CONF-9004125--3

3-28-90

SAND--89-2383C

DE90 009323

An acousto-optic image correlator with a throughput rate of 1000 templates per second

Perry A. Molley

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185
505-846-5672

ABSTRACT

A two dimensional image correlator based on acousto-optic (AO) and charge-coupled devices (CCDs) is described that can be built with existing technology to provide 1000 frames per second operation. In recent years, architectures have been developed that perform the two dimensional correlation utilizing one dimensional input devices. The input scene is loaded into the acousto-optic device (AOD) one line at time. This line is then correlated against all of the rows of a reference template introduced into the optical system using a one dimensional array of LEDs or laser diodes. However, it generally takes a much greater time to load the AO cell than it does to process the information. This latency time severely limits the maximum throughput rate of the processor. This paper introduces a new acousto-optic correlator implementation that overcomes this bottleneck so that processing can occur close to 100% of the time. A grayscale image correlator is proposed that can be built using present technology that can realistically achieve throughput rates on the order of 10^{12} operations per second. This translates to over 1000 correlations per second for input scenes with dimensions of 512 x 512 pixels and reference templates of size 64 x 64 pixels.

2. INTRODUCTION

Optical pattern recognition based on the Fourier transform property of lenses has been the subject of research for many years.[1,2] However, this technique requires that the data be modulated onto an incident beam of coherent light using a two dimensional spatial light modulator (SLM). Photographic film was initially used as both the input modulator and for holographic storage of the Fourier transform of the reference image. The need for real time recognition of multiple target classes in arbitrary rotations and scales quickly led to the requirement for a two dimensional SLM to rapidly change both the input scene and the reference function. Although much research has been done on new spatial light modulators, no single device has demonstrated the combination of resolution, space-bandwidth product, grayscale modulation, dynamic range and speed of operation required for many real-time image recognition applications.[3,4]

An alternate approach for performing optical pattern recognition utilizes a combination of acousto-optic device, light emitting diode (or laser diode) array, and charge-coupled device technology to perform two dimensional correlations.[5-7] Because of the relative maturity of these devices, acousto-optic and charge-coupled devices have the needed space-bandwidth product and dynamic range to accommodate large grayscale input scenes. These properties combined with the inherent speed of AO cells make them ideal for pattern recognition applications.

The next section of this paper briefly reviews the technique of performing two dimensional correlation utilizing one dimensional devices. Section 4 details a method for calculating the throughput rate of the AO correlator. Section 5 then introduces a new acousto-optic correlator implementation that will allow processing at over 1000 templates per second. In section 6, a different implementation is shown that optimizes the correlator for both input and output pixel clock rates. Section 7 compares all three of the architectures discussed in this paper in terms of output frame rates.

3. ACOUSTO-OPTIC IMAGE CORRELATOR DESCRIPTION

For image recognition, the two dimensional correlation function can be performed as a weighted sum of multiple one dimensional time-integrating correlations.[8,9] An input scene $f(x,y)$ is electronically scanned by a TV camera. The square root of the video signal is taken before hetrodyning it to the appropriate center frequency and introduction into the AOD. At small diffraction efficiencies, the diffracted light intensity is proportional to the input video voltage. As each input scene line propagates down the length of the AOD, it is correlated against all of the rows of a reference template. The reference template, $g(x,y)$ is stored in an electronic memory and is being read out in row parallel to drive an LED or laser diode array as shown in Figure 1. The light from each diode element is collimated and expanded in the horizontal direction to the width of the input scene lines in the AOD. In the vertical direction, the image of the diode array is demagnified so as to fit into the aperture of the cell. Using a Schlieren imaging system, the diffracted light from the AOD is then reimaged onto the rows of the CCD such that each diode is focused onto a different CCD row. Therefore, there is a one-to-one correspondence between the elements of the diode array and the rows of the CCD. Each row of the CCD will contribute a different portion of the 2-D correlation result corresponding to a different row of the template, i.e.,

$$c(\tau,m) = \int_T f(t-\tau)\; g(t,m)\; dt \qquad (1)$$

where $c(\tau,m)$ represents the result from correlating the present input signal in the AOD with the m^{th} row of the template $g(t,m)$. The delay τ is given by $\tau=x/v$ where x is the coordinate along the long axis of the AOD and v is the velocity of propagation in the cell. The integration time T is the duration of the reference template. The correlation in the direction perpendicular to the axis of the AOD is accomplished by operating the CCD in the shift-and-add mode. The two dimensional correlation is performed by shifting the charge in all rows of the CCD down by one pixel position before each new input scene line enters the AOD. The resulting charge that accumulates on the last or M^{th} row of the CCD as a function of horizontal position τ after the completion of the n^{th} input scene line in the AOD is proportional to:

$$c(\tau,n) = \sum_{n'=n-M+1}^{n} \int_T f(t-\tau,n')\ g(t,n'+M-n)\ dt \qquad (2)$$

This charge is then transferred to a serial shift register within the CCD where the signal is read out sequentially. The resulting output from the CCD can be displayed as a function of x (horizontal position) and the input scene line number n. If there is an occurrence of the reference template in the input scene centered at τ_o and line $n=n_o$, a correlation peak would occur on the output of the CCD at $x=v\tau_o$ in the horizontal direction and on the $(n_o - M)^{th}$ row read from the CCD. This represents a delay of M lines before the correlation peak reaches the output of the processor.

