Fast Correlation Registration Method Using Singular Value Decomposition

Mingchuan Zhang and Kai-Bor Yu
Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

Robert M. Haralick
Machine Vision International, Ann Arbor, MI 48104

A new, fast template-matching method using the Singular Value Decomposition (SVD) is presented. This approach involves a two-stage algorithm, which can be used to increase the speed of the matching process. In the first stage, the reference image is orthogonally separated by the SVD and then low-cost pseudo-correlation values are calculated. This reduces the number of computations to $2*N*L$ instead of N^2L^2, where $L \times L$ is the size of the reference image and $N \times N$ is the original image size. At the second stage, a small group of values near the maximum pseudo-correlation is selected. The true correlation for the small number of pixels in this group is then computed precisely in the second stage. Experimental and analytic results are presented to show how the computation complexity is greatly improved.

I. INTRODUCTION

The problem of registering one image to another image or model, possibly made at a different time or from a different perspective, is one of the major research problems in remote sensing, image processing and robotic vision.[1-3]

The registration problem can be considered as a transformation T that maps an arbitrary point (x,y) in the first (image) coordinate space to a corresponding point (u,v) in the second (model) coordinate space such that the corresponding points represent the same points of the viewed objects.

In the brute force technique, the parameters of the transformation are varied until some measure of "difference" between the images is minimized or some measure of "similarity" (e.g., cross-correlation) is maximized. The value of this measure is then used to determine the optimum position of registration.

If an appropriate a priori correction has been applied or if the patches are defined small enough, then any residual geometric error is pure translation. Matching, then, consists of determination of the translations offset of one subimage from another "reference" subimage corresponding to the same scene. The most widely used method of measuring the correspondence of two images is cross-correlation, which follows naturally from the mean square error criteria.

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. I, 181-194 (1986)
 CCC 0884-8173/86/030181-14$04.00

But the high costs of its implementation severely limits its utility. In the correlation method, each point in the correlation surface requires a fixed amount of computation. Each reference point, regardless of informational content, is therefore processed with very high precision. However, accuracy is required only for those relatively few points with values close to the maxima of the cross-correlation surface. Hence, there is considerable waste in performing high accuracy calculation for the vast majority of points within an image.

This article presents a new fast correlation registration method which involves the SVD of the reference image. This greatly reduces the computational redundancy in performing a two-stage template matching algorithm. In the first stage we only use the largest eigenimage which is obtained from the SVD of the reference image to evaluate cross-correlation. We call it the pseudo-correlation. Using it, the computation complexity reduced from $N^2 * L^2$ to $2NL$. After this stage a small group of values near the maximum is selected. The small group of pixels near the pseudo-correlation peak are chosen by looking for positions which have enough pseudo-correlation. High enough means here than a threshold. Suitable thresholds values are discussed in Section III. The true correlation for only the pixels in this group needed to be calculated precisely in the second stage. In Section III, the cross-correlation problem using the SVD derived eigenimage is formulated. In Section III, the correlation threshold is derived and statistical properties are presented. In Section IV, the computational requirement is discussed. In Section V, the implementation details and experimental results are included. Some conclusions are discussed in Section VI.

II. FAST-CORRELATION METHOD

A. Cross-Correlation

For each translation (i,j), the conventional discrete two-dimensional cross-correlation coefficient is defined by:

$$R(i,j) = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} W(l,m) * S(i+1, j+m)}{\left(\sum_{l=1}^{L} \sum_{m=1}^{L} W^2(l,m) \sum_{l=1}^{L} \sum_{m=1}^{L} S^2(l,m) \right)^{1/2}} \tag{2.1}$$

where $R(i,j)$ is the normalized cross-correlation surface with $(M \times M)$ pixels, $S(i,j)$ is a sensor image with $N \times N$ pixels, and $W(i,j)$ is a reference image with a $L \times L$ correlation kernel. The dimensions of R are given by $M = N + L - 1$, i.e., $(N-L+1)^2$ cross-correlation have to be computed. Among all the computed values, the one with the largest value corresponds to the position of the best match. The number of multiplications required for conventional computation is $N^2 * L^2$

B. The SVD Method

The underlying idea of this two-stage fast registration method proposed in this article involves using the SVD technique to reduce the computation complexity in the first stage. The true cross-correlation is calculated with very high precision only for those relatively few points which are close to the maxima in the second stage.

