Photonics and AI: a conversation with Professor Demetri Psaltis

Guohai Situ*
Chinese Academy of Sciences, Shanghai Institute of Optics and Fine Mechanics, Shanghai, China

Demetri Psaltis, Professor Emeritus at École Polytechnique Fédérale de Lausanne (EPFL). Prof. Psaltis sat for an *Advanced Photonics* interview with Prof. Guohai Situ. View the video at https://doi.org/10.1117/1.AP.6.5.050501

Situ: Hi, Demetri, thank you for agreeing to do this interview with us. It's been actually a few years since we met last time. I think it was at SPIE Photonics West in San Francisco if my memory is correct. I was with Lei Tian, and we had some discussion with you in a cafeteria. It's always an inspiration to talk with you. How's the life at EPFL those days?
Psaltis: Everything is good. In Switzerland, things don't change very quickly. It's very stable. It has been very cold for some reason this year and for this late in the season.

Situ: I know that before moving back to EPFL in 2007, you'd been a professor at Caltech since 1980. During your four-decade-long career, you have made ground-breaking contributions to the development of many fields in optics, like optical computing, holographic storage, opto-fluidics, optical imaging, and so on. Reflecting on your career, can you share some of the pivotal moments that have defined your works in these fields?
Psaltis: Like you say, I have worked broadly. I haven't focused on one thing throughout my career. I would say the common thread is the combination of information and optics. Somehow this has been a persistent thread throughout my career. I got into this field at Carnegie Mellon University where I did my PhD (also my master's degree). To decide what to do, I visited the laboratory of David Casasent who later became my PhD professor.

In his lab, I saw on the wall a red sinc function, $f(x) = \sin(x)/x$, which I recognized from communication courses. I was very impressed that the physical system could produce a sinc function. That's what caught my attention and introduced me to the link between optics and information. I was very young, 22 years old, seeing a sine function on the wall impressed me. To this day I am pursuing things like that.

Situ: Me too, I also recollect my first time using the spatial light modulator to produce something you can derive in mathematics. It's very beautiful at that time.
Psaltis: At that time, we didn't have spatial light modulators, just film. So, we had to make hologram we recorded on a photographic plate. We didn't have digital holography. It was exciting though because you record your hologram! You go in the darkroom to develop the plate and only then you will see the three-dimensional image. When cameras became available for digital holography, that made a lot of things possible.

Back to your questions about pivotal moments. When I moved to Caltech, Prof. Nabil Farhat, from the University of Pennsylvania, a very nice man, was visiting me at Caltech, doing a sabbatical with me. He said, "There's an interesting talk, would you like to come and listen to it with me?" I agreed though at that time I was very busy. The talk was by John Hopfield. He was talking about brains and the Hopfield model as it would become known later. John Hopfield was talking a model for the brain that we could understand. Nabil and I were very excited because at least I didn't know anything about the brain. We thought maybe we can implement this in optics. In fact we did implement it in optics. We spent many years working on how optics can be used to implement neural networks. So it was lucky because Nabil passed by my office and said: "Let's go and listen to the talk."

Situ: Yeah, somehow, that's a kind of very accidental event that makes a change to the career path of many people. I think this has also happened in our careers as optics researchers. As you mentioned just now, optical computing and optical neutral networks are very hot topics in optical information right now. You mentioned the Hopfield model, for which I think you were the first scientist to do the optical implementation. Is this the reason why you got into the field of optical computing back then, or just because you attended the lecture and met Hopfield?
Psaltis: Well, two reasons. One is the serendipity, like you said, because Nabil and I went and were very excited learning about the brain, which was something we didn't know, and we did the optical implementation. So this was one aspect.

The other aspect was that, at that time, there was a lot of work on digital optical computing. Joseph W. Goodman and others had published their paper on optical interconnects. For myself, I was more in the analog implementation. When I looked at the neural network models, I thought there was a good match with optics. A neuron in the brain is connected to more than a thousand other neurons, whereas in digital Boolean logic, each gate is connected to maybe two or three others. Still there are many wires in a silicon chip just to connect all the devices. But when you compare to a neural network, you have many more connections in the neural network.

In our paper we pointed out that neural networks are well matched as a computational architecture for optics, because there are so many connections. And now it is understood that optics has the advantage for making connections between units, energetically and otherwise (see Fig. 1).

