ENERGY NORMALIZATION OF TV VIEWED OPTICAL CORRELATION
(Automated Correlation Plane Analyzer for an Optical Processor)

Alex Grumet
Grumman Aerospace Corporation
Bethpage, New York

SUMMARY

An automatic digital correlation plane processor is described that permits rapid location and identification of up to 99 different normalized autocorrelations. Up to 99 unnormalized correlations are acquired in a single TV frame of 1/30 sec. Analog preprocessing circuits permit digital conversion and RAM storage of those video signals with the correct amplitude, pulse width, rising slope, and falling slope. TV synchronized addressing of 3 RAMs permits on-line storage of: 1) the maximum unnormalized amplitude, 2) the image x location, and 3) the image y location of the output of each of up to 99 matched filters. A fourth RAM stores all normalized correlations.

Only the maximum correlation for each of 99 matched filters is determined in one TV frame of 1/30 sec. Successive TV frames look for the next largest correlation output for each optical matched filter.

The silhouette of each image stored in the multiple optical matched filter is also stored in a pair of ROMs that represent the left and right side of the silhouette. Input image energy is collected through an optical high-pass filter of the same bandwidth as the corresponding optical matched filter by a gated, wideband video integrator. The image energy signal is routed to a ROM reciprocal look-up table. Normalization (division by image energy) is accomplished in 1 μsec by a dedicated digital multiplier. A partial out-of-field signal is available to abort normalization and to indicate the approach of loss of track.

INTRODUCTION

The optical matched filter (MF) (reference 1) has not found wide application because acceptable false alarm rates have not been realized by observation of the amplitude of the correlation peaks. When the illumination level associated with an unmatched input image greatly exceeds that of the matched image of comparable size, the MF generally yields a false alarm if MF output correlation peak is the only measure of acceptance. However, if we could divide the above autocorrelation and cross-correlation output peaks each by their corresponding input image energies, the resulting normalized correlations would then be independent of input image energy, and for this situation, the normalized autocorrelation peak would be larger.

The normalization approach described below does normalize the autocorrelation peak but yields an approximate normalization of the cross correlation peak. We, therefore, expect the normalized autocorrelation not to always be larger, but to always be unity, whereas the approximate normalized cross correlation will rarely be unity. The deviation from unity will be used as a measure of the uncertainty of a correct image match. To assist in reducing false alarms, other discriminants such as unnormalized correlation pulse width, rising and falling slopes, as well as unnormalized pulse amplitude are an integral part of the system.

PHENOMENOLOGIC MATCHED FILTER REVIEW

Because we (1) exploit the automatic tracking feature of the MF, (2) use multiple MFs, and (3) employ band-limited MF's to normalize the correlation outputs, a brief description of the MF is in order.

Figure 1 is one of the many possible MF configurations. The back focal plane of the transform lens contains the on-axis, two-dimensional Fourier transform $G(f_x, f_y)$ of $g(x,y)$, the amplitude of the matched input image. When we signify the input image we refer to the image of the object of interest in the input scene. During fabrication it will be the only image in the scene. The spectrum $G(f_x, f_y)$ is complex and two-dimensional and its phase can be recorded on high resolution spectroscopic film by use of a second phase-locked collimated laser beam at angle θ to the signal beam. The resulting stationary interference pattern shown magnified in figure 1 contains a grating, or "subcarrier," whose spacing depends on the angle between the signal and reference beams. Superimposed on the subcarrier grating are the very slight perturbations of the phase contributions of the Fourier transform of the input image. On playback with the reference beam removed and with the exposed and developed MF inserted in the identical fabrication position, the conjugate first order of the MF subcarrier is used as shown in figure 1. In figure 2 we fold the arrangement of figure 1 in line to demonstrate several points. First, for the matched input image the light exiting the MF is $G(f_x, f_y)\ G^*(f_x, f_y)$, where $G(f_x, f_y)$ is the amplitude of the input matched spectrum and $G^*(f_x, f_y)$ is the amplitude transmittance of the matched filter. Therefore, all rays exiting the MF, regardless of their input direction (phase), are parallel. The consequence of this fact is that they focus to a point at the correlation output plane. The focused point in the correlation plane is the summation of all spatial frequencies and, therefore, all the input image energy less system transmission losses. This is not true for an unmatched input image and the cross-correlation peak will be reduced as well as broadened. For a band-limited MF with the low spatial frequencies shown in the magnified MF of figure 1 saturated (overexposed), the autocorrelation function becomes a sharp point, as shown in figure 3(b) and compared to that of figure 3(a) for the full spectrum MF. In practice, the photographic MF plate has a dynamic range of the order of 1000:1 and the energy spectrum of most objects has a dynamic range of the order of 10^6:1, so that the full spectrum autocorrelation of figure 3(a) is not realizable. Second, several MFs can be processed simultaneously through the use of a multiple holographic lens array that is equivalent to several overlapping lenses as shown in figure 2(b). Third, figure 2(a) indicates a transverse displacement of the input image, and, for $F_o = F_1$, the correlation will experience the same displacement with an inversion in the x and y directions. Therefore, input image position in the input plane can be determined by observing the position of the correlation output. By making $F_1 < F_o$, the mapping of the input image plane into the output correlation plane can be reduced to fit the output of several filters into the TV camera tube viewing aperture.

