cell). The state of development of techniques for doing things with cellular automata is comparable with low-level logic in mainstream digital systems, such as shift registers, etc., and is far below the level of a microprocessor. (It has been shown how to simulate conventional logic operations in cellular automata, so that one can, in principle, do anything with a cellular automaton that one can do with conventional logic. However, this is a meaningless approach if the cellular automata is implemented by use of logic gates in the first place or simulated on a workstation. But things may change if cellular automata are implemented by virtue of some atomic-scale physical phenomena, etc.)

Regularly interconnected optoelectronic device planes correspond to the electronic integration technology that made the microprocessor possible. In the three subsequent sections, I try to provide the remaining ingredients for three platforms whose main characteristic is that they take as their starting point the capabilities of optics as manifested in our foundation architecture rather than some abstract computational paradigm or platform that was originally developed for electronics. They allow common application areas to be mapped directly onto the physical architecture we have argued is optimal. We do not claim that these are the only or the best application platforms and do not exclude other possibilities. For instance, database processing is one such application area that could be added to those below.[35–37]

## A. Digital Fourier Optics

The first platform or class of systems that is based on regularly interconnected device planes that we discuss may be referred to as digital Fourier optical systems. These are digital equivalents of analog Fourier optical systems, in which numbers are represented by digital bit streams. The basic architectural concepts of analog Fourier optics permit the specification of a platform enabling the con-
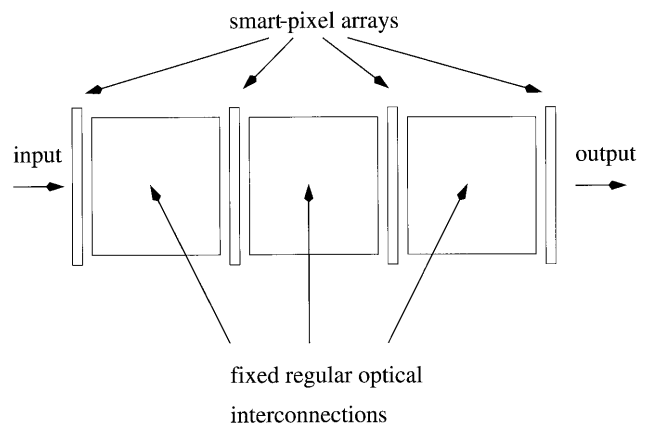


Fig. 7. Digital Fourier optics system.

struction of a broad range of interesting digital signal-processing systems. It is important to emphasize that we do *not* propose merely to insert smart-pixel arrays in existing Fourier optical systems. Rather, we propose a generic architecture with which the digital optical equivalent of essentially any analog Fourier optical system can be realized. More precisely, in Ref. 38 my colleague and I showed that the digital optical equivalent of any analog optical system consisting of lenses, complex spatial filters, holograms, masks, spatial light modulators, etc., can be constructed with properly chosen (or programmed) smart-pixel arrays interconnected in succession with fixed and regular connection patterns (Fig. 7).

Analog Fourier optical processing systems can perform important classes of signal-processing operations in parallel but suffer from limited accuracy. The digital optical equivalents of such systems share many features of the analog systems while permitting greater accuracy. There are many possible applications for such systems, as well as many alternative technologies for constructing them; in Ref. 38 we discussed the potential of free-space interconnected active-device-plane-based optoelectronic architectures as a digital signal-processing environment. Implementation of the active-device planes through hybridization of optoelectronic components with silicon electronics should permit the realization of systems whose performance exceeds that of purely electronic systems. A number of comparisons with state-of-the-art digital signal-processing chips indeed shows that digital Fourier optical systems can offer significant reduction of the overall time of computation.[38]

The major disadvantage of these systems is that, with current smart-pixel technology, the number of optical channels may be limited to approximately a thousand. Thus, they would be beneficial for applications in which amplitude accuracy is important, but limited resolution is sufficient. However, with the introduction of scaling schemes such as those suggested in Subsection 3.A, this restriction would be removed.

