Sequential Hierarchical Scene Matching

ROBERT Y. WONG, MEMBER, IEEE, AND ERNEST L. HALL, MEMBER, IEEE

Abstract—The general approach to matching two scenes by a digital computer is usually costly in computations. A match is determined by selecting the position of maximum cross correlation between the window and each possible shift position of the search region. A new approach which is logarithmically efficient is presented in this paper. Its logarithmic efficiency and computational savings will be demonstrated both theoretically and in practical examples. Experimental results are presented for matching an image region corrupted by noise and for matching images from optical and radar sensors. The significance of this approach is that scene matching can be accomplished by the use of a computer even in cases which are difficult for humans or standard correlation techniques, and can be accomplished with greatly reduced computations.

Index Terms—Hierarchical search, image processing, scene matching, sequential pattern recognition.

I. INTRODUCTION

The problem of matching two images of the same scene taken by different sensors under different viewing conditions is a challenging problem in the field of computer applications to image processing and pattern recognition. The scenes are usually transformed so drastically by the different viewing geometries and sensor characteristics that it is extremely difficult, if not impossible, to match the original images without the proper data preprocessing.

The general approach to matching two scenes involves locating the corresponding regions in the images. If several corresponding regions can be located, geometric and intensity transformations [1] can be performed on the window and the search region in preparation for a match.

The standard computer approach to scene matching by correlation is usually very costly in computation. A match is determined by selecting the position of maximum cross correlation between the window and each possible shift position of the search region. For a window of size \( M \times M \) and a search region size \( N \times N \), there are \((N - M + 1)^2\) possible test locations. Although fast correlation techniques such as fast Fourier transforms decrease the correlation computations, these techniques still require computations at each of the \((N - M + 1)^2\) locations. In this paper, a new approach to scene matching is described. This approach, which is computationally three orders of magnitude more efficient than methods currently in use, is capable of matching scenes of many types. Its logarithmic efficiency and computational savings will be demonstrated both theoretically and in practical examples.

II. HIERARCHICAL SEARCH

The proposed approach incorporates a hierarchical search for a possible match location starting at a low resolution level. During the search at each level of resolution, sequential testing and detection techniques [2] are applied to further minimize the amount of computations.

A process was developed to create a set of images which are decreasing lower in resolution and smaller in size. Two sets of these images are created, one for the window and the other for the search region. The low spatial frequencies preserved in the low resolution images are utilized to find the scene of interest at considerably lower cost. At the lowest level of resolution, the number of possible test locations is reduced to \( [(N/2^{L}) - (M/2^{L}) + 1]^2 \) where \( L \) is the search level. Comparing this to the possible test locations of \((N - M + 1)^2\) at the highest resolution when \( L = 0 \), there is a reduction of nearly \( 2^{2L} \) in possible locations of computations.

A threshold sequence and decision rules are developed to guide the search from a low resolution level to the next higher resolution level. The decision rules are designed to select the most promising test locations at each search level. Only the selected locations are tested at the next level. By this technique, only the most promising test locations are examined at the higher resolution levels. Therefore, this search technique is logarithmically efficient, i.e., the number of search positions is reduced to \( K \log (N - M + 1)^2 \) where \( K \) is a constant.

At each test location, an ordering algorithm is used to order the \((M/2^L)^2\) window pairs. An algorithm similar to one described in [2] can be used to select the window pairs in a random nonrepeating sequence or the ordering may be adapted in a data-dependent fashion. For example, for the matching of edge pictures, ordering may be done along the edge gradient. An error measure is established and the errors of the window pairs are accumulated at each test location. At the \( n \)th window pair computation of the \( k \)th search level, if the accumulated error is equal to or greater than \( T_k \), the computation terminates and the test location is eliminated.

Manuscript received February 10, 1977; revised August 14, 1977. This work was supported by the Space and Missile Systems Organization under Contracts F04701-76-C-0224 and F04701-77-C-0072.

