Two-Dimensional Discriminative Filters for Image Template Detection

Alexandre P. Mendonça
Instituto Militar de Engenharia &
PEE/COPPE/UFRJ
Departamento de Engenharia Elétrica
Rio de Janeiro, RJ, Brazil
alexmend@aquarius.ime.eb.br

Eduardo A. B. da Silva
Programa de Engenharia Elétrica
COPPE/DEL/UFRJ
Rio de Janeiro, RJ, Brazil
Cx. Postal 68504, CEP 21945-970
eduardo@lps.ufrj.br

ABSTRACT: The image template detection is usually a very important halfway step for a computational vision algorithm. In general, the basic idea of this type of algorithm is to receive an input image and generate a set of lines, borders, edges and other well known geometrical forms as outputs. Recently, Ben-Arie and Rao have proposed a linear filter that, given a template, maximizes the energy concentration in a single sample of its output. In this paper, we propose a two-dimensional generalization of this method, formulating the template matching problem as a multivariable optimization. Three different objective functions are investigated. Simulation results show that the proposed method performs well on real images.

1. INTRODUCTION

The template matching problem has been investigated using several different approaches, as, for example, in [1] and [2]. In these works, edge detectors are built based on local image statistics. Ben-Arie et al. have been working on a new type of template detectors, referred to as EXM (***Expansion Matching***) [3-7]. These detectors are based on the optimum decomposition of a signal on a frame constructed from a given template. A template is detected whenever the coefficient corresponding to it has magnitude larger than a threshold. The success of this method relies on the fact that such expansion is equivalent to filtering the input signal with a linear filter, which maximizes the energy concentration in one sample of its output. Details can be found in [3-7]. The formulation used there assumes a one-dimensional template that requires either a scanning or some form of rank reduction on the data in the case of multidimensional signals.

In this paper, we propose a two-dimensional generalization of the methods in [3]. In section 2, we describe the two-dimensional mathematical formulation. In sections 3 and 4, we propose two alternative objective functions for discriminative filter optimization. In section 5, we test the proposed method in a template detection system. In section 6, experimental results are given. Section 7 presents the conclusions.

2. TWO-DIMENSIONAL DISCRIMINATIVE FILTERS

When discriminative filters are used for template detection, one usually wants to maximize the energy of an output sample whenever a match is found. The discriminative signal to noise ratio (DSNR), defined in [3], is a measure that accounts not only for the maximum energy of a sample, but also considers its energy in relation to the other samples. For two-dimensional discriminative filters, we define a two-dimensional DSNR ($DSNR_2$) as:

$$DSNR_2 = \frac{c_{i,j}^2}{\left(\sum_m \sum_n c_{m,n}^2\right) - c_{i,j}^2} \quad (1)$$

The $c_{m,n}$ coefficients are obtained after a two-dimensional convolution between an input image window $u_{m,n}$ and a linear operator Θ having impulse response $\theta_{m,n}$, that may be computed for each template to be matched. The coefficient $c_{i,j}$ is the one where we wish to concentrate the output signal energy.

In our approach, we do not make any restrictions related to the dimensions of either the image window or the operator $\boldsymbol{\Theta}$. Let $\mathbf{U}$ be a $M_1 x N_1$ window and Θ an $M_2 x N_2$ kernel. Then, the resulting image $\mathbf{C}$ after the convolution between U and Θ has dimensions $M_1+M_2-1 x N_1+N_2-1$ and can be expressed as.

$$c_{m,n} = \sum_{m'} \sum_{n'} u_{m-m',n-n'} \theta_{m',n'} \quad (2)$$

It is easy to see that the maximization of $DSNR_2$ in equation (1) is equivalent to the minimization of the following expression:

$$f(\Theta) = \frac{\sum_m \sum_n c_{m,n}^2}{c_{i,j}^2} \quad (3)$$

From equation 2, the function $f(\Theta)$ to be minimized becomes:

$$f(\Theta) = \frac{\sum_m \sum_n \left(\sum_{m'} \sum_{n'} u_{m-m',n-n'} \theta_{m',n'}\right)^2}{\left(\sum_{m'} \sum_{n'} u_{i-m',j-n'} \theta_{m',n'}\right)^2} \quad (4)$$

