

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Optical neural networks

Demetri Psaltis, Michael Levene

Demetri Psaltis, Michael Levene, "Optical neural networks," Proc. SPIE 10277, Adaptive Computing: Mathematics, Electronics, and Optics: A Critical Review, 1027709 (1 March 1994); doi: 10.1117/12.171198

SPIE.

Event: SPIE's International Symposium on Optical Engineering and Photonics in Aerospace Sensing, 1994, Orlando, FL, United States

Optical Neural Networks

Demetri Psaltis and Michael Levene
California Institute of Technology
Department of Electrical Engineering
Pasadena, California 91125

ABSTRACT

The trade-off between the number of neurons that can be implemented with a single correlator and the shift invariance that each neuron has is investigated. A new type of correlator implemented with a planar hologram is described whose shift invariance can be controlled by setting the position of the hologram properly. The shift invariance and the capacity of correlators implemented with volume holograms is also investigated.

1. Introduction

Image processing is the primary application for optical neural networks. In this application shift invariance is an important consideration. Normally, we wish to recognize an image independently of its position in the field of view. Optical pattern recognition systems have traditionally been designed with shift invariance using the classical VanderLugt correlator [1] and its variants. A schematic diagram of the conventional correlator is shown in Fig. 1. The output is the 2-D correlation between the input image and a reference image stored in the Fourier transform hologram at the middle plane. If there is a match between the input and the stored reference then a correlation

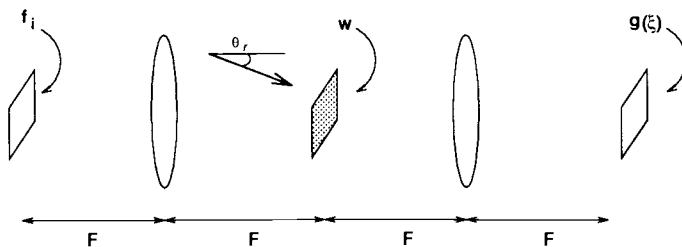


Figure 1. VanderLugt Correlator

Adaptive Computing: Mathematics, Electronics, and Optics: A Critical Review, edited by Su-Shing Chen,
H. John Caulfield, Proc. of SPIE Vol. 10277 (Vol. CR55), 1027709 · © (1994) 2017 SPIE
CCC code: 0277-786X/17/\$18 · doi: 10.1117/12.171198

peak forms at the output. The position of the correlation peak corresponds to the position of the input object. If the input object moves then the correlation peak follows it. In recent years such correlators have been used as building blocks to construct large optical neural networks [2]. Networks with several thousand units can be readily constructed optically. Each unit can be thought of as an optical correlator with the weights of the neuron encoded in the hologram.

It is not practical to duplicate the basic correlator hundreds or thousands of times in order to build a large network. Therefore, methods have been developed for storing the weights of multiple neurons on the same hologram [3]. Different pixels at the output plane are devoted to represent the response of different units. This raises the trade-off between shift invariance and the number of units that can be represented by a single correlator. If the response of multiple neurons can be represented by pixels at the output plane then the shift invariance for each has to be at least partially destroyed to make room for the additional units. In this paper we explore two methods for representing multiple neurons with a single correlator. The first method uses Fresnel holograms instead of Fourier holograms which destroys the shift invariance of the system and permits multiple locally tuned units to operate in parallel. The second method uses volume holograms in the middle plane. The Bragg selectivity of volume holograms destroys the shift invariance of the correlator. The thickness of the hologram and the geometry used determine the degree and the type of shift invariance obtained.

2. Fresnel Correlator

The Fresnel correlator is shown in Fig. 2. The only difference between the systems in Figures 1 and 2 is the fact that the hologram is positioned a distance d away from the Fourier plane. The hologram is recorded with a plane wave reference at an angle θ_r as shown in Fig. 2. The transmittance of the recorded hologram is

$$t(x) \propto e^{\frac{j2\pi}{\lambda}(\sin \theta_r x + \cos \theta_r d)} \sum_i f_i^* e^{-j2\pi i \Delta x / \lambda F} e^{j\pi i^2 \Delta^2 d / \lambda F^2} \quad (2)$$

where F is the focal length of the lens, Δ is the pixel spacing at the input plane, and λ is the wavelength of light.