R. Y. Wong is with the Department of Electrical Engineering, California State Polytechnic University, Pomona, CA 91768.
E. L. Hall is with the Department of Electrical Engineering, University of Tennessee, Knoxville, TN 37916.
If all window pairs are exhausted and the accumulated error is less than \( T^*_k \), the test location is recorded for further consideration at the next higher resolution level. It can be seen that for a properly chosen \( T^*_k \), many fewer than \((M/2^k)^2\) window pairs need to be computed for test locations at distances far from the match point.

At each search level \( k \), except for the highest level at the lowest resolution, a search is made only at test locations recorded in the \((k + 1)\)th level. As the search level \( L \) decreases (i.e., the resolution of the image increases), increasingly fewer test locations need be considered. This is due to the fact that as the search advances, the distances separating the true match location from the other test locations increase, and the threshold sequence \( T^*_k \) is designed to eliminate test locations which are furthest away from the true match point. By this technique, only the most promising test locations are examined at the higher resolutions. The search continues until either the match point is located or the highest resolution is reached at which the test location with the highest probability of a match is selected as the most likely candidate.

### III. Hierarchical Search Analysis

Although no previous work on hierarchical scene matching has been discovered, several previous experiments on hierarchical decomposition of pictures have been made. Klinger [3] and Klinger and Dyer [4] developed a regular decomposition of picture areas into successively smaller quadrants. Tanimoto and Pavlidis [5] utilized a hierarchical data structure for picture processing to speed up edge detection operations. Horowitz and Pavlidis [6] provided an algorithm for picture segmentation by a directed split-and-merge procedure. Ramapriyan [7] used multilevel search techniques for edge detection. A two-stage, coarse-fine template matching technique was described by Rosenfeld and Vanderbrug [8]. Digital template matching with ordered search techniques was used by Nagel and Rosenfeld [10].

In the hierarchical search technique, a structured set of pictures at different resolutions is created. The process of creating a lower resolution plane at the \( k \)th level involves two-dimensional low-pass filtering of the picture data of the \((k - 1)\)th level and then sampling the filtered data at \( \delta \) of the sample rate of that of the \((k - 1)\)th level. Unless the low-pass filter is perfectly ideal, spurious low spatial frequency components are introduced due to the effect of aliasing. In this section, the hierarchical structure is modeled mathematically and analyses are made to select the proper filtering characteristics to minimize aliasing errors. These analyses are made on a two-dimensional structured image as compared to the one-dimensional analyses made in [5].

Fig. 1 shows the mathematical model. An image at the \((k - 1)\)th level \( F_{k-1}(x, y) \) with Fourier spectrum \( F_{k-1}(\omega_x, \omega_y) \) is linearly filtered by a low-pass filter \( H(x, y) \) with the transfer function \( H(\omega_x, \omega_y) \). The filtered image is then spatially sampled at twice the spatial distances of the image sampling at the \((k - 1)\)th level. The resulting data \( F_k(x, y) \)

represent the picture function at the search level \( k \).

\[
F_k(x, y) = F_{k-1}(x, y) \odot H(x, y) \tag{1}
\]

where \( \odot \) is the convolution operation. The sampling signal is represented by

\[
S(x, y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(x - j_1 2\Delta x, y - j_2 2\Delta y) \tag{2}
\]

where \( \Delta x \) and \( \Delta y \) are the sampling intervals at the \((k - 1)\)th level. The Fourier transform is given by

\[
S(\omega_x, \omega_y) = \frac{\pi^2}{\Delta x \Delta y} \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(\omega_x - \frac{2\pi j_1}{\Delta x}, \omega_y - \frac{2\pi j_2}{\Delta y}) \tag{3}
\]

