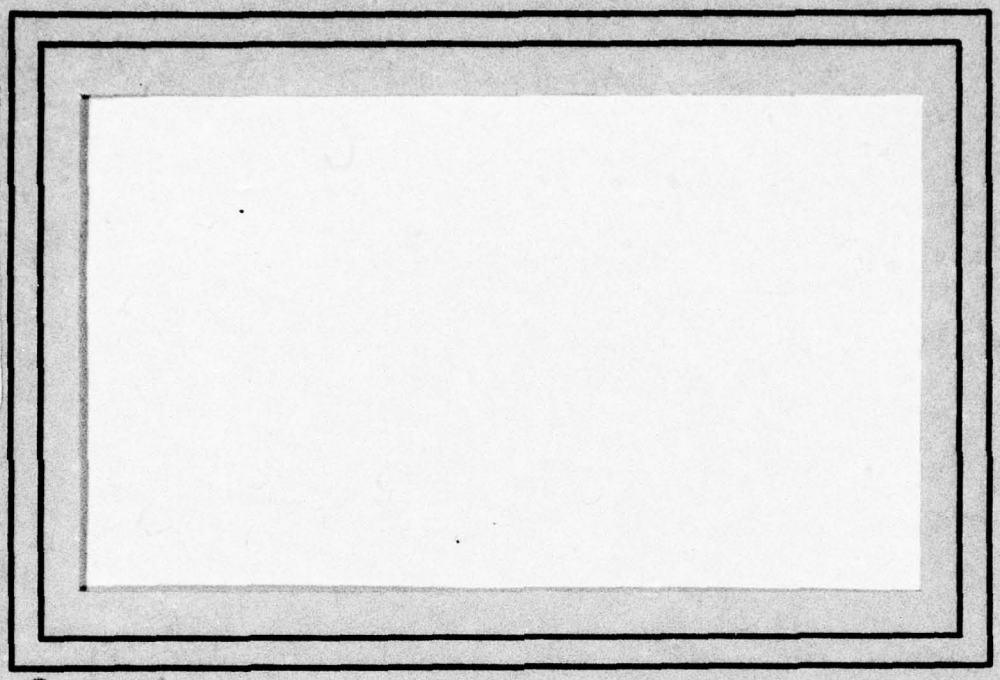


COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

ALGORITHMS AND HARDWARE TECHNOLOGY

FOR IMAGE RECOGNITION

Quarterly Report
on Westinghouse Subcontract

1 May-31 July, 1976

Contract DAAAG53-76C-0138
(DARPA under 3206)

A DISCUSSION OF DESIGN GOALS AND HARDWARE IMPLEMENTATION FOR AN AUTOMATIC TARGET CUEING SYSTEM

July 30, 1976

This is the first quarterly status report on a program for Recognition Technology for a Smart Sensor, conducted by Westinghouse for Maryland under Contract DAAG53-76-C-0138 with the U.S. Army Mobility Equipment Research and Development Command, Ft. Belvoir, Va. 22060

Prepared for

Computer Science Center
University of Maryland
College Park, Maryland 20742

By

Westinghouse Defense and Electronic Systems Center
Systems Development Division
Baltimore, Maryland 21203

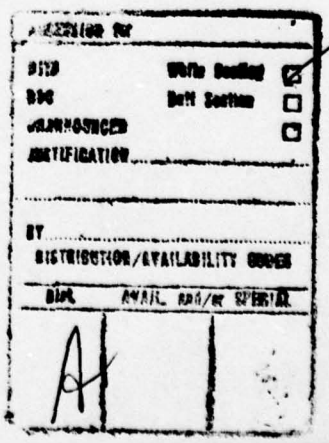


TABLE OF CONTENTS

next page

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

1.0 INTRODUCTION

This is the first quarterly status report on a program for Recognition Technology for A Smart Sensor, being conducted by the Westinghouse Systems Development Division for the Computer Science Center, University of Maryland. The program contains three phases, as follows:

Phase I Task and Technology Review (3 months)

Phase II Algorithm Selection and Test (9 months)

Phase III Hardware Development (9 months)

This report covers the Phase I effort. In addition, it describes initial activity on algorithms implementation which is a part of Phase II. The report was prepared by Mr. Tom Willett of Westinghouse. The Westinghouse program manger is Dr. Glenn E. Tisdale.

During the quarter, eight meetings were held between members of the Maryland and Westinghouse teams. Mr. John Dehne, NVL Program Manager, maintained close contact with the program, and attended some of the meetings. The meetings included a description by Maryland of the nature of the algorithms which are under consideration for the cueing function, together with preliminary test results, presentations by Westinghouse on the nature of automatic cueing, and the limitations involved in the hardware implementation of cueing algorithms. This material is summarized in this report.

During the quarter, Westinghouse supplied to Maryland a taped data base of digitized forward-looking infra-red images, together with ground truth, for use in testing algorithms. The taped data base had been furnished Westinghouse by the Night Vision Laboratory of the U. S. Army Mobility Equipment Research and Development Command under another contract.

1.1 DEFINITION OF THE AUTOMATIC TARGET CUEING FUNCTION

Before discussing the design goals and hardware implementation for an automatic target cueing system, it is useful to examine its character and purpose.

As the tools used in reconnaissance and target acquisition operations continue to improve, it is becoming increasingly clear that man himself is a performance bottleneck. Sensors provide greater resolution, dynamic range, and speed; and the techniques that make effective use of sensor information, such as communication links, and weapon delivery systems, are also being advanced. However, acquisition of target material largely depends on the ability of a human operator to scan sensor imagery in real time. If he misses targets, or identifies them incorrectly, mission performance is degraded.

One way to improve this situation is to provide the human with cues to locate, and possibly identify, targets on his displays. Such cues might originate with a priori information, or from the detection of peculiar target conditions such as motion, hot spots, or electromagnetic radiation. However, a more general approach is to apply automatic recognition techniques to the imagery. As shown in Figure 1-1, the image processor serves as an information filter on the image, alerting the human to the presence of potential targets, possibly by audible signals initially, and then by providing visual cues or overlays on his display. In this process, the final verification of target identity is reserved for the human operator. The image processor serves to assist him.

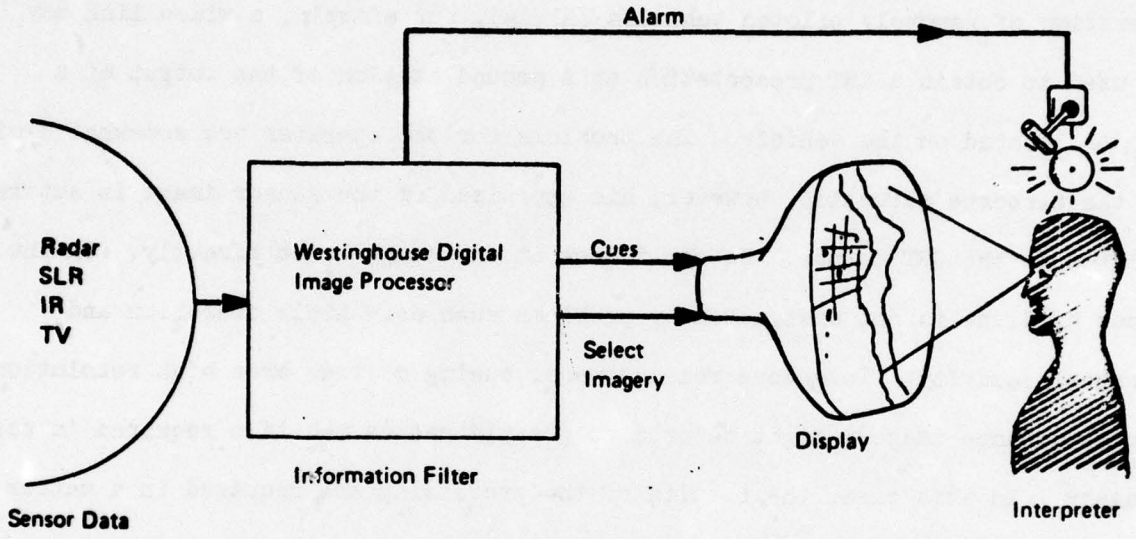


Figure 1-1. Information Extraction and Automatic Cueing

Automatic cueing can be carried out either in airborne or ground locations. In the airborne situation, the operator views a CRT-type display for acquisition of targets on a real-time basis. His determination may result in action in a matter of seconds, either offensive or defensive. On the ground, interpretation may be required in real-time, or on a more relaxed basis. In the proposed operation of remotely piloted vehicles (RPV's), for example, a video link may be used to obtain a CRT presentation at a ground station of the output of a sensor located on the vehicle. The problems for the operator are somewhat similar to the airborne situation; however, his appraisal of the sensor image is entirely limited to the CRT output. He can't look at the target area directly. On the other hand, he is not distracted by problems such as vehicle operation and personal security. In a more relaxed mode, cueing of wide area high resolution reconnaissance imagery might be used to greatly reduce the time required in scanning imagery. In this case, the results of the processing are required in a matter of minutes, so as to respond to the detection of mobile targets.

1.2 CONTENTS OF THIS REPORT

The report will proceed in Section 2.0 with a discussion of design goals for a cueing system. The goals will depend to an extent on the nature of the missions involved; accordingly, some typical missions will be listed. Next, performance goals will be considered, including the desired detection rate for targets, the acceptable false alarm rate, and the required speed of operation. The physical characteristics of the system will then be examined, followed by a discussion of allowable cost.

Section 3.0 will provide a general discussion of the constraints imposed on the cuer algorithms by requirements for high-speed, low cost hardware.

Section 4.0 will introduce the approach to the implementation of specific algorithms in circuit form. This work is a part of Phase II of the program, and will continue.

2.0 SYSTEM DESIGN GOALS

The key considerations in the design of an automatic target cueing system are its performance, physical characteristics, and cost. We will deal with each of these areas in this section. However, a quantitative determination of design parameters will depend on the manner in which the mission is implemented. Such implementation will be discussed first.

2.1 MISSION IMPLEMENTATION

As explained in Par. 1.1, the target cueing function might be performed aboard a vehicle, or at a ground station if imagery is relayed for analysis. In either case, the performance goals will tend to be comparable. As regards physical characteristics and cost, however, the vehicle location will provide much tighter restrictions. Our discussion will proceed on the basis of the vehicular application. Both helicopters and high-speed aircraft are airborne candidates. The RPV image, on the other hand, will be analyzed at a ground station; therefore the physical limitations within the RPV are not a problem. As the state-of-the-art in automatic target recognition develops, and high levels of performance are attained, it is anticipated that the human observer will eventually be eliminated in some applications. For example, recognition equipment might be placed aboard a missile for unaided terminal guidance. The requirement for high performance, small size and weight, low power consumption, and low cost will all apply in this case.

2.2 PERFORMANCE GOALS

The key performance parameters are the detection and recognition rates for targets of interest, the false alarm rate, and the speed of operation of the cueing system. A discussion of these parameters follows.

2.2.1 Detection and Recognition Rates

Detection of a target occurs when a target indication is signalled by the cueing system in response to the presence of a target of interest. Correct recognition occurs when the appropriate target class is selected from among several possible classes. Detection and recognition performance is expressed as a percentage of the targets actually present. Therefore, in testing the performance of a cueing system, it is important that "ground truth" be available for the test data base.

What levels of detection and recognition are required in order that the cueing system will be cost effective? Clearly, its use should provide a significant increase in the capability of the human interpreter. First, consider the interpreter's situation. His task is to locate and identify objects, such as vehicles or boats, that generally occupy a small area in the sensor field of view, and that may be imbedded in background clutter and noise. If there were no clutter or noise, the need for cueing would largely vanish. There is an advantage to target acquisition at maximum range, which implies that recognition will occur with a minimum of target resolution. In fact, attempts must be made to infer the nature of "blobs" on the basis of contextual information. The human interpreter of reconnaissance imagery is limited in several ways:

- His ability to serach imagery is limited by foveal vision and reaction time.
- He is subject to eye fatigue.
- His performance is degraded by the demands of other tasks and by his concern for platform navigation and safety.

What performance do humans actually provide? Although data are limited, it would appear that it is fairly low. Evidence is accumulating[1,2] that even under close scrutiny and laboratory conditions less than half of available targets are detected or recognized by trained interpreters. In the aircraft environment, the situation is evidently worse, more like 33% detection and 25% recognition.[3]

Consider, now, automatic recognition by image processing. Image processing and recognition machines offer several advantages:

- They can be wide band systems, capable of examining image elements at megasample per second (TV) rates;
- They may be designed to operate on sensor video signals, thereby avoiding loss of resolution and dynamic range caused by displays;

---

1. Jeffrey. T.E., Image Interpretation Research, from "Pattern Identification by Man and Machine," U.S. Army Tech. Memo 17-68, Human Engineering Laboratories, Aberdeen Proving Ground, Md., 12 December 1968, P. 35.
2. Evans, S.H. and Dansereau, D.F., Relating Man to Automatic Image Interpretation: It Pays to Know Your Competitor, paper presented at the EIA Symposium on Automatic Photointerpretation and Recognition, Washington, D.C., December 7, 1971.
3. Airborne Stano Systems Part II Test Report (U), U.S. Army MASSTER, Ft. Hood, Texas, 76544, July, 1971.

- They are expendable, capable of being stockpiled, and don't require training.
- They are not subject to fatigue or distractions.

What kind of performance may be expected from the machines? Based on preliminary test results, it appears reasonable to expect greater than 80% detection and 60% recognition of targets in clutter[4] under comparable conditions. These figures are more than double the figures for human performance stated above. This would indeed provide a good margin for cost-effective performance.

2.2.2 False Alarm Rate

A false alarm occurs when an indication of a target is signalled by the cueing system when, in fact, no target is present. The false alarm rate is expressed on the basis of a unit of elapsed time or area of coverage.

Where several different target classes are involved, it is possible to think about a quantity, P_{det}, that represents detection probability averaged over all the target classes of interest; P_{fa} is the false alarm counterpart. P_{fa} is based on one frame covering the whole field of view.

The false alarm rate[5] which the human operator must combat is peculiar to video displays. Since most false alarms are due to ground clutter, not noise, they will reappear in every frame until the scene changes. Once the operator identifies a cue as a target, it can be ignored thereafter. Or it can be eliminated after operator identification with a cursor. Consequently, the rate at which new data (with new false alarms) appears is the rate at which a totally new image

---

4. Final Report, Demonstration of Westinghouse Automatic Cueing Techniques Using NVL Images (U), Westinghouse Defense and Electronic Systems Center, Document No. 9780, May 1976.

5. Tisdale, G.E., Automatic Cueing of Reconnaissance Imagery - Performance vs. Cost, paper presented at the Symposium on Computer Image Processing and Recognition, University of Missouri, Columbia, August 24 - 26, 1972.

is in the field of view. Regardless of whether the application involves helicopters, high speed aircraft, missiles or RPV's, the geometry shown below is approximately representative. The platform is at altitude h, looking out at a

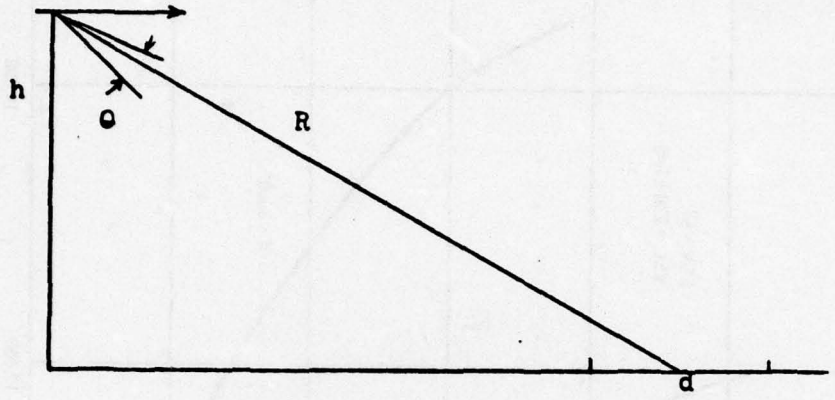


Figure 2-1. Downlook Geometry

slant range R, and having an angular field of view θ, and the area observed is d feet long. When the platform has travelled forward by d feet, a whole new image will be viewed. If the platform has a forward speed of v, the time between new frames is

$$T = h/v \{\tan[\cos^{-1}(h/R) + \theta/2] - \tan[\cos^{-1}(h/R) - \theta/2]\}$$

and the effective false alarm rate becomes

$$FAR_{eff} = P_{fa} \times \frac{1}{T} \text{ false alarms/sec.}$$

A graph of T is shown in Figure 2-2; the kind of mission considered will place limits on h, r, v, and T.

Mr. Donald Looft, Deputy Director of DARPA, at the sixth annual Automatic Imagery Pattern Recognition Sumposium of the EIA, in June of 1976, suggested

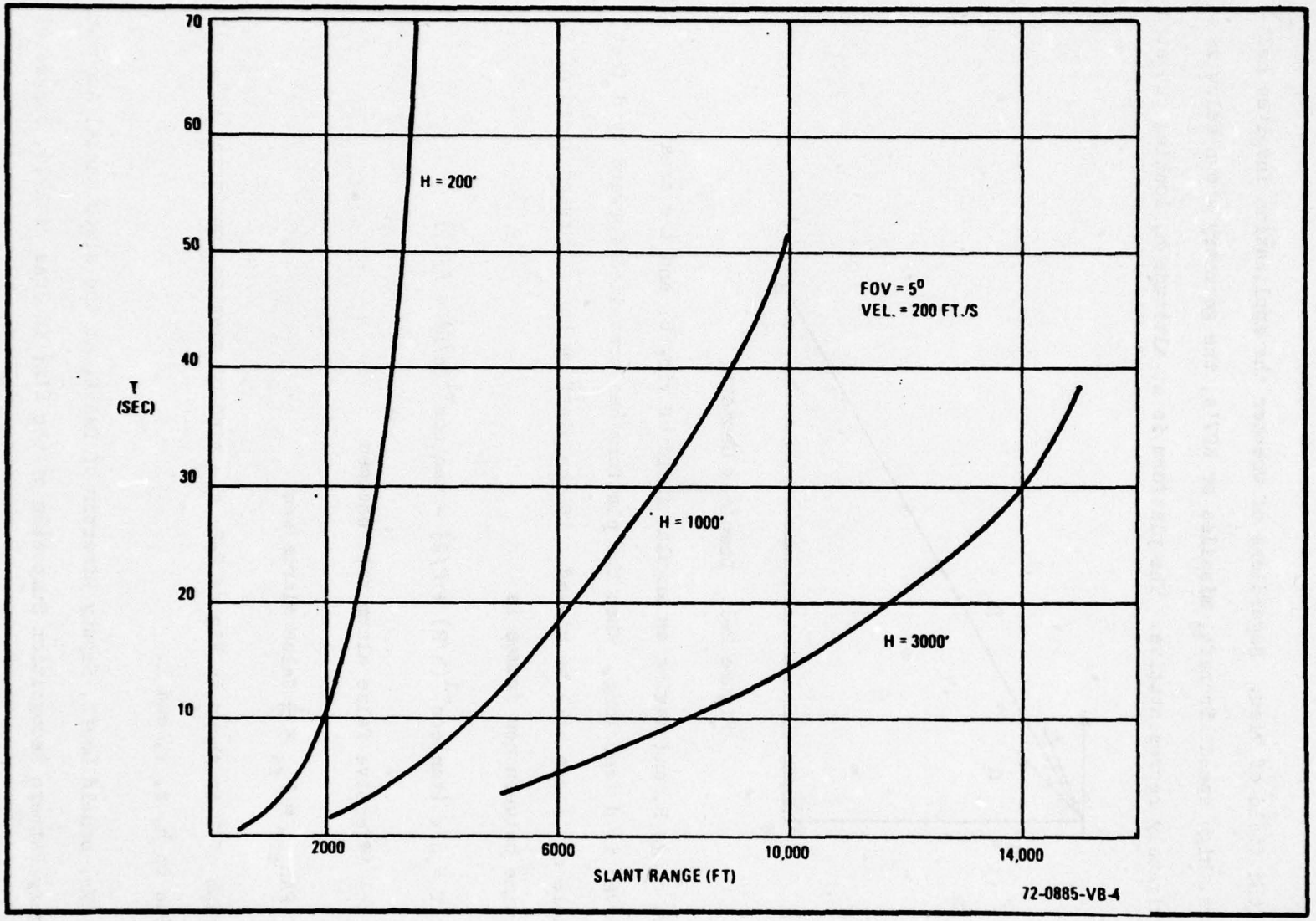


Figure 2-2. Time Between New Images vs. Sensor Slant Range

several scenarios during his banquet speech. The particular examples extracted for this project cover helicopters and high speed aircraft with airborne infrared imaging systems.