## B. Two-Dimensional Data-Flow Architectures

Our second proposal is what we call a two-dimensional data-flow architecture for a programmable image-processing machine.[39] Most image-processing applications involve a series of consecutive operations to be performed on an image. These operations may involve global flows of information throughout the extent of the image. (Operations involving only local flows of information can probably be implemented more efficiently electronically.) The operations to be performed on various parts of the image can be made parallel to a significant extent. For these reasons, the operations to be performed map nicely onto a multistage architecture, which we have argued is attractive from the hardware point of view as well. By the term programmable we mean that the system can be used to realize a large class of image-processing operations with adjustable parameters.

One does not design a microprocessor by the specification of its logical function and then use of a standard procedure to find the optimum circuit. The task is too difficult for such a one-step procedure. A successful approach is to employ a linear data path along which the words flow and are operated on in the process. This approach not only enables the desired functionality to be designed in a systematic way but also leads to an area-efficient implementation, well matched to the limitations of planar integration.

Three-dimensional integration technologies may inspire similar architectures. For instance, let us consider a three-dimensional integration technology that consists of a stack of several integrated circuit chips that have been connected in the third dimension by an array of vias. One way of using this technology would be simply to lay out on it a logic circuit designed in any conventional manner. A more efficient approach might be to use this technology to house a two-dimensional data-flow architecture, with the data flowing in the third dimension perpendicular to the active-device planes.

The two-dimensional data-flow architecture employing optical interconnections would provide vastly greater flexibility by permitting global connections. Depending on the application (logic, image processing, matrix algebra, etc.), we might select several functions by appropriately setting control lines, much as we can select between functions such as ADD, SHIFT, etc., in a microprocessor. Furthermore, several operations may be performed in pipeline fashion as the input propagates through several stages (Fig. 8). Thus the two-dimensional data-flow architecture is essentially a platform for parallel and pipelined processing of two-dimensional data streams.

In such a system we must exploit the potential for parallelism and pipelining by mapping the desired operations and data properly on the architecture. As an example, let us consider an object-recognition system. Various consecutive image-processing operations must be performed on the raw image. These might be preprocessing, edge detection, edge thinning, edge thresholding, segmentation, feature
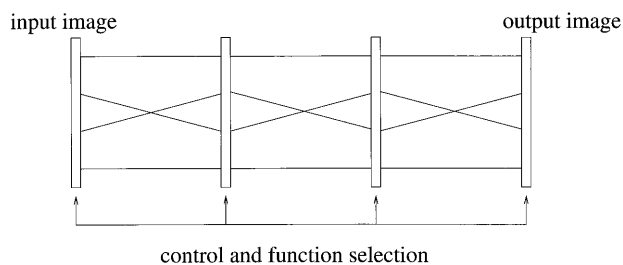


Fig. 8. Two-dimensional data-flow architecture.

extraction, and classification.[40] We might assign each operation to one or more device planes, depending on the amount of silicon needed for each as dictated by their relative complexity. By allocating more silicon for those operations that take more time, we can balance the workload through the pipeline. The preprocessing and edge operations require local communication only, so that optical interconnections are not necessary between the initial stages. The latter set of operations (feature extraction, classification) may require global connections, which can be provided optically. (Since it is beneficial to use optics between some stages, we might use it to provide the local connections between the initial stages as well, depending on whether this simplifies construction.) Notice that, in this example, the problem of parallelizing and pipelining the task is fairly straightforward because of the nature of the operations and the data. The problem of mapping the operations and data onto the physical architecture might be more challenging in other applications.

One of the most important aspects of this approach is its programmability. The several stages may each be designed to provide a certain set of operations with several free parameters, corresponding to the instruction set of a specialized high-level programming language, that can be used to realize a relatively general class of image-processing algorithms. This process would lead to the design of an advanced image coprocessor. (Of course, there is no reason why this general approach could not be used for other applications, such as a matrix processor or even a three-dimensional general purpose microprocessor.)

In conclusion, we see this type of platform as a candidate for similar success as the microprocessor. The strength of this concept lies in the excellent mapping of computational requirements onto a physically convenient architecture. Other work on optoelectronic architectures with image-processing applications include Ref. 41.