The Fourier spectrum of the sampled image field is

\[
F_k(\omega_x, \omega_y) = F_{k-1}(\omega_x, \omega_y) \odot H(\omega_x, \omega_y)
\]

\[
= \frac{\pi^2}{\Delta x \Delta y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_{k-1}(\omega_x, \omega_y) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(\omega_x - \omega - \frac{\pi j_1}{\Delta x}, \omega_y - \omega - \frac{\pi j_2}{\Delta y}) \, d\omega_x \, d\omega_y \tag{4}
\]

\[
= \frac{\pi^2}{\Delta x \Delta y} \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} F_{k-1} \left( \omega_x - \frac{\pi j_1}{\Delta x}, \omega_y - \frac{\pi j_2}{\Delta y} \right) \tag{5}
\]

The transfer function of the low-pass filter can be expressed as

\[
H(\omega_x, \omega_y) = \left[ \cos \left( \frac{\omega_x}{2} \right) \cos \left( \frac{\omega_y}{2} \right) \right]^n \tag{6}
\]

The numerical value of \( n \) is determined by the high frequency attenuation rate of the filter transfer function. Let the spectral energy passed through the low-pass filter be \( E_R \)

\[
E_R = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_{k-1}(\omega_x, \omega_y) H(\omega_x, \omega_y) \, d\omega_x \, d\omega_y \tag{7}
\]

For an ideal filter with infinitely sharp cutoff frequencies of \( \omega_x \) and \( \omega_y \), the filter output becomes

\[
E_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_{k-1}(\omega_x, \omega_y) H(\omega_x, \omega_y) \, d\omega_x \, d\omega_y \tag{8}
\]
An aliasing error can now be defined as
\[ E_A = \frac{E_K - E_0}{E_R}. \] (9)

Fig. 2 shows a plot of the aliasing error \( E_A \) for a low-pass filter having various values of \( n \). In this calculation, \( F_{k-1}(\omega_x, \omega_y) \) was assumed to have an amplitude of one at the frequencies of the passband.

A digital version of (5) can be written by letting
\[
\begin{align*}
F_{k-1}(N, N) & \quad \text{input image array} \\
F_k(M, M) & \quad \text{output image array} \\
H(l, l) & \quad \text{low-pass filter}.
\end{align*}
\]

Then
\[
F_k(m_1, m_2) = \sum_{n_1=m_1}^{m_1+l-1} \sum_{n_2=m_2}^{m_2+l-1} F_{k-1}(n_1, n_2) \cdot H(n_1 - m_1 + 1, n_2 - m_2 + 1). \] (10)

For a filter with \( n \approx 8 \), \( H(l, l) \) becomes
\[
H(l, l) = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.
\]

Equation (10) becomes
\[
F_k(m_1, m_2) = \sum_{n_1=m_1}^{m_1+l-1} \sum_{n_2=m_2}^{m_2+l-1} F_{k-1}(n_1, n_2) \cdot H(n_1 - m_1 + 1, n_2 - m_2 + 1). \] (11)

IV. SEQUENTIAL DECISION RULES

At each resolution level, the most promising test locations are selected for further testing at the next higher resolution level. The selection of the most promising test locations can be considered as a clustering problem. This problem is equivalent to maximizing the distance measures between the set of most promising test locations and the set of all other test locations. At the final level of testing, the problem is to maximize the distance measures between the true match location and all other test locations.

Let the window and search areas be scanned row-by-row such that
\[
S = \text{search area of size } M \times M \text{ in the search region of size } N \times N,
\]
\[
i.e., \text{search subregion superimposed by } W = (s_1, s_2, \ldots, s_{M \times M})^T
\]
\[
W = \text{window of size } M \times M
\]
\[
= (w_1, w_2, \ldots, w_{M \times M}). \] (12)

Then the test location map \( Q_k \) becomes
\[
Q_k \quad \text{test location map of size } (N - M + 1)^2 \text{ at search level } k
\]
\[
Q_k(u, v) \quad \text{test location, the top left corner of the window is located at } (u, v) \text{ of the search region.}
\]

Let \( e_k^i(s_i, w_i) \) be the error measure of the \( i \)th window pair of test location \( (u, v) \) at the resolution level \( k \). The accumulated error after \( n \) window pair computations becomes
\[
E_k(u, v) = \sum_{i=1}^{n} e_k^i(s_i, w_i). \] (15)

The sequential decision rules for hierarchical search can now be formulated. Let \( N_k \) be a set of test locations \( (u, v) \) at search level \( k \) such that
\[
N_k = \{(u, v) \mid E_k(u, v) < T_k^*, \quad 1 \leq n \leq M^2\} \] (14)

where \( T_k^* \) is the threshold computed after \( n \) window pair computations at search level \( k \).