4. PROCESSOR THROUGHPUT RATE

In the acousto-optic architecture just described, the throughput rate can be calculated in terms of the reference template and input scene sizes and the input pixel clock rate. Let us assume that a M x M pixel reference template is being correlated against a larger N x N pixel input scene. The time to load the AOD with one line of the input scene is equal to NT_{pix}, where T_{pix} is the period of the pixel clock. Once the information is in the cell, the amount of time it takes to correlate it against the rows of the reference template is MT_{pix}. If the shifting of charge on the CCD for the y-direction processing is overlapped with the loading of the next line of input scene, the time required to completely load and process each line of the input scene, called the AOD cycle time, can be described by:

$$\text{AOD cycle time} = (M + N)T_{pix} \qquad (3)$$

For now let us assume that the CCD output pixel rate is high enough to keep up with this processing rate. In a full 2-D correlation, each pixel out of the CCD represents the sum of M x M products accumulated from M pixels temporally integrated in the x-direction times M shift-and-add integrations in the y-direction. The number of equivalent operations performed on each row of the CCD after correlating all rows of the template with M input scene lines is:

$$\text{Operations/CCD line} = NM^2 \tag{4}$$

where an operation is defined as a multiply-add. The throughput rate of the acousto-optic processor is then given by equation (4) divided by equation (3) or:

$$\text{Throughput Rate} = \frac{N\,M^2}{(M + N)\,T_{pix}} \tag{5}$$

$$= \left[\frac{M}{M + N} \right] NM\, f_{pix} \tag{6}$$

where f_{pix} is the pixel clock frequency equal to $1/T_{pix}$. One important limiting factor affecting the throughput rate of acousto-optic processors is the time it takes to load the AOD. Since no processing takes place while loading the AOD with data, the throughput rate is limited by the ratio of processing time to the AOD cycle time. This is represented by the term in the brackets in equation (6). For an image recognition problem where the input scene dimension N is much larger than the reference dimension M, this ratio is small. Thus the throughput rate is restricted by the large amount of time spent flushing and reloading the AOD. For a typical recognition scenario, the input scene may have dimensions of 512 x 512 pixels and the reference template is of size 64 x 64 pixels. This produces a process-to-AOD cycle time ratio of only 0.125. Hence, only 12.5% of the time is the processor actually performing a correlation operation.

The number of frames per second that can be attained by an acousto-optic processor can be calculated by dividing the throughput rate from equation (6) by the number of operations required per frame. In the case of correlation, the number of operations per frame equals M^2N^2, producing a frame rate of:

$$\text{Frame rate} = \frac{\left[\dfrac{M}{M + N} \right] NM\, f_{pix}}{N^2\, M^2} = \frac{f_{pix}}{(M + N)\, N} \qquad (7)$$

To achieve higher throughput rates for a given size reference and input scene, the pixel clock must operate at greater clock frequencies. However, when the clock frequency is increased, the bandwidth requirements of the AOD are also increased. Currently available slow shear TeO_2 AODs can accommodate pixel clock frequencies of 50 MHz, or $T_{pix} = 20$ ns. From equation (6), a throughput rate of 1.82×10^{11} operations/second can be achieved at this clock rate. From equation (7), this represents 169 frames/second that can be processed. Faster clock rates could be used in order to reach our desired goal of 1000 frames/second, but other AOD material such as $LiNbO_3$ must be used in order to accommodate the greater bandwidths. However, other materials have lower diffraction efficiencies than do the slow shear TeO_2 devices, some as much as 100 times less per watt.[10] This can dramatically decrease the amount of light energy falling onto the detector. As shown by Molley and Stalker[9], diminishing the irradiance to the CCD can result in a correlation peak that falls below the sensitivity of current detector arrays. Furthermore, there is a limit to how fast the clock rate can be increased and yet still maintain a high dynamic range due to electrical crosstalk considerations. For our example, the pixel clock frequency would have to reach almost 300 MHz to attain a frame rate of 1000 correlations per second. This exceeds the practical limitations on the maximum operating speeds for memory chips and digital-to-analog converters used to generate the template data for the system. Especially in embedded system applications, clock frequencies greater than 100 MHz are prohibitive since they require emitter-coupled logic (ECL) technology. ECL circuits generally take up more space, need to be cooled, and require larger power supplies than their lower speed CMOS or TTL counterparts.