"Now visualize a likely scenario for the one-man pilot observer. He must direct the sensor and visually scan his display for targets which most assuredly are concealed by terrain or background clutter. He must fly low and as fast as possible to minimize exposure to radar controlled weapons, say in the helicopter case, speeds of 60 to 80 knots at altitudes of 100 to 200 feet are typical. In the case of attack aircraft, speeds of 400 to 500 knots at altitudes of two to three kms. (6560 - 9840 feet) are normal flight profiles, though the latter can also be "down on the deck". At the same time the man or the crew is performing these flight maneuvers which are "hairy" in broad daylight, let alone darkness, he must monitor his aircraft instruments, he must be wary of terrain obstacles and he may well be under enemy fire. Ideally, the pilot wants to acquire and engage targets on a single pass. If he must take several passes, his survivability is greatly reduced. If he has been directed to a probable target area and is able to take advantage of terrain masks he must "pop up", make his observations, acquire and engage targets in times like 15 or 20 seconds. He would always like to acquire and engage enemy targets at sufficient standoff distance to be out of range of radar controlled enemy air defense weapons. This means he is usually at or near to threshold recognition range of the sensor."

From these scenarios, it is possible to discern several bounds on the effective false alarm rate. The pilot observer who is performing difficult maneuvers, monitoring his platform instruments, and avoiding terrain obstacles is interested in a false alarm rate in the order of 10's of seconds per alarm, or higher. Given

a helicopter at 200 feet, $\theta = 5°$, v = 118 nmi/hr. (Figure 2-2) and looking ahead at a range of say 2300 feet, the time between new information frames is 15 seconds. For a $FAR_{eff} = 1/20$, $P_{fa} = 15/20 = .75$. An even better FAR_{eff} is 1/100, so that $P_{fa} = .15$.

For aircraft speeds of 5× the helicopter speed and an altitude of 6000 feet and slant range of 12000 feet, the time between frames $T = 1/5\ T_{100m} = 7$ seconds and the above numbers are reduced by 7/15.

The above has been concerned with a pilot observer who cannot be occupied full time with a display. In this case it has been suggested that an effective false alarm rate of one per 10 to 20 seconds is appropriate, although further examination by human factors specialists is clearly in order. Another case is that of a full time observer, such as an RPV operator. Here, it may be possible to tolerate a higher false alarm rate, perhaps up to one per second, since no additional tasks are involved.

2.2.3 Speed of Operation

The speed of operation of the cueing system depends on the number of sensor frames to be processed per second. The frame rate and frame size determines the sample or pixel rate.

To obtain design goals for frame rates and pixel rates, it seems reasonable to start with something familiar that is likely to be applied to automatic cueing systems: namely, the standard TV format. With 525 horizontal lines and assuming 635 pixels per line (which would correspond to high quality broadcast TV), there are 333,375 pixels per frame. This number may be high because it represents the picture sent to a home receiver by commercial TV. The home receiver projects a

picture, which due to interference, etc. is more like 525 lines by 490 pixels or 257,250 pixels per frame. The TV frame rate is 30 per second, which results in sample rates of 10 megapixels and 7.7 megapixels respectively, if each frame is processed.

What effect does the automatic cueing process have on frame rates? One factor is the slew rate of the sensor; assume, for example, that the sensor is responsible for 120 degrees of azimuth coverage and the platform is a helicopter in the "pop up" mode. If the allowable reaction time is 20 seconds[6], then the field must be scanned in, say, 8 seconds for a slew rate of 15 degrees/second and the human observer has another 12 seconds to verify and respond. Assume a FOV of 5 degrees, then 24 processing frames are necessary for azimuth coverage. For an eight second time interval, this reduces to three frames per second or every tenth TV frame. The pixel rate in this case becomes 1 megapixel per second to 0.77 megapixels per second. If the field of view is increased to, say, 10 degrees, the frame rate becomes 1.5 frames per second and a rate of 0.5 megapixels/second. Clearly, there is a trade off between slew rate, pixel rate, and field of view. The pixel rate, algorithm complexity, and machine speed are interdependent. Machine speed may in turn be increased by parallel processing which implies duplicate circuitry and increased expense. The design goal is difficult to define precisely without an operations analysis of each mission. Prime factors to be considered are the coverage desired, and the amount of time necessary to complete that coverage. In spite of the cursory nature of the above analysis, it does indicate a level of magnitude for frame rates and pixel rates.

---

6. Looft, D.J., Image Understanding -- A Perspective of DOD Needs, banquet speech presented at the Sixth Annual AIPR Symposium, College Park, Maryland, June 1, 1976.

There are several other factors which may be presented in support of frame rates lower than 30 per second. For example, once the system has detected a target, it is available to the operator to be dealt with; it is no longer necessary to reacquire at 30 frames/second. Further, as discussed in Par. 2.2.2, some work indicates that the process is not noise dependent; i.e., detections and false alarms do not change with repeated looks at the scene. This implies that the processing frame rate should be a function of the rate of change of the scene. From Figure 2.2., this may vary from 2 to 10 seconds which would imply a processing rate of 0.5 to 0.1 frames per second.

2.3 PHYSICAL CHARACTERISTICS

The physical characteristics of the automatic cueing system include the allowable size and weight, and the power consumption.

The size and weight of the cueing system are constrained by two limits. One limit is the allowable space and weight aboard the vehicle. The other limit is the minimum amount of hardware required to perform acceptable cueing. Hopefully, the latter value is less than the former. The advent of medium scale integration (MSI) techniques and large scale integration (LSI) techniques, coupled with special chips, will contribute substantially to a significant reduction in both size and weight. Examples of this are commercially available 1 megabit CCD serial memories on a 9 × 15 inch board. Another example is the field programmable logic array (FPLA) which places 24 inverters, 48 AND gates, and 8 OR gates on a 24 pin chip. Further, special chips are useful in that shift registers controlling data flow numbering perhaps 25 chips might become a single chip representing a special purpose stack register. Of course, bare chip packaging would further reduce size and weight figures. With

regard to hardware implemenation, it seems not at all unreasonable to expect the system to be in the neighborhood of 0.5 to 1.5 cubic feet, excluding displays, and to weigh a nominal 15 to 30 lbs. including power supply.

The requirements of a specific platform are, as mentioned, another consideration. The aircraft and helicopter could probably accommodate the upper end of the hardware estimates, i.e., 1.5 cubic feet and 30 lbs.; however,missile installation would produce more stringent requirements, such as 0.5 cubic feet and 15 lbs.

Power design goals, assuming the automatic cueing system has its own power supply,can be roughly bounded. Certainly even for the case of aircraft and helicopters, 1000 watts is too high; in fact, less than one fifth that number is probably more appropriate with a still lower power requriement associated with the missile. It appears that numbers in the neighborhood of 200 - 300 watts are reasonable. It must be cautioned that the size, weight and power quantities are considered to be design goals and not hard requirements; they represent a first attempt to define a set of acceptable numbers with which the users may not be completely satisfied, but which will be of some interest. Clearly, interest will increase by the amount that the design goals are met and exceeded.

2.4 ALLOWABLE COST

The allowable cost of an automatic cueing system will ultimately relate to its dollar value. Determination of this value will depend on the increase in the overall system performance resulting from its use.

Acquisition of targets is valuable primarily because it results in action of some kind. The cueing must take place in a system that can benefit from

increased information. Improvements in weapon delivery systems, recently described in the press, should, in turn, require more effective target acquisition.

Suppose the cueing machine could result in doubling the effective action of a particular mission. Then, it has doubled the mission value. To put it another way, half as many missions would be required to achieve the same result.

Placing a dollar value on the results of a military mission appears futile. However, it seems reasonable to assume that on the average, the value exceeds the cost; otherwise, such missions would not be attempted. The mission cost includes the mission operating expenses, and the appropriate depreciation of the total aircraft system.

The value of the cueing machine is, therefore:

Mission Cost × Total Missions Flown × Effective Increase in Performance

If we assume that the automatic cueing machine will perform for the life of the aircraft, then the first two factors above can be replaced by the total aircraft cost plus the expense of all operations. If the effective increase in performance caused by the cueing machine were, in fact, doubled, then clearly, its value would be very large. Since its cost will be only a small fraction of this amount, for any significant increase in performance, it will be highly cost effective.

In general, this line of reasoning can be carried a step further to infer the optimum level of complexity of the cueing machine. The solid curve in Figure 2-3 represents the machine performance as a function of circuitry complexity. Intersection with the horizontal axis at A indicates that a machine of some complexity is required to achieve any cueing performance at all. A rapid rise

from A to B is experienced as additional equipment is put to good use. Beyond B, however, saturation occurs, because no amount of equipment will produce performance above 100%. The dashed line represents the condition stated previously, that the value assigned to the machine is proportional to mission performance improvement. Since it touches the smooth curve only at B, it describes a mission that can make cost effective use of the automatic cueing machine only if its complexity (point C) provides performance approaching the maximum. Value lines above this will not result in cost effective operations while lines below it will offer solutions over a wide range of complexities. However, the most effective operating point will remain in the vicinity of B.

Although it may be argued quite convincingly that the cueing system could be priced in accordance with its value, as a practical matter the upper price limit will be dictated by other considerations. There is a strong tendency at present to provide a fixed amount of money for an overall military system (design-to-cost philosophy), such as an aircraft or helicopter. Accordingly, regardless of value, the cueing system must compete for acceptance with many other important items, such as sensors and navigation devices. Because the cueing function is a recent development, which is not fully understood or appreciated, it will suffer in the competition for funding. The real funding limit will depend on the circumstrances in each application.

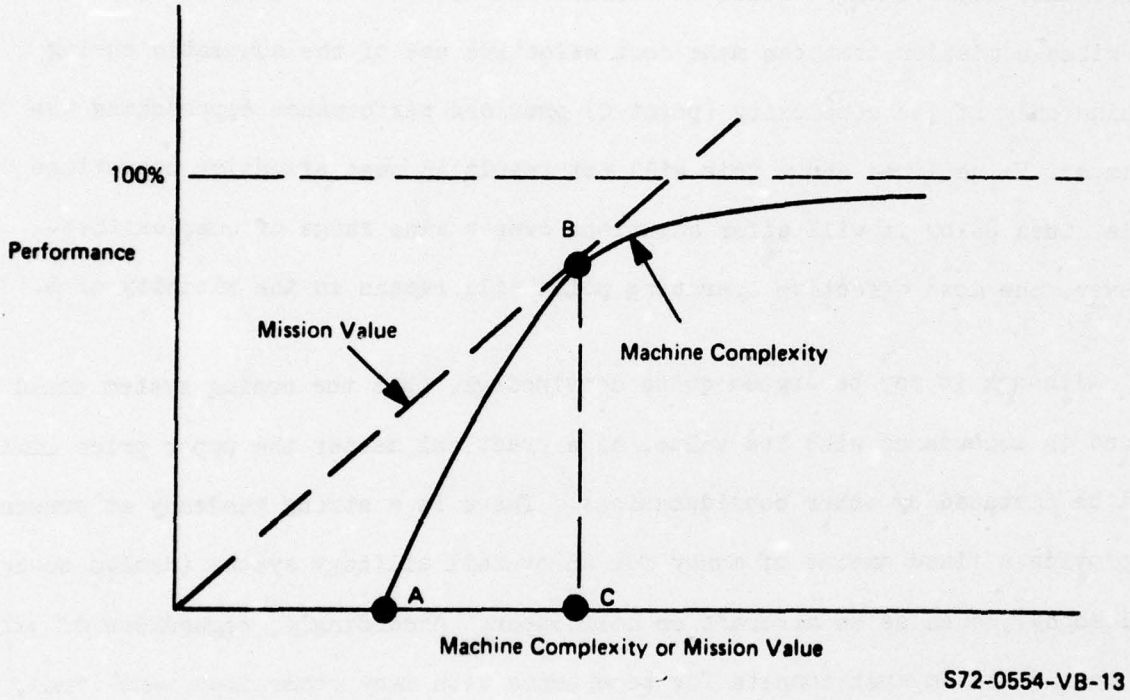


Figure 2-3. Performance vs Machine Complexity or Mission Value

3.0 HARDWARE CONSTRAINTS

This section discusses constraints placed upon the algorithms by the hardware; certain mathematical functions, logic, and memory structures have been found to be difficult, in terms of chip count and processing time, to implement and should be avoided by the analyst. Unfortunately, all too often, the hardware designer is given a completed and checked out algorithm to implement. To change the algorithm structure at that point would require that the analyst redesign the algorithm and repeat the checkout process which is usually not budgeted. This section is written to avoid that dilemma.

The system design goal section has suggested that an automatic cueing system should be capable of performing rates of roughly 1 megapixel/sec. This implies that there is a 1000 nanosecond time interval between pixels, which becomes a rough bound on machine speed. To complete an algorithm within a frame time, it probably must be repeated for each pixel which means that the algorithm must be completed on a per pixel basis in 1000 nanoseconds.

The use of microprocessors is generally ruled out with a 1000 nanosecond processing time because the faster bipolar microprocessors have instruction cycle times (through the arithmetic logic unit (ALU) and back) of 200 nanoseconds. In addition, time is involved in fetching instructions, decoding instructions, putting data into the RAM portion of the ALU, and finally, executing. Power requirements limit the microprocessors to transistor-transistor logic (T^2L) rather than emitter-coupled logic (ECL) which, of course, is much faster. So the hardware thrust seems to lie in the direction of clock-controlled sequential logic. Again, this is not limited to pipeline arrangements because parallel processing has obvious advantages in certain situations.

Clearly, the kinds of times involved prohibit other than very simple arithmetic such as sums and differences. Numerical integration, quadratic smoothing, etc., and techniques normally associated with larger machines, although not impossible to implement, will be costly to implement and less likely to succeed. Sums and differences will probably be performed with table look-up techniques if digital processing is desirable.

Another constraint acts on the kinds of numbers involved. In general, the fewer the number of bits to represent a number, the better. This is advantageous on two counts. First, the devices can be much simpler. Second, the devices can be faster.

The 1000 nanosecond constraint puts tight limits on the types of algorithms which one can consider applying to the whole image or major parts of it.

The problems of memory access are typical. In general, the memory content is shifting at the rate the image is scanned. If an algorithm calls for a change

in a variable memory location (pixel) based on algorithm results, then random access must be obtained to a shifting memory in a matter of 100 nanoseconds or so. The assumption, here, is that most of the allowable 1000 nanoseconds is consumed by the algorithm execution. If a small portion of the image is affected, then a RAM in parallel with a large shift register (4000 word CCD, for example) can accomodate the memory access. However, if the portion affected is large, the memory now must be composed of a large number of RAMs or a large shift register with intolerably long access times. In the former case, the chip count is substantially increased; in the latter case, the megapixel rate may not be achieved since access times are in microseconds. Some other guide lines are listed in the next several paragraphs.

In general, the algorithm should be as repetitive as possible in that it does the same operation of every point in the picture (like a simple FORTRAN DO loop without conditional statements).

In general the algorithm should require a minimal (e.g., 1) number of "passes" over the image to minimize storage, chip count, and time. An example would be algorithms which do not include edge and region growing approaches.

In general, the use of trigonometric functions should be limited, especially in analog and CCD implementations.

4.0 HARDWARE IMPLEMENTATION

The algorithms received fall into two classes which can be roughly classified as thresholding and noise filtering algorithms. The threshold algorithm is implemented with digital circuitry and the overall data flow is examined; analog implementation is considered for the same algorithm. Parallel processing and recursive formation of histograms, averages, and threshold comparisons are also discussed. The noise filtering algorithm is implemented with digital techniques since it is essentially a binary algorithm in its conception. A discussion of alternate hardware approaches to these algorithms follows with some indications of future directions. This section concludes with a discussion of CCD techniques and fabrications applicable to these algorithms for possible on-focal-plane processing. It should be emphasized that the intent of this section is to implement the specific algorithms and not to provide a general discussion of LSI techniques which are available in the open literature. Because we were able to start Phase II work during Phase I, it was felt that direct implementation was more beneficial to our colleagues of the Univeristy of Maryland and, ultimately, to the project.

4.1 THRESHOLD ALGORITHM

There are four variations of the thresholding concept , but central to all of them is the sequence of calculating the gradient and forming a histogram in order to find the gradient value corresponding to, say, 80% of the population in a frame. Then, using this amplitude , the lower 80% of the gradient population is discarded. The remaining pixels also have a gray level associated with them; the average amplitude is computed and used to threshold and discard those amplitude values below the average. The particular gradient threshold technique analyzed

is the "Differences of Averages over k × k Neighbors" where k = 2.

Consider the operator, $p(\tilde{X}) = \max\{|\overline{A} - \overline{B}|, |\overline{C} - \overline{D}|\}$; each term is shown in Figure 4-1 with the composite.

