High-capacity photorefractive neural network implementing a Kohonen topological map

Yann Frauel, Gilles Pauliat, André Villing, Gérald Roosen

High-capacity photorefractive neural network implementing a Kohonen topological map

Yann Frauel, Gilles Pauliat, André Villing, and Gérald Roosen

We designed and built a high-capacity neural network based on volume holographic interconnections in a photorefractive crystal. We used this system to implement a Kohonen topological map. We describe and justify our optical setup and present some experimental results of self-organization in the learning database. © 2001 Optical Society of America

OCIS codes: 070.460, 070.5010, 160.5320, 210.2860, 200.4700.

1. Introduction

Neural networks have proved to be good alternatives to traditional computers for solving certain problems such as those of classification and pattern recognition. They are particularly well adapted to processing complex and noisy data. Basically, they constitute a large number of interconnected elementary processors that work in parallel. Therefore the use of free-space optics may be a good way to implement such systems because of the high connectivity and massive parallelism of such a system. Several optical neural networks have already been built,[1–16] but few of them have actually achieved large capacities.[13–16] So our aim was to demonstrate the feasibility of such a system and to find the problems that are caused by high capacities.

Among the possible optical techniques, volume holographic interconnects are especially promising. Indeed, holographic memories with 10^9 data—that is, elementary holograms—have already been demonstrated.[17] If these holograms are read out every millisecond, we get the equivalent of 10^{12} operations/s. In particular, photorefractive crystals are an interesting choice for the recording material because they allow one to dynamically write, modify, or erase holograms.[18] This gives one the opportunity to change the neuron interconnections in real time. We previously built a simple photorefractive neural network that implemented a topological map.[9] This system was rather limited, but it allowed us to gain experience and thus to design a whole new setup with a greatly enhanced capacity. In this paper we describe the improved setup and the experimental results that we obtained with it.

2. Theoretical Background

A. Volume Holographic Interconnects

When N reference beams (with complex amplitudes R_i^W) interfere with M learning beams (with complex amplitudes L_k) in a volume holographic material, they write $M \times N$ elementary gratings whose index modulations δn_{ik} are functions of R_i^W, of L_k, and of the exposure time. Inasmuch as the amplitudes of the beams provide only $M + N$ degrees of freedom, it is not possible to record any set of δn_{ik} with only one exposure. To obtain a particular set of values, one must proceed to a series of several sequential exposures with different beam amplitudes. Taking advantage of the Bragg selectivity, or using fractal arrangements,[19] we set the reference beams such that each of them can read only the gratings that it previously recorded. When the learning beams are switched off and the holograms are read with the reference beams, M beams are diffracted in the direction of the learning beams. For transmissive thick holograms, the diffraction efficiencies vary as $\sin(\pi\delta n_{ik} d/\lambda)$, where d is the thickness of the holograms and λ is the wavelength of the light. Provided that $\delta n_{ik} \ll \lambda/d$, the diffraction efficiencies are sim-

The authors are with the Laboratoire Charles Fabry de l'Institut d'Optique, Unité Mixte de Recherche 8501 du Centre National de la Recherche Scientifique, Centre Scientifique, Bât. 503, B.P. 147, 91403 Orsay Cedex, France. G. Pauliat's e-mail address is gilles.pauliat@iota.u-psud.fr.

Received 13 October 2000; revised manuscript received 27 March 2001.

0003-6935/01/295162-08$15.00/0

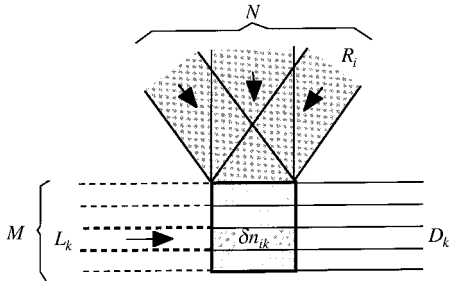

Fig. 1. Volume holographic interconnections.