NORMALIZATION APPROACH

Parseval's theorem equates image integrated energy to its spectral

$$\int_y \int_x g^2(x,y)\,dx\,dy = \int_{F_y} \int_{F_x} |G(f_x,f_y)|^2 df_x df_y$$

integrated energy. The left side of the equation, image energy, is measured from the input image and provides the denominator for the normalized correlation. The right side, spectral energy, is measured by the autocorrelation peak to provide the numerator. Provision is incorporated to collect image energy in the same restricted band of the MF. Image energy band limiting is accomplished by selecting one of the 100 holographic lenses for imaging. Instead of a MF at this beam we provide a high-pass or bandpass imaging filter of the same bandwidth as the MF, in the MF plane. The reference beam, on fabrication, is used to direct the band limited image on playback to an imaging lens to provide a different image plane scaling and a different TV pickup other than that used for the correlation plane. In a feasibility experiment now being assembled, a 2 x 2 holo-lens will provide three MF and one imaging lens that will share the same inverse transform lens for MF, as well as imaging, and will share the same TV camera for correlation and imaging.

To collect the image energy, the TV video is integrated only during the time the desired image is scanned. Detectors are square-law devices, and the TV video voltage is proportional to energy. An earlier version of this concept restricted video integration to a rectangle that was centered on the desired image. The x-y RAMs provide the integrating window position over the image. The particular MF, as determined by the RAM address, determines the size of the rectangle. In this manner the matched image will always fall within the rectangle window even though an unmatched image may fall partially outside. However, since the input image and its background are a high frequency version of the actual input scene, we can expect a non-negligible contribution by the background around the matched image within the rectangle. To avoid this situation, and to guarantee that the normalization of the autocorrelation will be accurate and free of background, the shape of the rectangle was altered to that of the silhouette of the matched image.

Two ROMs are used to store the silhouettes of all matched images. We will refer to one as the START ROM and the other as the STOP ROM. The silhouette in each interlaced field of a TV frame is located at contiguous addresses in the START and STOP ROMs. For silhouettes where a TV scan line leaves and reenters, an additional pair of ROMs will be required to avoid integrating background energy. A TV frame of 1/30 sec is required to collect the image energy of each selected unnormalized correlation. Preprocessing circuitry limits the number of unnormalized correlations to be normalized.

APPLICATION

The need for normalized correlations can be appreciated if we have to select or detect a possible autocorrelation out of thousands of correlations in a few seconds. This was the motivation for the development of the system to be described.

The OMFIC (Optical Matched Filter Image Correlator) is described in reference 2 and is presented in schematic form in figure 4. The system is capable of processing an input image through 100 optical matched filters simultaneously. Therefore, an image of an input scene containing 20 objects will yield 2,000 correlations in 2 or 3 ns at the OMFIC output plane. The output correlation plane is scanned by a conventional TV camera and all 2,000 correlations can be viewed on a TV monitor.

The problem addressed was to select any autocorrelations out of the array of correlations, and identify the image (associated matched filter) and the image x-y coordinates in a few seconds. Correlation normalization was one of the discriminators used to reduce false alarms. If the matched images were trucks parked side by side, the resulting autocorrelations could be a few microseconds apart, and this ruled out the use of microprocessors to process the TV video. TTL with 20 ns switching time are adequate and, therefore, the bulk of the system described below consists of integrated TTL circuits. A μP is used to provide slow information in a few 100 μs time frame and to supervise overall operations, conduct searches, and provide displays and printout messages.

SYSTEM DESCRIPTION

Figure 5 is an overall system block diagram. The TV video is preprocessed so that only video pulses of the correct amplitude, of the correct pulse width, and with the correct rising and falling slopes will be accepted to be converted to 8-bit digital data in a 600 ns A/D converter. Figure 6 shows the preprocessing gate timing. The unnormalized correlations are stored in a 256 x 8 bit RAM. There are three additional RAMs, all 256 x 8 bits, and two of these are synchronously addressed with the TV raster sweep, along with the unnormalized correlation RAM as described below. Two of the RAMs store image x-y coordinates, and the fourth RAM stores the normalized correlations.

