

**RECONHECIMENTO E LOCALIZAÇÃO DE PADRÕES EM
IMAGENS UTILIZANDO REDES NEURAIIS ARTIFICIAIS COMO
ESTIMADORES DE CORRELAÇÃO ESPECTRAL**

**DISSERTAÇÃO SUBMETIDA À PONTIFÍCIA UNIVERSIDADE CATÓLICA DO
RIO GRANDE DO SUL PARA OBTENÇÃO DO GRAU DE
MESTRE EM ENGENHARIA**

FERNANDO CÉSAR COMPARI DE CASTRO

PORTO ALEGRE, DEZEMBRO DE 1995

**RECONHECIMENTO E LOCALIZAÇÃO DE PADRÕES EM IMAGENS
UTILIZANDO REDES NEURAS ARTIFICIAIS COMO ESTIMADORES DE
CORRELAÇÃO ESPECTRAL**

CANDIDATO: FERNANDO CÉSAR COMPARSI DE CASTRO

Esta dissertação foi julgada para obtenção do título MESTRE EM ENGENHARIA e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Pontifícia Universidade Católica do Rio Grande do Sul

Prof. José Nelson Amaral, Ph.D.
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

BANCA EXAMINADORA:

Prof. Paulo Roberto Girardello Franco, Ph.D. - Presidente

Prof. José Nelson Amaral, Ph.D.

Prof. Roberto de Alencar Lotufo, Ph.D.

A todos os que, direta ou indiretamente, contribuíram para a conclusão a bom termo deste trabalho, meu sincero reconhecimento.

SUMÁRIO

Lista de Figuras	VII
Lista de Tabelas	X
Simbologia	XI
Resumo	XV
Abstract	XVI
1. Introdução	1
2. Definições Gerais	4
2.1. Imagem	5
2.2 Transformada Rápida de Fourier Direta	5
2.3 Transformada Rápida de Fourier Inversa	6
2.4 Transformada Rápida de Fourier Bidimensional	6
2.5 Energia e Densidade Espectral de Energia	6
3. Relações Espaço-Frequência	7
3.1 Derivada do Espectro	7
3.2 Correlação Circular e FFT	7
3.3 Espectro da Imagem Transladada	9
3.4 Separabilidade da FFT Bidimensionalem FFTs Unidimensionais	9
3.5 Derivada do Espectro Bidimensional na Direção Espectral Vertical	10
3.6 Derivada do espectro Bidimensional na direção Espectral Horizontal	11
3.7 Espectro da Imagem Rotacionada 180°	12
3.8 Derivada do Espectro na Direção Espectral Vertical da Imagem Rotacionada 180°	12
3.9 Derivada do Espectro na Direção Espectral Horizontal da Imagem Rotacionada 180°	13
4. Localização das Coordenadas do Padrão pelo Espectro de Fase	14

5. Redes Neurais Artificiais	17
5.1 Modelo de um Neurônio Artificial	18
5.2 Redes Neurais em Multi-Camadas Progressivas	19
5.3 Algoritmo <i>Backpropagation</i>	21
5.3.1 Definição do Algoritmo <i>Backpropagation</i>	22
5.3.2 Os dois Passos Computacionais do <i>Backpropagation</i>	26
5.3.3 A Derivada da Função de Ativação	27
5.3.4 Razão de Aprendizado e Momento	28
5.3.5 Sumário Operacional do <i>Backpropagation</i>	28
6. Processo de Detecção do Padrão na Imagem Alvo	30
6.1 Armazenamento das Características do Padrão	30
6.2 A Detecção e Localização do Padrão na Imagem Alvo	35
6.3 Sumário do Processo de Detecção e Localização	39
7. Simulações	41
7.1 Identificação do Caracter “A” <i>Times New Roman</i> Escalonado, Rotacionado e Transladado	41
7.2 Identificação do Caracter “V” <i>Times New Roman</i> Comparado ao Padrão “A”	44
7.3 Identificação do Caracter “C” <i>Times New Roman</i> Escalonado, Rotacionado e Transladado	47
7.4 Identificação do Caracter “O” <i>Times New Roman</i> Comparado ao Padrão “C”	50
7.5 Identificação de Composição de Formas Geométricas Escalonada, Rotacionada e Transladada	52
7.6 Identificação de Composição de Formas Diferenciadas por Detalhe Geométrico em Relação ao Padrão	54
7.7 Identificação do Caracter “A” <i>Times New Roman</i> sob Ruído Impulsivo Comparado a Ele Próprio	56
7.8 Identificação do Caracter “V” <i>Times New Roman</i> sob Ruído Impulsivo Comparado ao Padrão “A”	58
8. Conclusão	59

9. Apêndices	61
Apêndice A - Rotação Simultânea nos Domínios Espaço e Freqüência	61
Apêndice B - Ortogonalidade	62
Apêndice C - Energia e Densidade Espectral de Energia	63
Apêndice D - Diagrama do Processo de Localização e Detecção	64
10 . Referências.....	66

LISTA DE FIGURAS

1: Correlator espectral ótico.....	2
2: Matriz de Representação da Letra T.....	5
3: Modelo de um neurônio artificial.....	18
4: FMLP com arquitetura 10-4-2.....	20
5: Grafo de fluxo de sinal no neurônio j	23
6: Grafo de fluxo de sinal mostrando os detalhes do neurônio de saída k conectado ao neurônio escondido j	25
7: Letra “A” e módulo de seu espectro.....	31
8: Letra “C” e módulo de seu espectro.....	31
9: Estratégia geral de treino do FMLP.....	33
10: Estratégia de treino do FMLP. Curva de aprendizado para $500 < \text{época} < 2000$	34
11: Estratégia de treino do FMLP. Curva de aprendizado para $2000 < \text{época} < 8000$	34
12: Imagem padrão $p(x,y)$ - caracter “A” <i>Times New Roman</i> 128x128 pixels.....	41
13: Imagem alvo $a(x,y)$ - “A” da Figura 12 escalonado por 0.4 rotacionado de 60° e transladado de (40,30) pixels.....	41
14: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	42
15: Curva $\tilde{\mu}_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (75).....	42
16: Curva $\tilde{\mu}_{p_r}(\Psi)$ do <i>ensemble</i> $\mathbf{P}_T(\Psi)$ obtida através de (75).....	42
17: $R_{AT}(\theta)$ obtida de (76). O máximo de $R_{AT}(\theta)$ estima o ângulo de rotação θ_0	42
18: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	43
19: Módulo $\log+I$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 10%....	43
20: Imagem alvo $a(x,y)$ - caracter “V” <i>Times New Roman</i> com fator de escalonamento 0.4 rotacionado de 60° e transladado de (40,30) pixels.....	44

21: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	44
22: Curva $\tilde{\mu}_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (75).....	45
23: Curva $\tilde{\mu}_p(\Psi)$ do <i>ensemble</i> $\mathbf{P}_T(\Psi)$ obtida através de (75)	45
24: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0	45
25: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	45
26: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW= 10.6%..	46
27: Imagem padrão $p(x,y)$ - caracter “C” <i>Times New Roman</i> 128x128 pixels.....	47
28: Imagem alvo $a(x,y)$ - caracter “C” da Figura 27 com fator de escalonamento 0.6 rotacionado de -45° e transladado de (14,21) pixels.....	47
29: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	48
30: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0	48
31: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	48
32: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 9%.....	49
33: Imagem alvo $a(x,y)$ - caracter “O” <i>Times New Roman</i> com fator de escalonamento 0.6 rotacionado de -45° e transladado de (14,21) pixels.....	50
34: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	50
35: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0	50
36: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	50
37: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 8%.....	51
38: Imagem padrão $p(x,y)$ 128x128 pixels.....	52
39: Imagem alvo $a(x,y)$ - imagem padrão $p(x,y)$ da Figura 38 com fator de escalonamento 0.5 rotacionada de 90° e transladada de (21,14) pixels.....	52
40: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	52

41: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0	52
42: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	53
43: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -12dB BW =10%...	53
44: Imagem alvo $a(x,y)$ - imagem padrão $p(x,y)$ da Figura 38 com fator de escalonamento 0.5 rotacionada de 90° , transladada de (21,14) pixels e acrescida de detalhe na parte superior.....	54
45: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	54
46: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0	55
47: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	55
48: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -12dB BW =10%..	55
49: Imagem alvo $a(x,y)$ - caracter “A” da Figura 12 com fator de escalonamento 0.4 rotacionado de 60° , transladado de (40,30) pixels e com 1% da totalidade dos pixels corrompido por ruído impulsivo.....	57
50: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	57
51: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0	57
52: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	57
53: Imagem alvo $a(x,y)$ - caracter “V” <i>Times New Roman</i> com fator de escalonamento 0.4 rotacionado de 60° , transladado de (40,30) pixels e com 1% da totalidade dos pixels corrompido por ruído impulsivo.....	58
54: Curva $\mu_A(\Psi)$ do <i>ensemble</i> $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.....	58
55: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima θ_0	59
56: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$	59
B.1: Exponencial complexa de período $P=6$	63
D.1: Processo de detecção.....	65

LISTA DE TABELAS

1: Função mapeamento Valor p de pixel \rightarrow Atributo $M(p)$	4
2: Resultado do processo de detecção e localização - equações (77) a (82)	44
3: Resultado do processo de detecção e localização - equações (77) a (82)	46
4: Resultado do processo de detecção e localização - equações (77) a (82)	49
5: Resultado do processo de detecção e localização - equações (77) a (82)	51
6: Resultado do processo de detecção e localização - equações (77) a (82)	54
7: Resultado do processo de detecção e localização - equações (77) a (82)	56
8: Resultado do processo de detecção e localização - equações (77) a (82) Filtro com BW=10%	57
9: Resultado do processo de detecção e localização - equações (77) a (82) Filtro com BW=11.84% ($10\% \times 0.442 / 0.373 \Rightarrow 0.442 = \alpha_0$ item 7.7 , $0.373 = \alpha_0$ item 7.8)	59

SIMBOLOGIA

$a(x,y)$ - função no domínio espaço (imagem)

$A(u,v)$ - função no domínio frequência (Transformada de Fourier em 2D)

$A^*(u,v)$ - conjugado de $A(u,v)$

$\mathbf{A}(\Psi)$ - ensemble de saída do conjunto de RNAs

$A(\omega, \Psi)$ - espectro $A(u,v)$ em coordenadas polares

$|A(\omega, \Psi)|$ - módulo de $A(\omega, \Psi)$

$d_j(n)$ - resposta desejada para o neurônio j

E - energia total da imagem

$E(\Psi)$ - energia ao longo da direção espectral Ψ

$e_j(n)$ - sinal de erro na saída do neurônio j , para a iteração n

F_i - número total de entradas do neurônio

FFT - Transformada Rápida de Fourier

$F(u)$ - Transformada Rápida de Fourier de $f(x)$, espectro de $f(x)$

$|F(u,v)|^2$ - Densidade Espectral de Energia de uma imagem

IFFT - Transformada Rápida de Fourier Inversa

FMLP - *Feedforward Multilayer Perceptron*

MSE - erro médio quadrático

$M(p)$ - função mapeamento do atributo ótico

$o_k(n)$ - k -ésimo elemento do vetor de saída da RNA

$O_j(\Psi)$ - nível de ativação do neurônio de saída j

$R_{fg}(d)$ - correlação circular entre $f(x)$ e $g(x)$

$R_{AT}(\theta)$ - Correlação cruzada entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_r}(\Psi)$

R_j - rede neural de índice j

RNA - Rede Neural Artificial

$S_m(\omega)$ - vetor m com os ω valores de $|P(\omega, \Psi_j)|$

T_j - conjunto de treino da RNA j

$v_j(n)$ - nível de atividade interna líquida do neurônio j , na iteração n

v_k - nível de ativação do neurônio

$V\hat{x}_0$ - variância normalizada da estimativa de translação na direção x

$V\hat{y}_0$ - variância normalizada da estimativa de translação na direção y

W_j - matriz de pesos sinápticos

$w_{ji}(n)$ - peso sináptico conectando a saída do neurônio i à entrada do neurônio j , para a iteração n

w_{kp} - peso sináptico p do neurônio k

\hat{x}_0 - estimativa de translação na direção x

$x_i(n)$ - i -ésimo elemento do vetor de entrada

x_p - sinal p de entrada da rede neural

$x_0(u,v)$ - matriz translação relativa entre $a(x,y)$ e $p_T(x,y)$ na direção x com indexação de linhas e colunas no domínio frequência

\hat{y}_0 - estimativa de translação na direção I_y

$y_0(u,v)$ - matriz translação relativa entre $a(x,y)$ e $p_T(x,y)$ na direção y com indexação de linhas e colunas no domínio frequência

$y_j(n)$ - sinal resultante na saída do neurônio j , para a iteração n

$x_0(\omega, \Psi)$ - matriz translação $x_0(u,v)$ com indexação de linhas e colunas expressa em coordenadas polares

y_k - sinal de saída do neurônio k

$y_0(\omega, \Psi)$ - matriz translação $y_0(u,v)$ com indexação de linhas e colunas expressa em coordenadas polares

$\angle Z$ - ângulo do número complexo Z

α - constante de momento

α_0 - grau de escalonamento do padrão encontrado na imagem alvo

α_a - valor máximo de $\mu_A(\Psi)$

$\delta_j(n)$ - gradiente local do neurônio j

$\Delta w_{ji}(n)$ - correção aplicada ao peso sináptico $w_{ji}(n)$

$\varepsilon(n)$ - soma dos erros quadráticos no instante n

ε_{av} - erro médio quadrático

η - razão de aprendizado

$\varphi(\cdot)$ - função de ativação

$\varphi'(\cdot)$ - derivada da função de ativação

$\varphi_j(\cdot)$ - função ativação que descreve a relação funcional entrada-saída da não-linearidade associada com o neurônio j

$\mu_A(\psi)$ - curva de média do *ensemble* $\mathbf{A}(\psi)$

$\tilde{\mu}(\Psi)$ - variação da média $\mu(\psi)$

θ_0 - ângulo de rotação entre padrão original e padrão encontrado imagem alvo

θ_k - threshold do neurônio k

ψ - ângulo de varredura do espectro

Ψ_j - direção j tal que $E(\Psi_j)$ é máximo local de $E(\Psi)$

RESUMO

Redes Neurais Artificiais (RNA) são usadas para detectar a existência de um padrão geométrico em uma imagem alvo. O método estabelecido permite não apenas detectar a existência do padrão na imagem alvo, mas também a sua localização, o seu grau de escalonamento e o seu ângulo de rotação. Esta detecção é feita através da estimação da correlação espectral entre o padrão e a imagem alvo. As imagens no domínio espaço são levadas ao domínio espectral através da *Fast Fourier Transform* (FFT) em duas dimensões. No domínio frequência os padrões geométricos podem ser caracterizados pela representação da imagem em um reduzido número de direções espectrais onde a densidade de energia é máxima. A utilização de RNAs treinadas com características espectrais do padrão apenas ao longo destas direções reduz a complexidade computacional do método proposto. Esta técnica tem a vantagem de que a detecção do padrão é invariante à rotação, translação e escalonamento. Os resultados obtidos indicam que o método de detecção proposto é sensível mesmo quando compara-se formas geométricas diferenciadas apenas por detalhes.

1. Introdução

O problema do reconhecimento de um padrão isolado em uma imagem alvo bidimensional consiste em identificar a existência do padrão na imagem, seu grau de escalonamento, ângulo de rotação e posição [1]. Dentre os algoritmos utilizados para o reconhecimento de padrões encontram-se o Método dos Descritores de Fourier [2], o Método da Extração de Features [3], o Método dos Momentos da Região Limítrofe [4] e as Redes Neurais Artificiais [5].

Na área médica, reconhecimento de padrão é utilizado para a identificação de anomalias fisiológicas em imagens radiográficas [6]. Na indústria de defesa, o reconhecimento de padrão é um importante instrumento na localização de alvos em imagens de satélite e aerofotogrametria [7]. Na indústria metal-mecânica, imagens ultrassonográficas são usadas para detectar padrões de falha em peças submetidas a esforços.

Algoritmos para reconhecimento de padrões podem operar tanto no domínio espaço quanto no domínio frequência. Os algoritmos que operam no domínio espaço baseiam-se no processamento das coordenadas de posição (x,y) da imagem alvo. A operação destes algoritmos é dificultada pelas variações na posição, no ângulo de rotação e no grau de escalonamento da imagem padrão na imagem alvo. Os algoritmos que operam no domínio frequência utilizam o espectro da imagem obtido através da Transformada Rápida de Fourier (*Fast Fourier Transform* - FFT). Estes algoritmos usufruem do fato de que a informação de posição do padrão está contida no espectro de fase e a informação de formato está contida no espectro de módulo da imagem alvo.

Lugt foi um dos pioneiros no uso de filtragem espacial complexa no domínio frequência para o reconhecimento de padrões [8]. Utilizando um filme fotográfico exposto ao Plano de Fourier (Plano Frequência) de um sistema ótico, Lugt grava no filme o espectro conjugado da imagem padrão na entrada do sistema. O conjugado do espectro é obtido através de um interferômetro, já que é preciso operar no espectro de fase da imagem para realizar óticamente tal operação. O filme assim exposto passa a constituir um padrão espectral da imagem.

No método de Lugt, uma imagem alvo $f(x,y)$ contendo a forma geométrica padrão é submetida ao sistema ótico mostrado na Figura 1. Mediante o giro de 180° do

filme no Plano de Fourier $F(u,v)$, obtém-se na saída do sistema um máximo de intensidade luminosa na posição em que encontra-se o padrão na imagem alvo quando o ângulo de rotação do filme coincide com o do padrão encontrado.

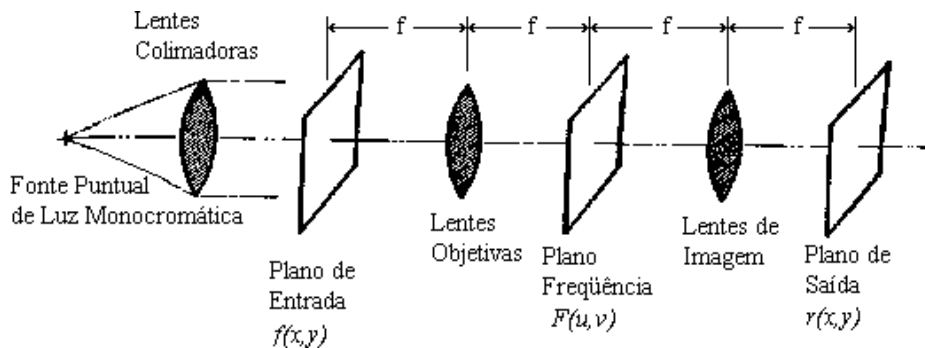


Figura 1: Correlador espectral óptico [8].

O problema do método de Lugt, além da dificuldade intrínseca de implementação devido ao sistema ser óptico, é sua sensibilidade ao grau de escalonamento. Se o padrão procurado não tiver o mesmo grau de escalonamento da imagem padrão utilizada na gravação do filme, o método falhará.

Tradicionalmente Redes Neurais Artificiais (RNA) têm sido usadas para o reconhecimento de padrões no domínio espaço [9]. Propomos aqui um novo método para a solução do problema de reconhecimento de padrões. Este método consiste em selecionar as m direções de maior energia espectral da imagem padrão e utilizar uma RNA para aprender as características do espectro de módulo em cada uma destas direções. Para o treinamento das m RNAs, imagens do padrão com diferentes graus de escalonamento são geradas, a FFT destas imagens é computada e o valor do espectro de módulo na direção j é utilizado para treinar a j -ésima RNA. Note que a RNA aqui desempenha função semelhante à do filme fotográfico no método de Lugt, armazenando informações espectrais do padrão.

A vantagem deste método de reconhecimento é a imunidade a variações de posição do padrão no alvo. Esta imunidade é obtida porque o espectro de módulo da representação da imagem alvo no domínio frequência é invariante à posição do padrão. A única informação contida neste espectro refere-se à forma do padrão. Como o treinamento das RNAs é feito antecipadamente, a única complexidade adicional no momento do reconhecimento do padrão é a transformação da imagem do domínio espaço para o domínio frequência através da FFT. A utilização apenas das informações

contidas nas m direções de maior energia espectral reduz a complexidade computacional do método proposto. Caso fosse utilizada a totalidade das informações espectrais no domínio frequência o método aqui sugerido poderia ser tão complexo quanto os métodos tradicionais que utilizam as informações no domínio espaço. Ainda, como as RNAs são treinadas com a variação do grau de escalonamento do padrão, ao final do processo de detecção é obtido o escalonamento do padrão encontrado no alvo. Semelhantemente ao método de Lugt, as entradas de cada RNA varrem circularmente por um ângulo $0^\circ < \psi < 180^\circ$ o espectro da imagem alvo de forma a se obter o ângulo de rotação do padrão encontrado.

O critério de existência do padrão na imagem alvo é estabelecido pela relação entre seus espectros nas direções de maior energia. Considera-se o conjunto de m saídas das RNAs submetidas aos espectros das imagens padrão e alvo, respectivamente, como dois *ensembles* $\mathbf{P}(\psi)$ e $\mathbf{A}(\psi)$ de m funções-membro cujo domínio é o ângulo ψ .

O máximo da curva de correlação cruzada entre a variação das curvas de média $\mu_P(\psi)$ e $\mu_A(\psi)$ de cada *ensemble* determina o ângulo θ_0 do qual encontra-se rotacionado em relação ao padrão original o padrão supostamente encontrado na imagem alvo; o valor de $\mu_A(\theta_0)$ determina o grau de escalonamento α_0 do mesmo.

A seguir, rotacionando a imagem padrão de θ_0 e escalonando a mesma por α_0 , obtém-se $p_0(x,y)$. Pela relação entre as derivadas dos espectros de $p_0(x,y)$ e da imagem alvo ao longo das direções espectrais de maior energia, estabelece-se um valor numérico de variância do conjunto resultante no domínio frequência, o qual indica a probabilidade da forma geométrica na imagem alvo não ser o padrão. Pela proporcionalidade entre estas derivadas, tomadas em direções ortogonais, obtém-se a translação da imagem alvo em relação ao padrão.

2. Definições Gerais

Ao longo deste trabalho, vários conceitos e operações matemáticas são utilizados. No sentido de dar embasamento a demonstrações e desenvolvimentos futuros, este item dedica-se a definir as operações e conceitos gerais a serem utilizados.

2.1 Imagem

Seja uma imagem bidimensional contínua com largura L e altura H . Seja um conjunto de N^2 células retangulares de dimensões individuais $L/N \times H/N$ obtido da subdivisão da imagem real em N partições por dimensão. Sejam $x = 0, 1, \dots, N-1$ e $y = 0, 1, \dots, N-1$ os índices associados às células respectivamente nas dimensões L e H .

Uma célula é caracterizada por um conjunto de atributos óticos. Um exemplo de atributos possíveis é mostrado na Tabela 1. Seja uma matriz de representação espacial f de dimensão $N \times N$, onde o valor numérico assumido pelo elemento $f(x_0, y_0)$ representa o mapeamento de um atributo ótico predominante internamente à célula retangular de índice (x_0, y_0) de uma imagem contínua sobre o elemento da matriz. Denomina-se o elemento $f(x_0, y_0)$ da matriz de *pixel*, o seu valor numérico é dito ser o “valor do pixel”, e os índices (x_0, y_0) são chamadas de “coordenadas do pixel”.

Por exemplo, a função mapeamento $M(p)$ representada na Tabela 1 e a matriz f mostrada na Figura 2 caracterizam a letra “T” com a travessa horizontal em cor branca, a travessa vertical em cor azul de baixa intensidade luminosa, desenhada sobre fundo preto.

TABELA 1: Função mapeamento Valor p de pixel \rightarrow Atributo $M(p)$	
p	Atributo ótico - $M(p)$
0	COR PRETA - INTENS. LUMINOSA NULA
10	COR VERDE C/ BAIXA INTENS. LUMINOSA
130	COR VERDE C/ ALTA INTENS. LUMINOSA
159	COR AZUL C/ BAIXA INTENS. LUMINOSA
255	COR BRANCA - MÁX. INTENS. LUMINOSA

$$f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 255 & 255 & 255 & 255 & 255 & 0 & 0 \\ 0 & 0 & 0 & 159 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 159 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 159 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 159 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 159 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figura 2: Matriz de Representação da Letra T.