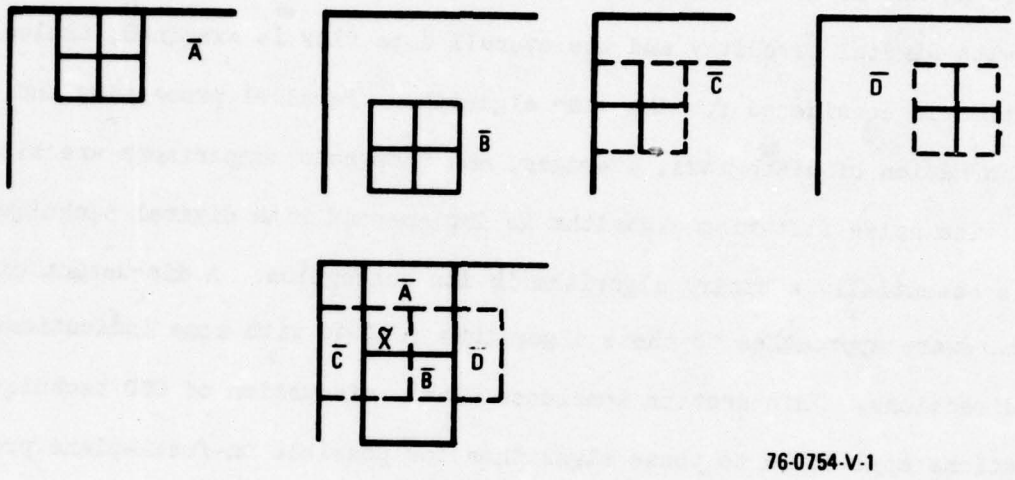


Figure 4-1. $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$ and Composite

where:

$$\tilde{A} = \sum_{i=1}^{4} a_i; \qquad a_i \text{ is a pixel element in } \overline{A} \ (2 \times 2)$$

$$\tilde{B} = \sum_{j=1}^{4} b_j; \qquad b_j \text{ is a pixel element in } \overline{B} \ (2 \times 2)$$

$$\tilde{C} = \sum_{k=1}^{4} c_k; \qquad c_k \text{ is a pixel element in } \overline{C} \ (2 \times 2)$$

$$\tilde{D} = \sum_{m=1}^{4} d_m; \qquad d_m \text{ is a pixel element in } \overline{D} \ (2 \times 2)$$

$\overline{A}$, $\overline{B}$, $\overline{C}$, and $\overline{D}$ each form a 2 × 2 matrix; $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$ each represent the sum of all four elements; each element corresponds to a picture element. Since the

value of each picture element is a gray level with a value $0 < e \leq 63$, the sums are formed with positive quantities which may be analog or digital. These sums may be formed serially by scanning up or down the frame or in parallel as in an array. Serial formation is discussed first. $\tilde{A}$ could be formed by scanning across two lines and placing the sum in an appropriate location. $\tilde{A}_1$ could be placed in the upper left pixel location for $\bar{A}_1$, or it could just have easily have been placed at the X position as in Figure 4-2.

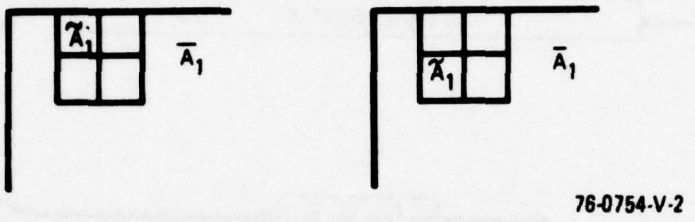
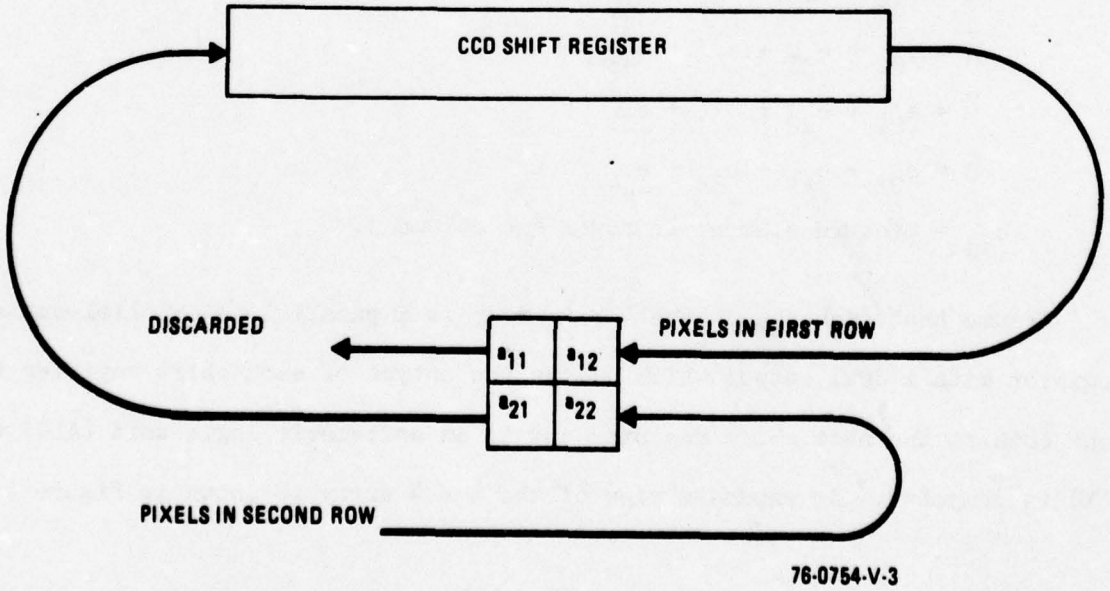


Figure 4-2. Alternate Placements of the Sum, $\tilde{A}_1$

In general, it would be convenient to place the sums $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$ such that indexing systems are simple.

Figure 4-3(a) shows how $\tilde{A}$ could be formed by conventional circulating CCD shift registers and some parallel in-parallel out shift registers.



Fgiure 4-3(a). Formation of $\tilde{A}$ with Shift Registers

Expanding on this idea, Figure 4-3(b) shows the formation of $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$.

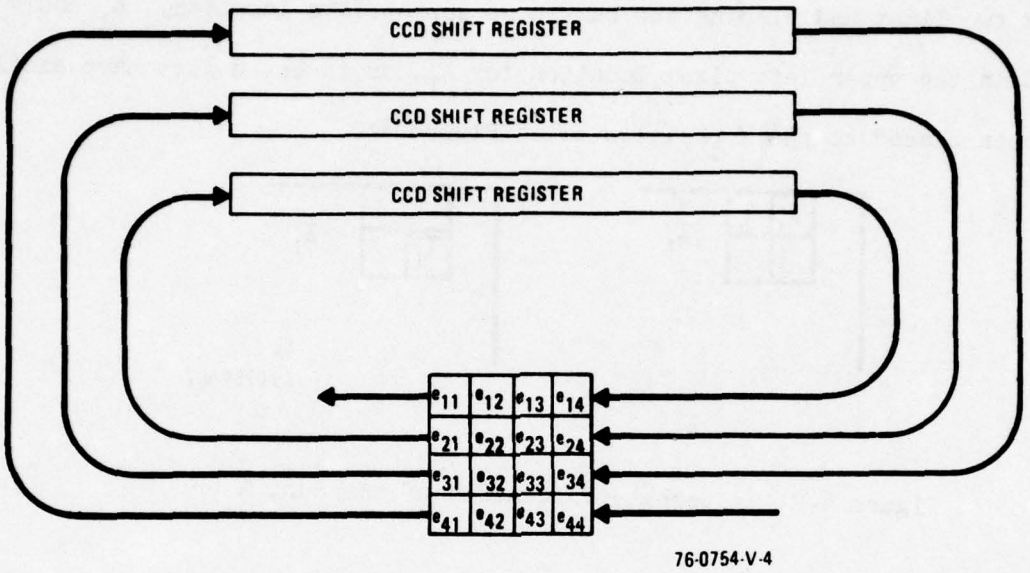


Figure 4-3(b). Formation of $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$ with Registers.

where:

$$\tilde{A} = e_{12} + e_{13} + e_{22} + e_{23}$$

$$\tilde{B} = e_{32} + e_{33} + e_{42} + e_{43}$$

$$\tilde{C} = e_{21} + e_{22} + e_{31} + e_{32}$$

$$\tilde{D} = e_{23} + e_{24} + e_{33} + e_{34}$$

e_{ij} = picture element in row i and column j.

Assume that each box in the 4 × 4 array is a parallel-in-parallel-out shift register with a dual output which allows the output of each shift register to be sent both to the next shift register and to an arithmetic logic unit (ALU) where p(X) is computed. An expanded view of the 4 × 4 array is shown in Figure 4-4.

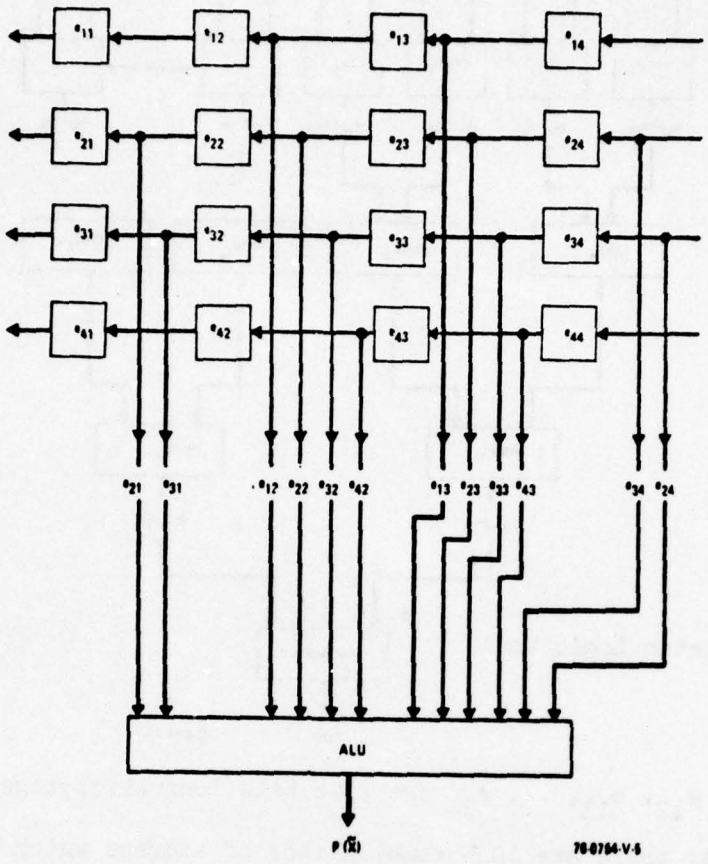


Figure 4-4. Expanded View of 4 × 4 Array

When the gray level amplitude located at e_{22} is shifted to the e_{21} position, it is also sent to the ALU where a new p(X) is calculated. The rate at which the frame is scanned and p(X) is formed is limited by the shift rate of the shift registers, which could be multiplexed to produce higher rates, and the speed at which p(X) is formed by the ALU. Recall that the present analysis is only concerned with a serial scan of the frame and the option to process data by analog or digital is now addressed.

4.1.1 Digital Implementation

Assume the pixels have been digitized and are represented by five bits, $0 \leq$ gray level ≤ 31. The ALU is formed in Figure 4-5.

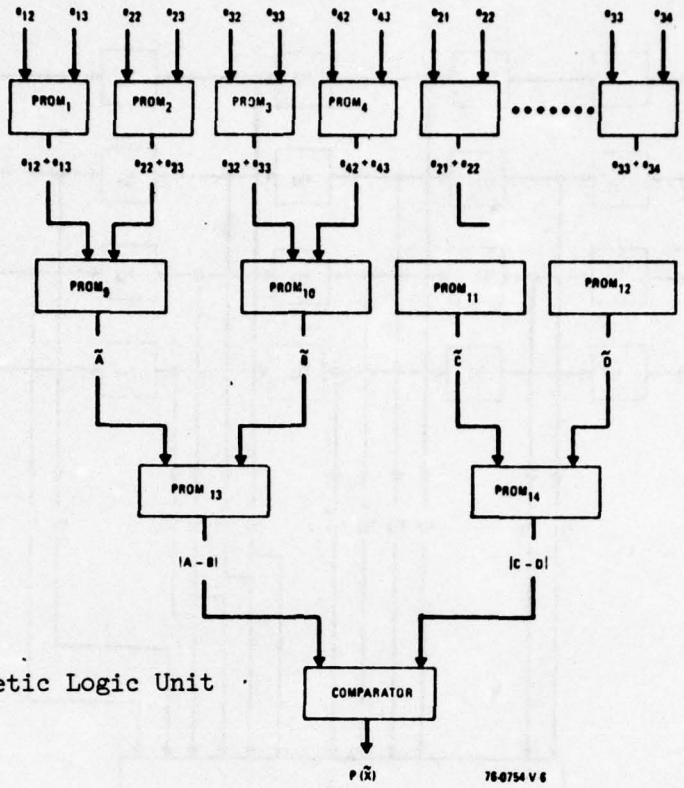


Figure 4-5. Arithmetic Logic Unit

The quantities e_{12}, e_{13}, ... e_{34} are five bits long which together form a ten-bit address; there are 1024 combinations of address which calls for a PROM which is 1024 × (at least) five bits wide. The sum is stored at each address and read out. The second stage of PROM's, 9 - 12, has the same characteristics as do PROM's 13 and 14. Typically, these PROM's have a propagation delay of 90 nanoseconds for Schottky and low power Schottky. Emitter coupled logic, ECL, will not be considered because of excessive power requirements. Assuming a settling time in the shift registers of 30 nanoseconds, and a comparator propagation time of 60 nanoseconds, the total time to produce p(X) is 360 nanoseconds. This time could be reduced by substituting adder chips for the PROM's, to 240 nanoseconds, but the number of chips would have increased from 15 to at least 30. Another way to reduce propagation time from e_{ij} to p(X) is to employ a single PROM stage to produce $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$ each. To form $\tilde{A}$, the address is 20 bits wide composed of 2_{12} e_{13} e_{22} e_{23}, or four PROM's wide and $32^4/1024$ PROM's long which is a substantial increase in chips for a reduction of 90 nanoseconds. Thus, perhaps an optimum combination of chips and speed is 15 chips and 360 seconds. Another consideration is the forming of a histogram composed of $p(\tilde{X}_{ij})$'s.

The operator $p(\tilde{X}_{ij})$ has a range of values $0 \leq p(X_{ij}) \leq 31$ and a five-bit binary to decimal decoder would place the $p(X_{ij})$ value in one of 32 slots. A counter and register for each decimal line would accumulate the population in each slot, as shown in Figure 4-6.

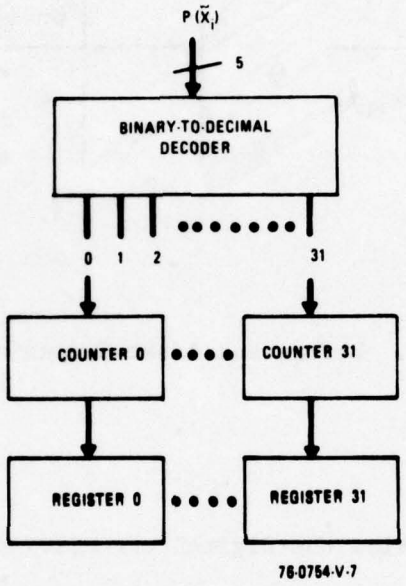


Figure 4-6. Histogram Formation

Over one frame, the number of pixels is a known quantity and the number of pixels corresponding to 80 percent of the pixels is also known, say C_1, so that the registers 0 - 31 need to be summed sequentially until C_1 is reached. A counter would track the registers to determine which register was being summed when C_1 was reached. This value $p(X_{C1})$ then serves as a threshold which is compared to all the $p(X_{ij}$'s) in a frame. Before and after comparison, the frame would look like Figure 4-7.

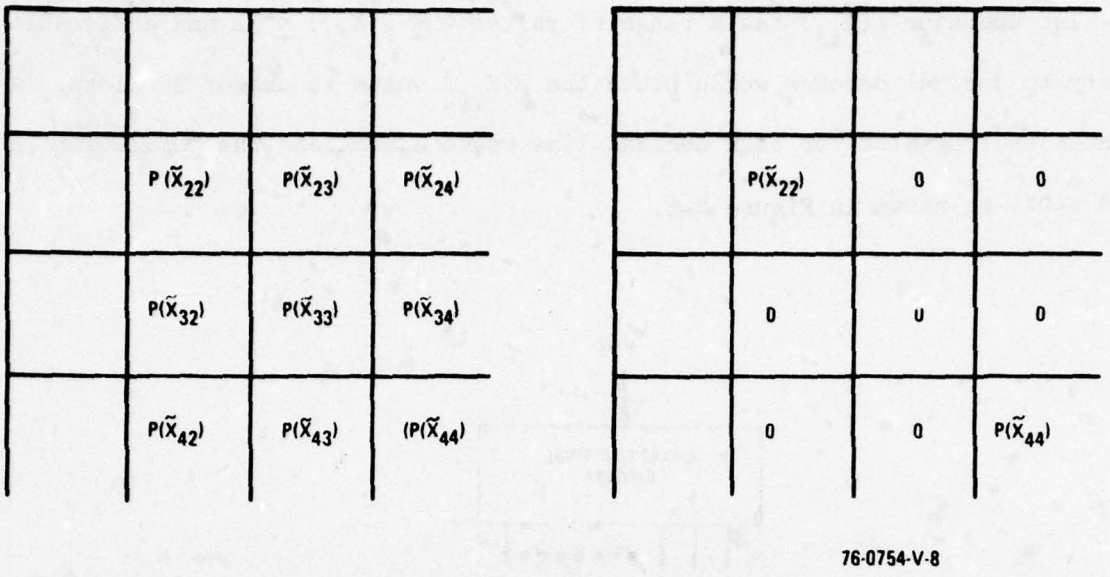


Figure 4-7. Before and After Comparison

4.1.1.1 Overall Data Flow

It is appropriate to review the digital circuitry derived so far and put it on one piece of paper. Note that after the computation of each $p(X_{ij})$, it is stored in a set of CCD shift registers at the left. When $p(X_{Cl})$ is obtained, all $p(X_{ij})$'s in the frame have been stored in the $p(X_{ij})$ registers. Now, each $p(X_{ij})$ value is shifted out and compared with $p(X_{Cl})$ in a comparator. At the same time that the operations to obtain $p(X_{Cl})$ are being performed, the gray level amplitudes are being stored in another set of shift registers, the e_{ij} registers. If $p(X_{Cl}) \leq p(X_{ij})$, then e_{ij} is stored in the e_{ij} registers at position ij, if $p(X_{Cl}) > p(X_{ij})$ then 0 is stored at position ij. At the same time, the value e_{ij}'s are being cumulated and counted to obtain an average $\bar{e}$. When $\bar{e}$ has been obtained, the remaining e_{ij}'s are then compared to $\bar{e}$, to eliminate another group of data.