*Address all correspondence to Guohai Situ, ghsitu@siom.ac.cn

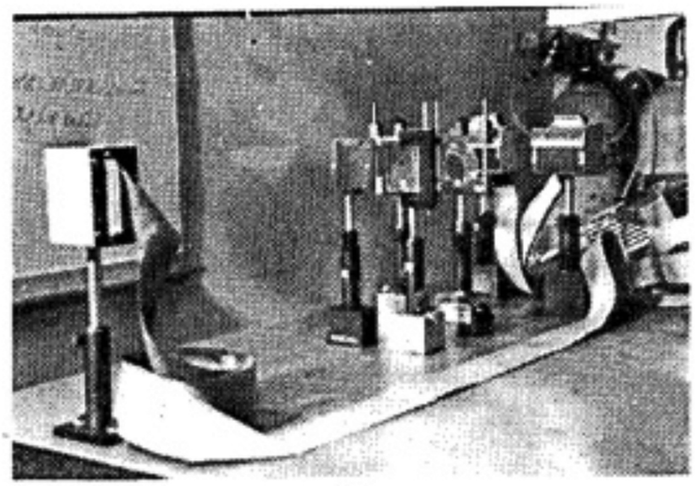

Fig. 1 Photo of the first optical neural network, built at Caltech in 1986. Reprinted with permission from N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," *Appl. Opt.* 24, 1469–1475 (1985).

Situ: I will come back to this analog computing and digital computing topic later on. That's a quite important question even for us at the moment. Let's go back to my questions. The field of optical computing went through a quiet period from the 1990s to the middle of 2010s. When looking back in history and comparing to what we have and where we are now, I understand that the techniques were not ready at all in those days. In your view, what were the main reasons behind that?
Psaltis: Caltech is operating a lab for NASA called Jet Propulsion Laboratory (JPL). I went up to visit and discovered that they had a wonderful optical computer. They used it to form synthetic aperture radar images, the only way to do it then.

In one of the occasions, the first image we did was my hometown in Greece, Thessaloniki. These optical computers were doing things that you could not do otherwise because digital signal processing had not been developed yet. Gradually, digital signal processing, came along, and it became very difficult for optics to compete. That's why many others and I moved on. What persisted was optical interconnects, which remains an important technology used in fiber networks for data centers. Some all-optical switching became very popular as well for a while. Now, all optical switching has come back. Google introduces the Apollo system where the switching is done all optically with MEMS. That's a form of optical information processing if not optical computing. But even more recently, optical neural networks came back in a big way.

The increase in the size of neural networks became exponential with time. As scientists, we know that anything that grows exponentially is going to have some limits. The problem in this case, is too much energy consumed and high cost. Therefore, we are looking for alternatives to the traditional digital techniques for doing neural computation. Optics came back because of the connectivity, the same argument as 40 years ago. With optics we can do the connectivity or the linear part of the neural network almost with very little energy.

Situ: But to improve an optical neural network, I think there are many ways to do that right now. In the past few years, what would you say has been the most transformative development in optical computing?
Psaltis: I believe we have two threads to pursue, both of which are important and very interesting. One is diffractive optics, in three-dimensional optical systems. You can use diffraction for connectivity and find a way to implement the nonlinearity (see Fig. 2). The other approach is integrated optic structures, either resonant structures or Mach–Zehnder interferometers, which can be programmed on a chip to do a certain task.

Both approaches are being pursued. There is a certain comfort for people to work with chips, because of the analogy with integrated electronic circuits. There's a question about how much integrated optics can be scaled (how many units we can put on a single chip). However, integrated optics have a huge advantage in terms of being able to use the bandwidth that became available to optics through fiber optics communications technology. Nowadays, you can modulate light at many Gbit/s, so you can build a computing structure at that rate.

These two approaches are being pursued and it's quite possible they will merge. Actually, I was just reading a paper from Prof. Qionghai Dai. It is very nice work. They use an integrated platform and diffractive optics.

Situ: I also know some people using some complex media to do this optical computing, so I think that is another way.
Psaltis: Yeah, like Sylvain Gigan and others who use multimode fibers. The idea is based on something called extreme learning machines, which was proposed by Professor Huang in Taiwan. He had the idea to multiply the data with a fixed random matrix followed by a simple trainable one-layer network. You can get a remarkably good performance. The technique is well suited for optics because you can do the linear map and then use a very simple digital processor. This technique doesn't give performance that's competitive with the best all digital networks, but it's very fast to train and uses very few digital weights.