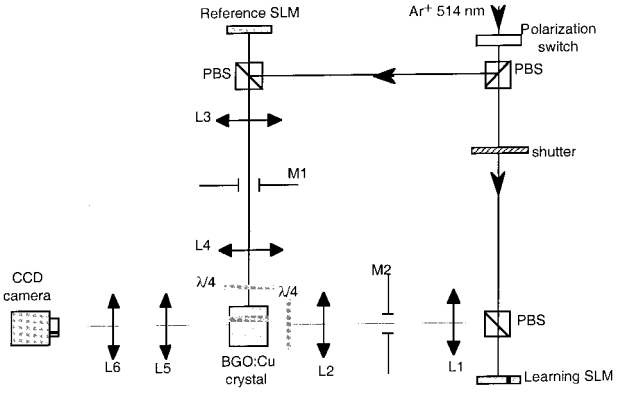

Fig. 2. Experimental setup.

ply proportional to δn_{ik}, and the diffracted amplitudes are thus

$$D_k \propto \sum_{i=1}^{N} \delta n_{ik} R_i, \quad 1 \le k \le M, \tag{1}$$

where R_i is the reference-beam amplitude during readout, which may differ from the amplitude during writing, R_i^W.

We define an input vector as $\mathbf{I} \propto (R_1, R_2 \ldots)$ and weight vectors as $\mathbf{W}_k \propto (\delta n_{1k}, \delta n_{2k} \ldots)$. Relation (1) can then be rewritten as

$$D_k \propto \mathbf{W}_k \cdot \mathbf{I}. \tag{2}$$

In other words, the amplitude of a diffracted beam is the inner product of the input and the weight vectors. This amplitude (or rather its square) can then be detected by a camera and sent to a computer, which can apply a threshold function. The result of such a computation corresponds to the usual output of a neuron with N weighted inputs.

We used the geometry depicted in Fig. 1. The N reference beams are angularly multiplexed, whereas the M learning beams are spatially multiplexed, i.e., set side by side. The M neurons are thus spatially separated. Although this geometry has a lower capacity than the more commonly used double-angular geometry, it has the advantage that a neuron can be updated with a learning beam without altering the state of other neurons too much. This property is of great importance for correct operation of the learning algorithm, as described below.

B. Topological Map

The unsupervised neural networks introduced by Kohonen[20] perform topological mapping of the vectors of a learning database. During a learning procedure, each vector of this learning database is presented to the system many times and at random. The network automatically classifies these input vectors according to their mutual correlations. Once this process is completed, each time an input vector is presented, a localized response appears in the output (a plane in our setup), and the locations of the responses reflect the correlations between the vectors. That is, if two vectors are similar, the corresponding responses must also appear similar.

In a holographic interconnect system as described above, when a set of reference beams (coding an input vector $\mathbf{I}$) is presented simultaneously with learning beam L_k, the holograms that correspond to neuron k are modified. In a photorefractive crystal, for instance, the interference of the beams results in an illumination pattern that causes a spatial redistribution of the carriers (electrons or holes) and hence a modification of the index modulations of the gratings. If exposure time Δt is short compared with the writing-time constant of the material, it can be shown[9,21] that the variation of the index modulations is such that the increment of the weight vector is

$$\Delta \mathbf{W}_k \propto \Delta t(\mathbf{I} - \mathbf{W}_k). \tag{3}$$

We took advantage of this law to implement a Kohonen neural network.

3. Experimental Setup

Our experimental setup is sketched in Fig. 2. A beam from an argon-ion laser at 515 nm is expanded, collimated, and divided into two paths by a polarizing beam splitter (PBS). This PBS is preceded by a ferroelectric liquid-crystal switch, which switches the optical polarization and thus changes the energy ratio between the paths. The beams recombine and interfere inside a copper-doped $Bi_{12}GeO_{20}$ photorefractive crystal (BGO:Cu), chosen for its high sensitivity.

The first path corresponds to the reference beams, that is, to the input vector. The crystal is illuminated with several angularly multiplexed reference beams, whose amplitudes are the components of the input vector. Each of these beams is created by a point source located in the object focal plane of lens L4. Because of the presence of lens L3, this plane is also the Fourier plane of a Displaytech binary ferroelectric liquid-crystal spatial light modulator (SLM). We generate the desired point-source pattern by displaying its inverse Fourier transform onto this SLM. We chose a ferroelectric SLM because of its short commutation time. The area of the 256×256 pixels is 3.84×3.84 mm. Each pixel of this reflective SLM switches the vertical (i.e., perpendicular to the plane of Fig. 2) input polarization into one of the two states that are linearly polarized at $\pm 45°$ to the vertical. The following PBS cube transforms this polarization

modulation into a $0-\pi$ binary phase modulation. This binary phase modulation produces, in the Fourier plane, the required point-source pattern as the +1-diffracted order. This binary phase modulation also produces a zero order and other orders that are blocked by the mask (M1, Fig. 2) set in this Fourier plane. This way to generate the reference beams was chosen because it automatically (roughly) normalizes the input vectors and allows one to use grayscale vectors even though the SLM is binary. Moreover, it can be shown that the residual noise obtained with this method varies randomly when the pattern is changed. This randomization strongly reduces the parasitic gratings inside the material.[22]