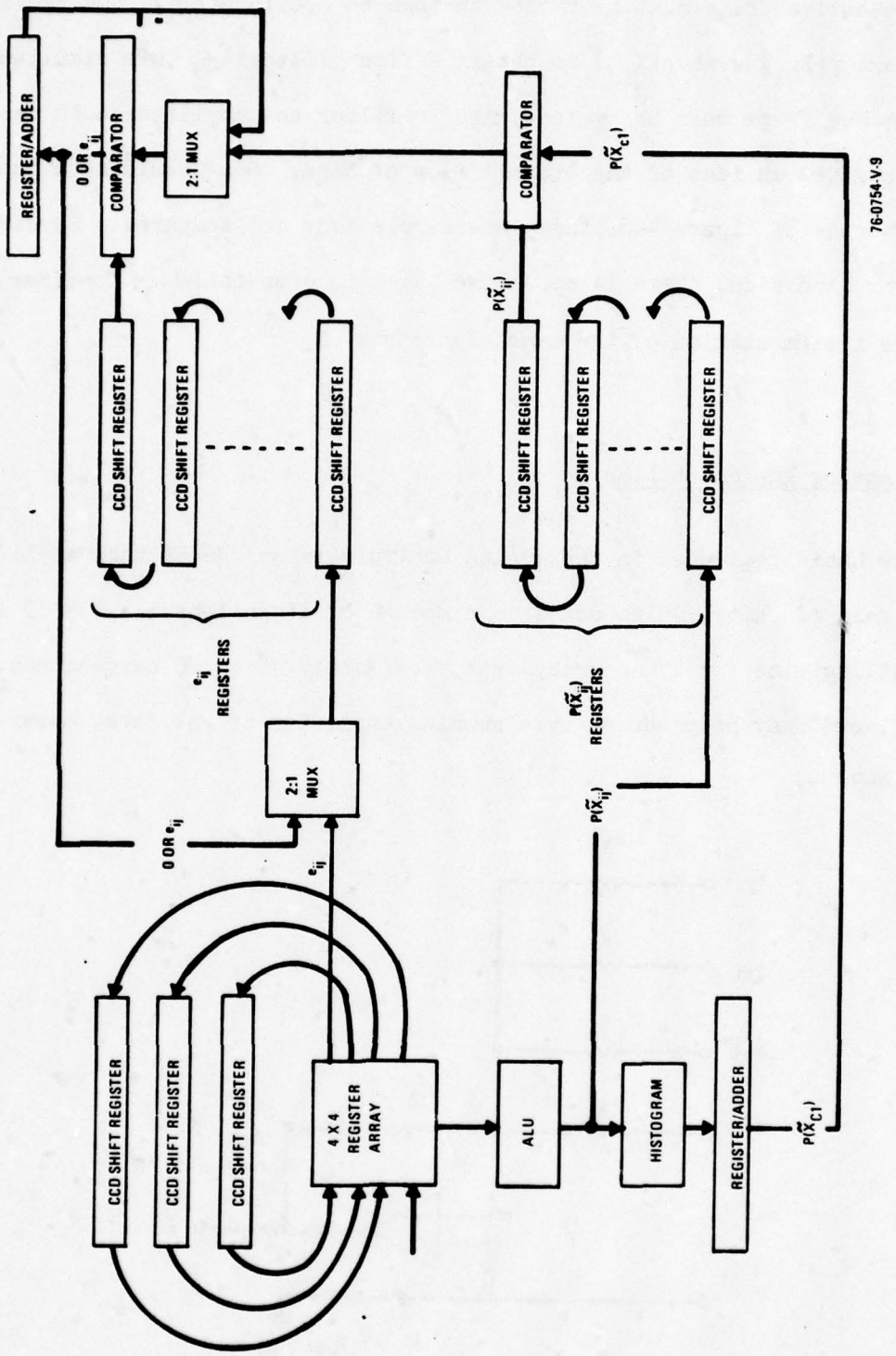


Figure 4-8. Overall Data Flow

The entire frame must be cycled through to obtain $p(X_{Cl})$, and recycled to compare $p(X_{Cl})$ with $p(X_{ij})$ to obtain a first filtering; this simultaneously gives $\bar{e}$. The frame must be cycled again to filter the amplitude data based on $\bar{e}$. This gives an idea of the overall flow of data. One should note that in the left side of Figure 4-8, there are simple adds and compares. However, in the right hand side, there is more time taken in computation. Consider an analogue implementation of the same algorithm.

4.1.2 Analog Implementation

The basic tool used in the analog analysis is an operational amplifier, with a gain of unity, which comes in a module of dimensions 1 × 2 × .5 inches. The settling time for 1% accuracy is approximately 40 to 50 nanoseconds. Either A, B, C, or D may be produced by a summing amplifier of the form, shown Figure 4-9.

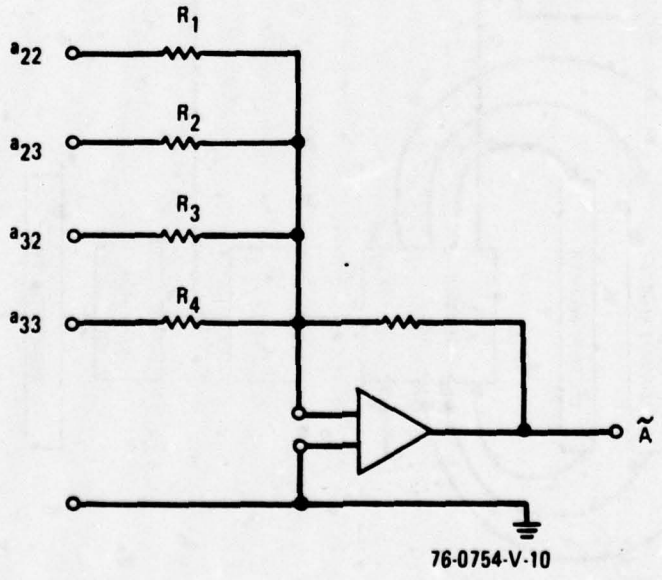


Figure 4-9. Summing Amplifier

$$\tilde{A} = -(a_{22} + a_{23} + a_{32} + a_{33}), \text{ where } R_0 = R_1 = R_2 = R_3 = R_4$$

and this amplifier, because of multiple inputs, may have a settling time of 80 nanoseconds. Note that settling time is the same as the propagation time in digital circuitry. Similarly, $\tilde{B} = -(b_{42} + b_{43} + b_{52} + b_{53})$ and four amplifiers are needed to obtain $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$. Since a_{22}, a_{23}, a_{32}, and a_{33} are positive quantities, as are all elements, $\tilde{A} - \tilde{B}$ may be obtained from a differential amplifier, (Figure 4-10), or a two stage invert and add. $E_0 = (\tilde{B} - \tilde{A})R_0/R_1$, where $R_0 = R_1$,

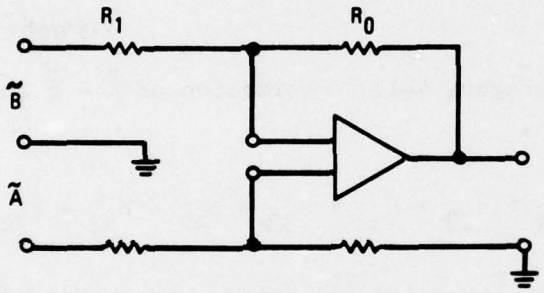


Figure 4-10. Differential Amplifier

but

$\tilde{A}, \tilde{B} < 0$ from the previous amplifier, therefore

$$E_0 = (\tilde{A} - \tilde{B}),$$

and E_0 will be positive or negative depending on the relative magnitudes of A and B. Before considering this, it is worth noting that $\tilde{A} - \tilde{B}$ may be obtained in one step,

Figure 4-10,

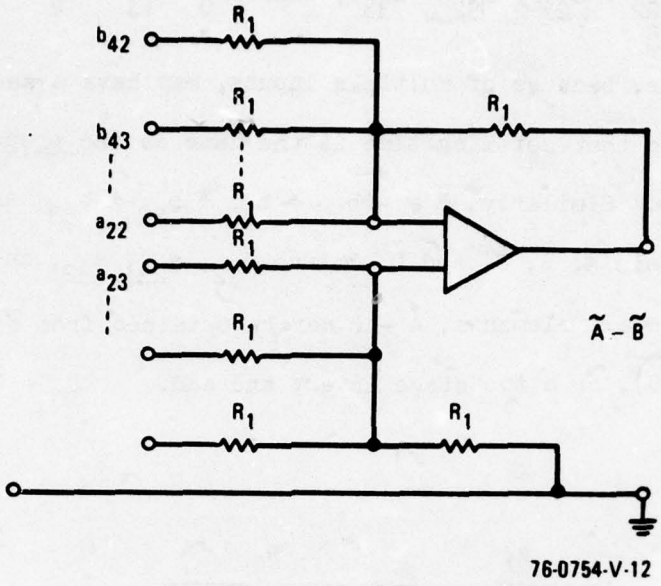


Figure 4-11. Formation of $\tilde{A} - \tilde{B}$ in One Step

$$\tilde{A} - \tilde{B} = E_0 = (a_{22} + a_{23} + a_{32} + a_{33} - b_{42} - b_{43} - b_{52} - b_{53}).$$

with a settling time in excess of 100 nanoseconds. For comparison of $|\tilde{A} - \tilde{B}|$ and $|\tilde{C} - \tilde{D}|$, it is necessary to obtain the absolute value of both $\tilde{A} - \tilde{B}$ and $\tilde{C} - \tilde{D}$ with the following type of circuit, Figure 4-12.

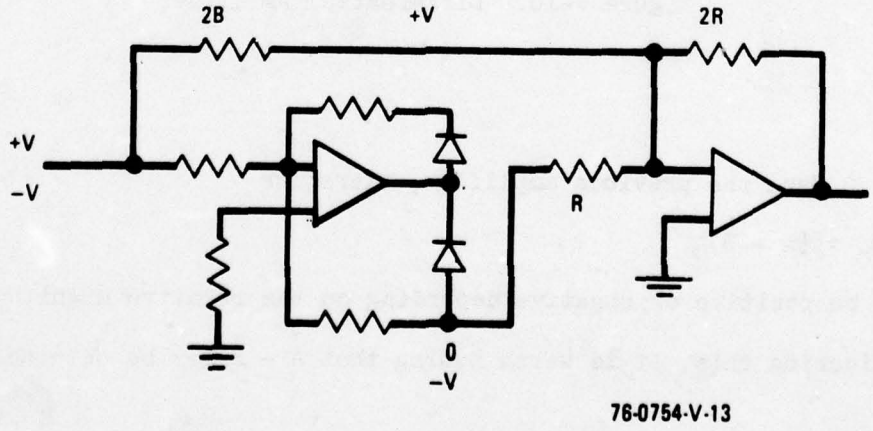


Figure 4-12. Absolute Value Circuit

If A - B = -V, then the circuit produces 2V - V = V, if A - B = +V the circuit produces V. The propagation time is roughly 130 nanoseconds. The output is fed to a comparator which contains a switch allowing $p(x_{ij}) = |\tilde{A} - \tilde{B}|$ or $|\tilde{C} - \tilde{D}|$. Thus far the propagation time is of the order of 310 nanoseconds which is close to the digital result. Forming the histogram is considered next.

Here a bank of voltage comparators can be selected with a switch which feeds a counter associated with each voltage comparator and the number of times a voltage comparator is triggered can be tabulated and the histogram formed. The counters are summed and the 80% value of $p(x_{ij})$ selected. It seems possible to accomplish the entire process in less than about 400 nanoseconds. Let us continue now and examine, briefly, the possibilities of parallel processing, not specifying whether it is analog or digital.

4.2 PARALLEL PROCESSING

Actually, a direct approach is to divide the frame into, say, horizontal slices of a third of a frame in the vertical dimension. Constructing the histogram and interfacing problems may contribute to the propagation time so a more conservative approach may be to divide the frame into four or five parts. It might be suggested that the price has increased, and, while this is true, the increments are small because a 4096 × 4 CCD digital shift register is about $30 to $40 in small quantities. Pictorially, Figure 4-13 shows

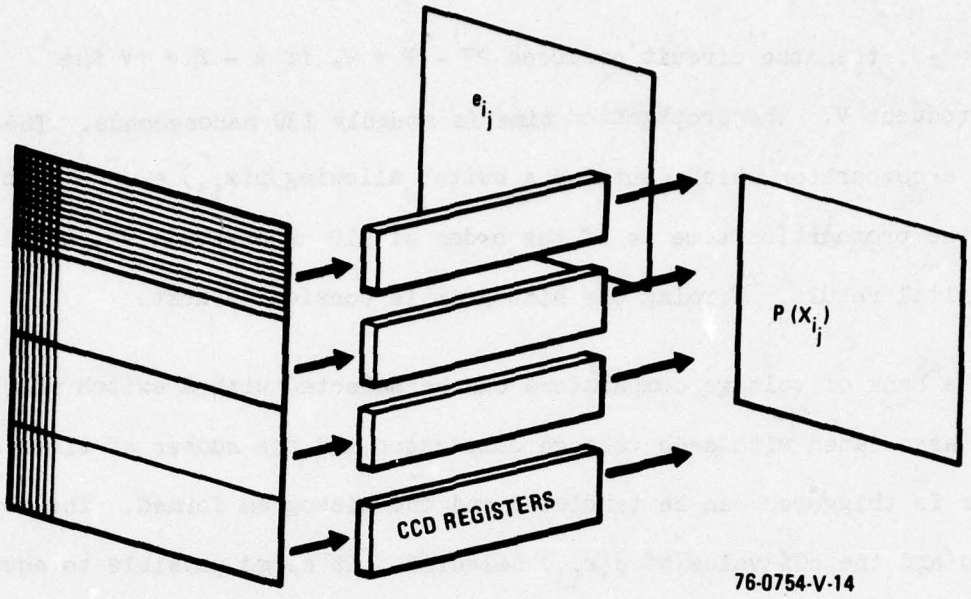


Figure 4-13. Parallel Processing

the frame divided into four horizontal sections. Recall that the $p(x_{ij})$'s must be compared with $p(X_{Cl})$, so the $p(x_{ij})$'s must be saved. Further the amplitude information in each frame must also be saved, so that additional memories of the entire frame in $p(x_{ij})$ and e_{ij} must be stored. The subsequent compares of $p(X_{Cl})$ vs. $p(x_{ij})$ and e vs. e_{ij} are limited by the shift rates of the registers and the comparison function itself. The longest task seems to be deriving $p(x_{ij})$. Another possibility is to update $p(X_{Cl})$ and e recursively and also perform the threshold comparisons at the same time.

4.3 NOISE REGION FILTERING

The Noise Region Filtering Algorithm acts to remove noise bursts which have survived the gradient and amplitude thresholds, but are small in area and isolated.

Input data consists of an entire frame whose pixels contain either 0's or gray level values which have exceeded past thresholds. For this test, the frame should be converted to 0's and 1's, otherwise the filtering will be more difficult to perform. Suppose that when the amplitudes were compared with an average amplitude threshold in the previous test, a "1" was stored in the appropriate pixel location when the test was successful and a "0" was stored when the test was unsuccessful. Thus, the input data for the Noise Region Filtering Test comprises only 1's and 0's. The shrink test may be stated as

$$q'(c) = \begin{cases} 1 \text{ if } c = 1 \text{ and } a + b + d + e = 4; \\ 0 \text{ otherwise} \end{cases}$$

where the arrangement of pixels a through e is

$$\begin{array}{ccc} & a & \\ b & c & d \\ & e & . \end{array}$$

The shrink operation is equivalent to an AND gate in that the test may be restated as

$$q'(c) = \begin{cases} 1 \text{ if } a \cap b \cap c \cap d \cap e = 1 \\ 0 \text{ otherwise.} \end{cases}$$

Since the data is already binary, and each pixel is represented by one bit, a five-bit AND gate is sufficient. The data flow could consist of two CCD shift registers holding the proper rows plus several smaller shift registers to hold the immediate a, b, c, d, e used in the calculation. The "c" shift register would also contain a dual input feature such that the result of the AND gate logic could be entered also. The block diagram appears in Figure 4-14. To

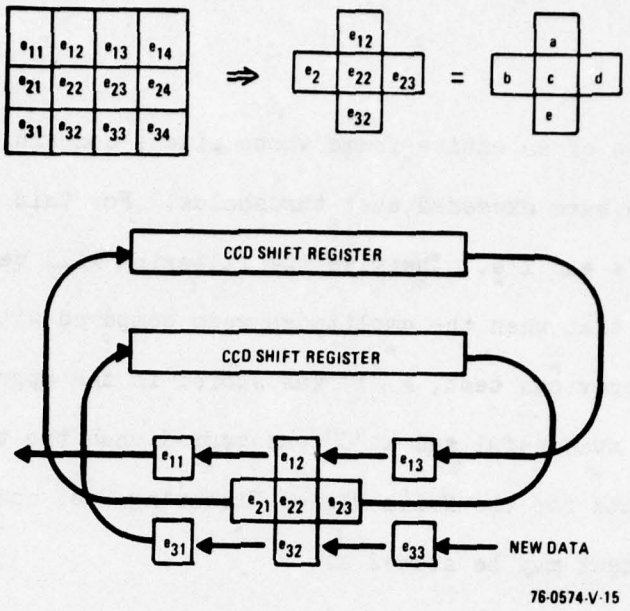


Figure 4-14. Register Block Diagram

allow symmetrical timing, four additional shift registers must be added to the first and third rows so that the set of individual registers becomes a 3 × 3, even though only two of the additional registers will be tapped. The contents of registers e_{12}, e_{21}, e_{22}, e_{23}, and e_{32} are tapped into the AND gate and the results placed in e_{22}. Figure 4-15 shows the data flow in

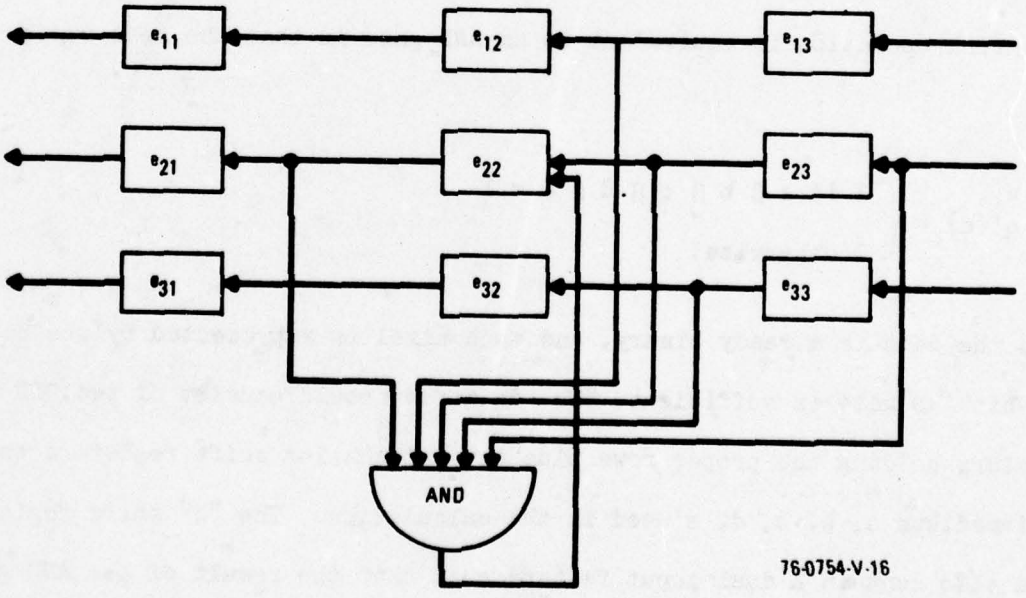


Figure 4-15. Data Flow through AND Gate

which e_{22} is the only shift register which requires a dual input; all the individual registers e_{11}, e_{12}, ... e_{33} are 1 × 1. The propagation time for the process will approximate 50 nanoseconds including the settling time for the entry of e_{22} from the AND gate. To produce a second shrink test a duplicate of the above is required.