In this section we first give the Singular Value Decomposition theorem as follows:

Let $A \varepsilon E^{m \times n}$ there exist orthogonal matrices $U \varepsilon E^{m \times m}$, $V \varepsilon E^{n \times n}$, and a $m \times m$ diagonal matrix D with diagonal elements $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n \geqslant O$, such that

$$U' A V = D \tag{2.12}$$

Note: The numbers $\mu_1, \ldots \mu_n$ are unique and are called the singular values of A.

From the above theorem, the singular value decomposition of A is given by

$$[A] = \sum_{i=1}^{R} \alpha_i u_i v'_i \tag{2.13}$$

where R is the rank of matrix $[A]$. The more traditional approach is given by the following definitions:

$$[A] = [U] [\Lambda]^{1/2} [V]' \tag{2.14}$$

and

$$[A] [A]' = [U] [\Lambda] [U]' \tag{2.15}$$

$$[A]' [A] = [V] [\Lambda]][V]' \tag{2.16}$$

where $[\Lambda]$ is the diagonal matrix of eigenvalues of $[A] [A]'$, and the columns of $[U]$ are the eigenvectors of $[A] [A]'$, and the columns of $[V]$ are the eigenvectors of $[A]' [A]$. Because $[A] [A]'$ and $[A]^t [A]$ are symmetric and square, the μ_i are real and the eigenvector sets $\{u_i\}$ and $\{v_i\}$ are self-orthogonal. Figure 1 is a graphical illustration of the SVD. From Eq. (2.13) and Figure 1 it is evident that for smaller R, fewer independent rows (columns) are required to define the matrix $[A]$. Ordering the singular values in monotonic decreasing order yields the most efficient least-square representation of the image in the fewest (truncated) set of retained components: $\{\mu^{1/2} u_i v_i^t\}$. The retention of only the $t \leqslant R$ largest eigenvalues in the expansion gives a normalized error energy of

$$E_t = 1 - \sum_{i=1}^{t} \alpha_i \Big/ \sum_{j=1}^{R} \alpha_j. \tag{2.17}$$

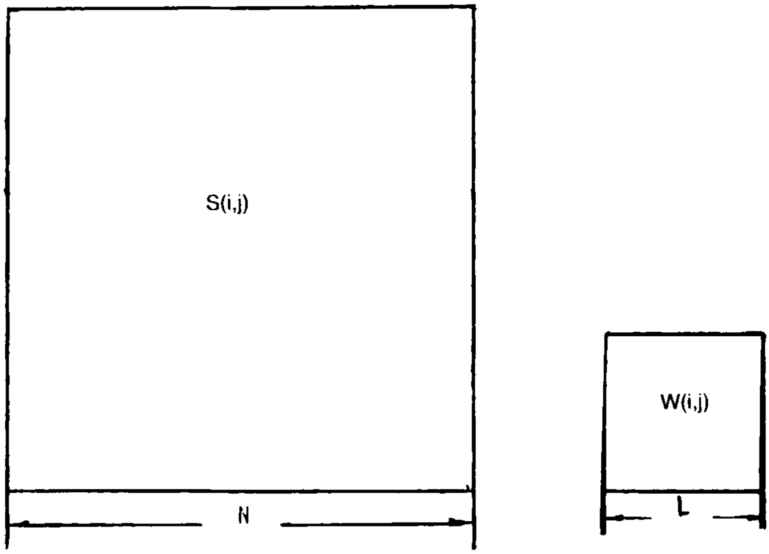

Figure 1. Translational Registration Parameters. $S(i,j)$—sensor image with $N \times N$ pixels. $W(i,j)$—reference image with $L \times L$ correlation kernel.

