

# Filtros Discriminativos para a Detecção de Gabaritos de Imagem

Alexandre P. Mendonça  
Departamento de Engenharia Elétrica  
Instituto Militar de Engenharia  
& PEE/COPPE/UFRJ  
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

**RESUMO:** A detecção de gabaritos da imagem é um passo intermediário importante para um algoritmo de visão computacional. Geralmente, a idéia básica deste tipo de algoritmo é receber uma imagem como entrada e gerar, como saída, um conjunto de linhas, bordas, quinas ou outras formas geométricas bem conhecidas. Recentemente, Ben-Arie e Rao propuseram um filtro linear que, dado um gabarito de sinal, maximiza a concentração de energia numa das amostras do sinal de saída do filtro. Neste artigo, nós propomos uma generalização deste método, formulando o casamento de gabaritos como um problema de otimização envolvendo múltiplas variáveis. Propomos e investigamos o uso de três funções objetivos. Os resultados das simulações mostram que o método trabalha muito bem com imagens reais.

## 1. INTRODUÇÃO

O problema do casamento de gabaritos tem sido encarado com as mais diversas abordagens, como, por exemplo, nos trabalhos [1] e [2]. Neles, foram construídos detectores de quinas baseados nas estatísticas locais da imagem. Já Ben-Arie e sua equipe têm trabalhado em um novo tipo de detectores, chamados por eles de EXM, iniciais de "*Expansion Matching*" [3-7]. Estes detectores são baseados em decomposição de um sinal numa base especial de funções para um determinado gabarito. Nestes trabalhos, um gabarito é detectado se uma determinada função objetivo, calculada sobre as amostras da saída do filtro, possuir um valor superior a um limiar. O sucesso deste método vem do fato de que ele equivale a maximizar a concentração de energia do sinal de saída em uma única amostra. Detalhes podem ser encontrados em [3-7]. A formulação usada refere-se a gabaritos de sinais em uma dimensão, ou seja, para seu uso em aplicações envolvendo imagens, é necessário algum procedimento de redução de posto no sinal da imagem.

Neste artigo, propomos uma generalização bidimensional para o método em [3]. Na seção 2, descrevemos uma formulação matemática em duas dimensões, incluindo uma função objetivo criada após

expandir a definição do caso em uma dimensão [3]. Nas seções 3 e 4, propomos outras duas funções objetivos para serem usadas na otimização dos filtros discriminativos. Na seção 5, testamos o método proposto em um sistema de detecção de gabaritos. Na seção 6, mostramos os resultados experimentais e, na seção 7, apresentamos as conclusões

## 2. FILTROS DISCRIMINATIVOS EM DUAS DIMENSÕES

Quando se usam filtros discriminativos para a detecção de gabaritos, geralmente se deseja concentrar a energia do sinal de saída em uma de suas amostras. A relação sinal-ruído discriminativa (DSNR), definida em [3], é uma medida que leva em conta não apenas a energia em uma amostra, mas também considera a energia em relação às outras amostras. Para os filtros discriminativos em duas dimensões, definimos uma função objetivo discriminativa (DSNR<sub>2</sub>) como sendo:

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

Os coeficientes  $c_{m,n}$  são obtidos após uma convolução bidimensional entre a janela de entrada do sinal de imagem  $u_{m,n}$  e um operador linear  $\Theta$ , com resposta ao impulso  $\theta_{m,n}$ , que vamos calcular para um determinado gabarito a ser casado. O coeficiente  $c_{i,j}$  é aquele nonde se deseja concentrar a energia do sinal de saída do filtro.

Em nossa abordagem, não fazemos qualquer restrição quanto às dimensões tanto da janela de entrada, quanto do operador  $\Theta$ . Seja  $U$  uma janela  $M_1 \times N_1$  e  $\Theta$  um filtro  $M_2 \times N_2$ . Portanto, a imagem resultante após a convolução entre  $U$  e  $\Theta$  tem dimensões  $M_1 + M_2 - 1 \times N_1 + N_2 - 1$  e pode ser expressa como:

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

É fácil ver que a maximização de  $DSNR_2$  (vide equação 1) é equivalente à minimização da seguinte expressão:

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

Da equação 2, a função  $f(\Theta)$  a ser minimizada torna-se:

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

$f(\Theta)$  é mínimo quando seu gradiente for igual a 0, ou seja:

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

A condição acima recai num sistema não linear de  $M_2 N_2$  equações. Para se obter uma solução numérica, usamos um método baseado em gradientes. A expressão para  $\nabla f(\Theta)$  é