The propagate test may be stated as

$$q'(c) = \begin{cases} 1 \text{ if } c = 1 \text{ or } a + b + d + e > 0 ; \\ 0 \text{ otherwise.} \end{cases}$$

This test states that q'(c) = 1 if any of a, b, c, d, e are equal to 1. That is, the AND gate in the shrink test is replaced with an OR gate. Similarly, a two step propagate requires duplicate OR gates.

4.4 ALTERNATE HARDWARE DIRECTIONS

In the previous sections, the purpose was to implement the algorithms in several different ways without regard to tradeoffs. This section is addressed to the tradeoffs process, conflicting considerations, and future directions.

CCD registers were employed throughout because their cost per bit is less than conventional shift registers and they can obtain the necessary shift rates. However, they are volatile; i.e., they lose memory when there is a power failure, whereas the Westinghouse (BORAM) Block Oriented Memory does not, and can be used as a shift register. However, it only requires a few frames to restore the data base with comparable shift rates.

Another factor is the array of smaller shift registers used to hold the appropriate e_{ij}'s, which are then sent to the (ALU) Arithmetic Logic Unit. This

array of smaller shift registers deletes an instruction set needed to draw the appropriate e_{ij}'s from the larger CCD shift registers. Another approach is the tapped CCD shift register with non-destructive read out, developed by Westinghouse, which would delete the need for a smaller array. However, this would call for a special chip.

When special chips are considered, it is possible, for example, to put the ALU of the threshold algorithm on one chip. Whether this is worthwhile or not will depend on size considerations of other parts of the hardware implementation which may be much larger than the ALU. The CCD shift register arrays necessary to hold a frame of e_{ij}'s and $p(x_{ij})$'s are now discussed.

Assume a frame size of 340,000 pixels × 6 bits per pixel; a CCD array of 2,000,000 bits capacity is needed. A 1 megabit array is commerically available of dimensions 128,000 words × 8 bits on a 9 × 15 inch board. Special designs and fabrications are required, and will be considered in Phase II, to reduce the word size to 6 bits. Assuming for the moment that the commercial version is used without modification, three boards would be necessary to hold e_{ij} and three more for $p(x_{ij})$. Further, assuming a spacing between boards of 2 inches, the memory volume becomes 12 × 9 × 15 inches or 1620/1728 = 0.94 ft^3 which is uncomfortably large at this early design stage. On the other hand, it can be suggested that analog memories on special chips would require smaller volumes. In fact, assuming a worst case wherein even the histogram is formed with analog devices which might require some 50-odd operational amplifiers yielding a volume of perhaps 200 cubic inches including support circuitry, the analog implementation may still be smaller than a digital implementation. It should be noted that the first cut has shown the speeds of analog

and digital implementations comparable and half the allowable time which is a good safety margin at this point. Another problem which occurs in analog implementation is signal degradation as a function of the shift rate and length of the register, e.g., a typical example of a 100-stage CCD delay line with an 80 dB dynamic range and transfer inefficiencies of 10^{-4} at 1 to 2 mHz clock rates will have an overall signal degradation of 1%, so there is some benefit to be gained by organizing the memory to reduce the number of shifts. Such organizations might include bringing an entire line into memory in parallel and shifting out in series or multiplexing out to reduce shift rates also.

The above discussion has considered the threshold algorithm and identified its memory requirement as the more interesting among other challenges. The noise filtering algorithm, by contrast, seems fairly straightforward.

Having identified the memory size as the more challenging of the Phase II hardware problems, the next section is a brief description of some of the Westinghouse analog techniques and devices in this area and on-focal-plane processing.

4.5 APPLICABLE CCD TECHNIQUES AND FABRICATION

Any signal processing system that involves the linear transformation of analog signals such as correlation, discrete Fourier transformation (DFT), filter banks, matched filters, multiplexing, array scanning, orthogonal scan transformation, time base translation, etc, can be realized with combinations of CCD basic building blocks. In discrete analog signal processing (DASP), analog data samples are stored, transferred, and operated upon by analog means, whereas in conventional digital signal processing (DSP), digital or quantized samples are handled with binary logic. A major advantage of DSP is retained by DASP, namely the precise transport delay, particularly in relation to coherent signal processing. The dyanmic range of an analog bit in DASP may be thought of as composed of 6-dB equivalent DSP digital bits. Thus, a typical example of 100

Stage (N = 100) CCD delay line with 80-dB dynamic range and transfer inefficiencies of $\varepsilon \sim 10^{-4}$ at 1-to-2 MHz clock rates will have an overall signal degradation of 1 percent (i.e., less than 0.1-dB insertion loss) without the need of A/D conversion.

One-dimensional basic building blocks (linear arrays) may be classified according to the characteristic information flow patterns:

- o Serial in/serial out (SI/SO)
- o Parallel in/serial out (PI/SO)
- o Serial in/parallel out (SI/PO)

which are shown in Figure 4-16.

These fundamental linear arrays may be combined to form area arrays (2-dimensional matrices) with increased signal processing capabilities. Table 4-1 provides a partial listing of applications for these basic building blocks.

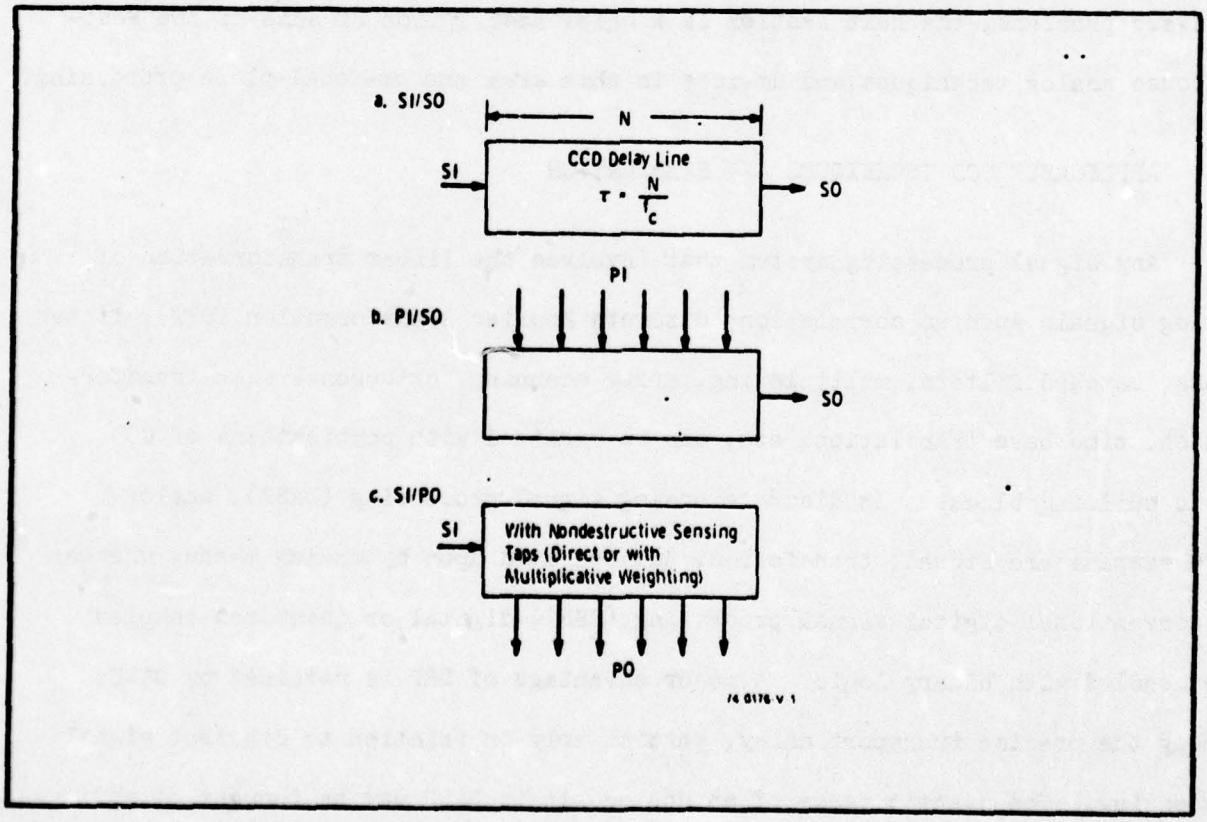


Figure 4-16. CCD Basic Building Blocks

TABLE 4-1.

BASIC INFORMATION FLOW

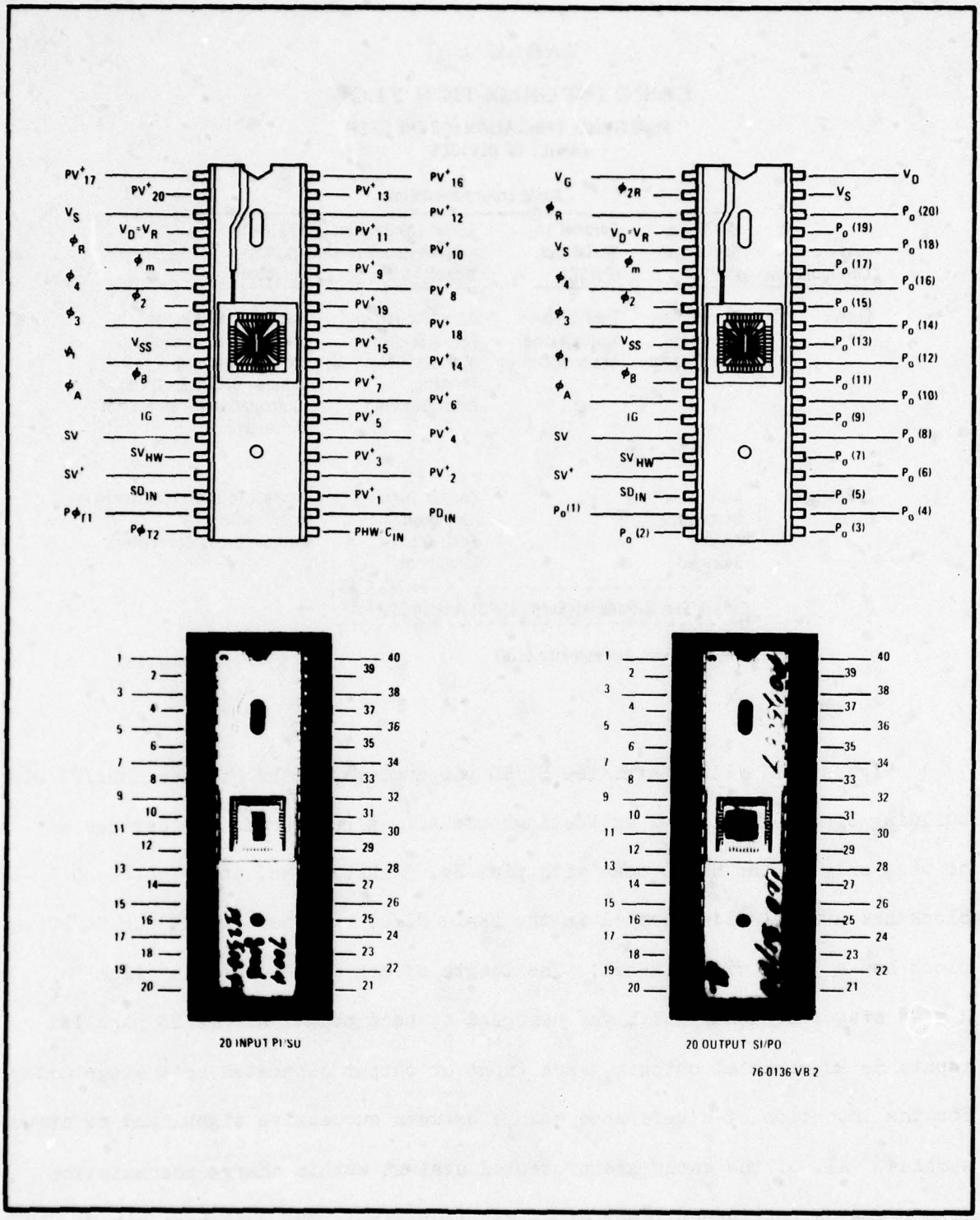
FUNCTIONAL TABULATION FOR THE DASP FAMILY OF DEVICES

Array Configuration	Basic Information Flow: Serial In; Serial Out (SI/SO)	Parallel In; Serial Out (PI/SO)	Serial In; Parallel Out (SI/PO) - Nondestructive Sensing Taps - Unweighted	Weighted
Linear	Pure Delay; Time Base Interchange	Time Division Multiplexing; Array Scanning	Beam Focusing; Focus Scanning Multiple-Beam Forming; Beam Steering	Transversal Filtering; Correlation/Convolution; Adaptive Filtering; Sampled Data Smoothing or Interpolation (Scan Format Converter)
2-d Matrix (Area)	Bulk Serpentine Analog Storage*		Corner Turn (Orthogonal Scan Transformation)	Discrete Fourier Transformers; Filter Banks; Multiple Cross Correlators
	Bulk Serial-Parallel-Serial (SPS) Analog Storage*			

*As For Video Refresh Memories

74-0178-VA-2

Figure 4-17 illustrates the SI/SO pin numbers on the PI/SO and SI/PO CCD building blocks fabricated at Westinghouse ATL, since the latter devices may be used only in the SI/SO mode with pins No. 3 through No. 18. The PI/SO block has a 2-mil wide channel in the basic SI/SO register, while the SI/PO block has a 5-mil wide channel. The length of the SI/SO CCD delay line is N = 44 stages of delay which was designed to accommodate either 20 parallel inputs or 20 parallel outputs, each input or output separated by a stage delay for the insertion of a reference sample between successive signal and reference samples. All of the gates are protected against static charge accumulation with a series resistance/Zener diode configuration. The substrate pin No. 10 should always be the most positive and the maximum substrate voltage is +35V with typical operation at +20V. CCD analog shift registers have been operated with substrate voltages of +8V; however, since the electrode lengths (i.e.,



20 INPUT PI/SO

20 OUTPUT SI/PO

76-0136 VB 2

Figure 4-17. PI/SO and SI/PO CCD Basic Building Blocks

polysilicon electrode length ≈ 12 μm, aluminum electrode length ≈ 8 μm) are not minimum size, the transfer efficiency improves with larger clock voltages due to the increase in electric fringe field beneath the electrodes.

4.5.1 Serial In/Serial Out (SI/SO)

The SI/SO block is a simple CCD analog shift-register with the characteristics discussed in paragraph 4.7. In a linear array configuration, the SI/SO block provides analog signal delay with the ability to provide time base translation. A typical dynamic range for present-day SI/SO blocks is 80 dB with 0.3 percent linearity and clock frequencies from 1 kHz - 1.0 MHz for a 64 analog bit delay line. The clock requirement may vary from device to device with voltages varying from TTL to MOS compatible. In general, MOS-type voltage swings are needed to obtain dynamic range and frequency response. The capacitance loading for the drivers is typically 0.2 pF/mil^2 of active clock electrode, e.g., clock electrode areas of 1.5 mil^2 will have 0.3 pF capacitance. Thus, a 64 bit delay line will offer a loading of ≈ 20 pF/driver. In general, CCD structures have not been built with interface/buffer circuits on the chip because of the advanced development nature of the work; however, CCD chips can be fabricated with MOS, CMOS, or bipolar interface circuits. In order to test SI/SO blocks without on-chip buffer circuits, a so-called "open collector" driver may be employed. This driver is relatively inexpensive and provides clock voltage swings of 30V up to 2-MHz clock frequencies. Clock shaping may be accomplished if desired by the use of a series resistor, which also protects the drivers in the pull-down transient. A CCD chip should have protective resistor diode combinations, similar to MOS-type circuits, to limit the displacement current and prevent shorting of the input electrodes.

For analog signal processing, as discussed in the introduction, a desirable

feature is the incorporation of an ac zero reference between successive signal samples, particularly for PI/SO and SI/PO blocks. In addition, sample and hold techniques are required for analog signal reconstruction which attenuates the response with a $\frac{\sin \pi f_c/f_s}{\pi f_c/f_s}$ shape factor. The input to a CCD is filtered with a $\frac{\sin \pi \Delta t f_c}{\pi \Delta t f_c}$ roll-off where Δt is the sampling window aperture. The output after sample and hold requires filtering with a low-pass filter with ideal "brick-wall" cut-off at $f_c/2$, the Nyquist limit. Figure 4-18 illustrates an analog output swept frequency response of a CCD SI/SO block with sample/hold and a 7-pole Butterworth filter (-3 dB at 750 kHz) to filter the clock and limit the aliasing of frequencies higher than $f_c/2$. Thus, in a properly designed CCD analog delay line the frequency response is limited by the sample/hold and low-pass filter characteristics. Since the input signal to the CCD is sampled, an input prefilter with cutoff at $f_c/2$ is required to reduce aliasing similar to the output filter.

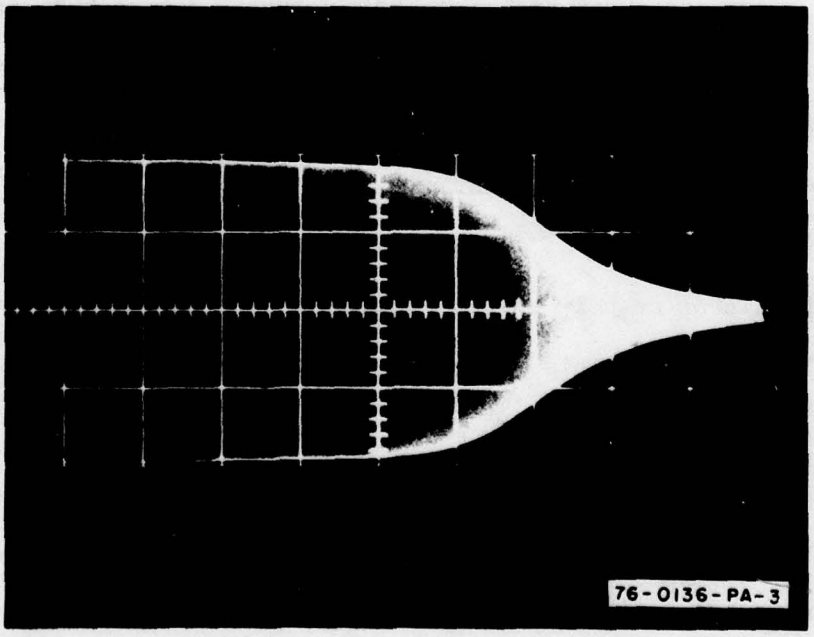


Figure 4-18. Frequency Response of Sampled CCD Analog Delay Line; 100 kHz/Div. Horizontal, f_c = 2.0 MHz; Transfer Inefficiency ε (f_c = 2.0 MHz) = 2×10^{-3} for L = 12 μm Electrodes; Sample/Hold and Filter Responses are Included in the Overall Response

4.5.2 Parallel In/Serial Out (PI/SO)

The PI/SO block may be used to time division multiplex a number of low data signal channels into a higher data rate output channel. The variations in electrical input may be minimized with the use of a stabilized charge injection circuit. An N-channel multiplexer converts N parallel input channels into a single-channel pulse amplitude modulated (PAM) signal. The input signals are synchronously sampled and the sampler information is entered into a unique spatial and temporal position in the CCD delay line. Applications include the multiplexing of many sensor input channels (e.g., electro-optical sensors, acoustical sensors) into a single video output channel.