Two timing counters accurately address the RAMs and locate the position of all correlations to the required precision on the TV raster. The horizontal counter consists of a digital phase-locked-loop (PLL) and provides the four least significant bits (LSB) for the RAM 8-bit address bus. The vertical counter is essentially a TV line counter providing the four most significant bits (MSB) for the RAM address. The H and V counts can be hardware or software altered to agree with the number and configuration of the matched filter array. A test pattern can be displayed on a TV monitor for system setup and assistance in matched filter fabrication output addressing. Such a pattern is shown in figure 7 for a 10 x 10 matched filter array. Each of the 100 boxes in the checkerboard test pattern represents a RAM address. The bright spot in the upper left-hand corner of each box (more visible in figure 3(b)) is the RAM erasure pulse of 100 ns duration of every other TV field. Since the RAMs are synchronously addressed with the TV raster, the data at any one address can be displayed on the TV monitor. The MSB of the data bus was connected to the TV monitor video input to demonstrate that the correlation peak was stored at the correct address. To ensure that only the peak of the video pulse is stored in each address, an 8-bit digital comparator tests the incoming 8-bit video peak with the RAM stored value and will dump the stored value if it is smaller. In this manner, only one correlation peak is stored in each address. If several correlation pulses are present for a particular MF (in the same box of the checkerboard), the stored value will be inhibited by the μP on the next TV frame after normalization. The next largest correlation peak of the particular filter will then be stored. After several TV frames all correlation peaks associated with each matched flter will be processed. The 8-bit comparator output pulse triggers a write pulse for the unnormalized and the x and y RAMs. Therefore, the maximum unnormalized correlation peak of each filter is stored in a RAM address assigned to that filter and the x-y coordinates associated with the correlation are stored at the same address in their respective x and y RAMs. At the end of each TV frame the RAMs are erased and the process is repeated.

During every other TV vertical blanking interval (once per TV frame) of approximately 1.7 ms all three RAMs (unnormalized correlation, x and y coordinates), together with the normalized correlation RAM, are scanned very rapidly by switching in a high-speed addressing counter. At each address the contents of the unnormalized and the normormalized RAMs are interrogated for any content. If no normalized number is stored where an unnormalized number has been found, then the normalization process will be initiated during the next TV frame, and the first step in the normalization process is the storage of 1) the RAM address, 2) x count, 3) y count, and 4) the unnormalized correlation.

In figure 8, we define the image height and location, referenced to its location, as stored in the x and y RAMs. The left shaded side of the image, designated SR, is stored in the START ROM at contiguous addresses for successive TV lines in a field, and the interlaced field is stored at another group of contiguous addresses. The same image is stored at the same addresses in the STOP ROM for the right edge of the image, designated SP. When the horizontal sweep and, therefore, the x counter of the timing chain is equal to the value stored in the START ROM, the associated comparator output will switch on a high-speed, wideband video integrator, triggered by SR, and, in a similar manner, the STOP ROM will switch the integrator off, triggered by SP. The height, TL + BL, of the image triggers a height FF to gate in only the correct SR and SP ROM addresses. To correctly address the START and STOP ROMs for: a) the correct image, b) correct field, and c) correct positon in x and y requires offsets on all position data. Some of these are slow (TV field intervals) and are provided by the microprocessor. The 8-bit counts that are required in less than a TV line interval of 63.5 μs are hard wired.

The RAM address identifies the MF and, hence, the image. The location of this image, in the START and STOP ROM, must be accessed as the TV sweeps through the image. The value $Y_o - T_L$ of figure 8 defines the image top line, and the μP provides the offset for the corresponding ROM location of this image. In this manner, the vertical timing chain line-counter is offset to correctly address the ROMs as the image is scanned. At the beginning of each TV scan line, the ROM stored value is subtracted from the X_o stored value to provide a count for the leading edge of the image (X_o - SR). An 8-bit comparison of this value with the horizontal PLL timing chain identifies the left edge of the image to start the video integrator. In a similar manner (X_o + SP) will stop the integrator for that particular TV line. After a complete TV frame, the integrator output is converted to 8 bits and a reciprocal ROM provides the denominator to a fast (1 μs) 8 bit by 8 bit multiplier with the unnormalized correlation. Storage of this value in the normal correlation RAM completes the normalization cycle. During the TV frame that normalization is underway, the X-Y RAMs, as well as the unnormalized RAM, continue to be updated as before. The μP reads the normalized RAM and checks for deviation from unity. An appropriate video overlay circle or square can encircle the correlation peak, and the blinking rate of the overlay can indicate deviation from unity.

The block diagram of figure 9 indicates how the operations described above are realized in a straightforward fashion. To determine if the image is partially cut off by the raster boarder, we evaluate four quantities. If $(Y_o - T_L) < 0$, then the image is partially off the top of the screen. Similarly, if $(Y_o + B_L) >$ 244/K (where K is the number of rows of MF, and, therefore, correlation address rows), the image is partially off the bottom of the screen. If (X_o - SR) < 0, then the image is partially off the left side of the raster, and if (X_o + SP) > 100 (for a 100 count raster), then the image is partially off the right side of the screen. In all four of these conditions, the normalization will be aborted and an appropriate message will be indicated by the μP. This same information could be used for an automatic tracker associated with the OMFIC that is about to lose its target and the direction to move to maintain the track is indicated.