These reference beams are regularly spaced in the horizontal plane. To reduce Bragg degeneracy cross talk, we make the angular spacing in the horizontal plane between two consecutive reference beams twice the angular Bragg selectivity. This spacing permits the use of 128 beams simultaneously. If these reference beams were regularly aligned in the same incident (horizontal) plane, the harmonics that resulted from the binarization for one reference beam would be superimposed upon the other reference beams, which would lead to highly nonuniform patterns that are difficult to predict. To reduce this problem, the vertical positions of the points in the object focal plane are chosen at random. The total vertical span of these points is half the horizontal span. Lenses L3 and L4 form a refractive telescope that images the SLM onto the crystal with a 1.2× magnification ratio.

The second path corresponds to the learning beams. A second Displaytech ferroelectric SLM is imaged onto the crystal. The effective area of the 256×256 pixels is also 3.84 mm × 3.84 mm. This reflective SLM changes the horizontal input polarization to either vertical or horizontal polarization. This new polarization direction results in a binary amplitude modulation after the PBS cube. Many diffracted orders are present in the Fourier plane of lens L1. With the help of filter M2, which selects the zero diffracted order only, we obtained an intensity contrast ratio of 2300:1. This high contrast is the second reason why we chose ferroelectric SLMs.

The refractive telescope, made with lenses L1 and L2, forms the image of this SLM onto the crystal with a unit magnification ratio. Each block of 2×2 pixels defines the amplitude of one learning beam. The crystal is thus illuminated with 128×128 spatially multiplexed signal beams. A mechanical shutter is located in this path to stop the light during the readout stage (see Section 4 below). The crystal is then imaged onto a Dalsa 256×256 fast CCD camera. That camera is used with the binning mode enabled; i.e., the pixels are grouped in blocks of 2×2 pixels, and their signals are added such that the apparent resolution of the CCD is 128×128 pixels. Operation of the CCD camera in this binning mode was required for increasing the signal-to-noise ratio that was limited by the low diffraction efficiency of the photorefractive crystal. The optical system is made such that each 2×2 pixel block corresponds to one learning–diffracted beam.

As can be seen from Fig. 2, the BGO:Cu crystal is set in a 90° geometry: The reference and learning paths are orthogonal. The crystal's dimensions are 5.9 mm × 7.3 mm × 7.5 mm along the $[110] \times [\bar{1}1\bar{1}] \times [\bar{1}11]$ crystallographic axes. The [110] axis is perpendicular to the incidence plane; the reference beams enter by the $(\bar{1}11)$ face, and the learning beams enter by the $(\bar{1}1\bar{1})$ face. Quarter-wave plates were used on both paths to yield circular polarization (with the same rotation directions). We did this to maximize the diffraction efficiency by taking into account the strong optical activity of the crystal. For this configuration, we measured an exponential time constant for the buildup of the photorefractive grating of 340 ms, for a total intensity of ~50 mW/cm^2. This relatively short photorefractive time constant shortens the learning procedure. It is typical of $Bi_{12}GeO_{20}$ samples and is gained at the expense of low diffraction efficiency. The measured photorefractive gain is indeed 0.3 cm^{-1}. With the intensity ratio between the reference and the image paths taken into account, the gain has a maximum diffraction efficiency of less than 0.3%, to be divided among the 128×128 pixels. This diffraction efficiency is even reduced relative to the square of the number of vectors to be classified, which can limit this number eventually.

This 90° geometry offers some advantages compared with a copropagating geometry: The diffraction efficiency is nearly independent of the reference beam's incident angle, the number of usable reference beams is greater,[23] and the scattering noise is less bothersome.

During writing, the illumination of the reference path is 1.2 mW cm^{-2} onto the crystal, whereas that of the learning path is 48 mW cm^{-2}. This high ratio limits the erasure of the holograms that correspond to neurons, which should not be modified and which are illuminated by the reference beams only. Before each readout, switching of the ferroelectric switch increases the illumination of the reference path to 3 mW cm^{-2} to improve the diffracted amplitudes slightly. The inverse switching restores the initial illumination after readout and before updating of the subsequent hologram.