Now generate a \((2N - 2M + 1)^2\) matrix \( G_{k-1} \), such that
\[
G_{k-1}(2i, 2j) = \begin{cases} 1, & \text{if } (i, j) \in N_k \\ 0, & \text{if } (i, j) \notin N_k \end{cases}. \] (15)

All other entries to \( G_{k-1} \) are set to zero. Matrix \( G_{k-1} \) is used as the test location map at the search level \( k - 1 \), and is used as a guide in the search for the match location at level \( k - 1 \). Tests are to be performed only at the locations \( (u, v) \) for \( G_{k-1}(u, v) = 1 \).

The search continues until one of two cases is encountered.

1) At the search level \( k \), \( G_k(u, v) = 1 \) for one value of \( (u, v) \), location \( (u, v) \) is declared the match location. If \( k \) is not the highest resolution level, a local search can be made to locate a point of minimal accumulated error. This is done by searching the four adjacent locations of the declared location at the next search level and selecting a point with minimal accumulated error as the match location. This process is repeated until the highest resolution level is reached.

2) At the search level \( k = 0 \), there exist several locations \( (u, v) \) such that \( G_0(u, v) = 1 \). Select the location with the smallest accumulated error as the most likely match location.

V. DERIVATION OF THRESHOLD SEQUENCE

Let the error measure be based on the \( L_1 \) norm, or
\[
e_k^i = |s_i - w_i|. \] (16)
At the location of true match \((u^*, v^*)\), the sum of \(e^*_k\) is a cumulative measure of the difference between the two images. In order to make a correct decision, one assumes that the total accumulated error is minimal at \((u^*, v^*)\). Therefore, the knowledge of the shape of the error growth curve and the density distribution of \(e^*_k\) at \((u^*, v^*)\) is important in determining the threshold sequence. Theoretical analyses and experimental results [2] indicated that the error density function can be approximated by an exponential function. Based on an exponential function, it can be shown that the threshold sequence can be determined by the following equation:

\[
\frac{T^*_n}{r_m} = (\sqrt{2})^{n-1}(n + g_k\sqrt{n})
\]

where \(T^*_n\) is the threshold to be used after the \(n\)th window pair computation in the \(k\)th search level. \(r_m\) is the averaged error measurement at \((u^*, v^*)\). \(g_k\) is a function of the probability of a match \(p_k\) at search level \(k\). If the number of error measurements is large, which is usually true in the case of scene matching, the distribution of the accumulated error \(E^*_k\) can be approximated by a Gaussian distribution. \(p_k\) can be expressed by the following equation:

\[
p_k = \frac{1}{\sqrt{2\pi} e^{\left(-\frac{1}{2}\right)}} \int_0^{\infty} \exp \left[-\frac{(E^*_k - \gamma_m)^2}{2\gamma_m^2}\right] d(E^*_k)
\]

\[
= 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-y(n + g_k\sqrt{n})} \exp \left[-\frac{(E^*_k)^2}{2}\right] d(E^*_k)
\]

where

\[
y = (n + g_k\sqrt{n})\gamma_m.
\]

\(p_k\) as a function of \(g_k\) was computed using (18) and (19) and the result is shown in the following table.

<table>
<thead>
<tr>
<th>(g_k)</th>
<th>(p_k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.900</td>
</tr>
<tr>
<td>1</td>
<td>0.841</td>
</tr>
<tr>
<td>2</td>
<td>0.977</td>
</tr>
<tr>
<td>3</td>
<td>0.998</td>
</tr>
<tr>
<td>4</td>
<td>0.999</td>
</tr>
<tr>
<td>5</td>
<td>0.999</td>
</tr>
</tbody>
</table>

It is seen that probability of a match increases with \(g_k\). However, the computational efficiency decreases as \(g_k\) increases. This is due to the fact that threshold \(T^*_n\) increases with \(g_k\) and more test locations require evaluation at the higher resolution levels.