NORMALIZATION OF CROSS CORRELATIONS

We can apply Parseval's theorem to the cross correlation situation

$$\int_y \int_x g_2(x,y) g_1(x,y)\, dxdy = \int_{F_y} \int_{F_x} G_2(f_x,f_y) G_1^*(f_x,f_y)\, df_x df_y$$

where $g_2(x,y)$ is an unmatched input image amplitude and $g_1(x,y)$ is the matched image. However, we do not have a single peak in the correlation plane that represents the integrated cross product because G_2 and G_1 do not have conjugate spectra, and the spectra do not have the same spatial frequencies. A cross correlation peak need not be an even function (as it is for an autocorrelation peak), and its location can be located almost anywhere within the input image if we map it back to the input plane. The ROM stored silhouette location is positioned by the x-y RAM stored correlation peak as indicated in figure 8. There are three different situations that we can encounter: 1) the extent of g_2 > the extent of g_1, 2) the extent of g_2 < the extent of g_1, and 3) g_1 and g_2 do not completely overlap. In case (1) where the silhouette g_1 falls entirely within g_2, the input image, we do collect less than the full input image energy, and the denomination of the normalization is smaller than for a perfect g_2 silhouette. It is, therefore, possible to have a normalized cross correlation greater than unity. In case (2) we can have a smaller than unity normalization; however, we do collect background, and this can alter the normalization. The same is true for case (3).

Therefore, although we can make a definitive statement that the normalized autocorrelation is unity, we can expect the normalized cross correlation to vary considerably with some slight possibility of a value of unity.

No effort was made to exploit the even symmetry of the autocorrelation peak in the preprocessor. If required, this extra feature can be implemented in a straighforward manner.

CONCLUSIONS

The automatic correlation plane processor described here can rapidly acquire, identify, and locate the autocorrelation outputs of a bank of multiple optical matched filters. The ROM stored digital silhouette of each image associated with each matched filter allows TV video to be used to collect image energy to provide accurate normalization of autocorrelations. The resulting normalized autocorrelations are, therefore, independent of the illumination of the matched input. Deviation from unity of a normalized correlation can be used as a confidence measure of correct image identification.

REFERENCES

1. Vander Lugt, A. B.: Signal Detection by Complex Spatial Filtering, IEEE Transactions on Information Theory, April 1964, pp 139-145. A more complete treatment is given in AD 411473, a University of Michigan Report 4594-22-T, July 1963, issued by the Radar Laboratory of the Institute of Science and Technology.

2. Grumet, A. and Reich, A.: Optical Matched Filtering for Missile Guidance, Final Report Contract N00019-73-C-0303, NASC, April 1974.

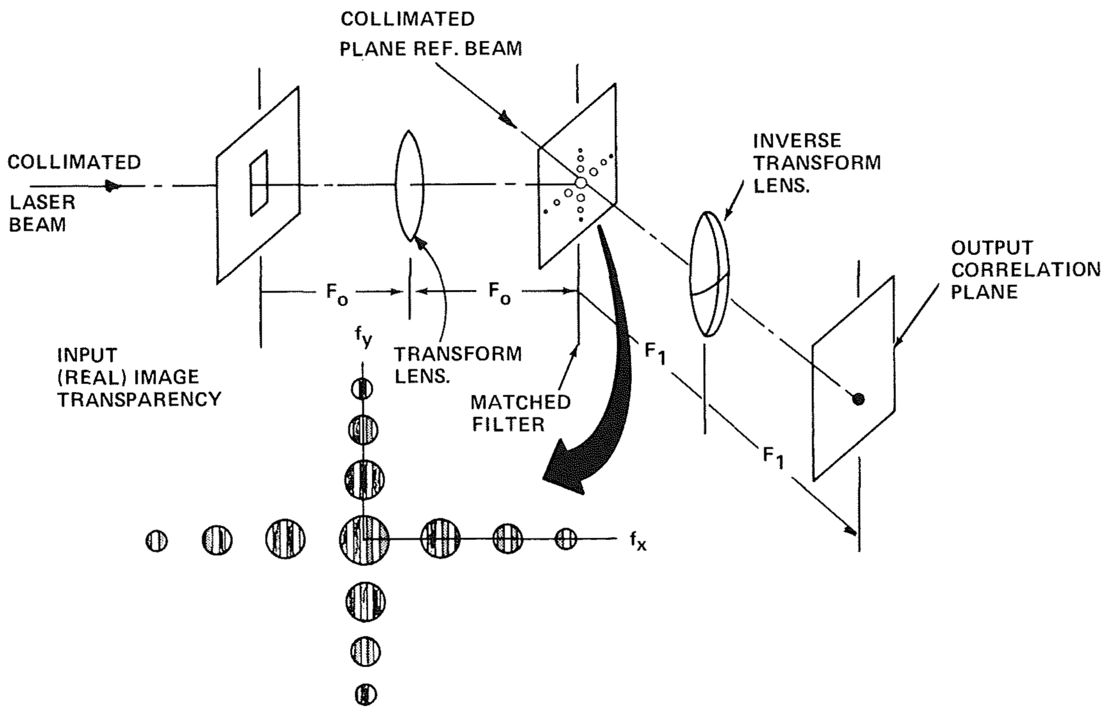


Figure 1.- Matched filter configuration.

