CONF-8904242--3

SAND--88-2911C

DE89 016416

A comparison of real-time optical correlators for pattern recognition*

K. Terry Stalker
Perry A. Molley
Bruce D. Hansche

Sandia National Laboratories
Albuquerque, New Mexico, 87185 USA

ABSTRACT

Two types of optical correlators have been built to investigate real-time pattern recognition. The first employs one-dimensional devices to perform the two dimensional correlation in real time. This architecture uses an array of light emitting diodes (LED's) to input an electronically stored reference image into the processor in parallel. The input scene data is introduced into the processor one line at a time using an acousto-optic device (AOD). Multichannel time integrating correlations are performed in the row direction using the AOD and in the column direction using a charge coupled device (CCD) operating in the time delay and integrate mode. A processor has been built using this technology which correlates a 64 x 44 pixel binary reference image with a 256 x 232 input scene at video rates.

The second correlator is a space integrating Fourier transform based correlator. A magneto optic-device (MOD) is used at the Fourier transform plane to rapidly change filter functions. The binary nature of the MOD device necessitates using either a binary phase or binary amplitude representation of the desired complex filter function. For this reason, several types of Binary Phase-Only Filter (BPOF) representations have been analyzed and experimentally investigated. Experimental correlation results have been obtained using both the Hartley BPOF and a newly developed class of complex binary filters, called Quad-Phase-Only Filters (QPOF).

The performance of the two systems will be compared on the basis of processing speed, space bandwidth product, processor size and light efficiency. The inherent differences between incoherent and coherent processing and their implications for filter design will also be discussed. Finally, estimates of future performance will be presented.

1. INTRODUCTION

Correlation based pattern recognition has been one of the chief research areas in optical information processing for many years. The Fourier transform based correlator shown in Figure 1 is the architecture which has received the most attention. These processors are based on the property that lenses perform the two-dimensional Fourier transform of the input light amplitude distribution placed

* This work performed at Sandia National Laboratories and supported by the U.S. Department of Energy under contract DE-AC04-76DP000789.

in the front focal plane of the lens. Correlation is obtained by placing a mask containing the complex conjugate of the Fourier transform of the object to be recognized at the rear focal plane of the lens. In this manner the Fourier transform of the input scene is multiplied in parallel with the Fourier mask and the cross correlation is obtained at the output plane after inverse transformation by a second lens. A spatially coherent, quasi-monochromatic light source is required since complex numbers are being represented. Initial concepts utilized composite film masks to represent the amplitude and phase portions of the Fourier transform of the desired pattern to be recognized.[1] The processing steps required to make these composite filters were quite involved and alignment of the phase and amplitude masks was critical. While this basic system has provided the motivation for much research over the past two decades, there are still difficulties in obtaining real-time operation.

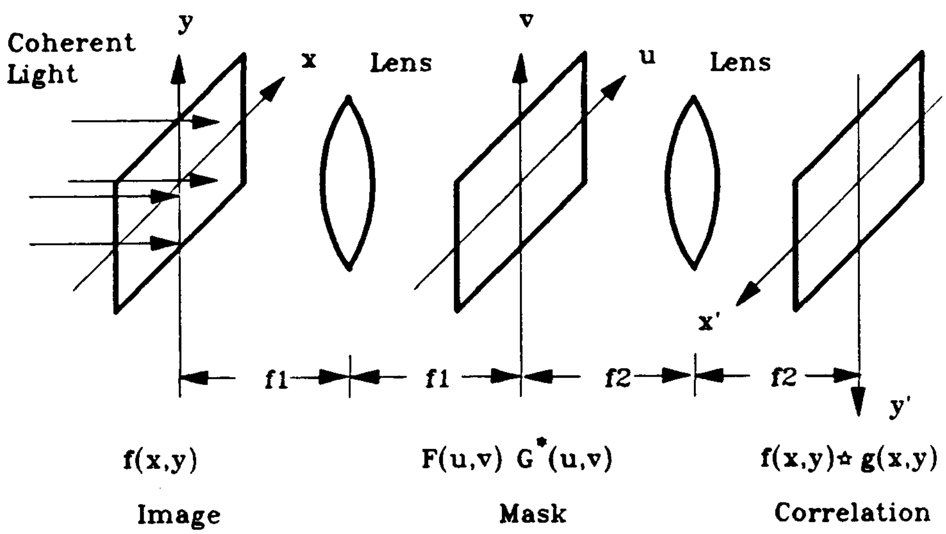


Figure 1. Fourier transform based optical correlator

The introduction of the holographic matched filter by Vander Lugt made coherent optical correlation much more practical.[2] The Vander Lugt filter utilizes an off axis plane wave which provides a phase reference so that both amplitude and phase information can be holographically recorded in an amplitude recording material. In this manner complex filters (containing both amplitude and phase information) can be recorded in media such as film which only modulate the amplitude of the optical field. Because of the relative ease of implementing this type of filter, Vander Lugt's filter has formed the basis of optical pattern recognition systems for many years.