4. Learning Algorithm

A. Principle

Our system must be trained to learn how to classify the input vectors correctly. This training is unsupervised because no classification is imposed by the experimenter. The system uses the learning algorithm to automatically group together vectors that are similar one to another. The algorithm consists in presenting a particular input vector, detecting the location of the higher response (called the winning neuron), and updating the weights of the neurons that are located in the neighborhood of the winner. The whole process is then repeated with a new input

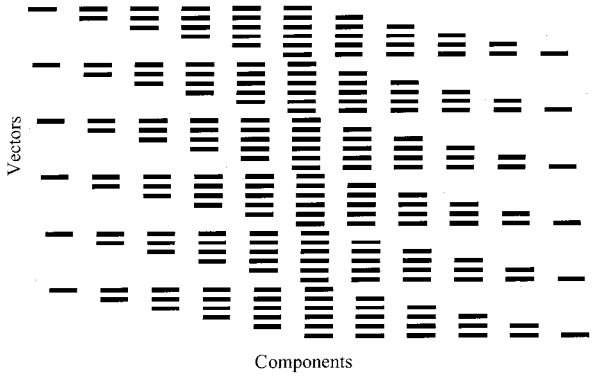

Fig. 3. Thirty-six grid-correlated vectors (bidimensional topology). Bright parts, 0s; dark parts, 1s.

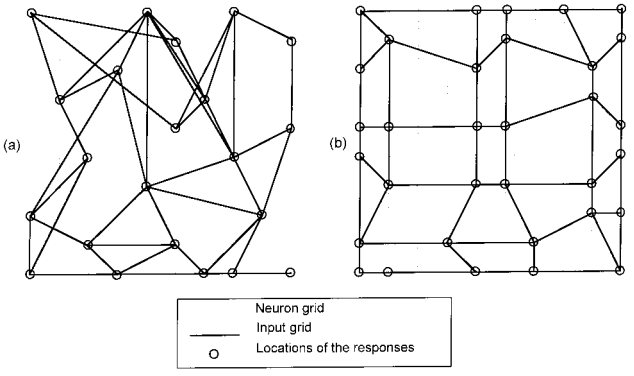

Fig. 4. Computer simulation: classification of a grid of vectors with metric discrepancy between readout and writing. (a) Kohonen's algorithm, (b) modified algorithm.

vector until the classification is completed. The radius of the updated neighborhood shrinks with time.

B. Metrics Discrepancy

In his description of the learning algorithm, Kohonen[19] specifies that the measurement metrics of the neuron responses must be consistent with the weight-vector updating law. Thus he uses the weight updating that is given in relation (3) and that corresponds to a reduction of the Euclidean distance between the weight and the input vectors. However, this consistency implies that the responses of the neurons must also be a measurement of the Euclidean distance between these two vectors. Such is not the case for our system because the responses are given by relation (2), which is a measurement of correlation instead of Euclidean distance. The metrics based on correlation may also be used for the topological map, but, in this case, every vector should be normalized.[20] In our setup, such use is impossible for the weight vectors that are defined by holograms. We thus have a metrics discrepancy between reading (correlation-based distance) and writing (Euclidean distance). We have demonstrated, with computer simulations, that this discrepancy results in wrong classifications of the input vectors. For instance, we used a 10×10 map to classify 36 vectors, each with 121 components. These vectors are shown in Fig. 3, where the bright parts correspond to 0s and the dark parts represent 1s. It can be seen that vector 1, for example, is decreasingly correlated with vectors 2–6 but also with vectors 7, 13, 19, 25, and 31. In mathematical terms, we can say that the topology of the input space is bidimensional. That is, the vectors can be seen as the nodes of a 6×6 grid and the mutual correlations are higher, as the corresponding nodes are spatially close to one another. At the end of the learning stage, we plotted the location of the winning neuron for each input vector, and we linked these locations according to the initial correlation grid. If the topological classification were correct, we should have reconstructed a kind of grid. As anticipated, Fig. 4(a) shows that this does not happen when Kohonen's algorithm is applied. In this algorithm, the update coefficient, which corresponds to exposure time Δt in relation (3), decreases with time. We showed that it is possible to take the same coefficient for every vector presentation, provided that the coefficient is sufficiently small to yield a limited modification of the weights (for instance, ~ 0.1). With this single small modification of the algorithm we obtained a better classification, even though there is a metrics discrepancy [Fig. 4(b)].