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

Uma boa estimativa inicial para o filtro é importante para este método baseado em gradiente. Sugerimos o uso da parte real da transformada inversa de Fourier do sinal obtido com o quociente, no domínio da frequência, entre a resposta em frequência desejada ( $DSNR_2$  infinito, ou seja,  $c_{m,n} = \delta_{i-m,j-n}$ ) e a transformada de Fourier de  $U$ . Calculamos o desempenho usando quinas em janelas  $7 \times 7$  como gabaritos (vide figura 1). O filtro operador  $\Theta$  também foi escolhido com a dimensão  $7 \times 7$  e as coordenadas da amostra onde se deve concentrar a energia do sinal de saída, ou seja,  $(i,j)$  nas equações (3), (4) e (6), foram escolhidas na amostra central do sinal de saída.

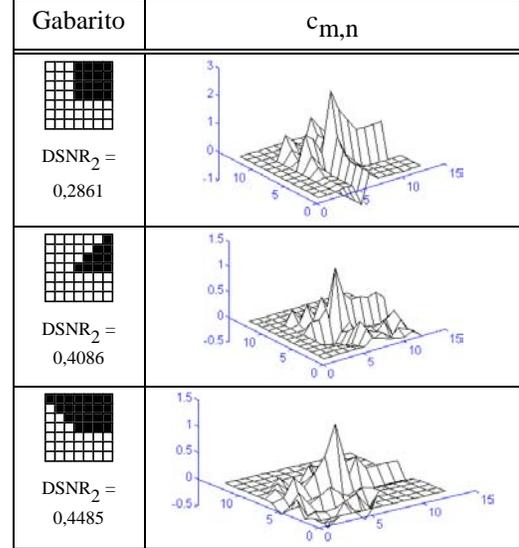


Figura 1: Coeficientes  $c_{m,n}$  obtidos após uma convolução linear entre o gabarito proposto e seu respectivo  $\Theta$  calculado.

### 3. ABORDAGEM ALTERNATIVA

Com a formulação da seção 2, encontra-se o  $\Theta$  que maximiza a  $DSNR_2$  para um dado gabarito. Contudo, ela não evita que um gabarito diferente forneça uma  $DSNR_2$  maior quando filtrado com o  $\Theta$ , o que pode ocasionar falsas detecções, como pode ser observado nos resultados da seção 6.

Um caminho possível para a solução deste problema é considerar a  $DSNR_2$  como uma função de  $\Theta$  e de  $U$  e procurar por um  $\Theta$  que maximize a  $DSNR_2$  quando  $U$  variar. Matematicamente, isto equivale a resolver:

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

Como o sistema de equações acima é de difícil solução analítica, criamos uma função objetivo, que chamamos de **potencial de discriminação** -  $g(\Theta)$ , que é igual à soma dos quadrados dos termos da equação 7, ou seja:

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

Minimizar  $g(\Theta)$  na equação (8) é equivalente e resolver o sistema da equação 7.

A equação (8) foi resolvida numericamente para os gabaritos apresentados na seção 2. A solução inicial foi o filtro gerado pelo método da seção 2. Os resultados obtidos estão apresentados na tabela 1. A aplicação dos filtros obtidos em imagens reais está mostrada na seção 6. Como será visto na seção 6, muitas falsas detecções aconteceram, mais até do que nas simulações utilizando o método da seção 2.

Gabarito	DSNR <sub>2</sub>	g(Θ)
quina de 90 <sup>0</sup>	0,1851	15,55
quina de 45 <sup>0</sup>	0,3721	3,97
quina de 135 <sup>0</sup>	0,0953	31,48

Tabela 1: DSNR<sub>2</sub> e g(Θ) obtidos com a abordagem alternativa.

#### 4. ABORDAGEM MISTA

Como mostrado nas simulações (vide seção 6), a abordagem da seção 3 não oferece o melhor desempenho em um sistema de detecção de gabaritos. Isto indica que um valor alto de DSNR<sub>2</sub> também é importante. As duas abordagens anteriores parecem ser conflitantes porque elas foram desenvolvidas a partir de dois indicadores diferentes: as funções f(Θ) e g(Θ). Um compromisso entre as duas deve então ser usado. Assim, definimos uma terceira função objetivo, chamada de **função de detecção - DF(Θ)**, que combina f(Θ) com g(Θ):

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

onde Z é um limitante superior para DSNR<sub>2</sub> e K é uma constante que pondera f(Θ) e g(Θ).