While Vander Lugt filters can be easily implemented using film or other photo sensitive media, they cannot generally be changed in real time. Thus, real-time optical pattern recognition has been hampered by the lack of real-time two-dimensional spatial light modulators (SLMs). Many candidate SLMs have been proposed and demonstrated over the past two decades. While any single candidate may have several desirable attributes, none have proven to have the desired combination of speed, modulation capability, space bandwidth product and cost to become a commercial success. Today the most widely used two-dimensional SLM devices are the liquid crystal light valve (LCLV), the liquid crystal television (LCTV), and the magneto-optic device (MOD). There are other promising but less widely used technologies including the deformable mirror device (DMD) and the microchannel spatial light modulator (MSLM).

Comparisons of these devices have been chronicled elsewhere.[3] For the purposes of this paper, the Fourier based correlator that will be reported on is based on the MOD. This particular device was chosen because of its potentially high refresh rate, up to 1000 frames per second, relatively large space bandwidth product, 128 x 128 pixels, and commercial availability.

Another approach to real-time correlation for pattern recognition utilizes 1-D input devices instead of using the Fourier transformation properties of lenses. It employs a combination of a one-dimensional acousto-optical device and a two-dimensional charge coupled detector (CCD) operating in the time delay and integrate (TDI) mode to perform the two-dimensional image correlation. Figure 2 illustrates the key features of this architecture which was first proposed by Psaltis.[4] It consists of multiple one-dimensional acousto-optic time integrating correlators operating in parallel. All lines of the reference image are simultaneously correlated with a single raster line of the input image. The reference image is stored in a electronic memory and introduced into the processor in row parallel fashion through the LED (or laser diode) array after digital to analog conversion. Correlation in the column direction is performed by operating the CCD array in the TDI mode to perform an analog shift-and-add operation. In this manner a full two-dimensional correlation can be performed on the input data after M lines have been read in, where M is the height of the reference image.

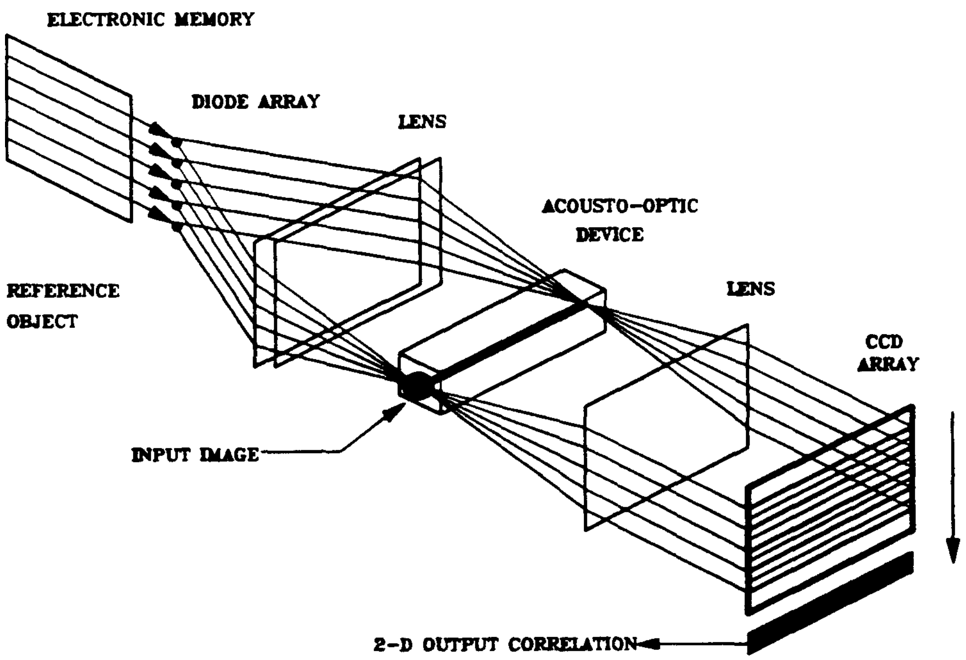


Figure 2. Acousto-optic image correlator

This architecture has several attractive features. First, the use of relatively mature AOD, CCD, LED and laser diode technologies means that operational correlators can be built in the near term.[5] The use of electronically stored reference templates allows a library of filters to be searched at a high rate. Also, there is the potential for very high processor throughput since laser diode, LED and AOD technologies support operation in excess of 100 Mhz. In addition, AODs have space-bandwidth products in excess of 1000 and so large (1000 pixel) input scenes can be processed. Finally, these technologies lend themselves to miniaturization and ruggedization as has been demonstrated by prototypes such as a compact optical spectrum analyzer.[6]

In the next section, an analysis of both systems will be performed to identify key operating parameters for each system. Then experimental results from the two systems developed at Sandia will be presented. Next the systems will be compared upon the basis of the current state of technology and potential developments. Finally, some future research directions will be outlined.