Quanto maior o número de partições N maior será a resolução da representação na matriz f . O número de atributos óticos representáveis é proporcional ao número de bits usado para armazenar um pixel. Neste sentido, f é uma representação discreta da imagem contínua, tanto quanto às coordenadas espaciais (x,y) como quanto ao atributo ótico mapeado através da função de mapeamento $M(p)$.

No contexto de reconhecimento de padrões, a função de mapeamento $M(p)$ pode ser construída de forma que os valores numéricos mais elevados sejam atribuídos a pixels que contenham informação de atributos pertencentes a regiões mais significativas para o reconhecimento da forma geométrica do padrão.

Embora $f(x, y)$ seja uma representação discreta da imagem quanto ao atributo ótico mapeado, para fins computacionais admitiremos que ela representa a imagem com valores numéricos discretos através de $M(p)$, mas a função $f(x, y)$ em si é contínua podendo assumir qualquer valor intermediário do domínio de $M(p)$.

Doravante nos referiremos à matriz $f(x, y)$ de dimensão $N \times N$, como sendo a própria imagem por ela representada.

2.2 Transformada Rápida de Fourier Direta

O algoritmo de detecção e localização a ser apresentado neste trabalho baseia-se em relações espectrais. A passagem do domínio espaço para o domínio espectral é feita através da FFT em duas dimensões. Esta passagem tem como unidade operacional básica a FFT unidimensional. Sendo assim, torna-se necessário definir esta operação.

Seja uma função contínua $f(x)$, sendo x uma variável que assume N valores discretos. A Transformada Rápida de Fourier (*Fast Fourier Transform*) de $f(x)$, $FFT\{f(x)\}$, pode ser definida pela operação [10]

$$F(u) = FFT\{f(x)\} = \sum_{x=0}^{N-1} f(x) e^{-j\frac{2\pi}{N}(ux)} \quad (1)$$

onde N é uma potência inteira de 2. $F(u)$ é dito ser o espectro de $f(x)$.

2.3 Transformada Rápida de Fourier Inversa

Seja $F(u) = FFT\{f(x)\}$. A Transformada Rápida de Fourier Inversa (*Inverse Fast Fourier Transform*) de $F(u)$, $IFFT\{F(u)\}$, pode ser definida pela operação [10]

$$f(x) = IFFT\{F(u)\} = \frac{1}{N} \sum_{u=0}^{N-1} F(u) e^{j\frac{2\pi}{N}(ux)} \quad (2)$$

2.4 Transformada Rápida de Fourier Direta Bidimensional

Seja uma função contínua $f(x,y)$ sendo x e y variáveis que assumem N valores discretos. A Transformada Rápida de Fourier Bidimensional (*2D Fast Fourier Transform*) de $f(x,y)$, $FFT\{f(x,y)\}$, pode ser definida pela operação [11] [12]

$$F(u,v) = FFT\{f(x,y)\} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (3)$$

onde N é uma potência inteira de 2.

Se $f(x,y)$ representa uma imagem bidimensional de $N \times N$ pixels, $F(u,v)$ é dito ser o espectro bidimensional da imagem $f(x,y)$.

2.5 Energia e Densidade Espectral de Energia

Seja uma função contínua $f(x,y)$ representativa de uma imagem bidimensional, sendo x e y variáveis que assumem N valores discretos. A energia total E da imagem pode ser definida pela operação [13]

$$E = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x, y) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} |F(u, v)|^2 \quad (4)$$

A equação (4) afirma que a energia total E de uma imagem pode ser obtida tanto no domínio espaço (x, y) como no domínio freqüência (u, v) . O termo $|F(u, v)|^2$ é denominado Densidade Espectral de Energia da imagem $f(x, y)$. Ver Apêndice C.

3. Relações Espaço-Freqüência

Vários resultados básicos relacionando sinais no domínio espaço com o domínio freqüência são necessários à técnica de reconhecimento de padrões a ser demonstrada neste trabalho. O objetivo básico destas relações é reduzir a complexidade global do algoritmo de detecção e localização. Este item dedica-se à apresentação formal de cada uma delas e a explicitar o seu uso.

3.1 Derivada do Espectro

Seja $F(u) = FFT\{f(x)\}$. A derivada de $F(u)$ em relação a u é dada por [14]

$$\begin{aligned} \frac{d}{du} F(u) &= \frac{d}{du} \left\{ \sum_{x=0}^{N-1} f(x) e^{-j\frac{2\pi}{N}(ux)} \right\} = \sum_{x=0}^{N-1} f(x) \frac{d}{du} \left\{ e^{-j\frac{2\pi}{N}(ux)} \right\} = \\ &= -j \frac{2\pi}{N} \left\{ \sum_{x=0}^{N-1} x \cdot f(x) e^{-j\frac{2\pi}{N}(ux)} \right\} = -j \frac{2\pi}{N} FFT\{x \cdot f(x)\} \end{aligned} \quad (5)$$

A equação (5) define a derivada do espectro sem necessariamente se valer de operações de derivação no domínio freqüência. Caso fosse utilizada a noção de derivada aplicada ao domínio freqüência, o procedimento envolveria operações de diferença entre números complexos fora do âmbito computacional da FFT, aumentando desnecessariamente seu grau de complexidade.

3.2 Correlação Circular e FFT

Sejam as funções contínuas $f(x)$ e $g(x)$, sendo x uma variável que assume N valores discretos e N uma potência inteira de 2. Seja $F(u) = FFT\{f(x)\}$ e $G(u) = FFT\{g(x)\}$.

Seja a correlação circular entre $f(x)$ e $g(x)$ definida por [15]

$$R_{fg}(d) = \sum_{x=0}^{N-1} f(x+d)g(x) \quad (6)$$

A relação entre $R_{fg}(d)$, $FFT\{f(x)\}$ e $FFT\{g(x)\}$ é dada por:

Sejam

$$F(u) = FFT\{f(x)\} = \sum_{x=0}^{N-1} f(x)e^{-j\frac{2\pi}{N}(ux)} \quad (7)$$

e

$$G(u) = FFT\{g(x)\} = \sum_{x=0}^{N-1} g(x)e^{-j\frac{2\pi}{N}(ux)} \quad (8)$$

Multiplicando a equação (7) pelo complexo conjugado da equação (8)

$$F(u)G^*(u) = \sum_{w=0}^{N-1} f(w)e^{-j\frac{2\pi}{N}(uw)} \sum_{x=0}^{N-1} g(x)e^{j\frac{2\pi}{N}(ux)} = \sum_{w=0}^{N-1} \sum_{x=0}^{N-1} f(w)g(x)e^{-j\frac{2\pi}{N}(w-x)u} \quad (9)$$

Aplicando a $IFFT$ em (9):

$$IFFT\{F(u)G^*(u)\} = \frac{1}{N} \sum_{u=0}^{N-1} \left(\sum_{w=0}^{N-1} \sum_{x=0}^{N-1} f(w)g(x)e^{-j\frac{2\pi}{N}(w-x)u} \right) e^{j\frac{2\pi}{N}(u)x} \quad (10)$$

Fazendo $w - x = d$ temos $w = x + d$. Portanto, de (10)

$$IFFT\{F(u)G^*(u)\} = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{d=0}^{N-1} \sum_{x=0}^{N-1} f(x+d)g(x)e^{-j\frac{2\pi}{N}(d)u} e^{j\frac{2\pi}{N}(u)x} \quad (11a)$$

$$IFFT\{F(u)G^*(u)\} = \frac{1}{N} \sum_{d=0}^{N-1} \sum_{x=0}^{N-1} f(x+d)g(x) \sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}du} e^{j\frac{2\pi}{N}xu} \quad (11b)$$

Pelo princípio da ortogonalidade (ver Apêndice B)

$$\sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(du)} e^{j\frac{2\pi}{N}(xu)} = \begin{cases} N & \text{para } x = d \\ 0 & \text{para } x \neq d \end{cases} \quad (12)$$

O que resulta em

$$IFFT\{F(u)G^*(u)\} = IFFT\left\{FFT\{f(x)\}FFT^*\{g(x)\}\right\} = \sum_{x=0}^{N-1} f(x+d)g(x) \quad (13)$$

E portanto

$$R_{fg}(d) = IFFT\left\{FFT\{f(x)\}FFT^*\{g(x)\}\right\} \quad (14)$$

A equação (14) informa o grau de similaridade entre $f(x)$ e $g(x)$ em função do deslocamento relativo d entre ambas. Esta expressão é a forma mais rápida, em termos computacionais, de se estabelecer o grau de semelhança entre duas funções.

3.3 Espectro da Imagem Transladada

Considere-se uma imagem $f(x,y)$ de $N \times N$ pixels. Seja $f_t(x,y)$ a função $f(x,y)$ transladada de (x_t, y_t) no sentido crescente de (x,y) , isto é, $f_t(x,y) = f(x-x_t, y-y_t)$. Seja $F(u,v) = FFT\{f(x,y)\}$ e $F_t(u,v) = FFT\{f_t(x,y)\}$. A relação entre $F(u,v)$ e $F_t(u,v)$ é dada por [16]

$$F_t(u,v) = FFT\{f(x-x_t, y-y_t)\} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x-x_t, y-y_t) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (15)$$

Sejam

$$\begin{aligned} \alpha &= x - x_t \Rightarrow x = \alpha + x_t \\ \beta &= y - y_t \Rightarrow y = \beta + y_t \end{aligned} \quad (16)$$

$$F_t(u,v) = \sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{N-1} f(\alpha, \beta) e^{-j\frac{2\pi}{N}(u(\alpha+x_t)+v(\beta+y_t))} \quad (17a)$$

$$F_t(u,v) = \left[\sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{N-1} f(\alpha, \beta) e^{-j\frac{2\pi}{N}(u\alpha+v\beta)} \right] \left[e^{-j\frac{2\pi}{N}(x_t u + y_t v)} \right] = F(u,v) \left[e^{-j\frac{2\pi}{N}(x_t u + y_t v)} \right] \quad (17b)$$

A equação (17b) define, a partir do espectro da imagem original, o espectro da imagem transladada, dispensando a necessidade de calculá-lo explicitamente através da FFT.

3.4 Separabilidade da FFT Bidimensional em FFTs Unidimensionais

Seja o espectro bidimensional da imagem $f(x,y)$ dado por $F(u,v) = FFT\{f(x,y)\}$. Separa-se a FFT bidimensional em duas FFT's unidimensionais através de [14]

$$F(u, v) = FFT\{f(x, y)\} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (18a)$$

$$F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(vy)} e^{-j\frac{2\pi}{N}(ux)} \quad (18b)$$

Seja

$$\sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(vy)} = FFT_y\{f(x, y)\} = g(x, v) \quad (19)$$

Então

$$F(u, v) = \sum_{x=0}^{N-1} g(x, v) e^{-j\frac{2\pi}{N}(ux)} = FFT_x\{g(x, v)\} = FFT_x\{FFT_y\{f(x, y)\}\} \quad (20)$$

A equação (20) postula que o espectro da imagem bidimensional $f(x, y)$ pode ser obtido através de duas FFTs unidimensionais sucessivas. A nível computacional, primeiro aplica-se a FFT unidimensional sobre os valores definidos ao longo das linhas da matriz $f(x, y)$. Em seguida, transpõe-se a matriz resultante e novamente aplica-se a FFT unidimensional sobre os valores definidos ao longo de suas linhas. A matriz resultante transposta desta última operação constitui o espectro $F(u, v)$.

3.5 Derivada do Espectro Bidimensional na Direção Espectral Vertical

Seja $F(u, v) = FFT\{f(x, y)\}$. A derivada de $F(u, v)$ em relação a u é dada por

$$\frac{d}{du} F(u, v) = \frac{d}{du} \left\{ \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(ux+vy)} \right\} \quad (21a)$$

$$\frac{d}{du} F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(vy)} \frac{d}{du} \left\{ e^{-j\frac{2\pi}{N}(ux)} \right\} \quad (21b)$$

$$\frac{d}{du} F(u, v) = -j \frac{2\pi}{N} \sum_{x=0}^{N-1} x \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(vy)} e^{-j\frac{2\pi}{N}(ux)} \quad (21c)$$

Seja

$$\sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(vy)} = FFT_y\{f(x, y)\} = g(x, v) \quad (22)$$

Então

$$\frac{d}{du} F(u, v) = -j \frac{2\pi}{N} \sum_{x=0}^{N-1} x \cdot g(x, v) e^{-j \frac{2\pi}{N}(ux)} \quad (23a)$$

$$\frac{d}{du} F(u, v) = -j \frac{2\pi}{N} FFT_x \{x \cdot g(x, v)\} \quad (23b)$$

$$\frac{d}{du} F(u, v) = -j \frac{2\pi}{N} FFT_x \{x \cdot FFT_y \{f(x, y)\}\} \quad (23c)$$

A equação (23c) define a derivada do espectro na direção vertical, sem necessariamente se valer de operações de derivação no domínio frequência e a partir do uso de somente FFTs unidimensionais.

3.6 Derivada do Espectro Bidimensional na Direção Espectral Horizontal

Seja $F(u, v) = FFT\{f(x, y)\}$. A derivada de $F(u, v)$ em relação a v é dada por

$$\frac{d}{dv} F(u, v) = \frac{d}{dv} \left\{ \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j \frac{2\pi}{N}(ux+vy)} \right\} \quad (24a)$$

$$\frac{d}{dv} F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j \frac{2\pi}{N}(ux)} \frac{d}{dv} \left\{ e^{-j \frac{2\pi}{N}(vy)} \right\} \quad (24b)$$

$$\frac{d}{dv} F(u, v) = -j \frac{2\pi}{N} \sum_{y=0}^{N-1} y \sum_{x=0}^{N-1} f(x, y) e^{-j \frac{2\pi}{N}(ux)} e^{-j \frac{2\pi}{N}(vy)} \quad (24c)$$

Seja

$$\sum_{x=0}^{N-1} f(x, y) e^{-j \frac{2\pi}{N}(ux)} = FFT_x \{f(x, y)\} = g(u, y) \quad (25)$$

Então

$$\frac{d}{dv} F(u, v) = -j \frac{2\pi}{N} \sum_{y=0}^{N-1} y \cdot g(u, y) e^{-j \frac{2\pi}{N}(vy)} \quad (26a)$$

$$\frac{d}{dv} F(u, v) = -j \frac{2\pi}{N} FFT_y \{y \cdot g(u, y)\} \quad (26b)$$

$$\frac{d}{dv} F(u, v) = -j \frac{2\pi}{N} FFT_y \{y \cdot FFT_x \{f(x, y)\}\} \quad (26c)$$

A equação (26c) define a derivada do espectro na direção horizontal, sem necessariamente se valer de operações de derivação no domínio frequência e a partir do uso de somente FFTs unidimensionais.

3.7 Espectro da Imagem Rotacionada 180°

Considere-se uma imagem $f(x,y)$ com centro geométrico em (x_c, y_c) . Seja $f_r(x,y)$ a imagem $f(x,y)$ rotacionada 180° em torno de seu centro geométrico, isto é, $f_r(x,y)=f(-x+2x_c, -y+2y_c)$. Seja $F(u,v)=FFT\{f(x,y)\}$ e $F_r(u,v)=FFT\{f_r(x,y)\}$. A relação entre $F(u,v)$ e $F_r(u,v)$ é

$$F_r(u,v) = FFT\{f(-x+2x_c, -y+2y_c)\} \quad (27a)$$

$$F_r(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(-x+2x_c, -y+2y_c) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (27b)$$

Sejam

$$\begin{aligned} \alpha &= -x+2x_c \Rightarrow x = -\alpha+2x_c \\ \beta &= -y+2y_c \Rightarrow y = -\beta+2y_c \end{aligned} \quad (28)$$

$$F_r(u,v) = \sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{N-1} f(\alpha, \beta) e^{-j\frac{2\pi}{N}(u(-\alpha+2x_c)+v(-\beta+2y_c))} \quad (29a)$$

$$F_r(u,v) = \left[\sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{N-1} f(\alpha, \beta) e^{j\frac{2\pi}{N}(u\alpha+v\beta)} \right] \left[e^{-j\frac{2\pi}{N}2(x_c u+y_c v)} \right] \quad (29b)$$

$$F_r(u,v) = \left[F^*(u,v) \right] \left[e^{-j\frac{4\pi}{N}(x_c u+y_c v)} \right] \quad (29c)$$

A equação (29c) define, a partir do espectro da imagem original, o espectro da imagem rotacionada de 180°, dispensando a necessidade de calculá-lo explicitamente através da FFT. Este resultado, juntamente com aqueles a serem obtidos nos itens 3.8 e 3.9 serão utilizados adiante no processo de detecção e localização do padrão.

3.8 Derivada do Espectro na Direção Espectral Vertical da Imagem Rotacionada 180°

Seja uma imagem $f(x,y)$ com centro geométrico em (x_c, y_c) . Seja $f_r(x,y)$ a imagem $f(x,y)$ rotacionada 180° em torno de seu centro geométrico, isto é, $f_r(x,y)=f(-x+2x_c, -y+2y_c)$. Seja $F(u,v)=FFT\{f(x,y)\}$ e $F_r(u,v)=FFT\{f_r(x,y)\}$. A relação entre $(d/du)\{F(u,v)\}$ e $(d/du)\{F_r(u,v)\}$ é

De (29)

$$\frac{d}{du}\{F_r(u,v)\} = \frac{d}{du}\left\{\left[F^*(u,v)\right]\left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right]\right\} \quad (30a)$$

$$\frac{d}{du}\{F_r(u,v)\} = \left[\frac{d}{du}\{F(u,v)\}\right]^* \left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right] + \left[F^*(u,v)\right]\left(-j\frac{4\pi}{N}x_c\right)\left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right] \quad (30b)$$

$$\frac{d}{du}\{F_r(u,v)\} = \left\{\left[\frac{d}{du}\{F(u,v)\}\right]^* + \left[F^*(u,v)\right]\left(-j\frac{4\pi}{N}x_c\right)\right\}\left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right] \quad (30c)$$

A equação (30c) define, a partir do espectro da imagem original e de sua derivada, a derivada do espectro na direção vertical da imagem rotacionada de 180°, dispensando a necessidade de calculá-lo explicitamente através da FFT.

3.9 Derivada do Espectro na Direção Espectral Horizontal da Imagem Rotacionada 180°

Seja uma imagem $f(x,y)$ com centro geométrico em (x_c, y_c) . Seja $f_r(x,y)$ a imagem $f(x,y)$ rotacionada 180° em torno de seu centro geométrico, isto é, $f_r(x,y) = f(-x+2x_c, -y+2y_c)$. Seja $F(u,v) = FFT\{f(x,y)\}$ e $F_r(u,v) = FFT\{f_r(x,y)\}$. A relação entre $(d/dv)\{F(u,v)\}$ e $(d/dv)\{F_r(u,v)\}$ é:

De (29):

$$\frac{d}{dv}\{F_r(u,v)\} = \frac{d}{dv}\left\{\left[F^*(u,v)\right]\left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right]\right\} \quad (31a)$$

$$\frac{d}{dv}\{F_r(u,v)\} = \left[\frac{d}{dv}\{F(u,v)\}\right]^* \left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right] + \left[F^*(u,v)\right]\left(-j\frac{4\pi}{N}y_c\right)\left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right] \quad (31b)$$

$$\frac{d}{dv}\{F_r(u,v)\} = \left\{\left[\frac{d}{dv}\{F(u,v)\}\right]^* + \left[F^*(u,v)\right]\left(-j\frac{4\pi}{N}y_c\right)\right\}\left[e^{-j\frac{4\pi}{N}(x_c u + y_c v)}\right] \quad (31c)$$

A equação (31c) define, a partir do espectro da imagem original e de sua derivada, a derivada do espectro na direção horizontal da imagem rotacionada de 180°, dispensando a necessidade de calculá-lo explicitamente através da FFT.

4. Localização das Coordenadas do Padrão pelo Espectro de Fase - Fundamentos Teóricos

Embora o método de detecção ainda não tenha sido exposto neste trabalho, o presente item dedica-se a desenvolver os fundamentos do mesmo, a ser visto por completo adiante. Demonstra-se aqui também a teoria da técnica de localização, assumindo que o padrão exista na imagem alvo e que sejam conhecidos o seu grau de rotação e fator de escalonamento. É oportuno tal desenvolvimento neste instante porque ele é baseado apenas em análise espectral, sendo independente da teoria de Redes Neurais a ser abordada adiante.

O objetivo básico aqui é determinar a translação (x_0, y_0) entre a imagem padrão e a imagem alvo a partir da relação entre as derivadas de seus espectros. É possível obter o mesmo resultado sem a utilização de derivadas, através da solução de um sistema algébrico de equações lineares. No entanto, as operações envolvendo fase espectral, resultantes desta última abordagem, conduzem à falhas no algoritmo de localização pelo fato da fase espectral estar restrita ao intervalo $[-\pi, \pi]$.