The overall probability of a match \(P_m\) is the product of the individual \(p_k\) computed at each of the search levels:

\[
P_m = \prod_{k=m}^{n} p_k
\]

where \(m\) is the lowest resolution level and \(n \geq 0\) is the level at which a match decision is declared.

VI. EFFICIENCY OF HIERARCHICAL SCENE MATCHING

Computational efficiency of scene matching using the hierarchical search technique can be evaluated based on theoretical considerations. To form a basis of comparison, the number of window pairs needed to be computed in the standard correlation method is calculated. For a search region of \(N \times N\) and a window size of \(M \times M\), this number becomes

\[
C_1 = (N - M + 1)^2(M)^2.
\]

When hierarchical search is applied, the number of window pairs needed to be computed is

\[
C_2 = K_1(M)^2 \log (N - M + 1)^2
\]

where \(K_1\) is a constant, the value of which depends on the scene content. In the process of creating the two sets of structured images, the number of picture elements which computations can be calculated by the following equation:

\[
C_3 = 2 \sum_{l=1}^{L-1} (N/2)^2
\]

where \(L\) is the total number of search levels. The amount of computational saving \(\eta\) is

\[
\eta = \frac{C_1}{C_2 + C_3}
\]

\[
= \frac{(N - M + 1)^2}{K_1 \log (N - M + 1)^2 + 2 \sum_{l=1}^{L-1} (N/2)^2/(M)^2}
\]

At each level in hierarchical search, sequential testing and capture methods [2] are applied to further minimize the amount of computations. Let the savings due to the sequential testing and capture be \(K_2\); (24) then becomes

\[
\eta = \frac{(N - M + 1)^2}{K_2 K_1 \log (N - M + 1)^2 + 2 \sum_{l=1}^{L-1} (N/2)^2/(M)^2}
\]

The numerical value of \(K_2\) can be evaluated by obtaining an expression for the average number of observations \(N\) required to terminate the sequential test as a function of the probability of a miss \(\beta\) and the probability of false alarm \(\alpha\) [9]. The relative savings in \(N\) as compared to the number of observations required by the nonsequential, optimal Bayesian test to achieve the same error probabilities \(\alpha\) and \(\beta\) can be used to estimate the numerical value of \(K_2\). The results [9] indicate that the savings of 30 to 90 percent are possible for a wide range of values of \(\alpha\) and \(\beta\). Let an average savings be 50 percent. \(K_2\) of (25) takes on the numerical value of

\[K_2 = 0.5.\]

Substituting this value into (25) yields
\[
\eta = \frac{2(N - M + 1)^2}{K_1 \log (N - M + 1)^2 + 4 \sum_{i=1}^{L-1} (N/2^i)^2/M^2}.
\] (27)

For a search region of 256 x 256 or N = 256, a window size of 64 x 64 and L = 4, the savings \( \eta \) as a function of \( K_1 \) was computed by (27) and the results are as shown in Table I. Therefore, the amount of computation for scene matching with the conditions as stated is expected to be reduced by a factor of about 1000 to 3000 using the hierarchical search technique.

VII. EXPERIMENTAL RESULTS

A sequence of experiments were performed to verify the performance of the hierarchical scene matching techniques described in this paper. The scenes used in the experiments are the 256 x 256 geometrically corrected optical image and intensity corrected radar image as shown on top of Figs. 3 and 4. Two sequences of structured images of sizes 128 x 128, 64 x 64, and 32 x 32 were created through the repeated applications of (11). These images are as shown in Figs. 3 and 4.

