Pattern Matching Based on a Generalized Transform
[Final Report]

Ram Rajagopal

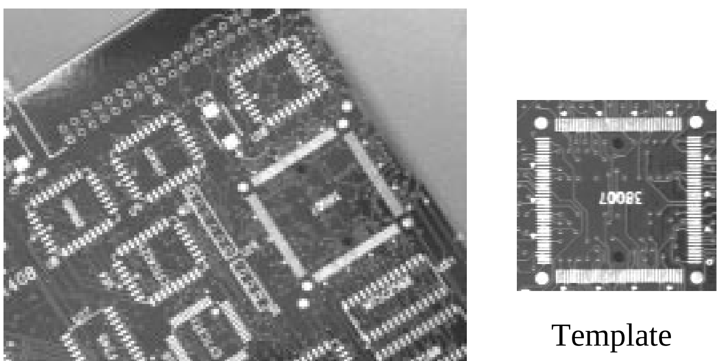
National Instruments
11500 N. Mopac Expwy., Building B,
Austin, TX 78759-3504
ram.rajagopal@ni.com

Abstract

In a two-dimensional pattern matching problem, a known template image has to be located in another image, irrespective of the template's position, orientation and size in the image. One way to accomplish invariance to the changes in the template is by forming a set of feature vectors that encompass all the variations in the template. Matching is then performed by finding the best similarity between the feature vector extracted from the image to the feature vectors in the template set. In this report we introduce a new concept of a Generalized Transform. The Generalized Transform offers a relatively robust and extremely fast solution to the described matching problem. An algorithm for scale invariant pattern matching based on the Generalized Transform is introduced.

1. Introduction

Pattern matching is an important technique in digital image processing. The evolution of computer technology has enabled many practical applications based on pattern matching, especially in industrial automation. An example of a process to be automated is the visual inspection of circuit boards. Typically, we are interested in finding a missing component in circuit boards on a production line. The procedure is based on a digital picture of the circuit board. In such image, one could search for a predefined template corresponding to the desired component. So given a test image I, we are interested in finding the location of the template I_t within this image. Typical test and template images are given in Figure 1.



Test Image I

Template Image I_t

Figure 1: Pattern matching application

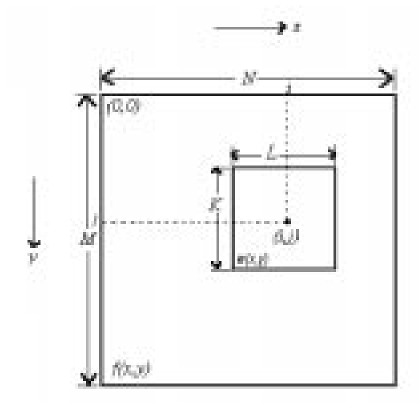


Figure 2: Classic Correlation

To properly define a pattern matching problem, all the valid transformations of the template should be clearly specified. In a majority of the applications, the template will appear shifted, rotated and scaled in the test image.

Approaches for solving the proposed problem can be divided into two categories: correlation based solutions and image understanding solutions [BB82, GMO99]. Correlation based solutions predominantly use a cross correlation to find the potential locations of the template, whereas image understanding solutions attempt to model the objects observed in the template.

In this paper we present a statistical sampling approach to pattern matching. We also introduce a new generalized transform and some of its properties. This transform provides the basis for a robust real-time scaling invariant pattern matching algorithm.