2. PRINCIPLES OF OPERATION

The operating principles of the two correlators will be reviewed in this section as a basis for comparing the performance. This will be only a brief review since both techniques are analyzed in depth in the literature. In both cases the goal is to perform the two-dimensional correlation of a reference image r(x,y) with an input scene f(x,y) given by

$$C(x,y) = \iint f(\zeta,\eta)\ r^{*}(\zeta-x,\ \eta-y)d\zeta d\eta. \qquad (1)$$

In the case of the Fourier transform based correlator C(x,y) can be calculated as

$$C(x,y) = F.T.^{-1}\ \{F\ R^{*}\} \qquad (2)$$

where F and R are the Fourier transform of f and r respectively and $F.T.^{-1}$ { } is the inverse Fourier transform. It should be pointed out that since F and R^{*} are both complex, even though f and r may be real and positive, coherent optical processing must be used to account for both amplitude and phase information.

Compromises in true complex operation have been made to achieve the real-time operation needed in many pattern recognition problems. For example, the MOD SLMs used in our experimental setup provide only two values of amplitude or phase. This has led to much interest in Binary Phase Only Filters (BPOFs) and their implementation.[7,8,9]

The second approach which was experimentally implemented is an incoherent, time integrating acousto-optic image correlator.[5] As was seen in Figure 2, this architecture consists of multiple one-dimensional correlators each performing the time integration given by:

$$C(x) = \int_T r(t)\ f(t-x/v)\ dt \tag{3}$$

where r modulates the LED source, f(t) is applied to the AOD, v is the acoustic velocity and T is the time duration of the reference image. There are M channels of input, each corresponding to a row in the reference image. Both the reference and the input image must be real positive functions since this is an incoherent correlation and is linear in intensity and not field amplitude. Thus both input signals may be represented as bipolar functions placed on a bias. Both f and r are assumed to be normalized with a maximum value of 1. The input signal introduced into the AOD is given by:

$$I_1(t,n) = \left[B_1 + A_1\ f(t-\tau,n) \right] \text{rect} \left[\frac{t-\tau-T_1/2}{T_1} \right] \tag{4}$$

and the reference signal modulating the LEDs is given by

$$I_2(t,m) = \left[B_2 + A_2\ r(t-T_1,m) \right] \text{rect} \left[\frac{t-T_1-T_2/2}{T_2} \right] \tag{5}$$

where T_1 is the duration of one line of the input image, T_2 is the duration of the reference image, n is the input image line number, m the row number of the reference image, and $\tau = x/v$, where v is the speed of sound in the AOD. If f and r are assumed to be zero mean, the output from the CCD after M rows of the input image have been processed is given by[5]:

$$C(\tau,n) = MT_2B_1B_2 + \sum_{i=0}^{M-1} \left[B_2A_1 \int_{T_1}^{T_1+T_2} f(t-\tau,n-i)dt \right.$$
$$\left. + A_1A_2 \int_{T_1}^{T_1+T_2} f(t-\tau,\ n-i)\ r(t-T_1,\ M-i)\ dt \right] \tag{6}$$

The last term is the 2-D correlation between f and r. The resulting output from the CCD can be displayed as a function of τ(horizontal position) and the input image line number n. If there is an occurrence of the reference object in the input image centered at $\tau = \tau_0$ and line $n = n_0$, a correlation peak would occur on the output of the CCD at $\tau = \tau_0$ in the horizontal direction and on the $(n_0 - M)$ row from the CCD. This represents a latency of M lines before the correlation peak reaches the output of the processor.

Equation 6 shows that the desired correlation is obtained at the processor output. However, both a signal dependent and constant bias term is also present. While several techniques including subtraction can be used to eliminate the unwanted terms, their presence affects the dynamic range required at the output detector. This will be discussed in more detail in a later section.

3. CORRELATOR PERFORMANCE COMPARISON

3.1 Processing speed. The relative processing speed of the two approaches will be analyzed in this section. In both cases let us assume that a M x M pixel reference image is being correlated against a larger N x N input scene. Since the Fourier transform based correlator uses optics to perform the transformation and multiplications in parallel, the speed is determined by the input speed of the SLMs involved and the output frame rate of the detector used at the correlation output plane. Thus the processing rate in operations per second (ops/sec) is given by

$$P = \frac{M^2 N^2}{T_L} = M^2 f_{pix} \tag{7}$$

where T_L is the limiting load time. It should be remembered that although M may be much smaller than N, the load time depends on the time required to load the larger image and read the N x N output plane. Thus the load time is given by N^2/f_{pix} when f_{pix} is the effective pixel clock rate.