We show next that if an approximate representation of the matrix A is formed by truncation, then

$$A_k = \sum_{i=1}^{k} \mu_i^{1/2} u_i v_i' \tag{2.18}$$

and the squared norm between $[A]$ and $[A_k]$ becomes

$$\|[A] - [A_k]\|^2 = \sum_{i=k+1}^{R} \mu_i \tag{2.19}$$

where the matrix norm is the Euclidean measure.

$$\mathrm{Tr}[A]'[A] = \|A\|^2.$$

The motivation for utilizing the SVD expansion is that, hopefully, $[A]$ admits a good low-ranked (small k) approximation; in that case, the storage requirements drop from N^2 to $K(2N+1)$ computer words (i.e., $2N$ words for two singular vectors and one word for the singular value). The approximation error is minimized by choosing the k largest singular values and the corresponding singular vectors.

The SVD matrix decomposition applies for any arbitrary matrix. Hence the SVD expansion can be applied directly to a discrete image represented as a matrix. Now we go back to the cross-correlation problem. Suppose the reference image H (i.e., matrix H) has been orthogonally decomposed by the SVD as below:

$$H = \sum_{i=1}^{R} \alpha_i u_i v_i^T$$

where R is the rank of H,α_i is the ith singular values and u_i, v_i are the corresponding singular vectors. We will show next that if H is of full rank, the number of operations required to perform the correlation is of the order of $(2N^2L)*L$. In many practical cases, the rank value R of the matrix H is much less than L, and the number of operations can be reduced accordingly. For example, if the rank of the reference image is 1, the number of operations required to perform the cross correlation is $2NL$.[4]

As discussed earlier in this section, the cross-correlation matching of a $M \times M$ sensor image with a $L \times L$ reference image requires evaluation at every one of the possible $(N-L+1)^2$ shift positions. Among all these computed positions, the one with the largest value gives the cross-correlation coefficient for the best match. The number of multiplications required for conventional computation is N^2*L^2.

If we can decompose a 2-D matrix into two separated 1-D vectors (i.e., we decompose the $L \times L$ reference image into two vectors with L components), the cross-correlation can also be decomposed into two vector correlation operations as:

$$R(i,j) = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} S(i+1,j+m)*W(l,m)}{M}$$

$$= \frac{\sum_{l=1}^{L} \left(\sum_{m=1}^{L} S(i+1,j+m) * A(l) \right) * B(m))}{M}$$

$$M = [(\Sigma\, \Sigma\, W^2(l,m))(\Sigma\, \Sigma\, s^2(l,m))]^{1/2}$$

where S is the sensor image, and W is the reference image

$$W = A_{L\times 1} * B_{1\times L'}$$

and A, B are vectors. From the above expression we calculate the one dimension correlation with vector A row by row, and then calculate the one dimension correlation with vector B column by column. Furthermore, if we use pipeline techniques,[7] the number of multiplications required for the row vector is N^2*L, and the number of multiplications required for the column vector is N^2*L. Thus, for a $L \times L$ reference image, the decomposition method used $2*L$ operations per pixel instead of L^2 operations per pixel.

In general, if H is of full rank, the number of operations required to perform the correlation is of the order $(2L)*L*N^2$. For most real world images, the eigenimages having largest singular value dominate the remaining eigenimages. Pseudo-correlation uses the largest eigenimage of the reference image to evaluate the cross-correlation. The computation complexity reduces from N^2*L^2 to $2N^2L$.