Situ: So, how do you evaluate all these different kinds of ways to do optical computing right now?
Psaltis: I think a lot of people are interested and are pursuing optical computing all over the world, but very few people are using nonlinear optics. A neural network is inherently a nonlinear system because its function changes in response to what it experiences. The interconnections are very favorable in optics, one photon per weight, that's nothing, the optical connections are practically free. But the opposite is true for

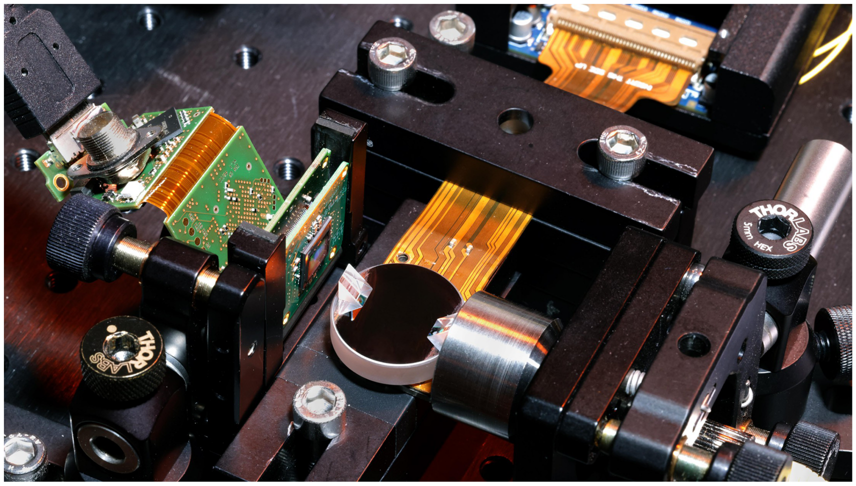

Fig. 2 Photo of optical compute developed in the EPFL lab, based on structural nonlinearity.

the nonlinearity. So, most of us rely on OEO (optical-to-electronic-to-optical conversion). You detect at some point, convert to the electronic domain, you perform some nonlinearity there and then you regenerate the light.

When I was young, we tried nonlinear optics. We used photorefractive crystals, nonlinear crystals, to implement the weights and different technologies, mostly based on gallium arsenide, for the nonlinearity, of the activation functions. It's very interesting, but it's challenging. Right now, most neural networks rely on OEO. Thankfully, modulation technologies have come along that are remarkable.

For example, Dirk Englund at MIT or Marko Lončar at Harvard, and here in Switzerland, Rachel Grange, make optical modulators very thin which can be modulated with low applied voltages. These devices can be integrated with silicon, and combined with detectors on silicon as well. So, you have an OEO interface that has very high bandwidth and low energy consumption because of the low voltage. I think this is the hope: combine linear optical interconnections, either with MZIs or resonators on silicon, or in free space with diffractive layers. And then combine that with OEO for the learning and the nonlinearity in the activation functions to be done in the electronic domain.

Situ: Right, but in this way, you did also need the optical and electric switching, and this kind of switching is quite slow, right? Because also in nonlinear optics, you need some kind of materials to respond to the field or to the voltage, and that also takes some time. So, if you put all these into an optical neural network, will that slow down the training process?

Psaltis: Well, like I said, I don't think nonlinear optics now is a good solution. We rely on OEO, optical to electronic, back to optical. Nonlinearities are implemented somehow either for learning or for decisions within the network, in the electronic domain, and then they transfer back to the optical. With nonlinear optics, either we have to slow down or use very high peak power, you have to attempt femtosecond laser, which means you lose the power advantage. There are promising technologies like phase change materials. But still, the training itself is done digitally and then transferred to the optical domain.

Situ: That means, at least now, what we have for optical computing, or optical neural network, still relies on digital or electric training. So, after all the training is done and all the parameters have been set, then we can transfer it into the optical domain, which is good for inference.

Psaltis: There are two approaches. One is that you train digitally completely. Use a digital twin to simulate what the optical system does and then transfer the information to the optical system. The other approach, which I think is also very exciting, is to use the optics-in-the-loop. In other words, you do the training in the digital system, but you do the inference, or you do the forward path in the optical domain. If the optical part makes some mistakes, the learning algorithm may correct those. Also, if there are drifts—because it's really an analog system—so if the system drifts or changes over time, then you can periodically or continuously retrain the system.

Situ: Well, that's good. And Demetri, back to the 1990s: You served as the executive officer of the computational neural system of Caltech, and later on you also acted as the director of the NSF Center for Neuromorphic Systems Engineering. I believe that you've seen optical computing evolve, not just as a scientist who works in this field, but also as an administrator who has the broader and further vision of how it develops. So, from this perspective, can you share with us some of your vision of how and why it evolved to the current state?

Psaltis: At Caltech as well as EPFL where I am now, I was involved almost continuously in some sort of administrative responsibility. I enjoyed that. I was in charge of the computation and neural system department. It was a department that was started together by Hopfield. He was the first department chairman. Hopfield had the idea that we should focus the neural networks activity at Caltech on education. The students would be interested and also it is a way to bridge biology, engineering, and computer science. I think it was a very good idea because it's easy for professors to come together and design an educational curriculum.