For the first sequence of experiments, Gaussian noise with zero mean and \( \sigma = 86 \) was introduced in the 256 x 256 optical image. The optical image which has been quantized into 256 levels has an averaged intensity level of 43. A sequence of structured images was then created from the noise-corrupted optical image as shown in Fig. 5. Three regions of interest were selected for the match. Three 8 x 8 subimages were extracted from the 32 x 32 noise corrupted optical image. These subimages are

1) upper part of the large stadium at the upper left corner
2) lower part of the large stadium at the middle left
3) small stadium at the lower right corner.

Parameters used in these experiments are summarized in Table II. The test locations which satisfied the decision rule in (14) are recorded for further tests at the next higher resolution level. The testing continues until the matching locations are found.

For the second sequence of experiments, an 8 x 8 subimage of the small stadium was extracted from the 32 x 32 radar image. A match was performed on the 32 x 32 optical image. The parameters used in these experiments are shown in Table II.

A. Scene Matching with Constant Threshold

For the experiments, a constant threshold was used in each search level \( k \). The numerical value of the threshold at each search level \( k \) was computed by setting \( n \) of (17) equal to the window size \( M \times M \). Table III shows the numerical value of \( T_{m} \) as a function of the search level \( k \) and the deviation \( \sigma_k \).

The results of these experiments are as summarized in Tables IV and V. The match location was correctly detected at each of the four matches. In the matching of the noisy optical regions to the optical image, only three levels of search were required for the lower part of the large stadium and the small stadium. For the matching of the upper part of the stadium, a search level at the highest resolution was needed to obtain a match. This is also true for the matching of the small stadium of the radar image to that of the optical image. Computational efficiency of the hierarchical scene matching technique can now be evaluated based on the experimental results. To form a basis of comparison, the

---

**TABLE I**

<table>
<thead>
<tr>
<th>( K_1 )</th>
<th>( \eta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2913</td>
</tr>
<tr>
<td>5</td>
<td>1699</td>
</tr>
<tr>
<td>10</td>
<td>1117</td>
</tr>
</tbody>
</table>

---

Fig. 3. Structured optical images. (a) 256 x 256. (b) 128 x 128. (c) 64 x 64. (d) 32 x 32.
number of window pairs needed to be evaluated in the standard correlation method was calculated by (21). For a search region of \(256 \times 256\) or \(N = 256\) and a window size of \(64 \times 64\) or \(M = 64\), (21) becomes

\[
C_1 = (N - M + 1)^2(M)^2 = 1.526 \times 10^4. \quad (28)
\]

In the process of creating the two sets of structured images, the number of picture elements \(C_2\) which require computations can be evaluated by (23) as follows:

\[
C_2 = 2 \sum_{i=1}^{L-1} (N/2)^i
\]

\[
= 43008 \quad (29)
\]

where \(L = 4\) and \(N = 256\). The number of window pairs evaluated, \(C_3\), in a match can be obtained from the experimental results in Tables IV and V. This is done summing the window pairs evaluated at each search level.

The computational efficiency \(\eta\), computed by (30), is shown in Table VI.

\[
\eta = \frac{C_1}{C_2 + C_3}. \quad (30)
\]
A good correlation is achieved between the theoretical and experimental results since the efficiency of each of the matches falls within the theoretical limits as shown in Table I.

### B. Scene Matching with Monotonic Increasing Threshold

In this sequence of experiments, the threshold \( T_k \) used in the scene matching is computed as a function of the search level \( k \) as well as \( n \), the number of window pairs of error accumulated. In order to establish a trend for the test points far from the true matching point and to reduce the probability of a miss, an initial threshold \( T_0 \) is established. The threshold equation (17) is modified as follows:

\[
T_k = \max \left[ T_0, \left( \frac{2}{n} \right)^m \cdot \left( n + g_k \cdot \bar{n} \right) \right].
\]  

(31)

Performance of scene matching with monotone increasing threshold is shown in Tables VII and VIII together with values of \( T_k \) used in the experiments. Comparisons of data in Tables VII and VIII and those in Tables IV and V indicated the following advantages of using monotone increasing threshold.

