Pattern Matching Based on a Generalized Transform
[First Report]

Ram Rajagopal

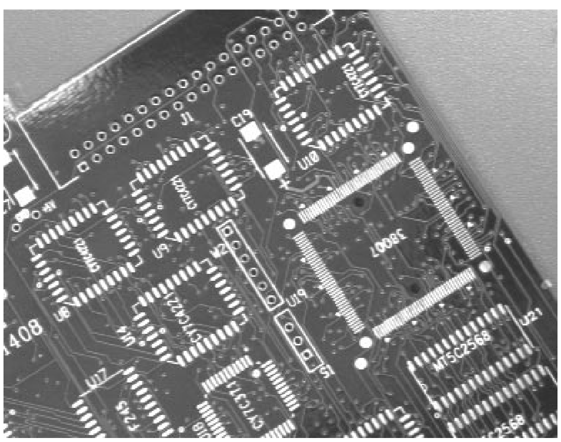
Department of Electrical and Computer Engineering
University of Texas at Austin
Austin, TX 78712
ram.rajagopal@ni.com

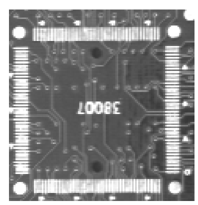
Abstract

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. The application of the Generalized Transform to scale invariant pattern matching 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. Then 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.





Template Image I_t

Test Image I

Figure 1: Pattern matching application

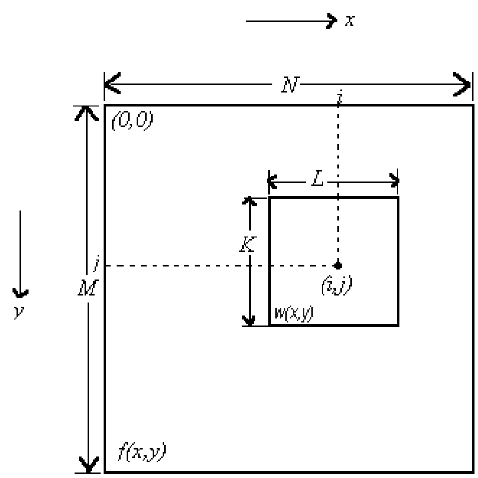


Figure 2: Classic Correlation

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 literature survey we present some classical approaches to correlation based pattern matching. Then we present an innovative approach to the same problem using a statistical sampling approach. Finally we introduce a new generalized transform, that when used in conjunction with statistical sampling provides the basis for a robust 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 $M \times \mathrm{N}$, where $K \times M$ and $L \times \mathrm{N}$. 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. For scale-invariant matching, you must repeat the process of scaling or resizing the template and then perform the correlation operation. This adds a significant amount of computation to the matching process.

A multidimensional Discrete Fourier Transform approach could be used to accelerate the correlation computation [UK97]. But, usually the size of the template offsets such advantages. Moreover the lack of robustness to deformations of the template turns out to be the main problem with correlation.

Another classical approach is to use Principal Component Analysis to extract information from the template and similarly extract information from the image at each bounding location [UK97]. The approach is not very successful because the learning phase is very slow (~2 hours) and the matching phase computationally intensive (lacking real time).

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. Thus, 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] and successfully used in a variety of applications.

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.

Shift Invariant Pattern Matching

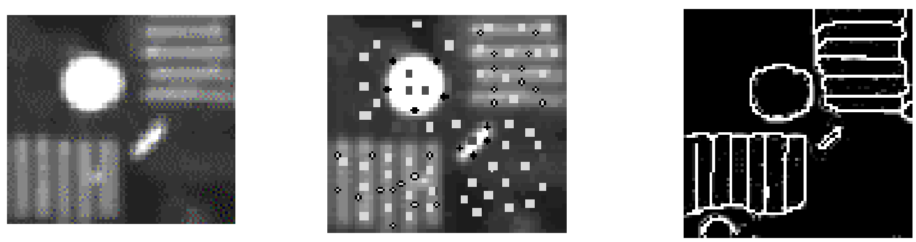


Figure 3: (a) Reference pattern, (b) Sampled pattern (c) Edges Information.

3.1 Shift Invariance

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

The selection process starts by generating the subsampling structure, whose first coordinate is always (0,0) [NW99b]. Then, the offset that should be used to determine the actual set of coordinates is chosen. The offset should maximize the robustness of the set, so that small pixel coordinate variations (up to 3 pixels) change the pixel intensities in the chosen coordinates as little as possible.

3.2 Rotation Invariance

The rotation invariant approach could be extended from the shift invariant approach as suggested in section 2. But there is a more efficient approach that still keeps the spatial subsampling idea. Instead of using a uniformly distributed sequence, a subsampling (discrete) circle could be used.