There were many good people involved: Carver Mead, Abu Mostafa, and Christof Koch. It was a very exciting time. People like Geoffrey Hinton and David Rumelhart would visit. I was young, so it was a great experience for me to have met all these people.

Situ: But you are very famous now, right?

Psaltis: Well, not like these guys, thank you. It was a good time, because it was the beginning. And I can say the following. Even the wildest ideas we had then are nothing compared to what actually happened in machine learning. I mean, the things that people do now with machine learning, like ChatGPT.

One difficulty I had as an administrator was the coordination of the biologists with the people from physical sciences. Initially, the

Computation of Neural Systems Department at Caltech was half-half. Computational scientists like myself learned about the brain and we were motivated. But at that time, there was a very strong resistance from the biologists to consider computational models to help them understand how the brain works. Gradually, engineers became less interested in understanding the details of how the brain works perhaps because we don't have a full model for how the brain works. Lacking that, engineers developed their own techniques, like error back propagation, more recently transformers and so forth that are not directly related to the brain. Researchers are starting to say, these computational models that we use now are not biologically plausible. So, they are looking for alternatives where the learning rules are local. Hinton calls it forward-forward, Bengio calls it equilibrium propagation. In biology, there's no evidence that there's feedback all the way back. Somehow the learning rules must be local.

Situ: I think the theory of neural computing is also evolving, right?
Psaltis: Definitely, every day there is something new. The open AI models like ChatGPT or all the generative models. I teach a class now in computational imaging. It's hard to keep track of all the new ideas that are coming out. Most of the new stuff is coming out of industry, not the universities.

Situ: Yeah, this also is the question, which is related to my following questions. The first one is what can the roles of optics be in this way? All these achievements, like ChatGPT and so on, are basically based on the electronic computing, the digital computing, but optics, so far, basically can only perform analog computing. So, I think, back to the history, many people tried to do some digital optical computation and optical logic cases. But still, do you know if there is any way to overcome these difficulties?
Psaltis: Digital computers of the type we use every day are based on Boolean logic. So you have very sharp nonlinearities, zeros and ones, there is little room for error. The brain doesn't work like that. Yet it makes decisions, it makes nonlinear transformation, does computations that can be very intricate and powerful.

So, again, the lessons from biology are many. One is connectivity. But the other is that you don't necessarily need extreme accuracy to do something well. The two are related because if a neuron receives input from a thousand or tens of thousands of other neurons, it has some room for error. If one of the inputs is off by a few percent, what the neuron responds to is the combination of all its inputs. So, it has error correcting capabilities built-in due to the connectivity. The same thing applies to analog VLSI, not just analog optics. Carver Mead, pursued analog VLSI beautifully.

Situ: Optics can accelerate AI, as we have discussed a lot. AI is also a transforming computation imaging, and we are also involved in this development quite a bit. So, how have you integrated AI into your work?
Psaltis: That's a very good point. We have optics for AI, which means building machines to compute AI related tasks. And vice versa. AI for optics is using new tools to help us build optical systems that are better. So, I've been involved in that as well, as a lot of other people. For example, I've been working on imagining 3D objects using tomography. You take a 3D object like a cell, and you try to reconstruct its 3D shape from its 2D projections.

That is inherently an ill-posed problem because you're trying to reconstruct the volume, a 3D object from two dimensions. When you try to do this reconstruction, you're missing some information, so you have to make guesses. You can make an educated guess by using a neural network and some data. We have used the neural network to do reconstructions of 3D shapes from 2D projections, using a technique we call learning tomography. It was published recently in *Nature Photonics*.

We are starting to think that the 3D images of cells, it may be a replacement for the genetic code, for the RNA or the DNA. Because for a cell, just like any living organism, the way it looks comes from the DNA. It was created by instructions from DNA, and if there is a problem like cancer or some other disease that changes its shape, then we can use machine learning to relate that to the genetic code.

Situ: Yeah, we just talk about the large language models like ChatGPT, which has shown to be remarkable in many aspects. So, what and how do you think that we, as optical researchers, can benefit from them?
Psaltis: I think ChatGPT is quite a challenge for all of us to adapt and adjust. For research, we have to see how it can be used in a productive way rather than in a way that diminishes creativity. Because if you work only with GPT, then it is guaranteed that you will be average because everybody can use GPT.