C. Experimental Implementation

The learning process starts by displaying on the reference SLM a pattern that corresponds to an input vector of the learning database. At that time, the shutter on the learning path is closed. The only light that goes in the direction of the camera is then the light that results from diffraction of the reference beams onto the holograms that are recorded inside the crystal. At the beginning of the learning stage, there is no hologram, and the camera detects only noise. The learning therefore starts from noise. The signal detected by the camera stands for the responses of the neurons. These responses are spatially nonuniform. They are sent to a personal computer, which detects the location of the higher response (winning neuron). The shutter is then opened, and a disk centered on the winning neuron is displayed on the learning SLM. The radius of this disk decreases when the number of iterations increases. Displaying this disk on the learning path results in illumination of only a small area of the crystal by the learning beams. In this area—and nowhere else—the holograms are updated. According to relation (3), the holograms that correspond to updated neuron k are reinforced if $\mathbf{W}_k$ is similar to the input vector; if not, they are erased. This explains why adjacent neurons tend to be sensitized to similar vectors and also why widely dissimilar vectors tend to activate the map in extremely distant locations.[21]

The whole process is reiterated with a new input vector and for the same exposure time. At the end of the learning stage (typically after 2000 vector presentations), the responses should be spatially organized according to the mutual correlations of the input vectors. Even though the process is long, the relative

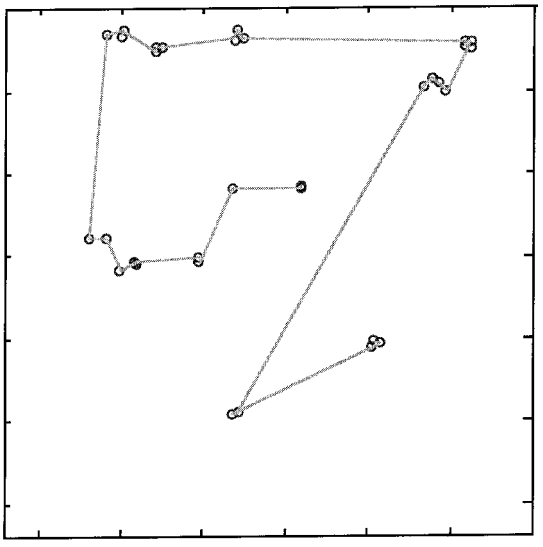

Fig. 5. Topological organization with 45 chain-correlated vectors. Circles, the winning neurons. They are linked by the straight lines in order of vector number.

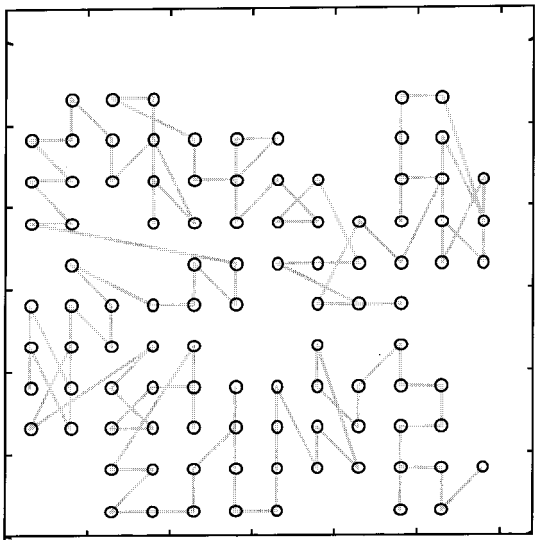

Fig. 6. Topological organization with 100 chain-correlated vectors. Circles, the winning neurons. They are linked by the straight lines in order of vector number.

path difference between the two arms of the interferometer does not have to be accurately maintained through time because the holograms are continually updated. The setup has just to be stable between two consecutive presentations of the same vector, or even between those of highly similar vectors, which means at most a few seconds.

5. Experimental Results

A. Chain-Correlated Vectors

The first set of vectors is a series of 45 binary vectors, each with 50 components. Six components are in the ON state, which corresponds to the maximum reflectivity of the SLM pixels; the remaining 44 components are OFF, i.e., there is no light. Vector N is defined as follows,

$$\mathbf{I}_N = [(N-1)\times 0,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ (45-N)\times 0)],$$

so the first vector is correlated with the second one, the second one with the third one, and so on. Each vector is thus correlated with several of its neighbors, but the correlation is greater as the vector numbers move closer to one another.