## C. General Purpose Parallel Computers

From the earliest days of work on optical interconnections it has been suggested that free-space optics may permit the construction of global computer interconnection networks. Our discussion implies that, among the various alternatives suggested, we should choose those based on regularly connected processor arrays rather than random-access models that require arbitrary dynamic connections. Al-
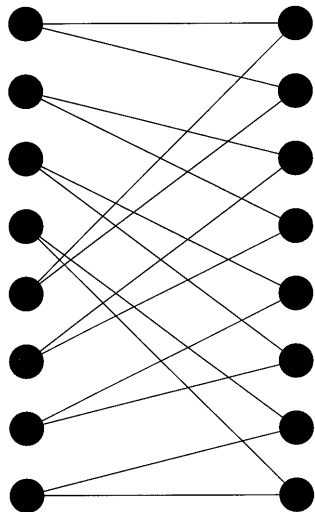
Fig. 9. General purpose parallel computer architecture. The nodes on the right-hand side represent the same eight processors as do the nodes on the left-hand side.

though many variations are possible, for clarity let us consider a single device plane that has been partitioned into some number of processors, such that each processor has two optical inputs and outputs. We may assume, for instance, that this array of processors is connected to itself according to a perfect-shuffle pattern (Fig. 9). (An immediate extension is to have several planes in cascade. Then the system exhibits the potential for both parallelism and pipelining.) The device plane may be a silicon wafer or large chip, or perhaps some kind of multichip assembly, furnished with optoelectronic input–output devices.

The use of a regular connection pattern such as the perfect shuffle might seem restrictive since each processor has a limited choice as to which other processors it can send information. However, for this type of architecture algorithms are developed such that the routing of information is interspersed with the processing more uniformly than in a case in which an arbitrary pattern of connections is available. Generally speaking, two planes of processors connected to each other with a regular connection pattern, a single plane of processors connected among themselves according to a regular pattern, or a multistage system in which each stage is connected to the next with a regular pattern can simulate a set of processors that are connected among themselves with a dynamic permutation network, with at most a $\log N$ slowdown in terms of the number of time steps. Since the increase in the number of time steps will be more than compensated by the reduction in the duration of a time step afforded by a regularly connected system, the overall performance will be superior for the case under consideration. As we have previously argued in a more general context, the intermediate stages of a multistage permutation network would require active devices anyway for the purpose of dynamic reconfigurability, signal regeneration, or both. Thus,

it makes sense to consider systems in which the processing power is distributed throughout the stages because this will permit more efficient use of the silicon in the many active-device planes.

Considerable effort has already been spent and more will be on the development of algorithms for such parallel-computing models. Thus combination of this effort with the technology under consideration may result in the realization of a powerful, optically interconnected multiprocessor computer. It might consist of a large number ($\sim 1000$) of processors on a large wafer or a multichip module. Faulty processors can be eliminated by individual testing prior to mounting or bypassed either by appropriate adjustment of the optical interconnection network (a redundant network should enable routing around faulty processors) or, preferably, adjustment at the software level. The optimal grain size of the processors can be determined from the results or by use of the methods from several studies that have already addressed this problem.[18,19]

D. Photonic Switching Fabrics

Another platform based on the regularly interconnected device-plane concept is that permitting the realization of multistage interconnection networks for telecommunications switching applications. Since this class of systems has been treated extensively elsewhere,[42] we do not discuss it here.

5. Discussion

It is my opinion that advanced smart-pixel technology such as that described in Refs. 2 and 3, combined with the physical foundation architecture described in this paper (either in its bare form or as developed on the basis of ideas similar to those discussed in Section 4), constitutes a promising platform that can be turned over to circuit, system, and computer designers for further development.

If a particular institution specializing in optoelectronic devices, systems, and packaging can make available a platform of the nature described above on which individuals or small groups can design systems for applications of their choosing, one of the important conditions for takeoff will have been satisfied. The optimality of the foundation architecture should assure such users that the end product will be one that provides more or less the best one can hope for. In other words, system designers are told just to design and optimize the algorithms and VLSI circuits and not to worry about the optics, since use of the foundation architecture described in this research gives us reason to believe that we should be doing fairly well in this respect.