5. NEW PROCESSOR IMPLEMENTATION FOR HIGHER THROUGHPUT RATES

The new architecture proposed by this paper suggests a way of achieving a very high throughput rate without relying on extremely high pixel clock frequencies. This architecture accomplishes the task by better utilizing the the information in the AO cell. Each time a new input scene line enters the cell, multiple templates are correlated with this data before it leaves. One way of implementing this is shown in Figure 2. Each diode element is imaged onto a different row of the charge-coupled device. However, these rows are separated by a number of other CCD rows which are shielded from the illumination by a mask. When an input line first enters the AO cell, it is correlated against all of the rows of template #1, the first set of information in the electronic memory to drive the diode array. This correlation is summed by the CCD rows exposed through the open spaces in the mask. Next, all of the charge in the CCD is shifted up one pixel position. The shifting of charge can be accomplished quickly enough that the input data in the AOD moves only a few (2 to 3) pixels. Template #2, the next set of data in the memory, is then correlated against the information in the AO cell. These individual correlations are then summed on the next exposed rows on the CCD. In this manner, the data in the AO cell is reused by more than one template with the result being formed on consecutive rows of the CCD. The number of CCD rows between the open areas of the mask correspond to the number of templates used, P, and hence the number of times the data in the AO cell is reused. After P correlations and P electronic shifts for the CCD, all of the correlation information for template #1 is again exposed to the illumination through the open areas in the mask ready to accumulate another part of the summation for the 2-D correlation described by equation (2). After correlation with M rows of reference template #1, the result is read from the CCD. Subsequently, the next (P-1) consecutive lines read from the CCD are formed by correlating the same input scene lines, (n-M+1) through n, with templates 2 through P.

Figure 3 shows this architecture from a slightly different perspective. The top rectangle in the AO cell represents an entire input scene line which has just been loaded; this is listed at time t1. For the first template, the data propagates to the left as the correlation process is taking place. This correlation produces charge in the CCD which is horizontally positioned as shown by the top solid rectangle on the CCD. This charge is then shifted up so that the next correlation can occur. For template #2, the position of the data in the AO cell is as shown at position labelled t2. This causes the correlation result to be shifted by the same amount on the CCD. However, because all of the intermediate correlation products on each CCD row for template #2 are shifted by the same amount, the y-direction processing is still correctly accomplished for each template. The position of the final correlation result for each template will be shifted to different segments on the CCD; however, this offset is known a priori. Therefore when the results are read from the CCD, only part of each CCD row will contain valid correlation information.

By making the processing time for the multiple templates larger than the load time for the AOD, the load time for the next input scene line can be overlapped with the processing time of the current line. The dead time, the time spent loading the AOD when no processing is taking place, can be reduced to zero thereby dramatically increasing the throughput rate of the processor. To meet this requirement, the processing time for all of the templates must be greater than the load time for an input scene line. The CCD and the AOD must both be long enough to handle the length of an input scene line plus the length of propagation for processing all of the templates. However, currently produced CCDs and AODs have the capability to accommodate at least 1024 pixels in the horizontal direction allowing input scene widths of 512 pixels. Hence, as soon as the one input scene line has been completely loaded in the AOD, the load for the next consecutive input scene line can be immediately started while the previous line is propagating down the cell. When the current input scene line has propagated down the entire length of the AO cell and has been processed by the different templates, the next line of data will have just finished being loaded. Therefore, it is possible to build an acousto-optic correlator that can process information for almost 100% of the cycle time (less the small time for shifting the charge up one pixel position after correlation with each template). Since the load time for the AOD is now overlapped with the processing time, the factor "N" in the denominator of equation (5) will be zero. If all of the charge in the CCD can be shifted by one pixel position in one pixel clock period, T_{pix}, then the AOD cycle time for one input scene line equals the time to process the line, MT_{pix}, plus the vertical shift time of the CCD pixels, T_{pix}. Equation (6) can then be rewritten for this new "high throughput" architecture as follows:

$$\text{Throughput Rate} = \left[\frac{M}{M + 1} \right] NM\, f_{pix} \tag{8}$$

Similarly, the frame rate for this new architecture is:

$$\text{Frame Rate} = \frac{f_{pix}}{(M + 1)\, N} \tag{9}$$

Compared to the original architecture, this represents a speedup factor of $(M + N)/(M + 1)$. For our example of a 512 x 512 pixel input scene and a 64 x 64 reference template, the new implementation can process information 8.86 times faster. Hence, the optical processing section can attain a frame rate of over 1502 correlations per second with presently available 1-D input devices operating at a pixel clock frequency of 50 MHz.