$f(\Theta)$ is minimum when its gradient is equal to zero, that is

$$\nabla f(\Theta)_{r,s} = \frac{\partial f(\Theta)}{\partial \theta_{r,s}} = 0 \qquad \begin{matrix} r=1,...,M_2 \\ s=1,...,N_2 \end{matrix} \qquad (5)$$

The above condition is a non-linear system of M_2N_2 equations. To get a numerical solution, we use a gradient-based method. The expression for $\nabla f(\Theta)$ is

$$\nabla f(\Theta)_{r,s} = 2 \{ c_{i,j} \sum_m \sum_n u_{m-r,n-s} c_{m,n} - u_{i-r,j-s} \sum_m \sum_n c_{m,n}^2 \} / c_{i,j}^3 \quad (6)$$

A good initial solution is of paramount importance for gradient-based methods. We have found that to use the real part of the inverse two-dimensional Fourier transform of the quotient between the desired response ($DSNR_2$ infinite, that is, $c_{m,n} = \delta_{i-m,j-n}$) and the Fourier transform of U leads to satisfactory results We evaluated the performance of the method using, as image templates, corners inside 7x7 windows (see figure 1). The support region of the filter Θ is the same as the one of the templates. The coordinates of the coefficient that concentrates the energy in case of a match, that is, (i,j) in equations (3), (4) and (6), have been chosen as the ones of the central sample of the output signal.

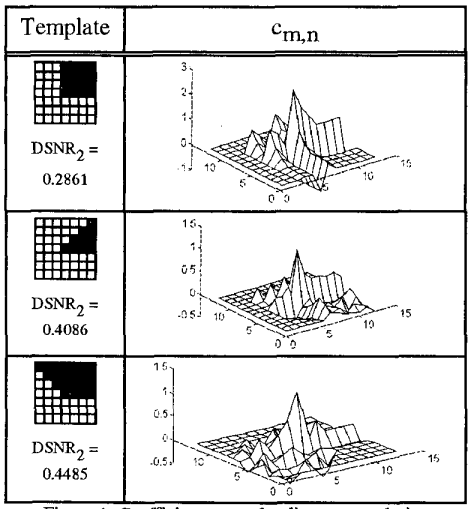

Figure 1: Coefficients $c_{m,n}$ after linear convolution between the shown template and the found Θ.

3. ALTERNATIVE APPROACH

The formulation in Section 2 finds the Θ that maximizes the $DSNR_2$ for a given template. However, it does not avoid that a different template might give an even larger $DSNR_2$ when applied to Θ, what can lead to false detections. This is indeed the case, as will be seen later in Section 6.

A possible way to solve this problem is to consider $DSNR_2$ as a function of Θ and U and to look for the Θ that gives the maximum $DSNR_2$ when U varies. More formally, we want to solve:

$$\partial f(U,\Theta) / \partial u_{r,s} = 0 \qquad r=1,...,M_1;\ s=1,...,N_1 \quad (7)$$

Since the above equations are of difficult analytic solution, we have created an objective function referred to as the **discriminative potential**, which is the sum of the squares of each of the terms in the above equation, defined below

$$g(\Theta) = \sum_r \sum_s (\partial f(U,\Theta) / \partial u_{r,s})^2 = \sum_r \sum_s \{ c_{i,j} \{ \sum_m \sum_n \theta_{m-r,n-s} c_{m,n} \} - \theta_{i-r,j-s} \{ \sum_m \sum_n c_{m,n}^2 \} \}^2 \quad (8)$$

To minimize $g(\Theta)$ in equation (8) is equivalent to solve the system in equation (7).

We have solved equation (8) numerically for the same templates as in Section 2. The initial solutions were the kernels generated with the method in Section 2. The results obtained are shown in table 1. The results for template matching on real images are shown in Section 6. As we will see in Section 6, we obtain more false detections and wrong detections then when using the method in Section 2.

Template	$DSNR_2$	$g(\Theta)$
90^0 corner	0.1851	15.55
45^0 corner	0.3721	3.97
135^0 corner	0.0953	31.48

Table 1: $DSNR_2$ and $g(\Theta)$ obtained by alternative approach.