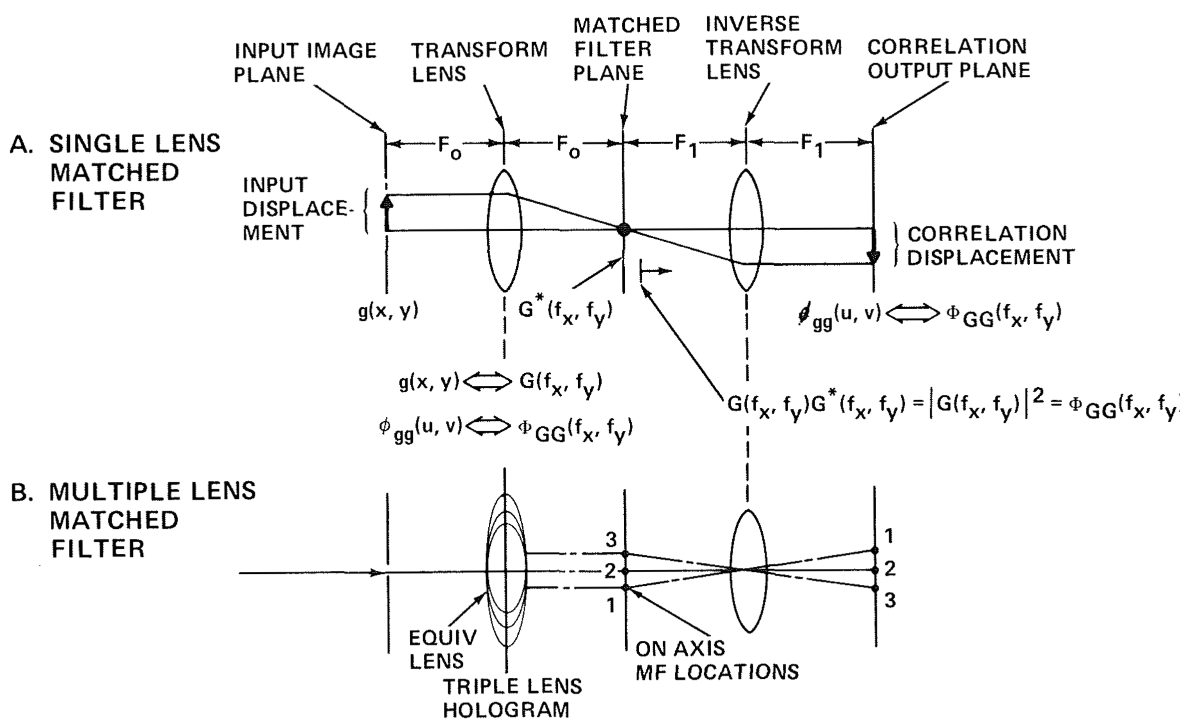


Figure 2.- Some aspects of matched filters.

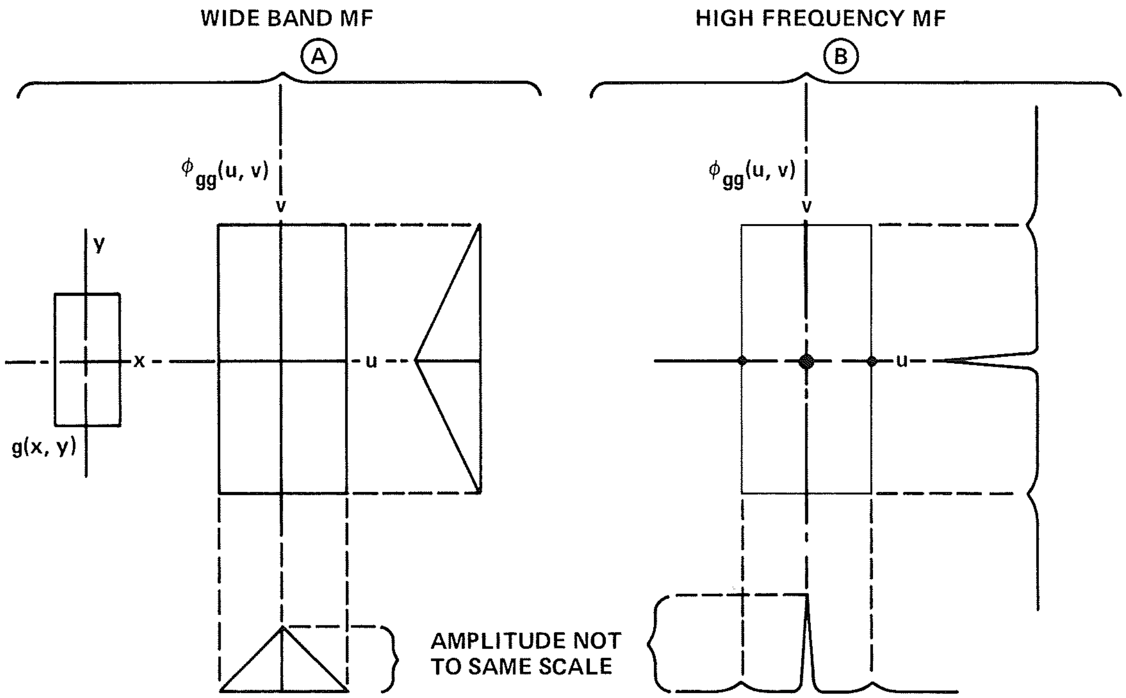


Figure 3.- Wide band and high frequency autocorrelation functions.