2. Classic Correlation Based Pattern Matching

Traditional pattern matching techniques include normalized cross correlation [BB82] and pyramidal matching [CC98]. Normalized cross correlation is the most common way to find a template in an image. The following is the basic concept of correlation: Consider a sub-image *w(x,y)* of size $K \times L$ within an image *f(x,y)* of size $M \times$ N, where $K \times M$ and $L \times$ N. The normalized correlation between *w(x,y)* and *f(x,y)* at a point *(i,j)* is given by

$$C(i,j) = \frac{\sum_{x=0}\sum_{y=0} (w(x,y) - \bar{w})(f(x+i, y+j) - \bar{f}(i,j))}{\left[\sum_{x=0}^{L-1}\sum_{y=0}^{K-1} (w(x,y) - \bar{w})^2\right]^{\frac{1}{2}} \left[\sum_{x=0}^{L-1}\sum_{y=0}^{K-1} (f(x+i, y+j) - f(i,j))^2\right]^{\frac{1}{2}}}$$

where $i = 0,1, \ldots M - 1$, j = 0,1 … $N - 1$, $\overline{w}$ (calculated only once) is the average intensity value of the pixels in the template *w*. The variable $\bar{f}(i,j)$ is the average value of *f* in the region coincident with the current location of w. The value of *C* lies in the range –1 to 1 and is independent of scale changes in the intensity values of *f* and w.

Figure 2 illustrates the correlation procedure. Assume that the origin of the image *f* is at the top left corner. Correlation is the process of moving the template or sub-image *w* around the image area and computing the value *C* in that area. The maximum value of *C* indicates the position where *w* best matches *f*. Since the underlying mechanism for correlation is based on a series of multiplication operations, the correlation process is time consuming. With new technologies such as MMX, multiplications can be done in parallel, and the overall computation time can be reduced considerably.

The basic normalized cross correlation operation does not meet speed requirements for many applications [BB82].

Normalized cross correlation is a good technique for finding patterns in an image as long as the patterns in the image are not scaled or rotated. Typically, cross correlation can detect patterns of the same size up to a rotation of 5° to 10° [NW99]. Extending correlation to detect patterns that are invariant to scale changes and rotation is difficult. Approaches based on multidimensional Discrete Fourier Transforms and Principal Component Analysis have been proposed [UK97]. But they are not very adequate, due to the slowness of the learning phase and requirements for non-integer operations [UK97].

3. Statistical Sampling Based Pattern Matching

Low discrepancy sequences have been successfully used in a variety of applications that require spatial or multidimensional sampling [C86, NW99]. A low discrepancy sequence can be described as a sequence that samples a given space as uniformly as possible. Thus, the density of points in relation to the space volume is almost constant.

Images typically contain a lot of redundant information. In a correlation based pattern matching a template image could be subsampled according to a two-dimensional low discrepancy sequence [NW99]. A set S of N coordinates of the template could be formed and the correlation computed only in relation to these coordinates. Such an algorithm has been proposed in [NW99b].

The algorithm has two stages. In the first, possible matches are computed based on a subsampled correlation. A threshold in the correlation value determines the exclusion or inclusion of a match. In the second, the edge information of the template is used to accurately locate the potential match indicated by the first stage. Typically, for a 100 X 100 template, a set of 61 points in enough to provide a robust correlation basis (160 times faster) for the first stage candidate list generation procedure [NW99].

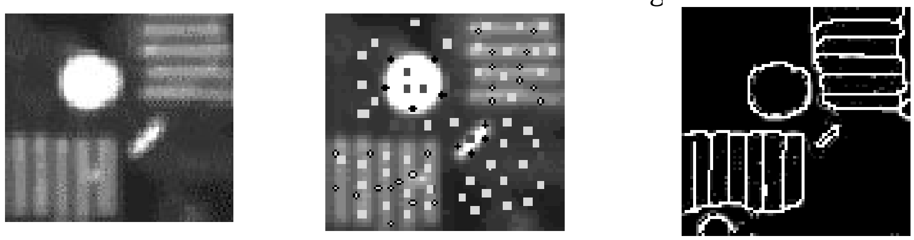


Figure 3: (a) Reference pattern, (b) Sampled pattern (c) Edges Information.

In a pattern matching application where only shift invariance is desired, a Halton low discrepancy sequence can be used [C86, NW99b]. Typically, 61-70 points from the template should be selected.