4. MIXED APPROACH

As shown in the simulations (see section 6), the approach in Section 3 does not give better template detection performance. This indicates that a large value of $DSNR_2$ is important in those cases. The two previous approaches seem to be in conflict because they were developed from two different indicators: the functions $f(\Theta)$ and $g(\Theta)$. A compromise would be to use, as objective function, a combination of the functions $f(\Theta)$ and $g(\Theta)$. We then define a Detection Function as

$$DF(\Theta) = (1 - K)\{ f(\Theta) - 1 - 1/Z \}^2 + K\ g(\Theta), \quad (9)$$

where Z is an upper limit for $DSNR_2$ and K is a constant which weights the contributions of both $f(\Theta)$ and $g(\Theta)$, found experimentally. In general, we tend to use K next to 1, tending to privilege the discriminative power over the energy concentration.

The training with some templates generated interesting curves. Figure 2 shows the results for the 3 types of corners. The initial solution was the same one used in Section 2.

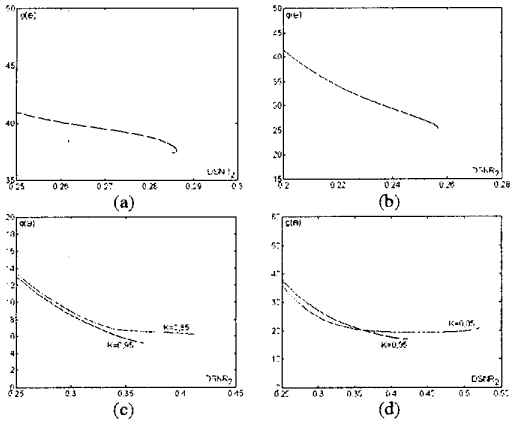

Figure 2: Compromise between $DSNR_2$ and $g(\Theta)$, with Z=1, for (a) 90^0 (K=0,85), (b) 90^0 (K=0,95), (c) 45^0 and (d) 135^0 corners.

Figure 3 shows the operator and the output signals for the 135^0 corner.

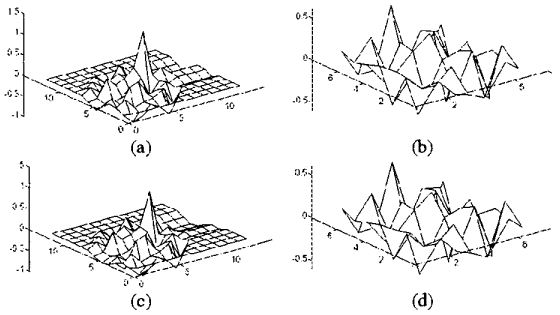

Figure 3: (a,b) C and Θ for the 135^0 corner with K=0.85 and K=0.95 (c,d).

5. PRACTICAL METHOD FOR TEMPLATE DETECTION

The natural application for the proposed algorithm is the detection of two-dimensional image patterns. As shown in figure 5, the basic idea is to compute Θ for a template. So, this window is used to filter an image signal that is generated by an appropriated transform of an input window image. The output window that has a $DSNR_2$ value larger than a threshold is considered morphologically matched to the original template (the one that generated Θ).

In order not to be necessary to compute a different Θ for each rotated version of a template, which would increase enormously the number of different filter kernels used, we pre-rotate the window containing each input image. We do this rotation such that the vector joining the window center and the centroid of the pattern is aligned with the same vector computed for the template. The block labeled "appropriate rotation" in figure 4 represents this. More specifically, this is done as follows:

- choose a template (fig.5.1);
- compute the center of the chosen template (fig.5.2);
- take the image window to be matched (fig.5.3);
- compute the window center (fig.5.4);
- compute the angle α, taking the two centers and using the central pixel as origin (fig.5.5);
- rotate the window of α, trying to superimpose the window over the template (fig.5.6).

