OPTICAL INFORMATION PROCESSING: PAST, PRESENT, AND FUTURE

Haldun M. Ozaktas

Department of Electrical Engineering, Bilkent University
TR-06800 Bilkent, Ankara, Turkey
phone: + (90) (312) 290 16 19, fax: + (90) (312) 266 41 92, e-mail: haldun@ee.bilkent.edu.tr

ABSTRACT

A historical overview of optical information processing, emphasizing the influence of and interactions with signal processing will be given. Analog optical signal processing, holography, digital optical computing, optical interconnections and optoelectronic computers, and optical switching systems are discussed. Optical and electronic devices and interconnections are compared.

1. INTRODUCTION

Optics is the branch of physics dealing with phenomena involving light. Light refers to electromagnetic radiation of frequencies within or close to the band which humans can sense. The term optical information processing can mean two things: the processing of optical information, or, the optical processing of information. (Naturally the two overlap when we are concerned with optical processing of optical information.) Here we will only briefly discuss the processing of optical information and then turn our attention to the optical processing of information. Optical information refers to the time- and space-dependent information carried by light fields. Mathematically, these can be expressed in the form of functions such as $f(x)$, $f(t)$, $f(x,y)$, $f(x,y,z,t)$. Here f represents the amplitude or intensity of the light field, t represents time, and x,y,z represent spatial coordinates. The polarization, wavelength, and other properties of light may also carry information which is of interest to us. An image coming from outer space and captured at the camera at the end of a telescope in astronomy, an image obtained by methods such as endoscopy in medicine, and more generally all images carried to our eyes or to a camera with light are examples of optical information. It is also possible to give non-image examples such as the time-dependent signals in an optical fiber or spectroscopic measurements showing the absorption or emission at different wavelengths. Optical information may be processed for the purpose of correcting or eliminating different kinds of distortion or noise, or transforming them to a form more amenable to human perception or interpretation. Today, optical information is processed often with digital computers, but also with analog optical systems. Here we will discuss analog optical processing in detail, but not concern ourselves with the processing of optical information by digital means.

The four key areas of information technology are the communication, processing, storage, and sensing of information. Optical fibers are commonly used to transmit telephone conversations and computer data from place to place. Wireless optical communication also has many applications, such as remote control units in consumer products. Optical technologies also have a prominent place in the storage of information, as evidenced by the widespread use of optical disks. As regards sensing, both point sensors such as photodetectors and image sensors such as cameras are widely used. So far the processing of information is the area where optics has been used the least.

In the area of communication and storage, optical technologies have increasingly replaced electrical and electronical technologies over the past decades. Some researchers believe that optics will likewise increasingly replace electronics in the processing of information. While this claim is controversial, most experts would agree that optics will play an increasingly more important role in electronic computers. The most widespread opinion is that optics will be used together with electronics in a complementary manner, rather than replacing electronics. For this reason, some prefer to use the term optics in computing, as opposed to optical computing.

In order for information to be processed inside a computer, it must be represented as a certain physical entity. For instance, the number 5 may be represented by the current through a wire or the charge on a capacitor, the rate of flow of water through a pipe or the amount of water in a tank, the speed of rotation or position of something, the intensity of mechanical or acoustical waves, or the amplitude or intensity of a beam of light. Information represented in this manner can be transmitted from one place to another inside a computer, meet other pieces of information and be subjected to the operations serving our purposes. What makes a computer electronic, hydrolic, mechanical, acoustical, or optical is how the information is represented and processed.

Researchers have dreamed of building computers which are based on optical principles for a long time. However, for various reasons, successes in this area have been limited to special-purpose applications, and the general-purpose optical computer has been evasive. Despite this, the successes in special-purpose applications as well as successes in the communication, storage, and sensing of information, the as yet unrealized but huge potential of hybrid optical-electronic computers, and the lingering dream of the general-purpose optical computer ensures optical information processing to remain an active area of research.

2. ANALOG OPTICAL INFORMATION PROCESSING

In analog optical information processing, information is usually represented with the amplitude or intensity of light, and less frequently, with parameters such as polarization or wavelength. To add two numbers optically, we may utilize two variable-amplitude light sources and one light meter. We simply set the sources to have an amplitude proportional to the numbers to be added and shine these on top of each other on the light meter, and we can read the result of the summation from the meter.