The pseudo-cross-correlation value at the optimal match position is not necessarily a maxima. Usually it falls into a small region around it which we will test further for the optimal matching. The pseudo-cross-correlation value at the optimal match position usually has high value closed to the maximum pseudo-cross-correlation value. Table I illustrates this fact for some experimental results. From the table we see that usually only 1–3% of the pixels have pseudo-correlation values higher than the pseudo-correlation value at the optimal match position. Thus only a very small number of pixels needs to be calculated precisely in the second stage.

III. THE CORRELATION VALUE THRESHOLD DETERMINATION

A. Correlation Threshold for Noise-free Image

The greatest difficulty in the two-stage algorithm is in the determination of an accurate correlation threshold value during the first stage. If the threshold value is too low, the size of the candidate group is too large. The optimal match position might be missed, however, if the threshold value is too high. In this section, we derive the correlation threshold value of the first stage, using the SVD definition and considering the noise-free and noisy images, respectively. As we showed in the previous section, the reference image, which is taken as an $L \times L$ array of digital picture elements, can be expressed as a sum of separable eigenimages by a singular value decomposition (SVD):

$$W = \Sigma\ \alpha_i a_i b_i^T = \Sigma\ \alpha_i E_i \tag{3.1}$$

Table I. Experimental results of pseudo-correlation.

Sensor image	Reference image	TH[a] value	Minimum group[b] size (%)
250 × 250	15 × 15	0.2918	5.6
250 × 250	11 × 11	0.1718	2.0
250 × 250	11 × 11	0.1485	1.5

[a]TH indicates the ratio between largest singular value and summation of all singular values.

[b]The minimum candidate size indicates that the number of pixels have pseudo-correlation values higher than the pseudo-correlation value of the optimal match position.

where α_i, E_i are the ith singular value and the eigenimage, respectively.

Let

$$R\,(i,j) = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} W(l,m) * S(i+1,j+m)}{\left(\sum_{l=1}^{L} \sum_{m=1}^{L} W^2(l,m) \sum_{l=1}^{L} \sum_{m=1}^{L} S^2(l,m)\right)^{1/2}} \tag{3.2}$$

and

$$R'(i,j) = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1 E_1(l,m) * S(i+1,j+m)}{\left(\sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1^2 E_1(l,m) \sum_{l=1}^{L} \sum_{m=1}^{L} S^2\,(l,m)\right)^{1/2}} \tag{3.3}$$

where $R(i,j)$ is the cross-correlation, and $R'(i,j)$ is the pseudo-correlation value, which is the cross-correlation between the original image and the largest referenced eigenimage

Suppose at the (i^*,j^*) position, $R(i^*,j^*) = \max R(i,j)$, which will be represented as R_M, and at the position (i',j'), $R'(i',j') = \max R'(i,j)$ which will be represented as R_M' (where $1 \leqslant i,j \leqslant M-L+1$, $1 \leqslant l,m \leqslant M$).

We have

$$R'_M = R(i*,j*) = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} W(l,m) * S(i*+1,j*+m)}{\left(\sum_{l=1}^{L} \sum_{m=1}^{L} W^2\,(l,m) * \sum_{l=1}^{L} \sum_{m=1}^{L} S^2(l,m)\right)^{1/2}} \tag{3.4}$$

and

$$R'_M = R(i',j') = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1 E_1(l,m) * S(i'+1,j'+m)}{\left(\sum_{l=1}^{L} \sum_{m=1}^{L} S^2(l,m) * \sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1^2 E_1^2(l,m)\right)^{1/2}} \tag{3.5}$$

and

$$R'(i*,j*) = \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1 E_1(l,m) * S(i*+1,j*+m)}{\left(\sum_{m=1}^{L} \sum_{l=1}^{L} \alpha_1 E_1^2(l,m) * \sum_{m=1}^{L} \sum_{l=1}^{L} S^2\,(l,m)\right)^{1/2}}$$

$$= \frac{\sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1 E_1(l,m) * W(l,m)}{\left(\sum_{l=1}^{L} \sum_{m=1}^{L} \alpha_1 E_1^2(l,m) * \sum_{l=1}^{L} \sum_{m=1}^{L} S^2(l,m) \right)^{1/2}} . \tag{3.6}$$