6. BALANCED INPUT-OUTPUT HIGH THROUGHPUT IMPLEMENTATION

To achieve the high frame rates just discussed, some consideration must be given to the CCD output rate needed to keep up with this processing speed. Regardless of how fast the input devices and processing section of the correlator can operate, if the CCD becomes the bottleneck of the system, the output pixel rate will determine the ultimate throughput rate for the correlator. For the high throughput architecture shown in figure 3, the CCD rows are 1024 pixels long even though only 512 pixels contain valid information. Therefore, 1024 x 512 pixels must be read out of the CCD for each correlation plane, or 524,288 pixels per frame. If the processor operates at 1000 frames/second, the CCDs equivalent output rate would need to exceed 524 x 10^6 pixels/second. This would require the CCD to have 32 parallel output channels at 16.4 Mpixels/second per channel in order to keep up with the processing section of the optical processor. This also requires a large amount of signal processing electronics sample and detect the resulting correlation peaks for each of the 32 channels of analog information.

To relieve this tremendous burden on the back end of the system, a modification can be made to the high throughput correlator architecture in order to balance the input data rate and processing speed with the output rate of the CCD and any post-processing electronics required. This consists of reintroducing the each input data line in the AOD so that all of the AOD is used and all of the CCD pixels are utilized for valid correlation results. This modified architecture will be referred to as the "balanced input-output high throughput" correlator. A diagram for this architecture is shown in figure 4. Similar to the previous design, this system contains a CCD mask to expose only certain rows to the illumination of the diodes. The rows between these exposed areas act as storage areas for subsequent correlation results of other templates. However, for this architecture, the AO cell and CCD are only as long as the input scene lines. The top bar in the AOD, listed at time t1 in the figure, represents the position of the input scene line for the start of the first template. This information propagates down the cell and is correlated against all of the rows of the first template. As this information is being correlated, a copy of the same input scene line is started directly behind it. After the correlation with the first template is completed, all of the charge in the CCD is shifted up by one pixel position. At this point, listed as time point t2 in the figure, a portion of input scene line has travelled out of the cell. Therefore, the second template is correlated with only part of the input scene line from the first loading. However, since the same input scene line is being reloaded at the beginning of the cell, a complete set of data for the entire input scene line will always be in the AO cell. The resulting correlation for the second template will be formed in two segments on the CCD according to the relative position of where the start of the input scene line is in the cell. The correlation results for all but the first template appear to "wrap-around" the ends of the CCD as shown by the broken bars on the CCD in Figure 4. The distance that each correlation result is horizontally shifted on the CCD for each template can be calculated. The order of the templates determines how far the start of the input scene line had propagated before correlation is performed and therefore where the correlation result will occur on the CCD for each template. So for each line read from the CCD, if a correlation peak detected at position $x-v\tau_o$, it must be shifted by a factor of $P_i(M + 1)$ pixels [modulo 512] to locate the actual location of the object in the input scene, where P_i indicates the template number being correlated for that particular CCD row. This can be easily accomplished by adding an offset to a counter that keeps track of the pixel positions within each row.

The advantage of this type of implementation is that all of the pixels on the CCD are utilized for valid information. The number of pixels per correlation plane is half of that for the previously describe architecture, only 262,144 pixels per frame, thereby reducing the required CCD output pixel rate. This is accomplished without severely affecting the throughput rate of the processor since the AO cell must now be completely flushed and reloaded only after an entire set of templates have been correlated. Furthermore, the width requirement for the CCD as well as for the AOD is reduced to that of the original architecture.

One factor affecting the throughput rate of this architecture is the number of templates used before needing to completely flush and reload the AO cell with the next consecutive line from the input scene. The greater the number of templates used before flushing the cell, the higher the throughput rate of the correlator. The maximum number of templates that can be used per input scene line is limited by the maximum number of rows possible on the CCD. For currently available CCDs, a realistic number of rows is 1024. This allows for 16 templates of size 64 x 64 pixels to be correlated with each input scene line before flushing the cell. The AOD cycle time for this architecture is equal to:

$$\text{AOD cycle time} = (N + PM + P)\ T_{pix} \tag{10}$$

where P is the number of templates used per input scene line. The throughput rate is now:

$$\text{Throughput Rate} = \left[\frac{M}{N + PM + P} \right] PNM\ f_{pix} \tag{11}$$

The factor P in the term outside the brackets is due to the fact that P templates are correlated per AOD cycle time. This factor helps offset the additional factor of N in the denominator signifying the need for a reload time. The frame rate is:

$$\text{Frame Rate} = \frac{P f_{pix}}{(N + PM + P)\,N} \qquad (12)$$

This produces a frame rate which is faster than the original implementation yet better utilizes the CCD. For example with a pixel clock frequency of 49.6 MHz, the balanced high throughput architecture can achieve a frame rate of 1000 frames per second. At this rate, a 512 x 1024 pixel CCD would need 16 parallel output channels each running at 16.4 Mpixels per second to keep up with the processing rate of the AO cell, a goal within the reach of present CCD technology.