During the learning stage, the holograms were reinforced for 50 ms at each iteration. However, the mechanical shutter did not withstand high switching frequencies; thus the speed was limited to ~3 iterations/s. The whole learning process required 2000 vector presentations and lasted for ~10 min. The radius of the disk update area shrank from 100 to 3 neurons. Figure 5 represents the output face of the crystal as seen by the camera. Each circle is the winning neuron for one particular vector. These circles have been linked in order of vector number, namely, 1–2–3–4–. . . . It can be seen that the system has self-organized, because the responses are not randomly distributed but rather are located according to the mutual correlations between vectors. One can yet note that the circles are grouped into clusters. This grouping results from the nonuniformity of the optical system, which tends to favor some areas of the crystal where there is more light. To reduce this problem in the following examples, we impose the condition that the responses for two different vectors be separated by at least 10 neurons. To do this, we perform the learning step by looking for the winning neuron on a grid with a 10-neuron step. The neurons that are not on this grid are not considered as we seek the maximum response. Moreover, we forbid one neuron from being the winner for two different vectors. That is, we memorize the position of the winning neuron for every vector, and, if the winner for a vector already happens to be the winning neuron for one of the other vectors, we discard it and look for the next-highest response on the grid of allowed winners. This use of a grid does not really reduce the capacity of the system because there are many more neurons than vectors to classify anyway. We use the grid merely to ensure that the whole map will be used. If we had a larger number of vectors we could reduce the separation distance to 1 neuron.

Figure 6 gives a classification of 100 vectors with 111 components. These vectors are constructed on the model described previously, except that 12 components are ON for each vector. Although we obtained some local errors, the overall organization is quite good.

B. Grid-Correlated Vectors

The learning database consists of the same 36 binary vectors described in Subsection 4.2 (Fig. 3). For each vector, 36 of the 121 binary components are ON at the same time. These vectors are correlated following a bidimensional geometry: Each of them may be considered one node of a 6×6 grid,

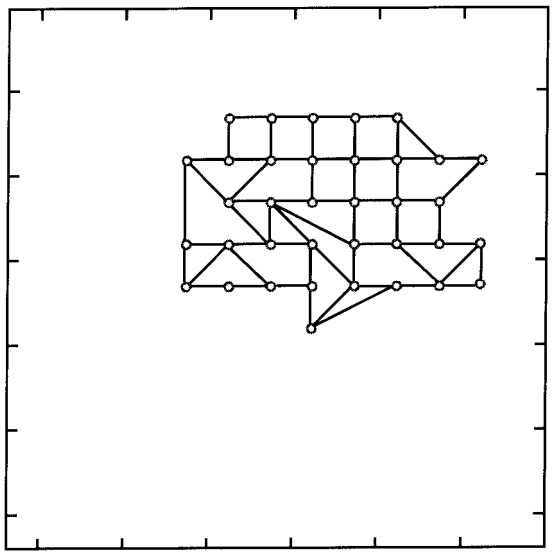

Fig. 7. Topological organization with 36 grid-correlated vectors. Straight lines link the winning neurons (circles) that respond to neighboring vectors in the initial correlation grid.

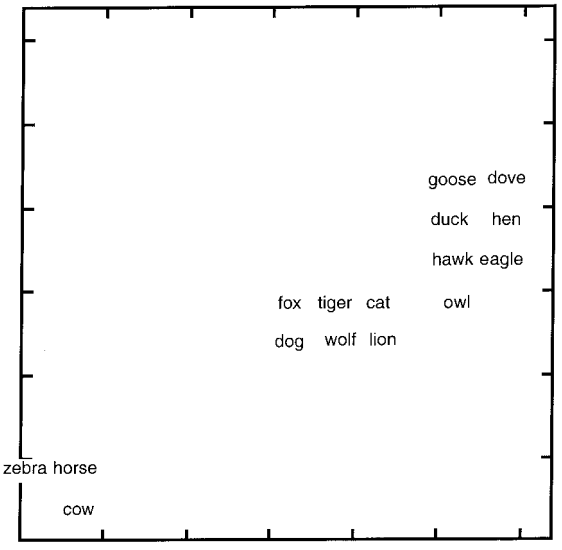

Fig. 8. Topological organization with 16 vectors that code characteristics of animals.

and the correlations are all the stronger as the vectors are closer to one other on this grid. The learning parameters were the same as above. The result of the self-organization can be seen in Fig. 7: The system successfully found and reconstituted the grid.

C. Classification of Animals

This learning database is taken from Ref. 19: Each of the 16 vectors describes an animal through a series of physiological and behavioral characteristics, such as the animal's size or its number of legs (Table 1). The vectors have 13 components, among which 4–6 are ON at a time. The other parameters are the same as above. We still imposed the condition that the responses repel each other by forbidding that a single neuron be the winner for two different vectors. Figure 8 gives the locations of the responses at the end of the learning stage. One can notice three main groups, which comprise, respectively, herbivorous mammals, carnivorous mammals, and birds. Moreover, the locations for predator mammals are close to those for predator birds. The system has thus correctly classified the animals according to their similarities. For this case as well as for the examples above, the classification is quite robust: The auto-organization is obtained at every attempt. Of course, because of the unsupervised nature of the process, the result is different each time. A few classification mistakes occur occasionally, but this happens also in computer-implemented neural networks.