For a noise-free image the correlation threshold can be expressed as below:

$$\mathrm{TH} = \frac{R'(i*,j*)}{R_M(i',j')}$$

$$= \frac{\text{Pseudo-correlation at optimal matching position } (i*,j*)}{\text{Pseudo-correlation at pseudo-optimal matching position } (i',j')} .$$

For noise-free image the intensities of the reference image and sensor image should be the same at the optimal matching position, that is $W(l,m) = s(l,m)$ for $(l,m=1, \ldots ,L)$, so the above is given by

$$= \frac{\sum_{i=1}^{L} \sum_{m=1}^{L} \alpha_1 E_1(l,m) * W(l,m)}{\left(\sum_{i=1}^{L} \sum_{m=1}^{L} W^2(l,m) * \sum_{i=1}^{L} \sum_{m=1}^{L} \alpha_1^2 E_1^2(l,m) \right)^{1/2} * R'_m} . \tag{3.7}$$

B. Statistical Properties of the Cross-Correlation

Before deriving the threshold for noisy images, we shall first discuss the statistical properties of the cross-correlation.[5,6] Without loss of generality, let the cross-correlation R be expressed as below:

$$R = \frac{\sum_{t=1}^{n} (X_t - E(X))\,(Y_t - E(Y))}{\left[\sum_{t=1}^{n} (X_t - E(X))^2 * \sum_{t=1}^{n} (Y_t - E(Y))^2 \right]^{1/2}} . \tag{3.8}$$

(i) $(X_1,Y_1) \ldots (X_t,Y_t) \ldots (X_n,Y_n)$ are samples from a bivariate normal distribution. $(t=1, \ldots ,n)$

(ii) ρ is the estimated population correlation coefficient,

$$\rho = \frac{E[\{X_t - E[X_t]\}\{Y_t - E[Y_t]\}]}{(\mathrm{VAR}(X_t)\mathrm{VAR}(Y_t))^{1/2}} . \tag{3.9}$$

The probability distribution of R can be expressed as below:[5]

$$P_R(r) = \frac{(1-\rho^2)^{(n-1)/2}(1-r^2)^{(n-4)/2}}{\sqrt{\Pi}\Gamma((n-1)/2)\Gamma(1/2n-1)} \sum_{j=1}^{n} A_j \tag{3.10}$$

$$A_j = \Gamma((n-1+j)/2]^2 * (2\rho r)^j 1/j!(-1 \leq r \leq 1) \tag{3.11}$$

when n is number of samples

It is clear that the distribution of R is so complicated that it is very difficult to use without a practical approximation or an extensive set of tables. Johnson[2] gave a practical approximation to the distribution of R, which can be obtained by the transformation:

$$Z' = \tanh^{-1} R = 1/2 \log((1+R)/(1-R)). \tag{3.12}$$

This transformation might be seen as a variance-equalizing transformation, and have properties as below:

$$\text{var}(R) = (1 - \rho^2)^2 n^{-1}$$

and

$$\int (1-\rho^2)^{-1} \, d\rho = 1/2 \log (1+\rho/1-\rho). \tag{3.13}$$

Johnson[5] pointed out that we can approximately regard Z' as normally distributed with an expected value of $1/2\log(1+\rho/1-\rho)$ and variance $(n-3)^{-1}$. Finally he proved the following approximation results:

$$P_R[R \leq r] = \Phi\,(B) \tag{3.14}$$

where

$$B = \frac{r(1 - r^2)^{1/2}(n-3/2)^{1/2} - \rho(1 - \rho^2)^{1/2}(n - 5/2)^{1/2}}{[1+r^2(1-r^2)^{-1}/2 + \rho^2(1-\rho^2)^{-1}]^{1/2}} \tag{3.15}$$

and

$$\phi(B) = \int_{-\infty}^{B} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