At this point, we believe it makes sense to provide the platform with the optics all manufactured in the form of fixed interconnections, perhaps with a choice among different optical-pixel to silicon-circuit ratios, and to let customers design the VLSI circuits without worrying about the optical design at all. Only minor modification to existing design and simulation tools that take into account the optical connections would

be needed, since application designers will not be engaged in optical design but merely in ordinary VLSI design. The modified design tools should simply account for the existence of global optical paths while routing the circuits in the most efficient way possible (the layout problem), and the modified simulation tools should be able to account for the delays, energy requirements, and parasitic effects associated with the optical interconnections.

It should be possible for foundries to provide a series of standard platforms with a choice of total silicon circuitry and number of optical connections, together with a complete map of the optical wiring provided, and software tools. Customers can then design and simulate systems and submit their VLSI layouts, on the basis of which the foundry can manufacture for them a packaged system. (At a later stage, it might also be possible to offer field-programmable semicustom platforms in the form of gate arrays or the like.) The most important point here is that the optics are always fixed and not subject to design by the customer. (It is perhaps worth restating one of the results of Ref. 1: We do not lose much by giving up the freedom of being able to realize arbitrary interconnection patterns, so there is no significant price to pay for hardwiring the optics.) The fact that customers do not have to worry about optical design but only about the VLSI design of circuits should make the whole idea of optoelectronic computing seem much less esoteric and more attainable.

## 6. Conclusion

It is clear that large arrays of fast and low-energy optical devices integrated with established electronic technology and interconnected with free-space optics has large computational power in the raw sense, but realizing this potential may not be so easy. The difficulty stems from the fact that a whole system of paradigms and levels of abstraction has been constructed around the capabilities and limitations of purely electronic systems, and the dominance of this system resists the introduction of a new technology with completely different capabilities and limitations. There does not seem to be much point in trying to build an optical microprocessor, and the user-level improvements obtained by replacement of the longer wires in conventional systems with optics may be limited. On the other hand, starting with an array of smart pixels, we are too many levels of abstraction away from being able to write a program that plays chess. Clearly, considerable research is needed to determine how optoelectronic computing systems should be contemplated and to develop platforms that could guide future efforts.

The foundation architecture we have argued for is depicted in Fig. 5. It consists of an array of primitive electronic elements lying in a plane, with a certain number of optical inputs and outputs corresponding to each element. The power dissipated by these elements is removed by fluid convection, with the paths of fluid flow perpendicular to the device plane. The thickness of the coolant paths is chosen according to the analysis in Ref. 4. The optical interconnection system provides regular global interconnections (such as the perfect shuffle or the most significant stage of the Banyan) among the optical pixels, perhaps by use of a nearly space-invariant optical system, as described in Refs. 6, 9, 12, 13, and 16 and shown in Fig. 4. What we have purported to show in this study is that, provided we use this type of architecture properly by employing the appropriate algorithm and so forth, no other architecture can do much better. We believe that development of the application platforms we have proposed on the basis of our foundation architecture will constitute promising and fruitful avenues for further research.