O treinamento com os gabaritos forneceu curvas interessantes. A figura 2 mostra alguns resultados para os 3 tipos de quinas. A solução inicial usada foi a mesma da seção 2.

A figura 3 mostra o operador e o sinal de saída gerados para quinas de 135<sup>0</sup>.

#### 5. MÉTODO PRÁTICO PARA A DETECÇÃO DE GABARITOS

A aplicação natural para o algoritmo proposto é a detecção de padrões bidimensionais de imagem. Como mostrado na figura 5, a idéia básica é calcular Θ para um determinado gabarito. Então, este operador é usado para filtrar uma imagem que é gerada após uma transformação apropriada de uma janela da imagem original. A janela de saída que fornecer um valor de DSNR<sub>2</sub> superior a um limiar é

considerada morfologicamente casada com o gabarito original (aquele que gerou Θ).

Com o objetivo de não necessitar de computar diferentes Θ para cada gabarito (um para cada rotação de U), o que aumentaria bastante o número de filtros a serem usados, é realizada uma pré-rotação da janela que contém o sinal de imagem. Fazemos esta rotação para alinharmos a janela da imagem com a janela utilizada no cálculo de Θ. O bloco "rotação adequada", da figura 4, representa esta etapa. Mais especificamente, esta rotação é feita da seguinte maneira:

- escolhemos o gabarito (fig.5.1);
- calculamos o centro do gabarito escolhido (fig.5.2);
- tomamos a janela da imagem a ser casada (fig.5.3);
- calculamos o centro da janela da imagem (fig.5.4);
- calculamos o ângulo α, tomando os dois centros e usando o pixel central como origem (fig.5.5);
- realizamos a rotação da janela de α, de forma a tentar sobrepor a janela por cima do gabarito (fig.5.6).

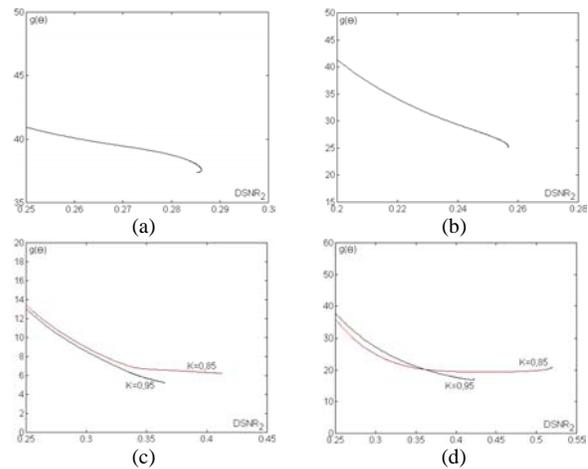


Figura 2: Compromisso entre DSNR<sub>2</sub> e g(Θ), com Z=1, para quinas de (a) 90<sup>0</sup> (K=0,85), (b) 90<sup>0</sup> (K=0,95), (c) 45<sup>0</sup> e (d) 135<sup>0</sup>.

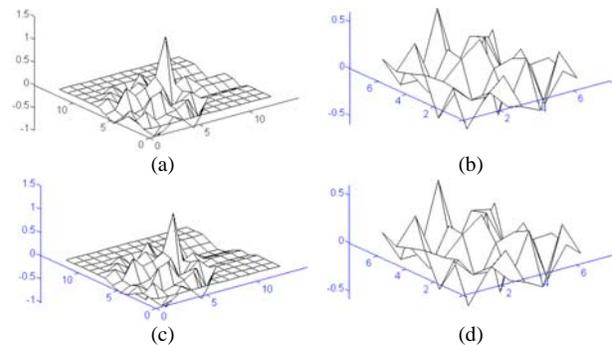


Figura 3: (a,b) C e Θ para a quina de 135<sup>0</sup> com K=0.85 e K=0.95 (c,d).

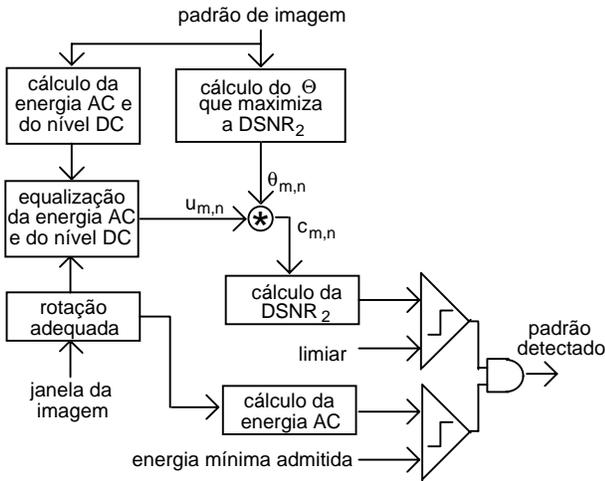


Figure 4: Diagrama em blocos simplificado para um detector de padrões de imagem.