5. Generalized Transform

Assume that *N* vectors (signals) of length *N* are given. Denote these vectors by f_i. A matrix *A* can be defined, such that $Af_0 = f_1$, $Af_1 = f_2$, …, $Af_{N-1} = f_0$, if the matrix *B* (*NxN*) formed by setting each of its columns to the corresponding vector f_i is regular (non-singular). Some properties that arise from the definition of *A* and *B* are that:

P1) *AB*=*B'*, where *B'* is the matrix *B* with a column-wise shift (i.e. $f_{i+1\,mod\,N}$ corresponds to the column *i* of *B'*). *B* is regular and so is *B'*. Thus $A = B' B^{-1}$.

P2) $A^N = I$ (*NxN* identity). Thus, eigenvalues of *A* are given by ($\lambda_k = exp\left(\frac{2\pi}{N} k\right)$, *k = 0, …, N-1*).

P3) The matrix *A* can be decomposed as $A = X_B V X_B^{-1}$, where *V* is the *NxN* diagonal matrix formed by the eigenvalues $\lambda_k = exp\left(\frac{2\pi}{N} k\right)$ [GL91].

From the stated properties it is clear that the *NxN* matrix X_B^{-1} expresses the desired Generalized Transform (GT). Theorem 1 proves the shift invariance property for the GT. Theorem 2 shows that if the vectors f_i are shifted versions of each other, then X_B^{-1} is the Fourier matrix. Theorem 3 provides a way to compute the GT in an efficient manner. For proofs see the appendix.

Theorem 1: The matrix X_B^{-1} defines a shift invariant transformation for the set of vectors f_i.

Theorem 2: If the vectors f_i are shifted versions of each other (i.e. $f_i = f([n+i]_N))$ then X_B^{-1} is the Fourier matrix (for a definition of the Fourier matrix see [GL91]).

Theorem 3: Given a regular matrix B the generalized transform can be computed as $X_B^{-1} = D^{-1} W_N B^{-1}$, where D is an arbitrary complex diagonal matrix. To define a unitary transform the diagonal elements of D should be set to $d_k = \sqrt{\sum_{i=0}^{N-1} \left|B_i^{inv}(k)\right|^2}$, where $B_i^{inv}(k)$ represents the Discrete Fourier Transform of the ith column of B^{-1}.

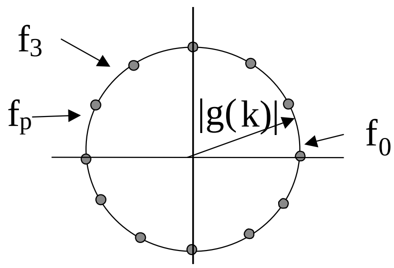


Figure 4: GFT Based Circle

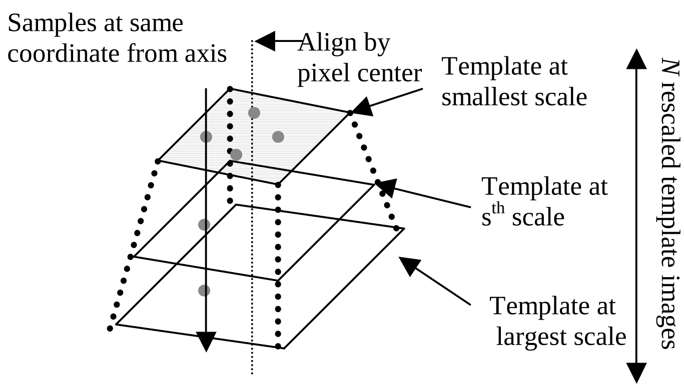


Figure 5: Pattern Matching Sampling Strategy

Choosing a frequency in the GT domain corresponds to selecting a line of the matrix X_B^{-1}. Due to the shift invariance property, for a fixed frequency, the set of vectors f maps to points in a circle in the complex plane (figure 4). If we set $g = X_B^{-1} f_0$, then $|g(k)|$ is the radius of the circle at frequency k. Theorem 3 states that for a unitary transform this radius is given by $1/d_k$. Moreover, the sequence $1/d_k$ $(k=0,\ldots,N-1)$ forms a spectrum equivalent to the Fourier spectrum.