1. The numbers of the search level required for a match are reduced.
2. There is a reduction in the window pairs of computation for a match. The percent reduction is greater at the high resolution than at the lower resolution.
3. The average reduction is 33 percent in the number of window pairs evaluated over those using constant threshold.
TABLE VII
OPTICAL-TO-NOISY OPTICAL SCENE MATCHING, MONOTONE INCREASING THRESHOLD

<table>
<thead>
<tr>
<th>Location</th>
<th>Search Level $k$</th>
<th>Number of Successful Test Locations</th>
<th>Number of Window Pairs Evaluated</th>
<th>Initial Threshold $T_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper part of large stadium</td>
<td>3</td>
<td>1</td>
<td>20,868</td>
<td>80</td>
</tr>
<tr>
<td>Lower part of large stadium</td>
<td>3</td>
<td>3</td>
<td>21,322</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>703</td>
<td>700</td>
</tr>
<tr>
<td>Small stadium</td>
<td>3</td>
<td>4</td>
<td>22,753</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>512</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1,885</td>
<td>2,000</td>
</tr>
</tbody>
</table>

TABLE VIII
RADAR-TO-OPTICAL SCENE MATCHING, MONOTONE INCREASING THRESHOLD

<table>
<thead>
<tr>
<th>Location</th>
<th>Search Level $k$</th>
<th>Number of Successful Test Locations</th>
<th>Number of Window Pairs Evaluated</th>
<th>Initial Threshold $T_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small stadium</td>
<td>3</td>
<td>22</td>
<td>26,746</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
<td>5266</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>4,816</td>
<td>2,000</td>
</tr>
</tbody>
</table>

VII. SUMMARY

The results of the theoretical analysis and experiments indicate that a large computational saving can be realized by the use of sequential scene matching with hierarchical search. For the matching of images described in this paper, the amount of computation in terms of the number of window pairs needed to be evaluated is reduced on the average by a factor of about 1900.

The true significance of this approach, however, is that scene matching can be accomplished even in cases which are difficult for humans or standard correlation techniques and can be accomplished with greatly reduced computation. This was demonstrated in the matching of the noisy small stadium.

REFERENCES


Robert Y. Wong (M’64) received the B.S.E.E. degree from the University of California at Los Angeles in 1964 and the M.S.E.E. and Ph.D. degrees from the University of Southern California, Los Angeles, in 1965 and 1976, respectively.

He was the principal designer of the successful Mariner 4 Planetary Scan System and one of the principal designers of the Mariners 6 and 7 on-board analog and digital television signal processing systems. He was a part-time faculty member in the School of Engineering, California State University at Los Angeles, where he has taught more than 20 courses in digital computer design, control systems, and electronics. He is now serving as a faculty member in the Department of Electrical Engineering, California State Polytechnic University at Pomona. His current interests include computer applications to sequential pattern recognition and image processing. He is also currently serving as a consultant to the University of Southern California and University of Tennessee.

Dr. Wong is a member of Sigma Xi.

Ernest L. Hall (S'63–M'65) received the B.S.E.E., M.S.E.E., and Ph.D. degrees in electrical engineering from the University of Missouri–Columbia in 1965, 1966, and 1971, respectively.

During 1971-1972 he served on the faculty of the University of Missouri–Columbia. In 1972 he joined the faculty of Yale University, New Haven, CT, and in 1973 moved to the University of Southern California, Los Angeles. In 1976 he joined the University of Tennessee, Knoxville, where he is currently an Associate Professor of Electrical Engineering. His research interests include computer image processing and pattern recognition, visual perception, computer graphics, and biomedical imaging. He was one of the principal developers of the DR ADAM system for automated chest X-ray analysis which is now in operation at the Appalachian Center for Respiratory Disease. He has served as a consultant to several professional organizations, including the Xerox Corporation, Jet Propulsion Laboratory, Loma Linda Hospital, and the National Institute for Occupational Safety and Health.

Dr. Hall is a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and several professional societies.