## References

1. H. M. Ozaktas, "Optimal foundation architecture for optoelectronic computing. Part I. Regularly interconnected device planes," Appl. Opt. **36,** 5682–5696 (1997).
2. K. W. Goossen, J. E. Cunningham, and W. Y. Jan, "GaAs 850 modulators solder-bonded to silicon," IEEE Photon. Technol. Lett. **5,** 776–778 (1993).
3. K. W. Goossen, J. A. Walker, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. D. Bacon, D. Dahringer, L. M. F. Chirovsky, A. L. Lentine, and D. A. B. Miller, "GaAs MQW modulators integrated with silicon CMOS," IEEE Photon. Technol. Lett. **7,** 360–362 (1995).
4. H. M. Ozaktas, H. Oksuzoglu, R. F. W. Pease, and J. W. Goodman, "Effect on scaling of heat removal requirements in three-dimensional systems," Int. J. Electron. **73,** 1227–1232 (1992).
5. J. Jahns and M. J. Murdocca, "Crossover networks and their optical implementation," Appl. Opt. **27,** 3155–3160 (1988).
6. M. R. Feldman, C. C. Guest, T. J. Drabik, and S. C. Esener, "Comparison between optical and electrical interconnects for fine grain processor arrays based on interconnect density capabilities," Appl. Opt. **28,** 3820–3829 (1989).
7. A. W. Lohmann and A. S. Marathay, "Globality and speed of optical parallel processors," Appl. Opt. **28,** 3838–3842 (1989).
8. A. W. Lohmann, "Image formation of dilute arrays for optical information processing," Opt. Commun. **86,** 365–370 (1991).
9. G. E. Lohman and K.-H. Brenner, "Space-variance in optical computing," Optik (Stuttgart) **89,** 123–134, 1992.
10. J. Giglmayr, "Locality and decomposition of regular optical interconnection patterns," Appl. Opt. **33,** 6157–6167 (1994).
11. T. J. Drabik, "Optoelectronic integrated systems based on free-

space interconnects with an arbitrary degree of space variance," Proc. IEEE **82,** 1595–1622 (1994).

12. H. M. Ozaktas and J. W. Goodman, "Lower bound for the communication volume required for an optically interconnected array of points," J. Opt. Soc. Am. A **7,** 2100–2106 (1990).

13. H. M. Ozaktas, Y. Amitai, and J. W. Goodman, "Comparison of system size for some optical interconnection architectures and the folded multifacet architecture," Opt. Commun. **82,** 225–228 (1991).

14. H. M. Ozaktas and D. Mendlovic, "Multistage optical interconnection architectures with the least possible growth of system size," Opt. Lett. **18,** 296–298 (1993).

15. H. M. Ozaktas, K.-H. Brenner, and A. W. Lohmann, "Interpretation of the space-bandwidth product as the entropy of distinct connection patterns in multifacet optical interconnection architectures," J. Opt. Soc. Am. A **10,** 418–422 (1993).

16. D. A. B. Miller, "Optical interconnection of devices on chips," U.S. patent 4,711,997, 8 December 1987.

17. H. M. Ozaktas, "A physical approach to communication limits in computation," Ph.D. dissertation (Stanford University, Stanford, Calif., 1991).

18. H. M. Ozaktas and J. W. Goodman, "Elements of a hybrid interconnection theory," Appl. Opt. **33,** 2968–2987 (1994).

19. A. V. Krishnamoorthy, P. J. Marchand, F. E. Kiamilev, and S. C. Esener, "Grain-size considerations for optoelectronic multistage interconnection networks," Appl. Opt. **31,** 5480–5507 (1992).

20. M. J. Little and J. Grinberg, "The 3-D computer: an integrated stack of WSI wafers," in *Wafer-Scale Integration* (Kluwer Academic, New York, 1988), Chap. 8.

21. D. A. B. Miller, "Device requirements for digital optical processing," in *Digital Optical Computing*, Vol. CR35 of SPIE Critical Review Series (SPIE Press, Bellingham, Wash. 1990), 68–76.

22. H. M. Ozaktas and H. Urey, "Space–bandwidth product of conventional Fourier transforming systems," Opt. Commun. **104,** 29–31 (1993).

23. H. M. Ozaktas, H. Urey, and A. W. Lohmann, "Scaling of diffractive and refractive lenses for optical computing and interconnections," Appl. Opt. **33,** 3782–3789 (1994).

24. W. K. Pratt, *Digital Image Processing*, 2nd ed. (Wiley, New York, 1991).

25. H. M. Ozaktas and J. W. Goodman, "Comparison of local and global computation and its implications for the role of optical interconnections in future nanoelectronic systems," Opt. Commun. **100,** 247–258 (1993).