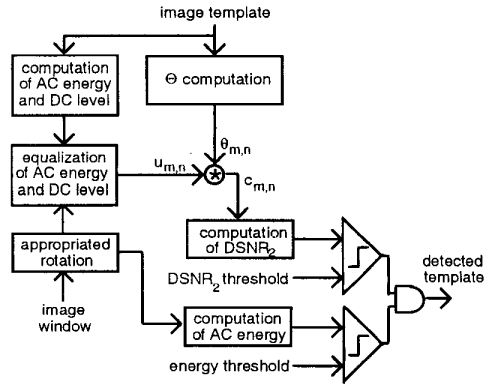

Figure 4: Simplified block diagram of the image pattern detector.

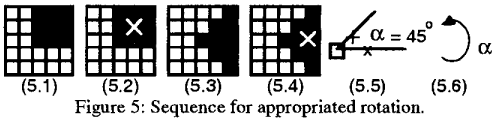

Figure 5: Sequence for appropriated rotation.

h

Note that, in order to compute the new pixel value of coordinates (i , j) after the rotation (see fig. 5.6), it is sufficient to copy the pixel of coordinates: (i cos α + j sin α , j cos α - i sin α).

A possible problem with this method occurs when we take a window that is equal to the complement of the template (see fig.6). If we apply the algorithm as in fig. 5, the window should match the template, because the computation of $DSNR_2$ is invariant by a multiplication by a constant ("window" = -1 x "template"). However, in the present implementation of the algorithm, a 180^0 rotation would be done, and there would not be a match. We have solved this problem by adding 180^0 to α whenever, after the DC level elimination, the number of pixels with negative values is smaller than the number of pixels with positive values.

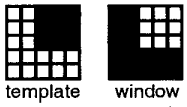

Figure 6: Window equal to the template complement.

6. SIMULATION RESULTS

To evaluate the performance of the proposed detection system, we used the templates in figure 1 and filters obtained using the three objective functions presented in section 2 (method A), section 3 (method B) and section 4 (method C). Table 2 lists the processing parameters.

From figure 7, we can see that the algorithm is very efficient for detecting templates on a two texture environment, as on the chimney and the sky boundaries. For the 90^0 corners, methods B and C give the same results, canceling one false corner detected in first method. Relating to 45^0 and 135^0 corners, we see that

method A detects a high number of false corners. For 45^0 corners, both methods B and C perform much better than method A. On the other hand, for the 135^0 case, method B causes a multiplication of the number of false corners, while method C performs well.

We can compare our results with the results obtained by Nandy and Ben-Arie [4], shown in figure 7l. In both cases, the results are very satisfactory. In [4], the windows suffered a KL transformation before the application of EXM detector. This transformation reduced the image dimension ($2 \rightarrow 1$) in order for the one-dimensional formulation to be applied.

Template	$DSNR_2$ (method A)	$DSNR_2$ (method B)		Thresh.	Minimum Gray Level Excursion
90^0 corner	0.2850	0.1851		0.16	11.7
45^0 corner	0.4163	0.3721		0.09	11.7
135^0 corner	0.4493	0.0953		0,09	11.7
Template	$DSNR_2$ (method C)	K	Z	Thresh.	Minimum Gray Level Excursion
90^0 corner	0.2470	0.95	1.0	0.16	11.7
45^0 corner	0.3611	0.95	1.0	0.09	11.7
135^0 corner	0,.5196	0.85	1.0	0,09	11.7
135^0 corner	0.4219	0.95	1.0	0.09	11.7

Table 2: Processing parameters table.

7. CONCLUSIONS

In this work, we first defined the two-dimensional extension of discriminative signal-noise ratio concept, defined in [3]. We have proposed three objective functions for the design of the discriminative filters. One of them (method A) is based on the maximization of the $DSNR_2$. Another one (method B) attempts to find the optimum filter, which gives the highest DSNR for the given template. Method C uses, as objective function, a combination of methods A and B. A potential application for the proposed filter is the pattern detection. An analysis of the computed filters for different templates was made for real images. We have found that method C generates discriminative filters with better performance than the other two.

The results were satisfactory, as good as the others presented in the literature. Moreover, the proposed method can be used to detect any two-dimensional pattern, even ones that cannot be adequately dealt with by the one-dimensional EXM in [3-7].