Seja $a(x, y)$ a imagem alvo e $p(x, y)$ a imagem padrão. Seja θ_0 o ângulo de rotação do padrão encontrado na imagem alvo com um fator de escalonamento α_0 . Seja $p_T(x, y)$ a imagem padrão $p(x, y)$ rotacionada de θ_0 e escalonada de α_0 . Seja $P_T(u, v) = \text{FFT}\{p_T(x, y)\}$ e $A(u, v) = \text{FFT}\{a(x, y)\}$. A diferença entre $P_T(u, v)$ e $A(u, v)$ é apenas no espectro de fase já que $a(x, y)$ e $p_T(x, y)$ diferem apenas por translação (x_0, y_0) . Assim, da equação (17) tem-se

$$P_T(u, v) = A(u, v) e^{-j \frac{2\pi}{N} (ux_0 + vy_0)} \quad (32)$$

A forma polar da equação (32) é

$$\left| \frac{P_T(u, v)}{A(u, v)} \right| e^{j(\angle\{P_T(u, v)\} - \angle\{A(u, v)\})} = e^{-j \frac{2\pi}{N} (ux_0 + vy_0)} \quad (33)$$

$$\angle\{P_T(u, v)\} - \angle\{A(u, v)\} = -\frac{2\pi}{N} (ux_0 + vy_0) \quad (34)$$

Como para quaisquer números complexos Z_1 e Z_2 , $\angle Z_1 - \angle Z_2 = \angle(Z_1 Z_2^*)$, têm-se

$$\angle \left\{ P_T(u, v) A^*(u, v) \right\} = -\frac{2\pi}{N} (ux_0 + vy_0) \quad (35)$$

Seja

$$P_T(u, v) A^*(u, v) = Z(u, v) + jW(u, v) \quad (36)$$

Tal que

$$\begin{aligned} Z &= \text{Re}\{P_T A^*\} \\ W &= \text{Im}\{P_T A^*\} \end{aligned} \quad (37)$$

Então

$$\angle \left\{ P_T(u, v) A^*(u, v) \right\} = \angle \{Z(u, v) + jW(u, v)\} = \arctan \left(\frac{W(u, v)}{Z(u, v)} \right) \quad (38)$$

Logo de (35) e (38)

$$-\frac{2\pi}{N} (ux_0 + vy_0) = \arctan \left(\frac{W(u, v)}{Z(u, v)} \right) \quad (39)$$

Derivando (39) em relação a u

$$\frac{d}{du} \left\{ -\frac{2\pi}{N} (ux_0 + vy_0) \right\} = \frac{d}{du} \left\{ \arctan \left(\frac{W(u, v)}{Z(u, v)} \right) \right\} \quad (40)$$

$$-\frac{2\pi}{N} x_0 = \left[\frac{1}{1 + \left(\frac{W(u, v)}{Z(u, v)} \right)^2} \right] \left[\frac{Z(u, v) \frac{d}{du} W(u, v) - W(u, v) \frac{d}{du} Z(u, v)}{Z^2(u, v)} \right] \quad (41)$$

$$x_0 = \frac{N}{2\pi} \left[\frac{W(u, v) \frac{d}{du} Z(u, v) - Z(u, v) \frac{d}{du} W(u, v)}{Z^2(u, v) + W^2(u, v)} \right] \quad (42)$$

Derivando (37) em relação a u

$$\frac{dZ}{du} = \text{Re} \left\{ \frac{dP_T}{du} A^* + P_T \frac{dA^*}{du} \right\} \quad (43)$$

$$\frac{dW}{du} = \text{Im} \left\{ \frac{dP_T}{du} A^* + P_T \frac{dA^*}{du} \right\} \quad (44)$$

Substituindo (43) e (44) em (42)

$$x_0 = \frac{N}{2\pi} \left[\frac{1}{\operatorname{Re}^2\{P_T A^*\} + \operatorname{Im}^2\{P_T A^*\}} \right] \left\{ \begin{array}{l} \operatorname{Im}\{P_T A^*\} \operatorname{Re}\left\{ \frac{dP_T}{du} A^* + P_T \frac{dA^*}{du} \right\} + \\ - \operatorname{Re}\{P_T A^*\} \operatorname{Im}\left\{ \frac{dP_T}{du} A^* + P_T \frac{dA^*}{du} \right\} \end{array} \right\} \quad (45)$$

Explicitando o domínio (u, v)

$$x_0(u, v) = \frac{N}{2\pi} \left[\frac{1}{\operatorname{Re}^2\{P_T(u, v) A^*(u, v)\} + \operatorname{Im}^2\{P_T(u, v) A^*(u, v)\}} \right] \left\{ \begin{array}{l} \operatorname{Im}\{P_T(u, v) A^*(u, v)\} \operatorname{Re}\left\{ A^*(u, v) \frac{d}{du} P_T(u, v) + P_T(u, v) \left[\frac{d}{du} A(u, v) \right]^* \right\} + \\ - \operatorname{Re}\{P_T(u, v) A^*(u, v)\} \operatorname{Im}\left\{ A^*(u, v) \frac{d}{du} P_T(u, v) + P_T(u, v) \left[\frac{d}{du} A(u, v) \right]^* \right\} \end{array} \right\} \quad (46)$$

Derivando (39) em relação a v :

$$\frac{d}{dv} \left\{ -\frac{2\pi}{N} (ux_0 + vy_0) \right\} = \frac{d}{dv} \left\{ \arctan\left(\frac{W(u, v)}{Z(u, v)} \right) \right\} \quad (47)$$

$$-\frac{2\pi}{N} y_0 = \left[\frac{1}{1 + \left(\frac{W(u, v)}{Z(u, v)} \right)^2} \right] \left[\frac{Z(u, v) \frac{d}{dv} W(u, v) - W(u, v) \frac{d}{dv} Z(u, v)}{Z^2(u, v)} \right] \quad (48)$$

$$y_0 = \frac{N}{2\pi} \left[\frac{W(u, v) \frac{d}{dv} Z(u, v) - Z(u, v) \frac{d}{dv} W(u, v)}{Z^2(u, v) + W^2(u, v)} \right] \quad (49)$$

Derivando (37) em relação a v

$$\frac{dZ}{dv} = \operatorname{Re} \left\{ \frac{dP_T}{dv} A^* + P_T \frac{dA^*}{dv} \right\} \quad (50)$$

$$\frac{dW}{dv} = \operatorname{Im} \left\{ \frac{dP_T}{dv} A^* + P_T \frac{dA^*}{dv} \right\} \quad (51)$$

Substituindo (50) e (51) em (49):

$$y_0 = \frac{N}{2\pi} \left[\frac{1}{\text{Re}^2\{P_T A^*\} + \text{Im}^2\{P_T A^*\}} \right] \left\{ \begin{array}{l} \text{Im}\{P_T A^*\} \text{Re}\left\{ \frac{dP_T}{dv} A^* + P_T \frac{dA^*}{dv} \right\} + \\ -\text{Re}\{P_T A^*\} \text{Im}\left\{ \frac{dP_T}{dv} A^* + P_T \frac{dA^*}{dv} \right\} \end{array} \right\} \quad (52)$$

Explicitando o domínio (u,v)

$$y_0(u,v) = \frac{N}{2\pi} \left[\frac{1}{\text{Re}^2\{P_T(u,v) A^*(u,v)\} + \text{Im}^2\{P_T(u,v) A^*(u,v)\}} \right]^x \left\{ \begin{array}{l} \text{Im}\{P_T(u,v) A^*(u,v)\} \text{Re}\left\{ A^*(u,v) \frac{d}{dv} P_T(u,v) + P_T(u,v) \left[\frac{d}{dv} A(u,v) \right]^* \right\} \\ -\text{Re}\{P_T(u,v) A^*(u,v)\} \text{Im}\left\{ A^*(u,v) \frac{d}{dv} P_T(u,v) + P_T(u,v) \left[\frac{d}{dv} A(u,v) \right]^* \right\} \end{array} \right\} \quad (53)$$

As equações (46) e (53) determinam respectivamente as matrizes $x_0(u,v)$ e $y_0(u,v)$ de tamanho $N \times N$ as quais definem a translação relativa entre $a(x,y)$ e $p_T(x,y)$ nas direções x e y . Se $a(x,y)$ tem a mesma forma geométrica e grau de escalonamento que $p_T(x,y)$ as matrizes $x_0(u,v)$ e $y_0(u,v)$ contêm respectivamente o mesmo valor de translação em todas as suas posições (u,v) , o que, aparentemente, sugere uma redundância de valores desnecessária. No entanto, quando a imagem alvo não contém a imagem padrão os valores diferem entre si, sendo a variância do conjunto de valores uma medida da similaridade quanto à forma entre $a(x,y)$ e $p_T(x,y)$. Esta técnica de detecção será explorada no item 6.2.

5. Redes Neurais Artificiais

Uma Rede Neural Artificial é um processador paralelo massivamente distribuído, constituído por unidades básicas de processamento chamadas neurônios, que têm uma tendência natural para armazenar conhecimento obtido de dados experimentais e fazê-lo disponível para uso [17].

Uma RNA assemelha-se ao cérebro em dois aspectos:

1- O conhecimento é adquirido pela RNA através de um processo de aprendizado.

2 - As transmitâncias das conexões inter-neurônios, conhecidas como pesos sinápticos, às quais está submetido o fluxo de informações através da rede, são utilizadas para armazenar o conhecimento.

O procedimento utilizado para efetuar o processo de aprendizado é chamado Algoritmo de Aprendizado, cuja função é modificar os pesos sinápticos de uma RNA, de forma ordenada, de maneira a atingir um determinado objetivo de projeto. O algoritmo mais usado para tanto é o algoritmo *backpropagation*. A descrição de funcionamento deste algoritmo já foi largamente publicada e encontra-se em vários livros textos [22] [24] [25]. Tendo em vista que o mecanismo utilizado neste algoritmo é fundamental para que se entenda o procedimento descrito no item 6, o algoritmo *backpropagation* é discutido neste item. Esta descrição foi construída a partir da apresentação desenvolvida por Haykin [22].

A teoria de Redes Neurais tem seu análogo mais próximo na Teoria de Filtros Lineares Adaptativos [18], a diferença sendo que uma RNA é um sistema essencialmente não-linear. Esta característica a torna adequada para extrair conhecimento de bases de dados geradas inclusive por fenômenos físicos inerentemente não-lineares.

5.1 Modelo de Um Neurônio Artificial

O Neurônio é a unidade de processamento de informações fundamental na operação de uma RNA. A Figura 3 mostra o modelo de um neurônio.

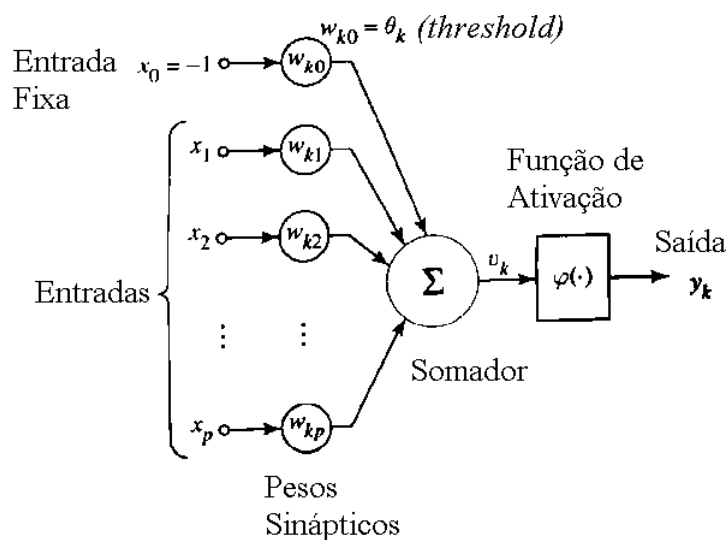


Figura 3: Modelo de um neurônio artificial.

Neste trabalho, o modelo matemático adotado para o k -ésimo neurônio de uma RNA pode ser descrito por (ver Figura 3):

$$v_k = \sum_{j=0}^P w_{kj} x_j \quad (54)$$

$$y_k = \varphi(v_k) \quad (55)$$

$$\varphi(v_k) = a \tanh(b \cdot v_k) \quad (56)$$

$$\begin{aligned} x_0 &= -1 \\ w_{k0} &= \theta_k \end{aligned} \quad (57)$$

Onde x_1, x_2, \dots, x_P são os sinais de entrada, $w_{k0}, w_{k1}, w_{k2}, \dots, w_{kP}$ são os pesos sinápticos do neurônio k , v_k é o nível da ativação do neurônio, θ_k é o chamado *threshold* do neurônio, $\varphi(\cdot)$ é a Função de Ativação e y_k é o sinal de saída do neurônio.

A função de ativação é responsável pela característica não linear de uma RNA. Para sinais de entrada variando no intervalo $[-1,1]$ as constantes a e b assumem os valores $a=1.716$ e $b=0.666$. Utilizou-se a tangente hiperbólica como função de ativação pela sua simetria ímpar, já que o algoritmo de aprendizado que utilizaremos (*backpropagation*), converge mais rapidamente se a função de ativação apresenta esta característica [19].

O *threshold* exerce função análoga à de um *DC-restorer* sobre o nível de ativação do neurônio. O *threshold* dá um grau de liberdade maior ao neurônio visto que v_k pode ser alterado por um parâmetro independente dos sinais de entrada, já que x_0 é fixo.

5.2 Redes Neurais em Multicamadas Progressivas - *Feedforward Multilayer Perceptrons*

Uma RNA do tipo *Feedforward Multilayer Perceptron* (FMLP) é constituída por um conjunto de nós de entrada os quais constituem a camada de entrada, uma ou mais camadas escondidas (*hidden layers*), e uma camada de saída (*output layer*). Com exceção da camada de entrada, todas as outras são constituídas por neurônios e portanto apresentam capacidade computacional.

O número de camadas escondidas e o número de neurônios em cada uma delas é determinado experimentalmente em função do tipo de conjunto dados a ser aprendido.

O número de nós de entrada e de neurônios na saída dependem respectivamente do tipo de sinal de entrada e de saída desejados. A rede é dita ser *feedforward* porque a saída dos neurônios em cada camada sempre se conectam às entradas dos neurônios da camada seguinte. A Figura 4 mostra um FMLP com arquitetura 10-4-2, isto é, 10 nós de entrada, 4 neurônios na única camada escondida e 2 neurônios de saída.

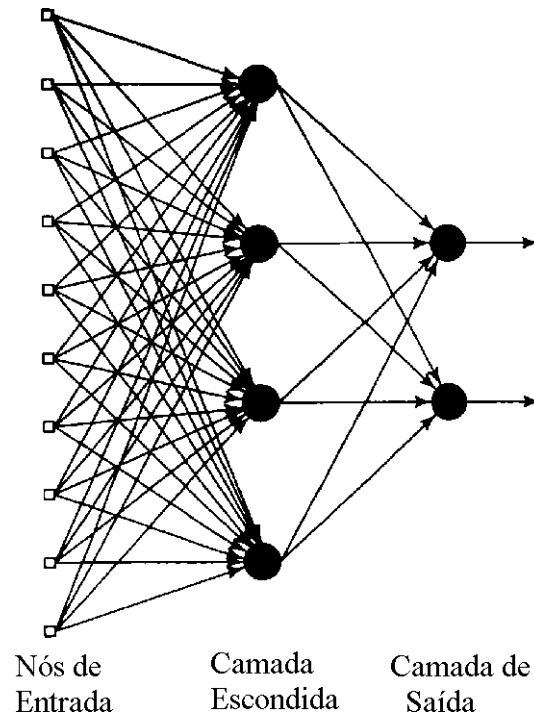


Figura 4: FMLP com arquitetura 10-4-2.

A função das camadas escondidas é influir na relação entrada-saída da rede de uma forma ampla. Uma RNA com uma ou mais camadas escondidas é apta a extrair as estatísticas de alta ordem do processo aleatório do qual a rede está tentando adquirir conhecimento. A RNA adquire uma perspectiva global do processo aleatório apesar de sua conectividade local em virtude do conjunto extra de pesos sinápticos e a dimensão extra de interações neuronais proporcionada pelas camadas escondidas. Tal rede é capaz de aprender o *underlying random process* do conjunto de dados de entrada.

FMLPs têm sido aplicados na solução de diversos e difíceis problemas através do chamado Treino Supervisionado. O algoritmo quase universalmente utilizado para tanto é o denominado *backpropagation*, que baseia-se em aprendizado por erro-correção. Este algoritmo utiliza os princípios do algoritmo *Least Mean Squared* (LMS) desenvolvido por Bernard Widrow [23].

O termo *backpropagation* surgiu após 1985. No entanto, a idéia básica foi primeiramente descrita por Werbos em sua tese de doutorado em 1974 [21]. Em seqüência, foi redescoberto por Rumelhart, Hinton e Williams em 1986 e popularizado através da publicação do livro *Parallel Distributed Processing* de Rumelhart e McClelland em 1986.

É importante citar que não existe prova da convergência do *backpropagation* nem demonstração de critério absoluto de parada de sua operação [22], no entanto, o desenvolvimento do *backpropagation* representa um marco fundamental em redes neurais no sentido que ele resolveu o problema de realizar a propagação reversa do erro em RNAs com múltiplas camadas.

5.3 O algoritmo *backpropagation*

Basicamente o *backpropagation* consiste em dois passos através das diferentes camadas da rede: o passo direto e passo inverso.

No passo direto um padrão de atividade do processo a ser aprendido (vetor de entrada) é aplicado aos nós de entrada da rede e o seu efeito se propaga na rede, camada por camada, produzindo na camada de saída a resposta da rede à excitação aplicada (vetor de saída). Durante o passo direto os pesos sinápticos são todos fixos.

Durante o passo inverso, por outro lado, os pesos sinápticos são todos ajustados, regidos pela regra de aprendizado erro-correção. Especificamente, a resposta da rede à excitação é subtraída de um padrão de resposta desejado para aquela excitação aplicada de forma a produzir um sinal de erro. Este sinal de erro é então propagado de volta através dos mesmos neurônios utilizados no passo direto, mas no caminho contrário do fluxo de sinal nas conexões sinápticas - daí o nome *backpropagation*. Os pesos sinápticos são então ajustados de forma a que a resposta obtida da rede mova-se para mais perto do padrão de resposta desejado.

5.3.1 Definição do algoritmo *backpropagation*

Seja o sinal de erro na saída do neurônio j da camada de saída na iteração n (isto é, na apresentação do n -ésimo padrão de treinamento) definido por

$$e_j(n) = d_j(n) - y_j(n) \quad (58)$$

Define-se o valor do erro quadrático para o neurônio j como $\frac{1}{2}e_j^2(n)$.

Correspondentemente o valor instantâneo da soma dos erros quadráticos $\varepsilon(n)$ é obtida somando $\frac{1}{2}e_j^2(n)$ sobre todos os neurônios da camada de saída; estes são os únicos neurônios “visíveis” para os quais os sinais de erro podem ser calculados. A soma instantânea dos erros quadráticos da rede é então escrita como

$$\varepsilon(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (59)$$

onde C inclui todos os neurônios na camada de saída da rede. Seja N o número total de padrões (exemplos) contido no conjunto de treino. O erro médio quadrático (MSE) é obtido somando $\varepsilon(n)$ sobre n e então normalizando com respeito ao tamanho N do conjunto de treino

$$\varepsilon_{av} = \frac{1}{N} \sum_{n=1}^N \varepsilon(n) \quad (60)$$

O valor instantâneo da soma dos erros quadráticos $\varepsilon(n)$, e portanto o MSE denotado por ε_{av} , são funções de todos os parâmetros livres (pesos sinápticos e *thresholds*) da rede. Para um dado conjunto de treino, ε_{av} representa a Função Custo do processo de minimização do erro de aprendizado [22], constituindo uma medida da performance do aprendizado a partir do conjunto de treino. Para minimizar ε_{av} atualiza-se os pesos sinápticos a cada apresentação de um novo padrão à rede através do vetor de entrada. O ajuste dos pesos é feito de acordo com os respectivos erros computados para cada padrão apresentado à rede. A média aritmética destas alterações individuais nos pesos sobre o conjunto de treino é portanto uma estimativa da verdadeira alteração que resultaria a partir da alteração de pesos baseadas na minimização da função custo ε_{av} sobre todo conjunto de treino.

Considere a Figura 5, a qual mostra o neurônio j sendo alimentado pelos sinais produzidos na saída dos neurônios da camada à esquerda.

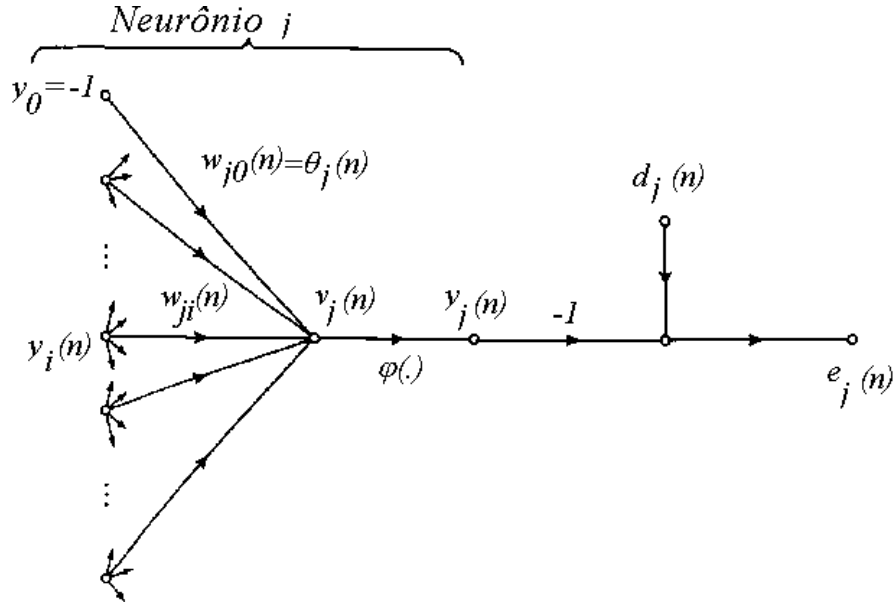


Figura 5: Grafo de fluxo de sinal no neurônio j .

O nível de atividade interna líquida $v_j(n)$ aplicado na entrada da não-linearidade associada ao neurônio j é portanto

$$v_j(n) = \sum_{i=0}^P w_{ji}(n) y_i(n) \quad (61)$$

onde P é o número total de entradas (excluindo o *threshold*) aplicado ao neurônio j . O peso sináptico w_{j0} (correspondente à entrada fixa $y_0 = -1$) define o *threshold* θ_j aplicado ao neurônio j . Portanto o sinal $y_j(n)$ que aparece na saída do neurônio j na iteração n é:

$$y_j(n) = \varphi_j(v_j(n)) \quad (62)$$

De maneira similar ao algoritmo LMS, o *backpropagation* aplica a correção $\Delta w_{ji}(n)$ ao peso sináptico $w_{ji}(n)$ tendo como base a direção contrária do gradiente da superfície de erro $\varepsilon(w)$ relativo ao peso sináptico. Se para uma dada variação no peso sináptico o algoritmo movimenta-se em uma trajetória ascendente na superfície $\varepsilon(w)$ então significa que esta variação deve ser aplicada com o sinal invertido sobre o peso sináptico, já que houve um aumento do erro, e objetiva-se uma diminuição do erro. Por outro lado, se para uma dada variação no peso sináptico o algoritmo movimenta-se em uma trajetória descendente na superfície $\varepsilon(w)$ então significa que esta variação deve ser aplicada com o sinal positivo sobre o peso sináptico, já que houve uma diminuição do

erro e portanto o movimento deve ser encorajado naquela direção. Este método de correção dos pesos sinápticos é denominado de Regra Delta [24].

O gradiente da superfície de erro $\varepsilon(w)$ relativo ao peso sináptico representa portanto um fator de sensibilidade, determinando a direção de procura no espaço de pesos sinápticos para o peso sináptico w_{ji} .

A correção $\Delta w_{ji}(n)$ aplicada a $w_{ji}(n)$, ditada pela Regra Delta, é definida por

$$\Delta w_{ji}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)} \quad (63)$$

onde η é uma constante que determina a razão de aprendizado. Ela é denominada *learning-rate* no algoritmo *backpropagation*. O uso do sinal negativo em (63) impõe a movimentação contrária à direção apontada pelo gradiente na superfície de erro definida no espaço de pesos sinápticos.

O algoritmo *backpropagation* postula que a Regra Delta em um FMLP pode ser implementada por [22]

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (64)$$

onde $\delta_j(n)$ é denominado de gradiente local do neurônio j , definido por

$$\delta_j(n) = \begin{cases} e_j(n) \varphi'_j(v_j(n)) & , \text{neurônio } j \text{ é de saída} \\ \varphi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj} & , \text{neurônio } j \text{ é escondido} \end{cases} \quad (65a)$$

$$\quad (65b)$$

Portanto, o gradiente local sugere as alterações requeridas pelos pesos sinápticos neurônio j .

De acordo com (65a) o gradiente local $\delta_j(n)$ para o neurônio de saída j é igual ao produto do respectivo sinal de erro $e_j(n)$ e a razão de variação $\varphi'_j(v_j(n))$ da função ativação associada. Neste caso o fator chave necessário no cálculo do ajuste dos pesos $\Delta w_{ji}(n)$ é o sinal de erro $e_j(n)$ na saída do neurônio j .

Quando o neurônio j está localizado em uma camada escondida, conforme mostra a Figura 6, mesmo não sendo diretamente acessíveis, tais neurônios dividem a responsabilidade pelo erro resultante na camada de saída. A questão, no entanto, é saber como penalizar ou recompensar os pesos sinápticos de tais neurônios pela sua parcela de

responsabilidade, já que não existe resposta desejada especificada neste local da rede e portanto não há como calcular o sinal de erro. A solução, dada pela equação (65b), é computar o sinal de erro recursivamente para o neurônio escondido em termos dos sinais de erro de todos os neurônios aos quais o neurônio escondido encontra-se conectado.

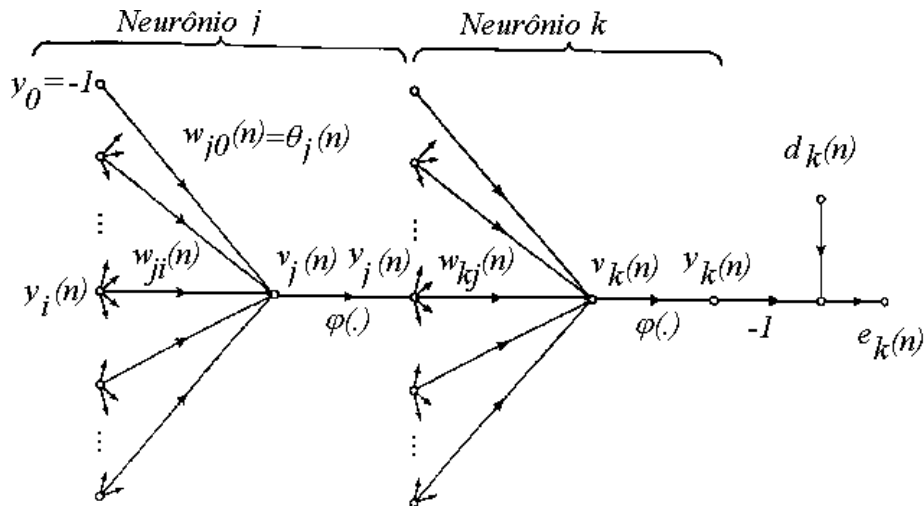


Figura 6: Grafo de fluxo de sinal mostrando os detalhes do neurônio de saída *k* conectado ao neurônio escondido *j*.