To multiply two numbers optically, the number to be multiplied may be represented by the amplitude of a beam, and the multiplier may be represented with the transmittance of a semi-transparent material, e.g. if the multiplier is 0, the material is opaque, if the multiplier is $1/2$, the material passes half the light. When the light beam representing the number to be multiplied passes through this material, we can read the product from a light meter. While this method restricts the multiplier between 0 and 1, this difficulty can be overcome easily through normalization. It follows that any problem which can be expressed in the form of consecutive additions and multiplications can be realized with optical systems. Moreover, beyond adding and multiplying individual numbers, very high resolution images and large matrices can be operated on at once.

The key result enabling most of analog optical information processing is the ease with which the Fourier transform can be realized by using a simple convex lens. Since the Fourier transform allows us to realize filtering operations by mere multiplication, the ease, low cost, and speed with which it can be optically realized paves the way for many applications, especially the high-speed filtering of large resolution images for purposes such as pattern recognition, recovery, restoration, and so forth.

The major weakness of this approach is that, as in all analog systems, when there are a large number of consecutive operations to be performed, the errors arising from noise accumulate in an uncontrolled manner.

To realize the Fourier transform optically, we position a lens of certain focal length f to the right of the image to be transformed, at a distance f. The transform is observed at a distance f to the right of the lens. This result is a direct consequence of the law of propagation of light and passage through lenses. What is noteworthy is that even very high resolution transforms are realized in the time it takes light to travel from one end of the system to the other. A typical filtering scheme consists of a filter mask realized in the form of a semi-transparent material, sandwiched between a Fourier transform and an inverse Fourier transform stage. The mask for a low-pass filter can be realized simply by cutting a hole in a piece of cardboard.

That such operations could be realized with optical systems was known for a long time. While we can witness examples of operations based on the Fourier transform in the fifties, the lack of quality light sources exhibiting the desired properties was an important obstacle. The lasers which became available in the sixties filled this need. This period also witnessed the application to optical systems of communications and signal theory concepts that were developed in electrical engineering. This development, together with the availability of lasers, resulted in fast progress that led to a significant cumulation of knowledge in the decades that followed.

Increasingly complex optical processing systems were developed during the sixties. Some of these were used to solve large problems for which the very low-capacity digital computers of the time were no where nearly capable of solving. Approaches based on using acousto-optical devices to convert electrical signals to optical ones and thus process them also date from this period. The sixties also witnessed the beginning of pattern recognition applications based on holographic matched filters.

The seventies not only witnessed the continuation of research in these areas, but also the development of matrix processing systems. These systems set an example for certain future digital optical systems, by the way they viewed the spatial coordinates as discrete variables. Iterative algorithms for phase retrieval from intensity information and other problems in image recovery and reconstruction received attention in this period with research in this area continuing actively well into the eighties.

Relatively speaking, the eighties was a period of stagnation for analog optical information processing. Despite successes with special-purpose systems, the lack of major commercial successes had led to a certain degree of disappointment in both researchers and their funders. During this period, most researchers turned their research efforts to digital optical computing and optical interconnections. Following this period, the subject receiving the greatest interest in the area of optical information processing was the subject of fractional Fourier transforms, which became popular throughout and beyond the nineties.

The fractional Fourier transform (FRT) is a generalization of the ordinary Fourier transform. The zeroth transform of a function is the function itself, the first transform of a function is the ordinary Fourier transform. The 0.5th transform is that operation, which when repeated, results in the ordinary Fourier transform; likewise the 0.25th transform results in the ordinary transform after four applications. The FRT can also be realized with a single lens, just like the ordinary Fourier transform. For this reason, many systems which find widespread use in analog optical information processing can be generalized to the fractional case. For instance, we can speak of fractional filtering and fractional convolution. These generalizations have introduced considerable flexibility to analog optical information processing.

Before closing this section, a few words about holography, which is an important part of analog optical information processing, will be appropriate. Holography is widely thought to be synonymous with three-dimensional photography. We perceive objects in daily life as a result of the light waves reflecting off of objects and reaching our eyes. Holography is the recording and reconstruction of these light waves. Thus, even in the absence of the object whose hologram was recorded, the same light waves reach our eyes and are perceived as the original object. Since these waves are ideally identical to those reflected off the original object, depth information which enables three-dimensional perception is also preserved. On the other hand, in ordinary photography, only the amplitude of the wave is recorded and the phase is lost, resulting in loss of information. Beyond recording of three-dimensional information, holography has important applications in high-density information storage and the routing of light beams in complex patterns. Three-dimensional television, an old objective of holography, may be a reality within the next ten or twenty years.