In the case of the acousto-optic processor, there is a latency time NT_{pix} involved in loading a line of input data into the AOD where T_{pix}, the pixel period is equal to the inverse of f_{pix}, the pixel clock frequency. Then it takes MT_{pix} seconds to perform the correlation. This is done on all M rows of the input template in parallel. Thus the effective processing speed is given by

$$P = \frac{M^2 N}{(M+N)} f_{pix} \tag{8}$$

Comparisons of Equations 7 and 8 show that the two-dimensional Fourier correlator has a higher potential throughput rate. However, when N is very large compared to M, i.e. the scene is very large and the template is relatively small, the speed of the acousto-optic correlator approaches that of the 2-D Fourier based correlator. Table I compares the relative speeds of the two approaches for various image sizes.

Table I - Processing Speed vs Image Size

Input Image Size	Electro-optic Correlator	Fourier Correlator
N = 128	8.2×10^9	1.02×10^{10}
N = 256	8.9×10^9	1.02×10^{10}
N = 512	9.4×10^9	1.02×10^{10}
N = 1024	9.7×10^9	1.02×10^{10}

Template Size, M = 32 x 32 pixels
Pixel Clock, $f_{pix} = 10^7$

Table II - Comparison of Processing Speed

Technology	Input Image Size (pixels)	Frame Rate (fps) [Pixel Rate (pps)]	Processing Speed (ops/sec)	Remarks
2-D FOURIER BASED CORRELATOR			32 x 32 Template	
MOD	128 x 128	100 fps [1.6×10^6 pps]	1.7×10^9	Current Implementation
Future MOD	512 x 512	100 fps [2.6×10^6 pps]	2.7×10^{10}	Near Future
Future (DMD)	1024 x 1024	1000 fps [1.0×10^9 pps]	1.0×10^{12}	Requires Advances in I/O Technology
2-D ACOUSTO-OPTIC CORRELATOR				
T_eO_2	512	$1.0 \times x10^7$ pps	9.7×10^9	Demonstrated (Binary)
T_eO_2	1024	1.3×10^8 pps	1.3×10^{11}	Technology Exists
L_iNbO_3	1024	1.0×10^9 pps	1.0×10^{12}	Requires Development of CCDs, Lasers

3.2 Performance estimates. Although the previous section derived an expression for processing speed in terms of image size, reference template size and pixel clock rate, attainable processing speed depends on device technology. Table II shows the expected processing rate for two candidate technologies for both a MOD based Fourier correlator and the acousto-optic image correlator. The input rate for the Fourier correlator is given in both frames/sec (fps) and the equivalent pixel rate in pixels/sec (pps). The input data rate for the acousto-optic correlator is given only as a pixel rate in pps.

The table shows that current implementations of the MOD based correlators provide processing rates of 1.7 x 10^{10} operations per second (ops/sec) where an operation consists of a multiply-add. Increasing the image size to 512 x 512 increases the processing rate to 2.7 x 10^{10} ops/sec for a 100 fps refresh rate. Although refresh rates of up to 1000 fps may be feasible, 80 to 100 fps have been attained using the current 128 x 128 devices.[10] Neither a 512 x 512 MOD device nor the ability to rewrite 2.56 x 10^5 pixels at 100 fps have been demonstrated. In order to attain the 10^{12} ops/sec potentially required for some automatic target recognition application, a 1000 x 1000 pixel device operating at 1000 fps would be required. While this is far beyond current capability, technologies are being developed which have the potential to achieve this level of performance. Since an effective pixel rate of 10^9 pixels/second is also required at the output, an extension of current detector technology would also be required. This could be achieved for example by reading out the CCD using 16 separate channels at 16 MHz per channel. Multi-region readouts are currently being offered by a few CCD manufacturers. This same multi-region technique could also be used to load a DMD device since the mirror deflection is caused by charge stored in wells similar to those in a CCD.

The acousto-optic image correlator currently attains a processing speed of 10^{10} ops/sec and should be able to reach 1.3 x 10^{11} in the near future through the use of recently developed CCD technology in combination with existing AODs. A processing rate of 10^{12} ops/sec would require a higher speed CCD than is currently available; however, the required 1000 point space bandwidth product is available in a variety of high bandwidth AODs.

The major conclusion from Table II is that the acousto-optic approach offers high speed and large image size. Image correlators with throughput rates of 10^{11} or 10^{12} ops/sec are attainable in the near term. This is not the case for the two-dimensional Fourier based correlators which will require significant device development before 10^{12} ops/sec is practical.

3.3 Output dynamic range

Next let us compare the output dynamic range required for both incoherent and coherent optical processing. Once again a smaller M x M reference template is to be correlated against a larger N x N image. First the incoherent acousto-optic correlator will be considered. A bias has been added to the bipolar input signals to make them real positive inputs as indicated in Equations 4 and 5. The incoherent correlator output at the detector plane is given by Equation 6.