7. Throughput rate comparisons

Table 1 compares the frame rates of the three architectures discussed in this paper for various input pixel clock frequencies. For these comparisons, we have assumed an input image of size 512 x 512 pixels and a reference template with dimensions of 64 x 64 pixels.

Table 1

Output Frame Rate Comparison

Input Pixel Clock Frequency (MHz)	Original Architecture (frames/s)	High Throughput Architecture (frames/s)	Balanced Input-Output High Throughput Arcitecture (frames/s)
10	33	300	201
30	101	901	604
50	169	1502	1006

As can be seen from Table 1, both high throughput architectures can achieve over 1000 correlations per second with a clock frequency that does not exceed the bandwidth limitations of present slow shear TeO_2 AO cells. The advantage of the "balanced input-output high throughput architecture" is that it requires fewer output channels on the CCD at the expense of needing a faster input clock frequency. However, as can be seen from the table, the "balanced" architecture has a frame rate that is 50% slower than the high throughput architecture introduced first.

8. CONCLUSIONS

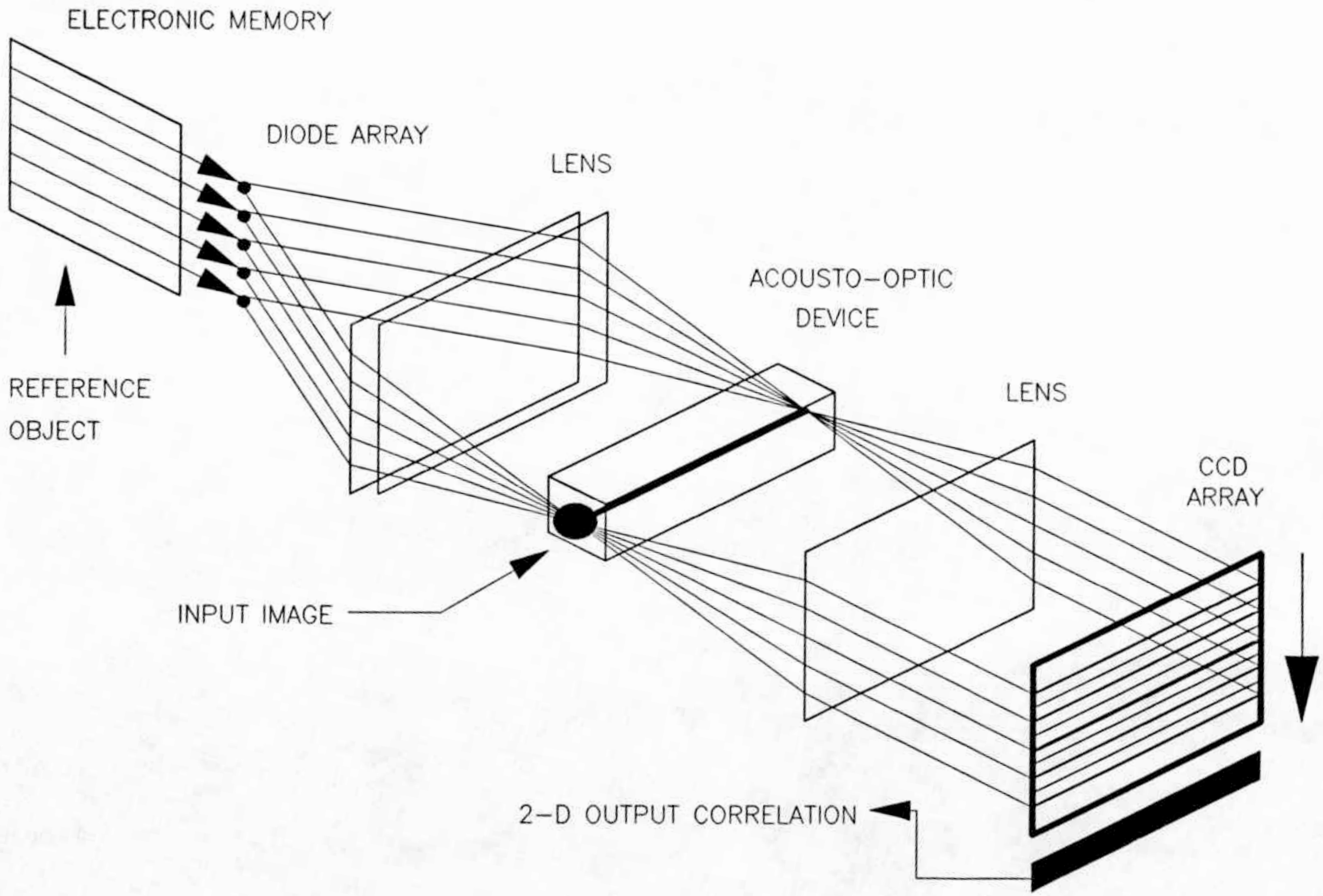
This paper introduced two new architectures that can attain over 1000 correlations per second for reference templates of size 64 x 64 pixels and input scenes of size 512 x 512 pixels. These architectures allow for the high speed necessary to detect objects that may be at arbitrary orientation (scale or rotation) in the input scene or provide discrimination capability between different classes of objects. The two proposed architectures utilize the fact that many image recognition problems require a number of different templates to be compared against the same input scene. Therefore, a larger percent of the cycle time can be spent processing the data instead of simply reloading the AOD. However, to operate at 50 MHz, the brightness of a laser diode array is needed in order to insure the resulting grayscale correlation is above the detector noise level. A high throughput optical processor, whether Fourier or acousto-optic-based, would also require a multi-channel CCD to read out the correlation result on parallel channels. The technologies presently exist to construct these needed devices. Hence, with these new acousto-optic implementations, there exists the ability to build an optical processor having a throughput rate of over 1000 correlations per second in a compact, light weight, power efficient package. This acousto-optic processor could fill an important need in automatic target recognition applications.