3. DIGITAL OPTICAL COMPUTING

In current usage, a computer usually refers to the now ubiquitous digital computer. The idea of building digital computers based on optical principles is not new. Originally it used to be thought that the defining characteristic of an optical computer would be its reliance on optical rather than electronic transistors and gates. Indeed, it is possible to manufacture in the form of large arrays, nonlinear optical devices that perform similar functions to electronic transistors and gates. Very large efforts have been put into the development of such devices. While the inputs and outputs of an electrical transistor or logic gate are electrical currents or voltages, the inputs and outputs of optical transistors and logic gates are light beams. Information is represented with the existence or non-existence of light beams rather than currents and the desired logic operations can be realized based on the physical principles governing the interaction of light with matter.

In principle, it is possible to realize optical computers based on a similar architecture to electronic computers, by bringing together a large number of optical logic gates in the form of complex optical circuits. However, it is not expected that such an approach would result in a very efficient result. It should not be forgotten that electronic computer architecture and design has evolved around the strengths and weaknesses of electronics technology. Maybe the most important example of this is the Von Neumann architecture. For reasons both physical and technological, the strengths and weaknesses of optics and electronics are very different. Therefore, optical computers must be designed according to an architecture matched to the strengths and weaknesses of optics.

Several optical computer architectures have been proposed with this understanding. While some of these are still based on logic gates, alternative approaches have also been considered. Approaches based on the shadowing of optical patterns, or approaches based on algebraic operations or the substitution of symbols according to predetermined rules are only a few of these. Physically, optical computers may consist of two-dimensional optical circuits, much like integrated electronics circuits, or may consist of discrete optical devices connected to each other with optical fibers. But the approach which seems most promising and which has received the most attention is that in which light beams travel in free space. Typically, in such systems the active devices (transistors or logic gates) are arrayed on a plane which has been manufactured in a similar manner to electronic integrated circuits (the integration of thousands of optical devices is possible). The inputs of these devices are on one side of the plane and their outputs are on the other side. On and off light beams representing the binary values of one and zero fall on the logic gate where they will be processed, and the result emanates from the other side. If we imagine all of the devices at once, at the input side of the device plane, a large number of beams are falling onto the logic gates and light beams representing the results of the logic operations are being emanated from the output side of the plane and then manipulated with mirrors, prisms and microlenses to be redirected to the input side of the active devices place in accordance with the circuit diagram of the computer. In contrast to electronic computers, the interconnections are realized without solid wires or cables; the information carried by light travels in free space and the connections are realized in the desired pattern by the help of optical elements such as mirrors, prisms, and lenses. Since there are no wires and the light beams freely pass through each other—in contrast to electronic circuits which are confined to two-dimensional planes—it becomes possible to realize three-dimensional complex circuits. Such optical structures exhibit qualities of parallel architectures and are thus especially suited to the implementation of parallel algorithms.

The major problem with optical transistors, logic gates, switches, or other nonlinear devices serving similar purposes, is their high energy consumption. There are limits to what can be done with linear systems; many problems we may wish to solve will require nonlinear devices such as transistors or logic gates. The underlying reason for the high energy consumption of optical devices is usually seen to be a consequence of the different nature of photons (which are bosons) and electrons (which are fermions). While there are strong interactions between electrons, there are no such interactions between photons. In practice photons can interact only through the mediation of matter.

Certain studies dating back about twenty to thirty years claimed that optical computers can never be a reality due to the large amount of heat that would be generated by the high energy consumption of optical devices. These claims had considerable impact. During the eighties and nineties, many researchers abandoned work on optical computers in favor of optical interconnections, which we will discuss below. Indeed, in the near future it is much more likely to see computers using optical interconnections than all-optical computers. Nevertheless, the dream of optical computing is far from dead. For one thing, although the architecture making the most of the strengths of optics is still subject to debate, there exists a considerable cumulation of knowledge regarding different alternatives. Also, the claims reducing interest in optical computing are now interpreted in a different light. Not only is it possible to make lower energy optical devices, it is now understood that the total energy consumption of computers are dominated not by the devices but by the interconnections, which totally changes, and in fact reverses the picture since optical interconnections can consume less energy than electrical ones. And finally, it is now understood that an optical transistor or logic gate is not different than an electrical transistor or logic gate with a light detector at its input and a light emitter at its output (the multifunctional and highly efficient SEED device is an example of this). Therefore, if it makes sense to make the more critical interconnections optical, there is little reason not to make the transistors or logic gates optical as well.