In addition to the desired correlation between the functions f and r, there are two bias terms present. The first term is a constant bias, $MT_2B_1B_2$, where T_2 is the time that it takes to read the reference template into the processor. This can be rewritten as $M^2T_{pix}B_1B_2$, where the relation $T_2 = MT_{pix}$ has been used to relate the time T_2 to the reference template size. Thus the magnitude of the constant bias term is directly related to the template size and the bias values added to the two input signals f and r.

The second term in Equation 9 is a signal dependent bias caused by integrating the larger input image, with a duration of $T_1 = NT_{pix}$ over the shorter interval T_2. This value can vary from 0, which implies that f integrates to 0 over the period T_2, up to a value of $A_1B_2M^2T_{pix}$ for the case where the expression A_1f has a constant value of A_1.

The last term in Equation 6 is the desired cross correlation between f and r. To determine an upper bound on the dynamic range required at the detector plane, both A_1f and A_2r will be assigned their maximum values A_1 and A_2 respectively. Thus the maximum value that this term can attain is given by $A_1A_2M^2T_{pix}$. Now the maximum and minimum values of $C(\tau,n)$ when f = r can be written as

$$C_{max} = M^2T_{pix}\,(B_1B_2 + A_1B_2 + A_1A_2) \tag{9a}$$

$$C_{min} = M^2T_{pix}\,(B_1B_2 + A_1A_2). \tag{9b}$$

The signal dependent bias can vary from 0 to a value of $M^2T_{pix}A_1B_2$.

Assuming $B_1 = A_1$ and $B_2 = A_2$, that is the input signal bias is just equal to the maximum amplitude of the time varying portion of the signal, Equations 9a and 9b become

$$C_{max} = 3B_1B_2M^2T_{pix} \tag{10a}$$

$$C_{min} = 2B_1B_2M^2T_{pix}. \tag{10b}$$

In the first case the constant bias, signal dependent bias and desired correlation each provide 1/3 of the output signal amplitude and the presence of bias terms would increase the needed dynamic range at the detector by a factor of three. The most favorable case would be when the input signal was not only zero mean over the input image size but also over the reference template size. Under the assumptions above, this would eliminate the second term in Equation 6 leaving a correlation level and a fixed bias level both equal to $B_1B_2M^2T_{pix}$. Still the use of a bias requires a factor of two increase in the dynamic range at the detector. In actual practice the requirements will be somewhat less than is given by Equation 9a and more than Equation 9b because it is not likely that the second term in Equation 6 will integrate to zero since that would require the input signal, f, to integrate to zero over both time interval T_1 and T_2.

The signal dependent bias term in Equation 6 can also be calculated by correlating the input function, f, with a template that consists of only the bias term B_1. This result can then be electronically subtracted from the previous output to remove the signal dependent bias and provide a bipolar output. While this does increase image quality, it does not reduce the dynamic range requirement on the detector and it requires an extra processing step.

Now let us look at a coherent correlator. The coherent correlator can represent both positive and negative filters in the system. Therefore, no bias need be added to the input image and reference functions. Thus the full dynamic range of the output detector is available to record only the desired correlation between f and r. This is in contrast to the incoherent case where the energy at the output due to bias terms amounts to 1/3 to 1/2 of the available dynamic range in the detector. In addition, if the input function is real and positive as is the case for most sensors such as film, video or IR sensors, the average value of the input scene can be removed by placing a d.c. block at the Fourier plane of the correlator. This results in a sharper correlation peak with a higher peak to sidelobe ratio without the need for a background subtraction operation.

4. EXPERIMENTAL RESULTS

Both a MOD based correlator and a two-dimensional acousto-optic image correlator have been implemented at Sandia. Figure 3 shows a picture of the acousto-optic image correlator.