From Eq. (3.15) we see that the distribution of R is a function of ρ, r, and n. The first and second moments of the distribution of R can be approximated as below:

$$\mu_1 = \rho - 1/2n\ \rho(1-\rho^2) \tag{3.16a}$$

$$\mu_2 = (1-\rho^2)^2/n. \tag{3.16b}$$

C. Correlation Threshold for Noisy Image

Suppose the sensor image is a digital image with Gaussian noise, which can be expressed as below:

$$S(i,j) = I(i,j) + N(i,j)$$

and we assume $N(i,j)$ is smaller than $I(i,j)$. As before, we have the expressions:

$$R_M = R(i*,j*) = \frac{\sum_{l=1}^{L}\sum_{m=1}^{L} W(l,m)*I(i*+l,j*+m)}{\left(\sum_{l=1}^{L}\sum_{m=1}^{L} W^2(l,m) * \sum_{l=1}^{L}\sum_{m=1}^{L} I^2(l,m)\right)^{1/2}} \tag{3.17}$$

and

$$R'_M = R(i',j') = \frac{\sum_{l=1}^{L}\sum_{m=1}^{L} \alpha_1 E_1(1,m)*S(i'+l,j'+m)}{\left(\sum_{l=1}^{L}\sum_{m=1}^{L} S^2(l,m) * \sum_{l=1}^{L}\sum_{m=1}^{L} \alpha_1^2 E_1^2\,(l,m)\right)^{1/2}}. \tag{3.18}$$

At the $(i*,j*)$ position, $R(i*,j*) = \max R(i,j)$, which is represented by R_M, and at the position (i',j'), $R'(i',j') = \max R'(i*,j*)$, which is represented by R'_M. We should note that the R_M corresponding to the optimal match position is the maximum cross-correlation value of true intensity value.

$$\begin{aligned} R'(i*,j*) &= \frac{\sum_{l=1}^{L}\sum_{m=1}^{L} \alpha_1 E_1(l,m)*S(i*+l,j*+m)}{\left(\sum_{l=1}^{L}\sum_{m=1}^{L} W^2(l,m)* \sum_{l=1}^{L}\sum_{m=1}^{L} S^2(l,m)\right)^2} \\ &= \frac{\sum_{l=1}^{L}\sum_{m=1}^{L} \alpha_1 E_1(l,m)*W(l,m)}{\left(\sum_{l=1}^{L}\sum_{m=1}^{L} W^2(l,m)* \sum_{l=1}^{L}\sum_{m=1}^{L} S^2(l,m)\right)^2}. \end{aligned} \tag{3.19}$$

For a noisy image the correlation threshold can be expressed as below:

$$\mathrm{TH} = \frac{E(R'(i*,i*))}{E(R_M(i',j'))}$$

$$= \frac{\rho_0 - 1/2n\ \rho_0(1-\rho^2{}_0)}{\rho_1 - 1/2n\ \rho_1(1-\rho^2{}_1)} \tag{3.20}$$
$$= \rho_0/\rho_1 \qquad \text{(when } n \text{ is large)}$$

where ρ_0, ρ are the estimations of the correlation coefficient (3.9) at pixels $(i*,j*)$ and (i',j'), respectively. Also we have

$$E(R'(i*,j*)) = \frac{E[(W+N'-E(W+N'))*(I+N-E(I+N)]}{[\mathrm{VAR}(W+N')*\mathrm{VAR}(I+N)]^{1/2}}$$
$$= \frac{E[(W-E(W))*(I-E(I))]+E[(W-E(W))*N]+E[N'*(I-E(I)]+E[N'N]}{[(\mathrm{VAR}(W)+\mathrm{VAR}(N'))*(\mathrm{VAR}(I)+\mathrm{VAR}(N))]^{1/2}}. \tag{3.21}$$