The first optical computers exhibiting the defining qualities of a digital computer were demonstrated in the laboratory in the early nineties. About the same time, switching systems for communications purposes were successfully demonstrated. It may not be very likely for general-purpose optical digital computers to gain widespread use. However, there is great interest in optical switching systems due to the fact that communications infrastructures which are increasingly dominated by optics are bottlenecked by electronic switching stages and thus working significantly under their potential. For this reason, optical switching systems integrated with the fiber optics infrastructure may well be the first examples of optical computers.

4. OPTICAL INTERCONNECTIONS AND OPTOELECTRONIC COMPUTERS

The solution of complex problems requires a large number of arithmetic-logic operations to be simultaneously performed and the results of these to be appropriately combined. For this reason, at least as important as how information is represented and operated on, is how information is percolated to and fro within the components of a complex computing system. Communication within computing systems, a subject attached less importance before the eighties, has become increasingly important as the interconnections (wires, cables) connecting the components of large computers to each other have become the key factor limiting the power and speed of such systems. This is a consequence of the electrical resistance and other undesirable qualities of materials such as copper and aluminum widely used for wires. Interconnections now consume more space and energy and contribute greater amounts of delay than the devices.

We have already mentioned that optics and electronics have different strengths and weaknesses. Electrons, unlike photons, are able to interact strongly so that electronic nonlinear devices can consume less energy than optical ones. They can also be smaller. On the other hand, electrical resistance and capacitance effects negatively effect the performance of electrical interconnections and slow them down. Electrical interconnections also consume more space and energy as system size increases, compared to optical interconnections. They also exhibit crosstalk and due to the possibility of short-circuiting, routing of electrical interconnections are more difficult. On the other hand, with optical interconnections it is possible to route three-dimensional complex circuits and parallel structures. (We may also note that use of superconducting wires eliminates only some of the negative qualities of electrical interconnections.) For these and similar reasons, during the first half of the eighties, many researchers started advocating the concept of the optically interconnected electronic computer where the nonlinear operations are realized by electronic transistors or logic gates, and the interconnections among these are realized optically. Such systems, also referred to as optoelectronic computers, are thought to allow optics and electronics to complement each others strengths and make up for each others weaknesses.

As already mentioned, some researchers have bended these arguments in favor of all optical computing: If the interconnections are dominant, having made the interconnections optical, there is little reason not to make the logic gates optical as well. However, this is not a widely accepted position. The more commonly asked question is, beyond what stage of the interconnection hierarchy should electrical interconnections give way to optical interconnections? Should connections between transistors be made optically, or should these be made electrically but inter-chip connections be made optically, or should optics be used only among larger subunits. While different modeling studies give different results, it can be said that there is data supporting the use of optics down to at least one centimeter.

Let us give a concrete example of how optical interconnections may be used inside a computer. In an electronic computer, large numbers of integrated circuits are connected to each other on a printed circuit board. Instead of the printed circuit board, we may realize optical connections among these if the circuits have light emitters and detectors situated on them. Much work has been done on how to situate optical emitters and detectors on or besides electronic circuits. The light leaving the emitters is directed to the detector at the target destination with the help of mirrors, prisms, and microlenses, and thus the desired connections can be realized. The number of connections that can thus be realized can not only be much larger than possible with a printed circuit board, but also their speed will be greater and their energy consumption lower.

5. ARTIFICIAL NEURAL NETWORKS, NANOCOMPUTERS, AND OTHER NEW APPROACHES

The use of optics in information processing is not limited to those approaches described above. Artificial neural networks are systems inspired by animal brains. The most important characteristic of neural networks is the simultaneous realization of many operations at once and the large number and complex network of interconnections among the units. We have seen that optics is very suitable for realizing such connections. Therefore, optics may play an important role in the realization of such computers.