Table 1. Vector Coding of Characteristics of Animals

	Characteristic												
Animal	Small	Medium	Big	Two Legs	Four Legs	Hair	Hooves	Mane	Feathers	Hunts	Runs	Flies	Swims
Dove	1	0	0	1	0	0	0	0	1	0	0	1	0
Hen	1	0	0	1	0	0	0	0	1	0	0	0	0
Goose	1	0	0	1	0	0	0	0	1	0	0	1	1
Duck	1	0	0	1	0	0	0	0	1	0	0	1	1
Owl	1	0	0	1	0	0	0	0	1	1	0	1	0
Hawk	1	0	0	1	0	0	0	0	1	1	0	1	0
Eagle	0	1	0	1	0	0	0	0	1	1	0	0	0
Fox	0	1	0	0	1	1	0	0	0	1	0	0	0
Dog	0	1	0	0	1	1	0	0	0	0	1	0	0
Wolf	0	1	0	0	1	1	0	1	0	1	1	0	0
Cat	1	0	0	0	1	1	0	0	0	1	0	0	0
Tiger	0	0	1	0	1	1	0	0	0	1	1	0	0
Lion	0	0	1	0	1	1	0	1	0	1	1	0	0
Horse	0	0	1	0	1	1	1	1	0	0	1	0	0
Zebra	0	0	1	0	1	1	1	1	0	0	1	0	0
Cow	0	0	1	0	1	1	1	0	0	0	0	0	0

6. Discussion of the Setup's Capacity

As our reference SLM has 256 × 256 pixels, it is able to address 256 × 256 points in its Fourier plane, which in principle allows it to use as many as 256 angularly multiplexed reference beams. However, we have already explained that to avoid cross talk we can use only 128 beams. This number is imposed by the dimensions of the crystal and of the SLM and cannot be increased easily.

In the learning path, we had to group the pixels 4 by 4 to increase the energy per neuron. This grouping is useful because the diffraction efficiency is low. Moreover, Filter M2 (Fig. 2) causes the diffraction spot in the crystal to have the same size as the image of one pixel. We thus have some interpixel cross talk. This is not too bothersome when a neuron is made by four pixels because the cross talk is then weak compared with the response of the neuron and because Konhonen's algorithm uses neighbor interactions anyway. However, if the neurons had only one pixel, the cross talk would become too strong for the neurons to be considered independent. The maximum number of neurons is thus 128 × 128. But we never used all the neurons at the same time. The reason is that usually one input vector is expected to activate one output neuron at most. So, although all the neurons are effectively present and modified simultaneously, we actually use as many neurons as we have input vectors. The problem is that the number of vectors is at present limited to ~100 because of the unwanted erasure of the holograms. Indeed, as we mentioned in Section 3, the holograms that correspond to one particular vector are partially erased by the reference beams when the other vectors are recorded. This erasure depends both on the total number of vectors and on the power ratio between the two arms of the interferometer. However, this ratio cannot be increased indefinitely because each increase reduces the diffraction efficiency, which is already low.

As we explained above, we use our setup at the maximum of its capacity. Anyway, with these particular components (crystal, camera) and a laser, we could hardly increase the capacity because the contrast on the camera (signal ratio between active and inactive neurons) can be estimated to be already ~1.1. The problem is that the diffraction efficiency of the crystal is low to begin with and decreases when the number of vectors increases. Moreover, the offset of the camera increases with the exposure time, which prevents integration of the signal over a long period.

7. Conclusions

We have designed, built, and demonstrated an optoelectronic neural network that uses volume holograms in a photorefractive crystal to implement a Kohonen self-organizing network. We have shown that the system is able to classify successfully as many as 100 vectors with more than 100 components each. As it is designed, the setup provides 128 × 128 neurons (although we have not used them all by now) and 128 input vector components. Because approximately three input vectors can be presented each second, the system is able to perform more than 10^6 weight updates each second. This result shows that the setup is among the best optoelectronic neural networks, in terms of capacity, that have been built. To improve it, some remaining problems, such as nonuniformity of the responses, low diffraction efficiency, and unwanted erasure of the holograms during exposures, still have to be solved.