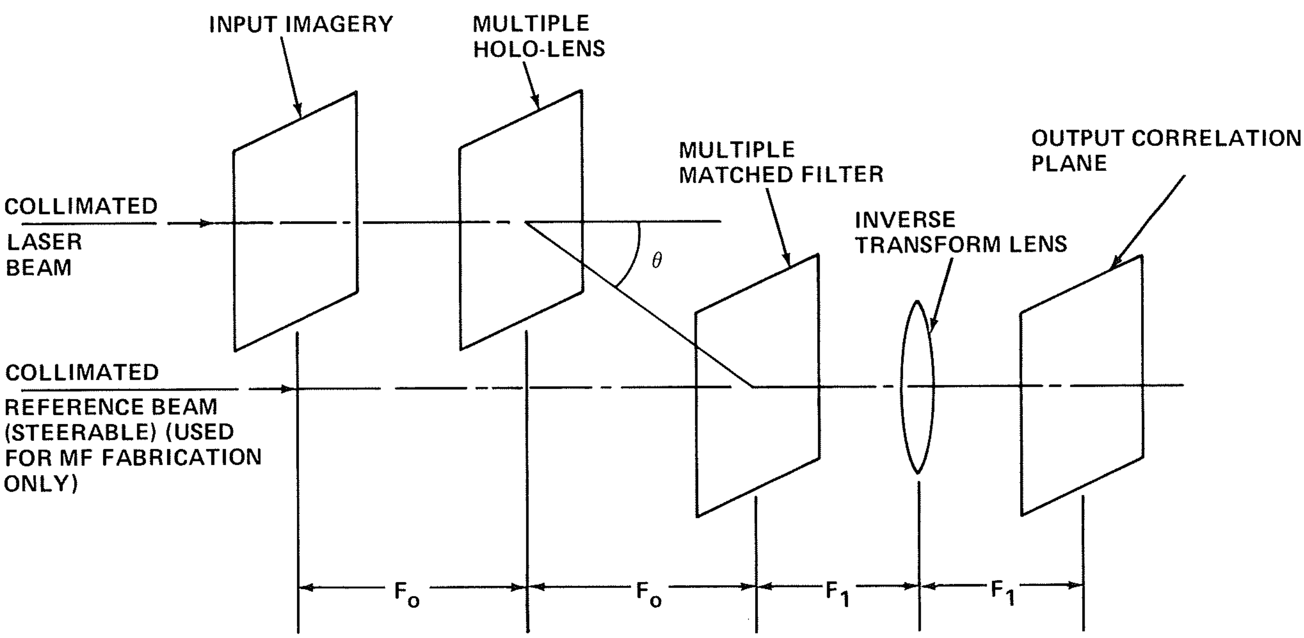


Figure 4.- Optical matched filter image correlator (OMFIC).