O fator $\varphi'(v_j(n))$ envolvido na computação do gradiente local $\delta_j(n)$ na equação (65b) depende somente da função ativação associada com o neurônio escondido *j*. Os demais fatores envolvidos no somatório sobre *k* dependem de dois conjuntos de termos. O primeiro, $\delta_k(n)$, requer conhecimento dos sinais de erro $e_k(n)$ para todos aqueles neurônios localizados na camada imediatamente à direita do neurônio escondido *j* e que estão diretamente conectados a ele (ver Figura 6). O segundo conjunto de termos, $w_{kj}(n)$, consiste de pesos sinápticos associados com estas conexões.

5.3.2 Os Dois Passos Computacionais do *backpropagation*

Na aplicação do algoritmo *backpropagation*, dois passos computacionais distintos podem ser identificados.

No passo direto (*forward pass*) os pesos sinápticos permanecem inalterados em toda rede e os sinais são propagados da entrada da rede para a saída de neurônio a neurônio. O sinal que resulta na saída do neurônio *j* é computado por

$$y_j(n) = \varphi(v_j(n)) \quad (66)$$

onde $v_j(n)$ é o nível de ativação interna líquida do neurônio j definido por

$$v_j(n) = \sum_{i=0}^P w_{ji}(n) y_i(n) \quad (67)$$

sendo P o número total de entradas (excluindo o *threshold*) aplicadas ao neurônio j , $w_{ji}(n)$ é o peso sináptico conectando o neurônio i ao neurônio j e $y_i(n)$ é o sinal de entrada do neurônio j ou equivalentemente o sinal na saída do neurônio i . Se o neurônio j está na primeira camada escondida da rede então o índice i refere-se a i -ésimo nó de entrada da rede, para o qual escreve-se

$$y_i(n) = x_i(n) \quad (68)$$

onde $x_i(n)$ é o i -ésimo elemento do vetor de entrada. Se, por outro lado, o neurônio j está na camada de saída, o índice j refere-se ao j -ésimo nó de saída da rede para o qual escreve-se

$$y_j(n) = o_j(n) \quad (69)$$

sendo $o_j(n)$ o j -ésimo elemento do vetor de saída. Esta saída é comparada com a resposta desejada $d_j(n)$ sendo obtido o sinal de erro $e_j(n)$ para o j -ésimo neurônio de saída. Portanto, o passo direto começa na primeira camada escondida pela apresentação do vetor de entrada a ela e termina na camada de saída com a determinação do sinal de erro para cada neurônio desta camada.

O passo reverso (*backward pass*) começa na camada de saída passando os sinais de erro na direção esquerda através da rede (de volta para a entrada), de camada em camada, e recursivamente computando os gradientes locais para cada neurônio. Este processo recursivo permite que sejam operadas correções nos pesos sinápticos de acordo com a Regra Delta (equação (64)). Para um neurônio localizado na camada de saída o gradiente local é simplesmente o sinal de erro daquele neurônio multiplicado pela primeira derivada de sua não-linearidade (equação (65a)). Em seguida, usa-se a equação (64) para computar as mudanças em todos os pesos das conexões que alimentam a camada de saída. Obtido os gradientes locais para os neurônios da camada de saída, usa-se a equação (65b) para computar os gradientes locais na penúltima camada e portanto operar as mudanças nos pesos sinápticos que alimentam esta camada.

Este procedimento é continuado recursivamente, propagando correções nos pesos sinápticos camada por camada, até a camada de entrada.

Note que durante cada ciclo passo direto - passo reverso ao longo da apresentação do conjunto de treino à rede o vetor de entrada para aquele ciclo é mantido fixo.

5.3.3 A Derivada da Função de Ativação

A determinação do gradiente local para cada neurônio do FMLP requer o conhecimento da derivada da função ativação $\varphi(\cdot)$ associada com o neurônio. Neste trabalho foi utilizada a tangente hiperbólica como não-linearidade neuronal. A derivada da função ativação pode ser obtida através de

$$\begin{aligned}
 \frac{d}{dv} \varphi(v) &= \frac{d}{dv} a \tanh(b \cdot v) = a \frac{d}{dv} \left\{ \frac{1 - e^{-bv}}{1 + e^{-bv}} \right\} = \\
 &= a \left\{ \frac{be^{-bv}(1 + e^{-bv}) - (1 - e^{-bv})(-be^{-bv})}{(1 + e^{-bv})^2} \right\} = a \left\{ \frac{be^{-bv}(1 + e^{-bv} + 1 - e^{-bv})}{(1 + e^{-bv})^2} \right\} = \\
 &= \frac{2abe^{-bv}}{(1 + e^{-bv})^2} = \frac{ab(4e^{-bv})}{2(1 + e^{-bv})^2} = \frac{ab(1 + 2e^{-bv} + e^{-2bv} - 1 + 2e^{-bv} - e^{-2bv})}{2(1 + e^{-bv})^2} = \\
 &= \frac{ab}{2} \left[\frac{(1 + e^{-bv})^2 - (1 - e^{-bv})^2}{(1 + e^{-bv})^2} \right] = \frac{ab}{2} \left[1 - \frac{(1 - e^{-bv})^2}{(1 + e^{-bv})^2} \right] = \frac{ab}{2} [1 - \varphi^2(v)] = \\
 &= \frac{ab}{2} [1 - y^2(v)]
 \end{aligned} \tag{70}$$

sendo $y(v)$ o sinal na saída do neurônio.

5.3.4 Razão de Aprendizado e Momento

O *backpropagation* provê uma aproximação da trajetória de movimento sobre a superfície de erro no espaço de pesos sinápticos a qual, a cada ponto da superfície, segue a direção de descida mais íngreme. Quanto menor a razão de aprendizado η menor serão as correções dos pesos sinápticos da rede de uma iteração para a próxima e mais suave será a trajetória no espaço de pesos. Isto é obtido sob o custo de uma lenta convergência do algoritmo até um valor de erro pequeno o suficiente para ser aceitável. Se, por outro lado, a razão de aprendizado η é feita grande, de modo a acelerar a convergência do algoritmo, as correções feitas nos pesos sinápticos podem resultar demasiadamente grandes, de modo que o algoritmo se torna instável (oscilatório). Um método simples que acelera a convergência e mantém a trajetória estável é a adição do chamado Momento à Regra Delta [22].

$$\Delta w_{ji}(n) = \alpha \cdot \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \quad (71)$$

A constante α é denominada de Constante de Momento e usualmente seu valor é mantido no intervalo $0 < \alpha < 1$. Seu efeito é aumentar a velocidade da trajetória no espaço de pesos na direção da descida mais íngreme. Da equação (71) nota-se que se a correção aplicada em determinado peso sináptico mantém o mesmo sinal algébrico durante várias iterações consecutivas, situação que ocorre quando a trajetória na superfície de erro desenrola-se ao longo de um caminho em descida íngreme, a correção do peso sináptico é acelerada pelo fator de momento, já que, sendo o caminho íngreme, o mínimo deve estar longe ainda. Por outro lado, se a correção aplicada em determinado peso sináptico troca o sinal algébrico durante várias iterações consecutivas, situação que ocorre quando a trajetória na superfície de erro desenrola-se ao longo de um caminho próximo a um mínimo, a correção do peso sináptico é freada pelo fator de momento já que um mínimo está próximo e uma alta velocidade poderia desestabilizar o algoritmo em torno do mínimo.

5.3.5 Sumário Operacional do *backpropagation*

1- Inicialização: Inicializa-se todos os pesos sinápticos com números aleatórios de distribuição uniforme. Usualmente a faixa de inicialização adotada reside no intervalo $[-2.4/F_i, +2.4/F_i]$ onde F_i é o *fan-in* ou o número total de entradas do neurônio [22].

Define-se o momento α e a razão de aprendizado η por camada da rede. Como os neurônios próximos da camada de saída tendem a ter maiores gradientes locais usualmente a eles atribui-se razões de aprendizado menores [22]. Outro critério também a ser considerado simultaneamente é que neurônios com muitas entradas devem ter η menores.

2 - Treinamento: Apresenta-se cada exemplo (vetor de entrada) do conjunto de treino à rede. Para cada exemplo executa-se completamente um ciclo passo direto - passo reverso. O final da apresentação de todos os exemplos do conjunto de treino define uma Época (*epoch*). Ao final de cada Época a ordem dos exemplos no conjunto de treino é alterada aleatoriamente e todo conjunto de treino é apresentado novamente à rede. Prossegue-se o treino da rede de Época em Época até que se atinja o Critério de Parada.

3 - Critério de Parada: O critério de parada no treino de um FMLP é subjetivo, já que não existe prova que o *backpropagation* tenha convergido para o mínimo global (se é que existe) da superfície de erro. Neste trabalho foi utilizado como critério de parada o seguinte procedimento:

As iterações de treino são terminadas se

- O valor do MSE atingiu um valor suficientemente baixo

ou

- A razão de variação do MSE atingiu um valor suficientemente baixo em valor absoluto e negativo.

Quando qualquer uma das condições acima é atingida, considera-se que a rede não necessita mais ser treinada.

6. Processo de Detecção do Padrão na Imagem Alvo

O processo de detecção da imagem padrão pressupõe o armazenamento prévio de suas características. Uma vez conhecidas tais características, estas podem ser utilizadas como referência no processo de detecção em uma imagem alvo genérica. Neste trabalho, as características do módulo do espectro do padrão são armazenadas em uma RNA do tipo FMLP a qual é utilizada no processo de detecção como referência espectral. Como as informações de módulo do espectro dizem respeito apenas ao formato do padrão, ficando as informações de localização no espectro de fase, obtém-se invariância à posição.

6.1 O Armazenamento das Características do Padrão

Seja $p(x,y)$ a imagem padrão de $N \times N$ pixels e $P(u,v) = \text{FFT}\{p(x,y)\}$.

Fazendo

$$\begin{aligned} u &= \omega \cos \Psi \\ v &= \omega \sin \Psi \end{aligned} \quad (72)$$

de tal forma que $P(\omega, \Psi)$ seja o espectro $P(u,v)$ em coordenadas polares.

Seja $E(\Psi)$ a Energia de $p(x,y)$ ao longo da direção espectral Ψ , definida pela equação (4) com u e v definido por (72) e ω tal que $0 < u < N-1$ e $0 < v < N-1$.

Inicialmente seleciona-se as m direções $\Psi_j, j=0,1..m-1$ tal que $E(\Psi_j)$ seja um máximo local de $E(\Psi)$.

A Figura 7 mostra, a título de exemplo, o caractere alfabético “A” do tipo de letra *Times New Roman* e o módulo de seu espectro. Esta imagem padrão apresenta $m=3$ máximos locais para $E(\Psi)$, claramente definidos nas direções $\Psi_0 = 156^\circ$, $\Psi_1 = 30^\circ$ e $\Psi_2 = 24^\circ$.

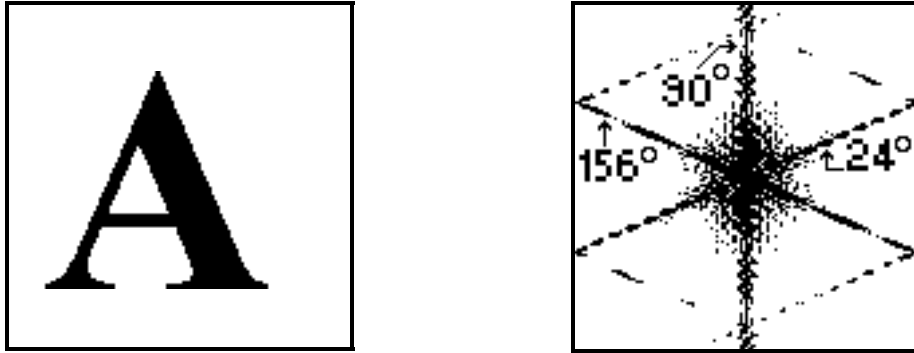


Figura 7: Letra “A” e módulo de seu espectro.

A Figura 8 mostra, o caractere alfabético “C” do mesmo tipo de letra e o módulo de seu espectro. Vê-se que não há máximo local para $E(\Psi)$ claramente definido. Nesta situação em que o espectro apresenta simetria rotacional convém adotar o mínimo de direções espectrais de máxima energia ($m=2$). No caso da letra “C” estas direções são $\Psi_0 = 5^\circ$ e $\Psi_1 = 179^\circ$.

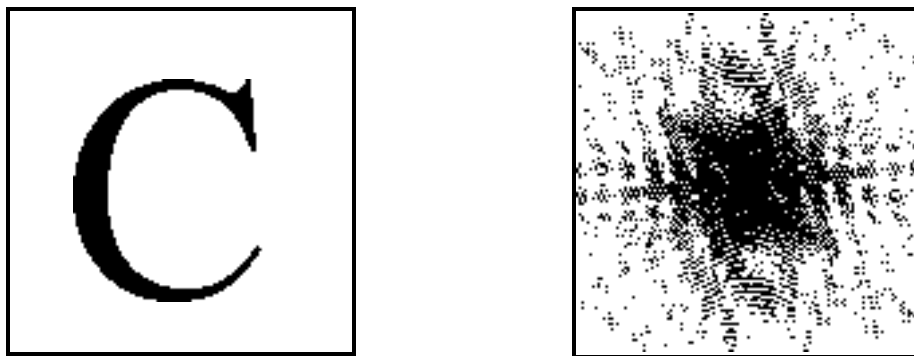


Figura 8: Letra “C” e módulo de seu espectro.

Como regra geral, o número m de direções espectrais de máxima energia deve ser idêntico ao número de máximos locais de $E(\Psi)$. Um número maior que este introduz informação irrelevante no processo de detecção resultando em uma incerteza maior na estimativa de existência ou não do padrão na imagem alvo. Por outro lado, um número m menor deixa de fornecer ao detector informação relevante sobre o padrão, incorrendo no mesmo problema.

Obtidas as direções espectrais de maior energia, o próximo passo consiste em treinar um conjunto de RNAs para reconhecer as características espectrais do padrão. Neste trabalho, o conjunto de treino foi formado considerando-se apenas variações do módulo do espectro $P(\omega, \Psi)$ ao longo das direções Ψ_j resultantes do escalonamento do padrão $p(x, y)$ no domínio espaço. Especificamente, foi utilizada para treino a seqüência

$|P(\omega, \Psi_j)|$ resultante do escalonamento de $p(x,y)$ pelo conjunto de fatores $\alpha=\{1, 0.98, \dots, 0.1\}$. É possível acrescentar ao treino do sistema de RNAs qualquer variação adicional considerada relevante, tais como ruído, efeitos de desfocalização do sistema ótico de aquisição da imagem alvo e variações quanto ao ângulo de perspectiva do padrão no alvo. Não é necessário incluir no treino o efeito de rotação de $p(x,y)$ no plano (x,y) visto que o sistema de detecção extrai amostras do módulo do espectro da imagem alvo utilizando um sistema de coordenadas polares (ver Apêndice A).

Especificamente, treina-se as m RNAs com os $\omega=0,1\dots N/2$ valores de $|P(\omega, \Psi_j)|$ extraídos ao longo da direção Ψ_j para cada $|P(\omega, \Psi_j)|$ resultante do escalonamento de $p(x,y)$ pelo conjunto de valores definido por α . O conjunto $\omega=0,1\dots N/2$ foi assim definido porque $|P(\omega, \Psi)|$ apresenta simetria par ao longo de Ψ_j .

Por exemplo, considerando a imagem $p(x,y)$ de $N \times N$ pixels, $N=128$, da letra “A” mostrada na Figura 7, e o módulo ($\log(|P(\omega, \Psi)|+1)$) de seu espectro juntamente com as m ($m=3$) direções Ψ_j ao longo das quais é máxima a energia $E(\Psi)$.

Para cada $|P(\omega, \Psi_j)|$ resultante do escalonamento de $p(x,y)$ por um dos valores de $\alpha=\{1, 0.98, \dots, 0.1\}$, extraem-se os 65 ($\omega=0,1\dots N/2$) valores de $|P(\omega, \Psi_j)|$ ao longo das direções $\Psi_0 = 156^\circ, \Psi_1 = 30^\circ$ e $\Psi_2 = 24^\circ$, a partir do centro do espectro ($\omega=0$), formando os vetores $S_0(\omega)$, $S_1(\omega)$ e $S_2(\omega)$ para cada valor de α .

Para cada direção Ψ_j forma-se o conjunto de treino T_j formado por V funções membro (vetores de entrada da RNA - ver item 5.3), cada uma delas constituída pelas 65 amostras espectrais contidas no vetor $S_j(\omega)$, mais um elemento adicional representando o valor do respectivo α necessário para determinação do erro instantâneo no aprendizado por *backpropagation*. O número de funções membro V é dependente dos valores máximo, mínimo e decremento adotados para α . A expressão (73) ilustra este cálculo para o conjunto $\alpha=\{1, 0.98, \dots, 0.1\}$.

$$V = \frac{(1 - 0.1)}{(1 - 0.98)} + 1 = 46 \quad (73)$$

Após a normalização de seus valores numéricos para a faixa $\{-1,1\}$, cada T_j é utilizado para treinar a respectiva RNA R_j com arquitetura 65-7-1, resultando na matriz de pesos sinápticos W_j .

O número de neurônios escondidos foi experimentalmente fixado em 7. Este número mostrou-se suficiente para armazenar as informações espectrais de caracteres do tipo de letra *Times New Roman*, bem como de formas geométricas simples (retas, polígonos, circunferências, curvas) e composições destas, os quais, neste trabalho, constituíram o conjunto de teste. O MSE de treino final para todas estas formas geométricas, em todas direções espectrais de maior energia foi menor que 0.00018.

Utilizou-se como estratégia de treino das RNAs o seguinte procedimento:

- Inicia-se o treinamento do FMLP com uma razão de aprendizado de 0.005 para a camada escondida e 0.01 para a de saída. O momento é fixado em 0.1.
- Quando o MSE cai abaixo de 0.0025 aumenta-se a razão de aprendizado para 1.00 nas camadas escondida e de saída. Simultaneamente aumenta-se o momento para 0.5.
- Quando o MSE cai abaixo de 0.00077 aumenta-se o momento para 0.8 e mantém-se a razão de aprendizado em 1.00 nas camadas escondida e de saída.

Com este procedimento o MSE final obtido é menor que 0.00018 e menos do que 8000 épocas são processadas. As figuras 9 a 11 ilustram esta estratégia de treino.

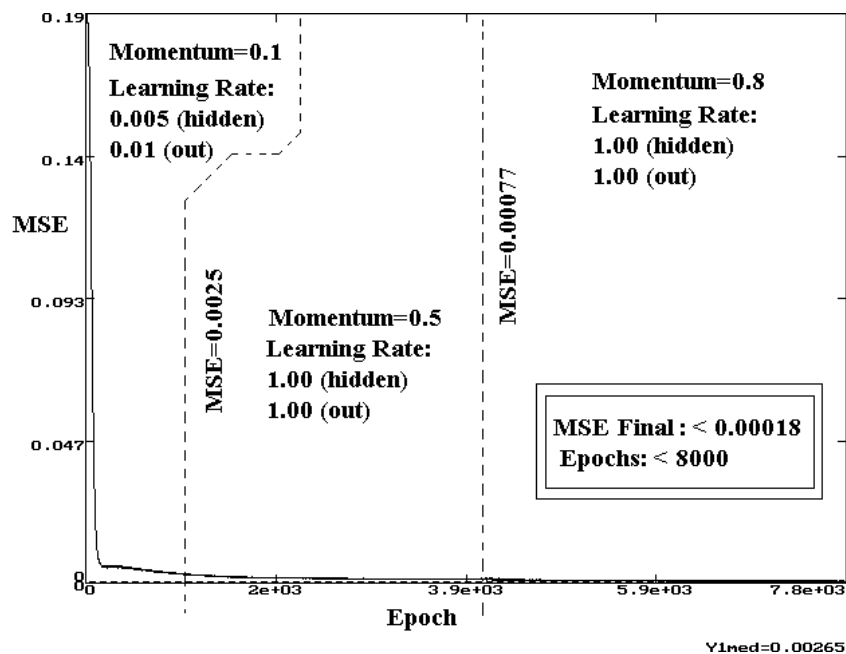


Figura 9: Estratégia geral de treino do FMLP.

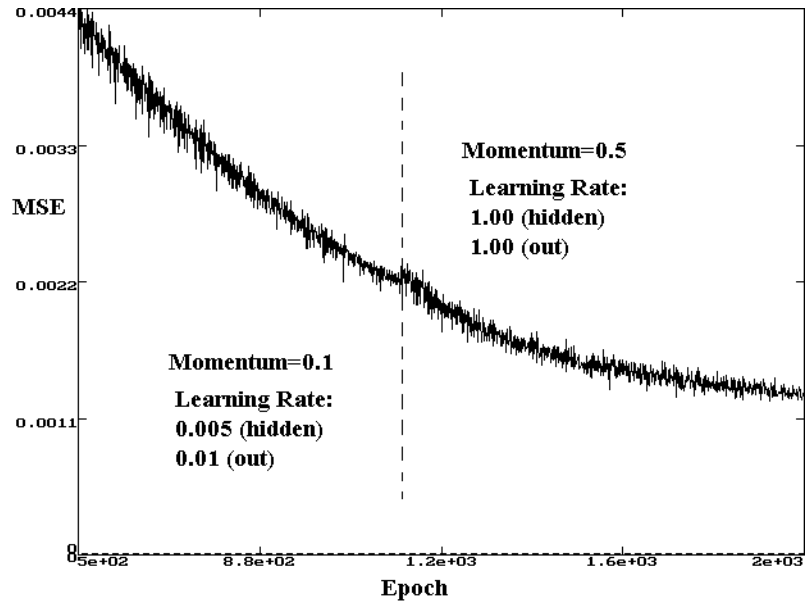


Figura 10: Estratégia de treino do FMLP. Curva de aprendizado para $500 < \text{época} < 2000$.

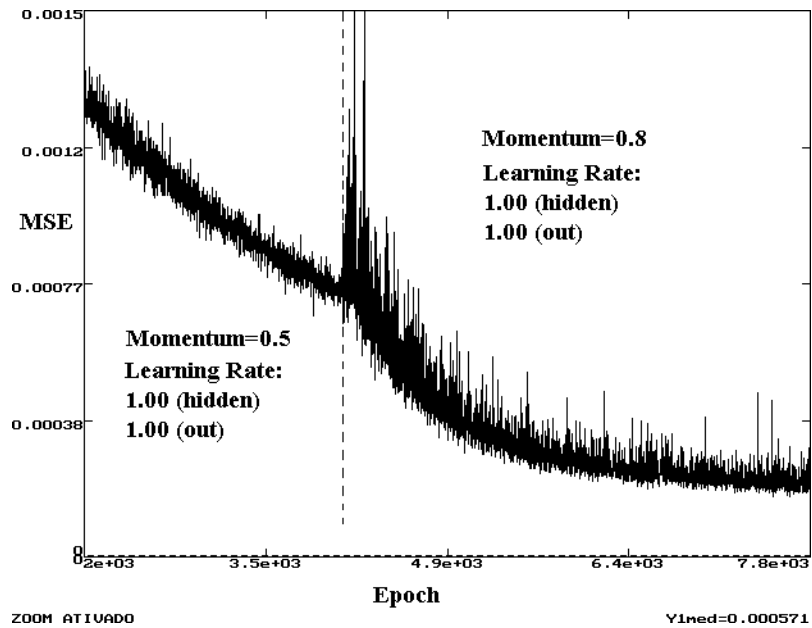


Figura 11: Estratégia de treino do FMLP. Curva de aprendizado para $2000 < \text{época} < 8000$.