Figure 4-19 illustrates the basic 20 input PI/SO CCD building block with pin diagram shown in Figure 4-17. Figure 4-20 shows an overall photomicrograph of the PI/SO block. Figure 4-21 shows the SI/SO input part of the PI/SO block prior to aluminum metalization with the polysilicon electrodes defined. Figure 4-22 shows a plotter schematic and pin-diagram labeling of the PI/SO

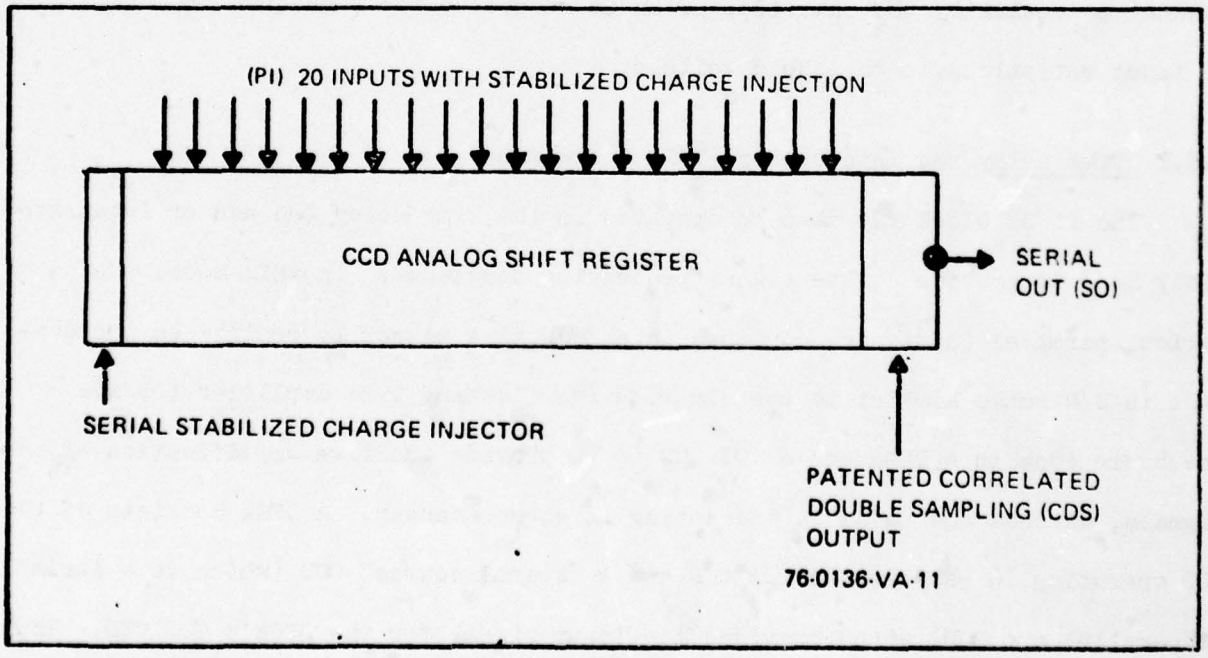


Figure 4-19. PI/SO (20 Input) CCD Basic Building Block

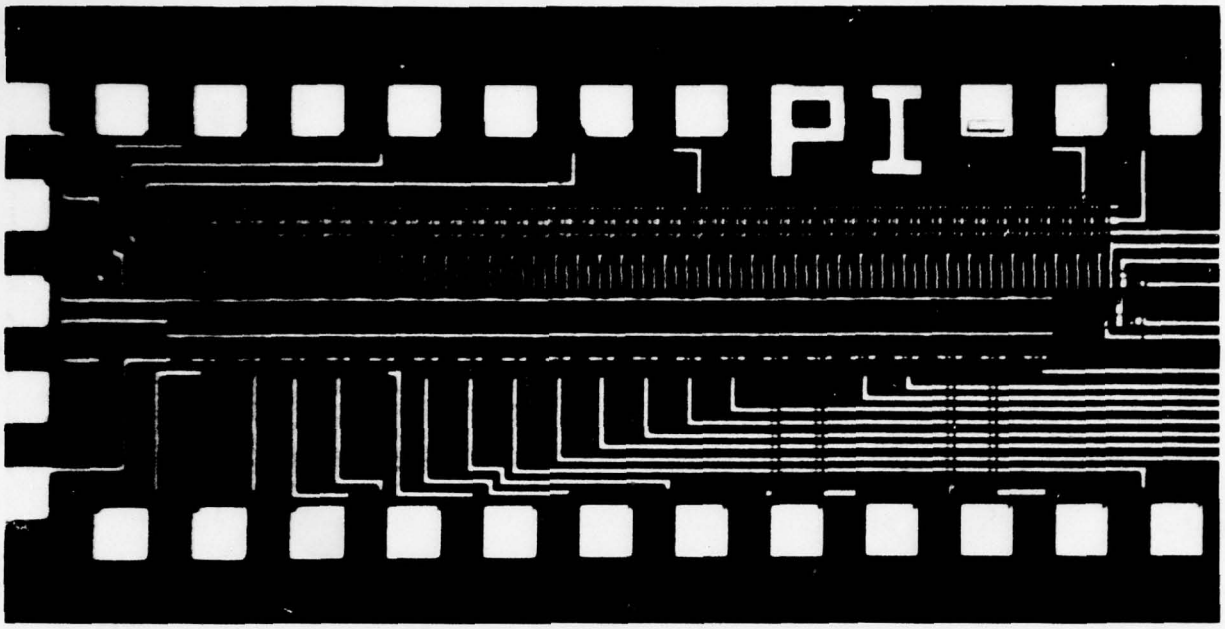


Figure 4-20. Photomicrograph of PI/SO Block

block and Figure 4-23 is a photomicrograph of the PI/SO block after aluminum metalization. The PI/SO block has stabilized charge injection at each of its inputs with parallel injection at alternate stage delays to minimize interchannel crosstalk by the insertion of an ac reference for subsequent removal of input variations in threshold voltage.

4.5.3 Time Delay and Integration (TDI) Operation

The PI/SO block may also be operated in the time delay and add or integrate (TDI) mode to provide unique signal processing functions. In this mode, the various parallel inputs are combined in an additive manner to realize an improvement in S/N ratio similar to the distributive floating gate amplifier (DFGA). The basic idea in a DFGA and a TDI CCD is to provide additive amplification of the signals, whereas the noise is increasing in an rms manner. A DFGA consists of two CCD operating in parallel, a TDI CCD and a "signal source" CCD (which is a serial IN/parallel out CCD) which provides the input signal for the DFGA's TDI CCD. Thus, for N stage DFGA or TDI CCD (i.e., N parallel inputs) the output signal will be N times larger and the noise (assuming uncorrelated noise sources) will be N times

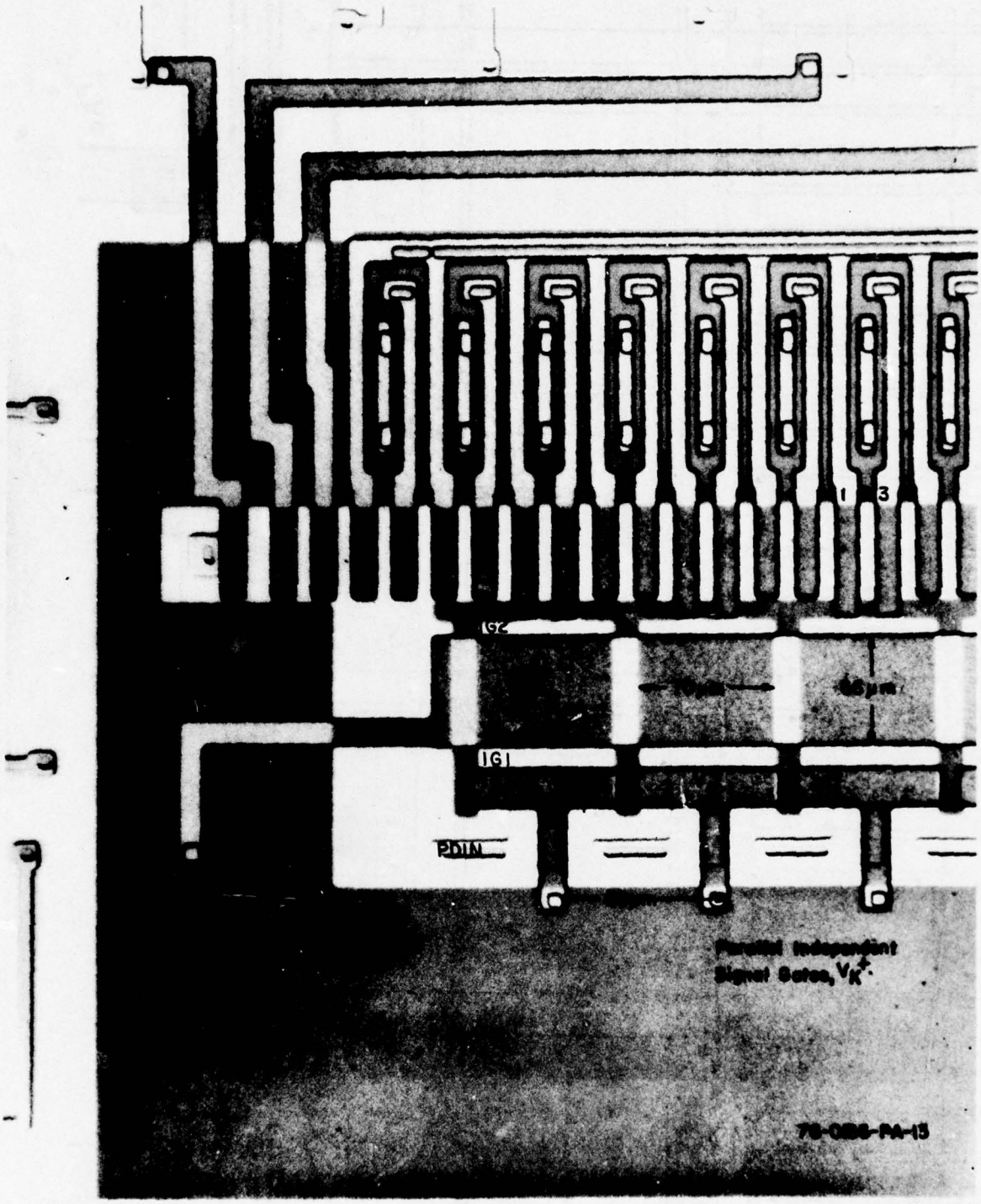


Figure 4-21. SI/SO Input Part of the PI/SO Block Prior to Aluminum Metalization

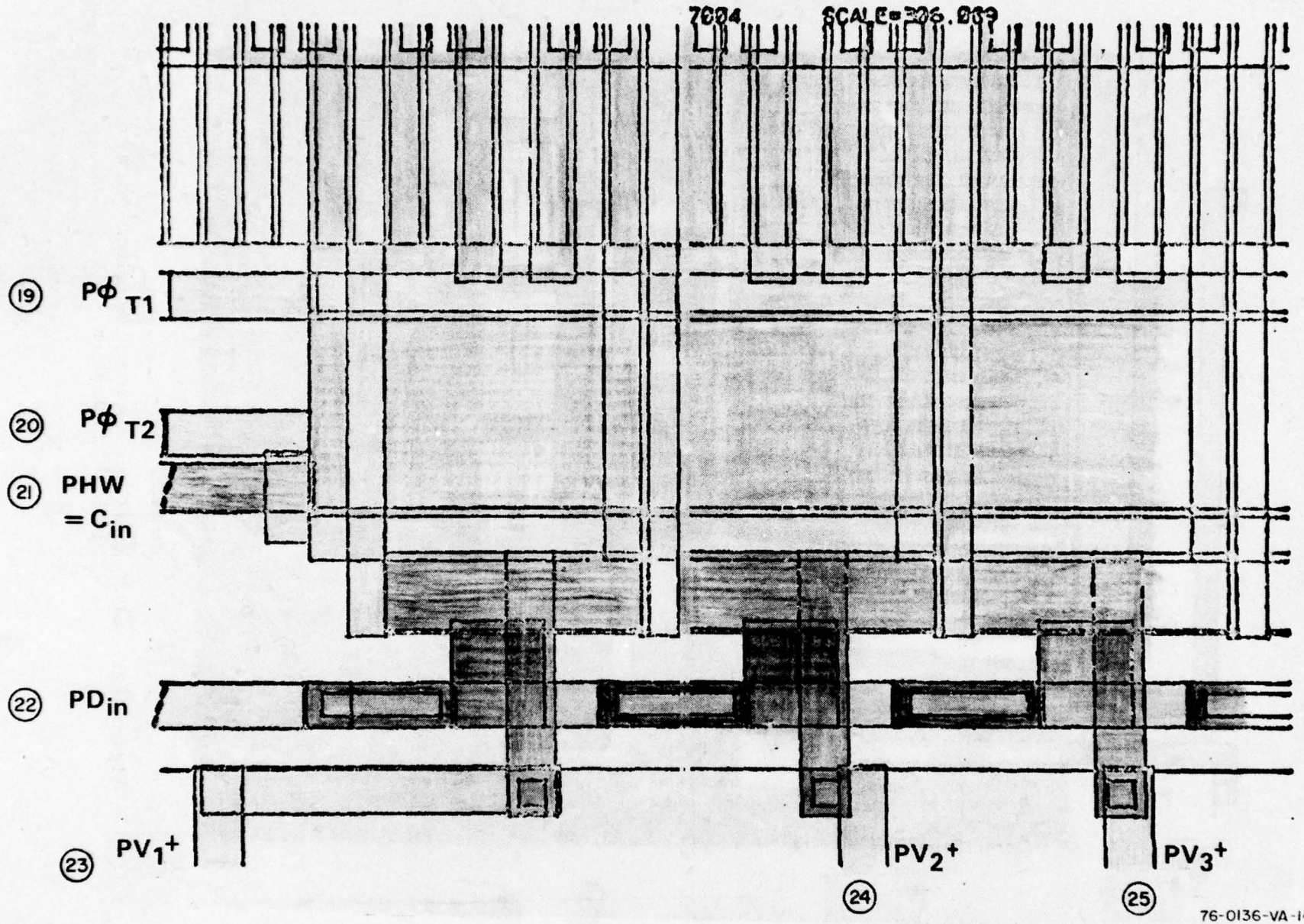


Figure 4-22. Plotter of PU/SO Block with Pin Number Designations
(See also Figure 2-16)

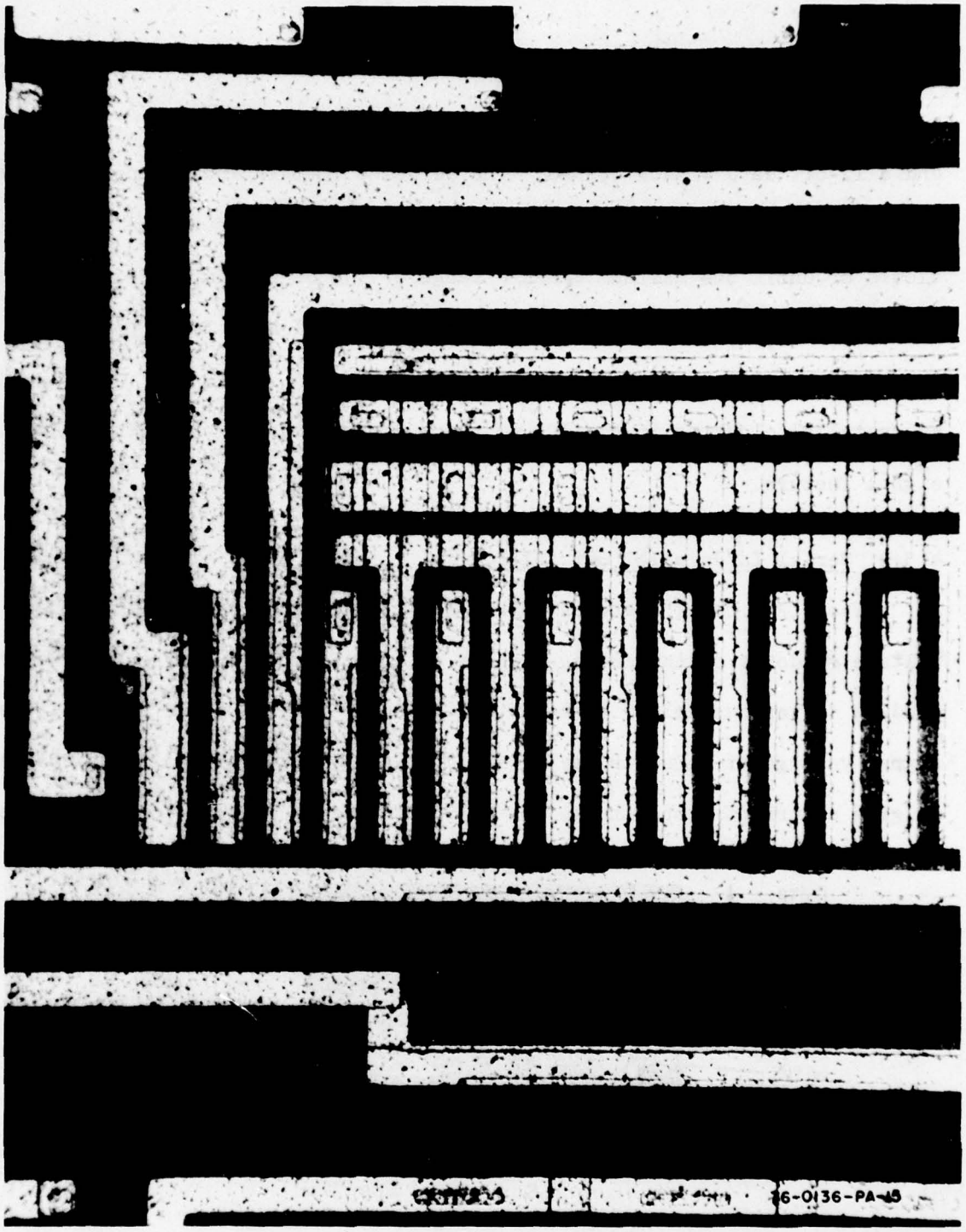


Figure 4-23. Photomicrograph of PI/SO Block

larger with an overall improvement of the $\sqrt{N}$ in S/N. Figure 4-24 illustrates an example of the TDI mode in which the parallel inputs are all tied together and a 12-bit data window opened. The serial output is increasing linearly as the analog bits are added to one another until the input data window is closed or until the maximum dynamic range of the device is exceeded. Figure 4-25 illustrates the successive addition in the TDI mode for 6 parallel inputs with a 100-Hz sinusoidal signal sampled at f_c = 100 kHz. The serial and parallel clocking in the TDI mode are synchronous to provide the additive signal feature.