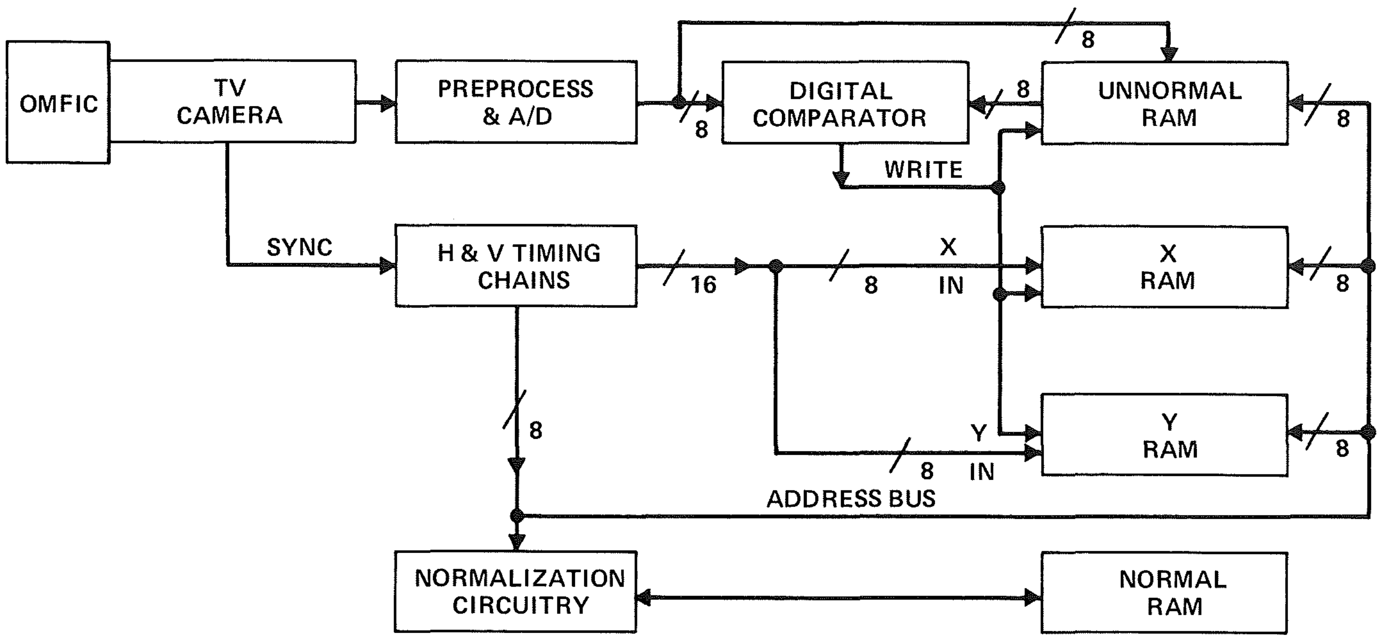


Figure 5.- Overall system block diagram.

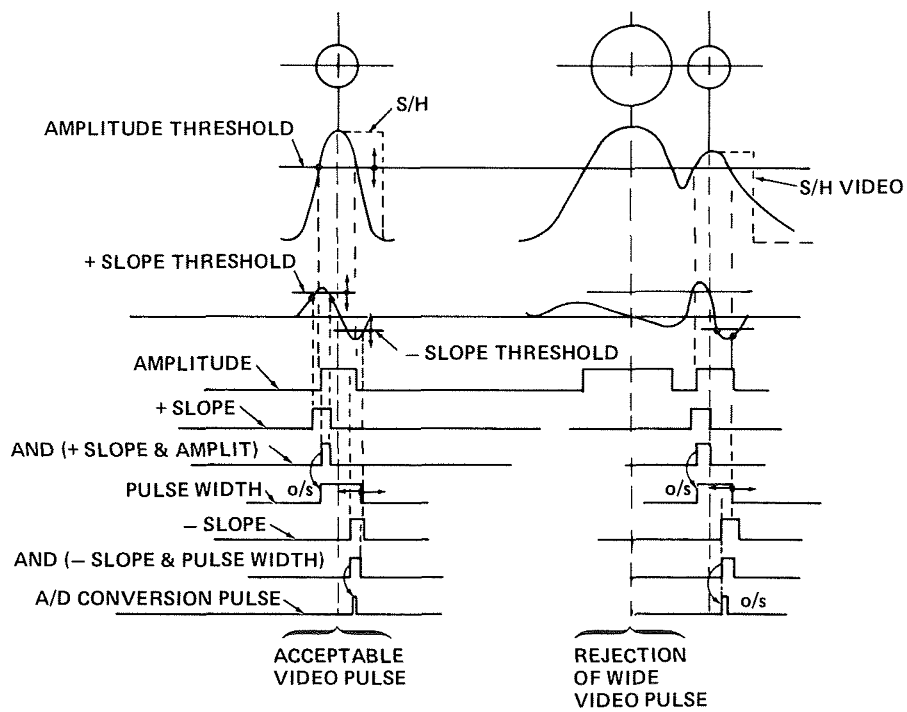
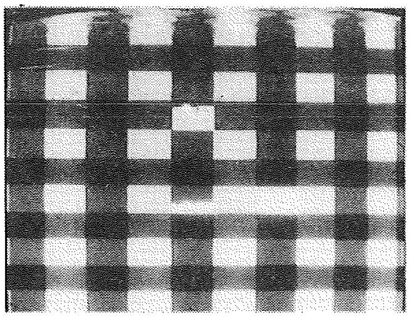
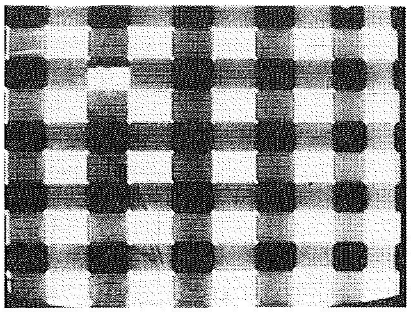


Figure 6.- Video preprocessing waveshapes (condition for A/D conversion).

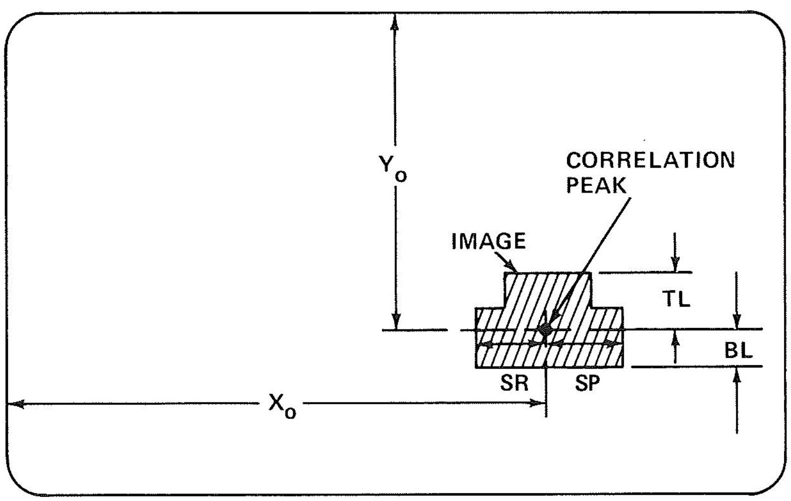


(a) THREE CORRELATIONS PRESENT



(b) ONE CORRELATION PRESENT

Figure 7.- Ten by ten test pattern.