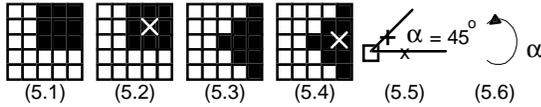


Figura 5: Sequência para a rotação adequada.

Note que, para o cálculo do novo pixel de coordenadas  $(i, j)$  depois da rotação (fig. 5.6), basta copiar o pixel de coordenadas  $(i \cos \alpha + j \sin \alpha, j \cos \alpha - i \sin \alpha)$ .

Um problema possível deste método aparece quando tomamos uma janela igual ao complemento do gabarito (fig. 6). Se aplicarmos o algoritmo da fig. 5, a janela deveria casar-se com o gabarito, pois o cálculo de  $DSNR_2$  é invariante à multiplicação por uma constante ("janela" =  $-1 \times$  "gabarito"). Contudo, na presente implementação, seria feita uma rotação imprópria de  $180^\circ$ , o que ocasionaria um não casamento. Resolvemos este problema ao adicionar  $180^\circ$  a  $\alpha$  se, após a eliminação do nível DC da janela, o número de pixels com valores negativos for menor que o número de pixels com valores positivos.

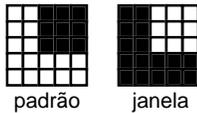


Figura 6: Janela igual ao complemento do gabarito.

## 6. RESULTADOS DE SIMULAÇÕES

Para verificar o desempenho do sistema de detecção proposto, usamos os gabaritos da figura 1 e os filtros obtidos usando as três funções objetivas apresentadas na

seção 2 (método A), na seção 3 (método B) e na seção 4 (método C). A tabela 2 lista os parâmetros de processamento.

Gabarito	$DSNR_2$ (método A)	$DSNR_2$ (método B)	Limiar	Excursão Mínima (níveis de cinza)	
quina de $90^\circ$	0,2850	0,1851	0,16	11,7	
quina de $45^\circ$	0,4163	0,3721	0,09	11,7	
quina de $135^\circ$	0,4493	0,0953	0,09	11,7	
Gabarito	$DSNR_2$ (método C)	K	Z	Limiar	Excursão Mínima (níveis de cinza)
quina de $90^\circ$	0,2470	0,95	1,0	0,16	11,7
quina de $45^\circ$	0,3611	0,95	1,0	0,09	11,7
quina de $135^\circ$	0,5196	0,85	1,0	0,09	11,7
quina de $135^\circ$	0,4219	0,95	1,0	0,09	11,7

Tabela 2: Tabela de parâmetros de processamento.

Observando a figura 7, podemos ver que o algoritmo é muito eficiente na detecção de gabaritos de quinas num ambiente com duas texturas apenas, como na chaminé da casa e nas fronteiras do céu. Para a quina de  $90^\circ$ , os métodos B e C forneceram os mesmos resultados, cancelando uma falsa detecção do método A. Quanto às quinas de  $45^\circ$  e de  $135^\circ$ , notamos que o método A detectou um número elevado de quinas falsas. Para a quina de  $45^\circ$ , ambos os métodos B e C trabalharam melhor do que o método A. Por outro lado, o método B multiplicou o número de quinas falsas de  $135^\circ$ , enquanto que o método C trabalhou bem.

Podemos comparar nossos resultados com os obtidos por Nandy e Ben-Arie [4], mostrados na figura 7-(1). Em ambos os casos, os resultados foram muito satisfatórios. Em [4], as janelas passaram por transformadas KL antes da aplicação do método EXM original. Esta transformação reduziu a dimensão da imagem ( $2 \rightarrow 1$ ) de forma a poder-se aplicar a formulação para uma dimensão.