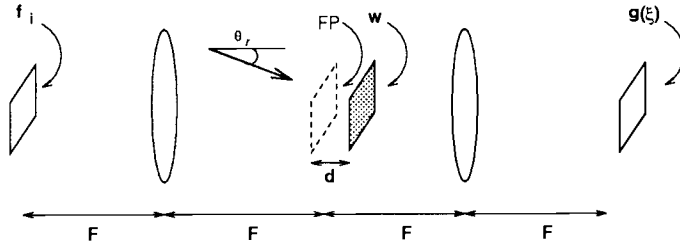


Figure 2. Fresnel correlator

When the same image f_i is placed at the input plane then the light at the output plane of the correlator is

$$\begin{aligned}
 g(\xi) &= \int t(x) \sum_i f_i e^{2j\pi i \Delta x / \lambda F} e^{-j\pi i^2 \Delta^2 d / \lambda F^2} e^{-2j\pi \xi / \lambda F} dx \\
 &\approx \sum_i |f_i|^2 \delta(\xi - F \sin \theta_r)
 \end{aligned}
 \tag{3}$$

In other words when the input is an exact replica of the image that was used to record the hologram then the output can be approximated by a delta function when the finite apertures are ignored. In practice, this means that a sharp autocorrelation peak appears in the Fresnel correlator just as in the conventional correlator. When the input is a shifted version of f_i then the output becomes

$$\begin{aligned}
 g(\xi) &= \int t(x) \sum_i f(i - i_0) e^{2j\pi i \Delta x / \lambda F} e^{-j\pi i^2 \Delta^2 d / \lambda F^2} e^{-2j\pi \xi / \lambda F} dx \\
 &\approx \sum_i |f_i|^2 e^{j2\pi i i_0 \Delta^2 d / \lambda F^2} \delta(\xi - F \sin \theta_r - i_0 \Delta)
 \end{aligned}
 \tag{4}$$

The correlation peak shifts at the output as in the conventional correlator however, the strength of the correlation peak diminishes due to the phase factor in the summation over i in Eq. (4). We can obtain an estimate for the allowable shift invariance by assuming that f_i is bandlimited by B . In this case we can show that the correlation peak will become zero when

$$i_o \Delta > 2B\lambda F^2/d \quad (5)$$

Therefore, by selecting d sufficiently large, we can limit the shift invariance to any desired value. In this way we can allocate a specified region of the output plane to each reference image recorded on the hologram. Multiple neurons can be recorded on the same hologram by changing the angle of the reference beam for each new hologram which relocates the position of the autocorrelation peak. There is an obvious trade-off between shift invariance and the number of reference images that can be accommodated in the same system. Complete shift invariance would allow only one reference, and if there is no shift invariance then each output pixel can be used for a separate neuron. There are other limits, however, that restrict the number of neurons that can be represented. One limit is the limited dynamic range of the recording medium. The diffraction efficiency of the hologram goes as $1/M^2$ where M is the number of holograms stored. Therefore, for large enough M , the output falls below the noise level and it cannot be reliably detected.

A second, more serious limit is the crosstalk that arises when multiple holograms stored in a planar medium are read out. In the above equations we approximated the output correlation pattern by a delta function whose amplitude drops as the input shifts or changes. In reality, a finite signal level will be distributed throughout the entire output plane due to each recorded hologram. We can obtain a rough estimate for the amplitude of the signal away from the peak by assuming that the amplitude of the crosstalk is the random sum of the signals from the N pixels of the input image diffracted by the M holograms stored in the hologram. The standard deviation of the crosstalk noise goes as $\propto \sqrt{MN}$. The amplitude of the autocorrelation peak is proportional to N , therefore the SNR is $\sqrt{N/M}$. Therefore, given a desired SNR and N , the crosstalk limited bound on M can be readily calculated. For instance, if the desired SNR=10, and the number of pixels N in 2-D is 10^6 , then $M = 10^4$.

3. Volume hologram correlator in the 90° geometry

The crosstalk obtained in the Fresnel correlator described in the previous section is drastically reduced when volume holograms are used. This is due to the Bragg selectivity of volume holograms. The diffraction efficiency drops when the angle of the read-out beam changes. In a correlator, the read-out beam is actually the Fourier transform of the input image. If the input image is a replica of the image used to record the hologram, then the read-out beam is automatically Bragg matched and a strong autocorrelation peak is produced. When the input shifts, the Fourier transform is multiplied by a linear phase term. Physically, this phase term describes a tilt in the beam that illuminates the hologram compared to the recording beam. Thus, a shift in the input leads to a Bragg mismatch and a reduction in the diffraction efficiency. The loss in shift invariance due to Bragg mismatching is fundamentally different from the mechanism in the Fresnel correlator. Bragg mismatching leads to a loss in diffraction efficiency and hence no crosstalk. In the Fresnel correlator there is crosstalk due to defocussing that redistributes the energy from a single correlation peak to the entire output plane.