4.5.4 Serial In/Parallel Out (SI/PO)

The SI/PO block features INDEPENDENT nondestructive, low-impedance voltage readouts of the analog signals at specified locations or taps corresponding to various delays through the CCD shift register. In general, the signal voltage at each tap may be multiplicatively weighted by conductance to give a current proportional to the PRODUCT of the signal voltage by the

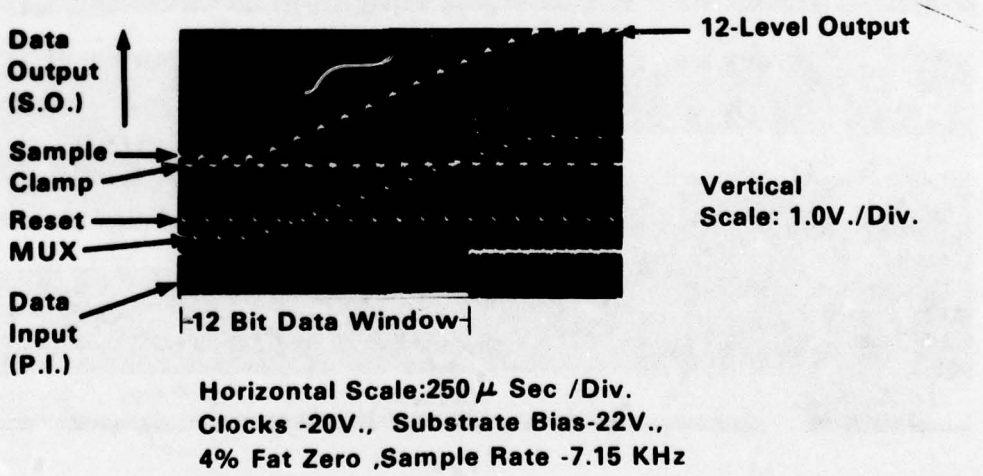


Figure 4-24. TDI Operation for a Constant Input Signal Common to All Parallel Inputs

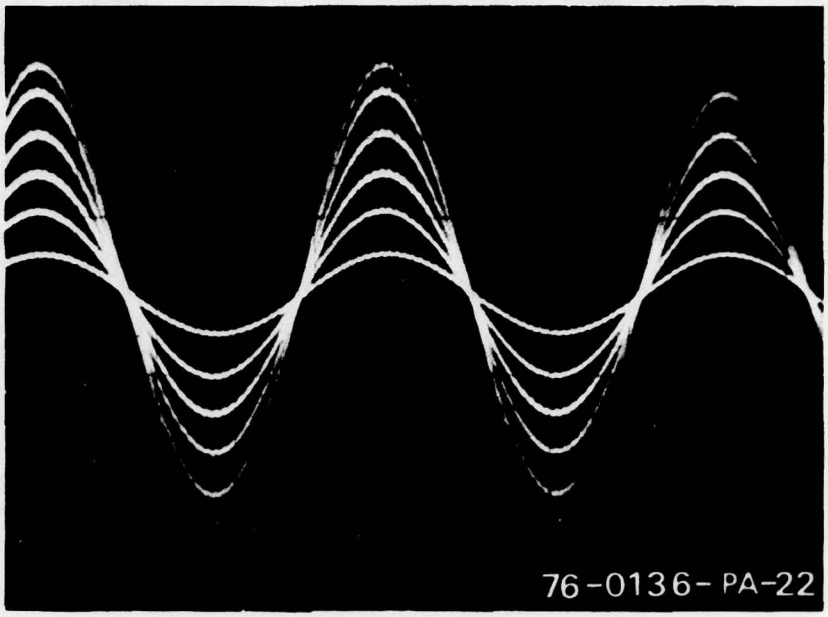


Figure 4-25. TDI Operation for a Sinusoidal (100 Hz) Signal Applied to 6 Parallel Inputs of a PI/SO Device (f_c = 100 kHz)

weighting conductance. Summation of the product currents provide such functions as transversal filtering, correlation, or sampled data smoothing/interpolation for line arrays. Two-dimensional weighting matrices driven by the independent low-impedance taps of the SI/PO block can give discrete Fourier transforms, filter banks, or multiple cross correlators. Figure 4-26 illustrates a photomicrograph of a SI/PO block (N = 20 outputs) which uses a floating clock electrode sensor at alternate stage delays along the CCD delay line. This permits the sensing of a "reference-only" and "signal and reference" signals whose difference yields the net "signal" and eliminates the nonuniformities in the SI/PO structure. Using such a format, numerous taps with multiplicative analog weighting can be accommodated on a small CCD area without signal amplitude degradation due to stray parasitic capacitance. Paralleling SI/PO blocks, each with independent nondestructive read out voltage taps,

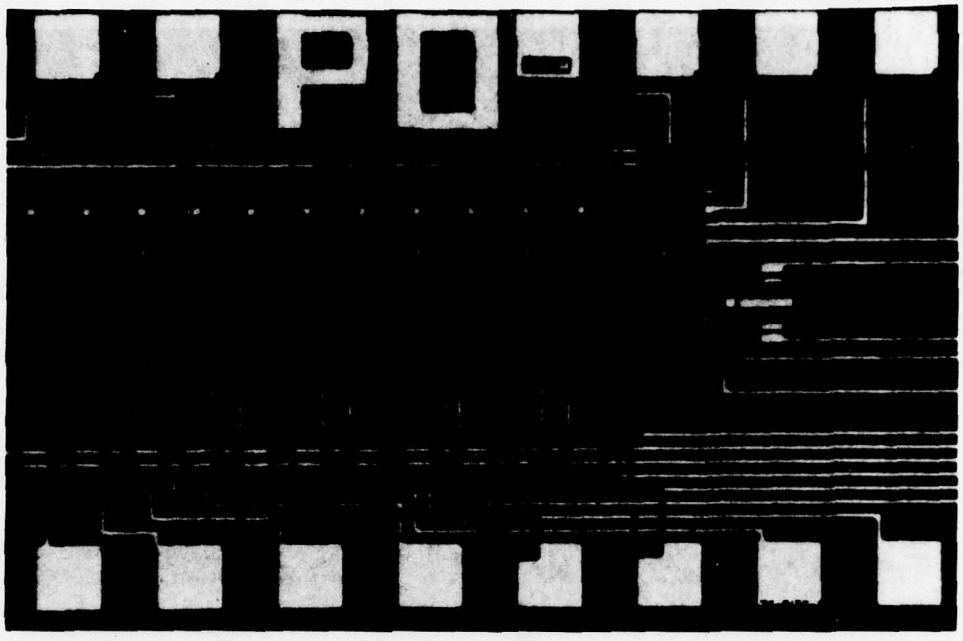


Figure 4-26. Photomicrograph of SI/PO CCD Basic Building Block

can provide the analog voltage signals needed by multipliers to achieve CCD real-time analog correlation. Use of programmable conductances such as the nonvolatile MNOS type or conventional MOS type permit such device applications as adaptive transversal line equalizer or programmable matched filter (or correlation detector) for secure voice/data communications systems.

4.5.5 SI/PO Transversal Filters

Transversal filters may be mechanized with tapped delay lines as shown in Figure 4-27. They are characterized by the summation of fractions of present and past input signals and the absence of feedback. This kind of filter is also referred to as a feed-forward, longitudinal or nonrecursive filter. The output of the filter is the sum of signal samples delayed by a different number of equal increments of time, and each of which may be weighted by a different multiplicative constant. A transversal filter having

a finite number of taps also has a finite impulse response so that it may also be referred to as a FIR filter. Furthermore, the absence of feedback paths prevents (frequency domain) poles in the transfer function of the transversal filter so that all frequency response shaping is done by judicious location of the N (frequency domain) zeroes of a filter having N-1 tapped delay stages. Transversal filters are more tolerant of arithmetic errors in the tap weights than are the recursive filters (which require feedback). The transversal filter can, in theory, be used to synthesize any linear function. In general, the signal input, tap weights, and signal output are complex, considering the frequency domain features of the transversal filter. Low pass (LP) and high pass (HP) filters as well as bandpass (BP) and band stop (BS) filters may be synthesized. The tap weights for LP and HP filters are real, so that inphase and quadrative signal channels can be independent; whereas, the tap weights for image rejection BP and BS filters are complex, which requires "cross-coupling" between the inphase and quadrature signal channels.

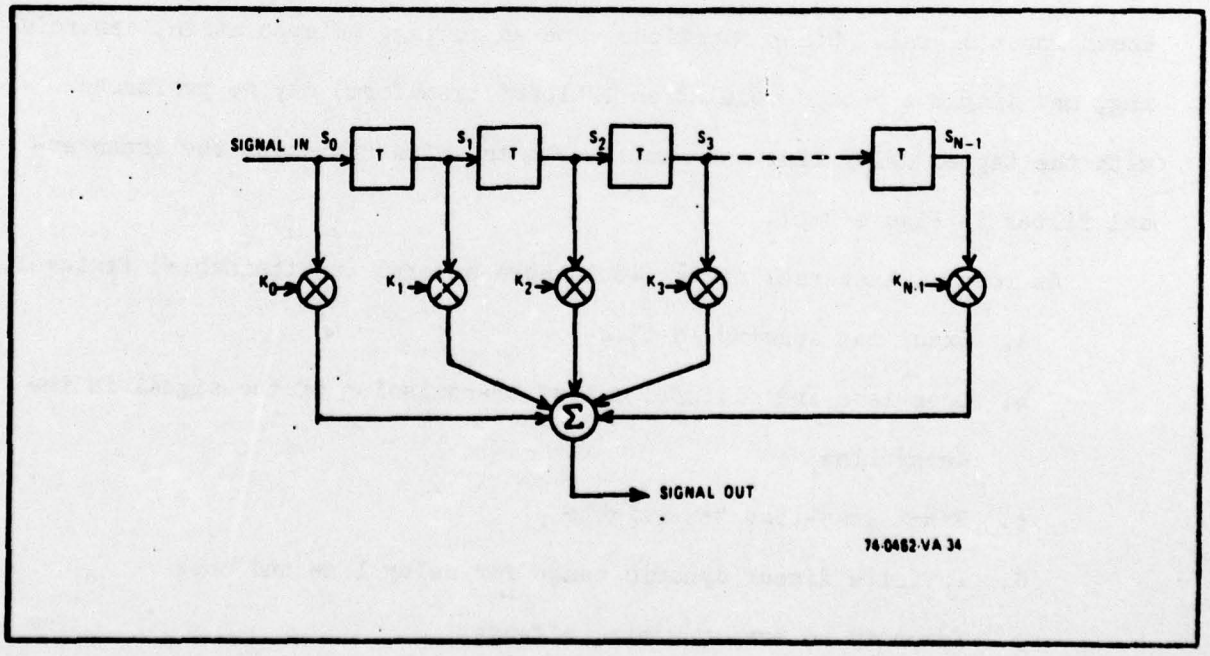


Figure 4-27. Transversal Filter

The bandwidths and center frequencies are normalized to the unit delay, T, shown in Figure 4-27 so that a given filter may be scaled from one center frequency and bandwidth to another by simply changing T. It follows then that the center frequency and bandwidth are limited by the minimum and maximum values of T. These filters have a linear phase versus frequency characteristic provided that the tap weights (amplitudes) have conjugate symmetry about the center of the weighting array.

4.5.6 Time Domain Characteristics

Considering the performance of the transversal filter in the time domain, correlation (or convolution) may be performed by considering the W_k's of Figure 4-28 to be the time reference function. A single impulse applied at the input to the transversal filter will propagate through the delay sections and produce an output proportional to the tap weight of each stage in turn; this describes the impulse response of the filter. Matched filters may be implemented by selecting tap weights which are the time inverse of the known input signal. Other functions such as coding, interpolation, centroiding, and single sideband modulation (Hilbert transform) may be performed with the tapped delay line and summed weights which represent the transversal filter in Figure 4-27.

An ideal transversal filter would have several (unattainable) features:

a. Exact tap spacing in time

b. Loss-less and distortion-less transmission of the signal in the delay line

c. Exact loss-loss tap weights

d. Infinite linear dynamic range for delay line and taps

e. Immunity to environmental effects.

While CCD-based transversal filters are not perfect in any of these respects, the tap spacing can be made significantly better than attainable with

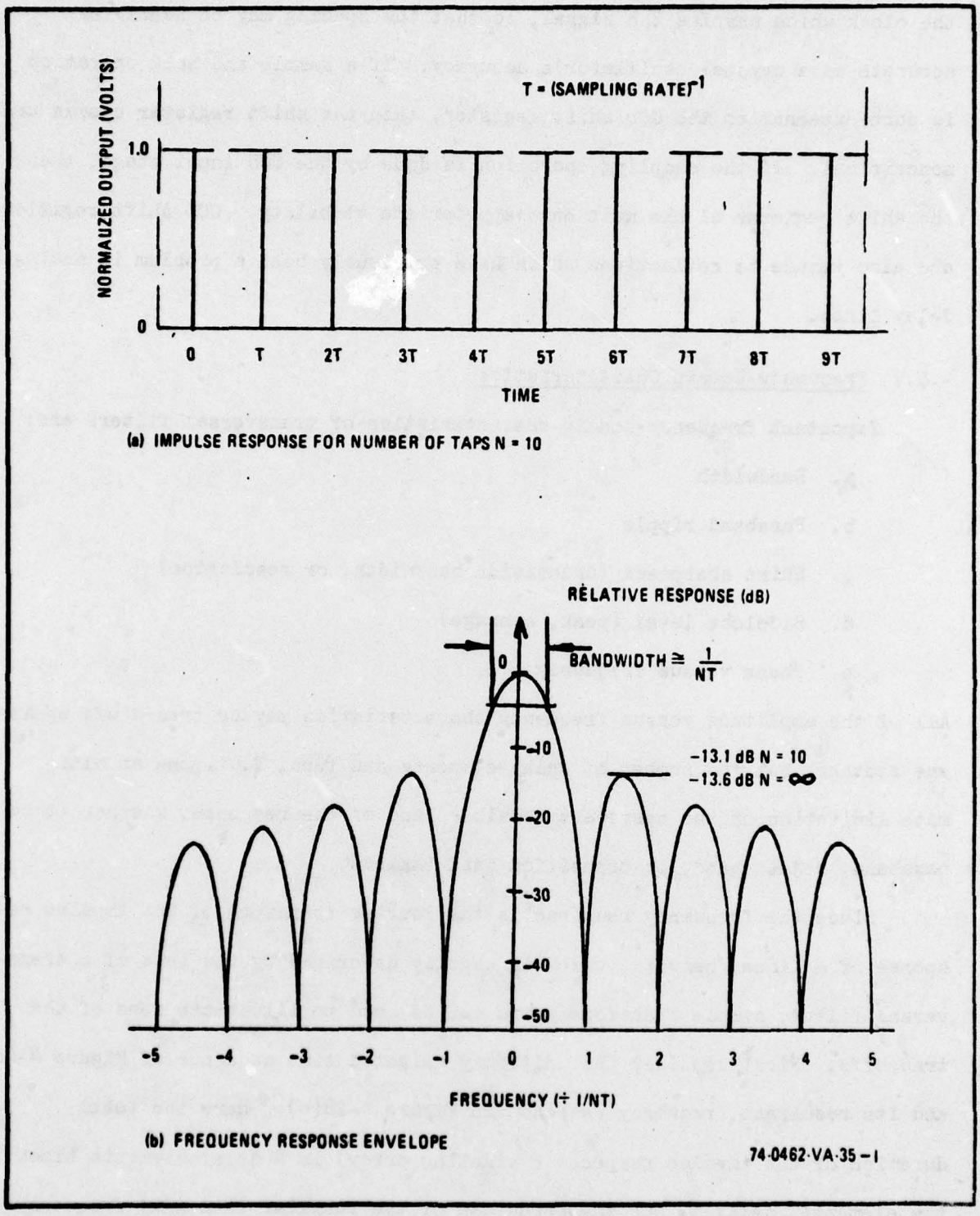


Figure 4-28. Uniform Time Function

any previous analog technique. The effective tap spacing is determined by the clock which samples the signal, so that the spacing may be nearly as accurate as a crystal oscillator's accuracy. If a sample and hold operation is done external to the CCD shift register, then the shift register clocks are noncritical. If the sampling operation is done by the CCD input stage, then the shift register clocks must have appropriate stability. CCD shift registers are also immune to reflections which have previously been a problem in analog delay lines.

4.5.7 Frequency Domain Characteristics

Important frequency domain characteristics of transversal filters are:

a. Bandwidth

b. Passband ripple

c. Skirt sharpness (transition bandwidth, or resolution)

d. Sidelobe level (peak, average)

e. Phase versus frequency.

All of the amplitude versus frequency characteristics may be traded off against one another, but the number of delay elements and taps, N, impose an ultimate limitation on the narrowest possible band of the response, whether it be passband, reject band, or transition band (skirt).

Since the frequency response is the Fourier transform of the impulse response of a linear network, which is exactly described by the taps of a transversal filter, simple transform pairs can be used to illustrate some of the tradeoffs. First consider the uniformly weighted time sequence of Figure 4-28(a) and its resultant frequency response in Figure 4-28(b). Here the total duration of the impulse response (weighting array) in N delay elements times the elemental delay T, and the bandwidth of the familiar (sin x/x) frequency response is approximately 1/NT. Other important characteristics of this frequency response are a total mainlobe width of 2/NT and peak sidelobe response of -13.6 dB, which are the narrowest mainlobe and (in a practical sense) the worst sidelobes attainable with an N point impulse response.

This is a reasonable starting point in considering a tradeoff between sidelobe levels and bandwidth (resolution) since -13.6 dB sidelobes are rarely adequate for any filter. The transform pair of Figure 4-29 represent the Hamming window function. In this case, much lower sidelobes are obtained by the use of a smoother time domain function but at the price of twice the mainlobe width. Many other weighting functions have been cataloged.

4.5.8 Correlation and Convolution

In addition to filtering, the SI/PO block provides convolution with the tap weight, as we have discussed for the PI/SO block, and correlation under the direction of the input signal is reversed as illustrated in Figure 4-30. The powerful operations performed with this structure should lead to the replacement of many digital multiples or adds currently required to achieve 8-10 bit accuracy.

4.5.9 Nondestructive Readout (NDRO) of the CCD Analog Signal

In the realization of a transversal filter, we must provide a low loss analog delay line (namely, the CCD) and NDRO of the signal for subsequent weighting and summation. Two approaches have been discussed in the literature:

a. Surface Potential Sensing with a Floating Diffusion

b. Displacement Charge Sensing with a Floating Clock Electrode Structure.