X_o = x COORDINATE OF THE CORRELATION STORED IN X–RAM
Y_o = y COORDINATE OF THE CORRELATION STORED IN Y–RAM
TL = TOP LINE OF IMAGE
BL = BOTTOM LINE OF IMAGE
SR = 8-BIT NUMBER STORED IN START ROM
SP = 8-BIT NUMBER STORED IN STOP ROM

Figure 8.- Definition of terms.

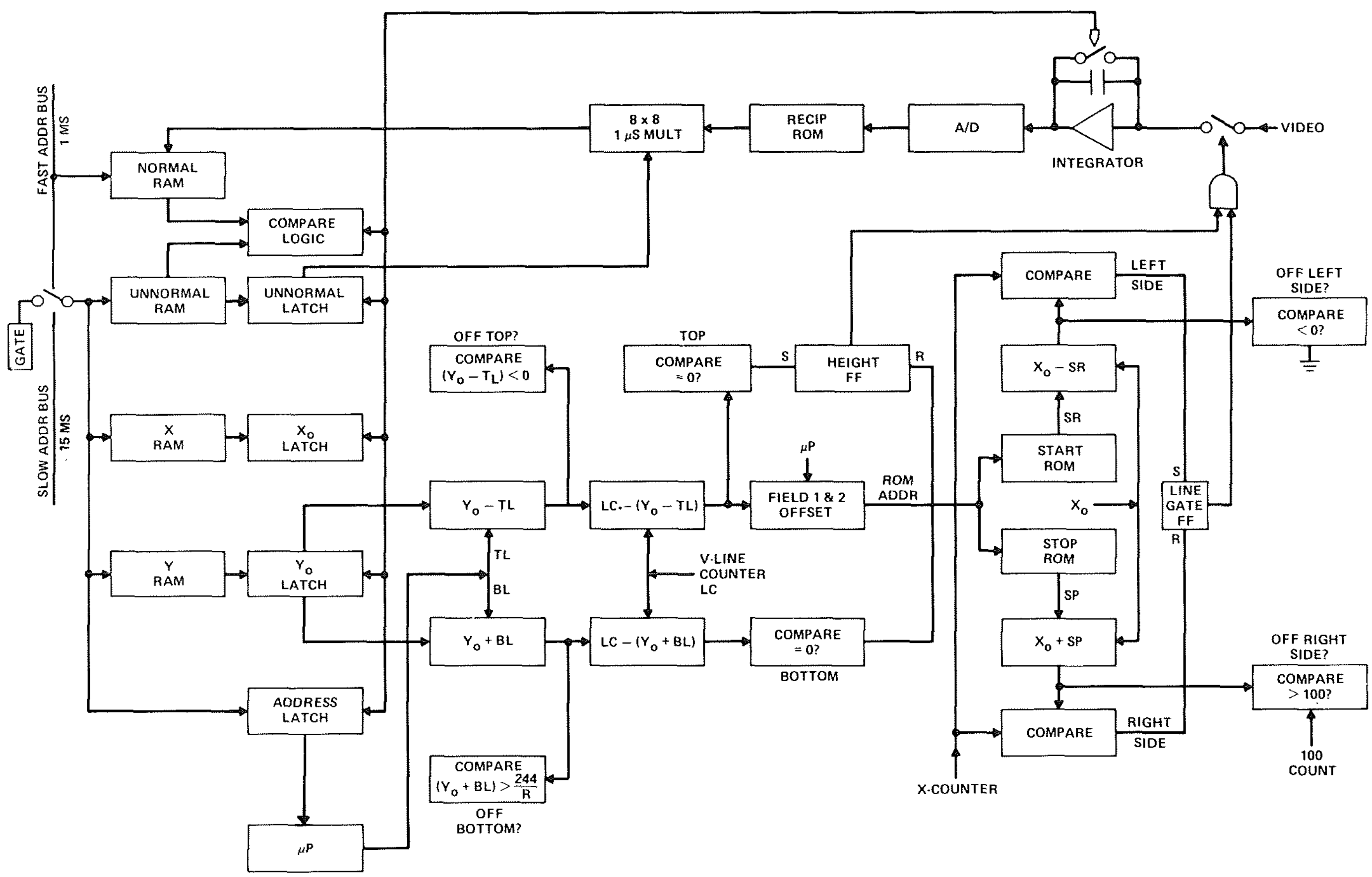


Figure 9.- Normalization block diagram.