Em resumo:

Dado uma imagem padrão, seleciona-se as m direções espectrais de máxima energia $E(\Psi_j)$, $j=0,1..m-1$. Portanto, o padrão é representado por um conjunto de m RNAs cujos conjuntos de pesos sinápticos são W_0, W_1, \dots, W_{m-1} . Cada matriz W_j é resultante do treinamento pelo algoritmo *backpropagation* em cada uma das m direções de maior energia espectral. Cada padrão a ser identificado terá um conjunto distinto de m direções de maior energia espectral.

6.2 A detecção e localização do Padrão na Imagem Alvo

Seja uma imagem alvo $a(x,y)$ de $N \times N$ pixels na qual desejamos encontrar o padrão $p(x,y)$, que pode encontrar-se escalonado, rotacionado ou transladado. Sejam as matrizes W_0, W_1 e W_{m-1} relativas às m RNAs representativas das direções de máxima energia espectral $\Psi_0, \Psi_1, \dots, \Psi_{m-1}$ de $p(x,y)$. Seja $A(u,v) = \text{FFT}\{a(x,y)\}$. Seja $A(\omega, \Psi)$ o espectro $A(u,v)$ em coordenadas polares.

Inicia-se o processo de detecção calculando o módulo do espectro de $a(x,y)$, isto é, $|A(\omega, \Psi)|$. A seguir, para a direção $\Psi_j + \Psi$, $\Psi = 0, 1, \dots, 179^\circ$, $j=0,1, \dots, m-1$, extraem-se os $N/2+1$ ($\omega=0,1..N/2$) valores de $|A(\omega, \Psi_j)|$ ao longo da direção em questão, e aplica-se estes valores na entrada da RNA definida por W_j . Para cada Ψ , atribui-se o valor resultante do único neurônio de saída da RNA R_j ao vetor $O_j(\Psi)$. Forma-se então o *ensemble* $\mathbf{A}(\psi)$, constituído pelas m funções membros $O_0(\Psi), O_1(\Psi), \dots, O_{m-1}(\Psi)$ e calcula-se a média $\mu(\Psi)$ e a variação da média $\tilde{\mu}(\Psi)$ através de

$$\mu(\Psi) = \frac{1}{m} \sum_{j=0}^{m-1} O_j(\Psi) \quad (74)$$

$$\tilde{\mu}(\Psi) = \mu(\Psi) - \frac{1}{180^\circ} \sum_{\Psi=0}^{179^\circ} \mu(\Psi) \quad (75)$$

A Figura D.1 mostrada no Apêndice C ilustra o processo descrito no parágrafo anterior para $m=6$. Cada um dos 6 triângulos representam uma RNA treinada com uma das $m=6$ direções de maior energia espectral do padrão. Para um dado Ψ , cada rede R_j , $j=0,1, \dots, m-1$, recebe em sua entrada os 65 valores do módulo do espectro da imagem alvo $a(x,y)$ resultando o valor $O_j(\Psi)$ na saída do único neurônio da rede R_j . O *ensemble* constituído pelas $m=6$ funções membros é formado girando cada um dos $m=6$ vetores

de entrada de cada RNA ao longo das direções $\Psi_j + \Psi$, $\Psi = 0, 1, \dots, 179^\circ$. Finalmente obtém-se a média $\mu(\Psi)$ e a variação da média $\tilde{\mu}(\Psi)$ do *ensemble* através de (74) e (75).

Sejam então $\mu_A(\Psi)$ e $\tilde{\mu}_A(\Psi)$ resultantes do processo descrito na Figura D.1 aplicado sobre o módulo do espectro $|A(\omega, \Psi)|$ da imagem alvo. O passo seguinte no processo de detecção consiste em obter $\alpha_a = \max\{\mu_A(\Psi)\}$, isto é, o valor máximo de $\mu_A(\Psi)$. O valor α_a é uma aproximação inicial do grau de escalonamento do padrão supostamente existente na imagem alvo. Isto ocorre porque cada RNA foi treinada com os valores de módulo do espectro na respectiva direção Ψ_j de máxima energia, o que implica nos valores de módulo serem máximos ao longo de Ψ_j . Sendo assim, a média do *ensemble* $\mathbf{A}(\psi)$ é máxima quando todos os m vetores de entradas das RNAs estiverem aproximadamente alinhados com suas respectivas direções Ψ_j .

Seja $p_T(x,y)$ a imagem padrão $p(x,y)$ escalonada pelo fator α_a . Seja $P_T(u,v) = \text{FFT}\{p_T(x,y)\}$ e $P_T(\omega, \Psi)$ o espectro $P_T(u,v)$ em coordenadas polares. Aplicando o processo de detecção descrito na Figura D.1 sobre $|P_T(\omega, \Psi)|$ obtém-se $\mu_{P_T}(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$ a partir das equações (74) e (75) e do *ensemble* $\mathbf{P}_T(\psi)$ resultante. A correlação cruzada entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$ é dada por (ver item 3.2):

$$R_{AT}(\theta) = \text{IFFT}\left\{\text{FFT}\{\tilde{\mu}_A(\Psi)\}\text{FFT}^*\{\tilde{\mu}_{P_T}(\Psi)\}\right\} \quad (76)$$

O valor θ_0 para o qual $R_{AT}(\theta)$ é máximo é uma estimativa do ângulo de rotação do padrão supostamente existente na imagem alvo e o valor $\mu_A(\theta_0)$ é uma estimativa de seu grau de escalonamento α_0 .

De posse da estimativa do grau de escalonamento α_0 e do ângulo de rotação θ_0 do padrão que se supõe existir na imagem alvo, resta decidir se a forma geométrica encontrada é realmente a forma geométrica do padrão, e localizá-la.

Seja $p_0(x,y)$ a imagem padrão $p(x,y)$ escalonada por α_0 e rotacionada de um ângulo θ_0 . Seja $P_0(u,v) = \text{FFT}\{p_0(x,y)\}$. Seja $P_0(\omega, \Psi)$ o espectro $P_0(u,v)$ em coordenadas polares.

No item 4 demonstrou-se que a translação (x_0, y_0) entre $a(x, y)$ e $p_0(x, y)$ é dada por

$$x_0(u, v) = \frac{N}{2\pi} \left[\frac{1}{\text{Re}^2 \{P_0(u, v) A^*(u, v)\} + \text{Im}^2 \{P_0(u, v) A^*(u, v)\}} \right]^x \left\{ \begin{array}{l} \text{Im} \{P_0(u, v) A^*(u, v)\} \text{Re} \left\{ A^*(u, v) \frac{d}{du} P_0(u, v) + P_0(u, v) \left[\frac{d}{du} A(u, v) \right]^* \right\} + \\ - \text{Re} \{P_0(u, v) A^*(u, v)\} \text{Im} \left\{ A^*(u, v) \frac{d}{du} P_0(u, v) + P_0(u, v) \left[\frac{d}{du} A(u, v) \right]^* \right\} \end{array} \right\} \quad (77)$$

$$y_0(u, v) = \frac{N}{2\pi} \left[\frac{1}{\text{Re}^2 \{P_0(u, v) A^*(u, v)\} + \text{Im}^2 \{P_0(u, v) A^*(u, v)\}} \right]^x \left\{ \begin{array}{l} \text{Im} \{P_0(u, v) A^*(u, v)\} \text{Re} \left\{ A^*(u, v) \frac{d}{dv} P_0(u, v) + P_0(u, v) \left[\frac{d}{dv} A(u, v) \right]^* \right\} + \\ - \text{Re} \{P_0(u, v) A^*(u, v)\} \text{Im} \left\{ A^*(u, v) \frac{d}{dv} P_0(u, v) + P_0(u, v) \left[\frac{d}{dv} A(u, v) \right]^* \right\} \end{array} \right\} \quad (78)$$

Sejam $x_0(\omega, \Psi)$ e $y_0(\omega, \Psi)$ as matrizes de translação $x_0(u, v)$ e $y_0(u, v)$ definidas em (77) e (78) com indexação de suas linhas e colunas expressas em coordenadas polares. Seja o intervalo de frequência $[\omega_1, \omega_2]$ tal que para valores de ω compreendidos na faixa $\omega_1 < \omega < \omega_2$ os valores de $|P_0(\omega, \Psi)|$ sejam máximos. Sejam $\Psi_j, j=0, 1, \dots, m-1$, as m direções espectrais de maior energia do espectro $P(\omega, \Psi)$.

Estando as matrizes de translação $x_0(\omega, \Psi)$ e $y_0(\omega, \Psi)$ expressas em coordenadas polares é possível aplicar filtragem direcional ao longo de Ψ_j no domínio frequência (ω, Ψ) . Seja \hat{x}_0 o valor numérico da estimativa da translação entre $p_0(x, y)$ e $a(x, y)$, determinado pelo filtro direcional definido através de

$$\hat{x}_0 = \frac{1}{m(\omega_2 - \omega_1 + 1)} \sum_{\omega=\omega_1}^{\omega_2} \sum_{j=0}^{m-1} x_0(\omega, \Psi_j + \theta_0) \quad (79)$$

A equação (79) computa a média dos valores de translação ao longo das direções de maior energia espectral Ψ_j para valores de frequência ω compreendidos no intervalo $[\omega_1, \omega_2]$ dentro do qual a relação sinal-ruído é máxima. Considera-se aqui como sinal as componentes espectrais de $p_0(x, y)$ presentes no espectro de $a(x, y)$ e como ruído todas as componentes no espectro de $a(x, y)$ que não sejam características do espectro de $p_0(x, y)$.

Como, na prática, não se sabe que tipo de ruído estará presente em $A(\omega, \Psi)$ escolhe-se $[\omega_1, \omega_2]$ de forma às componentes espectrais de $P_0(\omega, \Psi)$ serem máximas no intervalo.

O grau de certeza da estimativa de translação \hat{x}_0 é uma medida direta do grau de similaridade geométrica entre $a(x,y)$ e $p_0(x,y)$. Se $a(x,y)$ é geometricamente igual a $p_0(x,y)$ todos os elementos da matriz $x_0(\omega, \Psi)$ assumem o mesmo valor. Se $a(x,y)$ é geometricamente diferente de $p_0(x,y)$ os elementos da matriz $x_0(\omega, \Psi)$ tornam-se distintos entre si, mesmo que $a(x,y)$ e $p_0(x,y)$ estejam localizados nas mesmas coordenadas de posição (x,y) . Isto ocorre principalmente porque a relação entre as derivadas dos espectros nas direções u e v torna-se dependente de (ω, Ψ) quando as imagens $a(x,y)$ e $p_0(x,y)$ são geometricamente distintas. Este grau de semelhança pode ser medido através da variância normalizada do conjunto amostral, obtida através de

$$V\hat{x}_0 = \frac{1}{m(\omega_2 - \omega_1 + 1)(\hat{x}_0)^2} \sum_{\omega=\omega_1}^{\omega_2} \sum_{j=0}^{m-1} \left(x_0(\omega, \Psi_j + \theta_0) - \hat{x}_0 \right)^2 \quad (80)$$

Um baixo valor para $V\hat{x}_0$ indica uma alta semelhança entre $a(x,y)$ e $p_0(x,y)$ quanto à forma. Para cada padrão a ser detectado no cenário de possíveis imagens alvo $a(x,y)$ existe um limiar para $V\hat{x}_0$, a ser determinado experimentalmente, que indica a forma geométrica contida em $a(x,y)$ ser ou não a forma geométrica do padrão $p(x,y)$. Neste trabalho adota-se o valor percentual de $V\hat{x}_0$, isto é $V\hat{x}_0 \% = 100.V\hat{x}_0$.

Com exceção do espectro de uma forma geométrica impulsiva, todo espectro tem o máximo de seu módulo no centro em $\omega=0$. Portanto, para efeito de filtragem direcional em frequência, em geral, faz-se $\omega_1=0$ e define-se o círculo com centro em $\omega=0$ e raio ω_2 de forma ao somatório em (79) e (80) abranger uma região do plano (ω, Ψ) tal que em seu interior fiquem contidos os valores de $|P_0(\omega, \Psi)|$ maiores que D_{BW} , uma constante arbitrária que define a banda passante do filtro direcional. Considere, por exemplo, o espectro $|P_0(\omega, \Psi)|$ de uma imagem de 128x128 pixels. Ao definir o círculo $\omega_2 = 8$ tal que $|P_0(\omega, \Psi)|$ seja maior que $D_{BW} = 1/4$ de seu valor máximo em todos os pontos internos ao círculo, caracteriza-se um filtro com banda passante (BW) de -12 dB (-12db=1/4) de 16 pixels ($2\omega_2 = 16$). Neste trabalho, tal filtro é referido como um filtro -12dB BW = 12.5% ($(16/128).100=12.5\%$).

Semelhantemente têm-se para a matriz de translação $y_0(\omega, \Psi)$:

$$\hat{y}_0 = \frac{1}{m(\omega_2 - \omega_1 + 1)} \sum_{\omega=\omega_1}^{\omega_2} \sum_{j=0}^{m-1} y_0(\omega, \Psi_j + \theta_0) \quad (81)$$

$$V\hat{y}_0 = \frac{1}{m(\omega_2 - \omega_1 + 1)(\hat{y}_0)^2} \sum_{\omega=\omega_1}^{\omega_2} \sum_{j=0}^{m-1} \left(y_0(\omega, \Psi_j + \theta_0) - \hat{y}_0 \right)^2 \quad (82)$$

Como os espectros em consideração tem simetria par ao longo das direções Ψ_j , o método de detecção apresenta uma ambigüidade de 180° para θ_0 . Esta ambigüidade fica resolvida se compararmos os valores de $V\hat{x}_0$ e $V\hat{y}_0$ com aqueles obtidos a partir de um ângulo de rotação de $\theta_0 + 180^\circ$. A situação com menores $V\hat{x}_0$ e $V\hat{y}_0$ indicará o verdadeiro ângulo de rotação da imagem alvo em relação à imagem padrão. Não é necessário, no entanto, que o processo de detecção seja repetido duas vezes (uma para θ_0 e outra para $\theta_0 + 180^\circ$) para que se elimine a ambigüidade em θ_0 ; basta que se aplique as propriedades espectrais desenvolvidas nos itens 3.7, 3.8 e 3.9 às equações (77) e (78).

6.3 Sumário do Processo de Detecção e Localização

Seja uma imagem padrão $p(x,y)$ conhecida, caracterizada por m direções espectrais de maior energia. Seja o conjunto de m RNAs cujos conjuntos de pesos sinápticos são W_0, W_1, \dots, W_{m-1} treinadas com variações do módulo do espectro de $p(x,y)$ relativas ao seu escalonamento, ao longo das direções da máxima energia. Seja uma imagem alvo $a(x,y)$ de $N \times N$ pixels na qual desejamos encontrar o padrão $p(x,y)$.

Procede-se da seguinte maneira:

- Determina-se $A(\omega, \Psi) = \text{FFT} \{ a(x,y) \}$.
- Submete-se $|A(\omega, \Psi)|$ ao processo mostrado na Figura D.1 e obtém-se $\mu_A(\Psi)$ e $\tilde{\mu}_A(\Psi)$ a partir do *ensemble* resultante $\mathbf{A}(\psi)$ através de (74) e (75).
- Determina-se $\alpha_a = \max \{ \mu_A(\Psi) \}$.
- Escalona-se $p(x,y)$ por α_a e obtém-se $p_T(x,y)$.
- Determina-se $P_T(\omega, \Psi) = \text{FFT} \{ p_T(x,y) \}$.
- Submete-se $|P_T(\omega, \Psi)|$ ao processo mostrado na Figura D.1 e obtém-se $\tilde{\mu}_{p_T}(\Psi)$ a partir do *ensemble* resultante $\mathbf{P}_T(\psi)$ através de (74) e (75).

- Determina-se $R_{AT}(\theta)$ através de (76).
- Determina-se θ_0 tal que $R_{AT}MAX = \max \{ R_{AT}(\theta) \}$ para $\theta = \theta_0$.
- Determina-se $\alpha_0 = \mu_A(\theta_0)$.
- Escalona-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$.
- Determina-se $P_0(\omega, \Psi) = \text{FFT} \{ p_0(x,y) \}$.
- Determina-se $V\hat{x}_0, \hat{x}_0, V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$ (ver itens 3.7, 3.8 e 3.9).
- A partir do limiar adotado para $V\hat{x}_0$ e $V\hat{y}_0$ decide-se se a forma geométrica em $a(x,y)$ é o padrão ou não e levanta-se a ambigüidade de θ_0 .

7. Simulações

Este item dedica-se a demonstrar o funcionamento do algoritmo em diversas situações de identificação de caracteres alfabéticos e formas geométricas simples. O processo de identificação consiste em determinar se a imagem alvo contém ou não a forma geométrica do padrão, o escalonamento, ângulo de rotação e translação relativa do padrão encontrado. Sugere-se que se acompanhe a seqüência de operações executadas através do sumário do item 6.3.

7.1 Identificação do Caracter “A” *Times New Roman* Escalonado, Rotacionado e Transladado

Neste item é mostrado a identificação do caracter “A” *Times New Roman*. A Figura 12 mostra a imagem padrão e a Figura 13 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão.



Figura 12: Imagem padrão $p(x,y)$ - caracter “A” *Times New Roman* 128x128 pixels.

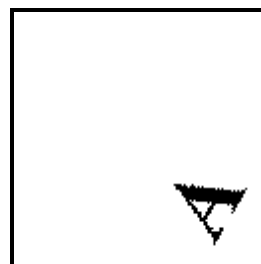


Figura 13: Imagem alvo $a(x,y)$ - “A” da Figura 12 escalonado por 0.4 rotacionado de 60° e transladado de (40,30) pixels.

Inicialmente, determina-se $A(\omega, \Psi) = \text{FFT}\{a(x,y)\}$ e submete-se $|A(\omega, \Psi)|$ ao processo mostrado na Figura D.1. O resultado são as curvas $\mu_A(\Psi)$ e $\tilde{\mu}_A(\Psi)$, mostradas nas Figuras 14 e 15, obtidas a partir do *ensemble* resultante $\mathbf{A}(\psi)$ e através de (74) e (75).

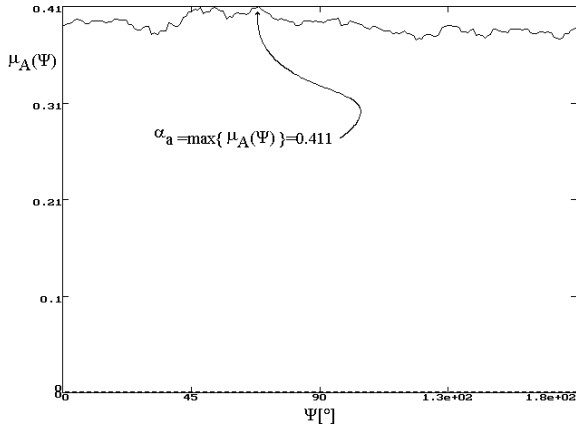


Figura 14: Curva $\mu_A(\Psi)$ do *ensemble* $\mathbf{A}(\psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

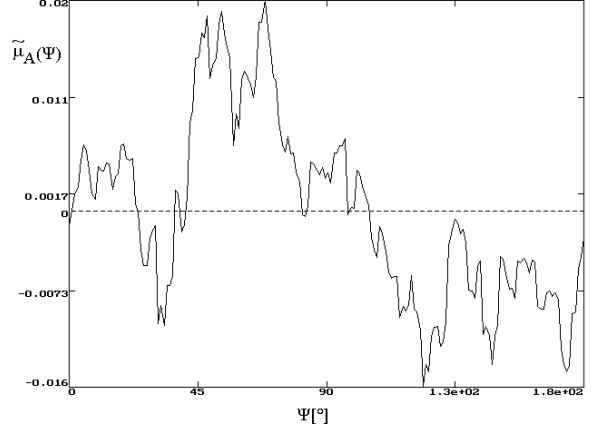


Figura 15: Curva $\tilde{\mu}_A(\Psi)$ do *ensemble* $\mathbf{A}(\psi)$ obtida através de (75).

Continuando, determina-se $\alpha_a = \max\{\mu_A(\Psi)\}$ (ver Figura 14), escalona-se $p(x,y)$ por α_a e obtém-se $p_T(x,y)$. Determina-se, então, $P_T(\omega, \Psi) = \text{FFT}\{p_T(x,y)\}$ e submete-se $|P_T(\omega, \Psi)|$ ao processo mostrado na Figura D.1. O resultado é a curva $\tilde{\mu}_{P_T}(\Psi)$, mostrada na Figura 16, obtida a partir do *ensemble* resultante $\mathbf{P}_T(\psi)$ e através de (74) e (75).

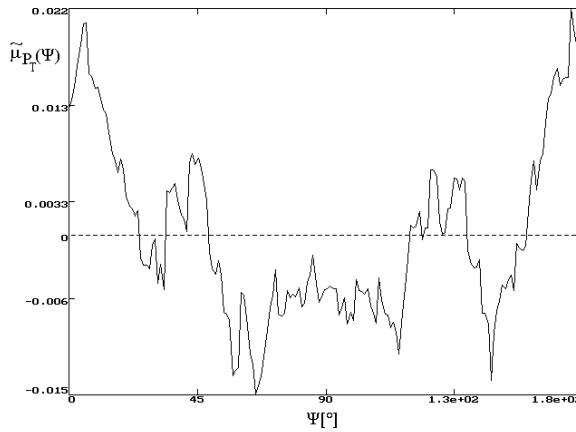


Figura 16: Curva $\tilde{\mu}_{P_T}(\Psi)$ do *ensemble* $\mathbf{P}_T(\psi)$ obtida através de (75).

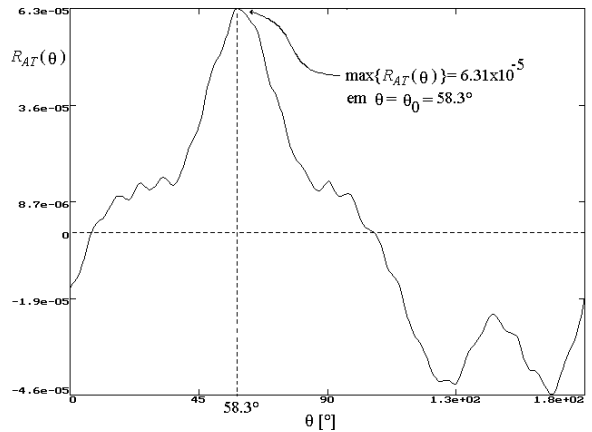


Figura 17: $R_{AT}(\theta)$ obtida de (76). O máximo de $R_{AT}(\theta)$ estima o ângulo de rotação θ_0 .

A partir das curvas $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, Figuras 15 e 16, determina-se a correlação cruzada $R_{AT}(\theta)$ entre elas através de (76), mostrada na Figura 17. A partir de $R_{AT}(\theta)$, determina-se θ_0 tal que $R_{AT}(\theta)_{MAX} = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

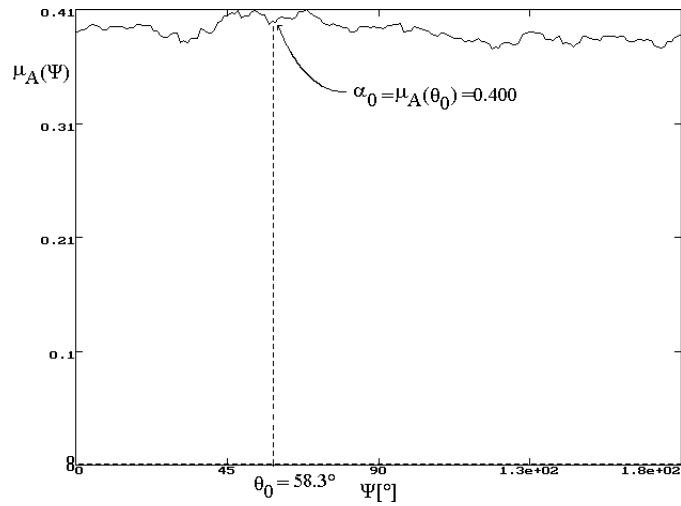


Figura 18: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

De posse da estimativa do ângulo de rotação θ_0 , determina-se $\alpha_0 = \mu_A(\theta_0)$ conforme mostra a Figura 18.