4. Scaling Invariant Pattern Matching

The requirement for scaling invariance might arise in applications where the distance between the camera and the imaging plane is variable. Usually, in scaling invariance applications the scaling range is fixed and finite due to physical constraints of the imaging system.

Given an arbitrary vector $\bar{f}$ and a set of vectors, represented by B (see section 3), a simple projection based algorithm for detecting the closest vector to $\bar{f}$, among the columns of B is presented in Table 1 (a). The procedure assumes that $\bar{f}$ is close enough to a vector in B, so that projecting to a lower dimensionality does not compromise accuracy [RW00b]. An optimal way of selecting the projection matrix, based on the shift invariant property of the GT, is presented in Table 1 (b). The procedure is based on an optimal procedure for detecting delays in signals [RW00b].

Projection Match Algorithm
STEP1: Compute and store $P = FB$, where F is an arbitrary *KxN* matrix**. (Done once at learn time)**
STEP2: Compute $\bar{p} = F\bar{f}$
STEP3: Find the closest line-vector to $\bar{p}$, among the lines of P.

Choosing a Projection Matrix *F*
STEP1: For the matrix B compute X_B^{-1} unitary, according to Theorem 3.
STEP2: Select *K/2* frequencies (K integer) of the GT of B according to the optimization below [RW00b]. Where *g(k)* is the GT of f_0 at frequency k (f_0 is the first column of B). $$\max_{k_1,\ldots,k_{K/2-1}} \min_{r} \sum_{i=0}^{K/2-1} \lvert g(k_i)\rvert^2 (1-\cos\frac{2\pi[k_i r]_N}{N}) \quad r, k_i \in \{1,\ldots,N/2-1\}$$
STEP3: Set the lines of F to be the real and imaginary parts of the selected *K/2* lines of X_B^{-1}. The k^{th} frequency corresponds to the k^{th} line of X_B^{-1}.

Table 1: (a) Projection Match Algorithm and (b) selecting the Projection matrix.

A scaling invariant pattern matching algorithm, based on the Projection Match algorithm and in statistical sampling is presented in Table 2. The algorithm explores the finite range of the scaling factor to create an efficient image matching process.

Figure 5 presents the sampling process that generates the set $\{f_0,\ldots,f_{N-1}\}$. The template is rescaled (using bilinear interpolation) to N different discrete scaling factors, evenly distributed in the finite range. At each scaling factor, the same N distinct Halton points (in reference to the center of the image) are sampled. Note that the template image rescaled to the smallest scale determines the extent of the sampling area.

The matching phase consists in sliding the sampling structure defined by the statistical sampling over the test image (as in figure 2), and at each pixel location extracting the corresponding vector $\bar{f}$.

Then, finding the vector f closest to $\bar{f}$ determines a match. A full normalized correlation between the chosen vector f and $\bar{f}$ determines a score. The best match is the match with highest score among all pixel locations. This procedure is presented in Table 2 (b).

Pattern Matching Learning Phase
Inputs: template Image, scaling factor range (s_0, s_1) and N (matching granularity). Define $\Delta s = (s_1 - s_0)/N$.
STEP1: Create 2D Holton set for rectangle of size $s_0 X$ by $s_0 Y$, where *(X,Y)* is the template size. Store the set of N reference coordinates $S=\{(x_0, y_0),...,(x_{N-1}, y_{N-1})\}$.
STEP2: For i=1 to N {   Rescale template image to scaling factor $s = s_0 + i\Delta s$ Extract the pixel values of the N Holton samples   (image center as reference) ➔results in f_i }
STEP3: Set each f_i as a column of B and compute the projection matrix F as suggested in Table 1 (b). Randomized Correlator for lines of P can be learnt in this step [RW00].