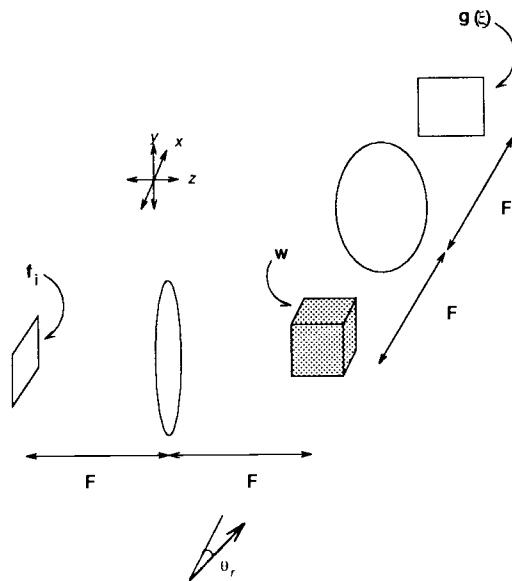


Figure 3. The 90° geometry

An optical correlator implemented in the 90° geometry is shown

in Figure 3. The hologram is approximately a cube. The signal beam enters the crystal from a different facet of the crystal than the reference beam, as shown in Fig. 3. The two beams intersect in the crystal at 90° . When a recorded hologram is read out with the Fourier transform of the input image, the diffracted signal is produced through a third facet of the crystal and the correlation signal forms at the output plane. If the input image shifts by a distance $i_0\Delta$ then the wave illuminating the crystal tilts by an angle $\Delta\theta = i_0\Delta/F$. The angular selectivity of the hologram in the 90° geometry is approximately λ/L_x , where L_x is the size of the crystal in the x direction. From these two relationships the maximum allowable shift is found to be

$$i_0 \leq \lambda F/L_x \Delta \quad (6)$$

pixels. The quantity on the RHS of the above equation is typically approximately equal to one, therefore this correlator has no shift invariance in the x direction. In the y direction, the volume hologram has no selectivity and therefore the shift invariance in y is approximately maintained. We can take advantage of the loss in shift invariance in x by recording multiple holograms. The reference beam is rotated in the $x - z$ plane by an angle $\Delta\theta$ before each new hologram is formed. The position of the correlation peak in the z direction is proportional to $\sin\theta_r$, which implies that the autocorrelation peaks due to different holograms will be horizontally displaced. The presence of a strong correlation peak unambiguously identifies the input image since a horizontal shift in the input that might move the correlation peak to the position of another reference image would be Bragg mismatched. Therefore there is almost no crosstalk in this architecture. The limit on the number of holograms that can be stored in this case is no longer limited by the SNR. Instead it is the geometrical limit due to the finite lens apertures that in practice imposes the upper bound on the number of holograms. Practical systems with $M = 1,000$ references can be readily implemented.

Figure 4. Volume hologram correlator in the reflection geometry

Although the 90° geometry allows one to record many independent correlators in one crystal, it is often desirable to have limited shift invariance in both dimensions rather than broad shift invariance in only one direction. This can be accomplished by recording holograms in the reflection geometry as shown in Fig. 4. The Bragg

selectivity function in the reflection geometry, within the paraxial approximation, is given by

$$\eta \propto \text{sinc}\left[\frac{Lz}{\lambda} \Delta\theta_s (\Delta\theta_s + \theta_s + \theta_r)\right]$$

where η is the diffraction efficiency, θ_r is the angle of the reference wave with respect to the z axis, θ_s is the angle of the recorded signal wave, $\Delta\theta_s$ is the shift in the reconstructing signal wave, and $\text{sinc}(x) = \sin(\pi x)/\pi x$.

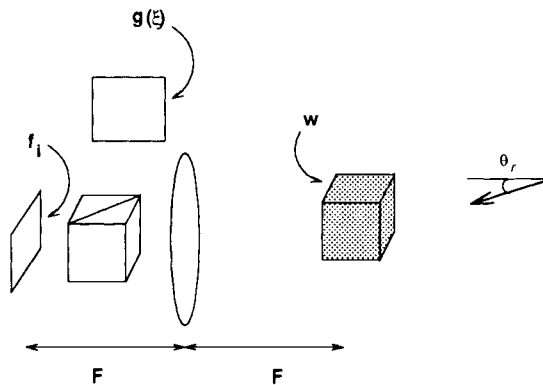


Figure 4. Reflection geometry correlator

The expression describing the field at the correlation plane due to the auto-correlation of $f(x)$ is

$$g(\xi) = \sum_i f(i\Delta) f^*(i\Delta + \xi + \theta_r F) \text{sinc}\left[\frac{Lz}{\lambda} \left(\frac{i\Delta}{F} + \theta_r\right) \left(\frac{\xi}{F} + \theta_r\right)\right]. \quad (12)$$

This is just the auto-correlation of $f(i\Delta)$ except for the *sinc* term that accounts for Bragg selectivity. The *sinc* term limits the shift invariance of the correlator. In this case, we get a similar windowing effect as in the 90° geometry. As images shift out of the center of the image plane the *sinc* term causes the correlation peak to diminish until, for a certain critical shift given by the first zeros of the *sinc*, the peak disappears completely. The *sinc* term also apodizes the image with increasing shift, so that features on the edge of the image contribute less to the correlation peak than those in the center.