## 7. CONCLUSÕES

Neste artigo, primeiramente definimos a extensão bidimensional de relação sinal-ruído discriminativa, definida em [3]. Propusemos então três funções objetivas para o projeto de filtros discriminativos. A primeira delas (método A) é baseada na maximização de  $DSNR_2$ . A segunda (método B) busca o filtro que forneça o maior valor de  $DSNR_2$  para um dado gabarito. O método C usa, no cálculo da função objetivo, uma combinação dos métodos A e B. Uma aplicação em potencial para o filtro proposto é a detecção de padrões de imagem. Fizemos

simulações com os filtros calculados em imagens reais e obtivemos melhores resultados discriminativos para o método C.

Os resultados foram bastante satisfatórios, tão bons quanto os outros apresentados na literatura. Contudo, o método proposto pode ser usado para detectar qualquer padrão bidimensional, mesmo para aqueles que não possam ser adequadamente transformados para o uso do EXM em uma dimensão [3-7].

## 8. REFERÊNCIAS

- [1] Ji, Q. e 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. e Pratt, W.K., "Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors", Proc. of IEEE, 67(5)753-763,1979.
- [3] Ben-Arie, J. e 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<sup>o</sup>.1, Fev.1993.
- [4] Nandy, D. e Ben-Arie, J. , "EXM Eigen Templates for Detecting and Classifying Arbitrary Junctions", IEEE International Conference on Image Processing, 1998.
- [5] Rao, K.R. e Ben-Arie, J. , "Multiple Template Matching Using the Expansion Filter", IEEE Transactions on Circuits and Systems for Video Technology, Vol.4, N<sup>o</sup>.5, out.1994.
- [6] Rao, K. R. e Ben-Arie, J., "Optimal Edge Detection Using Expansion Matching and Restoration", IEEE Transactions on Pattern Analysis and machine Intelligence, Vol.16, N<sup>o</sup>.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<sup>o</sup>.11, dez.1996.h

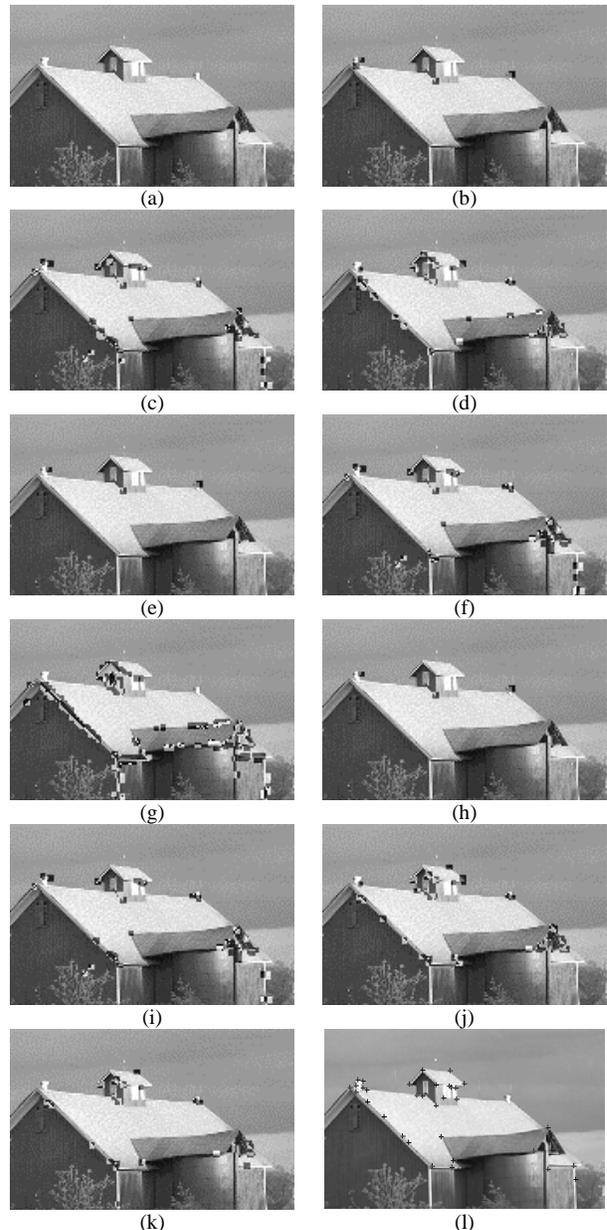


Figura 7: (a) Imagem original. (b)(c)(d) Quinas de  $90^\circ$ ,  $45^\circ$  e  $135^\circ$  detectadas (método A). (e)(f)(g) Quinas detectadas (método B). (h)(i)(j)(k) Quinas detectadas pelo método C. (vide tabela 2) (l) Resultados de [4].