It is possible that we will witness significant changes in the structure of computers within the coming decades. Much research is taking place on atomic-scale information processing systems based on quantum and biological effects and the individual or collective behavior of atoms. Such systems may exhibit features totally beyond the computer paradigm based on transistors (or other nonlinear devices) connected with wires (or optical connections). Nevertheless, given that electromagnetic waves play an important role in the transmission of information from one place to another and that in order to provide a certain information density at least optical frequencies would probably be used, we may presume that optics may have a role to play in such computers as well.

6. CONCLUSION

In the medium run, there is a reasonably high probability that we will witness the use of light for communication between the components of especially high-performance computing systems. This will enable computers faster and more powerful than possible today. As a consequence, optics will play an important role in the processing of information as well as its communication, storage, and sensing, and thus have an even more prominent place in information technology. Whether a general-purpose optical digital computer will ever become a reality is a more uncertain issue.

This work was originally prepared for [1]. The author acknowledges partial support of the Turkish Academy of Sciences.

REFERENCES

[1] Haldun M. Özaktaş. Optik bilgi işleme. In *Türkiye Bilişim Ansiklopedisi*, Türkiye Bilişim Vakfı, to appear. In Turkish.

[2] M. A. G. Abushagur and H. J. Caulfield, editors. *Selected Papers on Fourier Optics (SPIE Milestone Series 105)*. SPIE Optical Engineering Press, Bellingham, Washington, 1995.

[3] H. H. Arsenault and Y. Sheng. *An Introduction to Optics in Computers (Tutorial Texts in Optical Engineering 8)*. SPIE Optical Engineering Press, Bellingham, Washington, 1992.

[4] D. G. Feitelson. *Optical Computing*. The MIT Press, Cambridge, Massachusetts, 1988.

[5] J. W. Goodman. Four decades of optical information processing. *Optics and Photonics News*, pages 11–15, February 1991.

[6] J. W. Goodman. *Introduction to Fourier Optics*, second edition. McGraw-Hill, New York, 1996.

[7] H. S. Hinton. *An Introduction to Photonic Switching Fabrics*. Plenum Press, New York, 1993.

[8] J. Horner, editor. *Optical Signal Processing*. Academic Press, San Diego, 1987.

[9] K. Iizuka. *Engineering Optics*, second edition. Springer, Berlin, 1987.

[10] J. Jahns and S. H. Lee, editors. *Optical Computing Hardware*. Academic Press, San Diego, 1994.

[11] M. A. Karim and A. A. S. Awwal. *Optical Computing: An Introduction*. Wiley, New York, 1992.

[12] A. W. Lohmann. *Optical Information Processing (lecture notes)*. Optik+Info, Post Office Box 51, Uttenreuth, Germany, 1986.

[13] A. D. McAulay. *Optical Computer Architectures*. Wiley, New York, 1991.

[14] D. A. B. Miller. Computing with light. In *1995 Yearbook of Science and the Future*, pages 134–147, Encyclopedia Britannica, Chicago, 1994.

[15] H. M. Ozaktas. Toward an optimal foundation architecture for optoelectronic computing. Part I: Regularly interconnected device planes. Part II: Physical construction and application platforms, *Applied Optics*, 36:5682–5705, 1997.

[16] H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay. *The Fractional Fourier Transform with Applications in Optics and Signal Processing*. Wiley, New York, 2001.

[17] A. Papoulis. *Systems and Transformations with Applications in Optics*. McGraw-Hill, New York, 1968.

[18] B. E. A. Saleh and M.C. Teich. *Fundamentals of Photonics*. Wiley, New York, 1991.

[19] A. VanderLugt. *Optical Signal Processing*. Wiley, New York, 1992.

[20] F. T. S. Yu. *Optical Information Processing*. Wiley, New York, 1983.

[21] F. T. S. Yu and S. Jutamulia. *Optical Signal Processing, Computing, and Neural Networks*. Wiley, New York, 1992.

[22] F. T. S. Yu and S. Jutamulia, editors. *Optical Pattern Recognition*. Cambridge University Press, Cambridge. 1998.

[23] Special issues on Optical Information Processing, Optical Computing, Optics in Computing, and Optical Interconnections of the journals *Applied Optics*, *Optical Engineering*, and *Proceedings of the IEEE*.