Situ: Actually, we can argue in fact to develop an algorithm for anything like reconstruction or design anything, but I don't know if they can create something new. So, I had some kind of discussion with Wolfgang Osten recently about whether AI is actually intelligent, and we have some kind of agreement that the current state is not exactly intelligence. So, they just let people know a trend, they're determined to know. So, I think to do some creative jobs, I don't think something like ChatGPT can actually do that.
Psaltis: Some time ago, I visited Japan. I've been many times since then. But the first time I visited, I was talking to a professor about innovation, and he told me that, "in Japan, we don't say we innovate instead we say we discover." So, in some sense, when we say we're innovating, we're really discovering what is already there.

Situ: Okay, this is one way. I know some people in the industry, they can use ChatGPT to speed up the flow of work, I think this is also one way for people in academy to do that.
Psaltis: Yeah, I mean, you have to be careful, but there is a risk of that we all become lazy. And then we are content with mediocrity or being like everyone else.

Situ: So, to train a large language model needs so many resources that a research institute or a university can offer. Actually, there is some concern that this may hinder the further development of AI. What do you think about this and what optics actually can do about it?
Psaltis: Well, your question has two elements. One is whether it is possible to do something at a university related to large-scale models. That's a concern, right? As we said before, most of the work, most of the new ideas now seem to flow out of the big companies. So, it's a concern because they are somehow different than scientific research as far as improving the good of the whole society. They are a business, so they have to maximize their profit, that is what a business does.

Whether optics can provide some new diversity in the AI market, it's hard to tell. Optics provides a new way for scale up. I think most people, including me, don't believe that it's possible to continue just using GPUs. The power consumption is the key problem and something has to be done differently. It could be analog VLSI, quantum computing or optics. So, the transition may help to disperse a little, the concentration of all the activities to the current players that dominated.

Situ: So, as you look ahead, what are the research areas at the intersection of photonics and AI that you are most excited about?
Psaltis: Well, of course, I have my own work and my own work is focusing a lot on structural linearity. It is a way of implementing

nonlinear functions using linear optics. It is promising because we can reduce optical-to-electronic-to-optical conversions. This is an exciting area that I think will keep me busy for the next few years. For optics in general to take a foothold, we need the companies that build accelerators to be integrated into data centers. If they get a foothold, it will be a big step forward for all of us. An adoption of optics will not only help the data centers, but also open the floodgates for optics research, both in terms of resources, funding and people: more smart people being attracted to the field.

Situ: Yeah, so one last question, Demetri: I know you're a great scientist and we admire you so much, you also have to supervised so many graduate students who have also made significant contributions in their own fields of study or in industries. And how can you be so successful? What's your philosophy when it comes to nurturing the next generation of students?

Psaltis: Okay, I'm glad you ask this question, because I consider my students my number one accomplishment professionally. I think I've been very lucky, of course. At Caltech, we used to get fantastic students. But I've also had great students here at EPFL. I think it's an amazing gift to be a professor and have the young people come and work with you, listening to you. I don't have a formula other than to say I try to treat each person as an individual. Some people respond to encouragement, some people respond to criticism, some people need to be hand-held. So, different people respond differently. I learn a lot from the students and the interaction goes back and forth. The past students are all my friends now. So, it's a great gift of our profession. I always feel that I've never worked in my life. My job is my hobby. Only professional basketball or golf players can say that. But they have a short career.

Situ: All right, so I think I have all my questions, I would be finished, and I presently thank you so much for the interview with us. We also hope you to publish your next paper in *Advanced Photonics*, which is increasing its reputation also. Thank you for your time, Demetri. I hope I can see you in the near future in person.

Psaltis: Thank you very much.

Demetri Psaltis received his degrees from Carnegie-Mellon University, Pittsburgh, Pennsylvania, USA. In 1980, he joined the faculty at the California Institute of Technology, Pasadena, California, USA. He is a professor emeritus and director with the Optics Laboratory at EPFL in Lausanne, Switzerland. He moved to EPFL in 2006. His research interests include imaging, holography, biophotonics, nonlinear optics, and optofluidics. He has authored or coauthored more than 400 publications in these areas. He is a fellow of Optica, the European Optical Society, IEEE, and SPIE. He received the International Commission of Optics Prize, the Humboldt Award, the Leith Medal, the Gabor Prize, and the Joseph Fraunhofer Award/Robert M. Burley Prize.

Guohai Situ is a professor with the Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences. His research interests encompass a broad spectrum of computational optical imaging. Currently, his focus is on imaging through optically thick scattering media, computational ghost imaging, and phase imaging. To date, he has published 100 papers in leading journals. He is a fellow of Optica and a recipient of the National Science Fund for Distinguished Young Scholars (NSFC). He also serves on the editorial boards of *Advanced Photonics* and *Advanced Photonics Nexus*.