Because of the independence between the sensor image and the noise and also the white noise assumption, we have the following

$$= \frac{E[(W-E(W))*(I-E(I))]}{[(\mathrm{VAR}(W)+\mathrm{VAR}(N'))*(\mathrm{VAR}(I)+\mathrm{VAR}(N))]^{1/2}}. \tag{3.22}$$

Finally, the correlation threshold for a noisy image can be expressed as below:

$$\mathrm{TH} = \frac{E[(W-E(W))*(I-E(I)] * (1/R'_M)}{[(\mathrm{VAR}(W)+\mathrm{VAR}(N))*(\mathrm{VAR}(I)+\mathrm{VAR}(N))]^{1/2}}.$$

For the noisy image, we can simplify the expression for the threshold as follows:

$$\mathrm{TH} = \frac{\Sigma\, \Sigma\, \alpha_1 E_1(l,m)*W(l,m)}{[(\Sigma\Sigma W^2(l,m)+\mathrm{VAR}(N))*(\Sigma\Sigma\alpha_1{}^2 E_1{}^2(l,m)+\mathrm{VAR}(N))]^{1/2}*R'_M}. \tag{3.23}$$

IV. COMPUTATION REQUIREMENT

In the two-stage template matching process, the expected computational cost at each pixel of the picture (ignoring the cost of the SVD operations) is of the form $c+pd$, where

c: cost of applying the "first-stage," pseudo-correlation,
p: probability of the cross-correlation value above the threshold,
d: cost of applying the cross-correlation in the second stage.

From the previous analysis we know that c is equal to $2L$ multiplication operations, and d is equal to L^2 multiplication operations.

The probability p can be evaluated from Eqs. (3.14) and (3.15), if we know the reference image size, the correlation threshold TH, and the population correlation coefficient ρ. On the other hand, if the probability p is specified, we also can get the correlation threshold TH from Eqs. (3.14) and (3.15) using the typical statistical hypothesis test method.

V. IMPLEMENTATION DETAILS AND EXPERIMENTAL RESULTS

In this section we summarize the procedure of the fast matching algorithm, and then present some experimental results to demonstrate the feasibility of this algorithm.

A. Implementation Details

The steps necessary to perform the fast correlation match using SVD are summarized below:

(i) Decompose the reference image represented as a matrix using the SVD.
(ii) For noise-free image, Eq. (3.7) is used to evaluate the correlation threshold. For this equation, α_1, E_1 (l,m) can be obtained from the SVD, $\omega(l,m)$ is the graytone values of the reference image, which is known. R_M' is the pseudo-correlation value at the pseudo-optimal matching position. For the noisy image, Eq. (3.23) is used to evaluate the correlation threshold. For this equation, α_1, $E_1(l,m)$ $\omega(l,m)$ and R_M' can be obtained in the same way as the noise-free case, and we assume that the noise is zero mean and variance given by $\mathrm{VAR}(N)$.
(iii) Calculate the pseudo-correlation values for each pixel.
(iv) Pick out the pixels whose pseudo-correlation values are larger than the threshold as candidates.
(v) At the second stage, calculate the cross-correlation values of these candidate pixels again.
(vi) Find the maximum value of cross-correlation and determine the position, which is the optimal match position.

B. Experimental Results

The technique is illustrated using digital remote sensing data collected by the Landsat and some large-scale airphoto images. The results and comparison of regular cross-correlation with the fast algorithm are illustrated in Table II. These results show that this method significantly speeds up the registration procedure for typical images between 256×256 and 1000×1000 and template images of 15×15 to 21×21.

VI. CONCLUSION

A method to speed up the template matching process has long been desired. We have developed a fast correlation registration method using Singular Value

Table II. The computational cost comparison between the Fast Correlation Method and the Conventional Method.