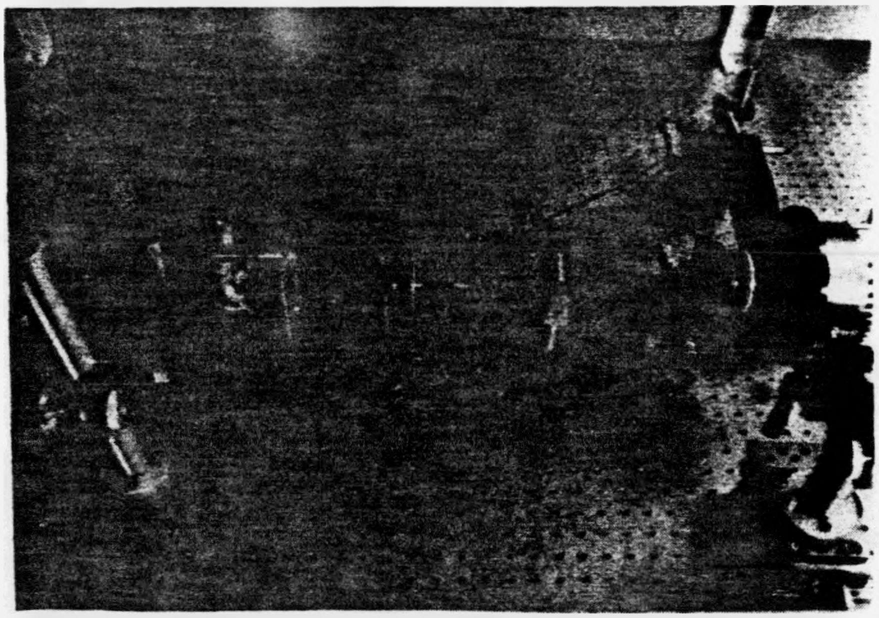


Figure 3. Real-time, acousto-optic correlator

Figure 4a shows a 256 x 232 pixel edge enhanced picture of Willow Run airport obtained using synthetic aperture radar. The reference image, shown in Figure 4b, consists of a 64 x 44 pixel template centered on the intersection of the four runways. The result of the optical correlation is shown in Figure 5. Bias subtraction can be performed to enhance processor performance as previously described.[5] Horizontal and vertical profiles through the correlation peak are

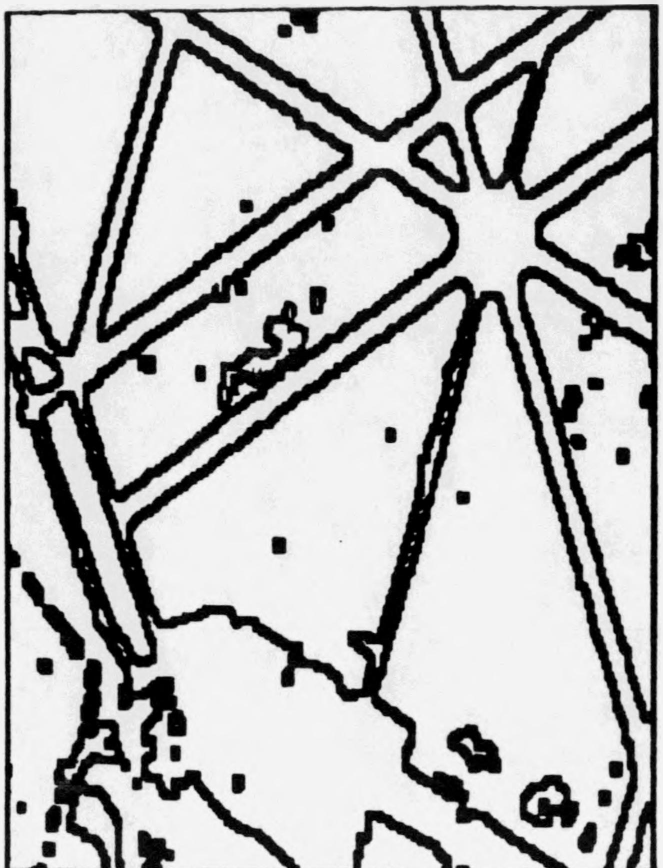
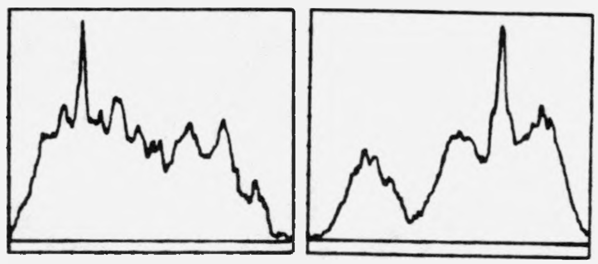


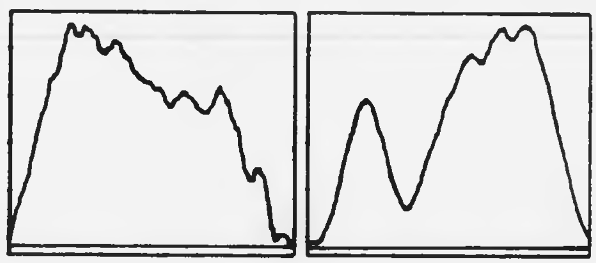
Figure 4 (a) Input image

Figure 4 (b) Reference image

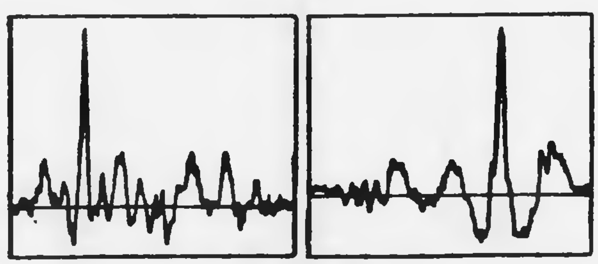
Figure 5. Optical correlation



(a)(b) Correlation peak



(c)(d) Bias term



(e)(f) Bias subtracted result

Figure 6. Correlation profiles

shown in figure 6a and 6b. The optically generated bias terms are shown in figures 6c and 6d at the location of the correlation peak. The profiles through the correlation peak after bias subtraction are shown in figures 6e and 6f. As can be seen from the figure, bias subtraction allows the optical correlator to produce a sharp peak at a location corresponding to the occurrence of the reference template in the input image. This improves the peak-to-sidelobe ratio of the result from 1.5-to-1 in figures 6a and 6b to 3.4-to-1 for the bias subtracted result.