26. Special issue on Smart Pixels, IEEE J. Quantum Electron. **29**(2) (1993).

27. S. Esener, "Smart pixels: technology and applications to parallel computing," in *Spatial Light Modulator Technology*, U. Efron, ed. (Dekker, Dordrecht, The Netherlands, 1994).

28. D. J. McKnight, M. A. Follett, and K. M. Johnson, "Liquid crystal over silicon spatial light modulators," Inst. Phys. Conf. Ser. **139,** 535–538 (1995).

29. M. P. Y. Desmulliez, J. F. Snowdon, A. J. Waddie, and B. S. Wherrett, "Critical issues in smart pixel design," in *Optical Computing*, Vol. 10 of OSA 1995 Technical Digest Series (Optical Society of America, Washington, D.C., 1995), pp. 96–98.

30. M. K. Smit, "Compact components for semiconductor photonic switches," in *Proceedings of the 1996 International Topical Meeting on Photonics in Switching* (1996), paper PWB1.

31. J. A. Neff, C. Chen, T. McLaren, C.-C. Mao, A. Fedor, W. Berseth, Y. C. Lee, and V. Morozov, "VCSEL/CMOS smart pixel arrays for free-space optical interconnects," in *Proceedings of the Third International Conference on Massively Parallel Processing Using Optical Interconnections* (*MPPOI '96*), (IEEE Computer Society, Los Alamitos, Calif., 1996), pp. 282–289.

32. T. Kurokawa and T. Ikegami, "Optical interconnection technologies based on vertical-cavity surface-emitting lasers and smart pixels," in *Proceedings of the Third International Conference on Massively Parallel Processing Using Optical Interconnections* (*MPPOI '96*) (IEEE Computer Society, Los Alamitos, Calif., 1996), pp. 300–305.

33. A. Kirk, H. Thienpoint, V. Baukens, N. Debaes, A. Goulet, P. Heremans, M. Kuijk, G. Borghs, R. Vounckx, and I. Veretennicoff, "Demonstration of parallel optical data input for arrays of PnpN optical thyristors," in *Proceedings of the Third International Conference on Massively Parallel Processing Using Optical Interconnections* (*MPPOI '96*) (IEEE Computer Society, Los Alamitos, Calif., 1996), pp. 360–366.

34. A. Z. Shang and F. A. P. Tooley, "Design of smart pixel receivers and transmitters for free-space optical backplane," paper presented at the Optical Society of America Annual Meeting, 20–25 October 1996, Rochester, N.Y.

35. P. B. Berra, A. Ghafoor, P. A. Mitkas, S. J. Marcinkowski, and M. Guizani, "The impact of optics on data and knowledge base systems," IEEE Trans. Knowl. Data Eng. **1,** 111–132 (1989).

36. P. B. Berra, K.-H. Brenner, W. T. Cathey, H. J. Caulfield, S. H. Lee, and H. Szu, "Optical database/knowledgebase machines. Appl. Opt. **29,** 195–205 (1990).

37. S. Akyokus and P. B. Berra, "A data/knowledge base machine based on an optical content addressable memory," Opt. Comput. Process. **2,** 179–187 (1992).

38. H. M. Ozaktas and D. A. B. Miller, "Digital Fourier optics," Appl. Opt. **35,** 1212–1219 (1996).

39. The use of the term data-flow architecture is not related to its use in computer engineering.

40. Y. Moon, N. Bagherzadeh, and J. Sklansky, "Macropipelined multicomputer architecture for image analysis," J. Opt. Soc. Am. A **6,** 951–962 (1989).

41. P. A. Mitkas, F. R. Beyette, Jr., S. A. Feld, L. J. Irakliotis, and C. W. Wilmsen, "Optoelectronic parallel processing with straight-pass optical interconnections and smart pixel arrays," in *Proceedings of the First International Workshop on Massively Parallel Processing Using Optical Interconnections* (*MPPOI '94*) (IEEE Computer Society, Los Alamitos, Calif., 1994), pp. 174–181.

42. H. S. Hinton, *Introduction to Photonic Switching Fabrics* (Plenum, New York, 1993).