9. ACKNOWLEDGEMENTS

This work performed at Sandia National Laboratories and supported by the U.S. Department of Energy under contract DE-AC04-76DP00789.

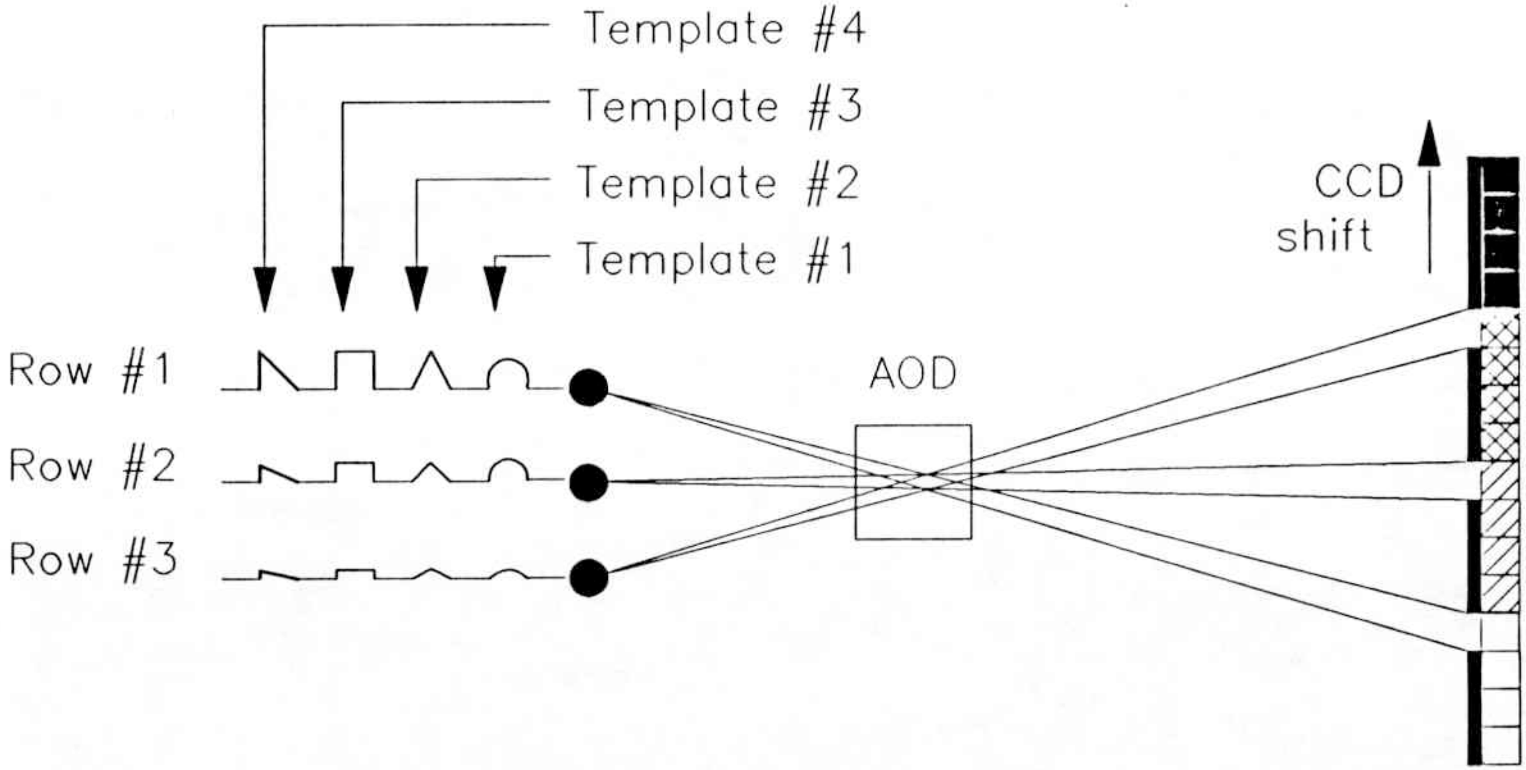
10. REFERENCES

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. Vander Lugt, "Signal detection by complex spatial filtering", IEEE Trans. Inform Theory, vol. IT-10(2), pp. 139-145, (1964).
3. K. T. Stalker, P. A. Molley, B. D. Hansche, "A comparison of real-time optical correlators for pattern recognition", Intern. Congress on Opt. Science and Engr, Proc. SPIE, Vol. 1134, April 1989.
4. A. D. Fisher, "A review of spatial light modulators", Proc. of the Topical Meeting on Optical Computing, TuA2, (1985).
5. D. Psaltis, "Two-dimensional optical processing using one-dimensional input devices", Proc. IEEE, 72(7), pp. 962-972, (1984).
6. D. Psaltis, D. Casasent, "General formulation for optical signal processing architectures", Opt. Eng. 19(2), pp. 193-198 (1980).
7. W. T. Rhodes, "Acousto-optic signal processing: convolution and correlation", Proc. IEEE 69(1), pp. 65-79 (1981).
8. D. Psaltis, "Incoherent electro-optic image correlator," Opt. Engr. 23(1), p. 12-15, (1984).
9. P. A. Molley and K. T. Stalker, "A high dynamic range acousto-optic image correlator for real-time pattern recognition", Proc. SPIE vol. 938, pp. 55-65 (1988).
10. Joseph L. Horner, **"Optical Signal Processing"**, San Diego, CA: Academic Press Inc., (1987).