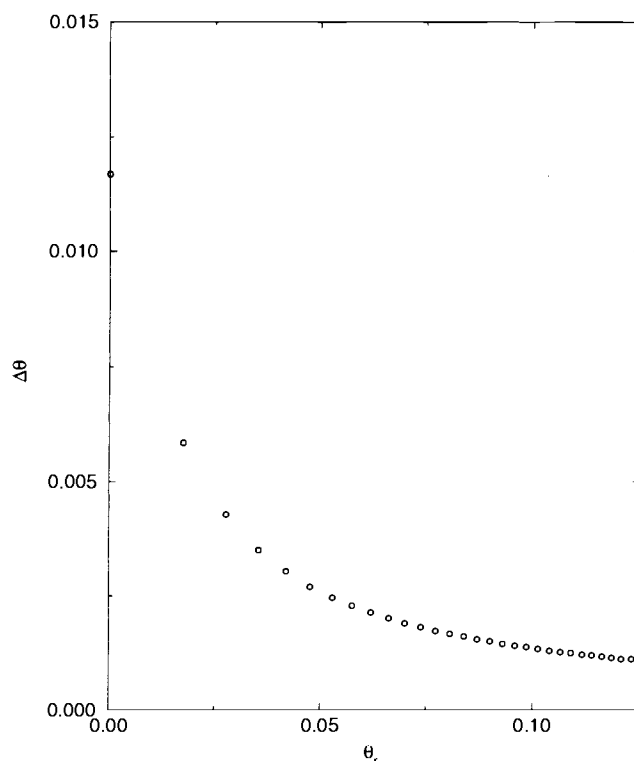


Figure 5. Shift invariance $\Delta\theta$ as a function of recording angle θ_r .

The limited shift invariance of the reflection geometry volume correlator creates the opportunity to tile the correlation plane with several correlators with limited shift invariance such that each correlator has its own square within the output plane. The center correlators (those recorded with on-axis reference waves) will have the most shift invariance, while those on the edges will have the least. The maximum allowable shift invariance for $\theta_r = 0$ is

$$\Delta\theta_s = \sqrt{\frac{\lambda}{L}}$$

The shift invariance decreases rapidly as θ_r increases. If we record adjacent correlators such that their first zeros overlap, we can be assured that there will be no confusion as to which correlator a given location on the correlation plane is responding to. Fig. 5 plots the shift invariance half angle versus recording angle in one dimension for such a recording scheme (angles are in radians external to an 8 mm crystal). If we assume a maximum $\theta_r \approx 0.08rad$, then a grid of 30 x

30 correlators can be formed. Larger images should be stored further from the center, since they are less likely to shift as much as smaller images.

5. Summary

The shift invariance and the capacity of the three architectures we discussed are summarized in Table 1, along with the conventional VanderLugt correlator. This simple comparison reveals the advantage of the reflection geometry which combines a large number of neurons with a limited shift invariance in both dimensions. The maximum allowable shift is \sqrt{N} pixels in both dimensions for the 180° geometry. The allowable shift invariance decreases rapidly as the reference beam angle is changed (see Fig. 5).

	Shift Invariance	Number of Neurons
Correlator	$N \times N$	1
Fresnel Correlator	$\sqrt{N}(SNR) \times \sqrt{N}(SNR)/(SNR)^2$	
90° Geometry	$N \times 1$	$\approx N$
Reflection Geometry	$< \sqrt{N} \times \sqrt{N}$	$\approx N$

Table 1. Comparison of the four architectures

Acknowledgements

This work is supported by ARPA. Michael Levene is partially supported by an NDSEG graduate fellowship. We also thank Sidney Li for his help.

References

- [1] A. VanderLugt, *Signal detection by complex spatial filtering*, IEEE Transactions on Information Theory, pp. 139 - 145, April 1964.
- [2] Y.S. Abu-Mostafa and D. Psaltis, *Optical Neural Computers*, Scientific American, Vol. 255, p.88, March 1987.
- [3] D. Psaltis, D. Brady, C. Gu, S. Lin. *Holography in Artificial Neural Networks*, Nature, vol. 343, no. 6256, p. 325, January 25, 1990.