The first method uses a floating diffusion under the CCD storage electrode to sense the surface potential which is given by the expression,

$$\phi_s = V_0 + V_1 - (V_0^2 + 2V_0 V_1)^{1/2} \qquad (4\text{-}1)$$

with

$$V_0 = \frac{eK_s \epsilon_o N}{2C_o^2} \; ; \; V_1 = V_G - V_{FB} - q_s/C_o \qquad (4\text{-}2)$$

where C_o is the oxide capacitance/unit area under the storage electrode, N the substrate doping, V_G the electrode voltage, and V_{FB} the flat-band voltage. The floating diffusion contacts a gate electrode of a MOS amplifier and is reset by a MOS switch. A major problem with surface potential sensing is

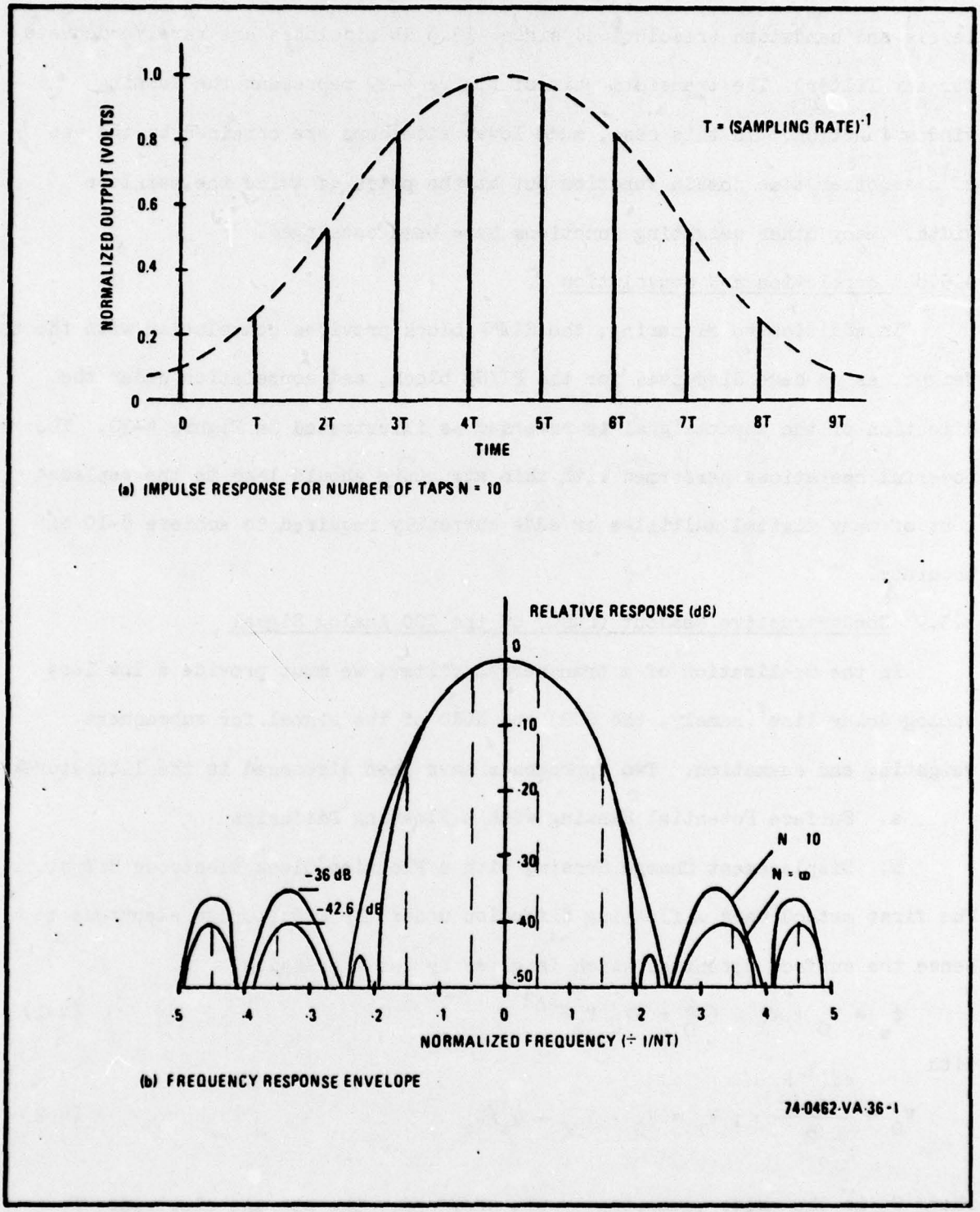


Figure 4-29. Hamming Time/Frequency Responses

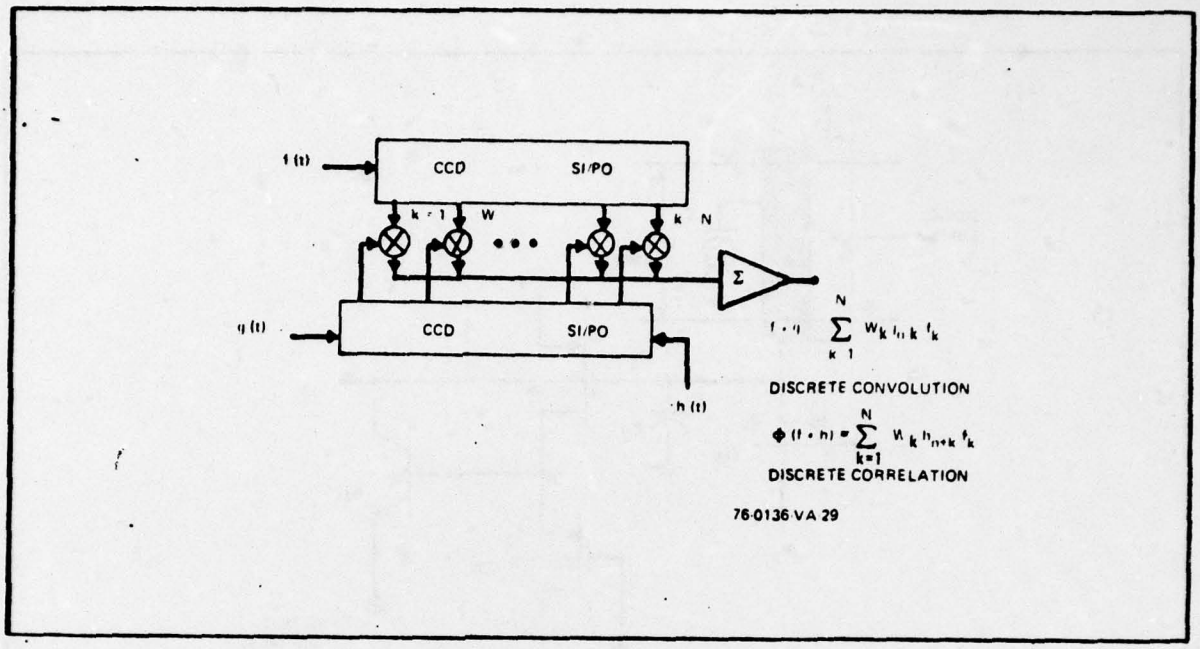


Figure 4-30. SI/PO Functions

the modulation of the storage well barrier height by adjacent fringe fields. The modulation of these barrier heights is responsible for the poor transfer efficiency in conductively coupled CCD and Bucket Brigade Devices structures. Recent awareness of such problems has resulted in the introduction of adjacent "protective electrodes" which are held at a fixed potential to counter the varying fringe fields under s-orage electrode. The voltage on the gate of the MOS amplifier is related to the signal charge by the expression,

$$V_G = \frac{C_O \, q_s}{C_S \, (C_O + C_G) + C_O \, C_G} \qquad (4\text{-}3)$$

where C_G is the external capacitance attached to the sensing node and C_S is the semiconductor space charge capacitance. The same expression is valid for the floating clock electrode sensor (FCLS) shown in Figure 4-31. The FCLS sensor does not require a diffusion beneath the storage electrode and

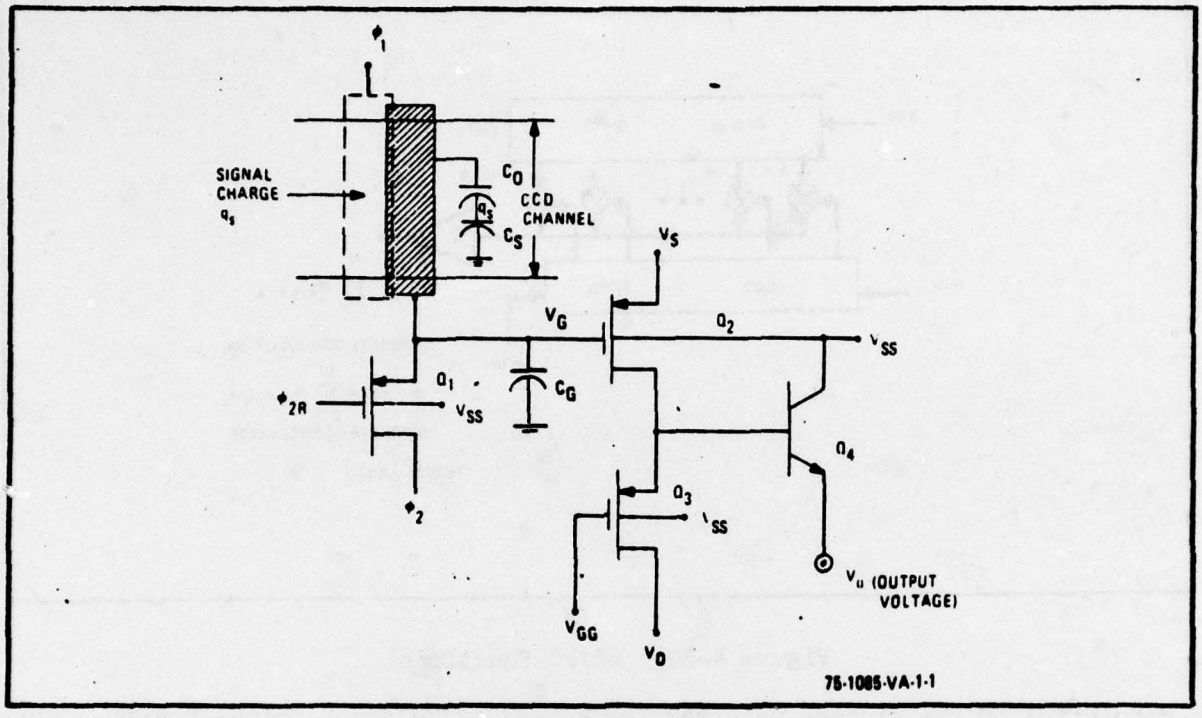


Figure 4-31. Floating Clock Electrode Sensor (FCES) Circuit for Nondestructive Signal Charge Sensing

is therefore not subject to fringe field nonlinearity effects. Because of relative immunity of the FCLS to fringe field effects and barrier height modulation we believe the best method to sense change in a CCD is the FCLS. The semiconductor space charge capacitance is related to the surface potential $_S$, and hence signal charge, by the expression

$$C_S = \left[\frac{qK_S \varepsilon_0 N}{2 \phi_S}\right]^{1/2} \qquad (4\text{-}4)$$

Combining equations (4-1), (4-2), and (4-3) provides a relationship between signal charge and voltage on the gate of the MOS amplifier.

4.5.10 PI/SO and SI/PO Filtering

A new concept has been developed in the analysis of the PI/SO circuit. The time delay and integration (TDI) mode of operation performs progressive or successive additions of the signal as shown in Figure 4-32. If the inputs are preweighted as indicated in the figure, then the output of the serial delay line is simply given as

$$e_o = \sum_{K=0}^{N-1} W_K \, e_{IN} \, (t-kT) \qquad (4-5)$$

assuming we inject at each stage delay. Actually, we inject at alternate sites to permit the incorporation of a reference signal (ac zero) together with the actual sampled signal. The above expression is identical to the transversal filter response of a SI/PO block as shown schematically in Figure 4-32. Thus, the PI/SO block can perform the same functions as the SI/PO without the need for nondestructive tapping, amplification, and buffering at each tap location. There are speed advantages since the electrodes can be placed

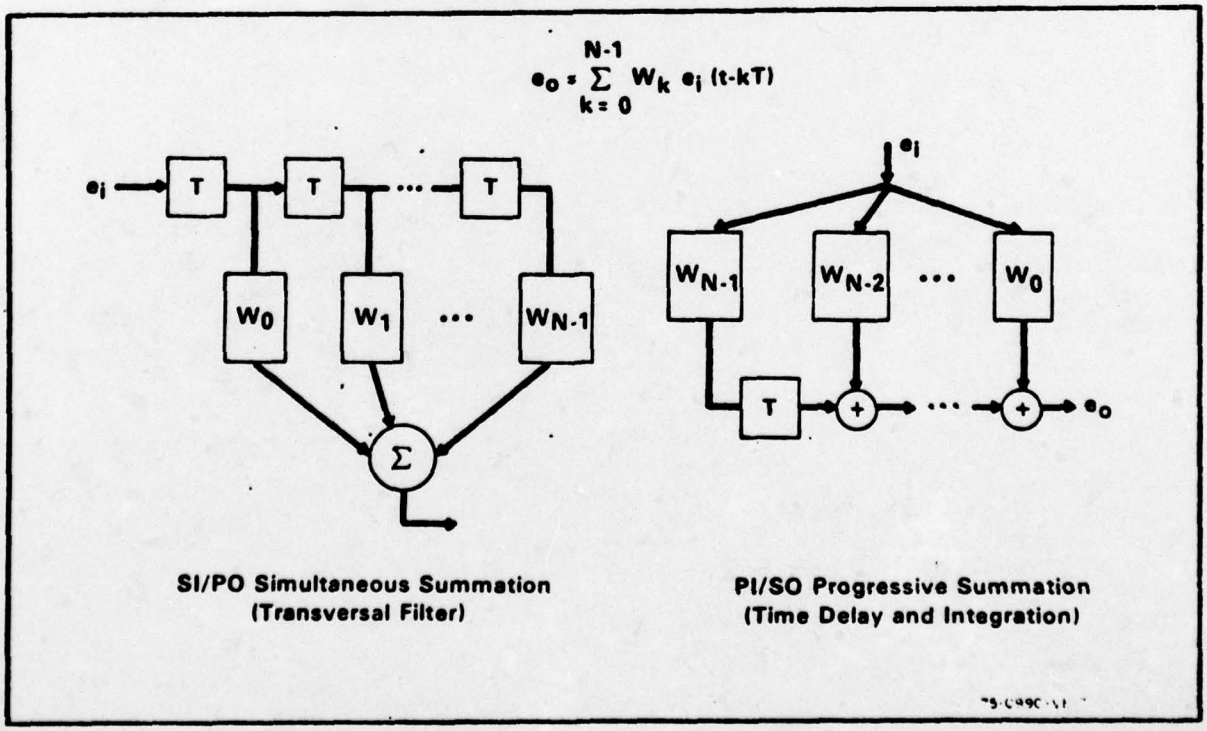
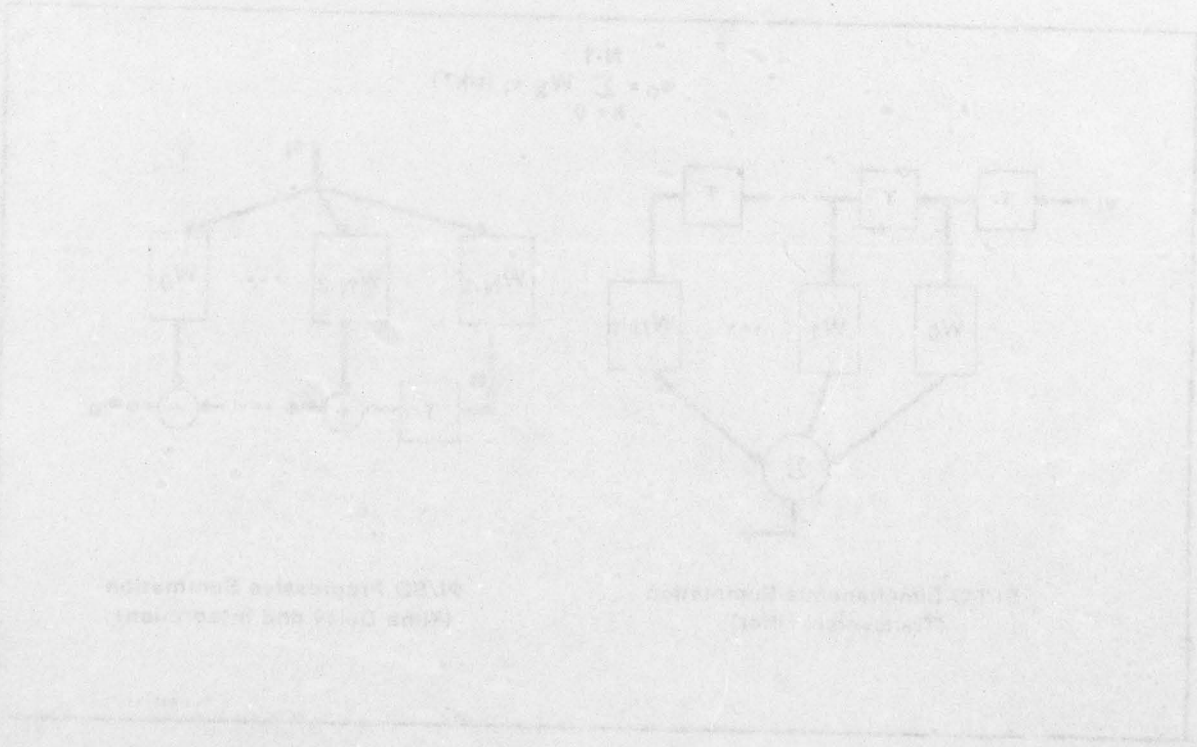


Figure 4-32. Comparison of PI/SO and SI/PO Filter Methods

close together; however, the biggest advantage appears to lie in the control of offsets with no amplifier gain variation to compensate. The disadvantage is the need to provide a tapered width to handle the dynamic range; however, this is not a big problem since the chip size remains almost unchanged (i.e., clocks and bonding pads take up more room than channel widening).

In addition to the filter operation the mathematical operation given by equation (4-5) is described by a convolution of the tap weights W_k with the delayed input serial e_{in} (t-kT). If the tap weights are controlled electrically, then the structure shown in Figure 4-32(a) becomes a reprogrammable filter/ correlator.



Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

Algorithms and Hardware Technology for Image Recognition

5. TYPE OF REPORT & PERIOD COVERED
Quarterly rept. no. 1,
1 May-31 July 1976

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

Tom Willett

8. CONTRACT OR GRANT NUMBER(s)

DAAG53-76C-0138 New

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Computer Science Center
Univ. of Maryland
College Pk., MD 20742

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
ARPA Order-3206

11. CONTROLLING OFFICE NAME AND ADDRESS

U. S. Army Night Vision Lab.
AMSEL-NV-VI
Ft. Belvoir, VA 22060

12. REPORT DATE
30 July, 1976

13. NUMBER OF PAGES
70

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

30 Jul 76 72p.

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

CCD architecture
Image understanding
Image Processing
Pattern recognition
Target detection
FLIR imagery

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This is the first quarterly status report on a program for Recognition Technology for a Smart Sensor, conducted by Westinghouse for Maryland under Contract DAAG53-76-C-0138 with the U.S. Army Mobility Equipment Research and Development Command, Ft. Belvoir, VA 22060

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

403018

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)