OPTICAL IMAGE CORRELATOR

Figure 1. Optical image correlator



HIGH THROUGHPUT CORRELATOR ARCHITECTURE

Figure 2. High throughput correlator architecture - side view

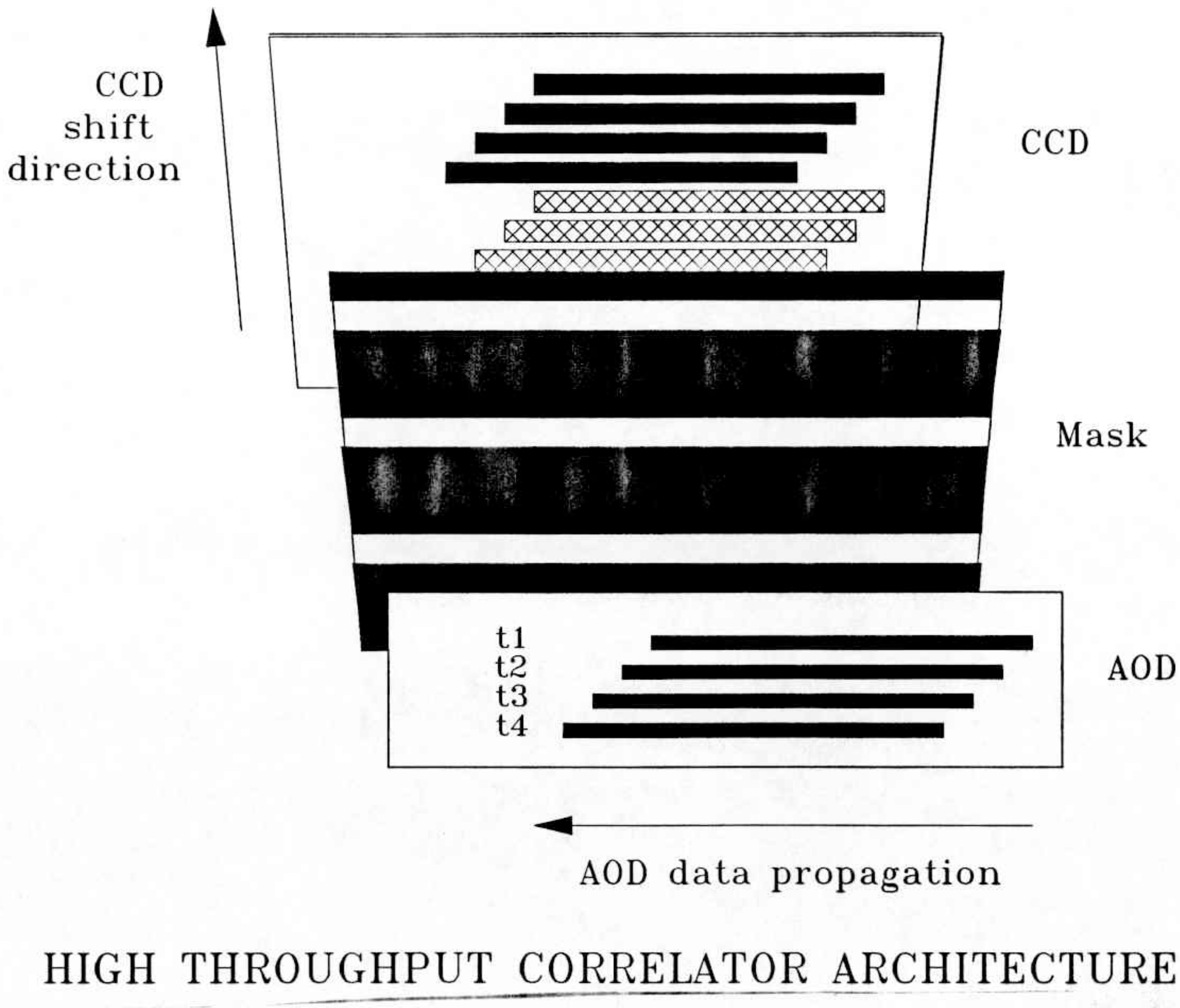


Figure 3. High throughput correlator architecture - top view

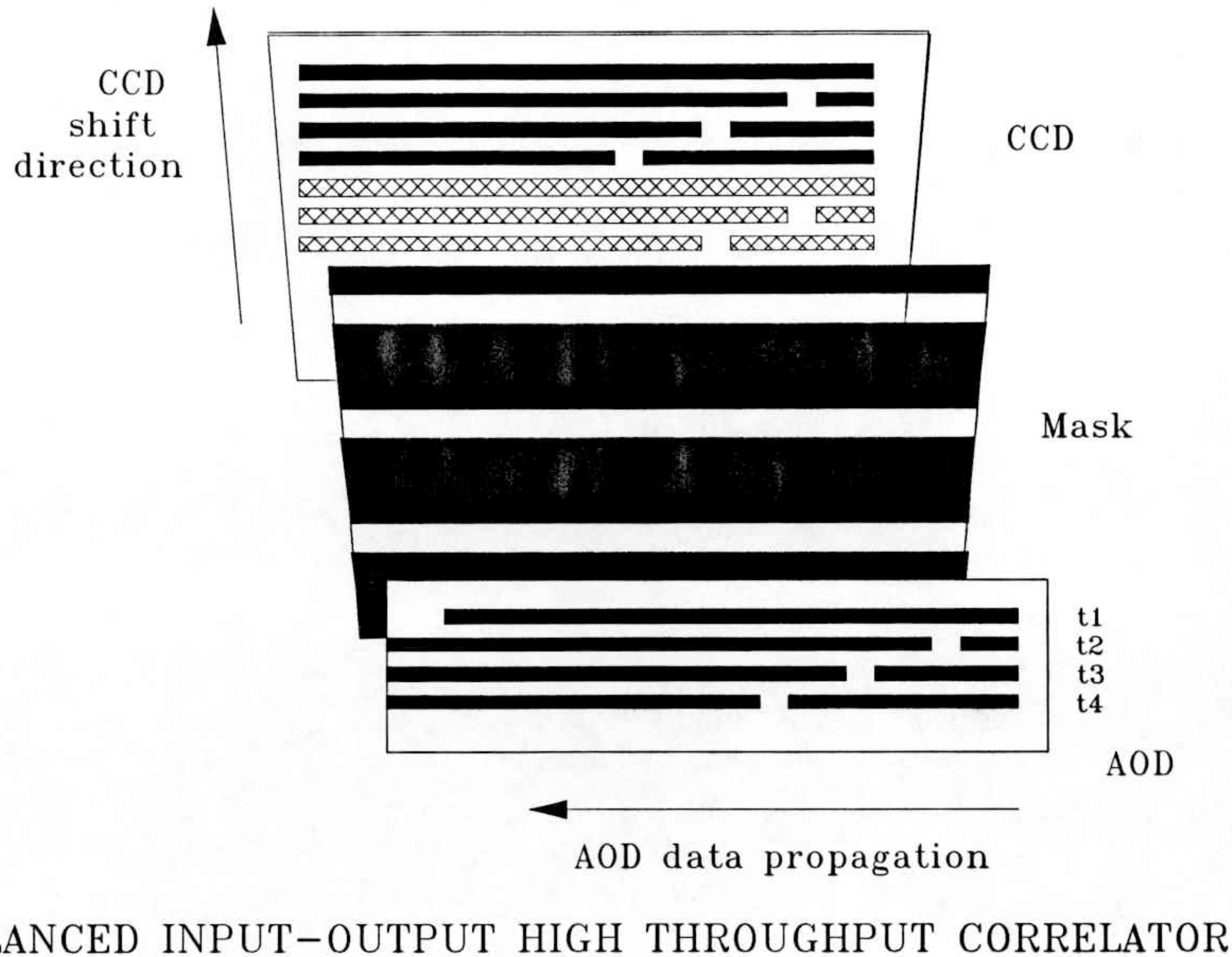


Figure 4. Balanced input-output high throughput correlator