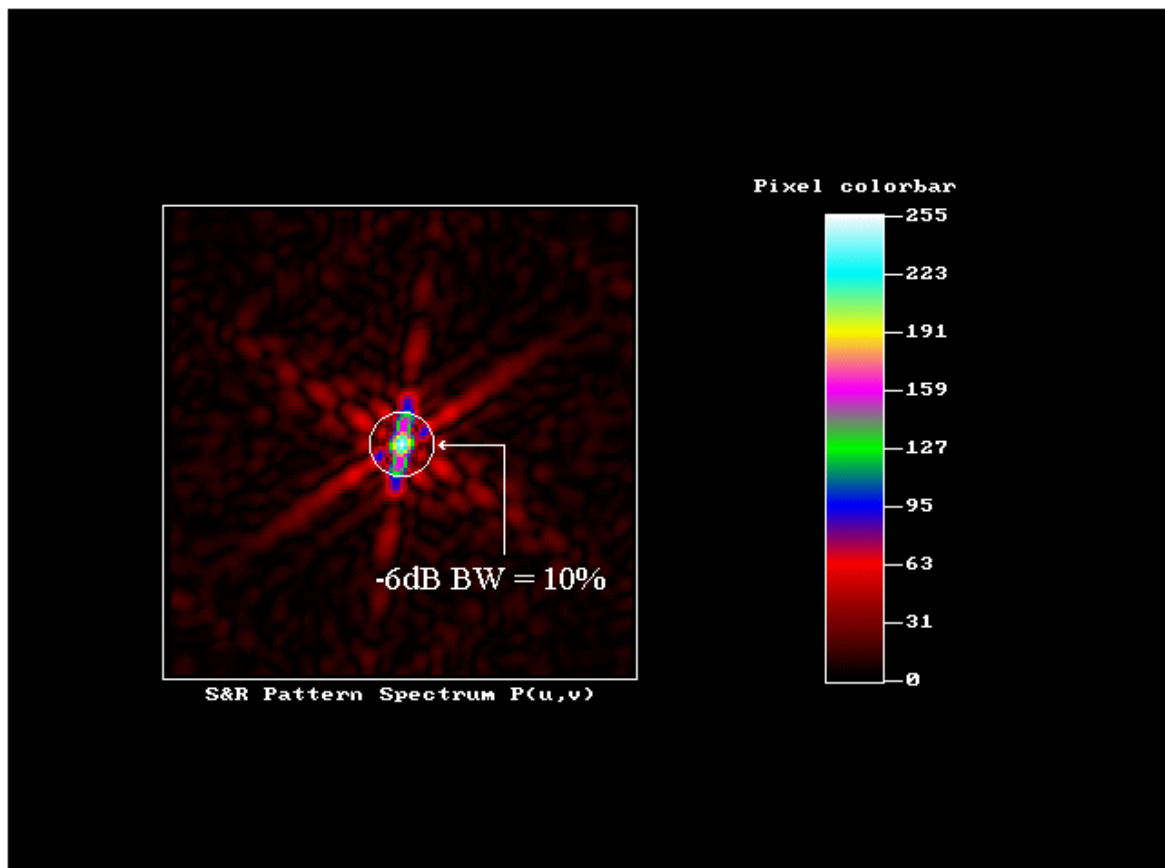


Figura 19: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 10%.

Finalmente, escalona-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$.

Determina-se $P_0(\omega, \Psi) = \text{FFT}\{p_0(x, y)\}$ (ver Figura 19) e obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 2.

Tabela 2: Resultado do processo de detecção e localização - equações (77) a (82)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como θ_0+180°
\hat{x}_0	40.5 pixels	48.7 pixels
\hat{y}_0	29.7 pixels	32.0 pixels
$V\hat{x}_0\%$	0.09%	14.11%
$V\hat{y}_0\%$	0.03%	5.32%

Como $V\hat{x}_0\% = 0.09\%$ e $V\hat{y}_0\% = 0.03\%$ para a estimativa de ângulo de rotação $\theta_0 = 58.3^\circ$ conclui-se que esta é a estimativa do ângulo, e não $\theta_0 + 180^\circ$, em que o padrão encontra-se rotacionado na imagem alvo. O fator de escalonamento estimado é $\alpha_0 = 0.400$.

7.2 Identificação do Carácter “V” *Times New Roman* Comparado ao Padrão “A”(Figura 12)

Neste item é mostrado a identificação do carácter “V” *Times New Roman* em relação à imagem padrão da Figura 12. A Figura 20 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão. A identificação apresentada neste item exemplifica a sensibilidade à detalhes do método, já que o carácter “V” difere do carácter “A” apenas pela barra horizontal.

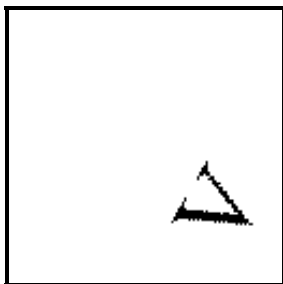


Figura 20: Imagem alvo $a(x, y)$ - carácter “V” *Times New Roman* com fator de escalonamento 0.4 rotacionado de 60° e transladado de (40,30) pixels.

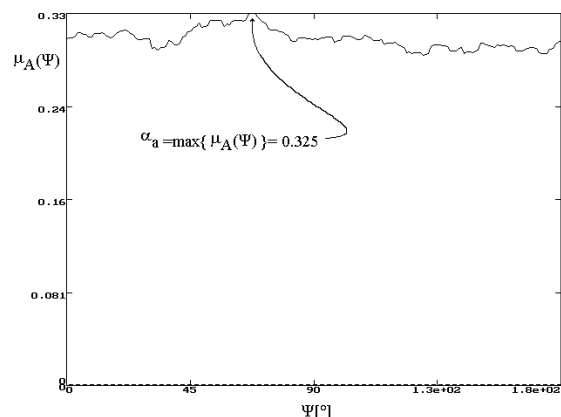


Figura 21: Curva $\mu_A(\Psi)$ do *ensemble* $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

Primeiro, determina-se $A(\omega, \Psi) = \text{FFT}\{a(x, y)\}$ e submete-se $|A(\omega, \Psi)|$ ao processo mostrado na Figura D.1. O resultado são as curvas $\mu_A(\Psi)$ e $\tilde{\mu}_A(\Psi)$, mostradas nas Figuras 21 e 22, obtidas a partir do *ensemble* resultante $\mathbf{A}(\psi)$ e através de (74) e (75).

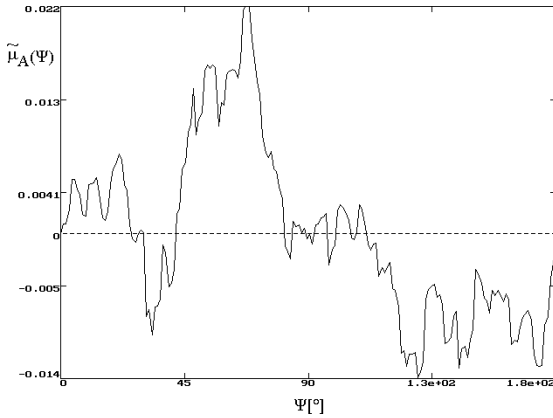


Figura 22: Curva $\tilde{\mu}_A(\Psi)$ do *ensemble* $\mathbf{A}(\psi)$ obtida através de (75).

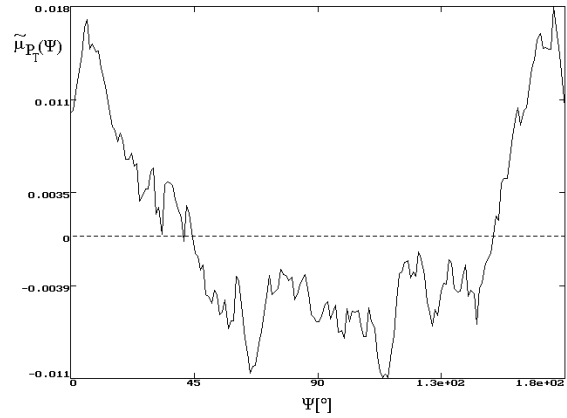


Figura 23: Curva $\tilde{\mu}_{P_T}(\Psi)$ do *ensemble* $\mathbf{P}_T(\psi)$ obtida através de (75).

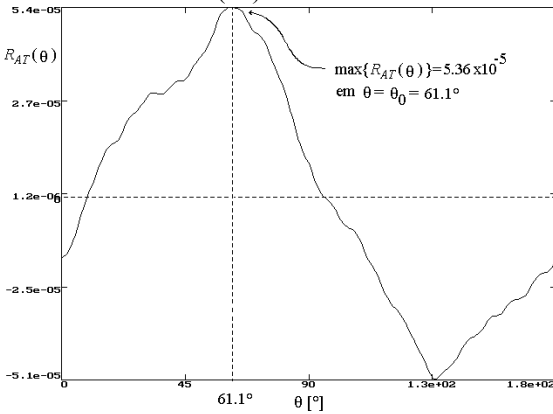


Figura 24: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

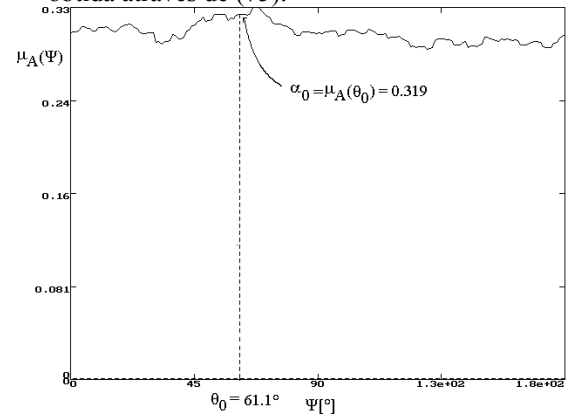


Figura 25: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

Continuando, determina-se $\alpha_a = \max\{\mu_A(\Psi)\}$ (ver Figura 21), escalona-se $p(x, y)$ por α_a e obtém-se $p_T(x, y)$. Determina-se, então, $P_T(\omega, \Psi) = \text{FFT}\{p_T(x, y)\}$ e submete-se $|P_T(\omega, \Psi)|$ ao processo mostrado na Figura D.1. O resultado é a curva $\tilde{\mu}_{P_T}(\Psi)$, mostrada na Figura 23, obtida a partir do *ensemble* resultante $\mathbf{P}_T(\psi)$ e através de (74) e (75).

A partir das curvas $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, Figuras 22 e 23, determina-se a correlação cruzada $R_{AT}(\theta)$ entre elas através de (76), mostrada na Figura 24. A partir de

$R_{AT}(\theta)$, determina-se θ_0 tal que $R_{AT}MAX = \max \{ R_{AT}(\theta) \}$ para $\theta = \theta_0$.

De posse da estimativa do ângulo de rotação θ_0 , determina-se $\alpha_0 = \mu_A(\theta_0)$ conforme mostra a Figura 25.

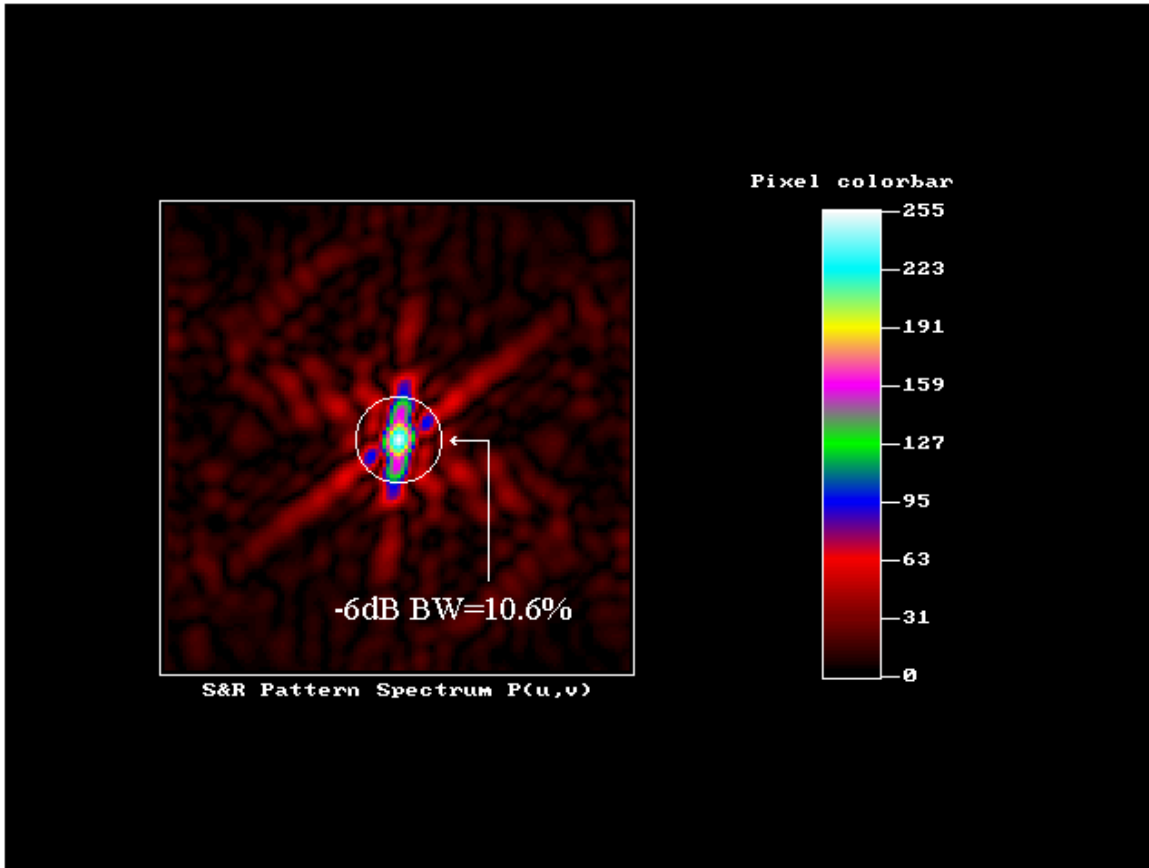


Figura 26: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 10.6%.

Finalmente, escalona-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$. Determina-se $P_0(\omega, \Psi) = \text{FFT}\{p_0(x,y)\}$ (ver Figura 26) e obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 3.

Tabela 3: Resultado do processo de detecção e localização - equações (77) a (82)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como $\theta_0 + 180^\circ$
\hat{x}_0	36.2 pixels	41.8 pixels
\hat{y}_0	26.4 pixels	26.9 pixels
$V\hat{x}_0\%$	12.5%	4.31%
$V\hat{y}_0\%$	5.95%	1.12%

Como os menores valores de variância normalizada percentual resultaram em $V\hat{x}_0\% = 4.31\%$ e $V\hat{y}_0\% = 5.95\%$ conclui-se que a forma geométrica na imagem alvo não é a do padrão. Quanto menores os valores de variância normalizada percentual, maior a semelhança geométrica entre padrão e alvo. Ao comparar os valores obtidos para $V\hat{x}_0\%$ e $V\hat{y}_0\%$ neste item (no qual a imagem alvo não contém o padrão) com aqueles obtidos no item 7.1 (no qual a imagem alvo contém o padrão) a razão entre as respectivas variâncias normalizadas é maior do que 35, indicando a sensibilidade do método à detelhes geométricos.

7.3 Identificação do Caracter “C” *Times New Roman* Escalonado, Rotacionado e Transladado

Analisemos agora a identificação do caracter “C” *Times New Roman*. A Figura 27 mostra a imagem padrão e a Figura 28 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão. Note que o caracter “C” é uma forma geométrica cujo espectro apresenta simetria rotacional, o que, a princípio, dificultaria o método por não haver direções espectrais de maior energia bem definidas.

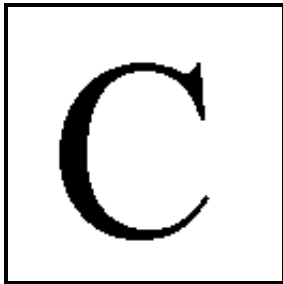


Figura 27: Imagem padrão $p(x,y)$ - caracter “C” *Times New Roman* 128x128 pixels.

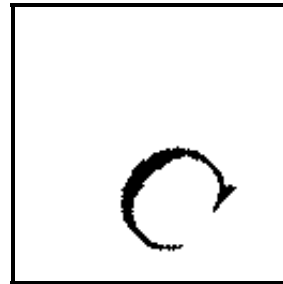


Figura 28: Imagem alvo $a(x,y)$ - caracter “C” da Figura 27 com fator de escalonamento 0.6 rotacionado de -45° e trasladado de (14,21) pixels.

Como anteriormente, primeiro determina-se $A(\omega, \Psi) = \text{FFT}\{a(x,y)\}$ e submete-se $|A(\omega, \Psi)|$ ao processo mostrado na Figura D.1. O resultado são as curvas $\mu_A(\Psi)$ (Figura 29) e $\tilde{\mu}_A(\Psi)$, obtidas a partir do *ensemble* resultante $\mathbf{A}(\psi)$ e através de (74) e (75). Determina-se $\alpha_a = \max\{\mu_A(\Psi)\}$ (ver Figura 29), escalona-se $p(x,y)$ por α_a e obtém-se $p_T(x,y)$. Calcula-se $P_T(\omega, \Psi) = \text{FFT}\{p_T(x,y)\}$ e submete-se $|P_T(\omega, \Psi)|$ ao processo mostrado na Figura D.1 resultando na curva $\tilde{\mu}_{P_T}(\Psi)$ obtida através de (74) e (75) aplicadas ao *ensemble* $\mathbf{P}_T(\psi)$.

Obtém-se a correlação cruzada $R_{AT}(\theta)$ entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, através de (76), mostrada na Figura 30. A partir de $R_{AT}(\theta)$, determina-se θ_0 tal que $R_{AT}MAX = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

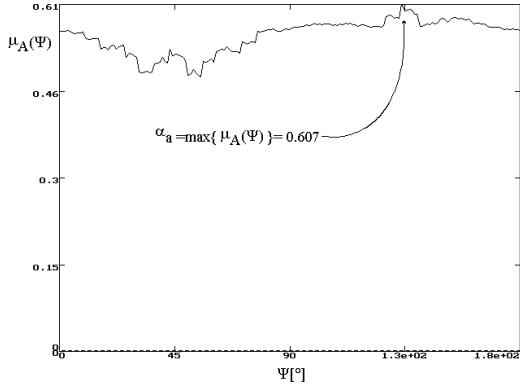


Figura 29: Curva $\mu_A(\Psi)$ do *ensemble* $\mathbf{A}(\psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

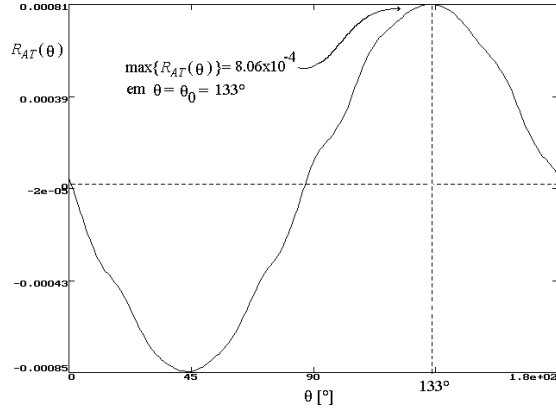


Figura 30: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

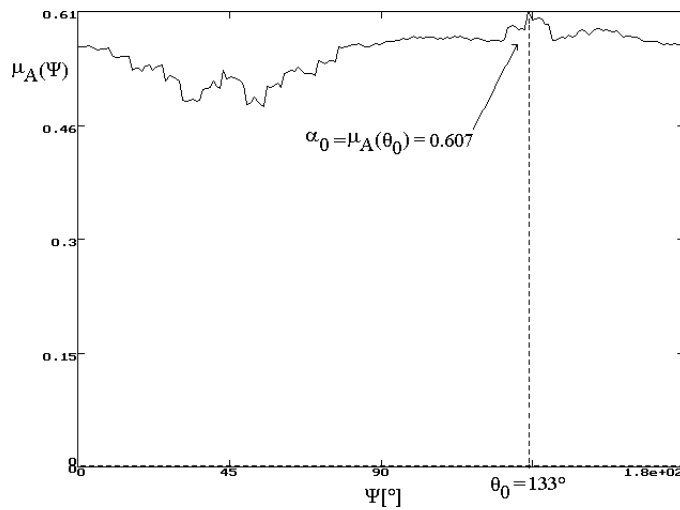


Figura 31: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

De posse da estimativa do ângulo de rotação θ_0 , determina-se $\alpha_0 = \mu_A(\theta_0)$ conforme mostra a Figura 31.

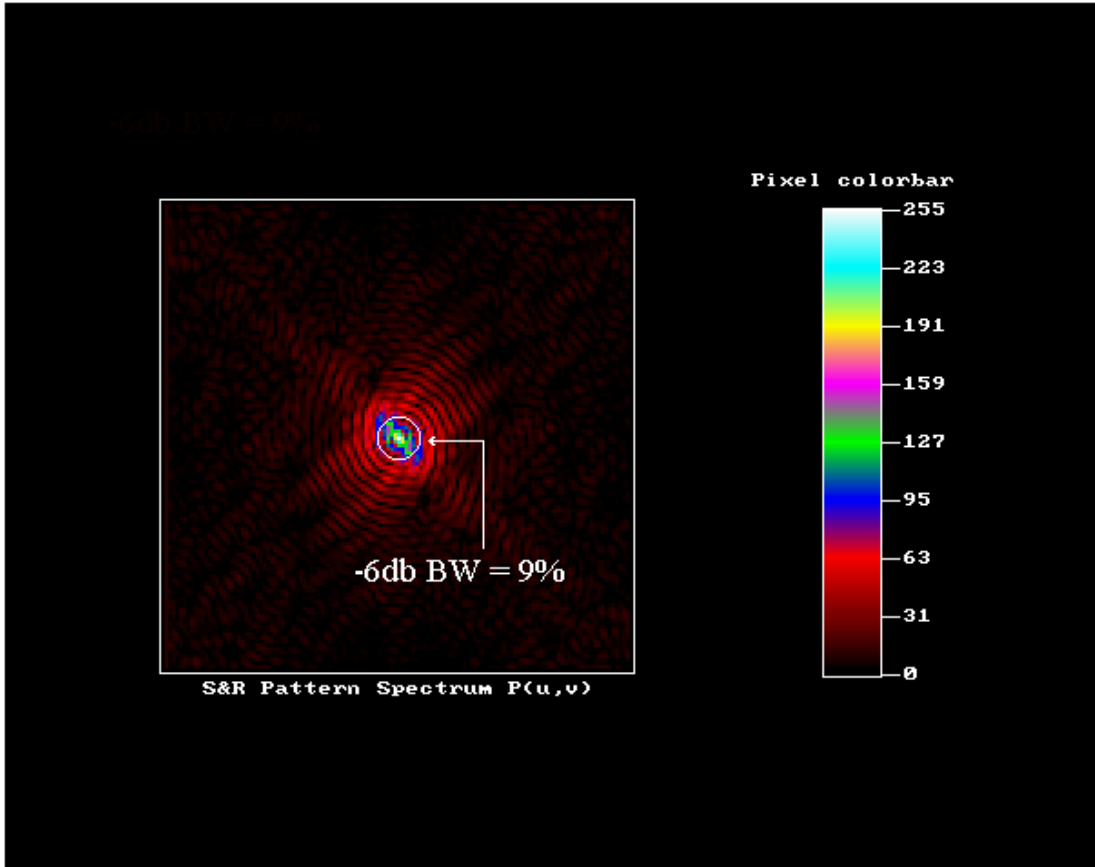


Figura 32: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 9%.

Finalmente, escala-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$. Determina-se $P_0(\omega, \Psi) = \text{FFT}\{p_0(x,y)\}$ (ver Figura 32) e obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 4.