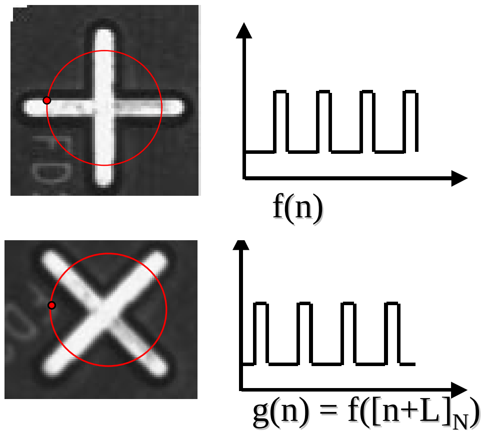


Figure 4: Rotation Invariant Pattern Matching

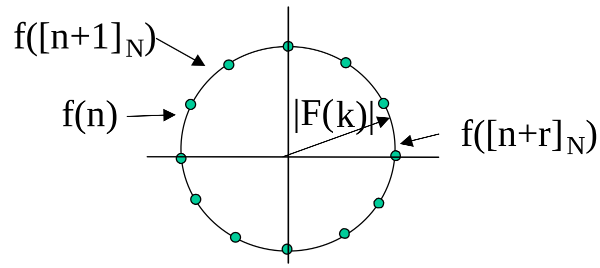


Figure 5: DFT Based Algorithm

Computing the candidate list is then simply computing a full circular cross correlation at each pixel point. If the chosen sequence of coordinate points is visualized as a 1D sequence of points *f(n)*, the correlation procedure is defined as $L = \arg\max_{n} \sum_{k=0}^{N-1} f([k+n]_N)g(n)$.

Figure 4 illustrates the procedure. The computation of the cross correlation can be accelerated if we assume that a match should always be sufficiently close to the given template. In this case an algorithm based on a Discrete Fourier Transform (DFT) can be used. The algorithm is presented in Table 1. The main idea is to explore the spatial distribution of the complex points of the DFTs of the shifted versions of *f(n)*, given a fixed frequency (Figure 5). If more robustness is required, the proposed procedure can be extended to include as many frequencies as desired.

Given the signal f(n), choose a frequency k that maximizes $\left\|F(k)\right\|$, where *F(k)* is the DFT of *f(n)*.
Compute and store the vectors $\left(Re(exp(\frac{2\pi}{N}k)F(k)), Im(exp(\frac{2\pi}{N}k)F(k))\right)$
(at match time) Compute the DFT at k of g(n) (sampling set S, at coordinate (i,j))
(at match time) Find the closest vector to $\left(Re(G(k)), Im(G(k))\right)$, in the stored list

Table 1: Algorithm for fast rotation invariant pattern matching

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.

A better approach is to proceed as suggested in the rotation invariant case. Given a set *S* of *N* subsampling points, sample scaled versions of the template and use this information to perform a complete correlation. The scaling factor range $[s_b, s_t]$ is given as part of the problem specification. This range can be discretized using a finite number of levels *P*. The template image is scaled and sampled at the coordinate locations defined by *S* and an offset, for each discrete value of the scaling factor. The offset can be chosen independently for each scale. The result of the proposed operations is a set of *P* vectors $F = \{f_0, ..., f_{P-1}\}$ of length *N*.

The matching phase consists in sliding the sampling structure defined by *S* over the test image and at every location computing the correlation with respect to each of the vectors in *F*. The best match indicates the potential scaling factor. A full normalized correlation with respect to the corresponding scaled template is then computed to obtain the matching score.

We can accelerate the matching process by using an idea similar to the one proposed for the rotation invariant case. There the vectors f_i corresponded to shifted versions of *f*, and the DFT mapped them to a circle if a single frequency was used. In the current situation we should define a more generic transformation that takes the vectors in *F* to a circle.

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). This property implies that all eigenvalues of A are the roots of unity ($\lambda_k = exp\left(\frac{2\pi}{N} k\right)$, *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 λ_k [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 1: The matrix X_B^{-1} defines a shift invariant transformation for the set of vectors f_i.

Proof: From the definition of the matrix *A*, it is clear that $f_p = A^p f_0$. Using property P3 we can write $f_p = X_B V^p X_B^{-1} f_0$, resulting finally in $X_B^{-1} f_p = V^p X_B^{-1} f_0$, which is the matrix form of the shift invariance property of the DFT (note that for a DFT $f_p = f([n+p]_N)$).

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.

Proof: For the specified set of vectors, $B' = I' B$ where *I'* is the column-wise shifted identity matrix (for example, for N=3, $I' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$). It is clear then that $A = I'$. If we set $X_B^{-1} = F$, where *F* is the DFT matrix [UK97], then $I' = F^{-1} V F$ which can be verified true by direct computation.

Choosing a frequency in the GT domain corresponds to selecting a line of the matrix X_B^{-1}. In this case too, for a fixed frequency, the set of vectors *f* maps to points in a circle in the complex plane. If $g = X_B^{-1} f_0$, then $|g(k)|$ is the radius of the circle at frequency *k*.

For scale and shift invariant pattern matching, the same algorithm presented for the rotation invariant case can be used, replacing the DFT by the GT. In fact if we allow for rotation of the template at different scales, a complete pattern matching solution can be built on the GT.

6. Summary

In this project we propose to implement the scale invariant pattern matching mechanism based on this new transform. We also propose to develop some additional properties of the GT. The GT based approach is promising from a computational efficiency aspect. Moreover it allows for easy integration towards a global solution that is able to handle template rotation, scaling and shift.

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.

[HKR93] D. P. Huttenlocher, G. A. Klanderman and W. J. Rucklidge, “Comparing Images Using the Hausdorff Distance”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 9, 850-863, 1993.

[K97] Donald E. Knuth, *The Art of Computer Programming: Fundamental Algorithms*, Addison-Wesley Pub. Co, 1997.

[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, Aug. 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, April 1999

[NW99] Dinesh Nair and Lothar Wenzel, "Image Processing and Low-Discrepancy Sequences", *Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, Proceedings of SPIE, Volume 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.