The authors are very grateful to Jean-Claude Launay for providing the BGO:Cu crystal and to Virginie Luyckx for cutting and polishing it.

References

1. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," Appl. Opt. **24,** 1469–1475 (1985).
2. J. H. Hong, S. Campbell, and P. Yeh, "Optical pattern classifier with perceptron learning," Appl. Opt. **29,** 3019–3025 (1990).
3. Y. Owechko, "Optical implementation of back-propagation neural networks using cascaded-grating holography," Int. J. Opt. Comput. **2,** 201–231 (1991).
4. J. Duvillier, M. Killinger, K. Heggarty, K. Yao, and J. L. de Bougrenet de la Tocnaye, "All-optical implementation of a self-organizing map: a preliminary approach," Appl. Opt. **33,** 258–266 (1994).
5. M. Barge, K. Heggarty, Y. Idan, and R. Chevallier, "64-channel correlator implementing a Kohonen-like neural network for handwritten-digit recognition," Appl. Opt. **35,** 4655–4665 (1996).
6. C. Berger, N. Collings, R. Völke, M. T. Gale, and T. Hessler, "A microlens-array-based optical neural network application," Pure Appl. Opt. **6,** 683–689 (1997).
7. M. Saffman, D. Montgomery, A. A. Zozulya, and D. Z. Anderson, "Topology-preserving mappings in a self-imaging photorefractively pumped ring resonator," Chaos, Solitons Fractals **4,** 2077–2092 (1994).
8. P. Aing, G. Pauliat, and G. Roosen, "Noise issues in holographic photorefractive interconnections: application to neural networks," Opt. Commun. **143,** 87–94 (1997).
9. Y. Frauel, T. Galstyan, G. Pauliat, A. Villing, and G. Roosen, "Topological map from a photorefractive self-organizing neural network," Opt. Commun. **135,** 179–188 (1997).
10. Y. Frauel, T. Galstyan, G. Pauliat, A. Villing, and G. Roosen, "Topological map from a photorefractive self-organizing neural network (Erratum)," Opt. Commun. **135,** 335 (1997).
11. G. W. Burr, S. Kobras, H. Hanssen, and H. Coufal, "Content-addressable data storage by use of volume holograms," Appl. Opt. **38,** 6779–6784 (1999).
12. K. Wagner and D. Psaltis, "Multilayer optical learning networks," Appl. Opt. **26,** 5061–5076 (1987).
13. G. A. Betzos, A. Lainé, and P. A. Mitkas, "Improved associative recall of binary data in volume holographic memories," Opt. Commun. **171,** 37–44 (1999).
14. H.-Y. S. Li, Y. Qiao, and D. Psaltis, "Optical network for real-time face recognition," Appl. Opt. **32,** 5026–5035 (1993).
15. K. Wagner and T. M. Slagle, "Optical competitive learning with VLSI/liquid-crystal winner-take-all modulators," Appl. Opt. **32,** 1408–1435 (1993).
16. F. H. Mok and H. M. Stoll, "Holographic inner-product processor for pattern recognition," in *Optical Pattern Recognition IV*, D. P. Casasent and T.-H. Chao, eds., Proc. SPIE **1701,** 312–322 (1992).

17. X. An, D. Psaltis, and G. W. Burr, "Thermal fixing of 10,000 holograms in $LiNbO_3$:Fe," Appl. Opt. **38,** 386–393 (1999).
18. N. V. Kukhtarev, V. B. Markov, S. G. Odulov, M. S. Soskin, and V. L. Vinetskii, "Holographic storage in electrooptic crystals," Ferroelectrics **22,** 949–964 (1979).
19. H. Lee, X. G. Gu, and D. Psaltis, "Volume holographic interconnections with maximal capacity and minimal crosstalk," J. Appl. Phys. **65,** 2191–2194 (1989).
20. T. Kohonen, *Self-organizing Maps* (Springer-Verlag, Berlin, 1997).
21. T. Galstyan, G. Pauliat, A. Villing, and G. Roosen, "Adaptive photorefractive neurons for self-organizing networks," Opt. Commun. **109,** 35–42 (1994).
22. Y. Frauel, G. Pauliat, and G. Roosen, "Improvement of holographic neural networks by reducing the deleterious influence of the limited contrast of spatial light modulators," Opt. Commun. **182,** 311–319 (2000).
23. H.-Y. S. Li, and D. Psaltis, "Three dimensional holographic disks," Appl. Opt. **33,** 3764–3774 (1994).