Tabela 4: Resultado do processo de detecção e localização - equações (77) a (82)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como θ_0+180°
\hat{x}_0	-7.4 pixels	14.9 pixels
\hat{y}_0	1.6 pixels	21.9 pixels
$V\hat{x}_0\%$	35.7%	0.13%
$V\hat{y}_0\%$	22.4%	0.018%

Como $V\hat{x}_0\%=0.13\%$ e $V\hat{y}_0\%=0.018\%$ para a estimativa de ângulo de rotação $\theta_0+180^\circ=133^\circ+180^\circ=313^\circ=-47^\circ$ conclui-se que esta é a estimativa do ângulo, e não θ_0 , em que o padrão encontra-se rotacionado na imagem alvo. O fator de escalonamento estimado é $\alpha_0=0.607$.

7.4 Identificação do Carácter “O” Times New Roman Comparado ao Padrão “C”(Figura 27)

Neste item é mostrado a identificação do carácter “O” Times New Roman em relação à imagem padrão da Figura 27. A Figura 33 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão. Note que o caractere “O” é bastante semelhante ao caractere “C” tanto no domínio espaço como no domínio frequência.

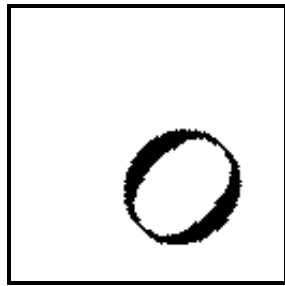


Figura 33: Imagem alvo $a(x,y)$ - carácter “O” Times New Roman com fator de escalonamento 0.6 rotacionado de -45° e transladado de (14,21) pixels.

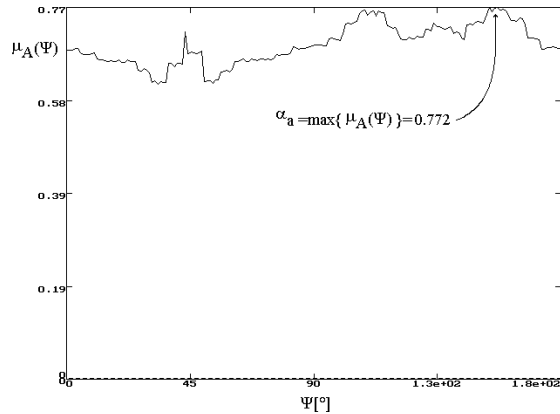


Figura 34: Curva $\mu_A(\Psi)$ do ensemble $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

A Figura 35 mostra a correlação cruzada $R_{AT}(\theta)$ entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, obtida através de (76) e θ_0 tal que $R_{AT}MAX = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

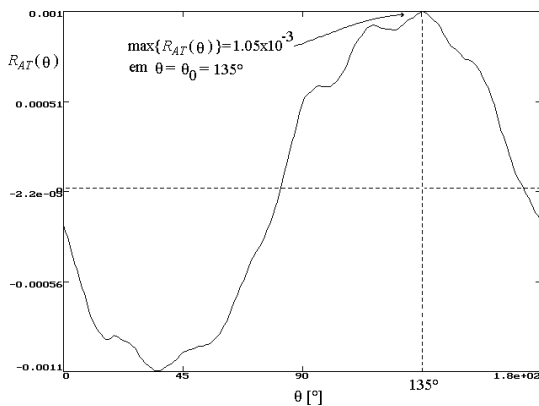


Figura 35: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

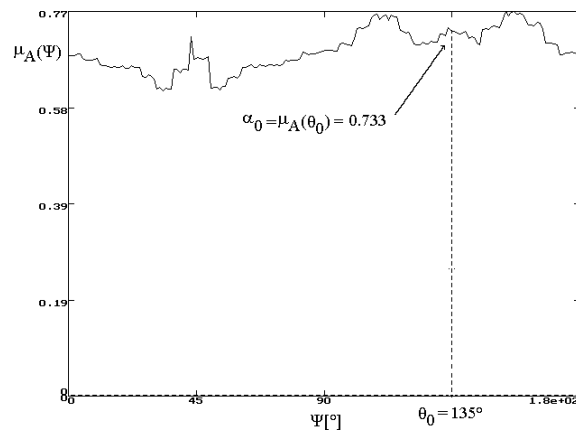


Figura 36: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

A Figura 36 mostra a estimativa do fator de escalonamento através de $\alpha_0 = \mu_A(\theta_0)$.

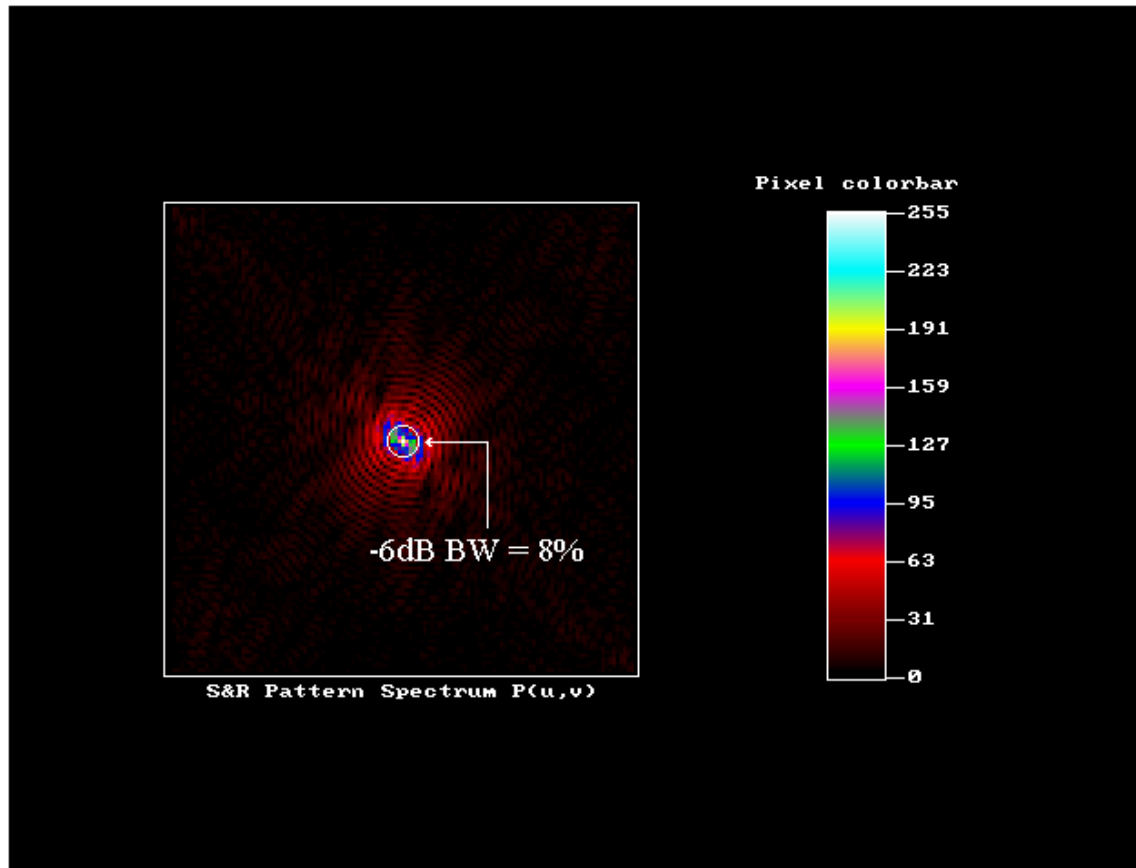


Figura 37: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -6dB BW = 8%.

Escalona-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$. Determina-se $P_0(\omega, \Psi) = \text{FFT}\{p_0(x,y)\}$ (ver Figura 37) e obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 5.

Tabela 5: Resultado do processo de detecção e localização - equações (77) a (82)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como θ_0+180°
\hat{x}_0	-2.3 pixels	26.9 pixels
\hat{y}_0	-0.9 pixels	23.8 pixels
$V\hat{x}_0\%$	19882.5%	135.1%
$V\hat{y}_0\%$	35267.1%	85.3%

Como os menores valores de variância normalizada percentual resultaram em $V\hat{x}_0\% = 135.1\%$ e $V\hat{y}_0\% = 85.3\%$ conclui-se que a forma geométrica na imagem alvo não é a do padrão (comparar com os valores obtidos para $V\hat{x}_0\%$ e $V\hat{y}_0\%$ no item 7.3).

7.5 Identificação de Composição de Formas Geométricas Escalonada, Rotacionada e Transladada

Neste item é mostrado a identificação de composição de formas geométricas em *gray-scale*. A Figura 38 mostra a imagem padrão e a Figura 39 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão.

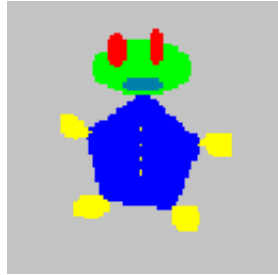


Figura 38: Imagem padrão $p(x,y)$ 128x128 pixels.

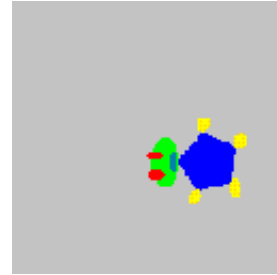


Figura 39: Imagem alvo $a(x,y)$ - imagem padrão $p(x,y)$ da Figura 38 com fator de escalonamento 0.5 rotacionada de 90° e transladada de (21,14) pixels.

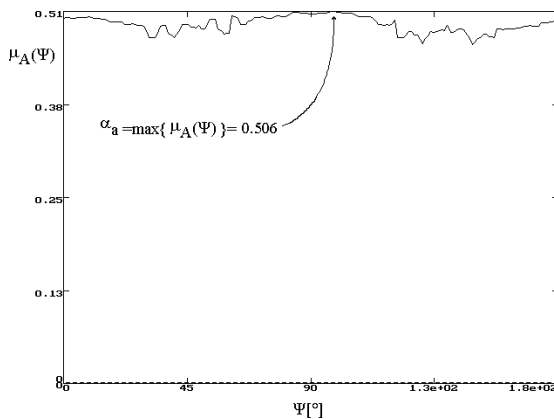


Figura 40: Curva $\mu_A(\Psi)$ do *ensemble* $\mathbf{A}(\psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

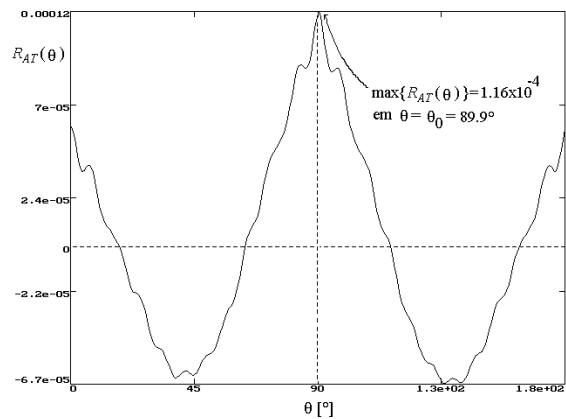


Figura 41: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

A Figura 41 mostra a correlação cruzada $R_{AT}(\theta)$ entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, obtida através de (76) e θ_0 tal que $R_{AT}MAX = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

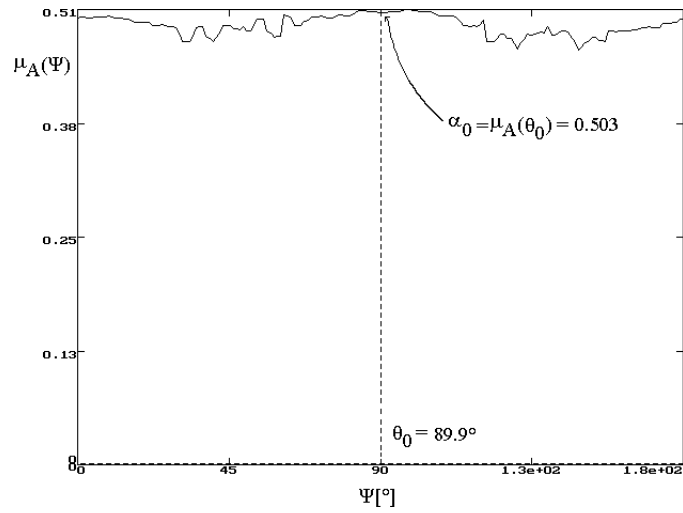


Figura 42: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

A Figura 42 mostra a estimativa do fator de escalonamento através de $\alpha_0 = \mu_A(\theta_0)$.

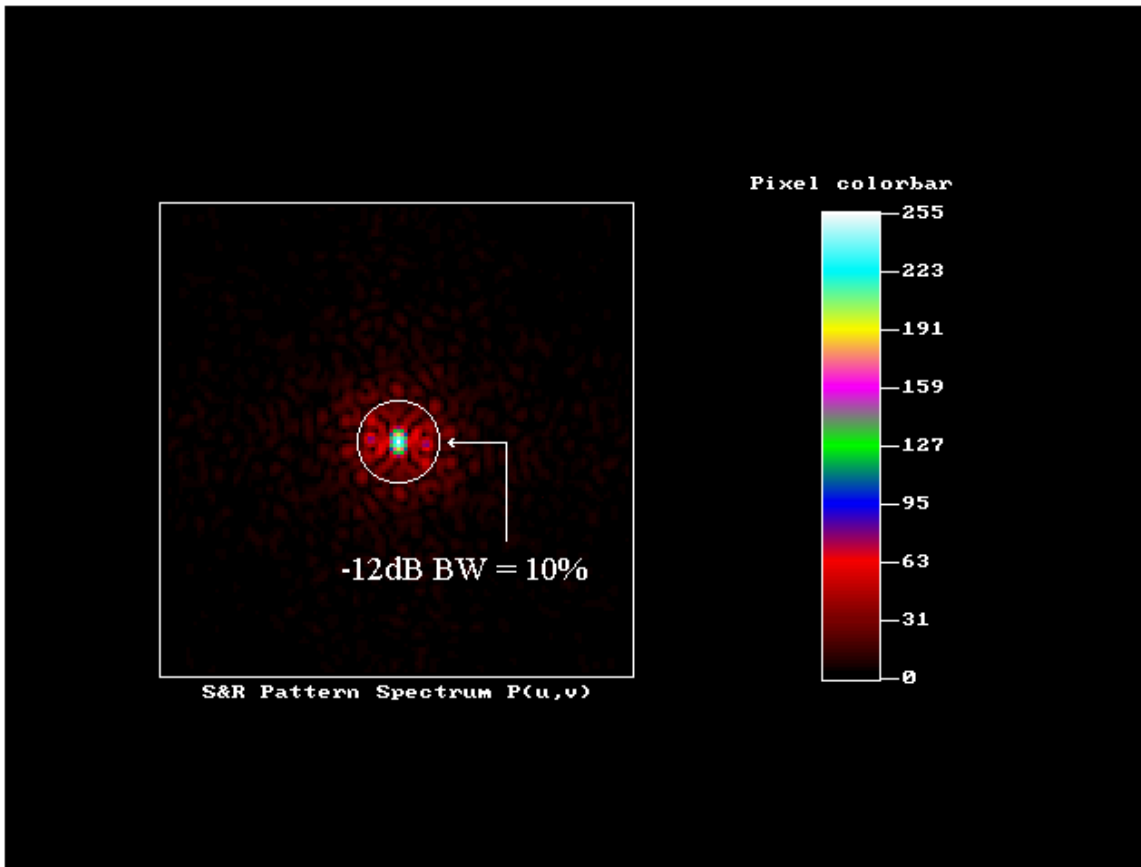


Figura 43: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -12dB BW =10%.

Escalona-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$. Determina-se $P_0(\omega, \Psi) = \text{FFT}\{p_0(x,y)\}$ (ver Figura 43) e obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 6.

Tabela 6: Resultado do processo de detecção e localização - equações (77) a (82)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como θ_0+180°
\hat{x}_0	21.0 pixels	21.4 pixels
\hat{y}_0	14.0 pixels	10.9 pixels
$V\hat{x}_0\%$	$3.7 \times 10^{-12}\%$	104.43 %
$V\hat{y}_0\%$	$2.2 \times 10^{-11}\%$	64.4 %

Como $V\hat{x}_0\% \approx 0\%$ e $V\hat{y}_0\% \approx 0\%$ para a estimativa de ângulo de rotação $\theta_0=89.9^\circ$ conclui-se que esta é a estimativa do ângulo, e não θ_0+180° , em que o padrão encontra-se rotacionado na imagem alvo. O fator de escalonamento estimado é $\alpha_0=0.503$.

7.6 Identificação de Composição de Formas Diferenciadas por Detalhe Geométrico em Relação ao Padrão

Neste item é mostrado a identificação de composição de formas geométricas em *gray-scale* distintas por detalhes. A Figura 38 é a imagem padrão e a Figura 44 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão.

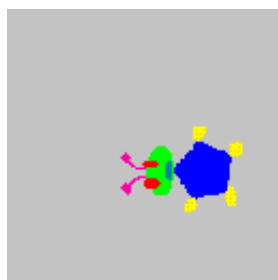


Figura 44: Imagem alvo $a(x,y)$ - imagem padrão $p(x,y)$ da Figura 38 com fator de escalonamento 0.5 rotacionada de 90° , transladada de (21,14) pixels e acrescida de detalhe na parte superior.

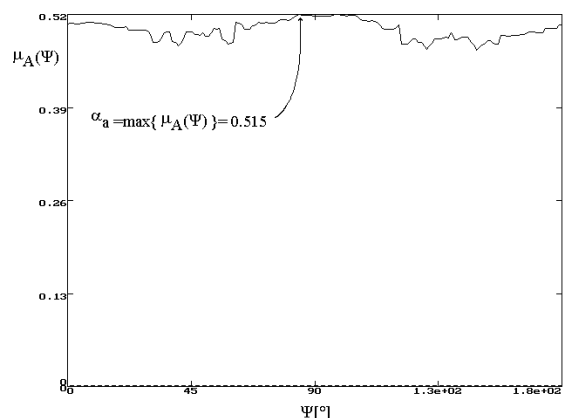


Figura 45: Curva $\mu_A(\Psi)$ do *ensemble* $\mathbf{A}(\psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

A Figura 46 mostra a correlação cruzada $R_{AT}(\theta)$ entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, obtida através de (76) e θ_0 tal que $R_{AT}MAX = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

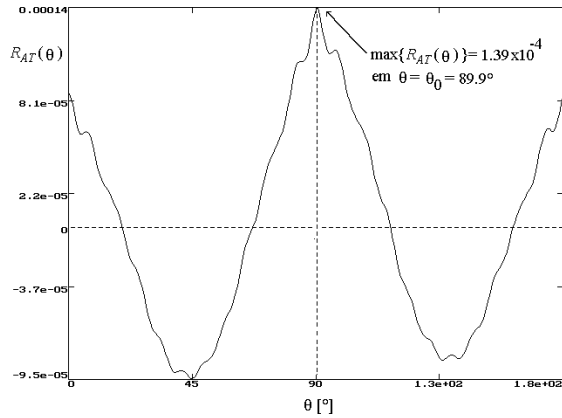


Figura 46: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

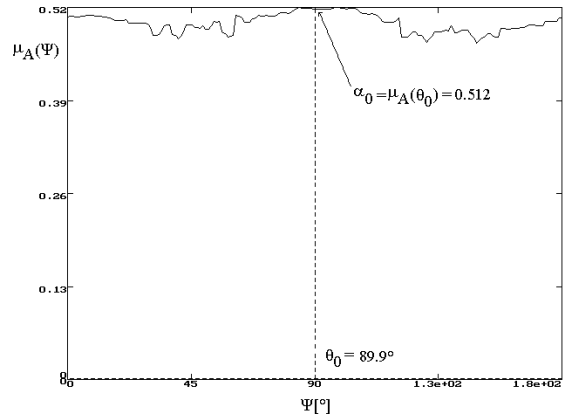


Figura 47: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

A Figura 47 mostra a estimativa do fator de escalonamento através de $\alpha_0 = \mu_A(\theta_0)$.

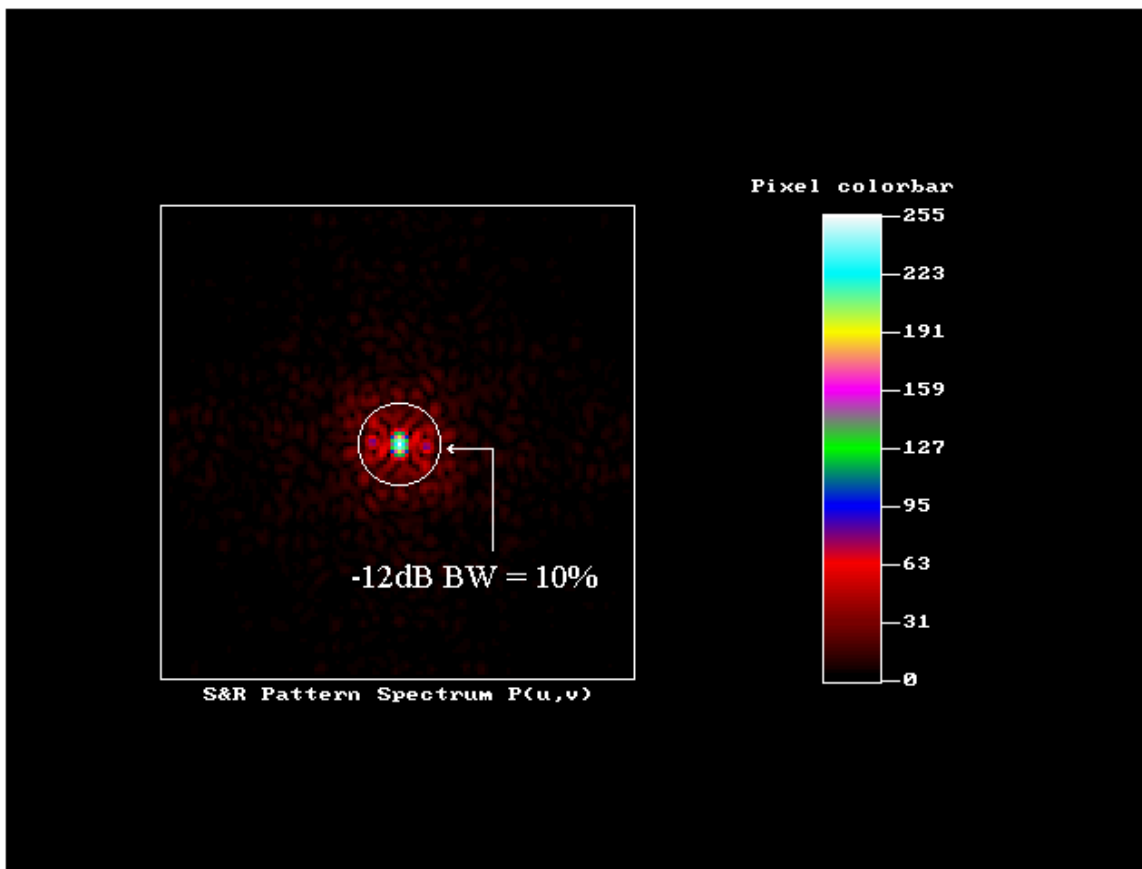


Figura 48: Módulo $\log+1$ do espectro de $p_0(x,y)$ e definição do filtro -12dB BW = 10%.

Escalona-se $p(x,y)$ por α_0 , rotaciona-se por θ_0 e obtém-se $p_0(x,y)$. Determina-se $P_0(\omega,\Psi)=\text{FFT}\{p_0(x,y)\}$ (ver Figura 48) e obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 7.