Experimental results were also obtained using the MOD correlator shown in Figure 7. Figure 8a through d show the reference image, input scene, Hartley BPOF and output correlation respectively. Comparison of Figure 8c and 5 show that the lower resolution of the MOD device (128 x 128 vs 256 x 256).

5. SUMMARY AND CONCLUSIONS

This paper has compared two types of real-time optical correlators for optical pattern recognition. They have been compared on the basis of processor throughput and dynamic range requirements at the correlator output. Performance of current systems as well as those capable of being operational in the near term were presented. A major conclusion of the paper is that the acousto-optic correlator offers both high speed and large image size. Acousto-optic image correlators with the 10^{12} operation/sec throughput required in complex pattern recognition problems are attainable in the near term. The two-dimensional Fourier based correlators will require more device development before attaining this throughput. However, the required spatial light modulator technology research is progressing at a rapid pace and devices with the required throughput are in the experimental stages.

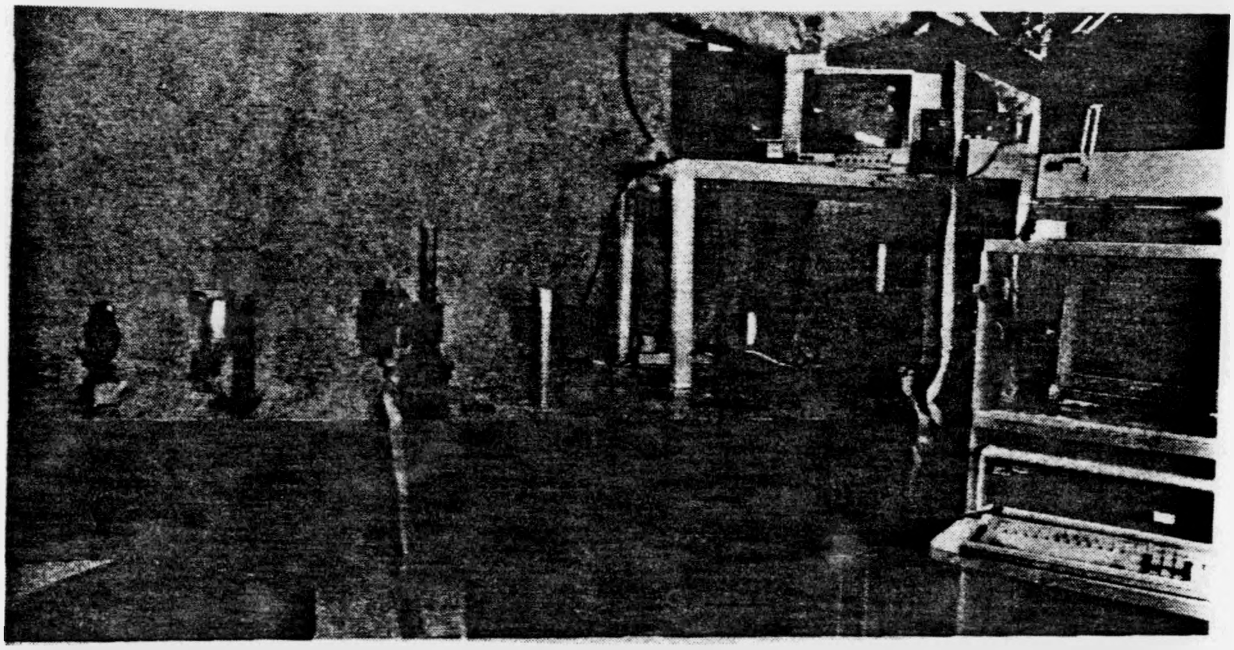


Figure 7. Experimental implementation of Fourier transform based correlator using MOD devices at input and Fourier planes.