Pattern Matching Runtime
Inputs: Test Image, learnt data.
For each pixel (i,j) do:
STEP1: Shift the set of reference coordinates to the pixel (i,j). Extract the intensity (pixel) values to $\bar{f}$.
STEP2: Compute $\bar{p} = F\bar{f}$
STEP3: Find line vector in P closest to $\bar{p}$ (Randomized Correlator could be used)
STEP4: Compute normalized correlation between f_i (corresponding to the line vector in STEP3) and $\bar{f}$. This is the score.
STEP5: If (score > threshold) match is found, exit search.

Table 2: Scale Invariant Pattern Matching Algorithm (a) learning and (b) matching

5. Computational Complexity and Performance

The main advantage of the pattern matching algorithm presented in section 4 is its relatively low computational complexity compared to classical procedures. In order to compute the computational complexity, assume that the template is of size *MxM*, the test image is of size *NxN* and that K discrete scale steps are used. In all cases, complexity will be measured as number of required multiplications. Assume also that $M << N$.

The classic correlation approach would be to compute the correlation of the test image with the K rescaled template images. The classic statistical sampling approach would incorporate statistical sampling into the correlation computation. Finally, the algorithm proposed in section 4 incorporates GT projection and a randomized correlator to reduce computation even further. By inspection we obtain the number of multiplications in Table 3.

Algorithm	Number of Multiplications
Classic Correlation	KM^2N^2
Classic Statistical Sampling	K^2N^2
Proposed Algorithm	$((p+1)K + O(p))N^2 \approx 2KN^2$

Table 3: Computational complexity of the proposed algorithm

The scale invariant pattern matching algorithm was implemented in the LabVIEW environment. Based on experiments, the required number of Halton samples was determined to be between 60-80 (*N*) and the number of projection vectors between 4 and 8 (*K*).

6. Summary

In this paper we presented a new algorithm for real-time scale invariant pattern matching. The proposed algorithm was based on the shift invariance property of a new transform. We showed how to compute this transformation for a given set of basis signals.

One advantage of the proposed approach is that any affine transform can be included as part of the algorithm. Transformed versions of the template image should be Halton sampled and these samples included in the estimation of the generalized transform. A continuation of this work would be to explore the relationships between the GT and other transforms, and study the application of the proposed pattern matching approach to image indexing.

7. References

[BB82] D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ, 1982.
[C86] R. L. Cook, “Stochastic Sampling in Computer Graphics”, *ACM Transactions on Graphics*, Vol. 5, No. 1, pp. 51-72, 1986.
[GL91] Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 1991.
[UK97] M. Uenohara and Takeo Kanade, “Use of Fourier and Karhunen-Love Decomposition for Fast Pattern Matching With a Large Set of Templates”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 8, pp. 891-898, 1997.
[CC98] V. Cantoni, L. Cinque, C. Guerra, S. Levialdi and L. Lombardi, “2-D Object Recognition by Multiscale Tree Matching”, *Pattern Recognition*, Vol. 31, No. 10, pp. 1443-1454, 1998.
[GMO99] Michael T. Goodrich, Joseph S. B. Mitchell, and Mark W. Orletsky, “Approximate Geometric Pattern Matching Under Rigid Motions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 4, pp. 371-379, 1999
[NW99] Dinesh Nair and Lothar Wenzel, "Image Processing and Low-Discrepancy Sequences", *Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, Proceedings of SPIE, Vol. 3807, pp. 102-111, July 1999.
[NW99b] Dinesh Nair and Lothar Wenzel, "Application of Low Discrepancy Sequences and Classical Control Strategies for Image Registration”, *Proceedings of the 2nd International Conference on Information Fusion*, Sunnyvale CA, pp. 706-714, July 1999.
[RW00] Ram Rajagopal and Lothar Wenzel, “Algorithms for Computing Fast Peak Correlation”, submitted to *SIAM Journal of Computation*, October 2000.
[RW00b] Ram Rajagopal and Lothar Wenzel, “Geometric Cross Correlation”, in preparation for submission to *IEEE Transactions in Signal Processing*, December 2000.