Tabela 7: Resultado do processo de detecção e localização - equações (77) a (82)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como θ_0+180°
\hat{x}_0	21.2 pixels	20.5 pixels
\hat{y}_0	13.6 pixels	40.6 pixels
$V\hat{x}_0\%$	0.964 %	100.1 %
$V\hat{y}_0\%$	0.952 %	40.6 %

Como a $V\hat{x}_0\%=0.964\%$ e $V\hat{y}_0\%=0.952\%$ para $\theta_0=89.9^\circ$ conclui-se que sob esta estimativa do ângulo de rotação deve haver algo geometricamente semelhante ao padrão na imagem alvo. Não é possível afirmar que seja exatamente a forma geométrica padrão se compararmos os valores de $V\hat{x}_0\%$ e $V\hat{y}_0\%$ com aqueles obtidos no item 7.5. Note, portanto, a grande sensibilidade do método ao comparar formas diferentes apenas por detalhes.

7.7 Identificação do Caracter “A” *Times New Roman* sob Ruído Impulsivo Comparado a ele Próprio (Figura 12)

Neste item é mostrado a identificação do caracter “A” *Times New Roman* corrompido com ruído impulsivo em relação à imagem padrão da Figura 12. A Figura 49 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão.

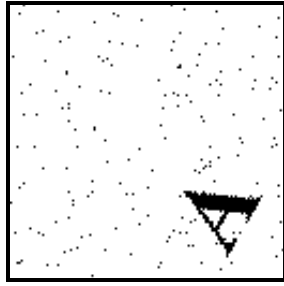


Figura 49: Imagem alvo $a(x,y)$ - caracter "A" da Figura 12 com fator de escalonamento 0.4 rotacionado de 60° , transladado de (40,30) pixels e com 1% da totalidade dos pixels corrompido por ruído impulsivo.

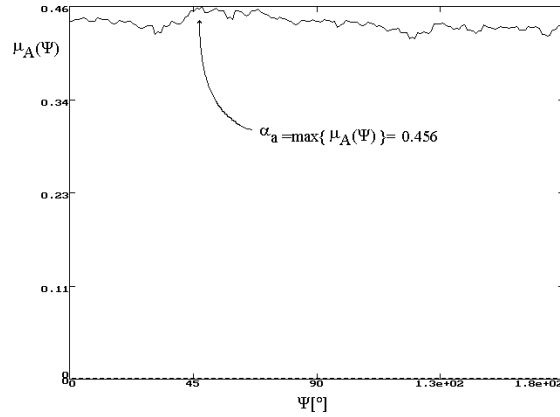


Figura 50: Curva $\mu_A(\Psi)$ do ensemble $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

A Figura 51 mostra a correlação cruzada $R_{AT}(\theta)$ entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, obtida através de (76) e θ_0 tal que $R_{AT}MAX = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

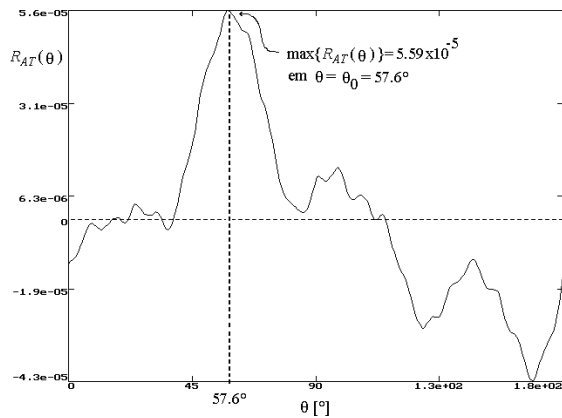


Figura 51: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

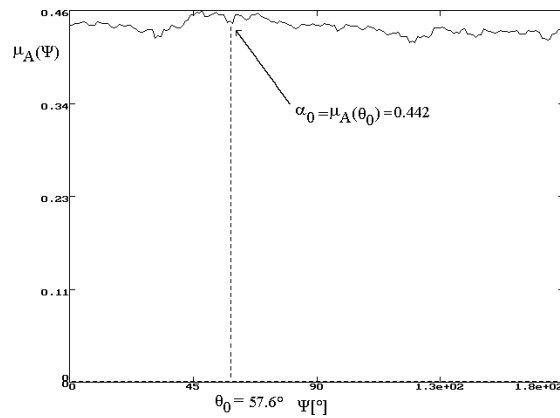


Figura 52: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

A Figura 52 mostra a estimativa do fator de escalonamento através de $\alpha_0 = \mu_A(\theta_0)$. Obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 8.

Tabela 8: Resultado do processo de detecção e localização - equações (77) a (82) Filtro com BW=10%		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como $\theta_0 + 180^\circ$
\hat{x}_0	41.0 pixels	49.9 pixels
\hat{y}_0	32.1 pixels	33.6 pixels
$V\hat{x}_0\%$	1.56 %	15.2 %
$V\hat{y}_0\%$	2.61 %	11.3 %

Nesta situação ruidosa, embora os valores de $V\hat{x}_0\%$ e $V\hat{y}_0\%$ não permitam concluir com precisão sobre a existência ou não do padrão na imagem alvo, os valores obtidos para \hat{x}_0 e \hat{y}_0 são bastante próximos aos reais. Isto sugere que o método possa, a partir dos valores obtidos para α_0 , θ_0 , \hat{x}_0 e \hat{y}_0 , isolar do ruído a região da imagem alvo $a(x,y)$ onde supostamente se encontra a forma geométrica do padrão e, sobre esta nova imagem alvo, aplicar recursivamente o algoritmo de detecção e localização.

7.8 Identificação do Caracter “V” *Times New Roman* sob Ruído Impulsivo Comparado ao Padrão “A”(Figura 12)

Neste item é mostrado a identificação do caracter “V” *Times New Roman* corrompido com ruído impulsivo em relação à imagem padrão da Figura 12. A Figura 53 mostra a imagem alvo na qual busca-se identificar a forma geométrica do padrão.

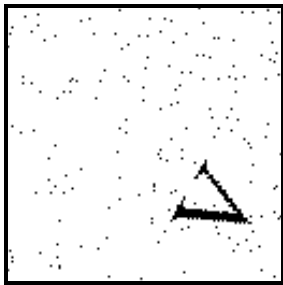


Figura 53: Imagem alvo $a(x,y)$ - caracter “V” *Times New Roman* com fator de escalonamento 0.4 rotacionado de 60° , transladado de (40,30) pixels e com 1% da totalidade dos pixels corrompido por ruído impulsivo.

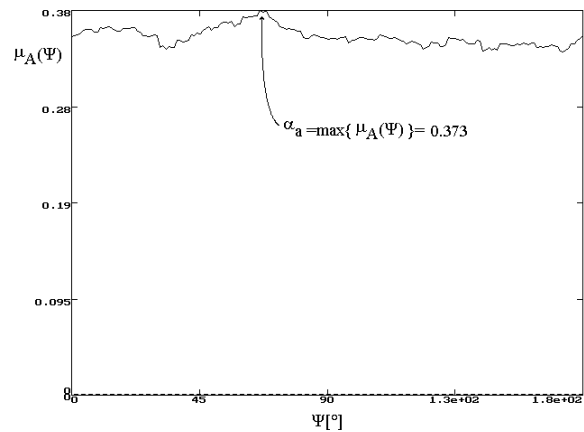


Figura 54: Curva $\mu_A(\Psi)$ do *ensemble* $\mathbf{A}(\Psi)$ obtida através de (74). O máximo de $\mu_A(\Psi)$ é a aproximação inicial do fator de escalonamento.

A Figura 55 mostra a correlação cruzada $R_{AT}(\theta)$ entre $\tilde{\mu}_A(\Psi)$ e $\tilde{\mu}_{P_T}(\Psi)$, obtida através de (76) e θ_0 tal que $R_{AT}MAX = \max\{R_{AT}(\theta)\}$ para $\theta = \theta_0$.

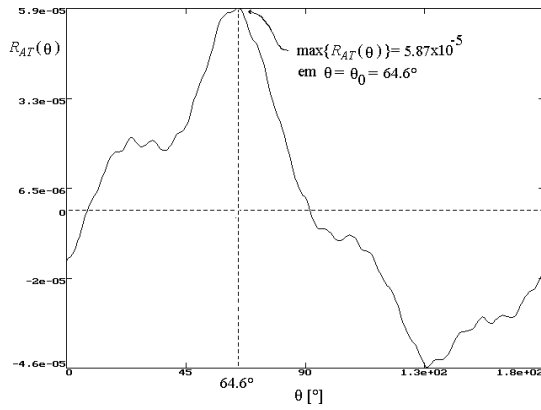


Figura 55: Curva $R_{AT}(\theta)$ obtida através de (76). O máximo desta curva estima o ângulo de rotação θ_0 .

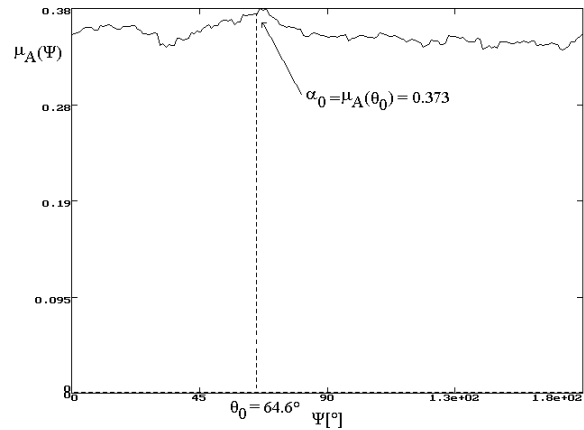


Figura 56: Fator de escalonamento estimado por $\alpha_0 = \mu_A(\theta_0)$.

A Figura 56 mostra a estimativa do fator de escalonamento através de $\alpha_0 = \mu_A(\theta_0)$. Obtém-se $V\hat{x}_0$, \hat{x}_0 , $V\hat{y}_0$ e \hat{y}_0 através de (77) a (82) para os ângulos de rotação θ_0 e $\theta_0 + 180^\circ$, conforme mostra a Tabela 9.

Tabela 9: Resultado do processo de detecção e localização - equações (77) a (82) Filtro com BW=11.84% ($10\% \times 0.442 / 0.373 \Rightarrow 0.442 = \alpha_0$ item 7.7, $0.373 = \alpha_0$ item 7.8)		
	Ângulo de rotação estimado como θ_0	Ângulo de rotação estimado como $\theta_0 + 180^\circ$
\hat{x}_0	34.8 pixels	44.5 pixels
\hat{y}_0	24.3 pixels	26.6 pixels
$V\hat{x}_0\%$	28.7 %	3.01 %
$V\hat{y}_0\%$	19.4 %	4.85 %

8. Conclusão

Este trabalho propõe um novo método para identificação de padrões isolados em imagens bidimensionais. O procedimento é capaz de identificar formas geométricas a partir das características espectrais mais significativas do padrão armazenadas em respectivas RNAs, previamente treinadas para tanto. O método é aplicável em tempo real já que todas as principais operações baseiam-se na FFT ou na característica de extrapolação da relação entrada-saída das RNAs.

O processo de detecção resulta em dois índices numéricos, $V\hat{x}_0\%$ e $V\hat{y}_0\%$, baseados na correlação espectral entre padrão e alvo. Para cada tipo de padrão deve-se determinar experimentalmente o limiar dos índices que decide quanto à existência ou não do padrão na imagem alvo. A necessidade de experimentação decorre da aleatoriedade do MSE final de treino das RNAs de padrão para padrão.

O algoritmo de detecção apresenta sensibilidade na comparação de formas geométricas diferenciadas apenas por detalhe, incluindo variações de nível de pixel (*gray-level* ou cor).

O filtro direcional utilizado no domínio frequência dá robustez ao processo de localização quanto à ruído impulsivo, no entanto, o limiar de decisão dos índices $V\hat{x}_0\%$ e $V\hat{y}_0\%$ fica comprometido.

O método assume que a imagem alvo tenha sido submetida a um processo de segmentação para fins de isolamento da forma geométrica nela contida do efeito de ruído ou outras formas geométricas coexistentes. No entanto, como ao final do processo consegue-se uma estimativa do escalonamento α_0 , rotação θ_0 e translação (\hat{x}_0, \hat{y}_0) da forma geométrica supostamente existente na imagem alvo é possível extrair da imagem alvo a região de interesse, isolando-a do ruído e demais objetos presentes.

Futuros estudos nesta linha de pesquisa incluem a possibilidade de tornar o algoritmo recursivo, isto é, pela propriedade do método apontada no parágrafo anterior, sucessivamente aplicar o procedimento sobre regiões por ele próprio isoladas no sentido de dispensar a necessidade de pré-segmentação.

9. Apêndices

A. Rotação Simultânea nos Domínios Espaço e Frequência

Seja

$$F(u, v) = FFT\{f(x, y)\} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (\text{A.1})$$

o espectro da imagem $f(x, y)$.

Passando a expressão (A.1) para coordenadas polares, temos:

Sejam

$$\begin{aligned} x &= r \cos \theta & u &= \omega \cos \phi \\ y &= r \sin \theta & v &= \omega \sin \phi \end{aligned} \quad (\text{A.2})$$

$$F(\omega, \phi) = FFT\{f(r, \theta)\} = \sum_{r=0}^{r_{\max}} \sum_{\theta=0}^{\theta_{\max}} f(r, \theta) e^{-j\frac{2\pi}{N}\omega r(\cos \theta \cos \phi + \sin \theta \sin \phi)} \quad (\text{A.3})$$

$$F(\omega, \phi) = \sum_{r=0}^{r_{\max}} \sum_{\theta=0}^{\theta_{\max}} f(r, \theta) e^{-j\frac{2\pi}{N}\omega r \frac{1}{2}(\cos(\theta+\phi) + \cos(\theta-\phi) + \cos(\theta-\phi) - \cos(\theta+\phi))} \quad (\text{A.4})$$

$$F(\omega, \phi) = \sum_{r=0}^{r_{\max}} \sum_{\theta=0}^{\theta_{\max}} f(r, \theta) e^{-j\frac{2\pi}{N}\omega r(\cos(\theta-\phi))} \quad (\text{A.5})$$

Girando a imagem no domínio espaço de um ângulo α , tem-se

$$FFT\{f(r, \theta + \alpha)\} = \sum_{r=0}^{r_{\max}} \sum_{\theta=0}^{\theta_{\max}} f(r, \theta + \alpha) e^{-j\frac{2\pi}{N}\omega r(\cos(\theta-\phi))} \quad (\text{A.6})$$

Seja

$$\theta + \alpha = \beta \Rightarrow \theta = \beta - \alpha \quad (\text{A.7})$$

Logo

$$FFT\{f(r, \theta + \alpha)\} = \sum_{r=0}^{r_{\max}} \sum_{\beta=0}^{\beta_{\max}} f(r, \beta) e^{-j\frac{2\pi}{N}\omega r \cos(\beta - (\phi + \alpha))} = F(\omega, \phi + \alpha) \quad (\text{A.8})$$

A expressão (A.8) indica que se a imagem no domínio espaço é rotacionada de um ângulo α seu espectro é rotacionado do mesmo ângulo no domínio frequência.

B. Ortogonalidade

Sejam x , w e N inteiros tal que

$$0 < x < N-1$$

$$0 < w < N-1$$

Seja $\delta = x-w$.

Logo $-(N-1) < \delta < (N-1)$.

Seja

$$\sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(ux)} e^{j\frac{2\pi}{N}(uw)} = \sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(u(x-w))} = \sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(\delta u)} = \sum_{u=0}^{N-1} e^{-j\frac{2\pi}{\left(\frac{N}{\delta}\right)}u} = \sum_{u=0}^{N-1} e^{-j\frac{2\pi}{P}u} \quad (\text{B.1})$$

Note que:

$P \Rightarrow$ Número de amostras em um período de revolução da exponencial $e^{-j\frac{2\pi}{P}u}$.

$N \Rightarrow$ Número de amostras totais computados em $\sum_{u=0}^{N-1} e^{-j\frac{2\pi}{P}u}$.

$\frac{N}{P} = \frac{N}{\frac{N}{\delta}} = \delta \Rightarrow$ Número de períodos de revolução da exponencial complexa (cada um

com P amostras) computados no somatório sobre todo intervalo de variação de u (de 0 a $N-1$).

A soma dos valores de uma exponencial complexa ao longo de seu período completo de revolução é sempre nulo, conforme podemos verificar pela soma zero dos 6 fasores complexos para o caso $P=6$, mostrado na Figura B.1 a seguir.

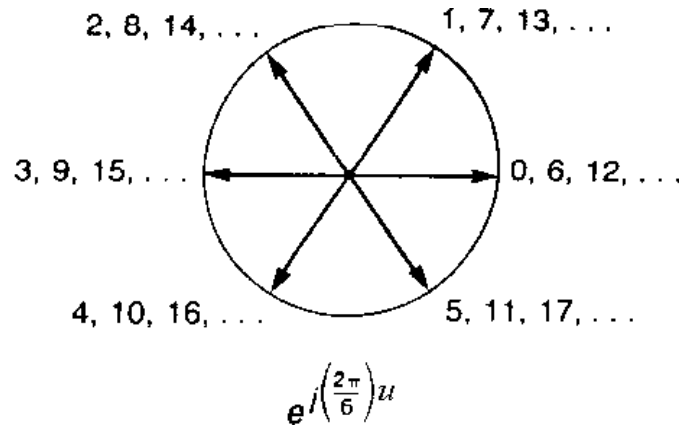


Figura B.1: Exponencial complexa de período $P=6$.

Como δ é um número inteiro, para o caso $\delta \neq 0$ um número inteiro de períodos completos da exponencial complexa são somados. Como a soma sobre cada um individualmente resulta em zero, o resultado global também será zero.

Para o caso $\delta=0$, $e^{-j\frac{2\pi}{N}(\delta u)} = 1$ e temos $\sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(\delta u)} = N$.

Portanto

$$\sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(ux)} e^{j\frac{2\pi}{N}(uw)} = \begin{cases} N & \text{para } x = w \\ 0 & \text{para } x \neq w \end{cases} \quad (\text{B.2})$$

As exponenciais $e^{-j\frac{2\pi}{N}(ux)}$ e $e^{j\frac{2\pi}{N}(uw)}$ são ditas serem ortogonais entre si.

C. Energia e Densidade Espectral de Energia

Seja a imagem $f(x,y)$ real.

Seja

$$F(u, v) = FFT\{f(x, y)\} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (\text{C.1})$$

e seja

$$F^*(u, v) = \sum_{w=0}^{N-1} \sum_{z=0}^{N-1} f(w, z) e^{j\frac{2\pi}{N}(uw+vz)} \quad (\text{C.2})$$

Multiplicando (C.1) por (C.2) e aplicando o somatório em relação a u e v em ambos lados da equação resultante temos:

$$\begin{aligned}
\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) F^*(u,v) &= \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} |F(u,v)|^2 = \\
&= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{w=0}^{N-1} \sum_{z=0}^{N-1} f(x,y) f(w,z) \left(\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} e^{-j\frac{2\pi}{N}(ux+vy)} e^{j\frac{2\pi}{N}(uw+vz)} \right) = \\
&= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{w=0}^{N-1} \sum_{z=0}^{N-1} f(x,y) f(w,z) \left[\sum_{u=0}^{N-1} e^{-j\frac{2\pi}{N}(ux)} e^{j\frac{2\pi}{N}(uw)} \right] \left[\sum_{v=0}^{N-1} e^{-j\frac{2\pi}{N}(vy)} e^{j\frac{2\pi}{N}(vz)} \right]
\end{aligned} \tag{C.3}$$

Pelo princípio de ortogonalidade (ver apêndice B), o 1º termo entre colchetes em (C.3) é nulo para $x \neq w$ e igual a N para $x=w$. Semelhantemente, o 2º termo entre colchetes é nulo para $y \neq z$ e igual a N para $y=w$. Assim (C.3) torna-se:

$$\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} |F(u,v)|^2 = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x,y) N^2 \tag{C.4}$$

ou

$$\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x,y) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} |F(u,v)|^2 \tag{C.5}$$

O termo a esquerda na equação acima define a energia total E da imagem obtida no domínio espaço. O termo a direita define também a energia E só que seu cômputo é feito no domínio freqüência. Assim, (C.5) postula que a energia total E de uma imagem pode ser obtida tanto no domínio espaço quanto no domínio freqüência. O termo $|F(u,v)|^2$ é denominado Densidade Espectral de Energia da imagem $f(x,y)$. Este teorema é denominado Teorema de Parseval em duas dimensões.

D. Diagrama do Processo de Localização e Detecção

A Figura D.1 a seguir, mostra o processo de detecção do padrão na imagem alvo para $m=6$. As m RNAs encontram-se treinadas com as m direções de maior energia espectral da imagem padrão $p(x,y)$ original.

10. Referências

- [1] K. S. Fu, *Syntatic Pattern Recognition and Application*, Englewood Cliffs, Prentice Hall, NJ, 1982.
- [2] E. Persoon and K.S. Fu, "Shape discrimination using Fourier descriptors", *IEEE Trans. Syst. Man, and Cybern.*, vol 7, pp. 170-179, 1977.
- [3] R. Gordon and R.M. Rangayyan, "Feature enhancement of film mammograms using fixed and adaptive neighborhoods", *Appl. Optics*, vol. 23, no. 4, pp. 560-564, 1984.
- [4] L. Gupta and M.D. Srinath, "Contour sequence moments for the classification of closed planar shapes", *Pattern Recognition*, vol 20, no.3, pp. 267-272, 1987.
- [5] E. Säcker, B.E. Boser, J. Bromley, Y. LeCun, and L.D. Jackel, "Application of the ANNA neural network chip to high-speed character recognition", *IEEE Transactions on Neural Networks* 3, pp. 498-505, 1992.
- [6] L. Shen, R. M. Rangayyan and J. E. Leo Desautels, "Application of shape analysis to mammographic calcifications", *IEEE Transactions on Medical Imaging*, vol 13, no 2, pp. 263-274.
- [7] G.K. Moore and F.A. Waltz, "Objective procedures for lineament enhancement and extraction", *Photogrammetric Eng. Remote Sensing*, vol 49, pp. 641-647, 1983.
- [8] A. V. Lugt, "Signal Detection By Complex Spatial Filtering", *IEEE Trans. on Information Theory*, April, 1964.
- [9] Boahen, K.A and A.G. Andreou, "A contrast sensitive silicon retina with reciprocal synapses", *Advances in Neural Information Processing Systems 4* (J.E. Moody, S.J. Hanson, and R.P. Lippmann, eds.), pp. 764-772, Morgan Kaufmann, San Mateo, CA, 1992.
- [10] P.M. Cooley and J.W. Tukey, "An Algorithm for the Machine Computation of Complex Fourier Series", *Mathematics of Computation*, vol 19, pp. 297-302, April, 1965.
- [11] T.Y. Young and K-S Fu, *Handbook of Pattern Recognition and Image Processing*, Academic Press, 1986.
- [12] E.L. Hall, *Computer Image Processing and Recognition*, Academic Press, 1979.
- [13] A.V. Oppenheim, *Applications of Digital Signal Processing*, Prentice Hall, 1978.
- [14] E.O. Brigham, *The Fast Fourier Transform and its Applications*, AvanteK, 1988.
- [15] S.D. Stearns and D.R. Hush, *Digital Signal Analysis*, Prentice Hall, 1990.
- [16] R. Gonzalez and P. Wintz, *Digital Image Processing*, Addison-Wesley, 1987.
- [17] I. Aleksander and H. Morton, *An Introduction to Neural Computing*, Chapman & Hall, 1990.
- [18] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1992.
- [19] I. Guyon, "Applications of neural networks to character recognition", *International Journal of Pattern Recognition and Artificial Intelligence*, vol 5, pp. 353-382, 1991.
- [20] B. Widrow and M.A. Lehr, "30 years of adaptive neural networks: perceptron, madaline and backpropagation", *Proceedings of the IEEE*, vol 78, pp. 1415-1442, 1990.
- [21] P.J. Werbos, "Beyond regression: New tools for prediction and analysis in the behavioral sciences", PhD Thesis, Harvard University, Cambridge, MA, 1974.
- [22] S. Haykin, *Neural Networks*, Macmillan College, New York, NY, 1994.
- [23] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning representations by back-propagating errors", *Nature*, London, vol 323, pp. 533-536, 1986.
- [24] M.H. Hassoun, *Fundamentals of Artificial Neural Networks*, MIT Press, 1995.
- [25] J.A. Anderson, *An Introduction to Neural Networks*, MIT Press, 1995.