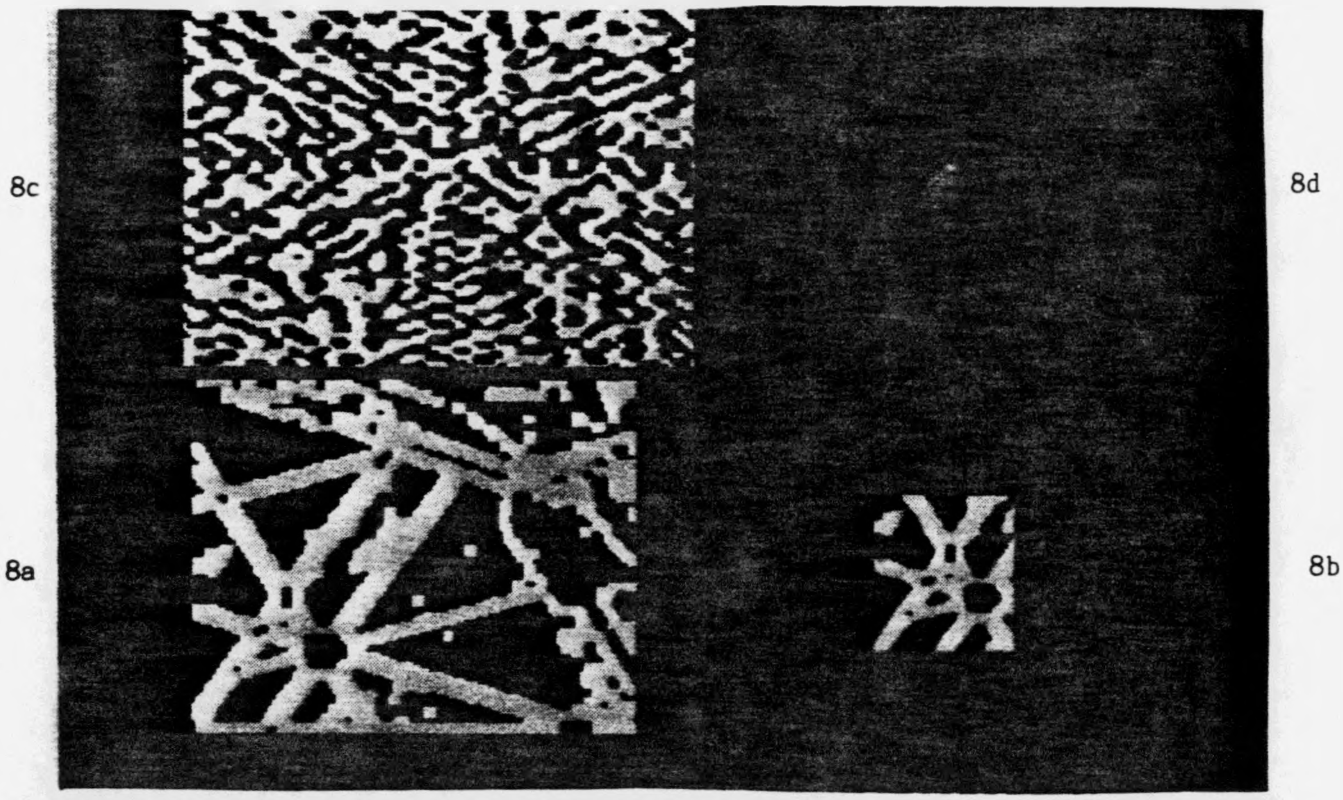


Figure 8. Experimental results obtained using MOD correlator.

The input scene shown in 8a is correlated with the reference image shown in 8b to produce the correlation result in 8d. The Hartley BPOF used at the Fourier plane is shown in Figure 8c.

6. REFERENCES

1. L. J. Cutrona, E. N. Leith, C. J. Palermo and L. J. Porcello, "Optical Data Processing and Filtering Systems," *IRE Trans. on Information Theory*, IT-6, pp. 386-400, 1960.

2. A. B. Vander Lugt, "Signal Detection by Complex Spatial Filtering," *IEEE Trans. Infor. Theory*, IT-10, 1964.

3. A. D. Fisher, "A Review of Spatial Light Modulators," Proc. of the Topical Meeting on Optical Computing, March 18-20, 1985, Incline Village, Nev., TuA2.

4. D. Psaltis, "Incoherent Electro-Optic Image Correlator," *Opt. Engr.* 23(1), pp. 12-15, 1984.

5. P. A. Molley and K. T. Stalker, "A High Dynamic Range Acousto-Optic Image Correlator for Real-Time Pattern Recognition," Proc. SPIE 938, pp. 55-65, 1988.

6. J. P. Lindley, "Applications of Acoust-Optic Techniques to RF Spectrum Analysis," in Acoust-Optic Signal Processing, edited by N. J. Berg and J. N. Lee, Marcel Dekker, Inc, New York, p. 87, 1983.

7. J. L. Horner and J. Leger, "Pattern Recognition with Binary Phase-Only Filters," *Appl. Opt.*, Vol 24, 609, 1985.

8. F. M. Dickey, K. Terry Stalker, and J. J. Mason, "Bandwidth Considerations for Binary Phase-Only Filters," *Appl. Opt.*, Vol. 27, p. 3811, 1988.

9. F. M. Dickey and B. D. Hansche, "Quad-phase Correlation Filters for Pattern Recognition," *Appl. Opt.*, Vol. 28, p. 1611, 1989.

10. B. Kast, Masters Thesis, New Mexico State University, Dec. 1988.