The proposed two-dimensional formulation lends more flexibility to the class of template matching approaches based on discriminative filtering, when compared with the one-dimensional approaches. Also, the use of method C, as opposed to the more straightforward extension of the EXM method (method A) has been shown to have superior performance.

8. REFERENCES

[1] Ji, Q. and Haralick, R.M., "Quantitative Evaluation of Edge Detectors Using the Minimum Kernel Variance Criterion", IEEE International Conference on Image Processing, 1999.

[2] Abdou, K.E. and Pratt, W.K., "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors", Proc. of IEEE, 67(5)753-763,1979.

[3] Ben-Arie, J. and Rao, K.R., "A Novel Approach for Template Matching by Nonorthogonal Image Expansion", IEEE Transactions on Circuits and Systems for Video Technology, Vol.3, N^0.1, Fev.1993.

[4] Nandy, D. and Ben-Arie, J. , "EXM Eigen Templates for Detecting and Classifying Arbitrary Junctions", IEEE International Conference on Image Processing, 1998.

[5] Rao, K.R. and Ben-Arie, J. , "Multiple Template Matching Using the Expansion Filter", IEEE Transactions on Circuits and Systems for Video Technology, Vol.4, N^0.5, out.1994.

[6] Rao, K. R. and Ben-Arie, J., "Optimal Edge Detection Using Expansion Matching and Restoration", IEEE Transactions on Pattern Analysis and machine Intelligence, Vol.16, N^0.12, dez.1994.

[7] Wang, Z. , Rao, K.R. e Ben-Arie, J. , "Optimal Ramp Edge Detection Using Expansion Matching", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.18, N^0.11, dez.1996.h

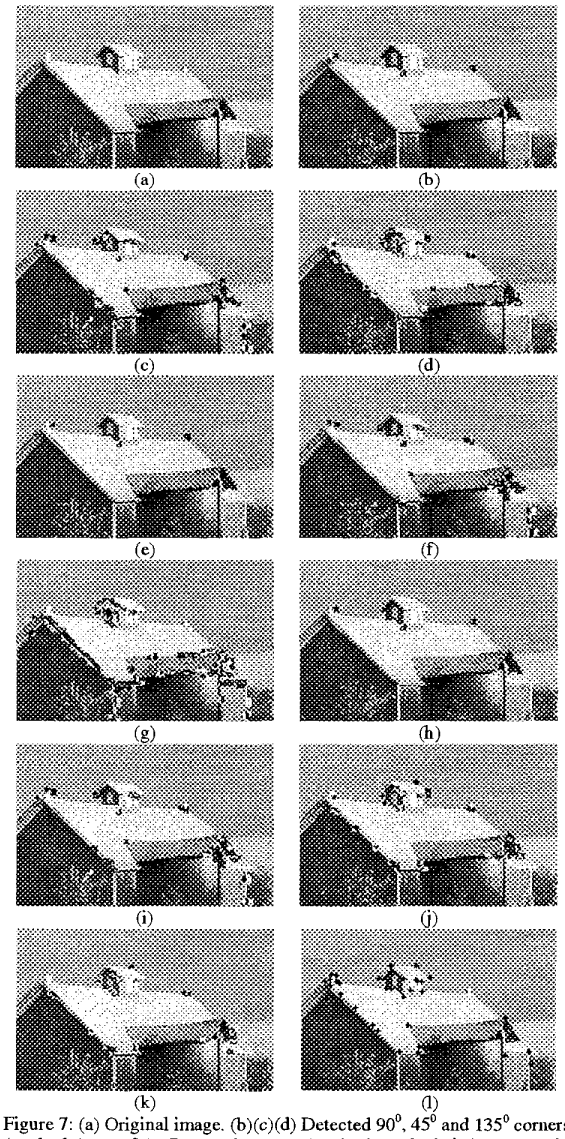

Figure 7: (a) Original image. (b)(c)(d) Detected 90^0, 45^0 and 135^0 corners (method A). (e)(f)(g) Detected corners (method B). (h)(i)(j)(k) Detected corners in method C. (see table 2) (l) Results of [4].