Method	Image	Reference image	CPU time stage 1	CPU time stage 2	SVD cost	Total cost
Regular	256 × 256	21 × 21	14m01s	/		14m01s
SVD			2m05s	1m10s	3.2s	3m15s
Regular	256 × 256	15 × 15	7m42s	/		7m42s
SVD			1m33s	1m08s	2.9s	2m48s
Regular	150 × 150	11 × 11	1m37.8s	/		1m37.8s
SVD			25 s	5 s	1.7s	30.1s
Regular	130 × 90	11 × 11	1m05s	/		1m05s
SVD			12.75s	10.2s	1.7s	25.5s
Regular	1000 × 1000	15 × 15	1h54m7s	/		1h54m7s
SVD			14m05s	9m02s	2.9s	23m10s
Regular	130 × 90	5 × 5	20.07s	/		20.07s
SVD			7.47 s	5.35s	1.7s	15.0 s

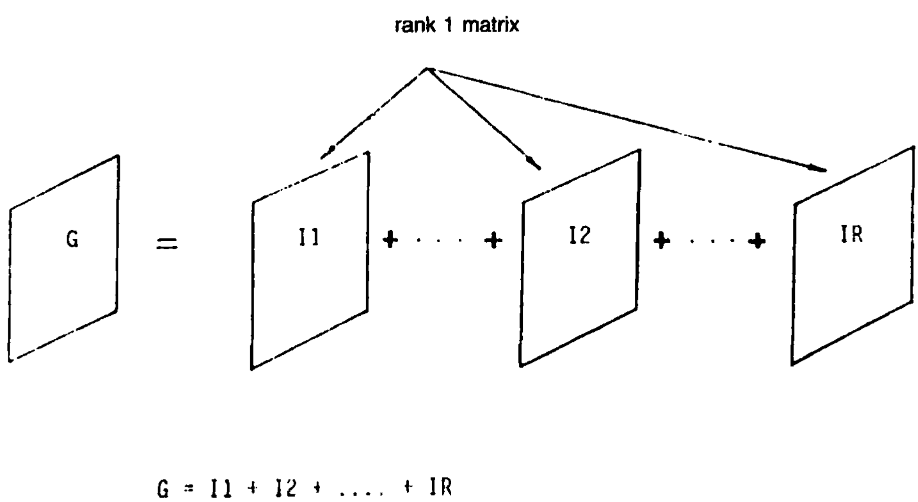

Figure 2. Singular Value Decomposition of G.

Decomposition. In this article the approach is a two-stage matching algorithm, which can be used to increase the speed of the matching process. In the first stage, the reference image is orthogonally separated by the SVD and then the low-cost pseudo-correlation values are calculated. This reduces the number of computations to $2NL$ instead of N^2L^2, where $L \times L$ is the size of the reference image, and $N\times$ is the original image size. Experimental and analytic results have been presented to show how the computation complexity is greatly improved.

References

1. A. Rosenfeld and G. J. Vanderbry, "Coarst-fine template matching," *IEEE Trans. Syst., Man Cybern.*, **SMC-7**(2), 104–107 (1977).
2. D. I. Barnea and H. E. Silverman, "A class of algorithms for fast digital registration," *IEEE Trans. Comput.*, **C-21**(2), 179–186 (1972).
3. R. Y. Wong and E. L. Hall, "Sequential hierarchical scene matching," *IEEE Trans. Comput.*, **C-27**(4) (1978).
4. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, 1977.
5. N. L. Johnson and S. Kotz, *Distributions In Statistics Continuous Univariate Distributions - 2*, Wiley-Interscience, New York, 1970.
6. E. L. Lehmaum, *Testing Statistical Hypothesis*, Wiley, New York, 1959.
7. S. Treitel and J. L. Shanks, "The design of multi-state separable planar filter," *IEEE Trans. Geosci. Electron.*, **GE-9**(1), 10–27 (1971).